Universidad de Valladolid
Escuela Técnica Superior de Ingeniería Informática
Departamento de Informática (ATC, CCIA y LSI)



Tesis Doctoral

Desarrollo de un Sistema Experto Simbólico Conexionista de Ayuda al Diagnóstico del Glaucoma

María Aránzazu Simón Hurtado
Junio 2002

Tesis dirigida por el Dr. D. Luis Alonso Romero

Dr. D. Luis Alonso Romero, Catedrático de Universidad del Departamento de Informática y Automática de la Universidad de Salamanca,

HAGO CONSTAR: que Dª Mª Aránzazu Simón Hurtado, Licenciada en Ciencias Físicas por la Universidad de Valladolid, ha realizado bajo mi dirección la Memoria que lleva por título **Desarrollo de un sistema experto simbólico conexionista de ayuda al diagnóstico del glaucoma**, con el fin de que sea presentada como Tesis para aspirar al grado de Doctor.

Y para que surta los efectos oportunos, firmo el presente en Valladolid, a 29 de abril de 2002.

A mis padres

Agradecimientos

Quiero expresar mi agradecimiento a Luis Alonso, sin cuya dirección, consejo y paciencia no se habría realizado este trabajo.

Esta Tesis ha sido realizada en el Departamento de Informática (ATC, CCIA y LSI) de la Escuela Técnica Superior de Ingeniería Informática de la Universidad de Valladolid. Sería difícil nombrar sin dejarme a alguien a todos aquellos que me han ayudado incondicionalmente, dedicando gustosamente su tiempo para aclarar mis dudas, hacerme oportunas observaciones, solucionarme todo tipo de problemas técnicos, y darme ánimos en todo momento.

Una parte fundamental del trabajo se ha realizado gracias a la colaboración de los componentes del Instituto de Oftalmobiología Aplicada (IOBA) de la Universidad de Valladolid, en especial de Alfonso Antón, José Carlos Pastor y José Andrés Maquet. Sin su apoyo y paciencia habría sido imposible llevar a buen término este trabajo.

Me gustaría también recordar a los alumnos que me han ayudado con sus preguntas, sugerencias y proyectos, especialmente a Daniel Bahillo, Eduardo de la Fuente, Antonio de la Cruz y David Hernández.

Agradezco a mi familia todo el apoyo y ánimo que me ha brindado, y la generosidad con la que ha prescindido de mí en muchas ocasiones, para que pudiera dedicar todo el tiempo necesario a esta Tesis. Sobre todo quiero agradecer a mis padres, a quienes dedico este trabajo, toda la ilusión que siempre han puesto en mis proyectos profesionales, y muy especialmente en éste.

Sería muy largo enumerar a todos aquellos que de algún modo han participado en el desarrollo de este trabajo. A todos ellos les estoy sinceramente agradecida.

Abreviaturas

ADRL	Análisis discriminante con regresión logística
BP	Retropropagación del error
CBR	Razonamiento basado en casos
CF	Factor de certeza
CLV	Varianza de la pérdida corregida
CPSD	Patrón de desviación estándar corregido
CV	Campo visual
dB	Decibelios
DL	Pérdida difusa
E	Especificidad
E/P	Relación excavación/papila
FFP	Fracción de falsos positivos
FN	Falsos negativos
FP	Falsos positivos
FVN	Fracción de verdaderos negativos
FVP	Fracción de verdaderos positivos
GCS	Glaucoma crónico simple
GHT	Test del hemicampo de glaucoma
GPI	Índice de patrón de glaucoma
GTE	Glaucoma de tensión elevada
GTN	Glaucoma de tensión normal
HRP	Perimetría de resolución high-pass
HTA	Retinopatía hipertensiva
HTO	Hipertensión ocular
HVA	Analizador de campos visuales Humphrey
IA	Inteligencia artificial
IC	Ingeniero de conocimiento
IOBA	Instituto de oftalmobiología aplicada
LTF	Fluctuación a largo plazo
LV	Varianza de la pérdida
LVQ	Cuantización vectorial adaptativa
MD	Defecto medio
MLP	Perceptrón multicapa
MS	Sensibilidad media
PC	Perimetría computerizada
PD	Precisión diagnóstica
PIO	Presión intraocular
PM	Perimetría manual
PSD	Patrón de desviación estándar
RBF	Redes de función de base radial
RNA	Redes neuronales artificiales
ROC	Curva de características operativas
S	Sensibilidad
SBC	Sistemas basados en conocimiento
SE	Sistemas expertos
SF	Fluctuación a corto plazo
SITA	Algoritmo sueco interactivo de estimación del umbral
SOM	Mapas autoorganizados
TOP	Perimetría orientada por tendencias
VN	Verdaderos negativos
VP	Verdaderos positivos
SWAP	Perimetría de longitud de onda corta
FDT	Perimetría de duplicación de frecuencia

Índice General

PREFACIO **XV**

1 INTRODUCCIÓN **1**

1.1 Planteamiento del problema **1**

1.2 Objetivo **3**

1.3 Técnicas empleadas **3**
1.3.1 Sistemas basados en conocimiento 3
1.3.2 Redes Neuronales Artificiales 5
1.3.3 Razonamiento Basado en Casos 5
1.3.4 Lógica Borrosa 6

1.4 Esquema seguido en esta Tesis **6**

2 ANTECEDENTES. ESTADO DEL ARTE **9**

2.1 Diagnóstico del glaucoma **9**

2.2 Los comienzos en la investigación sobre el diagnóstico automático del glaucoma **15**

2.3 Interpretación del campo visual **18**
2.3.1 Estudio sobre el deterioro del campo visual en el glaucoma 18
2.3.2 Perimetría computerizada 21
2.3.3 Métodos estadísticos 25
2.3.4 Distribución del campo en zonas 35
2.3.5 Sistemas Basados en Conocimiento o Sistemas Expertos 37
2.3.6 Redes Neuronales Artificiales 40

2.4 Diagnóstico global del glaucoma **49**
2.4.1 Sistemas basados en reglas de producción 50
2.4.2 Sistemas híbridos simbólicos-conexionistas borrosos 52

3 SISTEMAS EXPERTOS: PRIMERAS VERSIONES 59

3.1 Prototipo programado en M1 59
3.1.1 La adquisición del conocimiento 59
3.1.2 Arquitectura básica de un Sistema Experto 65
3.1.3 Software para construir Sistemas Expertos 66
3.1.4 Proceso de inferencia a lo largo de la base de conocimiento 68
3.1.5 Asignación de factores de certeza 71
3.1.6 Perspectivas de desarrollo 74

3.2 Sistema Experto Glaucom-Easy 75
3.2.1 Diseño arquitectónico 79
3.2.2 Proceso de inferencia a lo largo de la base de conocimiento 80
3.2.3 Asignación de factores de certeza 83
3.2.4 Validación del Sistema Experto 85
3.2.5 Perspectivas de desarrollo 88

4 ANÁLISIS DEL CAMPO VISUAL 91

4.1 Los datos 91
4.1.1 Estudio piloto 91
4.1.2 Estudio definitivo 92

4.2 Análisis de componentes principales 96

4.3 Distribución del campo en zonas 100

4.4 Clasificación de campos visuales mediante mapas autoorganizados (SOM) 100
4.4.1 Introducción. Algoritmo de aprendizaje SOM 100
4.4.2 Desarrollo de la aplicación GESL 104
4.4.3 Modificación de los algoritmos del SOM_PACK 105
4.4.4 Selección de los parámetros para el entrenamiento de las redes neuronales 107
4.4.5 Resultados 110
4.4.6 Conclusiones 112

4.5 Clasificación de campos visuales utilizando Razonamiento Basado en Casos 113
4.5.1 Introducción al Razonamiento Basado en Casos 113
4.5.2 Tratamiento de los datos y experimentos 120
4.5.3 Resultados 122
4.5.4 Conclusiones 132

4.6 Comparación de los experimentos SOM y CBR 133

4.7 Comparación de los resultados obtenidos en la clasificación de campos visuales con los de otros trabajos de investigación 135

5 SISTEMA HÍBRIDO DE AYUDA AL DIAGNÓSTICO DEL GLAUCOMA 139

5.1 Introducción. Lógica Borrosa 139
5.1.1 Conjuntos borrosos 140
5.1.2 Funciones de pertenencia de conjuntos borrosos 141
5.1.3 Particiones borrosas 143
5.1.4 Operaciones borrosas 144
5.1.5 Inferencia borrosa 146
5.1.6 Reglas borrosas 148
5.1.7 Sistemas de inferencia borrosa 148
5.1.8 Determinación de las funciones de pertenencia 150
5.1.9 Emborronador 154
5.1.10 Aclarador 154

5.2 Lenguaje de desarrollo: FuzzyClips 156
5.2.1 Técnica de inferencia de FuzzyCLIPS 156

5.3 Diseño de las funciones de pertenencia 160

5.4 Arquitectura del sistema híbrido de ayuda al diagnóstico del glaucoma 165

5.5 Inferencia a través de la base de conocimiento 166

5.6 Clasificador del campo visual 169
5.6.1 Desarrollo del Sistema Experto para el diagnóstico del campo visual 169
5.6.2 Elección y modificación de la RNA para clasificación de campos visuales 175
5.6.3 Sistema híbrido simbólico-conexionista para el diagnóstico del campo visual 178

5.7 Integración del clasificador del campo visual en el sistema de diagnóstico global 180

5.8 Modificación del factor de certeza del Sistema Experto Borroso 180

5.9 Validación del sistema híbrido de ayuda al diagnóstico del glaucoma 181
5.9.1 Validación de la clasificación del campo visual 182
5.9.2 Validación del consejo o pauta a seguir con el paciente 184

5.10 Ejemplo del seguimiento de un paciente con el sistema híbrido desarrollado 187

6 CONCLUSIONES Y TRABAJO FUTURO 189

6.1 Conclusiones 189

6.2 Trabajo futuro **191**

APÉNDICE A: SENSIBILIDAD, ESPECIFICIDAD Y PRECISIÓN DIAGNÓSTICA **195**

APÉNDICE B: CURVAS ROC **197**

BIBLIOGRAFÍA **201**

ÍNDICE DE FIGURAS **215**

ÍNDICE DE TABLAS **219**

Prefacio

Esta Tesis se presenta para la obtención del grado de Doctor por la Universidad de Valladolid. Ha sido realizada bajo la dirección del Doctor D. Luis Alonso Romero, Catedrático de Universidad en la Universidad de Salamanca.

El trabajo aquí presentado trata las siguientes áreas del conocimiento humano:

- El problema del diagnóstico del glaucoma.
- Los Sistemas Basados en Conocimiento como ayuda para resolver problemas.
- Las Redes Neuronales Artificiales como método de aprendizaje y clasificación.
- El Razonamiento Basado en Casos utilizado con el mismo objetivo que las redes neuronales artificiales.
- Sistemas híbridos simbólicos-conexionistas.
- Sistemas borrosos.

Se pretende lograr un buen acercamiento al problema del diagnóstico precoz del glaucoma utilizando técnicas de Inteligencia Artificial adecuadas para conseguir unos resultados similares a los del experto.

1 Introducción

1.1 Planteamiento del problema

El glaucoma puede definirse como una enfermedad ocular crónica, que se caracteriza por presentar una degeneración progresiva del nervio óptico que en un 60-70% de los casos cursa con elevación de la presión ocular. La pérdida de células ganglionares de la retina provoca alteraciones típicas en la función visual y en la papila del nervio óptico.

Su importancia radica en dos hechos:

- Su frecuencia. Aparece en aproximadamente un 2% de la población de más de 40 años y su prevalencia aumenta con la edad.

- El glaucoma crónico, el más frecuente, es en general una enfermedad asintomática hasta estadios terminales y, no tratada adecuadamente puede llevar a la ceguera. De hecho el glaucoma es la segunda causa de ceguera en los países desarrollados.

Su diagnóstico en estadios avanzados no resulta difícil para el oftalmólogo general, ya que normalmente va a encontrarse con una presión ocular elevada, una excavación papilar claramente glaucomatosa, y unas lesiones campimétricas típicas, que confirmarán el diagnóstico. Sin embargo, es en las fases iniciales del glaucoma, cuando suelen presentarse las dificultades diagnósticas, debido a que ninguna de las exploraciones que se realizan gozan de la sensibilidad y especificidad[1] suficiente como para confirmar por sí mismas y de forma indudable el diagnóstico.

Hoy en día, una de las principales vías de investigación en glaucoma está dirigida al diagnóstico precoz del mismo, ya que se sabe que cuando el campo visual muestra las primeras alteraciones ya se han lesionado de un 30 a 40% de las fibras

[1] La definición y cálculo de la sensibilidad y especificidad de una prueba diagnóstica, así como otros parámetros relacionados, se pueden consultar en el apéndice A de esta Tesis.

del nervio óptico. La perimetría computarizada es una prueba fundamental en el diagnóstico de la enfermedad, y aunque en la actualidad se tienen puestas las esperanzas en las nuevas técnicas de estudio de la papila, el campo visual constituye la prueba más importante para completar la exploración clínica.

En los inicios de la enfermedad ninguna de las exploraciones disponibles resulta definitiva para el diagnóstico. La tonometría puede hacer sospechar la existencia de un glaucoma cuando la presión intraocular determinada es superior a 21 mmHg, valor considerado como límite superior de la normalidad. Pero el hecho de que haya un porcentaje de la población normal con cifras superiores a ese límite, y que el 30% de los glaucomas no tienen la presión elevada, hace que su valor no permita ni confirmar ni descartar la enfermedad. La gonioscopia, una completa exploración ocular, las fotografías del nervio óptico y en ocasiones la curva tensional permiten llegar al diagnóstico de glaucoma, su sospecha o su tipo en la mayoría de los casos.

Junto al campo visual, el otro pilar básico del diagnóstico de glaucoma es la exploración de la papila y la capa de fibras nerviosas. El examen de la papila tiene sin embargo cierta dificultad dada la gran variabilidad en su forma y tamaño en la población normal. Si bien es cierto que la excavación de la papila aumenta de forma progresiva en el glaucoma debido a la pérdida de fibras nerviosas, no es menos cierto que la mayoría de las papilas normales tienen cierto grado de excavación y que un 10% de la población normal tiene una excavación mayor de 0.5, y algunos ojos normales tienen 0.8 y 0.9. Por todo ello, de nuevo no es posible considerar este parámetro de forma aislada, es necesario tener en cuenta todos los datos de la exploración del paciente y además considerar el tamaño de la papila y las características del anillo neuroretiniano.

La evaluación de la función visual es sin duda un elemento fundamental, si no el más importante del diagnóstico y seguimiento del glaucoma. El glaucoma altera distintos aspectos de la función visual que incluyen la sensibilidad al contraste, la detección de movimiento, la resolución y la percepción de los colores. Existen distintas pruebas funcionales útiles en el glaucoma, pero la prueba estándar es la perimetría de estímulo blanco sobre fondo blanco que evalúa la sensibilidad al contraste luminoso. Esta es la prueba utilizada en el presente estudio y a la que nos referiremos como campo visual o perimetría. No es una prueba específica para glaucoma y se altera también en otras enfermedades oculares y neurológicas que alteren la vía óptica a cualquier nivel. Por ello debe ser interpretada en el contexto general de la historia clínica y la exploración del paciente. Sólo así será útil en el diagnóstico del glaucoma.

En definitiva el diagnóstico del glaucoma debe realizarse considerando todos los datos de la exploración y la historia clínica, pero sobre todo debe basarse en el estado del nervio óptico y el campo visual. La integración de todos estos datos por un oftalmólogo experto, que aplique su experiencia adquirida es el mejor método para el diagnóstico de la enfermedad [141].

1.2 Objetivo

El objetivo general de este trabajo es conseguir un sistema de ayuda al diagnóstico precoz del Glaucoma Crónico Simple, que sea capaz de recoger y manejar gran parte de la experiencia del experto y alcanzar unos resultados lo más cercanos a éste posibles.

Dentro de este objetivo cabe destacar, como finalidad exclusiva, la realización de un **diagnóstico precoz**. Este sistema, por tanto, no pretende ayudar al diagnóstico del glaucoma en cualquier estado de progresión.

Nuestro sistema no sólo recogerá experiencia y conocimiento que provenga directamente del experto, sino que pretende incluir el conocimiento extraído a partir de los datos. Por este motivo, se usarán técnicas conexionistas además de simbólicas. Este conocimiento puede constituir una buena aportación para el experto o, al menos, una confirmación de sus conclusiones.

Aunque uno de los objetivos de una de las fases de desarrollo de este trabajo fue proporcionar al usuario una herramienta para diagnosticar cualquier tipo de glaucoma, nos centraremos en el diagnóstico precoz del Glaucoma Crónico Simple, puesto que es donde se encuentran las verdaderas dificultades de la investigación, y constituye el objetivo fundamental de la Tesis.

Este sistema será una ayuda importante en centros hospitalarios que no dispongan de especialistas, para el oftalmólogo general al que se le brinda la experiencia adquirida por una Sección de Glaucoma de un hospital Universitario, para el aprendizaje de aquellos oftalmólogos que estén especializándose en el tema, y como prueba de rastreo para detectar posibles casos de glaucomas tanto para oftalmólogos como para optómetras o médicos de familia.

Para conseguir este objetivo se van a aplicar diversas técnicas de aprendizaje y diagnóstico de Inteligencia Artificial (IA).

1.3 Técnicas empleadas

1.3.1 Sistemas basados en conocimiento

Hemos señalado que en los inicios de la enfermedad ninguna de las exploraciones disponibles resulta definitiva para el diagnóstico. Por tanto, el diagnóstico no puede basarse en una sola de las pruebas, sino en la interpretación global de todos los datos recabados en las consultas realizadas al paciente. Para esta interpretación lo más importante es la experiencia del oftalmólogo adquirida a lo largo de años de continuo contacto con este tipo de enfermos. Por todo ello, se ha visto conveniente utilizar esta técnica de Inteligencia Artificial para abordar el objetivo:

> "Los programas de IA que consiguen una capacidad a nivel de experto en la resolución de problemas mediante la reproducción de un cuerpo de conocimiento se denominan *sistemas basados en conocimiento* (SBC), o *sistemas expertos* (SE). A menudo, el término sistema experto se reserva para los programas cuya base de conocimiento contiene el conocimiento de expertos humanos, para diferenciarlos de los que adquieren el conocimiento de los libros de texto o de los no expertos. Más a menudo de lo habitual, los términos sistema experto y sistema basado en conocimiento se utilizan de forma sinónima." [41, 122].

Los SBC o SE se han aplicado a diversos campos del saber, entre ellos muy frecuentemente a la medicina y en alguna ocasión al diagnóstico del glaucoma [124, 134, 165, 166, 167].

Este trabajo se centra en los sistemas basados en reglas de producción. Por tanto, nuestra base de conocimiento estará compuesta por hechos y reglas, los cuales representan el conocimiento y la experiencia de los expertos en glaucoma de la Sección de Glaucoma del Hospital Clínico Universitario de Valladolid y del Instituto de Oftalmobiología Aplicada de la Universidad de Valladolid (IOBA).

El Sistema Experto desarrollado ha ido pasando por diversas **fases** a lo largo de este trabajo de investigación:

- Prototipo desarrollado utilizando un *shell* o herramienta de desarrollo de Sistemas Expertos: M1. Este software permite introducir la base de conocimiento sin preocuparse de programar los otros componentes principales del sistema (motor de inferencia e interfaz de usuario). Este Sistema Experto da como resultado un diagnóstico de la primera consulta realizada a un paciente. En este primer prototipo se incluyen factores de certeza.

- Sistema Experto programado en lenguaje C en su totalidad (base de conocimiento, motor de inferencia e interfaz de usuario), que realiza no sólo un diagnóstico de una primera consulta sino también el seguimiento del paciente. Además de ayudar en el diagnóstico del Glaucoma Crónico Simple, proporciona unas pautas para detectar otros tipos de glaucoma y aconseja un tratamiento para estos últimos.

- Nueva versión del Sistema Experto programado en C++ para MS Windows que añade a la anterior aplicación una interfaz gráfica más cómoda y agradable.

- Sistema Experto Borroso que transforma los parámetros que usa el oftalmólogo para realizar su diagnóstico en conjuntos borrosos, y recoge la información de uno de ellos, el campo visual, a partir de un sistema de clasificación basado en redes neuronales artificiales. Esta última fase reúne los dos enfoques de la IA: simbólico y conexionista.

1.3.2 Redes Neuronales Artificiales

Como ya se ha indicado, la perimetría computerizada es una de las pruebas más importantes para mejorar el diagnóstico del glaucoma. Esta prueba permite la exploración del campo visual (CV) y proporciona cada vez información más amplia y completa. Tanta información ha generado un nuevo problema: su interpretación. El análisis de los resultados es siempre difícil, ya que no existen criterios unánimes que definan un CV como glaucomatoso, pero aún es más complejo si lo que se pretende es realizar un diagnóstico precoz a través de lesiones incipientes del campo. Hay que tener en cuenta también que muchos de los procesos patológicos oculares más frecuentes (catarata, retinopatía diabética, retinopatía hipertensiva, miopía ...) pueden dañar el CV.

El descubrimiento de que las lesiones campimétricas en el glaucoma siguen generalmente unos patrones morfológicos, topográficos y evolutivos característicos, ha llevado en los últimos años a desarrollar sistemas expertos (Octosmart [14, 77, 90]), y a aplicar métodos estadísticos y técnicas de clasificación de patrones mediante redes neuronales artificiales [6, 56] para facilitar la ardua labor de interpretación del CV.

Durante los últimos años, las redes neuronales artificiales (RNA) han demostrado ser una herramienta muy flexible, capaz de proporcionar resultados satisfactorios en este tipo de problemas complejos de clasificación [111]. Esta técnica, que ya ha sido aplicada a diversos problemas dentro del campo médico y en particular al problema de clasificación del CV, va a ser abordada en este trabajo estudiando las soluciones ya aportadas y proponiendo otras nuevas.

1.3.3 Razonamiento Basado en Casos

El Razonamiento Basado en Casos (CBR) es otra técnica de IA que proporciona una metodología para reutilizar el conocimiento y la experiencia de forma eficaz. Cuando las personas resolvemos un problema, normalmente basamos la solución en otro similar que ya resolvimos en el pasado. Eso es en realidad lo que hace CBR: resolver problemas nuevos adaptando las soluciones dadas a otros resueltos con anterioridad [89, 129, 161].

Este enfoque ha sido también probado en este trabajo con vistas a facilitar la tarea de clasificación de campos visuales, y utilizar su resultado en el Sistema Borroso que realiza el diagnóstico final del paciente.

En el diagnóstico del CV, tenemos casos previamente diagnosticados y, por tanto, disponemos de una amplia información de cada caso que ya ha sido analizada. Podemos intentar extraer conocimiento a partir de esa experiencia almacenada, confrontando un nuevo CV de diagnóstico desconocido con esos casos que ya tenemos diagnosticados correctamente.

1.3.4 Lógica Borrosa

La lógica borrosa permite tratar información imprecisa, como *presión intraocular alta, relación excavación papila normal,* etc., en términos de conjuntos borrosos o difusos, que se combinan en reglas para definir acciones.

Los sistemas basados en lógica borrosa permiten utilizar fácilmente el conocimiento de los expertos, a veces ambiguo, formalizándolo de una forma realizable [111].

El conocimiento que manejan los expertos en el problema del diagnóstico del glaucoma es bastante impreciso, como se ha podido ver en el planteamiento del problema (por ejemplo, un mismo valor de presión intraocular puede ser normal para un paciente y alto para otro). Por esta razón, se propone un sistema borroso cuyo modo de razonamiento se asemeje más al del experto. Esto supondrá definir lo mejor posible los conjuntos borrosos de los parámetros en los que se basa el diagnóstico y dar la posibilidad de obtener varias soluciones ante un mismo problema.

Como siempre, **el sistema servirá de ayuda, pero no podrá sustituir al oftalmólogo**, que tendrá que tomar la decisión final sobre las acciones que se deberán seguir.

1.4 Esquema seguido en esta Tesis

En el capítulo 2 se expondrá con más detalle cómo se realiza el diagnóstico del glaucoma, y se presentará un resumen de los antecedentes de este trabajo. Se mostrará una recopilación de información de la mayor parte de los esfuerzos que se han hecho por facilitar la automatización tanto de los datos como del diagnóstico del glaucoma desde los inicios hasta la actualidad.

El capítulo 3 presentará los primeros sistemas expertos desarrollados en los inicios de este trabajo. Estos son la base sobre la que se ha ido construyendo el sistema final. Por su importancia para el resto de nuestra investigación serán examinados con cierto detalle.

Puesto que para el diagnóstico del glaucoma una prueba clave es la campimetría, se dedicará un capítulo entero a este tema: el capítulo 4. Primeramente se hará una exposición de las características exigidas a los datos recogidos para ser aceptados para realizar los experimentos, y un estudio estadístico de los mismos. Después se mostrarán con detalle todos los experimentos realizados con el fin de lograr la clasificación de los campos visuales, bien utilizando diversos tipos de redes neuronales artificiales, o bien usando razonamiento basado en casos.

El capítulo 5 presentará el sistema de diagnóstico final que se ha implementado utilizando lógica borrosa y en el que se reúnen los dos enfoques de la IA: simbólico, al implementar un sistema basado en reglas de producción, y

conexionista, al utilizar el mejor método de clasificación del CV mediante redes neuronales artificiales obtenido en nuestros experimentos e integrarlo en el sistema.

En cada uno de los tres capítulos anteriores se hará hincapié en los conceptos teóricos más importantes que se han aplicado en esta Tesis.

Por último, en el capítulo 6 se muestran las conclusiones alcanzadas en este trabajo, y los aspectos que se podrían abordar en investigaciones futuras.

2 Antecedentes. Estado del arte

2.1 Diagnóstico del glaucoma

Cuando definimos el glaucoma, al plantear el problema en el capítulo 1, dijimos que es una enfermedad ocular, con frecuencia crónica, que se caracteriza por presentar una degeneración progresiva del nervio óptico que en un 60-70% de los casos cursa con elevación de la presión ocular. La pérdida de células ganglionares de la retina provoca alteraciones típicas en la función visual y en la papila del nervio óptico.

En las fases iniciales del glaucoma es cuando suelen presentarse las dificultades diagnósticas, debido a que ninguna de las exploraciones que se realizan gozan de la sensibilidad y especificidad suficiente como para confirmar por sí mismas y de forma indudable el diagnóstico.

El diagnóstico del glaucoma debe realizarse, por tanto, considerando todos los datos de la exploración y la historia clínica, pero sobre todo debe basarse en:

A. Lesión anatómica del nervio óptico.
B. Lesión funcional del campo visual.
C. Aumento de la presión intraocular.

Las preguntas y pruebas a las que se somete al paciente que llega por primera vez a una consulta de glaucoma son las siguientes:

Anamnesis: Identificación de factores de riesgo

En primer lugar se hace una historia clínica del paciente, preguntándole principalmente sobre aquellos factores que pudieran influir con mayor peso en un posible diagnóstico de glaucoma.

Se consideran factores importantes los antecedentes familiares de glaucoma y la edad, ya que a partir de los 40 años la probabilidad de padecer la enfermedad es

mayor. Otros factores de menor importancia son: una miopía elevada, el padecer enfermedades cardiovasculares, y a veces se ha considerado como posible factor el padecer diabetes aunque no está claro que lo sea.

Tonometría

Es la medida de la presión intraocular (PIO).

Lo primero que necesitamos para poder considerar un valor de presión ocular como posiblemente normal o patológico es conocer cuál es la presión intraocular normal. Ésta se puede definir como aquella que no produce daño glaucomatoso en la cabeza del nervio óptico. Pero no se puede establecer un límite determinado de PIO que se considere normal debido a que todos los ojos no responden del mismo modo a determinados niveles de PIO.

Si la distribución de la PIO en la población general fuese Gaussiana podríamos establecer el límite superior normal de la PIO en 21 mmHg, cifra utilizada por muchos oftalmólogos para establecer en clínica el límite de la normalidad, por debajo de cuyo valor, desde el punto de vista estadístico debería quedar incluida el 95% de la población normal.

Pero la distribución de la PIO en la población general no es Gaussiana, sino que es bimodal, es decir, presenta cierta asimetría hacia los valores más altos. Esto plantea un problema diagnóstico importante, ya que a partir de ciertos valores de PIO, nos vamos a encontrar con una mezcla de poblaciones, normales y glaucomatosos. Ahora bien, como en clínica se necesita establecer un punto de referencia para poder clasificar a los pacientes, de forma generalizada se ha adoptado como límite superior de normalidad de la PIO el valor de 21 mmHg.

Por lo tanto, debemos tener en cuenta, que el valor de la PIO, generalmente, lo único que nos va a permitir es, ante una cifra determinada, considerar que el paciente puede ser normal si tiene menos de 21 mmHg, o que tiene mayor probabilidad de ser glaucomatoso si está por encima de ese valor.

Sin embargo, aun estando baja la PIO, no debemos olvidar que existe un cuadro clínico conocido con el nombre de Glaucoma de Tensión Normal (GTN), y que, aunque es cierto que cuanto más elevadas sean las presiones, mayor será el riesgo de padecer glaucoma, sólo el 1% de las personas con PIO mayor de 21 mmHg desarrolla glaucoma cada año. Por ello, si la PIO es mayor de 21 mmHg pero no existen lesiones de campo visual, tal paciente es catalogado como Hipertenso Ocular (HTO).

Este último concepto para algunos resulta peligroso, ya que puede llevar a una falsa sensación de seguridad tanto por parte del médico como del paciente. Además, no hay que olvidar que muchos glaucomas pasan por una fase que puede denominarse tonométrica, es decir, por aquella en la cual sólo existe un aumento de la PIO, pero falta que transcurra el tiempo para que se produzcan lesiones campimétricas y pase a ser considerado glaucomatoso.

Por todo lo que hemos visto hasta ahora, la decisión sobre la normalidad de una determinada cifra de PIO estará supeditada a otras exploraciones.

Relación Excavación/Papila

La siguiente exploración es la de la papila del nervio óptico. En el glaucoma se produce una atrofia de fibras nerviosas con desaparición de las mismas dejando en la papila un espacio vacío denominado excavación.

Sin embargo el determinar si una excavación papilar, aun siendo grande, es glaucomatosa, no resulta fácil dada la existencia de excavaciones fisiológicas. Estas últimas son más grandes cuanto mayor es el tamaño de la papila y este tamaño varía en la población normal entre 0.8 y 5 mm^2.

Se sabe que el número de fibras que proceden de las células ganglionares de la retina es más o menos constante, pero sucede que el canal escleral puede tener un diámetro variable de unas personas a otras. Por tanto, cuando este canal tiene un diámetro pequeño, las fibras pasarán apelotonadas, ocupando todo su espacio, con lo cual la papila tendrá un color rojizo más o menos uniforme en toda su superficie, al no tener en su centro ningún espacio vacío o excavación. En cambio cuando el canal escleral es amplio, las fibras nerviosas pasarán a través del mismo holgadamente, dejando un espacio central sin rellenar, a través del cual queda expuesta la lámina cribosa. Esta zona de palidez central recibe el nombre de **Excavación fisiológica**, cuya posible existencia es preciso tener en cuenta al estudiar al enfermo glaucomatoso.

Esta excavación fisiológica tiene una serie de características que nos pueden ayudar a distinguirla de las glaucomatosas, aunque no siempre resulta fácil hacer esta diferenciación:

- Su tamaño suele ser generalmente pequeño, la mayoría de las veces inferior a 1/3 del diámetro papilar, aunque excepcionalmente, en el 1-2% de la población, puede ser igual o mayor a 2/3 del mismo.

- Generalmente suelen ser simétricas entre los dos ojos del mismo individuo, observándose asimetría importante sólo en el 1% de la población normal.

- Los bordes de la excavación y del disco tienden a ser paralelos. Esto hace que la excavación fisiológica sea más o menos redondeada y los distintos sectores del anillo tienen distintos grosores en un determinado orden: inferior > superior > nasal > temporal.

- El límite entre la excavación y el reborde neural suele ser bastante preciso.

- Aun siendo las excavaciones amplias, nunca tocan el borde papilar.

La excavación glaucomatosa se produce como consecuencia de la destrucción progresiva de las fibras nerviosas a nivel del reborde neural.

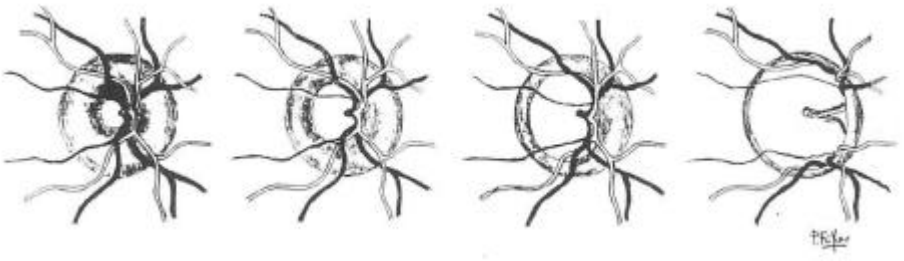


Figura 2.1: Esquema en el que se muestra la forma de crecimiento más frecuente de la excavación glaucomatosa [104].

Constituyen signos de alerta importantes el aumento de la excavación, su excentricidad, su ovalización vertical, las hemorragias en astilla o las modificaciones del trayecto de los vasos.

A diferencia de la excavación fisiológica, en la glaucomatosa a veces resulta difícil establecer los límites de la misma debido a una variación gradual entre la zona atrófica y la sana.

El método más corrientemente utilizado en clínica para medir la excavación es el denominado cociente excavación/papila o E/P, que relaciona el diámetro de la excavación (E) con el diámetro de la papila (P).

Para hallar este cociente, en un diámetro determinado se divide a la papila en una serie de partes, (10 ó 6) y se observa cuántas de estas partes abarca la excavación. Ésta es una forma de medir indirectamente el daño papilar.

De todo lo anteriormente expuesto se deduce que, al menos en estadios iniciales del glaucoma, no se puede establecer un diagnóstico basándose exclusivamente en la observación de las alteraciones de la papila.

Campimetría

La campimetría permite detectar en el glaucoma las alteraciones del campo visual originadas como consecuencia de la lesión de las fibras nerviosas a nivel de la cabeza del nervio óptico.

Es una exploración fundamental en el glaucoma, no sólo para el diagnóstico de la enfermedad, sino también para controlar su evolución, y poder estar seguros de la posible eficacia de un tratamiento.

Una de las principales características de las lesiones campimétricas es su forma de evolucionar:

- Inicialmente, aparecen defectos en el campo visual periférico en forma de retracciones concéntricas, y alrededor del punto de fijación -no afectándose, por este motivo, la agudeza visual del enfermo- los escotomas

aislados, que posteriormente, al unirse entre sí, dan lugar al escotoma arciforme de Bjerrum. Este último puede terminar en una línea horizontal, límite entre el campo visual superior e inferior, formando el denominado escalón nasal de Rönne.

- La progresión de las lesiones acabará aislando el campo central del periférico, dando lugar al llamado campo biinsular. En este estado avanzado se sigue conservando el campo central, y por ello el paciente sigue viendo bien, y no se da cuenta de la enfermedad que padece.

- En un estado terminal acaba perdiendo el campo visual central, quedando sólo una semiluna temporal y por ello su agudeza visual disminuye de forma rápida, y es cuando el paciente refiere alguna sintomatología, pero desgraciadamente tardía, pues se está quedando ciego, y el proceso es irreversible.

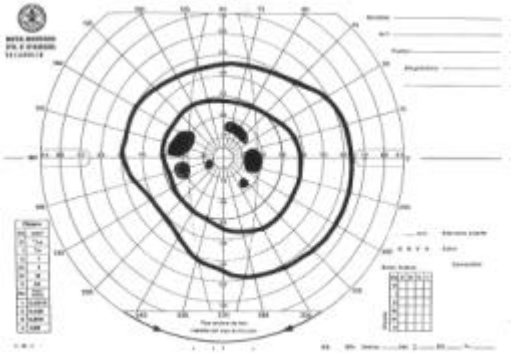


Figura 2.2: Escotomas aislados en el área de Bjerrum [104].

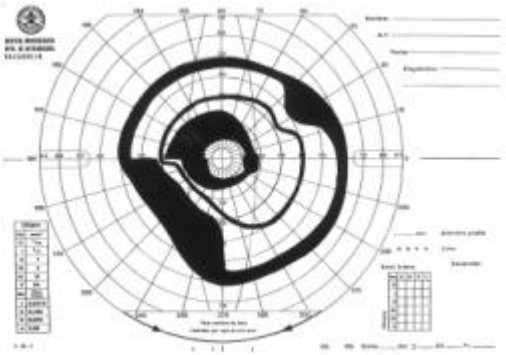


Figura 2.3: Escotomas arciformes paracentrales, asociados a estrechamientos periféricos [104].

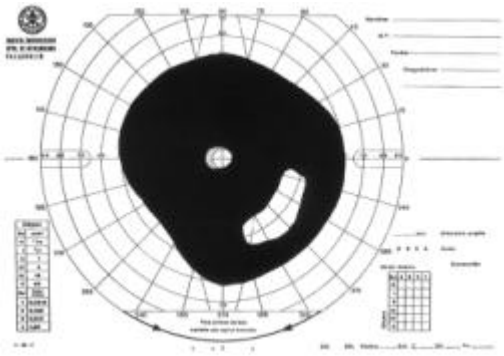


Figura 2.4: Campo biinsular [104].

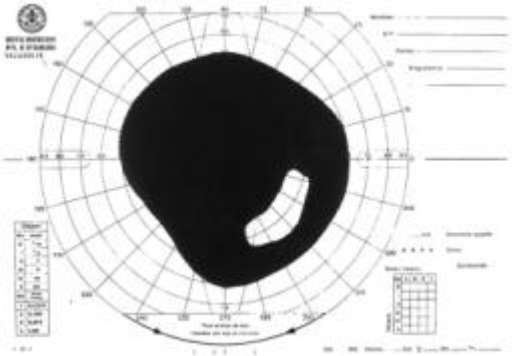


Figura 2.5: Lucerna temporal [104].

Si la presión intraocular es mayor de 21 mmHg, con el resultado de estas tres exploraciones (tonometría, oftalmoscopía, campimetría) el oftalmólogo puede dar un primer diagnóstico bastante acertado del paciente. Sin embargo, si la presión intraocular es baja o está próxima al límite de la normalidad, será en la mayor parte de los casos imprescindible realizar una curva tensional.

Curva tensional

Al igual que muchas otras funciones del organismo, la PIO no es un parámetro estable, sino que varía según las horas del día, siendo habitualmente mayor al despertar y menor en las primeras horas de la noche.

Al margen de cuál sea la causa de estas variaciones de la PIO, de todas las pruebas ideadas para una mejor comprensión y un más correcto diagnóstico de los Glaucomas Crónicos Simples (GCS), la curva tensional es quizá la única que mantiene en estos momentos plena validez, y puede incluso tener un valor de predicción del curso futuro que la PIO tomará en un sujeto determinado.

La ejecución de una curva tensional permite el establecimiento de diferentes parámetros, cada uno de los cuales puede resultar decisivo a la hora de establecer un diagnóstico.

Básicamente, tres son los valores que han sido objeto de atención por parte de los diferentes autores:

- la amplitud de las oscilaciones
- el valor de la PIO máxima
- el tipo de morfología

De todos ellos, el último es el que resulta menos valorable, por lo menos en lo que respecta a poder realizar un diagnóstico.

Respecto al valor de la PIO máxima, nos tendremos que regir por los mismos patrones establecidos para la PIO estática, que establece su límite superior en 21 mmHg. En este punto quizá la única matización importante sea resaltar el hecho de que un porcentaje muy elevado de pacientes tienen sus valores presionales máximos fuera del horario clínico habitual, lo que conduce de hecho a la demora de su diagnóstico, o a aceptar como efectivos tratamientos que no lo son.

Pero el dato más significativo de esta exploración es la amplitud de las oscilaciones presionales, es decir, la diferencia existente entre la PIO máxima registrada y la mínima. Se consideran normales las variaciones inferiores o iguales a 5 mmHg, dudosas las comprendidas entre 6 y 7 mmHg, y a partir de ese valor como netamente patológicas.

La curva tensional no es una exploración que se practique de rutina a todos los pacientes que llegan a la sección de glaucoma, sino sólo en aquellos en los que pueda aportar datos de interés, como son:

1. **Casos de diagnóstico dudoso**. Pacientes con campo visual y relación E/P normales o con alteraciones no claramente glaucomatosas. La curva tensional permitirá:
   - Descartar un glaucoma.
   - Confirmar el diagnóstico de glaucoma.
   - Catalogar al paciente como HTO.

2. **Pacientes con alteraciones de campo visual y papila típicamente glaucomatosas pero con la PIO, determinada durante el horario clínico, dentro de la normalidad**. La curva ayudará a establecer el diagnóstico de:
   - GCS, o de
   - GTN.

3. **Pacientes glaucomatosos sometidos a algún tipo de terapéutica y que, aunque desde el punto de vista presional se encuentran controlados, hay progresión de las lesiones campimétricas**.

Tonografía

La tonografía, desarrollada por Grant en 1950, es una de las técnicas de exploración del glaucoma más controvertidas en la actualidad.

El principio básico de los tonógrafos utilizados habitualmente en clínica, es producir un incremento de presión, mediante la colocación de un tonómetro de indentación. Este aumento de presión, va a provocar un mayor flujo de salida de humor acuoso lo que determinará una caída progresiva de la presión ocular.

La tonografía consiste en calcular el valor de ese descenso, que se supone es inversamente proporcional a la resistencia que se ofrece a la salida del acuoso, o a su inversa, la facilidad de drenaje o simplemente «C».

Esta prueba no tiene mucho valor clínico. Los valores patológicos tonográficos no concluyen el diagnóstico de glaucoma, únicamente alertan al oftalmólogo a controlar al paciente de manera más estricta.

2.2 Los comienzos en la investigación sobre el diagnóstico automático del glaucoma

Uno de los mayores problemas en medicina clínica e investigación ha sido el almacenamiento y recuperación de los datos. A finales de los años sesenta comenzó algún intento de desarrollar un sistema de base de datos para oftalmología. Desafortunadamente, muchos de estos intentos fueron tan sólo proyectos piloto o se quedaron muy pronto obsoletos debido a los rápidos avances que se producen en el campo de la informática.

En 1969, Metz y col. [113] realizaron un estudio para informatizar los datos de pacientes en una consulta de glaucoma. Las ventajas del sistema que proponen

son principalmente que se reduce del tiempo que utiliza el especialista en revisar la historia de un paciente, se mejora la calidad de la asistencia al paciente al poder acceder más fácilmente a los resultados de las últimas pruebas y poder hacer un seguimiento más eficaz del paciente, y se consigue una base de datos útiles para la investigación.

En 1977, Greenfield y col. [60] modificaron el sistema de base de datos MISAR, que había sido desarrollado en el Beth Israel Hospital de Boston, para uso oftálmico. Se introdujeron en el sistema aproximadamente 1165 registros de pacientes del Centro de Glaucoma de la Universidad de Washington. Las principales funciones de esta base de datos fueron, lógicamente, las mismas que en 1969 habían pretendido Metz y col. [113] con su sistema: almacenar de forma ordenada los datos, facilitar su recuperación, y permitir la investigación para mejorar el funcionamiento de ese Centro de Glaucoma.

Estos primeros intentos de informatizar datos médicos relativos al glaucoma fueron importantes como un comienzo para plantearse en años sucesivos utilizar métodos computerizados para ayuda al diagnóstico de esta enfermedad.

El diagnóstico ha sido a menudo considerado un problema de **reconocimiento de patrones**: encontrar a qué clase pertenece un patrón desconocido dentro de un conjunto de clases o hipótesis especificadas de antemano. La aplicación de estos métodos requiere un conjunto de características o medidas que pueden ser usadas para clasificar el patrón desconocido. Específicamente para diagnóstico médico, las características son un conjunto de síntomas, o resultados de pruebas de laboratorio descubiertos por el médico, y las clases específicas son las categorías de enfermedades. El problema consiste en decidir a qué categoría de enfermedad pertenece el patrón desconocido con unas características de un paciente concreto. Ejemplos de categorías de enfermedades podrían ser los diferentes tipos de enfermedades del corazón, mientras las características pueden incluir pruebas como la presión de la sangre o síntomas como dolor de cabeza [165].

En la década de los 70 Weiss y Kulikowski [165, 166, 167] desarrollaron en la Universidad de Rutgers un sistema experto para ayuda en el diagnóstico y terapia del glaucoma y otras enfermedades oculares llamado CASNET (Causal ASociational NETwork). Los procedimientos de razonamiento interpretan los hallazgos de un paciente particular en términos de un modelo de una red causal-asociativa (CASNET) o red semántica que caracteriza los mecanismos fisiopatológicos y el curso clínico de enfermedades tratadas y no tratadas. Las aportaciones de este programa son: la generación de interpretaciones complejas a partir de un modelo cualitativo de un proceso de enfermedad, el razonamiento sobre el manejo del seguimiento detallado de un paciente, la incorporación de opiniones alternativas de expertos sobre los temas bajo debate, y su prueba y actualización mediante una red colaborativa informática de investigadores en glaucoma.

Esta red se descompone en cuatro niveles o «planos» (Figura 2.6). En el plano de observaciones, hay nodos que representan las manifestaciones del paciente. Estos nodos se relacionan con los del plano de estados fisiopatológicos, que representan hipótesis elementales, y que están unidos entre sí por relaciones causales con factores de confianza asociados. En el plano de enfermedades se determinan los diagnósticos en función de los estados calculados y, en función de ellos, se establecen los tratamientos en el último plano.

CASNET dialoga con el usuario planteando una serie de preguntas sobre el paciente y sus manifestaciones. A partir de las respuestas, calcula sucesivamente, para cada estado fisiopatológico, un índice indicativo de la medida en que el sistema estima que ese estado está presente. Inicialmente, todos los estados están indeterminados y, conforme procede el sistema a considerar las respuestas, los va señalando como descartados o confirmados, según el índice sobrepase ciertos umbrales.

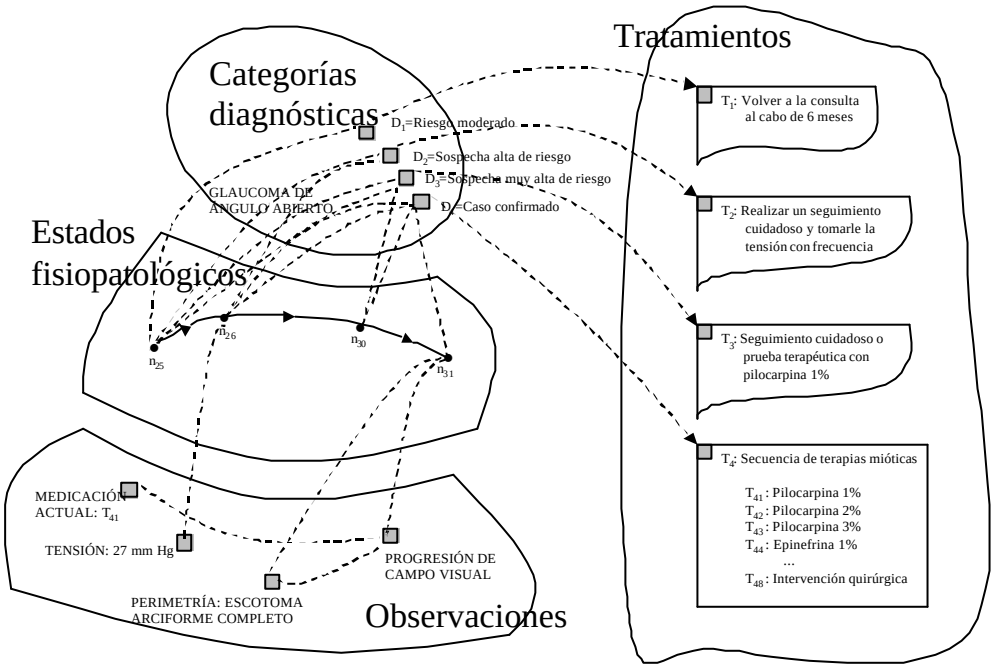


Figura 2.6: CASNET. Los cuatro niveles de la red semántica [165, 166 y 167].

El sistema incluye una estrategia para plantear la siguiente pregunta en función del coste de la exploración o de la prueba, y de la probabilidad de que la respuesta permita confirmar o descartar estados.

Este sistema de consulta sobre glaucoma fue sometido a una intensiva evaluación durante el Congreso de 1976 de la Academia Americana de Oftalmología y Otorrinolaringología. El programa de consulta fue validado estudiando los resultados y recomendaciones que daba frente a casos clínicos de pacientes y contrastándolos con la opinión de un jurado de expertos del Simposium de

Glaucoma. El jurado dio varias opiniones acerca de los casos, y en casi todos el programa incluía en sus alternativas la interpretación principal dada por el jurado. Los oftalmólogos rellenaron una encuesta y se les animó a examinar el programa con casos difíciles. Los resultados mostraron un 95% de aceptación de la habilidad clínica en la muestra que se examinó.

El sistema CASNET nunca llegó a usarse mucho. El dominio del problema fue extendido a muchos problemas oftalmológicos adicionales que no estaban cubiertos completamente en el modelo original. Sin embargo, no hubo ninguna demanda clara de un sistema experto en este área. Una posible causa puede ser que en ese momento, a mediados de los 70, muchos médicos tenían prejuicios respecto al uso de ordenadores en una consulta. Esto todavía persiste hoy en día aunque en menor grado.

2.3 Interpretación del campo visual

A partir de los años 70 muchos oftalmólogos han dirigido sus investigaciones hacia la interpretación del campo visual, que, como ya se ha señalado, es una de las pruebas fundamentales para lograr un diagnóstico precoz del glaucoma.

2.3.1 Estudio sobre el deterioro del campo visual en el glaucoma

Para el desarrollo de este apartado nos hemos basado principalmente en la Tesis Doctoral de Antón [6].

Se ha estudiado cómo se va deteriorando el campo visual (CV) y parece ser que la mayor parte de los autores confirman que se producen lesiones campimétricas en el glaucoma que también pueden aparecer en otras enfermedades; además en el glaucoma puede aparecer prácticamente cualquier tipo de lesión en el campo visual, es decir, las lesiones pueden tener muy diversas formas y estar situadas en cualquier parte del CV. Sin embargo, ciertos tipos de defectos y localizaciones se repiten con más frecuencia y tienden a progresar en extensión o intensidad de forma característica.

La mayor parte de estos autores han encontrado lesiones en el hemicampo superior con más frecuencia que en el inferior [27, 38, 68], excepto Aulhorn [12], quien observó que ambos hemicampos se alteraban con la misma frecuencia aunque la localización variaba considerablemente.

Básicamente hay dos tipos genéricos de lesiones campimétricas, que pueden y **suelen coexistir en las fases iniciales** del glaucoma: la **disminución generalizada de la sensibilidad**, **y** los escotomas o **defectos localizados**. La primera se produce por una pérdida difusa de axones, mientras que la segunda se debe a la lesión de un grupo de fibras nerviosas contiguas. Los escotomas se reconocen más fácilmente que las pérdidas difusas; sus bordes pueden ser más o menos abruptos, y pueden ser calificados como escotomas relativos, si existe una

disminución de la sensibilidad umbral, o absolutos, si la retina en esa zona no es capaz de detectar los estímulos de máxima intensidad luminosa que emite el campímetro.

Las zonas que se distinguen en el campo visual se pueden observar en la Figura 2.7.

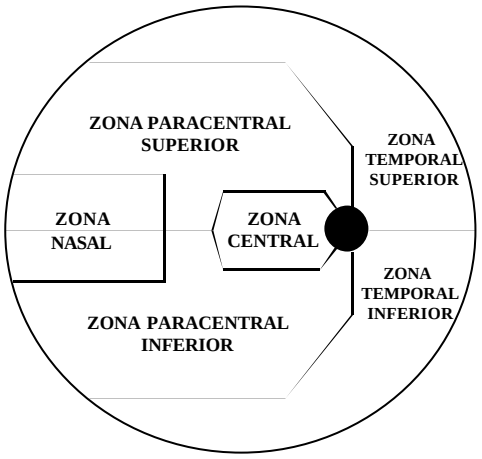


Figura 2.7: Zonas del campo visual.

Éstas son las lesiones que suelen aparecer en los pacientes con glaucoma en el orden en que habitualmente se producen:

- **Disminución generalizada de la sensibilidad**. Se expresaba en la perimetría manual cinética como una reducción concéntrica de las isópteras y puede ser la primera manifestación campimétrica de la enfermedad, pero también ocurre en sujetos con edad avanzada, miosis u opacificación de medios.

- **Escotomas paracentrales aislados**. Son disminuciones focales de la sensibilidad en la región paracentral del CV. Si la enfermedad avanza, se extienden, se profundizan y se unen para formar escotomas arciformes. Dado que la perimetría computerizada (PC) explora puntos separados entre sí por un intervalo de unos 6°, en ocasiones pequeños defectos podrían teóricamente no detectarse, por lo que en pacientes sospechosos puede estar indicada la exploración de ciertas zonas con un mayor número de puntos. Sin embargo, hay estudios que demuestran que la separación de los puntos explorados por los campímetros modernos es adecuada y suficiente para la detección incluso de los defectos precoces.

- **Defectos arciformes**. Son lesiones con disminución de la sensibilidad, que se extienden desde la mancha ciega en forma de arco alrededor de la zona de fijación hasta la parte nasal del rafe horizontal. Pueden llegar a afectar a la periferia nasal del campo.

- **Escalón o depresión nasal**. Defecto de sensibilidad en el hemicampo nasal que respeta el rafe horizontal. Puede ser central, periférico o ambos, y puede ser aislado o asociado a defectos paracentrales.

- **Escalón o depresión temporal**. Aunque es menos frecuente, pueden aparecer defectos en cualquiera de los lados del rafe horizontal en la zona del CV temporal a la mancha ciega.

- **Aumento o extensión de la mancha ciega**. Si la extensión se produce de manera arciforme, la lesión recibe el nombre de "escotoma de Seidel".

- **Campo bi-insular e insular**. En los estadios terminales del glaucoma en los que sólo persisten algunas fibras nerviosas en el haz papilo-macular o la región nasal de la retina, el paciente sólo conserva pequeñas "islas" de visión en el CV central o temporal.

Sin embargo, no todos los defectos que se acaban de describir son igualmente específicos del glaucoma. Así, la disminución generalizada de la sensibilidad y el aumento del tamaño de la mancha ciega, aunque aparecen frecuentemente en el glaucoma, son completamente inespecíficos y pueden presentarse también como consecuencia de la presencia de opacidades de medios, de una miosis, de defectos de refracción o de la edad avanzada. Por otro lado, los defectos en la zona de la mancha ciega pueden ser debidos a angioescotomas.

Los defectos más característicos del glaucoma son los que aparecen en la zona paracentral del CV (5° a 15° ó 10° a 26°). Las lesiones que con más frecuencia encontró Hart [66] al explorar el CV de pacientes en estadios iniciales de glaucoma mediante perimetría manual fueron: escotomas arciformes con o sin continuidad con la mancha ciega (50%), defectos paracentrales (41%), aumento de la mancha ciega (35%), reducción concéntrica de las isópteras (30%) y defectos en el sector temporal (3%). Más recientemente, en la Universidad de Yale se revisaron los resultados obtenidos con la PC en más de mil casos con glaucoma, y los defectos que se encontraron fueron los siguientes: pérdida difusa de sensibilidad (55%), escotomas arciformes superiores (45%), escotomas nasales superiores (22%), escotomas arciformes inferiores (12%), escotomas nasales inferiores (12%), CV terminal (8%), reducción concéntrica periférica (7%), aumento de la mancha ciega (5%), y defectos en el sector temporal (4%).

El CV periférico aporta una información muy limitada al diagnóstico y seguimiento del glaucoma. Por ello, en general, se acepta que en el estudio de esta enfermedad lo más adecuado es utilizar programas campimétricos que exploren los 26°-30° centrales. Sin embargo, entre un 4% y un 11% de los pacientes con glaucoma incipiente y CV central normal pueden presentar lesiones campimétricas en la periferia.

Mikelberg y col. [115] estudiaron el patrón de progresión de los defectos de campos visuales en glaucoma. Se dice que hay progresión del campo visual

cuando un escotoma aumenta en área o profundidad o cuando hay un escotoma nuevo. Esta progresión de las lesiones campimétricas glaucomatosas puede llegar a provocar la ceguera, pasando por un estadio terminal en que persiste sólo un islote de visión central o temporal. Mikelberg [115] encontró un empeoramiento del CV en el 87% de los casos tras un seguimiento de 8 años, y analizó las distintas formas de evolución de los escotomas: en el 68% de los CV aumentó su profundidad, en el 52% su tamaño, y en el 49% aparecieron nuevos defectos.

La velocidad de progresión de las lesiones depende de muchos factores: el tiempo de evolución [115], el estadio de la enfermedad, la presión intraocular, el control terapéutico, y otros no bien conocidos y probablemente de naturaleza vascular. La PIO es un factor importante, al menos en los denominados "glaucomas presionales". Algunos autores calcularon el tiempo que transcurría entre las lesiones incipientes y la pérdida absoluta del CV, y comprobaron que es inversamente proporcional al nivel de la PIO (14 años si la PIO es de 21 a 25 mmHg, 6 años con 20-30 mmHg, y 2 años con PIO mayor de 30 mmHg).

Pero la presión parece ser poco o nada importante en otros pacientes con defectos campimétricos y lesiones en la papila o capa de fibras del nervio óptico compatibles con el glaucoma pero que tienen valores de PIO que no superan los 23-25 mmHg. Y en los últimos años han surgido múltiples evidencias de la existencia de factores etiopatogénicos, distintos de la PIO, en el desarrollo del glaucoma.

Desde el punto de vista de la patogenia, se habla de "glaucomas presionales" y glaucomas no dependientes de la presión o "glaucomas vasculares". Desde el punto de vista clínico, se describen como "glaucomas de tensión elevada" (GTE) y "glaucomas de tensión normal" (GTN). En principio, se había aceptado que los defectos campimétricos que presentaban estos dos grupos de pacientes glaucomatosos eran similares, y muchas veces superponibles; sin embargo, en los últimos años se han encontrado diferencias que pueden apoyar la doble teoría etiopatogénica. Se ha observado que en el GTE es más frecuente la disminución difusa de la sensibilidad y, en cambio, en el GTN son más típicos los defectos localizados. Otros autores, aun a pesar de observar dicha tendencia, no han encontrado diferencias estadísticamente significativas. Por último, en otros trabajos se ha comentado que los escotomas son más profundos y con pendientes más pronunciadas en el GTN [6].

2.3.2 Perimetría computerizada

El estudio del CV en el diagnóstico del glaucoma se inició, hace más de cien años, de forma muy rudimentaria con una hoja de papel y unos puntos dibujados en ella, que servían de estímulos. Con el tiempo se han diseñado distintos instrumentos que facilitan el estudio del CV.

La perimetría manual (PM), sobre todo la cinética con perímetros de cúpula, permitió durante décadas una buena aproximación al estudio del CV para el diagnóstico y seguimiento del glaucoma. Sin embargo, en los años 60 se inició el

desarrollo de la perimetría computerizada (PC), que sin duda constituye hoy el método más adecuado para estudiar el CV de los pacientes glaucomatosos.

De los distintos sistemas perimétricos computerizados, el que se ha convertido en estándar es la perimetría estática, que explora la sensibilidad umbral al contraste entre el fondo homogéneamente iluminado de una cúpula y diversos estímulos luminosos proyectados sobre dicho fondo, con intensidad, tamaño y duración variable. Sus ventajas sobre la perimetría manual son múltiples; entre ellas destacan el mejor control de los parámetros de exploración (características del estímulo, localización de los puntos explorados y posición del globo ocular), el hecho de que los resultados sean expresados de forma cuantitativa, sean comparados con los valores normales para la edad del paciente y se calculen índices generales y de la fiabilidad de la prueba. Todo ello permite explorar el campo visual mediante la PC estática y umbral, de forma exhaustiva, en menos tiempo y sobre todo con una mayor sensibilidad que con la perimetría manual cinética.

Pero no debe pensarse que la PC ha resuelto todos los problemas del estudio del CV. Es cierto que ha aumentado de forma considerable la cantidad, precisión y sensibilidad de la información. Sin embargo, existen muchos factores que pueden modificar de forma significativa sus resultados, dificultando la valoración del daño verdaderamente producido por el glaucoma, y sigue siendo necesaria la interpretación subjetiva de los datos que ofrece, teniendo en cuenta dichos factores.

Entre ellos se encuentran los defectos de refracción, la experiencia perimétrica del paciente, el tamaño de la pupila, la fluctuación de la respuesta individual, la variabilidad entre distintos sujetos, la presencia de opacidades en el cristalino y cualquier otra patología oftalmológica (retinopatías diabética o hipertensiva, miopía...) o neurológica que pueda afectar al CV.

Además, no existen parámetros ni criterios definidos que permitan calificar un CV como glaucomatoso o no glaucomatoso. La decisión final está siempre en manos del experto que interprete los datos y pueden existir grandes diferencias en la evaluación que distintos oftalmólogos hacen de un mismo CV.

Existen actualmente en el mercado varios perímetros computerizados que presentan algunas diferencias entre ellos. Los datos de campos visuales que se han usado en los experimentos de esta Tesis se han obtenido utilizando el perímetro Octopus.

Dentro de las estrategias de exploración del perímetro computerizado Octopus existe un programa específicamente diseñado para el glaucoma, denominado G1. Este programa realiza una exhaustiva exploración de la sensibilidad de 59 puntos localizados estratégicamente dentro de los 30 grados centrales del campo visual, zona donde, como se ha indicado anteriormente, aparecen con más frecuencia las lesiones iniciales glaucomatosas. Utiliza la estrategia de perimetría umbral completo, la cual, pretende averiguar el nivel

máximo de sensibilidad de cada punto de la retina, investigando cual es la luz menos intensa capaz de producir una sensación visual.

Para ello, selecciona uno de los puntos de la retina e inicialmente varía la intensidad del estímulo de 0,4 en 0,4 unidades logarítmicas (0,1 unidades logarítmicas equivalen a 1 decibelio), incrementándola hasta que el punto luminoso comienza a ser visto por el paciente, o por el contrario, disminuyéndola hasta que deja de ser percibido. De esta forma el umbral del punto explorado quedará comprendido entre los valores del estímulo visto y del no visto, en un intervalo de 4 decibelios. Posteriormente los incrementos o disminuciones del estímulo serán de 0,2 unidades logarítmicas, de tal manera que utilizando una media de cinco exposiciones por punto retiniano quedará determinado su umbral de sensibilidad con un error inferior a un decibelio.

Los resultados aparecen reflejados en la denominada tabla numérica (Actual dB) (Figura 2.8), en la cual, las cifras expresan la sensibilidad exacta en decibelios de cada punto explorado. Sin embargo, la que mayor interés diagnóstico tiene es la tabla corregida según la edad (Diferencia dB) (Figura 2.9), ya que en ella, cada punto es comparado con el valor que debería tener según la edad del paciente, a partir de una base de datos que posee la memoria del perímetro. De tal forma que, si es normal, viene reflejado con una cruz, y en caso de ser patológico, si el defecto es relativo, aparece un número, que indica la magnitud del mismo, en decibelios, y cuando la pérdida es absoluta, es decir, la visión en un punto es nula, éste aparece con un cuadro negro.

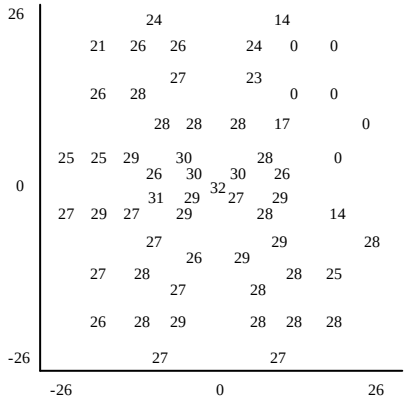


Figura 2.8: Actual (dB). Tabla numérica del programa G1 del Octopus 500 [104].

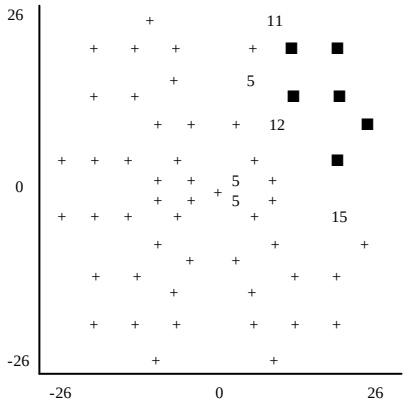
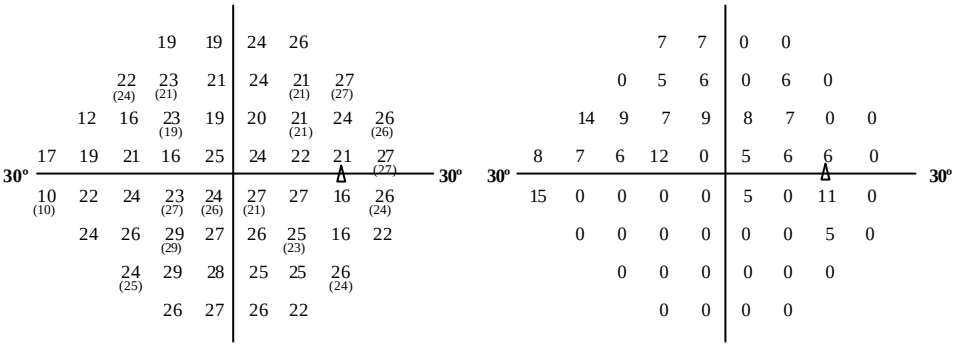


Figura 2.9: Diferencia (dB). Tabla corregida según la edad del programa G1 del Octopus 500 [104].

Otro de los perímetros usados para explorar el CV es el Analizador Humphrey. Éste ofrece varios patrones de puntos (central 24-1, 24-2, 30-1, 30-2, etc.) que pueden analizarse mediante diversas estrategias (test de umbral completo, de umbral rápido, etc.) [1]. Existen diversas diferencias respecto al perímetro Octopus como son la intensidad de iluminación del fondo y el tamaño de la cúpula; y aunque la forma de presentar al paciente los estímulos es similar,

sin embargo el número de puntos analizados y la posición de éstos es distinta. Por otra parte, estos dos últimos factores varían también en el analizador Humphrey según el programa de exploración de patrones que se ejecute.

Por ejemplo, si se realiza el programa central 24-2 podríamos obtener las siguientes tablas numéricas (Figura 2.10) que se interpretan de la misma forma que explicamos para el Octopus.



(a) Tabla numérica **(b)** Tabla corregida según la edad

Figura 2.10: Tablas de datos obtenidas con el programa central 24-2 en el analizador de campos visuales Humphrey [1].

Se puede observar que mediante este programa obtenemos información sobre 54 puntos del campo visual, a diferencia del programa G1 del perímetro Octopus que explora 59 puntos situados de forma diferente en los 30 grados centrales del campo visual (Figuras 2.8 y 2.9).

También se han desarrollado sistemas que permiten tratar computacionalmente los datos que proporcionan los perímetros. Hart y col. [64, 65] desarrollaron en 1981 un sistema que proporciona datos de campos visuales en tiempo real. Los datos de los campos son digitalizados mediante tabletas gráficas magnéticas, y se pueden recuperar en tiempo real y mostrar gráficamente en terminales de vídeo. Sponsel y col. [147] desarrollaron en 1984 un sistema automatizado para monitorizar pacientes con glaucoma. Ese sistema incorporó métodos nuevos y eficientes para transferir datos de la perimetría, recogidos tanto de forma manual como automática, a un formato digital. El sistema permitía reducir considerablemente la carga administrativa en las clínicas oftalmológicas. Filligoi y col. [43] en 1992 examinaron una técnica desarrollada recientemente para la adquisición automática de datos de pérdidas de campos visuales. Los algoritmos para la adquisición y procesamiento del campo visual central fueron implementados sobre un ordenador personal. Se diseñaron herramientas de procesamiento de imágenes para extraer información relativa al tamaño, forma, posición y número de escotomas del campo visual.

Actualmente los perímetros incorporan sus propios programas para recoger la información de los campos visuales de forma digital y poder tratar los datos mediante un ordenador para posteriores análisis.

La fatiga física de los pacientes ante una prueba perimétrica de larga duración llevó a los investigadores a buscar estrategias más rápidas de examen, pero solamente se habían conseguido reducciones de aproximadamente el 30% (programas Dynamic de Octopus y Fastpac de Humphrey). González de la Rosa y col. [58, 112] desarrollaron un programa para el estudio del glaucoma (programa Delphi) con una duración de un minuto tras observar que a partir de la sensibilidad de cuatro puntos críticos del campo visual se podían deducir los demás con una precisión suficiente. En 1997 el perímetro Octopus ha incorporado una nueva estrategia que ellos desarrollaron llamada Perimetría Orientada por Tendencias (TOP). Realizaron unos estudios en los que demostraron alcanzar con TOP unos resultados altamente comparables con los de la perimetría convencional. Confirmaron la buena correlación entre los resultados de ambos procedimientos en pacientes glaucomatosos y neurológicos. Redujeron el tiempo perimétrico en un 80% evitando el efecto fatiga que se producía con la perimetría convencional. Debido a esto, TOP produjo en los sujetos normales valores más centrados en la normalidad estadística que un programa largo. De forma simultánea, Humphrey desarrolló hace años otro programa corto de estimación de umbral denominado SITA que ofrece resultados análogos al programa de umbral completo en la mitad de tiempo. Hoy, SITA y TOP son las estrategias más utilizadas en las consultas.

En la actualidad se están ensayando nuevos tipos de perimetría, orientados fundamentalmente hacia el diagnóstico precoz del glaucoma. Se ha propuesto el estudio topográfico de la frecuencia crítica de fusión (Perimetría Flicker) [131], la perimetría cromática de estímulo azul y fondo amarillo (Perimetría SWAP), la perimetría de detección de movimiento [59], la perimetría de duplicación de frecuencia (FDT), etc. SWAP es una prueba muy sensible pero muy larga por lo que su utilidad clínica es algo limitada. FDT es un buen test de rastreo de glaucoma. El resto de las técnicas están en experimentación y no existe unanimidad respecto a que sus resultados superen a la perimetría convencional de estímulo y fondo blancos, aunque algunas de ellas parecen muy prometedoras a este respecto.

2.3.3 Métodos estadísticos

Puesto que la evaluación del campo visual no es una tarea sencilla, como se ha podido ir observando en los apartados anteriores, muchos de los esfuerzos en investigación que se fueron realizando inicialmente, iban encaminados a diseñar sistemas matemáticos que pudieran permitir una interpretación del CV más precisa y objetiva.

Se utilizaron como herramienta de trabajo métodos estadísticos que permitieran conocer el estado global del campo visual bien aplicando los datos

obtenidos a determinadas fórmulas estadísticas, de ahí surgieron los llamados índices del CV, o bien mediante análisis estadísticos más complejos.

2.3.3.1 Índices

El primer paso se dio al aplicar fórmulas que permitieron conocer el estado global del CV a través de los llamados "índices". Flammer [45, 46] diseñó los utilizados actualmente en el campímetro Octopus (sensibilidad media, defecto medio, varianza de la pérdida, varianza de la pérdida corregida, fluctuación a corto y a largo plazo) y Heijl desarrolló los utilizados por el campímetro Humphrey (sensibilidad media, desviación media, patrón de desviación estándar, patrón de desviación estándar corregido y fluctuación a corto plazo).

Estos índices se describen a continuación [6]:

La **sensibilidad media** (MS) es la media de los valores de sensibilidad umbral medidos en cada uno de los puntos explorados. Su utilidad clínica es bastante restringida, pues disminuye de forma progresiva con la edad.

El **defecto medio o desviación media** (MD) es la media de los defectos de sensibilidad presentes en cada punto, es decir, la media de la diferencia entre la sensibilidad medida y la sensibilidad normal (de acuerdo con grandes bases de datos de sujetos sin patología) para la edad del sujeto y cada punto concreto. Es un buen índice del estado general del CV y puede ser útil para calificarlo como normal o patológico, o comprobar su evolución. Sin embargo, se modifica tanto por defectos difusos como localizados, y no permite distinguir entre ellos, ofreciendo así una información incompleta.

La **fluctuación a corto plazo** (SF) expresa la variabilidad dentro de una misma exploración, al comparar dos medidas distintas de la sensibilidad umbral en varios puntos, con un intervalo temporal de tan sólo unos minutos. En sujetos normales no suele ser superior a 3 dB; puede aumentar como consecuencia de la falta de atención o concentración del paciente (generalmente coincide con un alto índice de falsos negativos), en casos de patología avanzada y puede alterarse bastante en casos de patología incipiente. Con frecuencia está elevada en el glaucoma, pero no es específica de esta enfermedad como su autor Flammer pensó en un principio. La falta de especificidad es, probablemente, la mayor limitación que tienen todos los índices; éstos son útiles para catalogar un CV como normal o patológico o evaluar su evolución, comparando dos o más exploraciones sucesivas de un mismo paciente, pero no son capaces de establecer un diagnóstico etiológico de las lesiones presentes en el CV.

La **fluctuación a largo plazo** (LTF) es un índice que evalúa los cambios de sensibilidad a lo largo de días o meses, es decir, la variabilidad entre distintas exploraciones, y refleja la dinámica fisiológica del CV. La LTF fisiológica es un cambio global del CV en su conjunto, de forma que en unos puntos aumenta la sensibilidad y en otros disminuye, siendo la suma de las variaciones cercana a cero. Modificaciones de 2-4 dB en sujetos normales no son clínicamente

significativas, y en los glaucomatosos disminuciones de hasta 7 dB pueden no significar progresión de las lesiones. Sin embargo, si estos cambios son focales, en una determinada zona, lo más probable es que expresen una progresión de la enfermedad. Este índice no lo calculan los campímetros.

Hay también unos índices que evalúan la pendiente del CV, es decir, la variación de la sensibilidad en toda su extensión. Son sensibles a los defectos localizados, y por ello completan la información del MD. Son la **varianza de la pérdida** (LV, Octopus) o el **patrón de desviación estándar** (PSD, Humphrey). Existe además una modificación de la LV y el PSD teniendo en cuenta el valor de la SF: la **varianza de la pérdida corregida** (CLV, Octopus) y el **patrón de desviación estándar corregido** (CPSD, Humphrey). La CLV se eleva con frecuencia en casos de glaucoma incipiente y puede ser útil para distinguir a los glaucomatosos de los sospechosos de glaucoma, pero no hay que olvidar que también aumenta en algunos sujetos normales de forma proporcional a la edad.

Marrafa [105] hizo un estudio comparativo del programa G1 del perímetro Octopus con otras tres estrategias de perimetría automatizada cuyos resultados en términos de sensibilidad y especificidad se muestran en la Tabla 2.1. Este estudio pone de manifiesto una mayor capacidad del programa G1 para evaluar campos visuales.

Perímetro	**S (%)**	**E (%)**
Henson	59.4	88
Humphrey (test de campo completo de Armaly)	64.2	64.2
Perikon (discriminación Genoa)	55	90.4
Octopus (programa G1)	92.1	83

Tabla 2.1: Estudio comparativo del programa G1 del perímetro Octopus con otras estrategias de perimetría computerizada (Marrafa [105]).

Funkhouser [52] encontró una correlación muy alta entre los índices del Octopus y los del Humphrey, afirmando que son equivalentes y, por ello, utilizables de forma indistinta.

Señalan que con estos índices, sobre todo el MD y la LV (PSD) o CLV (CPSD), es posible calificar el campo visual como normal, patológico o dudoso y compararlo con exploraciones previas en unos segundos. Por esta razón, indican que han desbancado a las gráficas de grises o representaciones de los datos individuales de cada punto explorado del CV. Sin embargo, si únicamente se tienen en cuenta estos índices, se está desperdiciando una información de gran importancia: **la distribución geográfica de los defectos**.

Chauhan y col. [24] realizaron un estudio en el que demuestran que los índices (MD, CLV, SIZ, CLUS y PCLUS) usados actualmente pueden no ser clínicamente confiables en la determinación de cambios en el campo visual.

La clave sobre la posible causa de cualquier defecto campimétrico no solamente se encuentra en sus características cuantitativas (valor del defecto, índices) sino en sus características cualitativas (localización, forma y extensión de los escotomas). Asman [10] comprobó que, mediante el empleo de métodos de análisis del CV que tienen en cuenta el factor topográfico (Glaucoma Hemifield Test y análisis de "clusters"), se obtiene una mejor eficacia diagnóstica a la hora de clasificar los CV como normales o glaucomatosos, sobre todo gracias a una especificidad mayor que los métodos tradicionales de interpretación (índices: MD y PSD).

También algunos autores intentaron resumir toda la información que proporciona el campo visual en un único índice que permitiera fácilmente su clasificación y seguimiento a lo largo de las diferentes pruebas perimétricas realizadas a un mismo paciente. Como ya se ha indicado, el problema fundamental de los índices generales del campo visual (defecto medio, varianza de la pérdida, patrón de desviación estándar) es su falta de especificidad, es decir, se alteran en el glaucoma pero también en otras enfermedades oculares. Algunos de estos índices que pretenden ser más específicos al tener en cuenta la distribución de los defectos se comentan a continuación.

Brechner y col. [17] describen la creación de un nuevo estadístico Q que se puede calcular a partir de los datos umbrales de un único campo visual computerizado. El estadístico Q se puede usar para decidir si el campo visual es normal o anormal. Es particularmente sensible a la situación en la que solamente un pequeño subconjunto de los valores del campo está afectado, siendo especialmente útil para el diagnóstico de glaucoma precoz. Desde finales de los años 60 se ha trabajado en investigación para detectar defectos de glaucoma de forma precoz en el campo visual.

El cálculo del estadístico Q no depende del uso de los valores normales corregidos de la edad, sino que debido a su construcción, depende solamente de la forma o contorno del campo visual, no de su valor absoluto. Por tanto, puede usarse el mismo conjunto de valores normales en el cálculo de todos los estadísticos Q, sin tener en cuenta la edad del paciente, suponiendo que la forma o contorno del campo visual normal no cambia con la edad. El estadístico Q no se ve afectado si el campo visual completo tiene puntos con defecto. Es independiente de pequeñas variaciones en la calibración de la máquina. Sí se ve afectado por grandes cambios en la iluminación del fondo que modificarán el contorno del campo visual normal. Los cambios en el tamaño del estímulo de la prueba también cambiarán el contorno del campo. Por tanto, los valores críticos específicos que recogen estos autores no se pueden generalizar a otros perímetros que usen diferentes niveles de iluminación ni a otros campos visuales obtenidos con diferentes diámetros de estímulos para la prueba.

Es importante distinguir entre pérdidas de campo difusas y localizadas. Funkhouser [53] introdujo un nuevo índice de pérdida difusa (DL) que sirve para decidir si está o no presente un componente de pérdida difusa, y si está, ayudar a estimar su cantidad. Comparan los enfoques utilizados para calcular mapas de

probabilidades empíricas y las salidas del Statpac de la desviación de patrones con los resultados obtenidos con el algoritmo del índice de pérdidas difusas. Ambos métodos producen grandes diferencias cuando se comparan con el índice de pérdidas difusas. Por esta y otras razones, concluyen que los mapas de probabilidades empíricas y de desviación de patrones pueden llevar al médico poco experimentado a ignorar información de valor relativa a la patología real.

Según Mandava y col. [102] los índices del campo visual MD, CLV, SF son parámetros estadísticos que no incorporan información suficiente sobre patrones típicos de pérdida de campo visual glaucomatoso. Ellos desarrollaron un nuevo índice sensible a los defectos del campo visual llamado Índice de Patrón de Glaucoma (GPI). Como su nombre indica es un índice específico para detectar glaucoma, que tiene en cuenta la distribución espacial de los defectos. En la Tabla 2.2 se muestra su estudio sobre la capacidad de distinguir los CV glaucomatosos incipientes de los CV normales de varios de los índices.

Índice	**Sensibilidad (%)**	**Especificidad (%)**
GPI	88	88
MD	89	80
CLV	85	80
SF	81	70

Tabla 2.2: Estudio comparativo de la capacidad de varios índices de discriminar campos normales de glaucomatosos incipientes (Mandava [102]).

2.3.3.2 Análisis de Agrupamientos (Clusters)

El análisis de cluster es el nombre de una técnica multivariable que se usa en estadística para identificar características similares en un grupo de observaciones. En análisis de cluster se identifican y clasifican observaciones o variables de tal forma que cada observación o variable de un cluster se parece a las demás de ese mismo cluster y difiere de las de otros.

Esta técnica es útil para estudiar y analizar el campo visual. De hecho se ha aplicado en varios estudios. Es interesante conocer el número de puntos contiguos con defecto del CV para establecer criterios para el diagnóstico del glaucoma, y se puede analizar el CV de forma automática utilizando paquetes estadísticos para encontrar clusters o grupos de puntos con características comunes.

Henson define cluster como el grupo de puntos (2 o más) del CV que presentan defecto de sensibilidad y se encuentran situados entre sí a una distancia no superior a 7’5º [6].

Chauhan y col. [25, 27] estudiaron los clusters de defectos que se producen tanto en ojos normales como glaucomatosos. Aproximadamente el 13% de los normales tienen clusters, la gran mayoría de estos individuos tienen un cluster de dos defectos. La mayor parte de los clusters en los ojos normales son debidos a angioscotoma o variaciones fisiológicas de la posición del punto ciego. Los clusters debidos a otros factores ocurren raramente. Como ya se ha mencionado,

se encuentran con iguales frecuencias en los hemicampos superior e inferior en los ojos normales, pero con mayor frecuencia en el superior en los ojos con glaucoma. El análisis de cluster en perimetría es una técnica usada en la evaluación de pérdida localizada en el campo visual. Ha sido aplicado previamente a datos supraumbrales y, a diferencia de los índices actualmente disponibles para indicar pérdidas localizadas, es influenciado por las posiciones relativas de defectos. Desarrollaron un análisis de cluster con datos procedentes del programa 31 del perímetro Octopus. Usando los resultados de la muestra que utilizan, y con la ayuda de otras propiedades del campo visual, es posible diseñar índices de probabilidad ponderados para clasificar campos visuales. Vieron que hay alteraciones localizadas en el campo que no afectan al índice CLV y sí que afectan a los parámetros de cluster computados. El uso de clusters en la cuantificación del campo visual es sensible y específico.

Mandava y col. [103] usaron un análisis estadístico de cluster para analizar patrones de pérdidas en campos visuales. Comprobaron que este análisis era efectivo para localizar pérdidas localizadas y fluctuación a largo plazo. Estudiaron el uso de clusters para distinguir campos normales de glaucomatosos así como los estables de los que siguen deteriorándose.

Asman y Heijl tuvieron en cuenta en su estudio de clusters del CV no sólo su forma sino también la importancia diagnóstica de cada punto afectado. Asignaron a cada cluster dos valores numéricos: uno que expresa su volumen y el otro su extensión. Definieron como extensión el número de puntos que lo constituyen y como volumen la suma de las probabilidades que asigna el Humphrey a cada punto (probabilidad con que el valor de defecto de sensibilidad se presenta en la población normal). Estos autores comprobaron que el estudio de clusters arciformes era más útil para identificar las lesiones glaucomatosas que el de los clusters tradicionales [6].

2.3.3.3 Análisis discriminante con regresión logística

Además de los métodos estadísticos utilizados por diversos autores y citados en los apartados anteriores, cabe destacar la investigación realizada en este campo por Antón recogida en su Tesis Doctoral [6]. En este apartado se recogen sus experimentos y resultados con bastante detalle puesto que la muestra que utilizó es la misma que se ha tomado para realizar el presente trabajo de investigación y más adelante se podrá hacer un estudio comparativo de las técnicas empleadas.

Aplicó la discriminación logística o análisis discriminante con regresión logística (ADRL) para clasificar campos visuales. En concreto usó el ADRL del procedimiento LOGISTIC del paquete SAS para construir unas reglas de clasificación que permitieran distribuir correctamente individuos entre diversas clases.

El ADRL es el método de discriminación que mejor se comporta cuando existen variables cuya distribución se aleja bastante de una curva normal, como es el caso de los datos del CV. Este método es capaz de analizar múltiples variables

cualitativas o cuantitativas, cercanas o no a una distribución normal. También es capaz de buscar por sí mismo las características de los defectos glaucomatosos incipientes para después reconocerlos en otros casos. De esta forma, el método de clasificación extrae la información de la muestra de aprendizaje sin tener que aplicar antes los conocimientos ya existentes, y así se pueden completar éstos.

Para su trabajo revisaron 1491 historias clínicas y según unos criterios de inclusión seleccionaron 96 ojos con glaucoma que constituyeron el grupo 1 de la muestra de entrenamiento. El grupo 2 se formó con 84 sujetos sin glaucoma. De éstos, 28 no presentaban ninguna patología oftalmológica (subgrupo 2a). Los otros 56 pacientes del grupo 2 integran el subgrupo 2b: 27 ojos con catarata, 14 con retinopatía diabética y 15 diagnosticados de retinopatía hipertensiva.

Con la finalidad de evaluar la precisión diagnóstica de los métodos de interpretación se seleccionó una muestra de prueba independiente. Se siguió el mismo procedimiento que con la muestra anterior.

La muestra de prueba estuvo formada por 48 ojos, 20 diagnosticados de glaucoma (grupo 3) y 28 sin glaucoma (grupo 4). De éstos últimos, 8 sujetos no tenían patología oftalmológica (subgrupo 4a), 7 sufrían catarata, 7 tenían retinopatía diabética y 6 presentaban retinopatía hipertensiva (subgrupo 4b).

El objetivo era poder clasificar campos visuales en cinco categorías: glaucoma, normal, cataratas, retinopatía diabética y retinopatía hipertensiva.

El procedimiento Logistic del SAS seleccionó los puntos que separaban de forma significativa los individuos glaucomatosos de los demás. Teniendo en cuenta su distribución en la representación en variables canónicas, se obtuvieron primero reglas de clasificación para diferenciar los CV glaucomatosos de los que pertenecían a las demás categorías diagnósticas por separado, con excepción de los diabéticos e hipertensos que se consideraron como única categoría.

En una primera parte se tuvo en cuenta exclusivamente la presencia o ausencia de defecto y no la profundidad del mismo, esto es, se intentó discriminar la etiología de las lesiones campimétricas considerando únicamente su distribución topográfica. La precisión diagnóstica de cada una de las ecuaciones parciales, calculada mediante validacion cruzada y teniendo en cuenta distintos tipos de datos de entrada, se expresan en la Tabla 2.3.

	59 PUNTOS		**59 PUNTOS + MD y LV**		**7 ZONAS**		**7 ZONAS + MD y LV**	
	S	**E**	**S**	**E**	**S**	**E**	**S**	**E**
GL vs. NOR	79	77	78	73	92	75	94	67
GL vs. D+H	91	62	83	62	100	20	97	24
GL vs. CAT	81	57	81	57	96	29	96	29

Tabla 2.3: Sensibilidad (S) y especificidad (E) de las ecuaciones de regresión a la hora de clasificar los CV como glaucomatosos (GL), normales (NOR), con catarata (CAT) o diabéticos o hipertensos (D+H).

En la Tabla 2.4 se expresan los valores de sensibilidad y especificidad obtenidos con la misma muestra utilizada para obtener las ecuaciones (método de resustitución) y con una muestra de prueba independiente, así como las variables seleccionadas en cada caso. La variación de los valores de sensibilidad y especificidad de las reglas de clasificación en función de los puntos de corte seleccionados se expresa en las curvas de características operativas (ROC)[2] (Figuras 2.11-2.15).

		MUESTRA ENTRENAMIENTO		MUESTRA PRUEBA	
VARIABLES DE ENTRADA	**VARIABLES SELECCIONADAS**	**S**	**E**	**S**	**E**
59 PUNTOS	2,3,5,6,7,8,10,11,12,13,15, 16,17,19,20,21,22,23,24,2 7,30,32,34,35,36,37,41,43, 44,46,47,51,52,55,56 y 58.	83	92	65	60
59 PUNTOS, MD y LV	1,3,5,6,7,8,9,10,12,13,15, 16,17,19,21,22,23,27,28, 30,32,34,35,36,37,41,42, 43,46,51,55,56,57,58,MD y LV.	83	85	65	64
7 ZONAS	A1,A2,A6,A7	80	70	85	75
7 ZONAS, MD y LV	A1,A2,A3,A6,A7,MD y LV.	84	70	80	92

Tabla 2.4: Validez diagnóstica de las reglas de clasificación, obtenidas mediante regresión logística, según los distintos tipos de datos de entrada. Estos datos se consiguieron teniendo en cuenta la afectación o normalidad en la sensibilidad de cada punto.

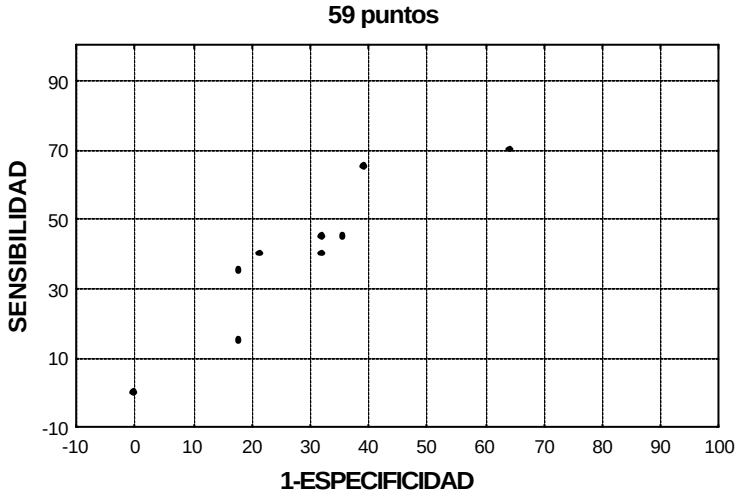


Figura 2.11: Curva de características operativas de la regla de clasificación desarrollada a partir de los 59 puntos del CV y evaluada con la muestra independiente.

[2] El apéndice B de esta Tesis muestra brevemente en qué consiste una curva de características operativas (ROC).

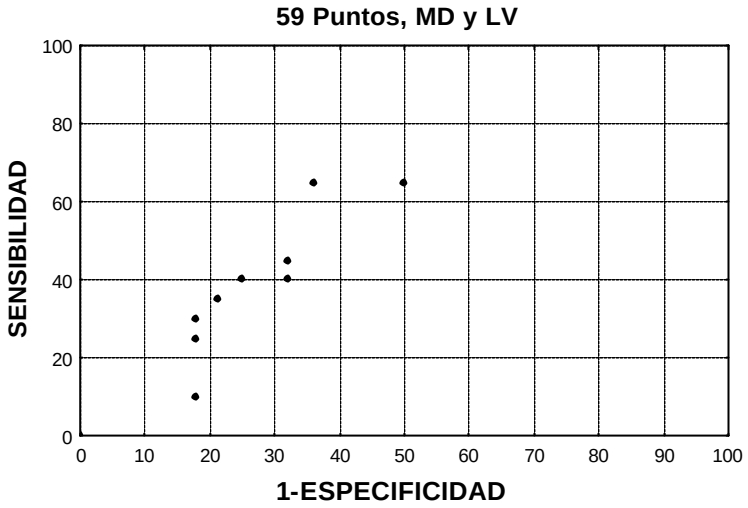


Figura 2.12: Curva de características operativas de la regla de clasificación desarrollada a partir de los 59 puntos del CV, MD y LV, evaluada con la muestra independiente.

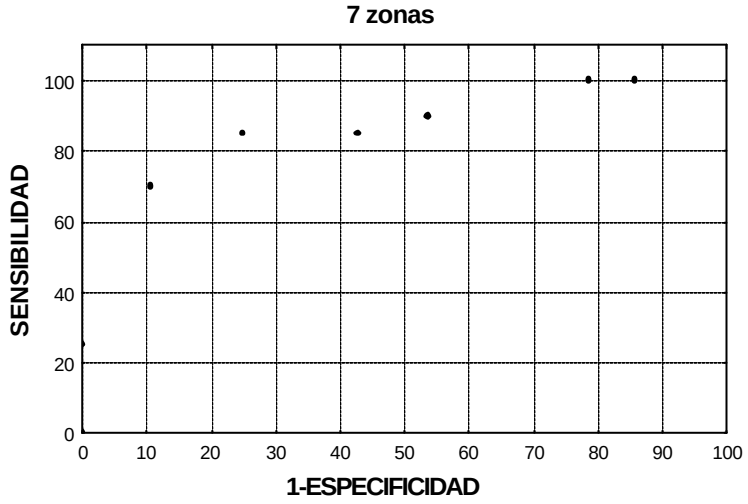


Figura 2.13: Curva de características operativas de la regla de clasificación desarrollada a partir de las 7 zonas del CV y evaluada con la muestra independiente.

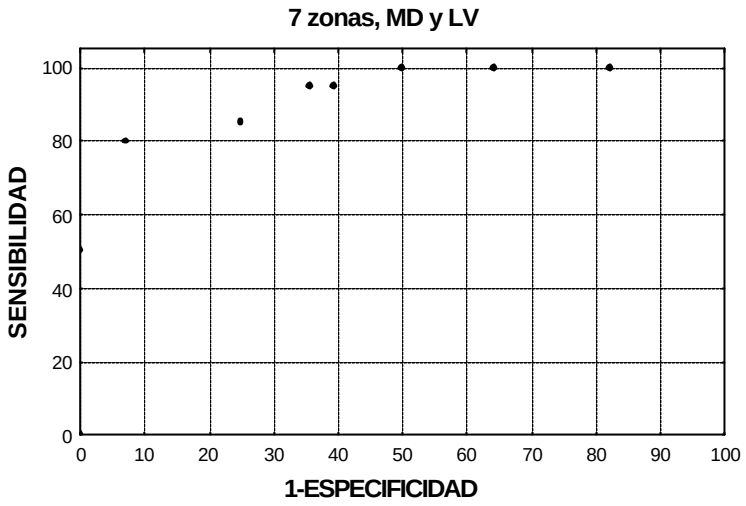


Figura 2.14: Curva de características operativas de la regla de clasificación desarrollada a partir de las 7 zonas del CV, el MD y la LV, evaluada con la muestra independiente.

En una segunda parte se introdujo el valor del defecto en cada punto y se aplicó el mismo procedimiento de ADRL. El procedimiento Logistic del SAS tuvo dificultad para seleccionar las variables significativas cuando se tuvieron en cuenta los valores unitarios de los 59 puntos, y no fue posible obtener una regla de clasificación. Al utilizar la distribución de los puntos en 7 zonas (Figura 4.2), el programa aportó tres reglas de clasificación y con ellas se estableció una regla general.

Esta regla de clasificación fue capaz de clasificar los CV con una sensibilidad del 84% y una especificidad del 60% (Tabla 2.5, Figura 2.15).

7 ZONAS	VARIABLES SIGNIFICATIVAS	S (%)	E (%)
GL vs NOR*	ZONAS 1 Y 7	93	67
GL vs D+H*	ZONAS 3 Y 7	99	20
GL vs CAT*	ZONAS 6 Y 7	97	22
GLOBAL**	ZONAS 1,3,6 Y 7	84	60

* Validación cruzada

** Resustitución

Tabla 2.5: Validez diagnóstica del algoritmo obtenido mediante regresión logística, agrupando los puntos en 7 zonas, teniendo en cuenta el valor cuantitativo del defecto en cada punto.

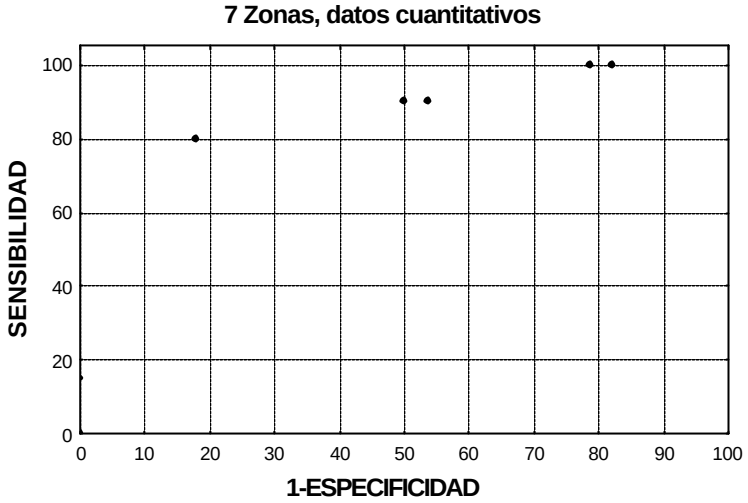


Figura 2.15: Curva de características operativas de la regla de clasificación desarrollada a partir de las 7 zonas del CV, teniendo en cuenta el valor de los defectos de sensibilidad, y evaluada con la muestra independiente.

El ADRL ya había sido utilizado con anterioridad para el tratamiento de los datos del CV. Sin embargo, nunca con casos de glaucoma tan precoces como los incluidos en este estudio ni para discriminar entre lesiones campimétricas de distintas enfermedades.

Este método permitió construir reglas de clasificación capaces de calificar los defectos incipientes presentes en el CV como de origen glaucomatoso o no glaucomatoso con un elevado porcentaje de acierto. En este estudio se alcanzaron sensibilidades entre el 65 y 80%, y especificidades del 60 al 92%. Son más precisas las reglas basadas en una división del CV en 7 zonas, y los índices generales, muy útiles para distinguir campos normales de otros patológicos, no mejoran de forma significativa la sensibilidad y especificidad cuando se trata de establecer el origen de los defectos campimétricos. El valor cuantitativo de los defectos, que sin duda tiene importancia en el diagnóstico de glaucomas más avanzados frente a casos normales, aporta escasa información a la hora de realizar el diagnóstico etiológico de las lesiones en casos incipientes.

2.3.4 Distribución del campo en zonas

Muchos autores han dividido el campo visual en zonas para facilitar su estudio e interpretación. Entre ellos están Schwartz y col. [138] que comparan los campos visuales contiguos a través de los resultados en **siete zonas** (Figura 2.16) mediante el test de la t de Student con datos apareados. La selección de los datos de cada zona se basa en estudios anteriores de otros autores y en su experiencia. Excluyen los valores periféricos. Seleccionan el grupo nasal para detectar cambios en ese área (por ejemplo un escalón nasal), y las áreas superior e inferior se seleccionan para detectar escotomas arciformes y paracentrales.

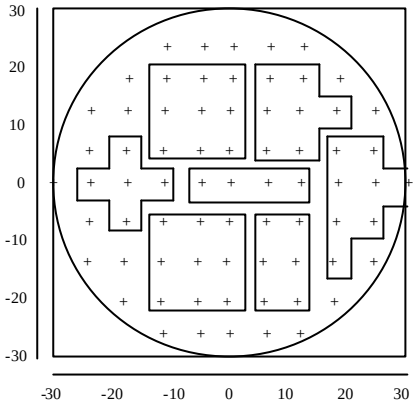


Figura 2.16: Siete zonas de Schwartz [138].

Posteriormente Wu y col. [169] realizan un estudio basándose en estas mismas 7 zonas. Si hay 5 o más campos, hacen un análisis de la tendencia sobre los valores umbrales medios de los campos visuales respecto al tiempo, sobre cada una de las 73 localizaciones del test respecto al tiempo, y sobre los valores umbrales medios de 7 regiones seleccionadas en el campo visual respecto al tiempo (análisis de regresión). Para pacientes con menos de 5 campos, desarrollan el test de la t.

Suzuki y col. [149] sectorizan los 30 grados centrales del campo visual. Los dividen en 15 sectores que tiene cada uno al menos 3 puntos. El patrón y los valores de sector obtenidos se consideraron útiles para estudiar los datos del campo visual en el glaucoma. (Usan el programa 30-2 del campímetro Humphrey).

Asman y col. [10], como ya se ha indicado anteriormente, desarrollaron un algoritmo, el Test del Hemicampo de Glaucoma (GHT), para automatizar la evaluación de resultados de campos visuales en glaucoma. El GHT usa límites de normalidad determinados empíricamente para las diferencias superior e inferior de los mapas de probabilidades del Statpac del Analizador de Campos Humphrey para detectar pérdidas de campo localizadas. También permite detectar pérdidas de campo simétricas respecto al eje horizontal. El análisis compara cinco pares de sectores correspondientes al hemicampo superior e inferior (Figura 2.17) que están basados en la anatomía normal de la capa de fibras del nervio óptico. Las desviaciones a partir del umbral normal corregido con la edad en las porciones más sensibles del campo se usan para detectar reducciones generales de sensibilidad o sensibilidades anormalmente altas. Asimetrías entre el hemicampo superior e inferior son típicas de glaucoma y clasificadas por el GHT como fuera de límites normales.

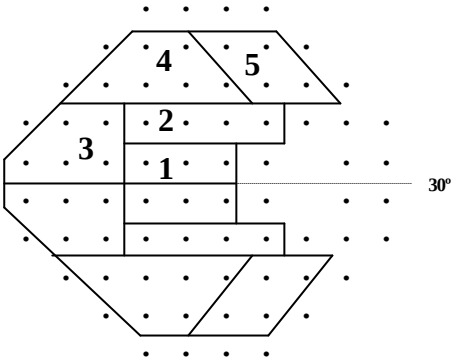


Figura 2.17: Sectores de Asman [10].

2.3.5 Sistemas Basados en Conocimiento o Sistemas Expertos

Los métodos de análisis del CV vistos hasta ahora no proporcionan un diagnóstico del mismo, sino que aportan nuevos datos que el experto debe interpretar. Además, existe un considerable desacuerdo entre los oftalmólogos en la interpretación del CV [107]. Shapiro y Jonson en 1991 [139] realizaron un estudio que mostró que dos expertos en perimetría no están de acuerdo en su interpretación del CV más que en un 40% de los casos, y en un 80% cuando se les pide tres posibles diagnósticos, coinciden en uno. Similares conclusiones mostraron otros estudios (Hirsbrunner y col. 1990 [77], Kaufmann y col. 1990, Funkhauser y col. 1991). Todo esto subraya la necesidad de utilizar métodos más formales en la ayuda al diagnóstico del CV. De ahí que surgieran otras herramientas que facilitaran esta tarea como son los Sistemas Expertos (SE) o Sistemas Basados en Conocimiento (SBC), y las redes neuronales artificiales (RNA) que se tratarán en el siguiente apartado.

Un Sistema Experto es un programa que intenta emular al experto humano aplicando métodos de inferencia a un cuerpo específico de conocimiento también llamado dominio [31].

Este tipo de programas tiene sus ventajas al aplicar sus reglas de diagnóstico y terapia con gran rigor. También es útil a los médicos que no son expertos aún. Un inconveniente de este tipo de programas es que podría suceder que el experto estuviera equivocado en algunos puntos y entonces el programa respondería siempre de forma incorrecta en esos aspectos del problema. Es necesario una realimentación que haga posible rectificar, si es necesario, las reglas de acuerdo a la experiencia.

Uno de los primeros que se conoce para el diagnóstico del CV es el de Krakau y col. [90] que analiza los defectos de campos visuales, reconoce su patrón y examina el desarrollo de una secuencia de campos. La base de datos contiene solamente casos de glaucoma.

Posteriormente Hirsbrunner, Bebie y col. [14, 76, 77] desarrollaron un SE perimétrico llamado OCTOSMART. Las dificultades asociadas a la tarea de distinguir entre resultados de campos visuales normales, dudosos y patológicos no deben ser subestimadas. Por ello este programa se puede ver como un paso importante hacia la automatización y la estandarización de la compleja tarea del diagnóstico del CV. El programa corre en cualquier ordenador personal compatible y analiza los resultados obtenidos con el programa G1 para el Octopus, es decir, se basa en la curva de defecto acumulado, los índices (MD, CLV y SF) y la influencia de los falsos positivos y negativos sobre dichos índices. El programa aporta una serie de comentarios sobre las características del CV y lo clasifica como normal, dudoso o patológico en distintos grados.

El programa ofrece a los principiantes ayuda en la toma de decisiones. Especialmente en la fase de aprendizaje dirige su atención a cualquier peculiaridad presente en los datos de partida y lo libera de tener en mente todas las reglas técnicas y la base estadística de la evaluación de los campos visuales.

El programa tiene sus limitaciones:

- No tiene en cuenta los resultados de campos visuales realizados anteriormente al paciente.
- Ha sido diseñado solamente para tratar datos derivados de exámenes hechos con el programa G1, y sólo cuando se han realizado las dos fases de este programa.

Según los autores, este programa no debe confundirse con un sistema experto que intenta llegar al diagnóstico basándose en patrones de defectos claros, sino que es como un prerrequisito para ese otro enfoque. Es la evaluación primera de los resultados de los exámenes médicos. Esta fase es normalmente la más difícil en el proceso de evaluación.

Los autores también afirman que **no es fácil evaluar la eficacia de un sistema experto ya que los propios expertos no se ponen de acuerdo en el diagnóstico de los casos dudosos**. De hecho, compararon la salida de su SE con el diagnóstico de tres oftalmólogos distintos encontrando importantes diferencias. Elaboraron posteriormente un índice global basado en el defecto medio y en la varianza de la pérdida corregida, que era el que mejor se correlacionaba con los diagnósticos emitidos por los médicos.

Martin y col. [106, 107] desarrollaron un programa informático para realizar una evaluación preliminar del campo visual, obtenido usando el sistema perimétrico de resolución high-pass (HRP). Éste es un método computerizado para medir umbrales de resolución en los 30 grados centrales del campo visual. Como con cualquier otra técnica perimétrica computerizada, HRP genera un gran número de datos que se utilizan cuando se evalúan los resultados de las pruebas. Inicialmente programaron un prototipo utilizando el lenguaje de programación de Sistemas Expertos **CLIPS.** CLIPS proporciona **encadenamiento hacia delante**, que se encontró apropiado para este propósito. (Si fuera necesario, el

encadenamiento hacia atrás podría ser emulado). En esta aplicación el razonamiento es deductivo y el encadenamiento hacia delante.

El prototipo estaba constituido por 120 reglas del tipo if-then y evaluaron 19 de entre más de 35 parámetros de cada ojo calculados con el sistema HRP. Los parámetros restantes no los juzgaron necesarios para el desarrollo del prototipo del programa. Las reglas se obtuvieron a partir del conocimiento estándar de perimetría acerca de los defectos de hemicampo y de cuadrantes. Como en la mayor parte de las pruebas fisiológicas, los bordes entre valores de HRP normales y patológicos son bastantes borrosos.

El programa proporciona una salida con una presentación gráfica de los resultados de ambos ojos. Basado en los valores actuales umbrales corregidos por la edad, el programa produce una explicación sobre la capacidad neuronal, las desviaciones de la normalidad en cualquier otro de los índices proporcionados por el sistema HRP, y, si es posible, una opinión concluyente del campo visual de cada ojo. Para poder detectar cualquier defecto bitemporal o binasal homónimo se hace un análisis de los datos de ambos ojos juntos. Las declaraciones están basadas en el análisis de los datos del campo visual solamente y no proporcionan un diagnóstico clínico. **La decisión del diagnóstico final debería siempre tener en cuenta más factores que los parámetros del campo visual y es responsabilidad del oftalmólogo.** Si todos los índices confiables caen dentro del rango normal la confiabilidad es juzgada como normal. El programa da los indicadores de confiabilidad que caen fuera de los valores recomendados y el examinador puede elegir continuar la evaluación o rechazar los campos.

Los resultados de la interpretación se muestran en la pantalla, junto con la pregunta de si el examinador quiere saber la explicación del programa sobre cómo llegó a la conclusión.

Tras este prototipo pensaron que podía ser útil aplicar alguna técnica disponible de los SBC para "**razonamiento bajo incertidumbre**". Compararon el funcionamiento del prototipo con la evaluación hecha por los observadores clínicos con varios grados de experiencia en el sistema HRP. Se usaron para este estudio 60 campos visuales de 30 pacientes que no eran de diagnóstico trivial. El número de diagnósticos correctos hechos por el programa fue más alto comparado con los observadores clínicos al evaluar campos visuales individuales e igual al mejor observador al evaluar pares de campos visuales.

Los resultados obtenidos fueron:

80% de aciertos en campos normales.
63% en glaucomatosos
64% en neurológicos
54% campos de pacientes con neuropatías ópticas menores.

Esto fue una señal de buen funcionamiento del prototipo dado su limitado tamaño.

Finalmente, desarrollaron un programa para uso clínico y una versión beta. Se hizo una nueva versión del programa. Se probó una versión beta durante 6 meses con aproximadamente 1000 campos visuales. Después de esto se hicieron algunas modificaciones y se añadieron unas pocas reglas más. Esta versión consistió en 200 reglas y tiene en cuenta parámetros de confiabilidad y datos del campo visual de cada ojo.

La versión actual del programa identificó 99.5% de los sujetos normales, 64% de pacientes con lesiones quiasmáticas y 83% con lesiones retroquiasmáticas. La aparente baja detección de lesiones de quiasma se puede explicar por el hecho de que no se pueden esperar en todos los pacientes de una población clínica defectos típicos de campo visual, especialmente en casos con bajo grado de daño. El grado de daño del sistema visual en desórdenes neurológicos puede ir desde daño muy pequeño, con campos aparentemente normales, hasta lesiones extensas con defectos absolutos.

Entre las conclusiones que destacan estos autores está la siguiente:
La interpretación asistida por computador de los resultados del HRP se puede esperar que dé un apoyo considerable a la decisión del diagnóstico final. No obstante, la decisión del diagnóstico final debería siempre **tener en cuenta más factores que los parámetros del campo visual** y queda bajo la **responsabilidad del oftalmólogo.**

Madsen y col. [101] en 1993 desarrollaron un Sistema Experto para analizar los patrones de pérdida del campo visual producidas por glaucoma usando para ello una herramienta de desarrollo de Sistemas Expertos disponible comercialmente.

2.3.6 Redes Neuronales Artificiales

A principios de los 90 se empezaron a utilizar Redes Neuronales Artificiales para mejorar el diagnóstico del campo visual.

Las RNA, las cuales son capaces de aprender a partir de ejemplos, poseen una serie de atributos que las hacen particularmente útiles en las tareas de diferenciación y clasificación de patrones espaciales. Tienen la habilidad de aprender similitudes entre patrones directamente a partir de ellos, no tienen nociones preconcebidas, y pueden generalizar cuando se les presenta para procesar datos que no han sido vistos previamente.

Hay dos clases amplias de RNA: con aprendizaje supervisado y no supervisado. La RNA supervisada aprende a partir de un conjunto de datos de entrenamiento en el que cada ejemplo ya ha sido clasificado. Entonces aprende a reproducir esta clasificación. El perceptrón multicapa entrenado mediante el algoritmo de retropropagación del error, que ha sido usado por un número de investigadores interesados en desarrollar sistemas automáticos para el reconocimiento y clasificación de defectos del campo visual, es una red típica de este tipo de aprendizaje.

La RNA no supervisada no requiere que los datos de entrenamiento estén previamente clasificados. La clasificación es desarrollada por la red. Los mapas autoorganizados (SOM) utilizan un método de aprendizaje no supervisado.

Debido a que las RNA aprenden a partir de los ejemplos del conjunto de entrenamiento la composición de este conjunto es un factor muy importante en la conclusión final.

Estas técnicas son prometedoras en problemas de reconocimiento de patrones como el que estamos estudiando. La mayor parte de los primeros intentos de clasificación de CV mediante RNA se hicieron utilizando redes multicapa (MLP) entrenadas con el algoritmo de retropropagación del error (BP), aunque también se han utilizado en algunos proyectos mapas autoorganizados de Kohonen.

Liu y col. [94] en 1994 identifican en el diagnóstico del glaucoma el ruido en los datos del campo visual debido al efecto de aprendizaje, falta de atención, fatiga, etc., mediante mapas autoorganizados de Kohonen. El método que proponen proporciona un modo de localizar y rechazar el ruido en los datos, y supone una mejora sobre los métodos estadísticos convencionales. También proporciona a los oftalmólogos una visión clara del comportamiento del paciente durante el examen, y les facilita el proceso de toma de decisión en el diagnóstico.

Los artículos publicados antes del año 94 sobre la aplicación de redes neuronales a la clasificación del campo visual demuestran que se pueden obtener buenos resultados, pero no exploran los efectos que el conjunto de entrenamiento puede tener sobre el funcionamiento de la red. Spenceley y col. [146] consideran el problema de diferenciar campos normales de glaucomatosos y exploran diferentes características del conjunto de entrenamiento usando campos recogidos de un perímetro Henson CFS2000, como son el tamaño de la muestra, el balance entre normales y glaucomas, la extensión de la pérdida de campo y la localización espacial de los defectos glaucomatosos.

Utilizan una red multicapa con 132 nodos de entrada, 20 en la capa oculta y 2 de salida, entrenada usando el algoritmo de retropropagación del error. Se calculan los valores de sensibilidad y especificidad durante la parte de test. Los resultados demuestran que los conjuntos aleatorios grandes son mejores que los pequeños ya que la sensibilidad mejora con el tamaño y la especificidad no empeora. La variabilidad en el funcionamiento también se reduce a medida que el tamaño del conjunto de entrenamiento aumenta. Además, los conjuntos que son sesgados hacia ejemplos de glaucoma son más sensibles y menos específicos, mientras que los conjuntos que son sesgados con ejemplos normales son más específicos y menos sensibles que los conjuntos balanceados. Por lo tanto es deseable tener una muestra grande con clases balanceadas para obtener buenas sensibilidades y especificidades. La distribución espacial de los defectos es crucial. Es importante presentar a la red casos que cubran prácticamente todas las localizaciones posibles. En conclusión el conjunto de entrenamiento ideal debería

contener muchos ejemplos de defectos tempranos que representen el rango completo de localizaciones donde estos defectos pueden ocurrir.

Goldbaum y col. [56], también en 1994, realizan un estudio aplicando **retropropagación del error** a la clasificación de campos visuales. Su objetivo es discriminar entre ojos normales y glaucomatosos. Para ello utilizan una muestra de 60 ojos normales y 60 glaucomatosos y recogen los datos obtenidos por el Analizador Humphrey.

El glaucoma fue diagnosticado basándose en la apariencia del nervio óptico (E/P de al menos 0.7, muescas en el anillo del disco, o progreso en la excavación en revisiones sucesivas) y PIO por encima de 22 mmHg en dos revisiones distintas no muy seguidas.

Se examinaron los 24º centrales del campo visual con 6º de resolución espacial usando el programa 24-2 o 30-2 eliminando las posiciones de prueba extra. La sensibilidad de la retina en cada posición fue examinada con una resolución de 2 dB con una estrategia umbral completa en personas que tenían al menos un campo visual previo. Se excluyeron los CVs con un error mayor del 25% en pérdidas de fijación, respuestas de falsos-positivos, o de falsos-negativos.

Se construyó una red multicapa con las características que se indican a continuación. Debido a que el valor de entrada para cada unidad está restringido entre 0 y 1, los valores de decibelios del campo visual fueron normalizados entre 0 y 1. La sensibilidad máxima absoluta medida en los ojos examinados fue 48 dB; por tanto, las unidades del CV fueron normalizadas dividiendo por 48. La edad fue también considerada debido a que afecta al CV en personas con o sin glaucoma. Se tomaron 56 valores de entrada: 54 del CV (inicialmente se tomaron los valores de sensibilidad umbrales absolutos para que la red hiciera sus propias asociaciones con datos no procesados previamente), el 55 fue la edad normalizada entre el valor mínimo (40 años) y el máximo (84 años), y el 56 toma valor 0 para el ojo izquierdo y 1 para el derecho. La capa oculta tenía 4 unidades. Experimentalmente se comprobó que muy pocas unidades ocultas alargaban el tiempo de aprendizaje y muchas unidades disminuían la precisión en la clasificación de nuevos patrones. La capa de salida tenía una unidad. La salida esperada era 0 para normal y 1 para glaucoma. La clasificación de un nuevo patrón fue asignada a normal si la salida para ese patrón era menor de 0.5 y a glaucoma si era mayor de 0.5. Se eligió el umbral 0.5 por ser el valor intermedio.

La red se consideró entrenada cuando el error de la salida medido era menor del 1% del error de una red sin entrenar o cuando el error no disminuía más con iteraciones posteriores. El número de iteraciones fue disminuido usando una razón de aprendizaje de 0.1 para las primeras 500 iteraciones, 0.03 para las 500 siguientes, y 0.01 para las últimas 500. 1500 iteraciones de un periodo fueron suficientes para todos los conjuntos de entrenamiento.

Una red multicapa, con salida logística y capas ocultas del tipo al descrito, usa superficies curvadas en el espacio de entrada adaptadas para separar los datos

en clases. Una red con salida logística pero sin capa oculta (una única capa de conexiones) separa los datos por hiperplanos. Comparando los resultados de la red de dos capas, que es análoga al análisis discriminante, a los resultados obtenidos con la red de tres capas, se puede asegurar la mejora lograda mediante las técnicas de separación no lineal.

En general, cuanto más grande sea el conjunto de entrenamiento la clasificación de datos no vistos previamente es más precisa.

Para evaluar la bondad de clasificación de la red se optó por utilizar el método estadístico de validación cruzada.

Este método usa eficientemente todos los patrones (entrenamiento y test) en diferentes veces. El conjunto entero de patrones se divide en k particiones. La primera partición se toma para test, y las otras k-1 particiones se combinan para formar el conjunto de entrenamiento. El proceso se repite con cada una de las demás particiones.

En este caso se divide el grupo de 120 individuos en 10 partes mutuamente exclusivas, y cada una contiene 6 ojos con glaucoma y 6 sin glaucoma. Para cada conjunto de entrenamiento se cuentan los errores en el conjunto de test. La razón de error es el número total de errores dividido por 120.

Los resultados obtenidos fueron los siguientes (Tabla 2.6):

	Experto 1	**Experto 2**	**Media de expertos**	**Red de 3-capas**	**3-capas Mitad del conj. Entr.**	**Red 2-capas**
Sensib.(%)	67	51	59	65	57	48
Especif (%)	65	83	74	72	65	60
Clasif. correcta global (%)	66	68	67	68	61	54

Tabla 2.6: Resultados obtenidos por Goldbaum [56] en la clasificación de campos visuales.

En esta tabla se puede apreciar que los expertos humanos tienden a ser más específicos (se ve en el 2º experto y en la media de expertos).

Los peores resultados fueron los de la red de 2-capas ya que los datos no fueron linealmente separables.

Usaron la estadística de desacuerdo de McNemar para comparar dos métodos de clasificación y no encontraron desacuerdo significativo entre los dos expertos o entre un experto y la red.

Algunos de los estudios realizados mostraron que los patrones de una tercera parte de los campos del grupo con glaucoma se parecían a una tercera parte de los campos normales, es decir, había cierto solapamiento en la muestra.

El porcentaje de acuerdo entre los 2 expertos y entre cada experto y la red de 3-capas está entre 72 y 75%. Esta similitud sugiere que fueron los datos, y no el método de clasificación, los que limitaron la precisión de la misma.

Se ha probado que las redes neuronales pueden funcionar bien comparadas con clasificadores estadísticos que usan análisis discriminante o regresión para tareas de clasificación similares. Las RNA pueden encontrar relaciones complejas en los datos para ayudar a hacer una clasificación correcta. Una desventaja es que a los médicos no les gusta confiar en una caja negra que da una respuesta sin una explicación. La base de conocimiento está formada por los pesos de las conexiones. Desafortunadamente, es difícil saber lo que un MLP hace con los pesos para lograr la clasificación. El análisis del proceso de razonamiento de las redes entrenadas es un área actual de investigación dentro de las RNA.

El método de aprendizaje BP de RNA es al menos tan preciso como un experto en glaucoma al clasificar CV a partir de un perímetro automatizado cuando a ambos se les da la misma información.

Mutlukan y col. [121] en el año 94 diseñaron una RNA para clasificar campos visuales a partir de video-campímetros basados en ordenador personal y facilitar la interpretación diagnóstica de los resultados de la campimetría por no-expertos. La red utiliza BP y tiene 3 capas con 110 unidades en la capa de entrada (cada unidad corresponde a un punto de test del campo visual), una capa oculta de 40 unidades de procesamiento, y una capa de salida de 27 unidades (cada una correspondiente a un tipo particular de patrón del campo visual). La red se entrenó con 540 patrones simulados del campo visual que incluían casos normales, glaucomatosos y otros con defectos neuro-oftalmológicos durante 20000 ciclos. La precisión de clasificación de la red se calculó inicialmente con un conjunto de 135 campos simulados no vistos previamente por la red y posteriormente se examinó con un conjunto de resultados reales que contenían 100 campos neurológicos y 200 glaucomatosos. Obtuvieron una precisión de clasificación de 91-97% con los resultados de campos simulados y 65-100% con resultados de campos reales.

Henson y col. [73] en 1996 desarrollan un sistema de clasificación objetivo para los patrones espaciales de pérdidas en el campo visual encontradas en el glaucoma.

Se sabe que el glaucoma produce patrones característicos de pérdida de campo visual. La clasificación subjetiva está basada en la localización, forma, y extensión de la pérdida.

Los patrones espaciales de pérdida dentro de los hemicampos superior e inferior muestran un grado de independencia debido a la organización anatómica de las fibras visuales en la cabeza del nervio óptico. Cuando se clasifica espacialmente el campo visual completo esta organización puede resultar en un cierto número de clases redundante en las cuales ocurran combinaciones

diferentes de los mismos patrones superior e inferior. Este problema puede ser solucionado clasificando independientemente los hemicampos superior e inferior.

Este trabajo también enfatiza la dependencia del sistema de clasificación de los criterios de inclusión del conjunto de entrenamiento.

La clasificación está basada en la salida de un mapa autoorganizado de Kohonen (SOM) entrenado con 560 datos tomados a partir del programa 24-2 del analizador **Humphrey**.

Tomaron 50 puntos del campo visual (quitan los 2 del punto ciego y 2 que están a más de una excentricidad de 24 grados en el campo temporal). Los dividieron en dos partes y por tanto toman 25 para cada hemicampo.

Los valores umbrales del HVA fueron transformados en valores binarios (normal o con defecto). Las pérdidas de más de 6 dB se consideraron defectos en el campo visual. Este proceso de conversión es similar al usado por el paquete de software STATPAC para analizar los datos de CV tomados con el HVA. **La entrada binaria a la RNA significa que las clasificaciones resultantes están basadas únicamente en la localización espacial y en la extensión de los defectos y no en su profundidad.**

La estructura de la red SOM puede ser concebida como una capa de nodos de entrada X_1-X_m (cada uno correspondiente a una localización de test dentro del campo visual) conectada completamente a una capa de salida bidimensional (el mapa) de nodos K_1-K_p. Durante el entrenamiento los pesos asociados a las conexiones entre las capas de entrada y de salida W_{11}-W_{mp} son modificados basándose en la información contenida en los datos de CV y en las variables de entrenamiento. Después del entrenamiento un patrón de un campo visual dado causará que un nodo de salida se active más que los otros. Este nodo representa la clase espacial del defecto. (Usaron el software NeuralWorks Profesional II para entrenar la red). Cada red tenía 25 nodos de entrada (uno por cada localización del test) y una capa de salida de 5 por 5. Durante el entrenamiento se le pasó a la red 30 veces el conjunto de entrenamiento. Se usó una función de vecindad cuadrada que comenzó por la dimensión máxima (es decir, 5) y decreció en tamaño hasta un único elemento de procesamiento durante el entrenamiento.

Cada uno de los 25 nodos de salida sobre el mapa viene a representar una clase de defecto del campo visual. Esto se representa gráficamente en este trabajo por la intensidad de los pesos que conectan cada localización de test al nodo individual. La intensidad de los pesos es representada por una escala de grises donde las áreas oscuras representan pesos fuertes (defecto) y las áreas claras representan pesos débiles (no defecto). Los 25 nodos/clases son numerados 0-24 desde la izquierda arriba hasta la derecha abajo; esta numeración no representa ninguna característica de los datos.

El SOM clasificó con éxito los defectos sobre la base de los patrones de pérdida. Los mapas muestran una continuidad de cambio a medida que uno

avanza por ellos con pérdida temprana en una esquina y avanzada en la esquina opuesta (Figura 2.18).

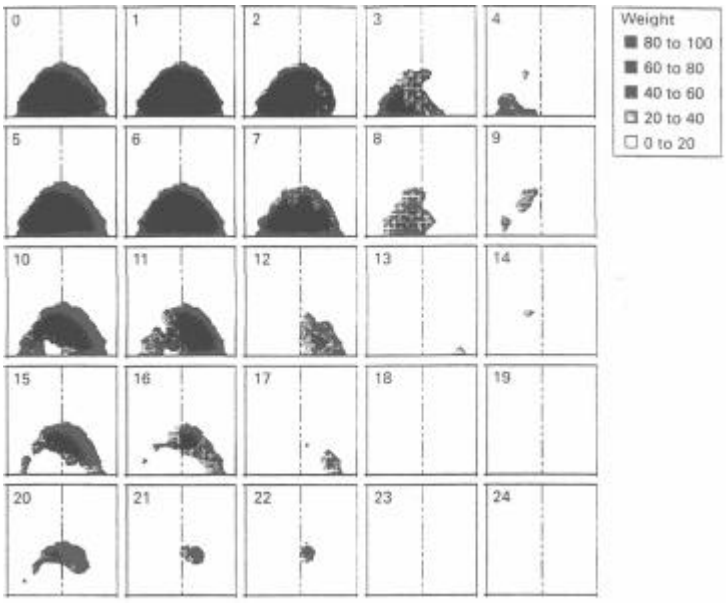


Figura 2.18: El mapa de dimensión 5x5 de una red de Kohonen entrenada con 560 campos visuales glaucomatosos y normales. Los datos de entrenamiento contienen los defectos de sensibilidad de los 25 puntos del hemicampo superior examinados por el perímetro Humphrey. Cada escala de grises representa los pesos para cada nodo sobre una escala que va de 0 a 100 organizada espacialmente para representar el campo visual [73].

Aulhorn y Harms listaron subjetivamente ocho clases de pérdida glaucomatosa, cinco de las cuales cayeron dentro de la región central del campo visual. Una de estas clases centrales, alargamiento del punto ciego, no es probable que sea detectada con una técnica estática cuya resolución espacial es 6 grados de ángulo visual. Las otras cuatro clases (escotoma pequeño, escotoma arciforme conectado al punto ciego, escotoma arciforme no conectado al punto ciego, y escotoma en forma de anillo) están todas representadas en los mapas que mostraron Henson y col. En ciertos ejemplos varios nodos son asignados a una única clase de Aulhorn y Harms, las distinciones a menudo son hechas sobre la base de tamaño y localización. Por ejemplo, la clase de Aulhorn y Hams de escotoma pequeño incluía defectos existentes casi en cualquier lugar dentro del campo central. En contraste, la RNA subdivide esta clase basándose en la localización del defecto –un escotoma pequeño en el campo superior nasal lo sitúa en una clase diferente de un defecto de tamaño similar en el campo temporal superior. La RNA ha sido configurada para producir 25 clases de pérdida en el hemicampo superior y 25 clases de pérdida en el hemicampo inferior; **esta graduación más fina de patrones espaciales de pérdida es útil para la monitorización de defectos progresivos.**

Es posible entrenar SOMs con casi cualquier número de nodos de salida. La elección de 25 nodos para cada hemicampo está basada en la consideración de un número de diferentes factores. Uno de éstos es el número posible de clases que existen dentro de los datos. Entrenando un SOM con más nodos que clases de datos hay, puede resultar bien un número de nodos redundantes (nodos cuyos pesos son muy similares a sus vecinos pero que no son nunca activados) o nodos que son específicos a ejemplos individuales dentro de los datos. Por otra parte, si el número de nodos fuera menor que el de clases, la red sería forzada a generalizar, poniendo juntas clases similares y limitando la habilidad de la red para discriminar. Un segundo factor que hay que tener en cuenta cuando se decide el número de nodos de la red es el número de entradas. Si hay solamente unos pocos nodos de entrada entonces el máximo número de patrones binarios potenciales es también pequeño. Un tercer factor es el tamaño del conjunto de entrenamiento. Si el conjunto de entrenamiento es pequeño y el mapa de salida es grande entonces hay el peligro de que algunos nodos tengan muy pocos ejemplos y que por tanto representen ejemplos específicos en lugar de clases.

También es posible alterar la forma del mapa de salida a partir de la forma cuadrada usada en este estudio a una rectangular o lineal. Cuando decidieron la forma apropiada pensaron en el progreso de los defectos y en la característica de la red de intentar organizar las salidas para formar un continuo desde un extremo hasta el otro. La forma cuadrada usada ofrece más opciones para el progreso uniforme que un mapa de salida rectangular o lineal.

Henson y col. concluyen que las RNA pueden clasificar los datos de campos visuales según los patrones de pérdida. Una vez entrenada la RNA puede usarse para clasificar los datos del campo visual longitudinal que pueden ser de valor al monitorizar la pérdida del campo visual y así poder realizar un seguimiento de los campos visuales de un paciente. Este sistema también permite comprobar si campos de diferentes pacientes son equivalentes respecto a su pérdida espacial.

Cho y col. [29] en 1998 usan un método para realizar las pruebas del campo visual en el que integran redes neuronales artificiales autoorganizadas y heurísticas empíricas, con el que logran reducir el número de veces que se deben examinar los puntos en un test perimétrico. Los experimentos desarrollados muestran que es posible reducir de un 20 a un 30% el número de pruebas por test sin afectar casi la precisión de los tests.

Reyes [128] en 1998 en su Tesis Doctoral usa técnicas de *soft computing* en el estudio del glaucoma en estado incipiente. Una de las técnicas que aplica son las RNA para predecir el diagnóstico de un campo visual. Utiliza redes *feedforward* (perceptrón multicapa) con el algoritmo *backpropagation* de aprendizaje supervisado.

Los datos de campos visuales proceden de la unidad de glaucoma del Hospital Clínico Universitario de Valladolid y son los mismos que anteriormente había usado Antón en su Tesis Doctoral [6]. Para la recogida de datos se usó el programa G1 del campímetro Octopus 500. En todos los experimentos utilizó las

medias de los defectos de sensibilidad de las siete zonas establecidas en el campo visual teniendo en cuenta la distribución de las fibras nerviosas en la retina, la experiencia clínica de la Unidad de Glaucoma del Hospital Clínico Universitario de Valladolid y las ya empleadas en el sistema experto Glaucom-Easy [6, 7, 142]. Por tanto, los vectores de entrada a la red tienen 7 componentes, a los que añade en algunos experimentos el MD o la LV, dando lugar a entradas de 8 ó 9 componentes.

Divide el estudio en tres fases. En la primera estudia un conjunto de redes que discriminan los campos visuales en tres clases: normal, glaucoma y otras patologías distintas al glaucoma. En la segunda clasifica en dos categorías: glaucoma y no glaucoma. En la tercera fase el objetivo es el mismo que en la segunda, pero dividiendo el problema en dos procesos diferentes: en el primero discrimina entre normales y patológicos, y éstos últimos se clasifican en glaucoma y otras patologías en el segundo proceso. Uno de los objetivos de la tercera fase es mejorar la sensibilidad y especificidad en la detección del Glaucoma Crónico Simple.

En la fase 1 la topología $7 \rightarrow 4 \rightarrow 3$ es la que proporciona los mejores resultados (Tabla 2.7):

	S (%)	**E (%)**
Glaucoma	90	82.14
Normal	75	77.50
Otras patologías	50	85.71

Tabla 2.7: Resultados de sensibilidad y especificidad obtenidos en la clasificación de campos visuales en tres categorías utilizando la red 7 ® 4 ® 3.

Los mejores resultados en la fase 2 fueron obtenidos con las siguientes topologías (Tabla 2.8):

Topología de la red	**S (%)**	**E (%)**
$7 \rightarrow 4 \rightarrow 1$	100	82.14
$9 \rightarrow 3 \rightarrow 1$	95	92.86

Tabla 2.8: Resultados de sensibilidad y especificidad obtenidos en la clasificación de campos visuales en dos categorías: glaucoma y no glaucoma.

En la fase 3 se entrenan de forma independiente redes para clasificar entre campos normales y patológicos (subfase 3.1), y entre glaucoma y otras patologías distintas de glaucoma (subfase 3.2). En la Tabla 2.9 se muestran los resultados de una configuración de red para cada subfase manteniendo fijo el número de entradas.

Topología de la red	subfase	S (%)	E (%)
$7 \rightarrow 4 \rightarrow 1$	Normal vs. patológicos	75	82.5
$7 \rightarrow 3 \rightarrow 1$	Glaucoma vs. Otras patologías	100	70

Tabla 2.9: Resultados de sensibilidad y especificidad obtenidos en la clasificación de campos visuales entrenando redes de forma independiente.

La fusión de ambas subfases se realiza de la siguiente forma: la entrada se presenta primero a la red que discrimina entre campos normales y patológicos; los que resultan clasificados como patológicos se muestran a la red que clasifica entre glaucoma y otras patologías. Los resultados obtenidos para la clase glaucoma son: sensibilidad 100%, especificidad 82.14% y precisión diagnóstica 89.6%.

2.4 Diagnóstico global del glaucoma

Hasta aquí hemos revisado los esfuerzos que se han hecho hasta nuestros días para clasificar el campo visual. Según Flammer y col. [46] el diagnóstico del glaucoma crónico de ángulo abierto está basado en la PIO, en los cambios de las características morfológicas de la cabeza del nervio óptico de la capa de fibras nerviosas y en los cambios típicos en los resultados de las pruebas psicofísicas. Se ha demostrado que varias pruebas psicofísicas son útiles en el diagnóstico del glaucoma, como el test de los colores o el test de sensibilidad al contraste. Ninguna de estas pruebas, no obstante, ha podido reemplazar la perimetría que actualmente es todavía el parámetro psicofísico más importante para el diagnóstico y seguimiento del glaucoma. Por eso todas las mejoras y avances que se logren en este tema son importantes.

Sin embargo, Brechner y col. [17] señalan que debe tenerse en cuenta que la declaración de un campo visual como patológico no implica que el paciente tenga esa enfermedad. Para decidir si el paciente tiene o no glaucoma, el médico debe tener en cuenta todos los datos recabados en la consulta. Y Martin, L. [107] indica que la decisión del diagnóstico final debería siempre tener en cuenta más factores que los datos del campo visual y que siempre es responsabilidad del oftalmólogo.

Por otro lado, Sponsel y col. [147] señalan que aunque se ha demostrado claramente que la presión intraocular no es un parámetro definitivo a la hora del diagnóstico del glaucoma y que el análisis del campo visual es mucho más seguro que el seguimiento de la presión intraocular, sin embargo en las consultas oftalmológicas, donde se trabaja bajo la presión del tiempo, no sorprende que la PIO sea muchas veces la única prueba que se le practica al paciente para descartar un posible glaucoma.

También se han realizado estudios (Zeyen y col. [177] y Iester y col. [79]) que ponen de manifiesto la relación temporal entre la progresión del campo visual y el deterioro del disco óptico en el glaucoma incipiente. Estos estudios documentan el daño progresivo del disco, anterior a la pérdida del campo visual en el glaucoma incipiente y la correlación entre parámetros de ambas pruebas.

En los siguientes apartados vamos a estudiar los trabajos de investigación que se han realizado hasta la fecha en los que se reúnen todos los datos recabados en una consulta para llegar a un diagnóstico final o global del paciente.

2.4.1 Sistemas basados en reglas de producción

Paycha y col. [124] desarrollaron en 1986 un programa de ayuda al diagnóstico por identificación (PADI), un sistema experto para glaucoma en adultos. El motor de inferencia genera un algoritmo general de comparaciones en serie. A partir de unos principios simples se formaliza el problema. Se definen unos símbolos para describir por una parte los síntomas y la enfermedad, y por otra, los procesos de comparación. Se establecen dos formas de razonamiento: hipótesis única e hipótesis múltiples. Los casos de hipótesis única sólo pueden tener una respuesta. Los casos de hipótesis múltiples pueden dar varias soluciones simultáneamente. En estos últimos casos la decisión se tomará teniendo en cuenta que si se duda entre dos afecciones se inclina por la más frecuente, y ante enfermedades con la misma frecuencia de aparición se escoge aquella cuyo tratamiento es más eficaz a la hora de lograr una mejoría.

Rouleau y col. [134] en 1991 desarrollaron un Sistema Experto, el prototipo CLERTORE, aplicado al diagnóstico y tratamiento del glaucoma primario de ángulo abierto (glaucoma crónico simple). Utiliza reglas de producción. El generador del SE, que comprende el motor y la interfaz, está escrito en Pascal. La base de conocimiento tiene 500 reglas. Un árbol de decisión guía la consulta del sistema. El motor de inferencia trabaja en lógica de proposiciones y en lógica monótona.

Martin [108], que ya había desarrollado en años anteriores un sistema experto para la interpretación de campos visuales, desarrolló en 1996 un sistema para ayuda al diagnóstico del glaucoma en el que tiene en cuenta más factores que los datos del campo visual.

Los resultados de algunas pruebas normalmente muestran una gran variabilidad, y los bordes entre los valores normales y los de enfermedad no están bien definidos. Debido a esto, y a que la progresión de la enfermedad es lenta, es muy difícil determinar el estado de la enfermedad (estable o progresivo). Por este motivo, ve la conveniencia de desarrollar un sistema borroso.

La metodología que Martin [108] sigue al desarrollar su sistema basado en conocimiento es como una espiral que va recorriendo 4 cuadrantes progresivamente y sin final. Estos cuadrantes son planificación, adquisición del conocimiento (incluye elicitación y representación), codificación del conocimiento, y evaluación del sistema basado en conocimiento.

Planificación. En esta fase se debe elegir una herramienta apropiada para construir el sistema. Martin usó CLIPS como ya había hecho en su anterior programa. CLIPS genera código C e incluye el algoritmo Rete. Este algoritmo es

un método eficiente para confrontar reglas contra hechos y producir un programa compacto y rápido.

Adquisición del conocimiento y codificación. Manejar incertidumbre es una tarea habitual en glaucoma. Hay muchas formas de manejar incertidumbre en un sistema basado en conocimiento. Este programa usa una estrategia similar a la de un controlador borroso. El programa también puede describirse como una red neuronal especializada con pesos predefinidos y fijados en todos los nodos y conexiones. El programa usa información sobre la edad, campo visual, antecedentes glaucomatosos familiares, PIO, la ocurrencia de pseudoexfoliaciones o lesiones vasculares, y si el paciente tiene o no tratamiento contra el glaucoma (7 informaciones). En el prototipo del programa los dos primeros hechos pueden recuperarse automáticamente a partir de la prueba del campo visual, y los demás se introducen a través del teclado.

El programa decide el grado de glaucoma calculando un índice, un número entre 0.00 (que indica ausencia total de signos de Glaucoma Crónico Simple) y 1.00 (que indica el cuadro clínico completo de glaucoma). Esta parte del programa es algorítmica y usa procedimientos de la teoría de conjuntos borrosos. Además, el programa identifica Hipertensión ocular. El índice es recombinado con las siete piezas originales de información. Esta parte del programa consiste en 50 reglas de producción aproximadamente. La salida es una de 25 posibles acciones (es decir, ninguna acción, tratamiento recomendado, volver a revisión al cabo de un año). Las bases para las recomendaciones fueron extraídas de un documento de consenso de distintos expertos y a partir de una entrevista de estudios empíricos recientes sobre la correlación entre la PIO y el daño del nervio óptico.

Resultados. La concordancia entre el experto y el programa no fue muy buena especialmente en algunos casos, pero si se ajustaba el programa al modo de actuar de cada uno, los resultados fueron muy buenos (concordancia de 92% a 100%). En conclusión, un programa informático parece proporcionar un buen soporte de decisión en la gestión del glaucoma.

Posteriormente, Martin [109] realizó un estudio con el objetivo de averiguar cómo el soporte de decisión por ordenador influye en la eficacia del cuidado y en la satisfacción subjetiva del trabajo compartido entre oftalmólogos y enfermeras para la gestión de pacientes con glaucoma. El programa ya desarrollado por el autor en CLIPS puede proporcionar ayuda en este sentido, es decir, en la toma de decisiones.

Este tipo de programas ha de ser aceptado por el usuario final. Eso depende de varios factores como la facilidad de uso, accesibilidad, tiempo de respuesta y beneficios prácticos en la situación del mundo real. Este programa se ha integrado en el perímetro de resolución comúnmente usado en Escandinavia para la prueba del campo visual, y se ejecuta inmediatamente después de esta prueba. Requiere introducir muy pocos datos a partir del teclado y el tiempo de respuesta es menor de 1s. Por tanto, no se espera que el procedimiento tenga ningún impacto negativo en el trabajo clínico.

2.4.2 Sistemas híbridos simbólicos-conexionistas borrosos

Zahlmann y col. [171, 172] en 1996 describen un SBC de monitorización de glaucoma y retinopatía diabética que han diseñado para el oftalmólogo privado. El monitor detecta situaciones críticas en el estado del paciente por medio de un proceso de decisión modular, que consiste en un conjunto de varias reglas borrosas, un clasificador de redes neuronales jerárquico para perimetría y una parte de procesamiento de imagen interactivo para imágenes del fondo de ojo. La salida es un mensaje al usuario conectando el sistema local con un sistema de información basado en la web para posteriores recomendaciones.

Estas enfermedades se producen con mucha frecuencia, y es importante, por tanto, la detección precoz y la monitorización a largo plazo de los cambios que va sufriendo el paciente y las decisiones sobre la terapia que se debe aplicar.

El enfoque que dan es el de un sistema de información basado en el WWW, una aplicación de tele-medicina para poder facilitar un servicio actualizado más rápido a todos los usuarios conectados a este sistema.

Sistema de monitorización basado en conocimiento

Las principales tareas son:

1. Clasificación del estado del ojo, detección de situaciones críticas.
2. Presentación de resultados críticos y recomendaciones para acciones relacionadas.

Para resolver estas tareas son necesarias cuatro componentes principales de un sistema de monitorización basado en conocimiento:

- unidad de entrada de datos/información
- unidad de procesamiento de información
- unidad de presentación de resultados
- unidad de recomendaciones

El glaucoma se diagnostica por medio de la PIO, exámenes de la papila y perimetría automática. La PIO se da en forma de números reales, los datos de la papila se describen verbalmente y el software de la perimetría produce unos ficheros ASCII con información sobre las localizaciones de los estímulos y la pérdida de sensibilidad detectada.

El conocimiento sobre la PIO y los datos de la papila pueden describirse por medio de reglas borrosas. Los conjuntos borrosos se usan debido a la traslación directa del conocimiento médico de un modo verbal a un conjunto de reglas (“relación excavación/papila regular”, “PIO dentro de un rango normal”). Para conseguir esta información es necesario traducir los datos originales a este meta-nivel usando procedimientos de emborronado. La clasificación de los datos de perimetría tiene que responder dos cuestiones diferentes:

- ¿qué diagnóstico diferencial está relacionado con este campo visual?
- ¿hay progresión en el tiempo de los defectos del campo?

La primera pregunta será respondida usando un clasificador basado en redes neuronales con los valores de sensibilidad como entradas y el diagnóstico diferencial como salida. La clasificación se realiza mediante una red jerárquica que consta de dos redes diferentes. Las dependencias con el tiempo se manejan usando redes también, pero con una estructura diferente. Los cambios se detectan de una forma más individual y se evalúan en paralelo en la parte de diagnóstico diferencial. La salida es entonces procesada en el conjunto de reglas borrosas para llegar al diagnóstico diferencial final.

La decisión sobre si este diagnóstico diferencial final requiere una acción por el usuario ("situación crítica") se hace usando un conjunto de reglas especiales. Si tal reacción es necesaria, el sistema monitor da un mensaje al usuario en la pantalla y presenta el resultado. Da un primer consejo usando la misma base de reglas. Además, presenta conocimiento más detallado y explícito en forma de recomendaciones. La estructura estática del modelo de dominio usado por el sistema de monitorización es básicamente el mismo que el usado por el sistema de información basado en conocimiento oftalmológico (O-KIS) mejorado por el conocimiento relativo al tiempo.

El objetivo del sistema O-KIS es proporcionar información oftalmológica de contexto específico. El contexto específico de la información depende de los datos de los pacientes, del tipo de usuarios y de la intención de éstos sobre el contenido de la información (tipo de pregunta).

El clasificador basado en redes neuronales consiste en un perceptrón entrenado mediante retropropagación del error (BP). Para investigar la importancia de la información espacial usan tres enfoques diferentes:

1. "Caja negra": todos los puntos de estímulo del perímetro dan una información de entrada al perceptrón (no importa qué patrón estándar de CV pudiera existir, los defectos pueden ocurrir en cualquier lugar).
2. Grupos de nervios según Weber o según una definición propia, ambas basadas en el conocimiento sobre los caminos neurales dentro del sistema visual.
3. Cuadrantes superior/inferior. Clasificación según los cambios en una u otra parte, separadamente.

El número de nodos de salida lo determinaron según las clases que tuviera que diferenciar. Tres nodos si tenía que clasificar entre "no patológico", "patológico no glaucomatoso" y "patológico glaucomatoso" en un solo paso. En otra unidad de entrenamiento diseñaron dos redes neuronales con dos nodos de salida cada una dividiendo la tarea de clasificación en dos subtareas. La primera red clasificaba entre normales y patológicos, y después éstos eran clasificados por la segunda red en glaucomatosos o no glaucomatosos.

Para entrenar la red usaron conjuntos de datos preclasificados por expertos médicos (2/3 para entrenamiento, del 1/3 restante 2/3 para prueba y el resto para evaluación).

El primer prototipo de este programa se probó con 182 pacientes (126 para entrenamiento/prueba de la red y 56 para evaluación). Los resultados obtenidos fueron los siguientes (Tabla 2.10):

	Tasa de aciertos (%)		
	red 1: "caja negra"	**red 2: grupos de nervios**	**red 3: cuadrantes superior/inferior**
clasificación en 1 paso			
Normal	88%	88%	71%
no-glaucomatoso	42%	25%	50%
glaucomatoso	83%	89%	83%
dos pasos (2 nodos de salida en cada paso)			
normal	88%	88%	75%
patológico	87%	83%	87%
no-glaucomatoso	90%	80%	95%
glaucomatoso	72%	89%	78%

Tabla 2.10: Resultados obtenidos por Zahlmann [171, 172] en la clasificación de campos visuales utilizando MLP con BP.

Las reglas borrosas que usaron para este primer prototipo sirvieron únicamente para seleccionar el paciente de la base de datos como sospechoso de glaucoma y después la red neuronal se encargó de su clasificación.

En 1997 y después en el 2000, Zahlmann y col. [173, 176] mejoraron el anterior prototipo. Desarrollaron ese sistema monitor basando la parte de redes neuronales en un algoritmo de selección de características especial implementado por ellos y en una red de función de base radial (RBF), y combinando los resultados de esta parte conexionista con todos los demás datos de los pacientes mediante un conjunto de reglas borrosas para obtener un diagnóstico final.

Los datos que obtuvieron de la perimetría se los dieron a tres expertos para su clasificación y comprobaron que no coincidía el diagnóstico en muchos casos. Eso reforzó la idea de que el diagnóstico del CV es una tarea muy difícil que conlleva un grado alto de incertidumbre. Como consecuencia, los resultados de la clasificación realizada por una red neuronal basada en la decisión de un experto deben tomarse con la incertidumbre que tienen cuando se usen como entradas de las reglas borrosas.

Entre los datos del paciente que usaron están la relación excavación/papila (E/P), la localización de la excavación, y una comparación de las relaciones E/P de los dos ojos. La E/P y la diferencia entre las E/P de los ojos derecho e izquierdo se dan como números reales, la localización en términos lingüísticos como "central", "inferior", "superior". Las E/P y las diferencias son transformadas

por los expertos médicos con la ayuda del ingeniero del conocimiento a un meta-nivel de clasificación, evaluándolas como “normal”, “incrementada”, etc.

Otra fuente de información es lo que cuenta el propio paciente en la consulta.

A partir de todos estos datos el oftalmólogo debe decidir si se trata de un paciente con glaucoma o no, y qué tipo de reacción se requiere: volver a revisar al cabo de un tiempo más corto o más largo, aplicar un tratamiento, intervenir quirúrgicamente, etc.

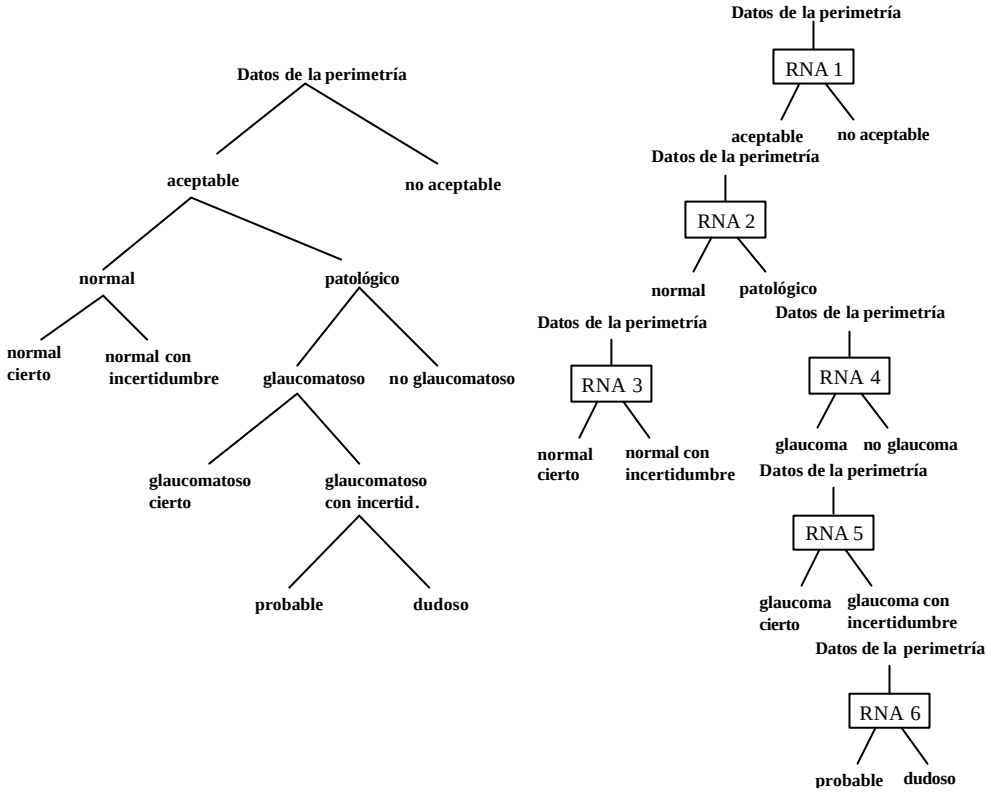


Figura 2.19: Red jerárquica RBF para datos de perimetría [173].

El objetivo de este SE es servir de ayuda a médicos que trabajen en atención primaria (oftalmólogos privados, optómetras, hospitales pequeños de ojos).

Un problema general para diseñar RNA con propiedades de generalización válidas es la dimensionalidad, es decir, el número de ejemplos necesarios para el entrenamiento aumenta exponencialmente con la dimensión de entrada. Los vectores de entrada de los datos perimétricos tienen 59 componentes, mientras que la base de datos para entrenamiento/evaluación contenía inicialmente sólo 1300 ejemplos preclasificados. Por tanto, vieron necesario desarrollar un método para reducir dimensionalidad y aplicarlo como un paso de preprocesamiento a los datos perimétricos. Desarrollaron una función tipo que tiene su mínimo global cuando las distancias de los vectores de características que pertenecen a la misma clase (distancias intra-clases) son cero, y las distancias de los vectores de clases

diferentes (distancias inter-clases) son máximas según las restricciones del espacio de medida.

El diseño del clasificador está basado en el árbol de la Figura 2.19. El motivo de usar ese árbol es introducir varios estados de decisión, que van desde decisiones bastante generales como la clasificación entre "normal"/"patológico" hasta decisiones refinadas como la clasificación "sospechoso"/"probablemente glaucomatoso". De esta forma, diseñan una RNA especializada para cada nivel de decisión desde la raíz hasta las hojas del árbol de clasificación. Por tanto, cada red discrimina entre dos clases. Cada ejemplo va pasando por todas las RNA desde la 1 hasta la 6 para su clasificación. Si por ejemplo el clasificador RNA 2 da que es más "normal" que "patológico", se presta más atención a los resultados de la RNA 3 que a los de la 4, 5 y 6. Este método de interpretación es posible debido a que usan redes RBF, que hacen una interpolación entre los ejemplos de una región del espacio de características definido. Una ventaja de esta clasificación jerárquica, además de la especialización, es que se cubre el espacio completo de características, es decir, que cada caso puede ser asociado correctamente a su clase de la jerarquía. Además, uno de los refinamientos se refiere a la parte de mayor interés que es la clase de casos glaucomatosos.

Al igual que con el prototipo anterior que habían desarrollado, entrenaron con 2/3 de los casos y tomaron el tercio restante para la prueba/evaluación. Los resultados obtenidos se muestran en la Tabla 2.11:

Jerarquía	**Todas las características. Sens/Espec**	**Caract. Reducidas**	**Nº de características rechazadas**
RNA 1	85/76 %	83/81 %	5
RNA 2	82/93 %	85/93 %	2
RNA 3	91/91 %	91/92 %	1
RNA 4	71/51 %	72/71 %	29
RNA 5	71/52 %	79/56 %	16
RNA 6	69/61 %	62/74 %	36

Tabla 2.11: Resultados de la red especializada.

Los datos de PIO, E/P y diferencias de E/P entre ambos ojos se introdujeron directamente como números reales y después fueron emborronados mediante funciones de pertenencia basándose en el conocimiento del experto médico. La PIO tenía una función miembro en forma de Z y la E/P fue emborronada usando una función miembro en forma de S. El emborronado de la entrada es el primer nivel de abstracción que transforma valores en variables lingüísticas. Ya se indicó que la localización de la excavación es directamente usada como una variable lingüística de entrada. Los datos del CV se dan en términos lingüísticos seguidos por ".yes" o ".no" según la clase de salida.

La inferencia borrosa de la parte IF de las reglas se da usando los operadores MIN/MAX. La inferencia de la parte THEN usa un grado de soporte

(DoS) dado por el experto médico a cada regla, ponderando su influencia específica sobre la decisión de subclase: $\mu_{THEN} = \mu_{IF} * DoS$.

Las variables de salida de los pasos de decisión del nivel inferior se usan como entradas al siguiente nivel de abstracción. La salida final puede usarse como salida borrosa o nítida. Los métodos de aclarado que utilizan son el del centro de los máximos y el de la media de los máximos.

Nivel 1:

Las primeras subclases describen la PIO final y la primera combinación de PIO y E/P con las subclases "glaucoma" y "glaucoma normotensivo". Estas subclases de situación se usan como salida al usuario para simulación y propósitos de desarrollo. También son las entradas a los niveles de abstracción más altos y, por tanto, parte de la evaluación de la situación final.

Nivel 2:

Este nivel combina las subclases del primer nivel con los resultados de las redes neuronales de las clases de perimetría. La perimetría se puede considerar sola o en combinación con los parámetros antes mencionados para obtener una visión global y por tanto un diagnóstico final.

Nivel 3:

Este nivel describe una decisión difícil en otro nivel de abstracción, considerando la combinación del segundo nivel de abstracción de todas las categorías de entrada (PIO, E/P, CV) evaluadas de nuevo con las subclases de la combinación PIO y E/P (nivel 1). Eso quiere decir que retrocedemos en la evaluación de la situación ya que la tarea de decisión ahora experimenta una modificación hacia un "diagnóstico diferencial diferenciado" en la aplicación médica, averiguando el tipo especial de glaucoma. A este paso le sigue la decisión final entre las categorías de normal, sospechoso de glaucoma, glaucomatoso y otras patologías. La conclusión final se da en su forma borrosa, o mediante aclarado.

La parte de diagnóstico del CV se entrenó y evaluó finalmente con 2225 casos (Tablas 2.12 y 2.13). La precisión diagnóstica del clasificador se mejoró un 5% con el preprocesamiento de datos.

Sensibilidad	**Especificidad**	**Subtarea**
98%	95%	normal/patológico
98%	97%	normal cierto/normal con incertidumbre
82%	81%	glaucomatoso/patológico no glaucomatoso
79%	84%	glaucomatoso cierto/glaucomatoso con incert.

Tabla 2.12: Resultados de la clasificación de la red RBF (entrenamiento).

Sensibilidad	**Especificidad**	**Subtarea**
98%	95%	normal/patológico
98%	97%	normal cierto/normal con incertidumbre
72%	70%	glaucomatoso/patológico no glaucomatoso
71%	72%	glaucomatoso cierto/glaucomatoso con incert.

Tabla 2.13: Resultados de la clasificación de la red RBF (prueba).

Las reglas borrosas se refinaron con 58 casos clínicos de prueba. La estimación de la situación dada por el sistema fue evaluada por los expertos médicos (oftalmólogos clínicos y privados) utilizando otra muestra distinta de casos clínicos (Tabla 2.14).

	Sensibilidad	**Especificidad**
Clínica de ojos	98%	100%
Oftalmólogos privados	82%	67%

Tabla 2.14: Sensibilidad y especificidad de la estimación final de la situación.

Una conclusión a la que llegan es que la diferenciación entre los casos glaucomatosos ciertos y con incertidumbre es muy difícil, pero si a los datos del CV se añaden los de la PIO y papila del nervio óptico se pueden diferenciar esas dos clases mejor.

En algunos casos, es difícil diferenciar entre glaucomatosos y otras patologías. Están refinando el sistema para lograr mejores resultados en este aspecto.

El trabajo futuro lo tienen planteado principalmente en dos puntos: conseguir detectar cambios dependientes del tiempo en los CV, y conseguir una base de casos mayor.

3 Sistemas Expertos: Primeras versiones

En este capítulo se exponen los inicios de este trabajo de investigación. Se comenzó implementando un prototipo de sistema experto mediante una herramienta de desarrollo (*shell*) llamada M1. En este primer desarrollo se adquirió gran parte del conocimiento sobre el diagnóstico del glaucoma. Posteriormente, se desarrolló otro sistema experto más completo y flexible utilizando para ello un lenguaje de programación convencional. Estos comienzos en la investigación han tenido gran influencia en el último sistema desarrollado, objetivo de esta Tesis. Por este motivo, vamos a exponer los aspectos que consideramos más relevantes con cierto detalle.

3.1 Prototipo programado en M1

La exposición de este sistema experto que se desarrolló en 1990 como trabajo de licenciatura [141], se ha estructurado de la siguiente forma: primeramente comenzamos comentando cómo se ha realizado la adquisición del conocimiento. Posteriormente recordaremos cuál es la arquitectura básica de un sistema experto. Puesto que se ha utilizado un *shell* de desarrollo de sistemas expertos, no se ha introducido ningún módulo que no estuviera ya implementado en M1. A continuación se indican los motivos de la elección de este software de desarrollo. Finalmente se aborda el proceso de inferencia que sigue para dar un diagnóstico y cómo se han asignado los factores de certeza en este proceso.

3.1.1 La adquisición del conocimiento

La adquisición del conocimiento ha sido reconocida ampliamente como un arte y como un cuello de botella en la construcción de Sistemas Expertos [132]. Es el problema central de la elaboración de la base de conocimiento.

El ingeniero del conocimiento (IC) juega un papel crítico en la construcción de un SE. Aunque lo que se debe modelar es el conocimiento del experto, el IC es el que debe construir realmente el sistema. El proceso de llevar el conocimiento del experto al sistema actuando de intermediario es crítico porque al final el conocimiento del IC, más que el del experto, es el reflejado realmente en un SE [132].

El primer paso a la hora de construir un SE es la comprensión del dominio, familiarizarse con el tema que se va a tratar. Para ello hay varios caminos. Siempre se cuenta con una bibliografía especializada que puede proporcionar el experto. De este modo los conceptos que éste utiliza se hacen propios y esto facilita mucho las conversaciones posteriores con él.

Sin embargo no se puede basar toda la adquisición del conocimiento en la lectura de libros adecuados puesto que está comprobado que los expertos en la clínica diaria no siguen totalmente lo que enseñan a los alumnos. En la clínica diaria lo que cuenta principalmente es la experiencia, esto hace que el experto muchas veces no sepa analizar los pasos que sigue para elaborar un posible diagnóstico que puede ser contrario a lo que la bibliografía enseña.

En el caso que nos ocupa -diagnóstico del glaucoma- se han empleado muchos métodos de adquisición y elicitación del conocimiento, como son:

- La lectura de bibliografía.
- La realización de un curso de glaucoma impartido por oftalmólogos del Hospital Clínico Universitario de Valladolid. El conocimiento adquirido de esta manera fue muy interesante puesto que todos los expertos hablaban de su experiencia frente a casos reales y no frente a casos de libros. También se pudo observar que el tema es muy controvertido y que las opiniones acerca del límite entre los casos normales y patológicos son diversas ya que es un límite muy borroso, así como respecto al momento en que se debe tratar farmacológicamente al paciente.
- **Entrevistas.** Conversaciones con el experto. Éstas inicialmente sitúan al "ingeniero del conocimiento" en el tema pero a medida que el conocimiento es mayor va tomando conciencia de la complicación del problema. Si la fase del glaucoma es avanzada el diagnóstico es fácil para el oftalmólogo y, por lo tanto, fácil para un Sistema Experto, pero si lo que se pretende es un diagnóstico precoz el problema deja de ser sencillo, pues el nivel de imprecisión e incertidumbre es muy alto.
- **Técnicas observacionales.** La observación del comportamiento del experto en la consulta diaria.
- **Análisis de protocolo.** Se le pide al experto que explique lo que hace mientras lleva a cabo la revisión de cada paciente.
- **Casos de estudio.** Se le presentan casos al experto y se discuten con él. Es importante cubrir todas las eventualidades.

Todos estos métodos de adquisición del conocimiento han intervenido positivamente, cada uno en el momento y con la finalidad para la que se usó. La lectura de bibliografía fue una forma rápida de situarse al principio, pero si he de destacar algún método son las entrevistas con el experto y los casos de estudio: sin ellos no hubiera sido posible el desarrollo del este sistema.

Las primeras entrevistas con el experto pueden ser de carácter general para situarse en el tema y consolidar los conocimientos adquiridos paralelamente mediante la bibliografía, pero después, el "ingeniero del conocimiento" necesita concretar para poder iniciar un prototipo. Para ello se elaboró una pequeña encuesta a la que se sometió al experto, es decir, se realizó una **entrevista estructurada**. Esta encuesta es la siguiente:

1. Factores de riesgo que pueden influir con mayor peso en un posible diagnóstico de glaucoma. ¿Con qué peso influyen?
2. Preguntas que se le hacen a un paciente que llega por primera vez a una Sección de Glaucoma.
3. Pruebas a las que se le somete.
4. Parámetros con los que se cuenta para dar un diagnóstico.
5. Relaciones entre los parámetros según sus magnitudes y posibles diagnósticos (grafos de decisión).

Respuestas

		Pesos:
1.	Factores de riesgo:	
	a) Antecedentes familiares de glaucoma. Herencia.	***
	b) Miopía elevada.	**
	c) Diabetes.	**
	d) Enfermedades cardiovasculares.	**
	e) Edad mayor de 40 años.	*

2. Preguntas al enfermo:
   a) Relacionadas con los factores de riesgo.

3. Pruebas:
   a) Agudeza visual. Refracción.
   b) Presión intraocular.
   c) Relación Excavación/Papila.
   d) Tonografía (opcional).
   e) Gonioscopia.
   f) Campo visual.
   g) Curva tensional (opcional).

4. Parámetros:

a) Agudeza visual:	No tiene ningún valor, sólo afecta en el glaucoma terminal

b) Presión intraocular:	• Normal	hasta 20
	• Dudosa	21-24
	• Patológica	≥ 25 (cuanto más alta, más posibilidades de diagnóstico de glaucoma).
c) Relación E/P:	• Normal	0-2/6
	• Dudosa	2-3/6
	• Patológica	≥ 3/6
d) Tonografía. C:	• Normal	≥ 0.15
	• Dudosa	0.11-0.14
	• Patológica	≤ 0.1
e) Gonioscopia	• Amplitud del ángulo irido-corneal	
	• Pigmentación	
f) Campo visual:	• Normal	
	• Dudoso	
	• Patológico incipiente	
	• Patológico avanzado	
g) Curva tensional:		
- Presión máxima	• Normal	≤ 20
	• Dudosa	21-24
	• Patológica	≥ 25
- Oscilación	• Normal	≤ 5
	• Dudosa	6-7
	• Patológica	> 7

5. Relación entre los parámetros:
   El experto ha elaborado unos grafos de decisión conjugando las respuestas de todas las preguntas de la encuesta. Los grafos iniciales se han ido mejorando a medida que se construía la base de conocimiento y surgían los problemas, y después, en la fase de validación, al introducir datos de pacientes reales. Estos grafos se muestran en las Figuras 3.1, 3.2, 3.3 y 3.4.

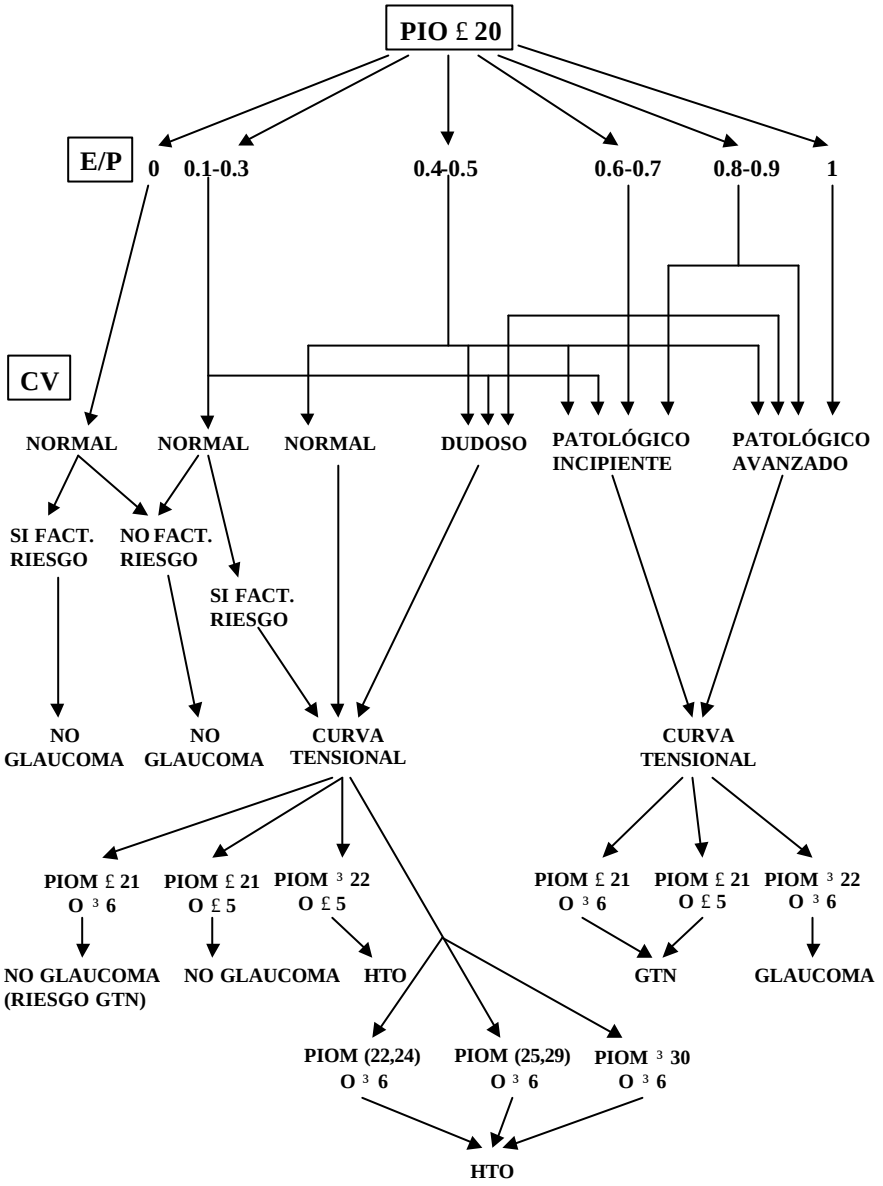


Figura 3.1: Grafo de decisión 1.

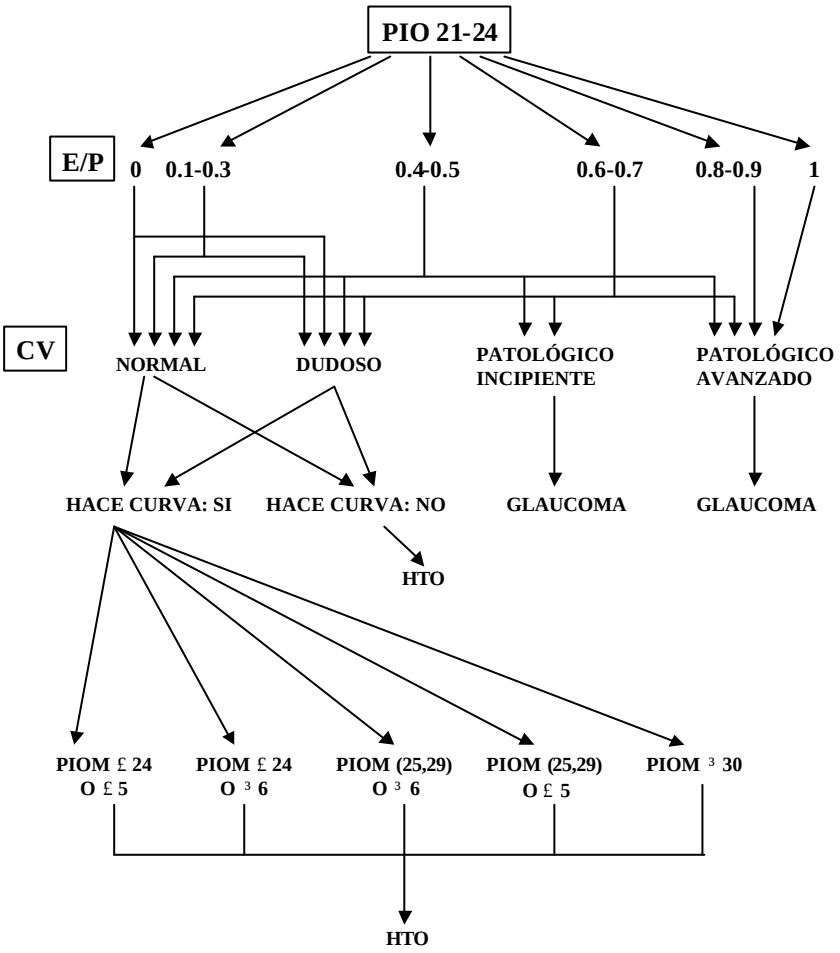


Figura 3.2: Grafo de decisión 2.

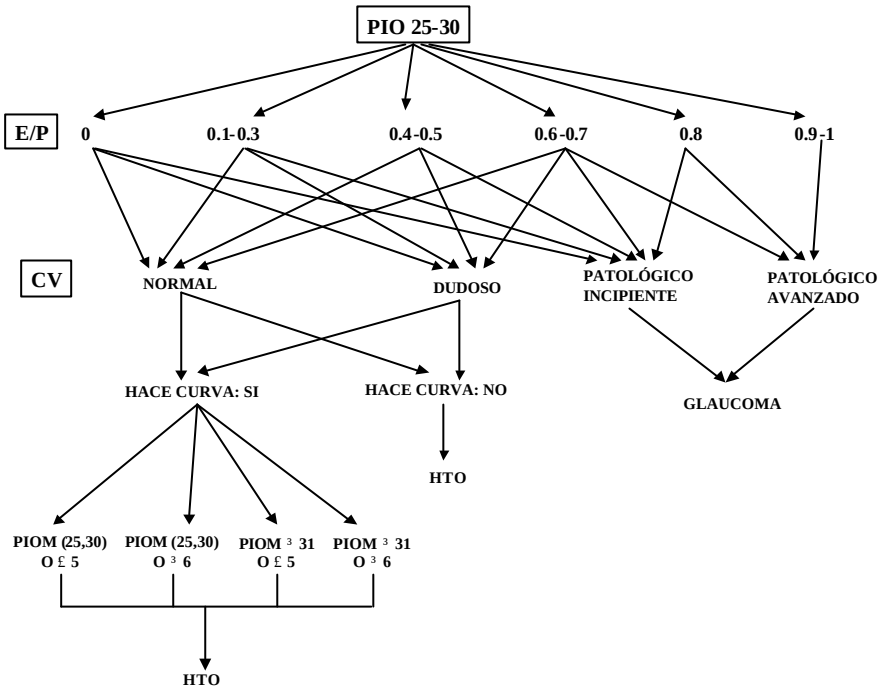


Figura 3.3: Grafo de decisión 3.

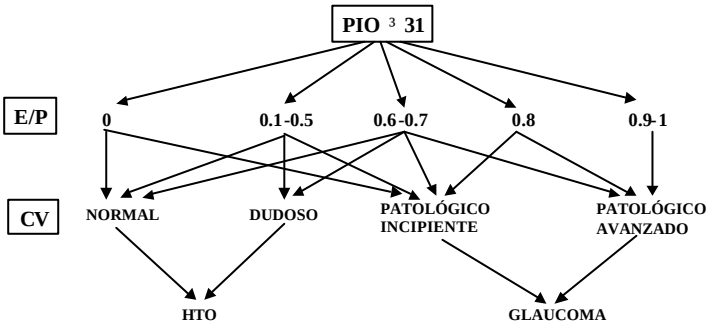


Figura 3.4: Grafo de decisión 4.

3.1.2 Arquitectura básica de un Sistema Experto

Un Sistema Experto consta de los módulos siguientes (Figura 3.5):

- **Base de conocimiento**: contiene el conocimiento del dominio necesario para resolver problemas codificado en forma de reglas. Éstas son un

paradigma general para representar el conocimiento, pero también hay otros tipos de SE que usan diferentes representaciones.

- **Interfaz de usuario**: es el mecanismo por el que el usuario y el SE se comunican.
- **Subsistema de explicación**: explica el razonamiento del sistema al usuario.
- **Motor de inferencia**: hace inferencias decidiendo qué reglas son satisfechas por los hechos u objetos, prioriza las reglas satisfechas, y ejecuta la regla con la prioridad más alta.
- **Subsistema de adquisición del conocimiento**: proporciona un modo automático para el usuario de introducir conocimiento en el sistema en lugar de tener que codificarlo explícitamente el ingeniero del conocimiento.

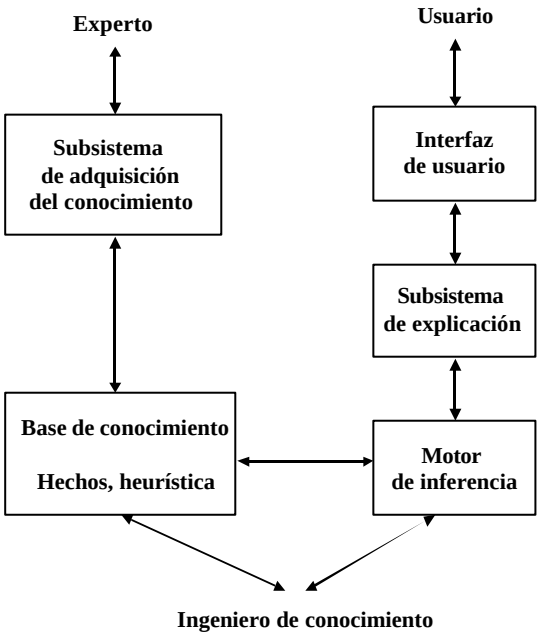


Figura 3.5: Componentes de un Sistema Experto [122].

3.1.3 Software para construir Sistemas Expertos

Las principales herramientas software de que se dispone para desarrollar SE se pueden dividir en las siguientes categorías [31]:

- Lenguajes de programación.
- *Shells* de desarrollo de SE.
- Herramientas de Inteligencia Artificial.

Los lenguajes de programación a su vez pueden ser:

- **Lenguajes convencionales,** también llamados orientados al problema, como C, Smalltalk, Ada, etc. Una de sus principales ventajas es la posibilidad de disponer de interfaces convencionales con bases de datos u hojas de cálculo. De todas formas, muchas de las herramientas específicas de desarrollo de SE también disponen de esas facilidades.
- **Lenguajes de IA,** también llamados de manipulación de símbolos porque han sido diseñados para aplicaciones de IA. Los más comunes son Lisp y Prolog.

Desarrollar SE usando lenguajes convencionales de programación puede llevar mucho tiempo porque se debe desarrollar la interfaz de usuario e implementar el motor de inferencia usando las estructuras disponibles en el lenguaje. Este tiempo de desarrollo puede disminuir mucho si se utilizan herramientas de desarrollo específicas como son los *Shells de Sistemas Expertos*.

Los shells son SE en los se han eliminado las reglas. Esto permite a los desarrolladores que se puedan centrar sólo en la base de conocimiento evitándoles la tarea de programar el motor de inferencia, la interfaz de usuario y el módulo de explicación. No obstante, usar este tipo de herramientas puede sobresimplificar el dominio de la aplicación ya que los shells son inflexibles y es difícil modificar o cambiar su modo de funcionamiento respecto a la representación del conocimiento y al mecanismo de inferencia.

Entre los shells disponibles comercialmente están AM, Leonardo, EXSYS, M1, etc.

Además existen otras herramientas para contruir SE como son las *Herramientas de Inteligencia Artificial.* Son herramientas híbridas bastante sofisticadas que contienen típicamente estructuras de código para un rango de tareas de SE. Pueden usar reglas, marcos, programación orientada a objetos, y lógica o redes semánticas. Pueden utilizar encadenamiento hacia delante o hacia atrás, Razonamiento Basado en Casos, y una gran variedad de técnicas de herencia. Debido a su complejidad, requieren más habilidades del programador que los shells o los lenguajes de programación. Así como los shells son muy apropiados para pequeñas aplicaciones, éstas son más apropiadas para aplicaciones más grandes cliente/servidor. Entre éstas está ART-IM, Level 5 Object, etc.

La selección de las herramientas apropiadas para construir SE es más difícil que con sistemas convencionales ya que muchas no son específicamente adecuadas para clases particulares de problemas. La construcción de un SE es generalmente considerado como un proceso iterativo. Esto también se aplica a la elección de la herramienta, por lo que muchas veces la herramienta seleccionada para el prototipo inicial tiene que ser cambiada después como parte del proceso iterativo. Esto podría ser debido a las siguientes razones:

- El programador no era consciente de las limitaciones de la herramienta seleccionada.
- Pueden haberse olvidado algunos factores importantes durante el análisis de requisitos.
- Es posible que se haya elegido una herramienta para desarrollar el prototipo inicial con la intención de cambiarla, si fuera necesario, al implementar el sistema completo.

En nuestro sistema seleccionamos el shell M1 para desarrollar el prototipo debido a las ventajas que ofrecía en el momento en el que se utilizó. Por otra parte, el disponer de experiencia en el grupo de investigación sobre ella facilitaba enormemente su uso. La intención inicial fue la de que sirviera para hacer el prototipo y así ir también adquiriendo el conocimiento de una forma relativamente sencilla, para más tarde elegir otra para desarrollar el sistema completo.

M1 era una de las herramientas de desarrollo de sistemas expertos más potentes. Tenía capacidad para desarrollar y usar sistemas expertos de más de 1000 reglas en un IBM PC, XT, AT o compatible, con sistema operativo MS-DOS 2.0 o superior.

La búsqueda en la base de conocimiento se realiza con la técnica del encadenamiento hacia atrás.

La base de conocimiento, constituida por hechos y reglas fundamentalmente, se puede crear usando un editor de textos estándar y cargarla después en el sistema. M1 proporciona el motor de inferencia y la interfaz de usuario.

Una característica importante de M1 es la posibilidad de representar y utilizar conocimiento dudoso. El grado de credibilidad de un hecho se indica con su factor de certeza (CF).

M1 permite la creación de un fichero preprocesado que contiene una base de conocimiento y las conclusiones alcanzadas por el sistema en el momento de la creación del fichero. Este se carga junto a M1 en su versión "runtime" formando lo que se llama SE para el usuario final.

Tiene también la posibilidad de desactivar comandos para proteger la integridad y confidencialidad de la base de conocimiento.

M1 puede llamar a subrutinas escritas en lenguaje ensamblador o en lenguajes de alto nivel como C.

3.1.4 Proceso de inferencia a lo largo de la base de conocimiento

Siguiendo la sintaxis propia de M1 se desarrolló un prototipo de SE para ayuda al diagnóstico del glaucoma. M1, como cualquier herramienta de desarrollo de Sistemas Expertos, evita al programador la tarea de programar el motor de

inferencia, la interfaz de usuario y el módulo de explicación, permitiéndole centrarse en la introducción del conocimiento en la base de hechos y reglas del sistema.

En este apartado vamos a ver cómo se estructuró todo ese conocimiento utilizando el formalismo propio de la herramienta de desarrollo seleccionada.

Al comenzar una consulta, el Sistema Experto interroga acerca de los factores de riesgo que podrían tener un peso importante a la hora del diagnóstico.

Estos factores de riesgo son los antecedentes de glaucoma en la familia, tener una miopía elevada, padecer diabetes o enfermedades cardiovasculares y la edad. El de mayor peso es la herencia.

Seguidamente M1 pregunta de qué ojo se quiere obtener el diagnóstico, ya que este sistema analiza cada ojo por separado, y va disparando las reglas necesarias para obtener el diagnóstico final.

La mayor parte de las reglas de la base de conocimiento tienen una estructura similar y van encaminadas a diferenciar en lo posible el diagnóstico final según se combinen los diferentes datos obtenidos de las exploraciones. Se han obtenido expresando cada rama de los grafos de decisión mediante la sintaxis propia del M1.

Estas reglas dan como conclusión un diagnóstico 1 del ojo correspondiente. Hasta llegar al diagnóstico final habrá que pasar por tres diagnósticos (Figura 3.6).

Se estructuró el conocimiento en base a cuatro grupos de parámetros, cada uno de los cuales aportará un mayor acercamiento al posible diagnóstico.

El grupo 1 lo constituye la presión intraocular, la excavación de la papila, el campo visual y la curva tensional (opcional en la mayor parte de los casos). Con este grupo M1 elabora el primer diagnóstico.

El grupo 2 está formado por la tonografía (opcional en todos los casos). Éste da lugar al diagnóstico 2.

El grupo 3 lo constituyen los factores relacionados con la excavación de la papila del nervio óptico como son la ovalización vertical, la excentricidad, el tener alguna muesca, cambios en la trayectoria vascular, hemorragias en astilla o tener bordes netos. Con éste se elabora el diagnóstico 3.

Finalmente el grupo 4 está formado por los factores de riesgo citados anteriormente. M1 dará ya un diagnóstico final.

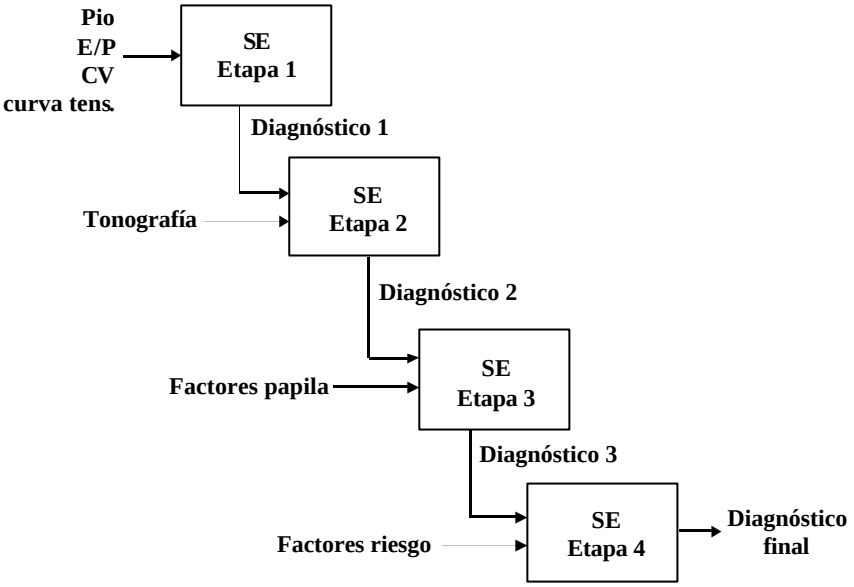


Figura 3.6: Etapas en que se realiza el diagnóstico.

Como ya se ha indicado, M1 usa la técnica del encadenamiento regresivo. Empieza por buscar el diagnóstico final y ve que necesita conocer el diagnóstico 3 y para conocer éste ha de buscar el diagnóstico 2 y así llega al diagnóstico 1. Entonces empieza a interrogar al usuario sobre los parámetros de cada una de las proposiciones que forman la premisa de las reglas.

M1 antes de preguntar sobre la curva tensional en los casos de presión intraocular mayor de 20, y sobre la tonografía, se asegurará de que estas pruebas han sido efectuadas.

Cuando M1 ha encontrado el diagnóstico final y lo ha sacado por pantalla repite este proceso para obtener el diagnóstico del otro ojo si así lo requiere el usuario.

Es importante observar que en este caso el usuario no es interrogado de nuevo sobre los factores de riesgo, puesto que están almacenados ya en la caché.

Si el diagnóstico final del otro ojo ha sido buscado o si la respuesta a la pregunta *¿Va a examinar el otro ojo?* ha sido negativa, termina la consulta.

Finalmente, M1 pregunta al usuario si desea salir del programa o comenzar una nueva consulta. En el primer caso sale automáticamente al sistema operativo MS-DOS, en el segundo caso M1 limpia la caché de todos los datos de una consulta previa, busca los datos iniciales y vuelve a repetir todo el proceso.

Este proceso de razonamiento que sigue el motor de inferencia del M1 se puede alterar. Esto ha sido utilizado en esta base de conocimiento para que las preguntas hechas al usuario sobre el paciente tuvieran un orden lógico.

Concretamente, se ha usado para alterar el orden en que se interroga sobre los factores de la papila, ya que lo más lógico es que estas preguntas se hagan a continuación del tamaño de la excavación de la papila. De otra forma no hubieran sido efectuadas hasta que M1 hubiese tenido que buscar el diagnóstico 3.

En el transcurso del proceso de razonamiento hacia el diagnóstico final, M1 va sacando mensajes por pantalla intercaladas en las premisas de algunas reglas. Estos mensajes de texto aparecen como hechos a lo largo de la base de conocimiento.

3.1.5 Asignación de factores de certeza

Para entender cómo se han elegido los factores de certeza, es necesario seguir el proceso de razonamiento del motor de inferencia de M1 a lo largo de la base de conocimiento y ver en qué momentos de este proceso se van asignando. Por ello será preciso analizar cada nivel de conclusiones obtenidas. Estos niveles se corresponden prácticamente con los grupos mencionados anteriormente, es decir, cada vez que un diagnóstico es concluido, el factor de certeza se modifica o permanece sin alterar, atendiendo principalmente a leyes totalmente empíricas.

Diagnóstico 1

Como ya hemos visto, para elaborar este diagnóstico, M1 necesita conocer la presión intraocular, la excavación de la papila y el campo visual en todos los casos, y en los casos de baja presión es imprescindible también la curva tensional.

Debido a estas diferenciaciones nos encontramos dos tipos diferentes de reglas y por tanto dos formas distintas de asignar factores de certeza.

Un tipo de reglas es aquél en que son necesarias todas las pruebas mencionadas o bien por su conveniencia han sido efectuadas. Son necesarias e imprescindibles en el grupo de presiones de hasta 20 mmHg.

La curva tensional es conveniente en muchos casos de los grupos de presiones entre 21 y 24, y mayores o iguales de 25. Las reglas en estos casos tienen la misma forma pero con una proposición más que asegura que la curva tensional se ha efectuado.

Hay que observar que la curva tensional deja de ser necesaria cuando el campo empieza a ser patológico, pues entonces el diagnóstico ya es claro.

En todas estas reglas los factores de certeza se asignaron mediante un análisis por parte del experto de la confianza que le merecía cada uno de los diagnósticos que concluían y haciendo un estudio comparativo de todas las reglas en conjunto para que no hubiese contradicción entre ellas.

En cada regla se valoraron todos los parámetros en conjunto teniendo en cuenta que cuanto mayor sea la presión intraocular, cuanto mayor sea la relación E/P, cuanto más patológico sea el campo, y cuanto mayores sean la presión máxima

en la curva y la oscilación, mayor será el riesgo de glaucoma y menor la certeza de que el paciente no lo padece.

El otro tipo de reglas es aquel que no cuenta con la curva tensional para dar el diagnóstico 1.

Obtención del factor de certeza en este tipo de reglas:

El factor de certeza varía con la presión intraocular de la forma indicada en la Figura 3.7.

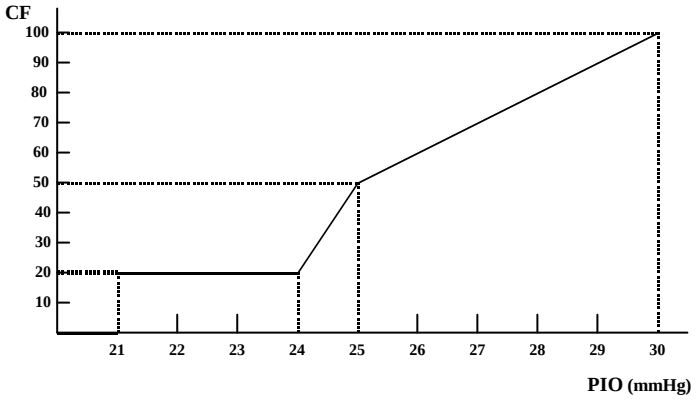


Figura 3.7: Factor de certeza según la PIO.

La variación del factor de certeza respecto a la relación E/P es lineal a trozos al igual que la de la presión intraocular y se muestra en la Figura 3.8.

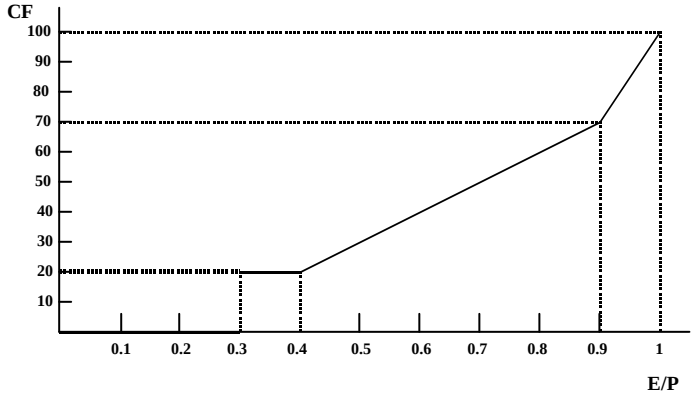


Figura 3.8: Factores de certeza según la E/P.

Con los factores de certeza de la PIO y de la E/P se obtiene el factor de certeza del diagnóstico 1.

Diagnóstico 2

Se elabora, como ya sabemos, a partir del diagnostico 1 y la tonografía.

M1 pregunta si se le ha hecho la prueba tonográfica al paciente, si la respuesta es negativa, el diagnóstico es el mismo y el factor de certeza no varía.

Si se ha realizado la prueba tonográfica:

- Si C es menor de 0.11, es decir, el resultado es patológico, el factor de certeza del diagnóstico_2 queda disminuido en 10 en los casos de no glaucoma, ya que la seguridad de que el paciente no padece la enfermedad es menor.
- Si C está comprendido entre 0.11 y 0.15 el diagnóstico permanece con igual factor de certeza en los casos de no glaucoma.
- Si el diagnóstico tiene un factor de certeza entre 90 y 100 y C es menor de 0.11, el factor de certeza es 100 en los casos distintos al no glaucoma. No se puede sobrepasar ese valor. Sale por pantalla un mensaje que dice que la tonografía apoya un posible diagnóstico de glaucoma.
- Si C es menor de 0.11, se aumenta en 10 el factor de certeza en los diagnósticos distintos al no glaucoma, a no ser que ocurra lo señalado en la regla anterior.
- Si C está comprendido entre 0.11 y 0.15, el sistema actúa igual que con las dos reglas anteriores en los casos distintos al no glaucoma, con la diferencia de que el factor de certeza sólo queda aumentado en 5.
- Si C es igual o mayor de 0.15, la tonografía es normal y por tanto no se modifica el factor de certeza. La tonografía confirma en caso de dar patológica, de otra manera no tiene mayor validez.

Diagnóstico 3

Después de preguntar M1 por la relación E/P, pregunta si hay algún factor de la papila.

- En los casos de no glaucoma todos los factores de la papila disminuyen en 2 el factor de certeza final excepto el tener bordes netos que lo aumenta en 10.
- En los casos distintos al no glaucoma se actúa de manera inversa a lo descrito en el punto anterior: todos los factores aumentan en 2 el factor de certeza final excepto el tener bordes netos que lo disminuye en 10.
- Si no se introducen los factores de la papila el diagnóstico final tiene el mismo factor de certeza que el diagnóstico 3.

Diagnóstico final

- En los casos de no glaucoma los factores de riesgo disminuyen el factor de certeza aplicando la misma ley que en el caso de haber factores de la papila. De esta manera conseguimos que la herencia disminuya en 5 el factor de certeza final y los demás factores de riesgo solamente en 2.
- En los casos distintos al no glaucoma se actúa de manera inversa a lo descrito en el punto anterior: la herencia aumenta en 5 el factor de certeza final y los demás factores de riesgo en 2.
- Si no hay factores de riesgo el diagnóstico final tiene el mismo factor de certeza que el diagnóstico 3.

3.1.6 Perspectivas de desarrollo

Este prototipo de SE da un diagnóstico a un paciente que llega por primera vez a una Sección de Glaucoma bien porque haya sido enviado por un Ambulatorio o bien debido a una revisión rutinaria.

Se ha validado con una batería de casos simulados comparando la salida que da el prototipo con la del experto que ha proporcionado el conocimiento, y en todos los casos, lógicamente, coinciden, puesto que las reglas se han construido siguiendo sus pautas.

Esto es un inicio de lo que en realidad se pretendía y que se ha ido realizando posteriormente. Ante este primer sistema desarrollado se establecieron las siguientes mejoras:

1. Una cuestión importante es el seguimiento del paciente. El SE debería analizar la variación de los parámetros de una consulta actual respecto a las anteriores y dar un pronóstico respecto a la enfermedad.
2. Habría que mejorar también las salidas gráficas de forma que la introducción de los datos fuera más sencilla para el usuario.
3. Por otra parte sería conveniente adaptar el SE para la docencia de forma que los médicos residentes pudieran poner a prueba sus conocimientos y adquirir la experiencia de los especialistas.
4. Convendría también introducir en la base de conocimiento los datos clínicos que diferenciasen el diagnóstico de Glaucoma en sus diversos tipos.
5. El dato del campo visual se introduce como normal, dudoso, patológico incipiente o patológico avanzado desde el teclado. Lo ideal sería que el mismo SE lo dedujese a partir de las observaciones del médico y de los parámetros que da el campímetro.

3.2 Sistema Experto Glaucom-Easy

En 1992 se comenzó el desarrollo de este sistema experto, patrocinado por el proyecto PETRI (Plan del Estímulo a la Transferencia de Resultados de Investigación) en el Plan Nacional de I+D y la Empresa Bloss (PTR 91-0027), que finalmente quedó registrado (Marca Registrada GLAUCOM-EASY, nº 1761499, BOE 1/7/93).

El objetivo era acometer las líneas de desarrollo que habían quedado abiertas al construir el prototipo anterior. Se abordaron los aspectos 1, 2 y 4 del apartado 3.1.6.

Se pretendió implementar un SE que sirviera como prueba de rastreo para ser utilizada por ópticos, facultativos de atención primaria o incluso oftalmólogos no especializados en glaucoma para el diagnóstico precoz de esta enfermedad.

La parte de ingeniería del conocimiento imprescindible para implementar un SE, se había desarrollado inicialmente con el proyecto anterior, aunque, lógicamente, hubo que completarla para poder adquirir conocimiento y experiencia que no había quedado reflejada en ese sistema.

Hubo sucesivas entrevistas con los expertos, en las que se fueron concretando los grafos de decisión necesarios para realizar el seguimiento del paciente a través de sus consultas oftalmológicas, y posteriormente se fueron refinando. Esos grafos de decisión se muestran en las Figuras 3.10 y 3.11.

Se eligió para desarrollar este sistema un lenguaje convencional. Inicialmente fue C, y más tarde se usó C++ para mejorar la interfaz. Ya se ha indicado que una de sus principales ventajas es la posibilidad de disponer de interfaces convencionales con bases de datos u hojas de cálculo. Pero el principal motivo por el que se usó un lenguaje convencional y no un shell fue el de poder tener una mayor flexibilidad para programar cada uno de los componentes del SE según las necesidades concretas de nuestra aplicación, a pesar del trabajo añadido que esto suponía.

Los valores de entrada a este SE son los factores de riesgo (edad, antecedentes glaucomatosos, miopía elevada, diabetes, enfermedades cardiovasculares), la presión intraocular, la relación excavación/papila, los factores de la papila (ovalización vertical de la excavación, excentricidad, alguna muesca en el borde, cambios bruscos en la trayectoria de los vasos, hemorragias, excavación con bordes netos), y el campo visual. La salida del sistema es un diagnóstico y consejo sobre la siguiente revisión o la conveniencia de instaurar un tratamiento o acudir al especialista en glaucoma (Figura 3.9).

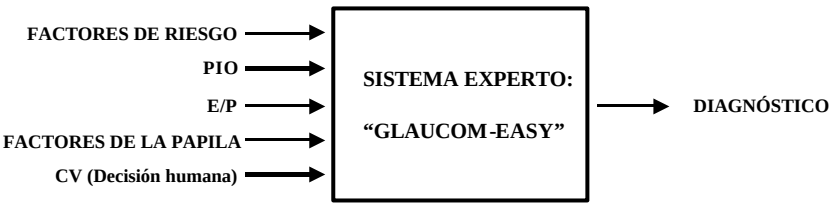


Figura 3.9: Entradas y salidas del SE GLAUCOM-EASY.

A los posibles consejos que proporciona el SE Glaucom-Easy se les ha asignado un número para facilitar el diseño de los grafos de decisión y su programación. Estos son los siguientes:

- 2, 2', 2'': Presión dudosa. Tomarla de nuevo en el plazo de 7 días, preferiblemente por la mañana.
- 3, 3', 3'', 3''': Repetir la tonometría y el campo visual a los 3 meses.
- 4: Revisión al año.
- 4': Revisión a los 6 meses.
- 5: Bajo riesgo. Puede volver al año si lo desea.
- 6: Alto riesgo. Conviene instaurar un tratamiento o enviarlo al especialista en glaucoma.
- 6': Muy alto riesgo. Conviene instaurar un tratamiento o enviarlo al especialista en glaucoma.

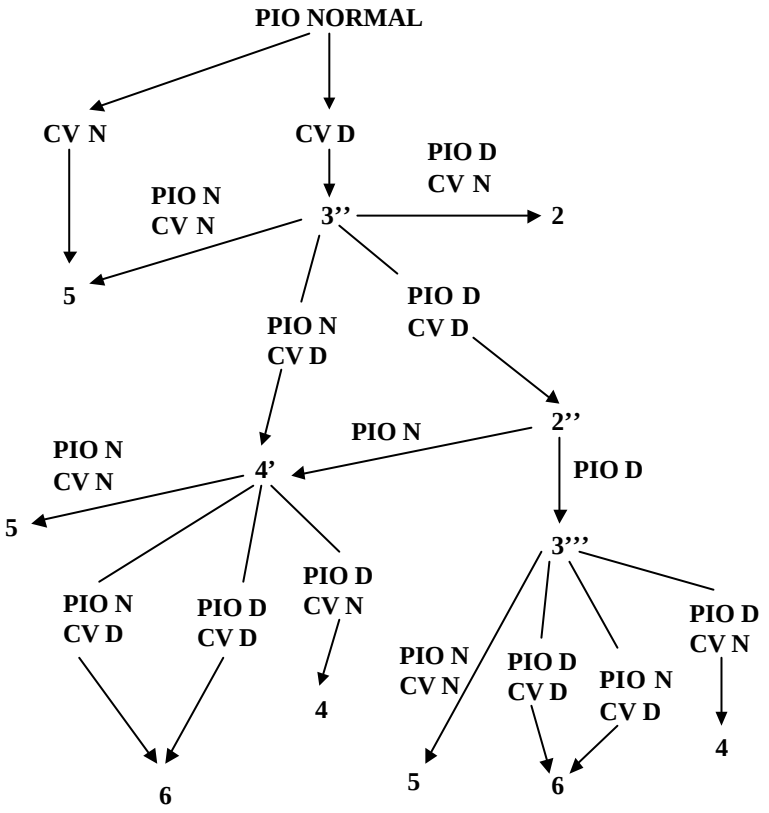


Figura 3.10: Grafo de decisión 1. GLAUCOM-EASY.

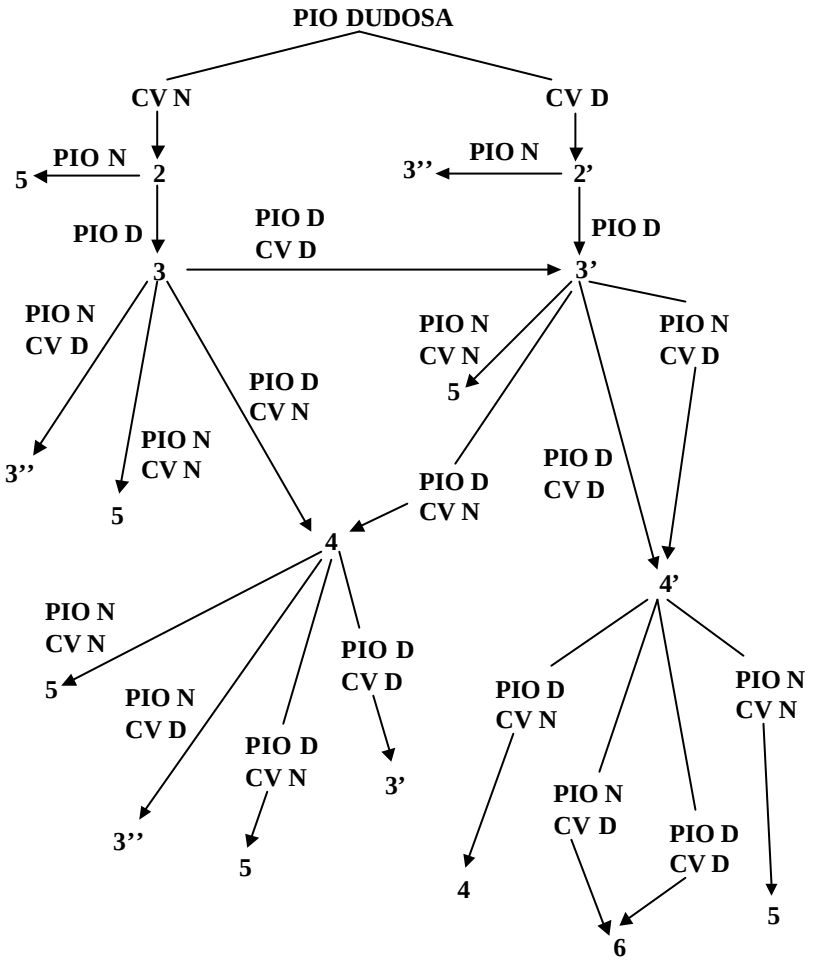


Figura 3.11: Grafo de decisión 2. GLAUCOM-EASY.

3.2.1 Diseño arquitectónico

Se ha utilizado para diseñar el SE una arquitectura del tipo *filtro-tubería*. La aplicación "Glaucom-Easy" consta de los siguientes módulos que interaccionan unos con otros según se puede ver en la Figura 3.12:

- **Sistema Experto.** Tiene la arquitectura básica de un SE. Se ha programado la base de conocimiento y el motor de inferencia y se han omitido algunos componentes opcionales (apartado 3.1.2).
- **Interfaz de usuario.** Es el módulo que permite al SE comunicarse con el usuario.
- **Ayuda.** El sistema proporciona dos tipos de ayuda para guiar al usuario en la consulta, y un tercer tipo que permite clasificar el tipo de glaucoma:
  - Ayuda contextual a petición del usuario.
  - Ayuda que se dispara siempre al llegar a alguna fase concreta de la consulta. El SE toma la iniciativa al considerar esta ayuda imprescindible.
  - Ayuda al diagnóstico diferencial y consejos terapéuticos. Recoge todos los tipos de glaucoma distintos al Crónico Simple que es el que ayuda a diagnosticar el SE. Mediante una serie de preguntas sobre las observaciones y síntomas del paciente, va dirigiendo al usuario hacia la información buscada.
- **Histórico.** Base de datos de los pacientes que almacena sus datos personales y todos los datos recogidos en cada consulta realizada.
- **Gestor de históricos.** Es un gestor de base de datos. Es el módulo que se encarga de gestionar los datos de los pacientes.
- **Validación de datos.** Este módulo se comunica con la interfaz de usuario para validar los datos que introduce el usuario, con el gestor de históricos para comprobar si los datos introducidos pertenecen a algún paciente dado ya de alta en la base de datos de pacientes, y en ese caso, interacciona con el SE para asegurarse de que los datos nuevos no se contradicen con otros de ese paciente ya almacenados.
- **Informes.** Es el módulo que se encarga de proporcionar, a petición del usuario, informes de pacientes o gráficas de su evolución. Se comunica, por tanto, con el gestor de históricos.

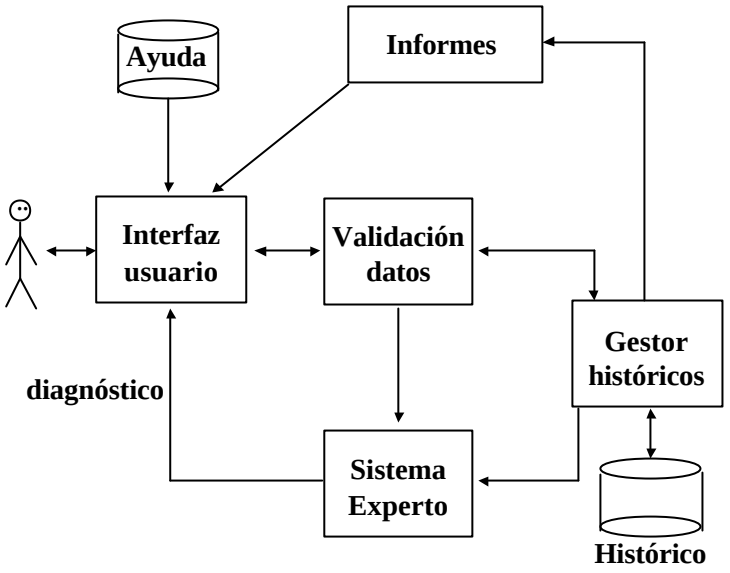


Figura 3.12: Diseño arquitectónico del Sistema Experto Glaucom-Easy.

3.2.2 Proceso de inferencia a lo largo de la base de conocimiento

El proceso que se sigue al realizar una consulta con este SE es el siguiente:

```
salir=no

mientras salir = no hacer

    Pedir el/los apellido/s del paciente
    Buscar en la base de datos de pacientes

    Si no lo encuentra entonces       //nuevo paciente
        Pedir sus datos de filiación
        Pedir los factores de riesgo
    si no
        Modificar los factores de riesgo (opcional)
        Proporcionar informes consultas previas (opcional)
    fin si

    Pedir los parámetros de la consulta
    //PIO, E/P, Factores de la papila (opcional), CV
    Dar un diagnóstico y una pauta de seguimiento del
    paciente.
    Mostrar una gráfica de seguimiento de la PIO y E/P
    (opcional).
    Ofrecer un informe de la consulta actual (opcional).
    Preguntar si salir o comenzar otra consulta.
fin mientras
```

El modelo de ficha que usa este SE es el de la Figura 3.13.

Apellidos y Nombre
Número de historia
Fecha de nacimiento
Fecha de consulta
Antecedentes familiares de glaucoma: ___SI ___NO
Antecedentes personales:
    Miopía ___leve ___moderada ___severa
    Enfermedades cardiovasculares ___SI ___NO
    Diabetes ___SI ___NO
Presión intraocular: OD ___OI ___
Excavación/papila: OD ___OI ___
Campo visual: OD ___OI ___

Figura 3.13: Datos solicitados por el Glaucom-Easy

Es interesante observar que, como es lógico, se piden los datos de los dos ojos a la vez, no como en el prototipo anterior. De esta forma, el modo de realizar la consulta por parte del oftalmólogo u optómetra se asemeja más a la seguida por el SE y, por tanto, se facilita la recogida de datos en tiempo real.

Como ya se ha indicado, todavía no se realiza en este sistema un diagnóstico automático del campo visual, sin embargo se le da al usuario una serie de pautas a seguir para clasificar el campo en normal, dudoso o patológico. Esos criterios vienen indicados en la ayuda contextual correspondiente a la introducción del dato del CV en la ficha que muestra el sistema (Figura 3.14).

La arquitectura que se ha elegido, del tipo *filtro-tubería*, facilita el poder pasar a los distintos módulos del sistema los datos por partes.

Es necesario validar los datos a medida que se van introduciendo antes de solicitar los siguientes. No es suficiente una validación del tipo de dato y rango, sino que es importante comprobar que el dato no se contradice con los que ya se han recogido en consultas anteriores. Por ejemplo, no sería coherente que la relación E/P o la miopía del paciente hubiesen disminuido respecto a la consulta anterior, etc. El SE debe avisar al usuario de un posible error en alguna de las dos consultas.

Para completar el diagnóstico de un paciente, no siempre son necesarios todos los datos que puede recoger la aplicación, por este motivo, el sistema pasa al módulo específico de SE una parte de los datos, y si es necesario, o el usuario lo desea de todas formas, solicita el resto de los datos.

NORMALES:
- Ningún defecto en los 30 grados centrales.
- Hasta 5 puntos aislados con defecto < 10 dB en los 30 grados centrales.
- Hasta 2 puntos de pérdida < 10 dB en zona II.

DUDOSOS:
- Más de 5 puntos aislados con pérdidas < 10 dB en los 30 grados centrales.
- Algún punto aislado con pérdida ≥ 10 dB en zona I.
- 2 puntos juntos < 10 dB alineados horizontalmente en zona I.a.
- Más de 2 puntos con pérdida en zona II.
- Más de 3 puntos juntos < 10 dB en zona III.

PATOLÓGICOS:
- 3 puntos juntos ≥ 3 dB en zona I.
- 2 puntos juntos ≥ 10 dB en zona I.

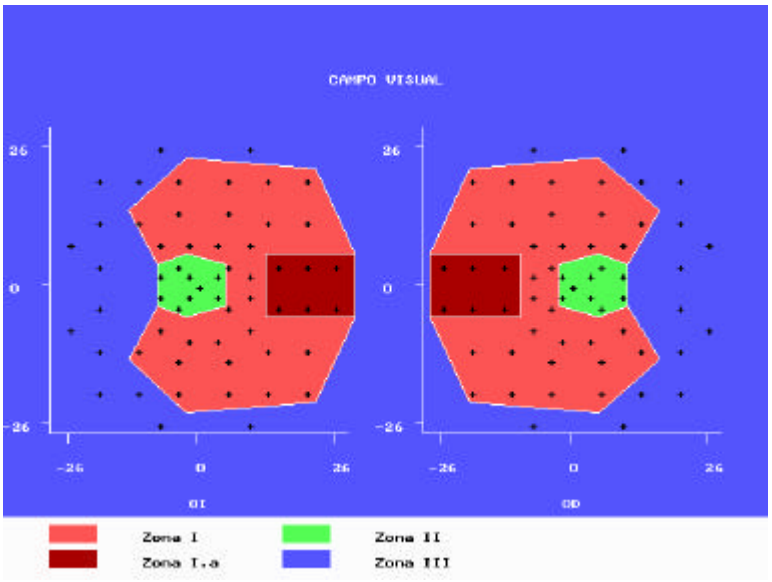


Figura 3.14: Criterios propuestos por el Glaucom-Easy para calificar el campo visual como normal, dudoso o patológico.

El sistema utiliza básicamente dos tipos de reglas para llegar al diagnóstico:

- **Reglas instantáneas**: sólo requieren la información que se introduce cuando se realiza la consulta.
- **Reglas progresivas**: utilizan históricos del paciente para determinar el diagnóstico actual, es decir, realizan un seguimiento del paciente.

Las reglas se recorren en el orden en que se introducen en la base de conocimiento y sólo se activa una de ellas para obtener el diagnóstico. Se han situado antes las reglas instantáneas que disparan un diagnóstico de alto riesgo o muy alto riesgo, y que no necesitan usar datos de consultas anteriores.

Se utiliza un encadenamiento hacia delante, se va de los datos a la meta. Normalmente lo más adecuado en diagnóstico médico es utilizar un método de búsqueda dirigido por la meta puesto que con los mismos síntomas se pueden dar

muchas enfermedades distintas y el número de nodos que se deben explorar crece mucho, sin embargo si se establece una hipótesis y se intenta demostrar o descartar, se debe explorar un número menor de nodos [100]. En nuestro caso no tenemos una gran variedad de enfermedades entre las que elegir, sino solamente una, por lo que es más adecuado partir de los datos y estudiar la posibilidad de que el paciente pueda o no tener la enfermedad.

Alguna de las reglas instantáneas se muestra a continuación en pseudocódigo:

```
si PIO_derecha >= 30
o PIO_izquierda >= 30
entonces
    diagnóstico=Muy alto riesgo. Conviene instaurar un
    tratamiento o enviarlo al especialista en glaucoma
fin si

si campo_visual_derecho patológico
o campo_visual_izquierdo patológico
entonces
    diagnóstico=Muy alto riesgo. Conviene instaurar un
    tratamiento o enviarlo al especialista en glaucoma
fin si
```

Un ejemplo de las reglas progresivas es el siguiente:

```
si (diagnóstico_anterior es “Bajo riesgo. Puede volver al
año si lo desea.”
y PIO_derecha >= 21
y PIO derecha <= 24
y campo_visual_derecho dudoso)
o (diagnóstico_anterior es “Bajo riesgo. Puede volver al
año si lo desea.”
y PIO_izquierda >= 21
y PIO_izquierda <= 24
y campo_visual_izquierdo dudoso)
entonces
    diagnóstico=Presión dudosa. Tomarla de nuevo en un
    plazo de 7 días, preferiblemente por la mañana.
fin si
```

3.2.3 Asignación de factores de certeza

La asignación del factor de certeza final del diagnóstico que ofrece el Glaucom-Easy se efectúa en un proceso paso a paso, dependiendo de los datos recogidos en las exploraciones realizadas al paciente.

Los factores de riesgo de padecer glaucoma no tienen un peso importante sobre el diagnóstico final, por esta razón, en caso de tener miopía media, enfermedades cardiovasculares, diabetes, o ser mayor de 40 años, se incrementará en 2 el factor de certeza total, y si tiene antecedentes familiares de glaucoma, al ser el factor de riesgo más importante, se incrementará en 5.

Eso mismo ocurre con los factores de la papila. Si la excavación tiene una ovalización vertical, es excéntrica, presenta su borde alguna muesca, hay cambios bruscos en la trayectoria de los vasos, o hemorragias dentro de la papila, se incrementa en 2 el factor de certeza total. Por el contrario, si lo que se observa es una excavación con bordes netos, disminuye en 5, ya que indica que posiblemente no se hayan empezado a atrofiar las fibras del nervio óptico.

El factor de certeza total se incrementa según la presión intraocular medida al paciente como indica la Figura 3.15.

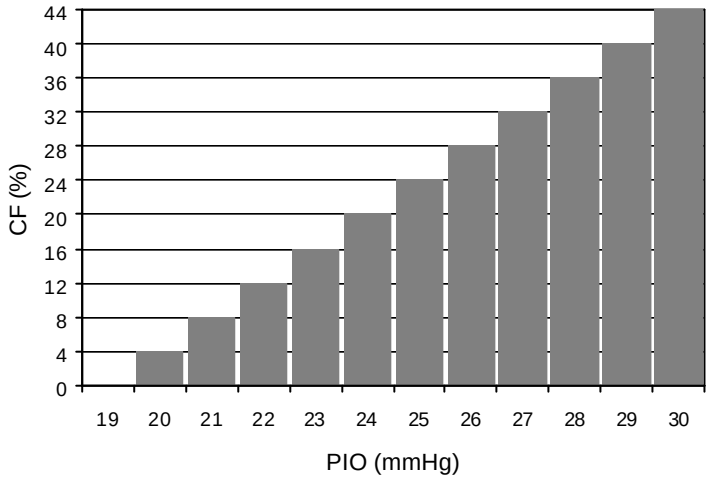


Figura 3.15: Factor de certeza según la PIO. Glaucom-Easy.

Si la PIO es menor de 20 mmHg no se incrementa el factor de certeza, pero si es 20 o mayor aumenta linealmente según la función de la línea siguiente:

El factor de certeza que aporta el valor de la relación Excavación/Papila al factor de certeza total se puede observar en la Figura 3.16. Si es menor de 0.3 no incrementa el valor de CF total y si es 0.3 o mayor se incrementa como indica esa gráfica.

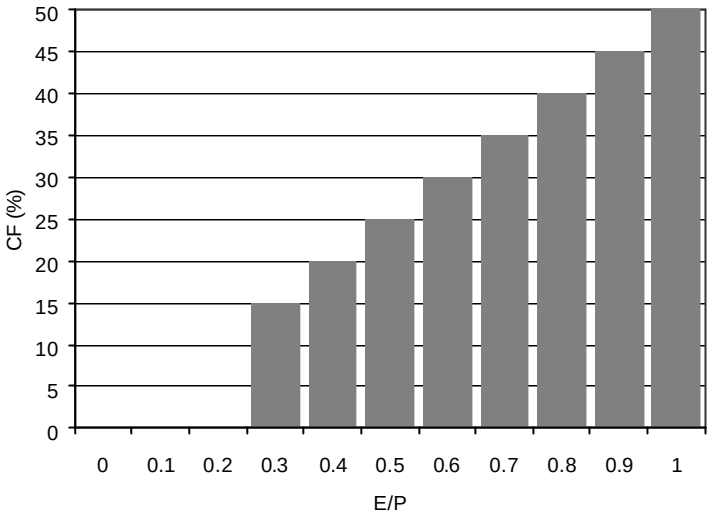


Figura 3.16: Factor de certeza según la E/P. Glaucom-Easy.

Un campo visual dudoso incrementa en 20 el factor de certeza total y un campo visual patológico en 40.

Se calcula el factor de certeza total de cada ojo sumando los factores de certeza parciales que se han ido comentando anteriormente.

Como solamente se va a dar un factor de certeza por diagnóstico de paciente aunque se hayan examinado los dos ojos, se tomará el mayor que se haya obtenido bien con los datos del ojo derecho o con los del izquierdo.

Si la asimetría entre ambos ojos de un paciente es mayor o igual a 0.2 se añade 20 al factor de certeza.

Si la diferencia entre las presiones intraoculares de ambos ojos es mayor de 3 se incrementa en 10 el factor total.

Después de todas estas operaciones e incrementos, podría pasar que obtuviésemos un factor negativo o mayor de 100, ninguno de los dos casos sería un valor aceptable. Por ello, si es negativo se sustituye por 0 y si es mayor o igual a 100 se sustituye por 99. Se considera que un riesgo de glaucoma con un CF 100% no conviene darlo en este sistema experto concreto.

3.2.4 Validación del Sistema Experto

En 1995 se validó el SE Glaucom-Easy [7]. Se incluyeron en el estudio 100 pacientes con edad igual o superior a 40 años, en los que fuera posible explorar la papila del nervio óptico y que pudieran realizar un campo visual con un perímetro computerizado. Fueron distribuidos en dos grupos según su estado clínico:

- **Grupo 1**: Pacientes diagnosticados de glaucoma (55).
- **Grupo 2**: Pacientes sin patología oftalmológica o con patología distinta del glaucoma (45).

En todos los casos se realizó una completa anamnesis (antecedentes familiares de glaucoma, antecedentes personales de miopía, enfermedad cardiovascular o diabetes) así como una exploración oftalmológica (agudeza visual, refracción subjetiva, presión intraocular, nervio óptico, campo visual).

Se introdujeron los datos de las historias clínicas en la base de datos del Glaucom-Easy. La presión intraocular debe expresarse en mmHg, la excavación papilar en décimas, y el campo visual debe ser calificado por el explorador como normal, dudoso o patológico, según los criterios que expresa el propio programa y que pueden observarse en la Figura 3.14.

Se registraron los dos tipos de resultados emitidos por el Glaucom-Easy:

- Este sistema, por un lado, expresa el riesgo que la persona explorada tiene de padecer glaucoma calificándolo como bajo o alto o identificándolo con un porcentaje entre 0 y 100.
- Por otro lado, aconseja al usuario del programa sobre la pauta a seguir: periodo de revisión, repetición de alguna exploración, necesidad de tratamiento o envío al especialista en glaucoma.

Finalmente se calcularon la sensibilidad, especificidad, falsos positivos, falsos negativos y valor predictivo del Glaucom-Easy en su configuración actual, así como los valores que podrían obtenerse si se modificara el porcentaje de riesgo de padecer glaucoma que se considera como límite de la normalidad. Para calcular estos índices se consideraron normales aquellos pacientes para los que el SE recomienda una revisión a los 12 meses (bajo riesgo o riesgo de padecer glaucoma inferior o igual a 35%), y patológicos aquellos que según el programa deben ser tratados o enviados al especialista (alto riesgo o riesgo superior o igual a 46%).

3.2.4.1 Resultados

La distribución de los pacientes según su diagnóstico previo y el resultado emitido por el Glaucom-Easy puede observarse en la tabla Tabla 3.1. En su configuración actual, este programa califica a los pacientes según expresa la Tabla 3.2 y tiene una sensibilidad del 96%, una especificidad del 77%, un valor predictivo positivo del 84% y un valor predictivo negativo del 94%.

	Bajo riesgo	**0-10%**	**11-20%**	**21-30%**	**31-40%**	**41-50%**	**51-60%**	**61-70%**	**71-80%**	**Alto riesgo**	**Total**
Grupo 1	2					1	2	1	1	48	55
Grupo 2	24	2	1	5	3	4	3			3	45

Tabla 3.1: Distribución de los pacientes según el diagnóstico previo y el riesgo de padecer glaucoma calculado por el Glaucom-Easy. En el Grupo 1 se incluyeron los pacientes diagnosticados de glaucoma y en el Grupo 2 los sanos o con distinta patología.

	Glaucom-Easy patológico	**Glaucom-Easy normal**	**Total**
Glaucoma	53	2	55
No glaucoma	10	35	45
Total	63	37	100

Tabla 3.2: Distribución de los pacientes según su diagnóstico previo y el resultado obtenido con el Glaucom-Easy.

Como es lógico, la sensibilidad y especificidad del sistema se modifican según donde establezcamos el límite entre la normalidad y la patología, y así lo expresa la Figura 3.17. Los valores de sensibilidad, especificidad, resultados falsos y valor predictivo que se podrían obtener tras la modificación del valor del riesgo que el Glaucom-Easy considera indicación de tratamiento o visita al especialista, se expresan en la tabla Tabla 3.3.

Límite de normalidad	**Sensibilidad**	**Especificidad**	**Falsos positivos**	**Falsos negativos**	**Valor predictivo positivo**	**Valor predictivo negativo**
Bajo riesgo	96	57	43	4	73	92
10%	96	60	40	4	74	93
15%	96	60	40	4	74	93
20%	96	60	40	4	74	93
25%	96	68	32	4	79	93
30%	96	71	29	4	80	94
35%	96	77	23	4	84	94
40%	96	77	23	4	84	94
45%	96	77	23	4	84	94
50%	94	86	14	6	89	92
55%	90	91	9	10	92	89
60%	90	93	7	10	94	89
65%	89	93	7	11	94	87
70%	89	93	7	11	94	87
75%	89	93	7	11	94	85
80%	89	93	7	11	94	85
85%	89	93	7	11	94	85
90%	89	93	7	11	94	85
Alto riesgo	89	93	7	11	94	85

Tabla 3.3: Valores de sensibilidad, especificidad, falsos positivos, falsos negativos, valor predictivo positivo y valor predictivo negativo, según el valor del riesgo que el Glaucom-Easy considera indicación de tratamiento o visita al especialista.

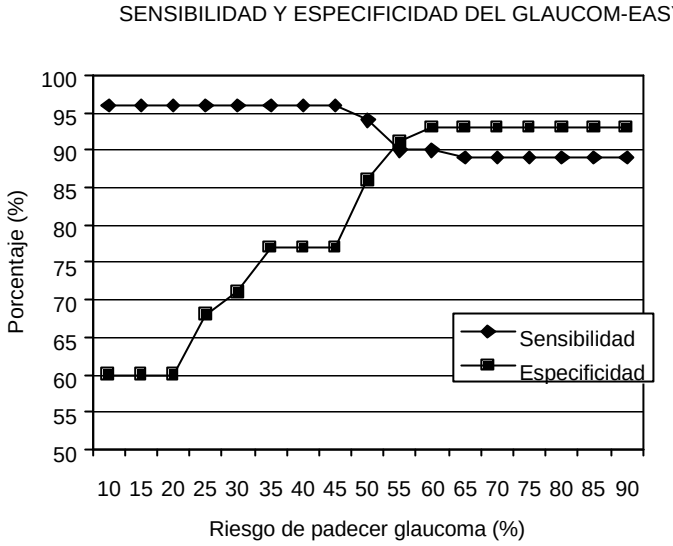


Figura 3.17: Variación de la sensibilidad y especificidad si se modificara el riesgo de glaucoma a partir del cual el Glaucom-Easy indica la necesidad de tratamiento o visita al especialista.

Si el límite de normalidad fuera situado en un riesgo de padecer glaucoma del 60% (según lo calcula el programa) este SE tendría una sensibilidad del 90%, una especificidad del 93%, un valor predictivo positivo del 94% y un valor predictivo negativo del 89%.

3.2.4.2 Conclusiones

Glaucom-Easy recomienda tratamiento o visita al especialista, es decir, considera altamente sospechosos de sufrir glaucoma a todos aquellos pacientes con un riesgo de padecer glaucoma igual o superior al 45%, y aconseja revisión a los 12 meses para aquéllos con un riesgo inferior o igual al 33%. Con estos criterios este sistema tiene una alta sensibilidad (96%) pero una especificidad algo inferior al 80% (77%), valor que se considera recomendable para una buena prueba de rastreo. Como consecuencia existe un porcentaje de resultados falsos positivos algo elevado (23%). Sin embargo, con esta configuración se obtiene el valor predictivo negativo más alto que el sistema es capaz de obtener (Tabla 3.3), cualidad muy valorada en una prueba de rastreo que debe ser utilizada por ópticos, facultativos de atención primaria o incluso oftalmólogos no especializados en glaucoma. Podemos concluir entonces que el Glaucom-Easy puede ser particularmente útil para el rastreo del glaucoma y el diagnóstico en fases iniciales de la enfermedad.

3.2.5 Perspectivas de desarrollo

Ya se ha indicado que el Glaucom-Easy mejoró el prototipo inicial desarrollado mediante M1 en tres aspectos fundamentales, como son la interfaz, la

ayuda al diagnóstico diferencial de los demás tipos de glaucoma que no son el Glaucoma Crónico Simple, y en el seguimiento del paciente a través de sus diversas consultas.

Sin embargo todavía hay un aspecto clave que no se ha abordado, y es el diagnóstico automático del campo visual. El dato del campo visual se introduce como normal, dudoso, patológico incipiente o patológico avanzado desde el teclado. Al igual que se indicó en las mejoras del prototipo M1, lo más adecuado sería que el mismo SE lo dedujese a partir de las observaciones del médico y de los parámetros que da el campímetro, para que de verdad fuera útil para el optómetra u oftalmólogo general que no son especialistas en esta enfermedad.

Además, parece conveniente aplicar técnicas de lógica borrosa para mejorar el diagnóstico del sistema. El conocimiento que manejan los expertos en el problema del diagnóstico del glaucoma es bastante impreciso, como se ha podido ver en el planteamiento del problema (por ejemplo, un mismo valor de presión intraocular puede ser normal para un paciente y alto para otro, y lo mismo con el resto de parámetros). Un sistema borroso, que maneje todo este conocimiento impreciso, tiene un modo de razonamiento que se puede asemejar más al del experto.

4 Análisis del campo visual

Con este estudio nos proponemos varias metas:

1. Obtener un buen clasificador de campos visuales que diferencie entre glaucoma y no glaucoma, para luego integrarlo al sistema de diagnóstico, en nuestro caso borroso, que realiza el diagnóstico global del glaucoma teniendo en cuenta también el resto de los parámetros.

2. Conseguir un clasificador que nos permita distinguir las cinco clases contenidas en los datos, de forma que pueda aportar información objetiva al experto sobre las diferencias entre el daño incipiente producido en el campo visual por distintas patologías.

   Como ya se analizó en el capítulo 2, este estudio fue abordado por primera vez por Antón [6] en 1995 en su Tesis Doctoral, y después de algún otro trabajo en esta línea, se ha querido ahora completar.

4.1 Los datos

Los datos usados en este trabajo son los mismos que usó Antón en su Tesis Doctoral [6], por ello los criterios de selección de los mismos han sido obtenidos de esa investigación.

4.1.1 Estudio piloto

Su objetivo fue calcular el tamaño de la muestra y poner a punto tanto el protocolo como el método informático-estadístico de análisis de los resultados. Para este estudio se utilizaron dos grupos de pacientes:

- **Grupo 1**.- Se eligieron pacientes glaucomatosos (40 ojos) seleccionados de forma retrospectiva a partir del archivo de historias clínicas de la unidad de glaucoma del Hospital Clínico Universitario de Valladolid (H.C.U.) y de acuerdo con los criterios que se especifican más adelante (apartado 4.1.2.1).

- **Grupo 2**- Cuarenta pacientes sin glaucoma (40 ojos) seleccionados de forma prospectiva entre los que acudieron a las consultas del H.C.U. y de acuerdo con los criterios que se definen en el siguiente apartado (4.1.2.1):
  - 10 Sujetos normales
  - 10 Cataratas seniles
  - 10 Retinopatías diabéticas no proliferativas
  - 10 Retinopatías hipertensivas

  A todos ellos se les pidió consentimiento verbal para realizar la exploración tras explicarles la naturaleza del estudio.

En todos los casos se obtuvieron los datos de la historia clínica (grupo 1) o se realizó (grupo 2) una anamnesis, exploración oftalmológica completa (agudeza visual, refracción, biomicroscopía anterior y posterior, PIO) y al menos dos exploraciones con el programa G-1 del perímetro Octopus 500 (Interzeag AG, Suiza).

Se diseñó una base de datos en MS Access (Microsoft) y se almacenaron todos los datos de estos pacientes.

4.1.2 Estudio definitivo

Se estudió el campo visual de 180 ojos (incluidos los 80 casos del estudio piloto), 96 de ellos diagnosticados de glaucoma (grupo 1) y 84 ojos sin glaucoma (grupo 2). Se incluyó un solo ojo de cada paciente. En caso de que los dos ojos cumplieran los criterios de inclusión, se seleccionó uno de ellos de forma aleatoria.

4.1.2.1 Selección de pacientes

Grupo 1: Glaucoma

Se seleccionaron de forma retrospectiva entre 1491 historias del archivo de la unidad de glaucoma del Hospital Clínico Universitario de Valladolid y de acuerdo con los siguientes criterios, que también se emplearon para seleccionar los glaucomas de la muestra de prueba (grupo 3).

A. Criterios de inclusión:

- Mayores de 40 años.
- Ojos fáquicos.
- Ojos diagnosticados de glaucoma crónico de ángulo abierto con presión elevada y lesiones campimétricas incipientes. Previamente se definió como tal todo paciente que presentara:
  1. PIO > 21 mmHg medida con el tonómetro de Perkins y con el paciente sentado. Se realizaron un mínimo de tres mediciones y se calculó el valor medio.
  2. Ángulo irido-corneal abierto.

3. Valor del índice excavación/papila (E/P) o signos asociados que hicieran sospechar el diagnóstico de glaucoma, es decir que presentara alguna de las siguientes características:
   - E/P > 0,5.
   - Asimetría en la E/P de ambos ojos mayor del 10% del diámetro papilar.
   - Cualquier valor de E/P con adelgazamientos focales del anillo neuro-retiniano, ovalización vertical de la excavación, desplazamiento nasal de los vasos o hemorragias en el nervio óptico.
4. Campo visual con lesiones incipientes, fiable y compatible con el diagnóstico de glaucoma. Se definió como tal todo CV que cumpliera las siguientes condiciones:
   - Falsos positivos y negativos < 15%.
   - Congruencia con CV posteriores. Al menos el 30% de los puntos con defecto debían estar también presentes, con un valor en dB igual o superior (defecto igual o más profundo), en el siguiente CV realizado por el paciente.
   - Lesiones incipientes con la certeza de su origen glaucomatoso. Se incluyeron ojos en los que la evolución de los CV hasta estadios intermedios y avanzados no ofrecía ninguna duda sobre el diagnóstico de glaucoma. De ellos se eligieron CV en estadios incipientes (al menos 1 defecto de 5 dB en los 26° centrales) con MD no superior a 6 dB y un número de defectos mayores de 4 dB inferior a 10.

B. Criterios de exclusión:

- Hipertensos oculares. Sujetos con presiones intraoculares superiores a 22 mmHg que no presentaban lesiones campimétricas o cambios en la cabeza del nervio óptico.
- Glaucomas de "tensión normal". Se definió como tal a todo paciente que presentara valores de presión intraocular no superiores a 21 mmHg (comprobado mediante curva tensional), ángulo abierto, cambios glaucomatosos en la cabeza del nervio óptico (adelgazamiento o pérdida del anillo neural, discrepancia palidez/excavación) y CV patológico. Se excluyeron estos casos por existir estudios que indican que sus alteraciones del campo visual son distintas de las presentes en glaucomas con presión elevada.
- Patología coexistente que pudiera alterar el campo visual (patología neurológica, retinopatía diabética o hipertensiva, miopía con atrofia coriorretiniana,...).
- Catarata que no permitiera observar un fulgor pupilar uniforme con el oftalmoscopio directo (+10 dioptrías y a 20 cm del ojo) o que presentara opacidades visibles a través de una pupila de unos 4 mm.
- Afáquicos y pseudofáquicos.

Grupo 2.- Ojos sin glaucoma

Estuvo integrado por dos subgrupos, el subgrupo 2a formado por sujetos normales y el subgrupo 2b de pacientes con patología no glaucomatosa. A todos

ellos se les pidió consentimiento verbal para realizar la exploración tras explicarles la naturaleza del estudio. Los mismos criterios que se describen a continuación fueron empleados para seleccionar los sujetos sin glaucoma de la muestra de prueba (grupo 4).

A. *Criterios de inclusión:*

- Edad, mayores de 40 años.
- Buena colaboración para una exploración oftalmológica completa y para la realización del CV.
- Para formar el subgrupo 2a se eligieron sujetos sin patología oftalmológica. Los pacientes del subgrupo 2b debían cumplir una sola de las siguientes situaciones:
  2. Catarata de cualquier tipo y grado de intensidad con agudeza visual > 0,5.
  3. Retinopatía hipertensiva de grado I o II según la clasificación de Keith-Wagener-Barker.
  4. Retinopatía diabética inicial. Microaneurismas, microhemorragias o exudados duros aislados que no ocuparan más de 1/3 de la superficie de la mácula o hasta tres cuadrantes de la retina.

B. *Criterios de exclusión:*

- Glaucoma crónico diagnosticado previamente o en el momento de la exploración.
- Hipertensos oculares.
- Glaucomas de "tensión normal".
- Catarata con agudeza visual inferior a 0,5.
- Retinopatía hipertensiva grado III o IV de la clasificación de Keith-Wagener-Barker.
- Retinopatía diabética preproliferante o proliferante.
- Ojos en los que coexistieran dos o más patologías consideradas como criterio de inclusión.
- Cualquier otra patología ocular o neurológica que pudiera alterar el campo visual y no se encontrara entre los criterios de inclusión.
- Afáquicos o pseudo-fáquicos.

4.1.2.2 Obtención de datos clínicos

Muestra de entrenamiento

En todos los casos se realizó o se había realizado previamente una anamnesis, exploración oftalmológica completa (agudeza visual, refracción, biomicroscopía anterior y posterior, PIO) y al menos dos exploraciones con el programa G-1 del perímetro Octopus 500. Las dos exploraciones del CV se realizaron en días distintos.

Para la selección de los casos incluidos en el grupo 1 se revisaron, como ya se ha dicho, de forma retrospectiva las historias clínicas de pacientes de la unidad de glaucoma del Hospital Clínico Universitario de Valladolid.

Los ojos incluidos en el grupo 2 fueron seleccionados de forma prospectiva. Se intentó estudiar a todos los pacientes que cumplieran los criterios de inclusión y exclusión entre los que acudieron a las consultas de oftalmología del Hospital Clínico Universitario y se incluyeron sólo aquellos que, tras recibir información sobre la naturaleza del estudio, dieron su consentimiento.

De las 1491 historias clínicas revisadas sólo 116 cumplieron los criterios establecidos para su inclusión en el grupo 1. Las primeras 96 constituyeron el grupo de glaucomas de la muestra de entrenamiento. El grupo 2 se formó con 84 sujetos sin glaucoma. De éstos, 28 no presentaban ninguna patología oftalmológica (subgrupo 2a). Los otros 56 pacientes del grupo 2 integran el subgrupo 2b: 27 ojos con catarata, 14 con retinopatía diabética y 15 diagnosticados de retinopatía hipertensiva (HTA). (Tabla 4.1).

Muestra de prueba

Con la finalidad de evaluar la precisión diagnóstica de los métodos de interpretación se seleccionó una muestra de prueba independiente. Se siguió el mismo procedimiento que con la muestra de entrenamiento.

El conjunto de muestra de prueba estuvo formado por 48 ojos, 20 diagnosticados de glaucoma (grupo 3) y 28 sin glaucoma (grupo 4). De éstos últimos, 8 sujetos no tenían patología oftalmológica (subgrupo 4a), 7 sufrían catarata, 7 tenían retinopatía diabética y 6 presentaban retinopatía hipertensiva (subgrupo 4b). (Tabla 4.1).

	Muestra de entrenamiento	**Muestra de prueba**
Glaucoma	96	20
Normal	28	8
Cataratas	27	7
Diabetes	14	7
HTA	15	6
TOTAL	180	48

Tabla 4.1: Distribución inicial de casos según clases.

Al estudiar detenidamente cada campo visual con su diagnóstico correspondiente, se pudo observar que campos totalmente idénticos tenían asignado distinto diagnóstico. Eso era así porque esos pacientes con el tiempo han degenerado en distintas patologías o han permanecido normales, pero inicialmente el campo era completamente normal. Se decidió cambiar el diagnóstico de los campos a normal siempre y cuando no tuvieran alterado ningún punto. De otra

manera eso despistaría, lógicamente, a cualquier clasificador, y estaríamos subestimando los resultados.

Finalmente la distribución de los casos de campos visuales según las distintas clases es la mostrada en la Tabla 4.2:

	Muestra de entrenamiento	**Muestra de prueba**
Glaucoma	96	20
Normal	37	9
Cataratas	22	7
Diabetes	12	7
HTA	13	5
TOTAL	180	48

Tabla 4.2: Distribución final de casos según clases.

4.2 Análisis de componentes principales

Antes de aplicar distintas técnicas de clasificación a los conjuntos de ejemplos de que disponemos, conviene señalar un punto muy importante que va a condicionar todos los experimentos de este trabajo: **el número de patrones de aprendizaje disponibles.** Claramente nos encontramos con el problema de no disponer de suficiente número de patrones para entrenar una determinada arquitectura. El número de datos necesarios en general crece exponencialmente con la dimensión del espacio de entrada. A este problema se le llama *la maldición de la dimensión* (*curse of dimensionality*) [111]. En nuestro caso, el espacio de entrada puede tener desde 59 a 61 variables (los 59 puntos de pérdida de sensibilidad, el defecto medio y la varianza de la pérdida) y no disponemos más que de 180 casos de entrenamiento y 48 de prueba.

Cuanto mayor sea una red neuronal, dispondrá de más grados de libertad (pesos) y se precisará también un mayor número de patrones de aprendizaje para su correcto ajuste. Una solución para reducir el número de parámetros de una arquitectura consiste en la reducción de la dimensión de los vectores de entrada, ya que dispondremos del mismo número de ejemplos para un número menor de pesos. Un modo de conseguir esto es aplicando análisis de componentes principales a las variables de entrada.

El análisis de componentes principales permite analizar una tabla de k observaciones con n variables con el fin de reducir las dimensiones en cuanto al número de variables a estudiar. Esta nueva representación ha de conservar la máxima información sobre los datos originales. Se trata, por tanto, de evitar redundancia prescindiendo de aquellas variables que sean innecesarias para el conocimiento del fenómeno, bien porque la información que aportan ya esté contenida en la información de otras variables, bien porque su información sea independiente del fenómeno en estudio [120, 135].

En nuestro trabajo tenemos una serie de pruebas perimétricas de pacientes, y cada una de esas pruebas está formada por una serie de valores que corresponden a los datos de campos visuales recogidos (apartado 4.1). Las variables cuantitativas de este estudio son, por tanto, los datos de pérdida de sensibilidad en la retina que proporciona el perímetro.

A cada punto del campo visual explorado por la fase 1 del programa G-1 del Octopus 500 (59 puntos en los 26° centrales) se le asignó un número correlativo empezando a contar por el punto situado más arriba y a la izquierda en el campo del ojo derecho. Los puntos del campo del ojo izquierdo fueron numerados de forma simétrica, con la finalidad de que los puntos con la misma posición relativa tuvieran asignado el mismo número en los dos ojos (Figura 4.1). De esta forma los datos de ambos ojos fueron analizados de forma conjunta.

Hemos utilizado el análisis de componentes principales sobre esta matriz de datos con el fin de obtener un número pequeño de combinaciones lineales de las 59 variables que reúnan la mayor parte de la variabilidad que hay en los datos.

Se usó el paquete estadístico STATGRAPHICS que ofrece dos formas de realizar este análisis:

1. Estandarizando los valores de las variables a media cero y varianza uno.
2. Sin estandarizar los datos.

Los resultados obtenidos fueron los siguientes:

En el **caso 1**:

Número de componente	Autovalor	Porcentaje de Varianza	Porcentaje acumulado
1	22.0592	37.389	37.389
2	3.04284	5.157	42.546
3	2.65444	4.499	47.045
4	2.29699	3.893	50.938
5	1.87156	3.172	54.11
6	1.82957	3.101	57.211
7	1.55506	2.636	59.847
8	1.48351	2.514	62.361
9	1.29095	2.188	64.549
10	1.20268	2.038	66.588
11	1.07303	1.819	68.407
12	**1.02517**	**1.738**	**70.144**
13	0.997118	1.69	71.834
14	0.92636	1.57	73.404
15	0.895816	1.518	74.923
16	0.864544	1.465	76.388
17	0.836012	1.417	77.805
18	0.806622	1.367	79.172
19	0.723853	1.227	80.399
20	0.679461	1.152	81.551
21	0.671984	1.139	82.69
22	0.646222	1.095	83.785
23	0.643548	1.091	84.876
24	0.598983	1.015	85.891
25	0.557814	0.945	86.836
26	0.487977	0.827	87.663
27	0.45738	0.775	88.439
28	0.426095	0.722	89.161
29	0.414468	0.702	89.863
30	0.398878	0.676	90.539
31	0.375876	0.637	91.176
32	0.366049	0.62	91.797
33	0.338872	0.574	92.371
34	0.327523	0.555	92.926
35	0.314817	0.534	93.46
36	0.306605	0.52	93.98
37	0.299673	0.508	94.487
38	0.275024	0.466	94.954
39	**0.258469**	**0.438**	**95.392**
40	0.250919	0.425	95.817
41	0.235622	0.399	96.216
42	0.21806	0.37	96.586
43	0.208538	0.353	96.939

44	0.190752	0.323	97.263	53	0.098024	0.166	99.327
45	0.181027	0.307	97.57	54	0.0814235	0.138	99.465
46	0.168451	0.286	97.855	55	0.0803698	0.136	99.601
47	0.159069	0.27	98.125	56	0.074827	0.127	99.728
48	0.146733	0.249	98.373	57	0.0635262	0.108	99.835
49	0.131373	0.223	98.596	58	0.0498731	0.085	99.92
50	0.12168	0.206	98.802	59	0.0471902	0.08	100
51	0.108703	0.184	98.986				
52	0.102749	0.174	99.161				

El principal objetivo del análisis de componentes principales es la determinación de los ejes principales de inercia de la nube de puntos alrededor de su centro de gravedad. Estos ejes marcan las direcciones predominantes a lo largo de las que se sitúan la mayoría de las observaciones, por lo que permitirán conseguir la mejor descripción de la variabilidad de los datos u observaciones con el menor número de variables.

El problema consiste en obtener los autovectores y autovalores de la matriz de covarianzas de la nube de datos. Los autovectores determinan las direcciones de máxima dispersión, es decir, variables o combinaciones de variables que aportan la máxima información. Los autovalores son una medida de la información contenida en la dirección indicada por el correspondiente autovector.

Como se puede observar en la tabla anterior, los autovalores están ordenados de mayor a menor.

En este caso, el programa estadístico ha extraído 12 componentes ya que sus autovalores son mayores o iguales a 1.0. Esas 12 componentes dan cuenta del 70.14% de la variabilidad de los datos originales. Si quisiéramos recoger alrededor de un 95% de la variabilidad deberíamos quedarnos con 39 componentes.

En el **caso 2**:

Número de componente	Autovalor	Porcentaje de Varianza	Porcentaje acumulado				
1	119.602	34.142	34.142	16	5.1133	1.46	78.824
2	25.7414	7.348	41.49	17	4.54775	1.298	80.123
3	18.7588	5.355	46.845	18	4.49622	1.284	81.406
4	16.3553	4.669	51.514	19	4.09986	1.17	82.577
5	12.7264	3.633	55.147	20	3.98891	1.139	83.715
6	11.2889	3.223	58.37	21	3.72795	1.064	84.779
7	10.3383	2.951	61.321	22	3.61049	1.031	85.81
8	9.76433	2.787	64.108	23	3.39704	0.97	86.78
9	8.36404	2.388	66.496	24	3.30396	0.943	87.723
10	7.83101	2.235	68.731	25	3.02599	0.864	88.587
11	6.89919	1.969	70.701	26	2.86995	0.819	89.406
12	6.56843	1.875	72.576	27	2.71502	0.775	90.181
13	5.95449	1.7	74.276	28	2.45295	0.7	90.881
14	5.56412	1.588	75.864	29	2.30234	0.657	91.539
15	5.25737	1.501	77.365	30	2.22951	0.636	92.175
				31	2.07906	0.593	92.768
				32	1.9451	0.555	93.324

33	1.85282	0.529	93.853	48	0.655565	0.187	98.848
34	1.77693	0.507	94.36	49	0.578266	0.165	99.013
35	1.61536	0.461	94.821	50	0.551586	0.157	99.17
36	1.56127	0.446	95.267	51	0.4868	0.139	99.309
37	1.42489	0.407	95.673	52	0.449582	0.128	99.437
38	1.38679	0.396	96.069	53	0.401492	0.115	99.552
39	1.33308	0.381	96.45	54	0.359671	0.103	99.655
40	1.23121	0.351	96.801	55	0.339172	0.097	99.752
41	1.17122	0.334	97.136	56	0.263257	0.075	99.827
42	**1.12652**	**0.322**	**97.457**	57	0.239412	0.068	99.895
43	0.981125	0.28	97.737	58	0.208633	0.06	99.955
44	0.884433	0.252	97.99	59	0.159121	0.045	100
45	0.847085	0.242	98.232				
46	0.77156	0.22	98.452				
47	0.730642	0.209	98.66				

En este caso, el programa estadístico ha extraído 42 componentes ya que sus autovalores son mayores o iguales a 1.0. Esas 42 componentes dan cuenta del 97.5% de la variabilidad de los datos originales.

Como conclusión de este estudio de análisis de componentes principales, podemos decir que para recoger una variabilidad de los datos originales adecuada, debemos seleccionar un número elevado de componentes. Por ello, la reducción de dimensionalidad que se podría lograr utilizando esta herramienta estadística no es suficiente para suplir la falta de casos para el entrenamiento. Además tiene como inconveniente el que las variables iniciales que llevan consigo un significado, se transforman en el nuevo sistema coordenado en una combinación lineal de todas las variables disponibles, y por tanto es más difícil interpretar los resultados que se obtengan utilizando esas nuevas variables, en caso de que sean interpretables. Debido a todo esto no se aplicarán los resultados obtenidos sino que se intentará reducir la dimensionalidad aplicando otro método.

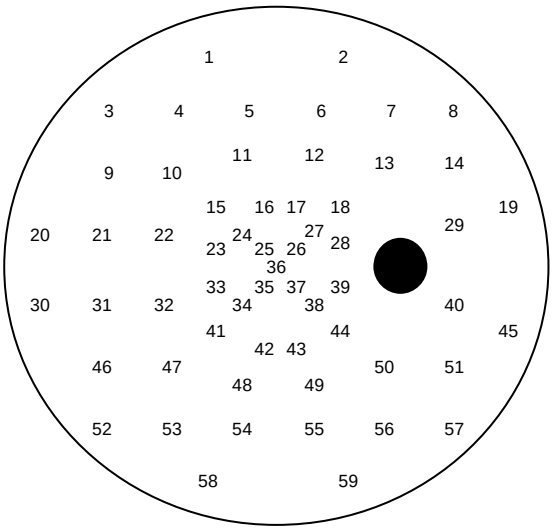


Figura 4.1: Número asignado a cada punto del CV.

4.3 Distribución del campo en zonas

Con el objetivo de reducir el número de variables que describieran las lesiones presentes, se dividió el CV en 7 zonas teniendo en cuenta la distribución de las fibras nerviosas en la retina, la experiencia clínica de la Unidad de Glaucoma del Hospital Clínico Universitario de Valladolid y las ya empleadas en el sistema experto Glaucom-Easy [6, 7, 142]. La Figura 4.2 representa las 7 zonas. Tanto para el entrenamiento de redes neuronales como para razonamiento basado en casos se utilizaron los valores de defecto de sensibilidad en los 59 puntos del CV y los datos referidos a las 7 zonas descritas. Cada zona estuvo representada por la media de los defectos (7 datos) o bien por la media y la desviación estándar (14 datos).

El uso de estas 7 zonas, elegidas fundamentalmente debido a la experiencia de los expertos, será justificable finalmente si se observa que los resultados obtenidos son mejores que utilizando los 59 defectos de sensibilidad obtenidos directamente del perímetro.

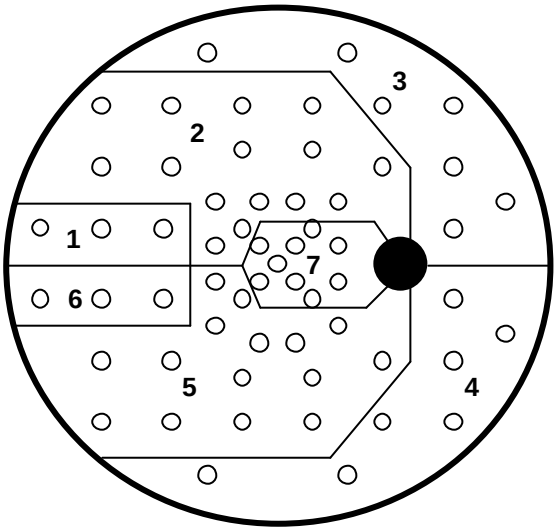


Figura 4.2: Distribución del campo visual en zonas.

4.4 Clasificación de campos visuales mediante mapas autoorganizados (SOM)

4.4.1 Introducción. Algoritmo de aprendizaje SOM

El propósito de un algoritmo de aprendizaje autoorganizado o no supervisado es descubrir patrones significativos o características en los datos de entrada sin la ayuda de un experto externo [67].

Existen varias clases de modelos autoorganizados. Una de ellas comprende las denominadas **redes no supervisadas competitivas**, en las que solamente una neurona, o un grupo de vecinas, pueden quedar finalmente activadas [111]. Dentro de esta clase está el modelo ART (*Adaptive Resonance Theory*), el Neocognitrón

o los mapas autoorganizados de Kohonen. El objetivo en todos ellos es el agrupamiento o categorización de patrones, ya que desarrollan clusters que agrupan patrones, en principio no etiquetados. También pueden servir para otras muchas tareas: cuantificación vectorial, aproximación funcional, procesamiento de imágenes, análisis estadístico y optimización combinatorial.

En los mapas autoorganizados (SOM), las neuronas se organizan en una arquitectura unidireccional de dos capas. La primera es la capa de entrada, que consiste en n neuronas, una por cada variable de entrada. El procesamiento se realiza en la segunda capa, que forma el mapa de características, y consiste habitualmente en un array bidimensional de $N_x * N_y$ neuronas que operan en paralelo, aunque también puede ser mono o tridimensional. El SOM define, por tanto, una correspondencia $\Re^n \rightarrow \Re^1$, $\Re^2$, $\Re^3$. Cada neurona o nodo *i* tiene un vector asociado $m_i \in \Re^n$. El tipo de array o mapa puede ser rectangular o hexagonal. Cada vector de entrada $x \in \Re^n$ se compara con el vector de pesos m_i de cada neurona o nodo del array, y aquella que proporcione un mejor resultado (menor distancia) en la comparación es la neurona ganadora. La entrada *x* se hace corresponder, por tanto, con esa neurona.

Esa comparación se realiza mediante una métrica, en este caso mediante la distancia Euclídea entre ambos vectores, de tal forma que el nodo que ofrezca una distancia menor, será considerado el ganador (*c*):

$$\|x - m_c\| = \min_i \{\|x - m_i\|\},$$
$$c = \arg\min_i \{\|x - m_i\|\}$$

Una vez que se ha seleccionado el nodo de salida, su vector de pesos correspondiente m_c se ajusta para minimizar la distancia entre éste y el vector de entrada. Para controlar este ajuste se usa un factor de aprendizaje $h_{ci}(t)$. El valor actualizado de *m* en la iteración *t* se calcula mediante la siguiente expresión:

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)]$$

El SOM "entrena" nodos ajustando sus pesos en una vecindad, N_c, alrededor del nodo activado, de forma que puedan reaccionar ante estímulos de entrada similares, es decir, los vectores de pesos de esos nodos pertenecientes a esa vecindad también son modificados. El radio de vecindad disminuye linealmente con el número de iteración durante la sesión de entrenamiento. El resultado es una red entrenada donde los nodos vecinos comparten propiedades similares [4].

Las vecindades que usa el SOM son de dos tipos:

Tipo *Burbuja*:

$$h_{ci}(t) = \begin{cases} a(t) & i \in N_c \\ 0 & i \notin N_c \end{cases}$$

donde $h_{ci}(t)$ es la cantidad de ajuste que el nodo i en N_c recibe en la iteración t. $a(t)$ (*ritmo de aprendizaje*) es una función monótona decreciente con la iteración ($0 < a(t) < 1$) y la vecindad también va disminuyendo según la iteración $N_c = N_c(t)$. Este método aplica un factor de entrenamiento constante a todos los nodos dentro de la vecindad del nodo activado.

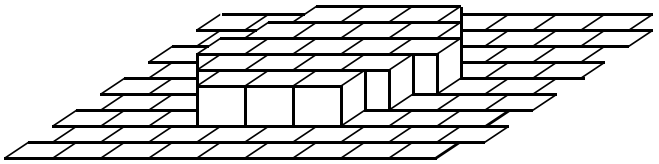


Figura 4.3: Vecindad tipo "burbuja" en un SOM.

Tipo *Gaussiana*:

$$h_{ci}(t) = \begin{cases} a(t) \cdot \exp\left(-\dfrac{\|r_c - r_i\|^2}{2s^2(t)} \right) & i \in N_c \\ 0 & i \notin N_c \end{cases}$$

donde $h_{ci}(t)$ es la cantidad de ajuste que el nodo i en N_c recibe en el instante t y $s(t)$ define la anchura de la vecindad. Este método aplica un factor de entrenamiento que decrece no linealmente a medida que la distancia a partir del nodo seleccionado aumenta.

Tanto $a(t)$ como $s(t)$ son funciones monótonas decrecientes con la iteración, y su forma exacta no es esencial, de hecho pueden ser lineales. En SOM_PAK [88] se usa una función del tipo $a(t) = A/(B+t)$, donde A y B son constantes.

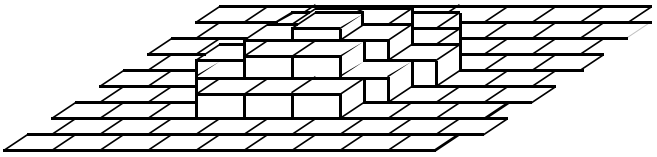


Figura 4.4: Vecindad tipo "Gaussiana" en un SOM.

Como ya se ha indicado, el array de nodos de la capa de salida o mapa SOM puede tener una topología rectangular o hexagonal. En la rectangular los nodos se disponen en filas y columnas. Cada nodo, m_c, puede interaccionar con los ocho nodos que lo rodean. Todas las distancias se miden comparando las coordenadas X e Y. Las vecindades en esta topología son regiones rectangulares alrededor del nodo m_c.

La topología hexagonal permite una interacción de nodos radial. Comenzando por la segunda fila, cada fila siguiente está trasladada la mitad de un nodo hacia la derecha. Esta traslación cambia la topología de forma que cada nodo interior está rodeado de seis, en lugar de ocho como en la rectangular. Las vecindades en esta topología son regiones hexagonales alrededor del nodo m_c.

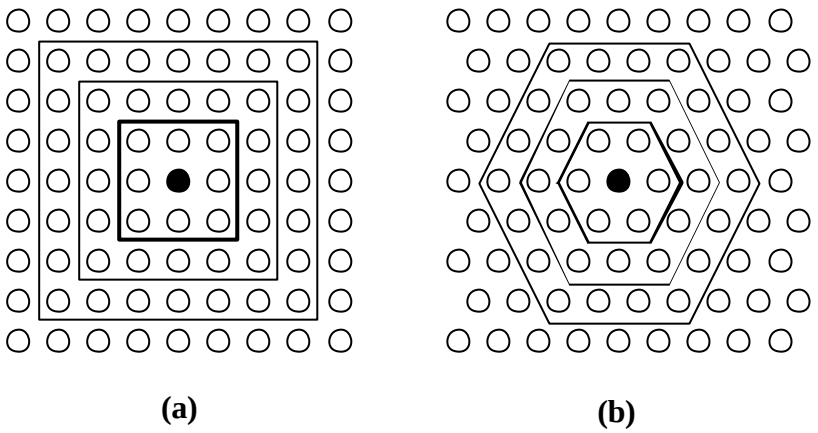


Figura 4.5: Topología rectangular (a) y hexagonal (b) de un SOM.

El algoritmo SOM consiste en los siguientes pasos para cada vector de entrenamiento:

1. Calcular la salida de cada neurona o nodo.
2. Seleccionar el mejor nodo m_c.
3. Determinar el radio de la vecindad de entrenamiento N_c.
4. Entrenar todos los nodos dentro de la vecindad N_c de m_c actualizando sus pesos.

El entrenamiento global de los mapas SOM se realiza mediante los siguientes pasos:

1. **Inicialización del mapa:** Los vectores de referencia del mapa se inicializan primero a unos valores tentativos. Se definen también la topología y la función de vecindad.
2. **Entrenamiento del mapa:** El mapa se entrena mediante el algoritmo de mapas autoorganizados SOM. Este entrenamiento se suele realizar en dos fases:

a. **Fase de ordenación:** Durante esta fase inicial del algoritmo tiene lugar la ordenación topológica de los vectores de pesos. El radio de vecindad se toma aproximadamente igual al diámetro del mapa y decrece a uno durante el entrenamiento, mientras el ritmo de aprendizaje decrece a cero.

b. **Fase de convergencia:** En esta fase se realiza un ajuste fino del mapa de modo que la distribución de pesos se ajuste más a las entradas. Normalmente esta fase es más larga que la anterior. El ritmo de aprendizaje es bastante más pequeño y el radio de vecindad es también más pequeño.

Después de estas dos fases de entrenamiento el mapa está listo para ser examinado y usarse para diversas aplicaciones.

3. **Evaluación del error de cuantización:** Éste se define como el error cuadrático que se comete al comparar el vector de entrada con el vector de pesos del nodo ganador. Se calcula el error de cuantización medio del mapa para un conjunto de ejemplos. El mapa mejor entrenado es el que tiene menor error de cuantización medio.

4. **Etiquetado del mapa:** En algunas aplicaciones prácticas de este tipo de mapas el objetivo final es lograr clasificar patrones nuevos, para ello es necesario asignar a cada nodo del mapa una etiqueta con la clase a la que pertenecen, de forma que los nuevos patrones que activen ese nodo pertenecerán a esa clase. Para lograr esto se utiliza un conjunto de ejemplos previamente clasificados y se compara el vector de pesos de cada una de ellas con el de cada neurona del mapa, seleccionando la mejor neurona. Hay muchas formas de llevar a cabo este proceso del etiquetado. Al final, al menos un subconjunto de los nodos se pueden llegar a etiquetar.

5. **Evaluación del mapa:** Una vez realizado el ajuste de pesos y el etiquetado, el mapa ya está preparado para clasificar ejemplos nuevos. A partir de esos nuevos conjuntos se puede validar la tasa de aciertos de cada mapa para poder elegir el mejor clasificador.

4.4.2 Desarrollo de la aplicación GESL

Para el entrenamiento de los Mapas Autoorganizados se partió del software SOM_PAK [88] desarrollado en la Universidad Técnica de Helsinki. Este entrenamiento requiere el uso de un elevado número de parámetros variables con amplio rango de valores y cuya interrelación no es clara, por lo que no es posible determinar a priori los valores óptimos con vistas a la obtención de resultados satisfactorios. Por ello es necesario realizar el entrenamiento de un número muy elevado de mapas que cubran el mayor rango posible de valores de los parámetros.

Para poder llevar a cabo esto en el menor tiempo posible se ha desarrollado una herramienta llamada GESL (Generador de Experimentos SOM y LVQ) [51]

que genera baterías de entrenamiento a partir de los rangos de valores asignados a los parámetros, y realiza de forma automática el entrenamiento de los mapas definidos en esas baterías. Así mismo asigna etiquetas a las neuronas de la capa de salida de la red neuronal minimizando, en la medida de lo posible, el número de neuronas de dicha capa que quedan sin etiquetar.

Debido al gran número de mapas entrenados, se necesita poder evaluar y filtrar los resultados para reducir el volumen de datos a analizar. Este filtrado debe realizarse de acuerdo a los parámetros de *sensibilidad* y *especificidad*, que se calculan para cada una de las clases de salida del mapa. Esta evaluación y filtrado también se realizan de forma automática con la herramienta GESL.

GESL ha sido implementada en lenguaje Visual C++ 5.0 para el entorno MS-Windows. Se ha tomado como referencia el paquete ya indicado (SOM_PAK 3.1) y se han hecho una serie de modificaciones de los algoritmos del mismo.

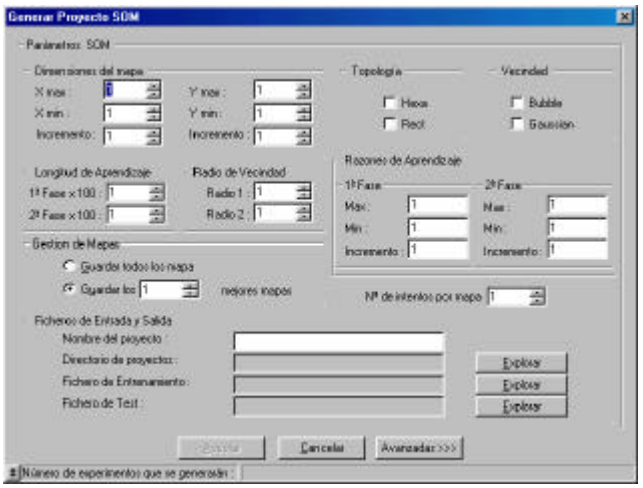


Figura 4.6: Ventana de GESL para generar una batería de entrenamientos SOM [51].

4.4.3 Modificación de los algoritmos del SOM_PACK

Inicialmente, intentamos clasificar nuestros datos [5] utilizando directamente el software SOM_PACK [88]. Nos referimos a los mismos datos de los que hemos hablado en el apartado 4.1, distribuidos de la misma forma: 180 ejemplos de entrenamiento y 48 de prueba.

Cuando se quisieron clasificar en cinco categorías se obtuvieron los siguientes resultados (Tabla 4.3):

Clase	**Tasa de aciertos (precisión diagnóstica)**
Glaucoma	95
Cataratas	43
HTA	0
Diabetes	0
Normal	0
Media global	45.7

Tabla 4.3: Resultados obtenidos en la clasificación del campo visual en cinco categorías utilizando el SOM_PACK [88].

Al lograr tasas de acierto tan bajas, se intentó lograr tan solo el objetivo de diferenciar entre dos clases, consiguiendo los resultados siguientes (Tabla 4.4):

Clase	**Tasa de aciertos (precisión diagnóstica)**
Glaucoma	100
No glaucoma	52
Media global	71.7

Tabla 4.4: Resultados obtenidos en la clasificación del campo visual en dos categorías utilizando el SOM_PACK [88].

Pensamos que con este tipo de redes, por sus propiedades características, debíamos poder conseguir mejores clasificaciones, y por ello modificamos los programas originales.

Para las tres primeras etapas del entrenamiento global de un mapa SOM explicadas en el apartado 4.4.1: inicialización, entrenamiento y evaluación del error de cuantización medio, hemos usado el software SOM_PACK [88]. Para las etapas 4 y 5 hemos creído conveniente diseñar nuestro propio software, para intentar mejorar los resultados que se habían obtenido con el SOM_PACK [5].

Además se ha añadido al software la capacidad de filtrado de la batería de experimentos SOM entrenados para seleccionar el mejor.

Etiquetado del mapa

Se utiliza para esta etapa un conjunto de ejemplos previamente clasificados y por lo tanto cada uno tiene su etiqueta correspondiente.

1. Se calcula la distancia entre el vector de pesos de un ejemplo de la muestra de entrenamiento y cada uno de los vectores de pesos del mapa entrenado.
2. En la neurona ganadora se almacena la etiqueta del ejemplo de entrenamiento que la ha activado.

Los pasos 1 y 2 se repiten hasta que ya no haya más ejemplos en la muestra de entrenamiento. Al final, tendremos en cada neurona del mapa el número de ejemplos de cada clase que la han activado.

3. Cada neurona se etiqueta con la clase que más ejemplos han caído en esa neurona.
4. Si hay dos clases con el mismo número de ejemplos que han activado la neurona correspondiente, se etiqueta con la clase cuya suma de distancias a esa neurona es menor.
5. Si no ha caído ningún ejemplo en una neurona se deja provisionalmente sin etiqueta.
6. Para disminuir el número de neuronas sin etiquetar, se calcula la distancia entre la neurona sin etiqueta y todas las demás del mapa. Se le asigna la etiqueta de la más cercana en cuanto a distancia euclídea, si ésta tuviera alguna.

Evaluación del mapa

Como nuestro objetivo es clasificar ejemplos nuevos de datos de campos visuales, necesitamos conseguir un clasificador cuya tasa de aciertos sea lo más elevada posible. Por este motivo no vamos a elegir el mejor mapa de acuerdo al menor error de cuantización. La evaluación se va a realizar calculando los parámetros de *sensibilidad* y *especificidad* para cada una de las clases de salida del mapa, como suele ser habitual en la práctica médica.

GESL está diseñado para poder entrenar y evaluar un mapa SOM o una batería de mapas. Si son varios los mapas entrenados, además de calcular la sensibilidad y especificidad, busca los *n* mejores atendiendo a la mejor precisión diagnóstica o tasa de aciertos.

4.4.4 Selección de los parámetros para el entrenamiento de las redes neuronales

El proceso de aprendizaje inmerso en la computación de un mapa de características es estocástico por naturaleza, lo que significa que la precisión del mapa depende del número de iteraciones del algoritmo SOM. Además, el éxito del mapa es críticamente dependiente de cómo sean seleccionados el parámetro ritmo o razón de aprendizaje y la función de vecindad, y del orden en que se proporcionen los ejemplos a la red. Desafortunadamente, no hay ninguna base teórica para la selección de estos parámetros. Se determinan normalmente por un proceso de prueba y error. No obstante, las siguientes observaciones proporcionan una guía útil. [67, 86, 88].

1. El parámetro razón de aprendizaje usado para actualizar el vector de pesos durante la fase de ordenación (las primeras 1000 iteraciones más o menos) debería comenzar con un valor cercano a la unidad y debería decrecer gradualmente, quedándose por encima de 0.1. La forma exacta de

variación de $a(t)$ con t no es crítica; puede ser lineal, exponencial, o inversamente proporcional a t.

2. Las iteraciones restantes del algoritmo, como ya se ha indicado, son necesarias principalmente para el ajuste del mapa. Para una buena precisión estadística, $a(t)$ debería mantenerse durante la fase de convergencia en un valor pequeño (sobre el orden de 0.01 o menos) para un periodo de tiempo moderadamente largo, que es típicamente de miles de iteraciones.
3. El mapa conviene que sea rectangular, es decir, $N_x \neq N_y$.

Para la función de vecindad se usa la función que implementa el SOM_PACK. El radio disminuye linealmente hasta 1, es decir, en ese momento el entrenamiento afecta a la neurona ganadora y a las inmediatamente vecinas.

A partir de los datos generamos los ficheros de entrenamiento y de prueba, ajustándonos al formato indicado en la descripción del paquete SOM_PAK 3.1 [88].

Con estos ficheros y utilizando la posibilidad de entrenamiento y evaluación de resultados de la aplicación desarrollada GESL, se realizó una tanda de entrenamientos de prueba con objeto de restringir en la medida de lo posible el rango de valores de algunos de los parámetros de entrada.

Al realizar estos entrenamientos de prueba, en los Mapas Autoorganizados (SOM) se observó que para dimensiones de mapa con menos de 21 neuronas, no aparecían representadas todas las clases de salida. Así, por ejemplo, en un mapa de dimensión 5x3, entrenado para clasificar entre las cinco posibles clases de salida: Glaucoma (G), Normal (N), Diabetes (D), HTA (H), y Cataratas (C), se puede observar que no aparecen neuronas asignadas a la clase Diabetes (Figura 4.7).

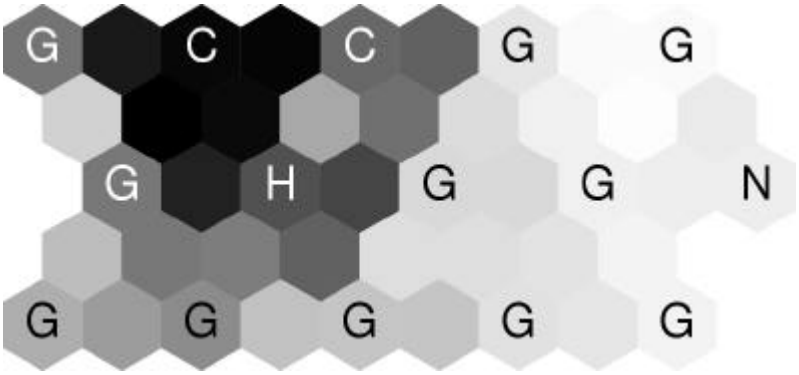


Figura 4.7: Mapa SOM para clasificar en 5 categorías. La clase Diabetes no está representada.

Debido a la falta de representación de determinadas clases de salida, se fijó la dimensión mínima de los mapas en 7x3, ya que a partir de este tamaño de mapa todas las clases de salida aparecen representadas en los mapas.

La dimensión máxima de los mapas autoorganizados se fijó en 11x5, ya que en aquellos mapas con más de 55 ó 60 neuronas, una gran cantidad de ellas quedaban sin representación en la capa de salida. Así, por ejemplo, en el mapa de la Figura 4.8, que tiene una dimensión de 12x8, y que se entrenó para clasificar en cinco categorías: Glaucoma (G), Normal (N), Diabetes (D), HTA (H), y Cataratas (C), podemos observar cómo algunas neuronas no están asignadas a ninguna clase de salida y tampoco han podido aproximarse a la etiqueta de la neurona más cercana.

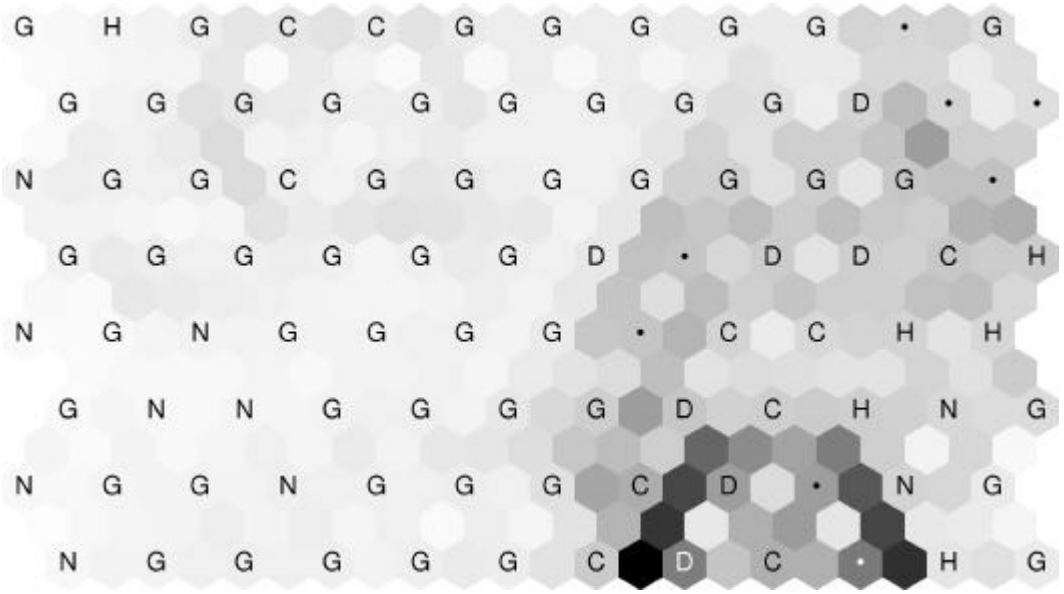


Figura 4.8: Mapa SOM para clasificar en 5 categorías. Hay varias neuronas sin etiquetar.

Para otro de los parámetros, los ciclos de aprendizaje, se eligió su valor basándose en la evaluación del error cuadrático medio (*q-error*) asociado a cada mapa. Este error da una orientación sobre el grado de aprendizaje de la red, de forma que a un *q-error* próximo a cero le corresponde un aprendizaje memorístico por parte de la red. Este aprendizaje memorístico es función del cociente entre los ciclos de aprendizaje, es decir, el número de vectores o ejemplos que se presentan a la red durante el entrenamiento, y el número de ejemplos disponible. Si el número de ejemplos es inferior a los ciclos de aprendizaje, éstos serán presentados repetidamente a la red, de forma cíclica o aleatoria, hasta completar los ciclos de aprendizaje, por lo que si el cociente entre estos dos valores es muy elevado la red realizará un proceso de memorización y no de aprendizaje de los datos de entrada.

Después de realizar estos entrenamientos de prueba para acotar los rangos de valores para los parámetros, en la primera fase del entrenamiento tomamos un radio de vecindad igual a 4, el número de iteraciones o ciclos de entrenamiento igual a 1000 y la razón de aprendizaje ($\alpha 1$) con valores de 1, 0.5, 0.4, 0.05, 0.04, 0.03. En la segunda fase se tomó un radio igual a 2, el número de iteraciones de entrenamiento se aumentó a 50000 y a la razón de aprendizaje ($\alpha 2$) se le asignaron valores de 1, 0.4, 0.3, 0.2, 0.1, 0.05, 0.04, 0.03, 0.02, 0.01, siendo

siempre $\alpha 1 < \alpha 2$. Para todos los mapas usamos topología hexagonal y vecindad tipo burbuja (*Bubble*) [51].

4.4.5 Resultados

A partir de las acotaciones de los rangos de valores de los parámetros de entrenamiento y haciendo uso de la aplicación desarrollada GESL, se entrenaron mapas capaces de clasificar el CV en una sola etapa. En esta primera fase se consideraron redes para la clasificación en dos categorías (glaucoma y no glaucoma), tres (glaucoma, normal y otras patologías), y cinco (glaucoma, normal, diabetes, HTA y cataratas).

Anteriormente se ha explicado que se ha tomado la decisión de disminuir la dimensión de los vectores de entrada a la red para poder solucionar en parte el problema ocasionado por el número pequeño de ejemplos en las muestras de entrenamiento y prueba, utilizando la experiencia de los oftalmólogos: división del campo en siete zonas y cálculo de la media de los defectos de sensibilidad de cada zona. Para poder valorar la conveniencia de esta decisión se han realizado baterías de experimentos primero con los 59 datos iniciales y después con las medias de las zonas. De esta forma se pueden comparar los resultados, analizando si la decisión tomada mejora la tasa de aciertos.

Resultados con los 59 puntos del campo visual

Los resultados obtenidos en esta fase, utilizando un único mapa y los 59 puntos como entradas, se observan en las Tablas 4.5, 4.6 y 4.7. En ellas se muestran las categorías de clasificación, la topología de los mapas SOM, y los resultados de sensibilidad, especificidad y precisión diagnóstica.

Topología*	a **1**	a **2**	**Variables**	**S(%)**	**E(%)**	**PD (%)**
7x3	0.04	0.01	59 puntos	95	82.14	87.5
7x5	0.05	0.03	59 puntos	85	89.29	87.5

Tabla 4.5: Los dos mejores mapas SOM que clasifican en glaucoma y no glaucoma entrenados con los defectos de sensibilidad de los 59 puntos del campo visual.

Topología*: 7x5	a **1: 0.04**	a **2: 0.04**	**Variables: 59 puntos**
		S(%)	**E(%)**
Glaucoma		95	71.43
Normal		88.89	92.31
Otras patologías		42.11	93.10
Precisión Diagnóstica (%)		72.91	

Tabla 4.6: El mejor mapa SOM entrenado con los defectos de sensibilidad de los 59 puntos del campo visual que clasifica en tres categorías: Glaucoma, Normal y Otras patologías distintas de Glaucoma.

Topología*: 11x4	a **1: 0.03**	a **2: 0.01**	**Variables: 59 Puntos**	
			S(%)	**E(%)**
Glaucoma			95	82.14
Normal			88.89	89.74
Cataratas			28.57	92.68
Diabetes			28.57	95.12
HTA			20	95.35
Precisión Diagnóstica (%)			66.6	

Tabla 4.7: El mejor mapa SOM entrenado con los defectos de sensibilidad de los 59 puntos del campo visual que clasifica en cinco categorías: Glaucoma, Normal, Cataratas, Diabetes y HTA.

*Dimensión de la matriz del mapa.
$\alpha 1$ y $\alpha 2$ = Razones de aprendizaje de las dos fases de entrenamiento.
Variables = Datos del vector de entrada.
S= Sensibilidad. E= Especificidad. MD= Desviación media. LV= varianza de la pérdida.

Resultados utilizando la división en siete zonas

Los resultados obtenidos en esta fase, utilizando un único mapa y las medias de los defectos de sensibilidad de las siete zonas del campo visual como entradas, se observan en las Tablas 4.8, 4.9 y 4.10. En ellas se muestran las categorías de clasificación, la topología de los mapas SOM, y los resultados de sensibilidad, especificidad y precisión diagnóstica.

Topología*	a **1**	a **2**	**Variables**	**S(%)**	**E(%)**	**PD (%)**
8x3	0.03	0.01	7 zonas y MD	95	92.86	93.75

Tabla 4.8: El mejor mapa SOM que clasifica en glaucoma y no glaucoma entrenado con las medias de los defectos de sensibilidad de las siete zonas del campo visual.

Topología*: 11x5	a **1: 0.04**	a **2: 0.01**	**Variables: 7 Zonas y MD**	
			S(%)	**E(%)**
Glaucoma			95	85.71
Normal			88.89	87.18
Otras patologías			57.89	96.55
Precisión Diagnóstica (%)			79.17	

Tabla 4.9: El mejor mapa SOM que clasifica en tres categorías: Glaucoma, Normal y Otras patologías distintas del Glaucoma, entrenado con las medias de los defectos de sensibilidad de las siete zonas del campo visual.

Topología*: 11x5	**a 1: 0.04**	**a 2: 0.01**	**Variables: 7 Zonas y MD**	
			S(%)	**E(%)**
Glaucoma			100	85.71
Normal			88.89	87.18
Cataratas			57.14	90.24
Diabetes			0	95.12
HTA			20	100
Precisión Diagnóstica (%)			68.75	

Tabla 4.10: El mejor mapa SOM que clasifica en cinco categorías: Glaucoma, Normal, Cataratas, Diabetes y HTA, entrenado con las medias de los defectos de sensibilidad de las siete zonas del campo visual.

*Dimensión de la matriz del mapa.
$\alpha 1$ y $\alpha 2$ = Razones de aprendizaje de las dos fases de entrenamiento.
Variables = Datos del vector de entrada.
S= Sensibilidad. E= Especificidad. MD= Desviación media. LV= varianza de la pérdida.

Conclusión

A la vista de los resultados obtenidos para los mejores mapas SOM en ambos tipos de experimentos: usando los 59 puntos del campo visual, o la división en siete zonas, se puede observar, como era de esperar debido a la experiencia de los expertos, que han mejorado en este último caso. La precisión diagnóstica en las mejores redes entrenadas con las medias de sensibilidad de las siete zonas es mejor en todos los casos que la precisión diagnóstica de las redes entrenadas con los defectos de los 59 puntos del campo visual.

Hemos comprobado que al introducir las categorías de cataratas, diabetes y HTA para diferenciar los campos visuales en cinco clases, las redes no son capaces de distinguir adecuadamente entre esas nuevas categorías, obteniéndose resultados de sensibilidad y especificidad muy bajos para considerar fiables los resultados proporcionados por el mapa.

4.4.6 Conclusiones

La reducción de dimensionalidad en los vectores de entrada a las redes SOM ha mejorado los resultados obtenidos con estos mapas en la clasificación de campos visuales, como se puede observar en la Figura 4.9 en la que se representa la variación de la precisión diagnóstica según los distintos experimentos realizados.

Se han obtenido unos resultados muy aceptables en la clasificación de campos visuales glaucomatosos con defectos incipientes.

El uso de la herramienta GESL ha facilitado enormemente la realización de los experimentos, y los algoritmos GESL basados en los del paquete SOM_PAK 3.1 han mejorado los resultados obtenidos en trabajos previos [5].

A pesar del reducido número de casos contenidos en las muestras de entrenamiento y prueba, se puede concluir que los Mapas Autoorganizados son una herramienta adecuada para discriminar campos visuales glaucomatosos.

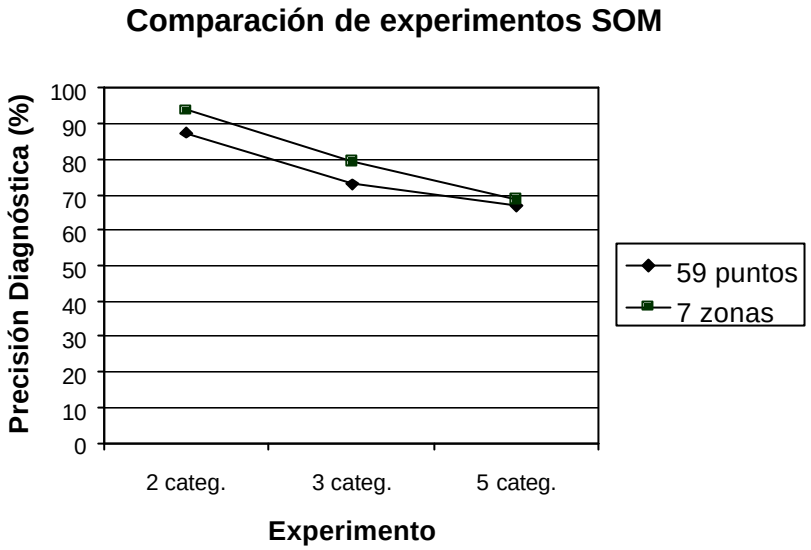


Figura 4.9: Comparación de los experimentos SOM, realizados con los defectos de sensibilidad de los 59 puntos del campo visual o con las medias de los defectos de sensibilidad de las siete zonas del campo visual, mediante la Precisión Diagnóstica.

4.5 Clasificación de campos visuales utilizando Razonamiento Basado en Casos

4.5.1 Introducción al Razonamiento Basado en Casos

El Razonamiento Basado en Casos (CBR) es un método para resolver problemas recordando situaciones previas similares y reutilizando la información y el conocimiento sobre esa situación [89, 96]. La idea original básica de Riesbeck y Schank es la siguiente:

> *Lo que hace CBR es resolver problemas nuevos adaptando las soluciones dadas a otros resueltos con anterioridad* [129].

Leake [63, 96] identifica cinco problemas importantes en IA que pueden mejorarse utilizando CBR: adquisición del conocimiento, mantenimiento del conocimiento, mejora de la eficiencia para resolver problemas, aumento de la calidad de las soluciones, y aceptación de éstas por el usuario.

Aamod y Plaza [3] describen CBR como un proceso que consta de cuatro pasos:

- *Recuperar* el caso o los casos más similares.
- *Reutilizar* la información y el conocimiento de ese caso para resolver el problema.
- *Revisar* la solución propuesta si es necesario.
- *Guardar* las partes de esta experiencia que se consideren útiles para resolver futuros problemas.

Estos pasos se ilustran en la Figura 4.10.

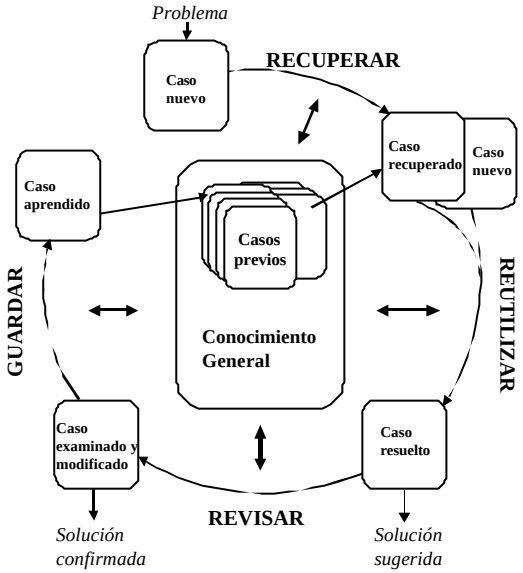


Figura 4.10: El ciclo CBR [3].

En la Figura 4.11 se muestra un diagrama de flujo clásico para CBR. Es básicamente el mismo que propusieron Riesbeck y Schank [63, 129].

El proceso que se sigue en este enfoque es el siguiente:

El usuario introduce el problema a través de la interfaz.

La evaluación de la situación es el proceso de determinar las características del problema introducido aplicables a la situación actual. Es decir, hay que interpretar un nuevo caso o elaborar su **representación** más adecuada para que su descripción esté lo más de acuerdo posible con los casos almacenados en la biblioteca de casos. Es uno de los retos importantes en CBR.

Kolodner [89] define un **caso** como un trozo contextualizado de conocimiento que representa una experiencia y que enseña una lección fundamental para lograr las metas del razonador.

Básicamente un caso está formado por

- El *problema*: el estado del universo cuando el caso ocurrió y la descripción del problema.
- La *solución* descrita para el problema específico o la reacción a la situación de que se trate.
- El *resultado* o estado del universo después de la aplicación de la solución o después de que ocurrió el caso.

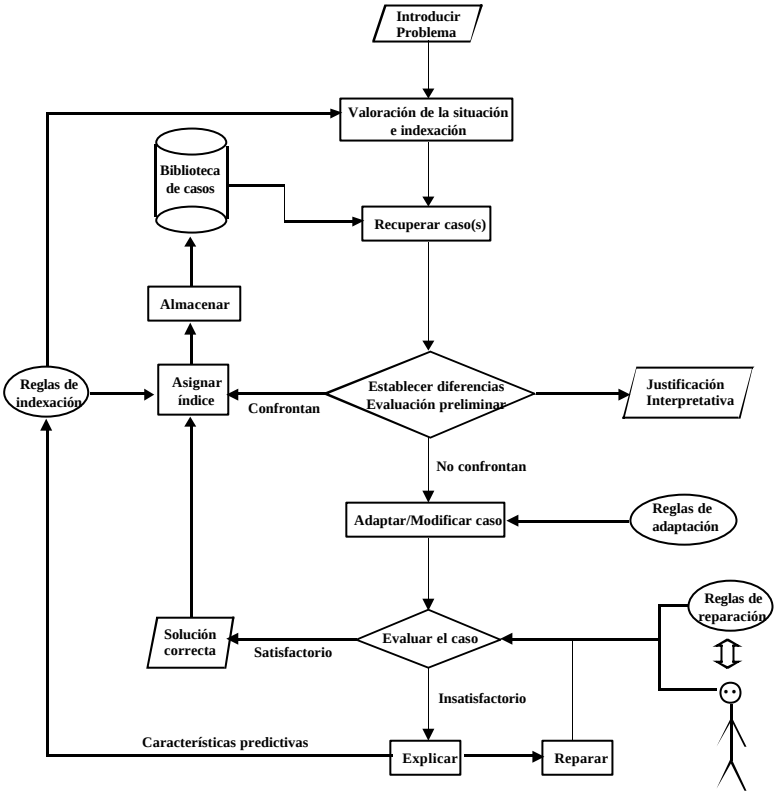


Figura 4.11: Diagrama de flujo de CBR [126].

También en este paso del proceso CBR tiene lugar la indexación de los casos, es decir, se les asigna unas etiquetas para identificar su naturaleza, contenido y contexto. Esto se hace para facilitar su posterior recuperación.

El proceso de indexación consta de dos partes:

- seleccionar el vocabulario de indexación apropiado que son los descriptores que describen el caso,
- asignar un índice: designar ciertas partes del vocabulario que identifican las características.

Distintos investigadores en CBR han sugerido varias guías para seleccionar índices para un caso particular [89, 161]. Son las siguientes: los índices deben ser predictivos de la parte de solución sobre la que tienen influencia, útiles para conseguir la finalidad para la que se usa el caso, lo suficientemente abstractos para poder usarse en una amplia selección de situaciones, y lo suficientemente concretos para no dificultar excesivamente el proceso de inferencia.

Al seleccionar los índices hay que tener en cuenta una serie de factores [89]:

1. El rango de tareas de las que es responsable la herramienta basada en casos.
2. El rango de casos disponibles para esas tareas.
3. El grado y direcciones en que el sistema será ampliado en el futuro.

Kolodner [89] describe dos enfoques principales para seleccionar un índice:

1. El enfoque funcional: examina un corpus de casos disponibles y la tarea que deben realizar, mirando para qué puede usarse cada caso y el modo en que debe ser descrito para conseguirlo.
2. El enfoque de recuerdos: examina las clases de recuerdos que son normales entre los expertos que hacen las tareas designadas, buscando similitudes relevantes entre las nuevas situaciones que los expertos introducen y los casos que recuerdan, para averiguar qué clases de descriptores son los más importantes para juzgar la similitud y en qué circunstancias.

El proceso de indexación explicado anteriormente se sigue normalmente en el indexado manual. La clave de este enfoque es identificar cual de las características de un caso es útil. El indexado manual es útil cuando los casos son bastante complejos, los índices de los casos deben ser muy precisos, no se dispone del conocimiento necesario para entender los casos lo suficientemente bien como para elegir los índices con precisión, las características de los índices son demasiado complejas para determinarlas automáticamente.

Existen también distintos métodos automáticos para indexar los casos. En ellos se utilizan unas reglas de indexación dinámicas para estructurar los casos según sus características. Entre los métodos de indexación automáticos están los siguientes:

- Indexación *basada en una lista de chequeo*. El que desarrolla el sistema elige las características clave (dimensiones) que van a usarse para indexar y crea una lista de chequeo. Cada caso es indexado mediante las dimensiones encontradas en la lista. Las mejores listas de chequeo son aquellas cuyas características tienden a predecir soluciones y consecuencias a lo largo de todo el dominio.
- Indexación *basada en diferencias*. Se apoya en las características extrínsecas que diferencian el nuevo caso de los demás almacenados, y selecciona los índices según esas características. Este método se puede

combinar con el anterior para lograr elegir índices con características predictivas, además de características que permitan discriminar casos.

- Indexación *basada en explicación*. Los dos tipos anteriores proporcionan un mecanismo computacionalmente eficiente pero, en general, no analizan cada caso de modo individual buscando características predictivas. Este método, sin embargo, elige índices apropiados para casos individuales. El razonador intenta analizar por qué funciona o no la solución y luego generaliza la explicación. Los índices se eligen a partir del contenido de la explicación generalizada, es decir, por las características que fueran importantes para la resolución. La mayor desventaja de este método es que tiene elevada complejidad computacional y es altamente dependiente del conocimiento del dominio.
- Métodos de *generalización basada en similitud y explicación*. Proporcionan un conjunto de índices para casos abstractos creados a partir de casos que comparten algún conjunto de características comunes. Las características que no comparten se usan como índices de los casos originales.
- Métodos de *aprendizaje inductivo.* Identifican características predictivas que usan como índices. Esta técnica y sus variantes (algoritmo ID3) son ampliamente usadas en herramientas CBR como por ejemplo en Cognitive Systems ReMind que es la usada en los experimentos de este trabajo.

Después, usando los índices creados, se recupera el caso o los casos que más se parezcan al actual. Los algoritmos de recuperación de casos usan heurísticas para restringir y dirigir la búsqueda. Entre los métodos de recuperación de casos están los siguientes:

- Confrontación mediante el *vecino más próximo*. Es el más simple y el que más se usa. Calcula la similitud entre los casos almacenados y el nuevo, basándose en la comparación de una suma ponderada de características. Cada característica del caso de entrada se compara con su correspondiente del caso almacenado, se calcula el grado de similitud, teniendo en cuenta también la importancia asignada a cada una, y finalmente cada caso almacenado tendrá un valor asignado que permitirá establecer un orden de similitud.
- Recuperación por *inducción*. Este método utiliza algoritmos de inducción como el ID3 para determinar las características que mejor discriminan los casos y las usan para generar un árbol de decisión que se utiliza para guiar la recuperación sin tener que examinar la base de casos entera. En ese sentido es más eficiente que el del vecino más próximo.
- *Inducción guiada mediante conocimiento.* Este proceso considera conocimiento específico del dominio y lo aplica al proceso de inducción. Como guía se usa un modelo causal cualitativo que utiliza jerarquías de

clasificación y redes de inferencia. Esto es útil para refinar un árbol de decisión y también en dominios donde los datos tienen ruido.

- Recuperación mediante *plantilla*. Se llama también recuperación mediante prototipos. El caso nuevo se compara con los almacenados y se obtienen todos aquellos que tienen un conjunto de rasgos característicos. Este método se usa muchas veces como una primera etapa, previa a otro de los anteriores, para limitar el espacio de búsqueda a un conjunto relevante de casos.

Una vez obtenido el caso más parecido al nuevo, se hace una evaluación preliminar de él para identificar las diferencias. Cuando los casos se usan para interpretar situaciones, el caso recuperado sirve para justificar cualquier decisión.

Cuando hay errores de confrontación entre la especificación del problema (caso introducido) y el caso recuperado, es decir, cuando el caso recuperado no proporciona la solución exacta al problema, es necesaria una adaptación del mismo. Puede hacerse mediante la intervención humana o de forma automática.

La adaptación busca diferencias importantes entre el caso recuperado y el actual y aplica unas fórmulas o reglas que tienen en cuenta esas diferencias y sugieren soluciones.

El caso adaptado se evalúa para ver si responde a los estándares esperados. Si la solución propuesta por el caso adaptado no es satisfactoria, entonces explica el fallo. El usuario puede reparar la solución modificándola para que sea satisfactoria, o se pueden usar procedimientos de reparación (reglas, funciones, etc.). Después se vuelve a evaluar.

Finalmente se obtiene un nuevo caso. El sistema ha aprendido de su propia experiencia. Se usan las reglas de indexación para asignar índice a la solución satisfactoria y se almacena en la biblioteca de casos para usar en el futuro.

En los últimos años CBR ha sido aplicado a numerosas tareas, como son planificación, diseño, diagnosis, justificación, explicación, predicción, etc. Dependiendo del tipo de tarea de razonamiento, los sistemas se clasifican según Kolodner [89] en dos categorías: interpretación y resolución de problemas. Althoff [126] proporciona una clasificación más actualizada: tareas de clasificación y de síntesis. Estas clasificaciones se muestran en la Figura 4.12. El proceso de inferencia de las tareas de clasificación normalmente lleva consigo la recuperación de una solución similar a partir de una base de casos y, bien usándola directamente o con una mínima adaptación, proporciona una justificación para subsiguientes acciones. Es inherentemente una tarea más simple que la síntesis. Las tareas de síntesis intentan crear una nueva solución combinando partes de las soluciones previas. Son tareas complejas debido a la interacción entre componentes. Los sistemas CBR que desarrollan este tipo de tareas recuperan soluciones a partir de una biblioteca de buenas soluciones previas, las modifican o adaptan para adecuarse a las situaciones actuales, examinan las nuevas soluciones y las validan.

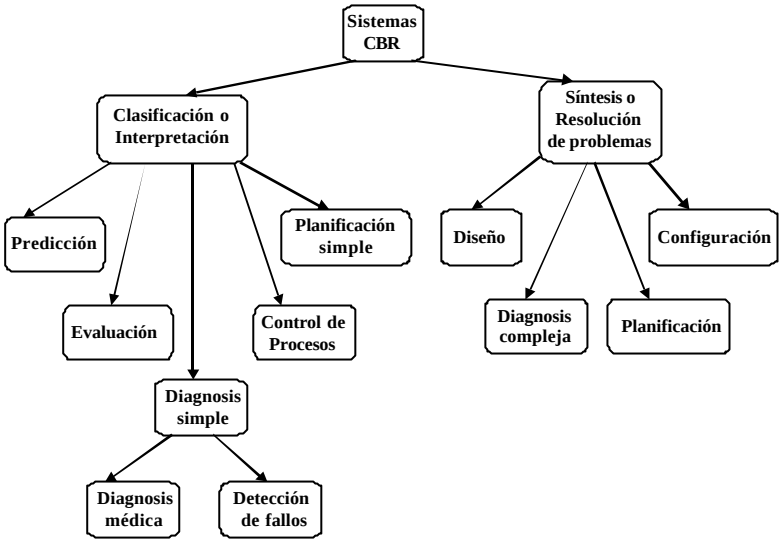


Figura 4.12: Tareas de aplicación CBR [126].

El enfoque de CBR en cuanto a tareas de clasificación ha sido también probado en este trabajo de investigación con vistas a facilitar la tarea de clasificación de campos visuales, y poder utilizar su resultado, si fuera aceptable, en el Sistema Híbrido que realiza el diagnóstico final del paciente.

En el diagnóstico del CV, tenemos casos previamente diagnosticados, y por tanto disponemos de una amplia información de cada caso que ya ha sido analizada. Podemos intentar extraer conocimiento a partir de esa experiencia almacenada, confrontando un nuevo CV de diagnóstico desconocido con esos casos que ya tenemos diagnosticados correctamente.

Para hacer eso, necesitamos una buena representación de nuestros casos. Se debe decidir qué información debería aparecer en un caso. Como ya se ha comentado, los casos se deben indexar, es decir, debemos asignar índices a los casos para facilitar su recuperación. Después usamos un algoritmo de recuperación que se apoye en esos índices para dirigir la búsqueda hacia casos útiles potencialmente. Los métodos para recuperación de casos que hemos usado para nuestro problema actual son recuperación mediante *el vecino más próximo* y recuperación por *inducción*.

Los algoritmos del vecino más próximo determinan la similitud del caso destino respecto al caso fuente, confrontando una suma ponderada de características. Las similitudes están generalmente normalizadas dentro de un rango de 0 a 1 o como un porcentaje de similitud donde 100% significa igualdad. Un algoritmo típico para calcular la confrontación por el vecino más próximo es el usado por el software Cognitive Systems ReMind donde w es el peso de importancia de una característica, *sim* es la función de similitud, y f_i^I y f_i^R son los valores para la característica i de los casos de entrada y el recuperado respectivamente.

$$Valor\ de\ similitud = \frac{\sum_{i=1}^{n} w_i \times sim\left(f_i^{\,I}, f_i^{\,R}\right)}{\sum_{i=1}^{n} w_i}$$

Los algoritmos de inducción determinan qué características son las mejores para discriminar casos, y generan una estructura tipo árbol de decisión para organizar los casos en la memoria.

Hemos usado ReMind para nuestros experimentos. ReMind, producido por Cognitive Systems, fue desarrollado con la ayuda de la Agencia de investigación del Ministerio de Defensa de los Estados Unidos, US DARPA. Fue desarrollado originalmente para Macintosh y posteriormente para las plataformas MS Windows y UNIX. Está disponible como una biblioteca C para embeberlo en otras aplicaciones, y como un entorno interactivo. La versión usada para nuestros experimentos fue la 1.1. Los casos en ReMind se representan como atributos: pares de valores. ReMind ofrece recuperación por el vecino más próximo, inductiva, y otros tipos. La recuperación por el vecino más próximo está informada por los pesos de la importancia definida por el usuario para las características de los casos. La recuperación por inducción lleva consigo la construcción de un árbol de decisión que indexa los casos.

Cuando los casos se recuperan por inducción, ReMind puede dar una explicación de por qué se ha recuperado uno en concreto. Esta explicación describe el camino a través del árbol de decisión inducido que ReMind siguió para recuperar casos similares. La explicación también indica cómo se pueden adaptar los casos recuperados.

4.5.2 Tratamiento de los datos y experimentos

Creamos una base de datos con el defecto de sensibilidad (en decibelios) de cada localización examinada (59 puntos) por el perímetro Octopus. El CV fue dividido en 7 zonas como ya se ha indicado anteriormente y así lo muestra la Figura 4.2. Calculamos la media y la desviación estándar para cada zona.

Diseñamos varias representaciones del caso. Cada una produjo una base de casos diferentes. Estas representaciones fueron:

1. El defecto de sensibilidad (en decibelios) en cada localización examinada (59 puntos).
2. El valor medio de los defectos de sensibilidad en cada una de las 7 zonas en las que fue dividido en CV.
3. La desviación estándar de los defectos de sensibilidad en cada una de las 7 zonas en las que fue dividido el CV.
4. Los dos últimos parámetros juntos (puntos 2 y 3).

Todas las características eran valores reales, excepto el diagnóstico, que era de tipo texto. Usamos el vecino más próximo e inducción para indexar y recuperar los casos de la base de casos.

Los experimentos siguientes fueron algunos de los llevados a cabo para clasificar los casos de prueba en 5 clases diferentes (glaucoma, normal, cataratas, diabetes, HTA):

- Representación 1 con inducción – primero en anchura.
- Representación 1 con *el vecino más próximo*. Todas las características con la misma importancia.
- Representación 2 con inducción – primero en anchura.
- Representación 2 con *el vecino más próximo*. Todas las características con la misma importancia.
- Representación 2 con *el vecino más próximo*. La importancia de las características fue obtenida a partir de la **experiencia de los oftalmólogos y de la bibliografía consultada**, y de acuerdo con ésta, asignamos un peso, de entre los ofrecidos por ReMind, para cada zona:
  - zonas 1 y 6: peso 8
  - zonas 2 y 5: peso 4
  - zonas 3 y 4: peso 2
  - zona 7: sin importancia
- Representación 2 con *el vecino más próximo*. La importancia de las características fue, al igual que en la representación anterior, obtenida a partir de la experiencia de los oftalmólogos, y de acuerdo con ésta, asignamos un peso, de entre los ofrecidos por ReMind, para cada zona:
  - zonas 1 y 6: peso 16
  - zonas 2 y 5: peso 8
  - zonas 3 y 4: peso 4
  - zona 7: peso 2
- Representación 3 con inducción – primero en anchura.
- Representación 3 con inducción – primero en profundidad.
- Representación 4 con inducción – primero en anchura.

Para clasificar los casos de prueba en tres clases (glaucoma, normal y otras patologías) fueron llevados a cabo los siguientes experimentos:

- Representación 1 con inducción – primero en anchura.
- Representación 1 con *el vecino más próximo*. Todas las características con la misma importancia.
- Representación 2 con inducción – primero en anchura.

- Representación 2 con *el vecino más próximo*. Todas las características con la misma importancia.

Fueron realizados otros experimentos para clasificar los casos de prueba en dos clases (glaucoma y no glaucoma):

- Representación 1 con inducción – primero en anchura.
- Representación 1 con *el vecino más próximo*. Todas las características con la misma importancia.
- Representación 2 con inducción – primero en anchura.
- Representación 2 con *el vecino más próximo*. Todas las características con la misma importancia.

4.5.3 Resultados

Los mejores resultados en la clasificación de los casos de prueba en glaucoma, normal, cataratas, diabetes y retinopatía hipertensiva (HTA) fueron obtenidos con la representación de las medias de los defectos de sensibilidad de las siete zonas del campo visual y el método del vecino más próximo. A todas las características de los casos se les asignó la misma importancia. En este experimento, los porcentajes de sensibilidad y especificidad de cada clase correctamente clasificada fueron: glaucoma 80% y 78%, normal 77.8% y 89.74%, cataratas 57.14% y 90.24%, diabetes 0% y 95.12%, HTA 40% y 93.02. La precisión diagnóstica fue 60.41% (Tabla 4.14).

En las tablas siguientes se muestran los resultados obtenidos según el tipo de representación utilizada para clasificar en cinco categorías.

Clase	**Casos prueba (48)**	**S (%)**	**E (%)**
Glaucoma	20	70	64.28
Normal	9	77.8	87.18
Cataratas	7	42.86	90.24
Diabetes	7	14.28	97.56
HTA	5	40	95.35
Precisión Diagnóstica (%)		56.25	

Tabla 4.11: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método de recuperación de casos por inducción. Clasificación en cinco categorías.

Clase	Casos prueba (48)	S (%)	E (%)
Glaucoma	20	55	71.43
Normal	9	66.7	74.36
Cataratas	7	28.57	92.68
Diabetes	7	0	95.12
HTA	5	40	90.7
Precisión Diagnóstica (%)		43.75	

Tabla 4.12: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en cinco categorías.

Clase	Casos prueba (48)	S (%)	E (%)
Glaucoma	20	65	78
Normal	9	66.7	89.74
Cataratas	7	42.86	82.93
Diabetes	7	14.28	90.24
HTA	5	20	93.02
Precisión Diagnóstica (%)		50	

Tabla 4.13: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método de recuperación de casos por inducción. Clasificación en cinco categorías.

Clase	Casos prueba (48)	S (%)	E (%)
Glaucoma	20	80	78
Normal	9	77.8	89.74
Cataratas	7	57.14	90.24
Diabetes	7	0	95.12
HTA	5	40	93.02
Precisión Diagnóstica (%)		60.41	

Tabla 4.14: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en cinco categorías.

Clase	Casos prueba (48)	S (%)	E (%)
Glaucoma	20	75	35.71
Normal	9	11.1	97.43
Cataratas	7	14.28	90.24
Diabetes	7	0	90.24
HTA	5	20	93.02
Precisión Diagnóstica (%)		37.5	

Tabla 4.15: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Las características tienen importancias 8, 4, 2, 0 según la zona del campo visual. Clasificación en cinco categorías.

Clase	Casos prueba (48)	S (%)	E (%)
Glaucoma	20	65	64.28
Normal	9	77.8	87.18
Cataratas	7	28.57	87.80
Diabetes	7	0	95.12
HTA	5	40	93.02
Precisión Diagnóstica (%)		50	

Tabla 4.16: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Las características tienen importancias 16, 8, 4, 2 según la zona del campo visual. Clasificación en cinco categorías.

Clase	Casos prueba (48)	S (%)	E (%)
Glaucoma	20	70	82.14
Normal	9	66.7	84.62
Cataratas	7	42.86	87.80
Diabetes	7	14.29	87.80
HTA	5	20	95.35
Precisión Diagnóstica (%)		52.08	

Tabla 4.17: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las desviaciones estándar de los defectos de las siete zonas del campo visual y el método de inducción tanto con búsqueda primero en anchura como primero en profundidad. Clasificación en cinco categorías.

Clase	Casos prueba (48)	S (%)	E (%)
Glaucoma	20	60	78
Normal	9	66.7	84.62
Cataratas	7	42.86	85.37
Diabetes	7	0	90.24
HTA	5	20	90.7
Precisión Diagnóstica (%)		45.83	

Tabla 4.18: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias y las desviaciones estándar de los defectos de las siete zonas del campo visual y el método de inducción con búsqueda primero en anchura. Clasificación en cinco categorías.

Clase	Casos prueba (48)	59 puntos (1)	59 puntos (2)	Media (1)	Media (2)	Media (3)	Media (4)	SD (1 y 2)	M y SD
Glaucoma	20	70.00	55.00	65.00	80.00	75.00	65.00	70.00	60.00
Normal	9	77.80	66.70	66.70	77.80	11.10	77.80	66.70	66.70
Cataratas	7	42.86	28.57	42.86	57.14	14.28	28.57	42.86	42.86
Diabetes	7	14.28	0.00	14.28	0.00	0.00	0.00	14.29	0.00
HTA	5	40.00	40.00	20.00	40.00	20.00	40.00	20.00	20.00
Precisión Diagnóstica		56.25	43.75	50.00	60.41	37.50	50.00	52.08	45.83

Tabla 4.19: Comparación de la sensibilidad y precisión diagnóstica de los experimentos CBR realizados para clasificar en cinco categorías.

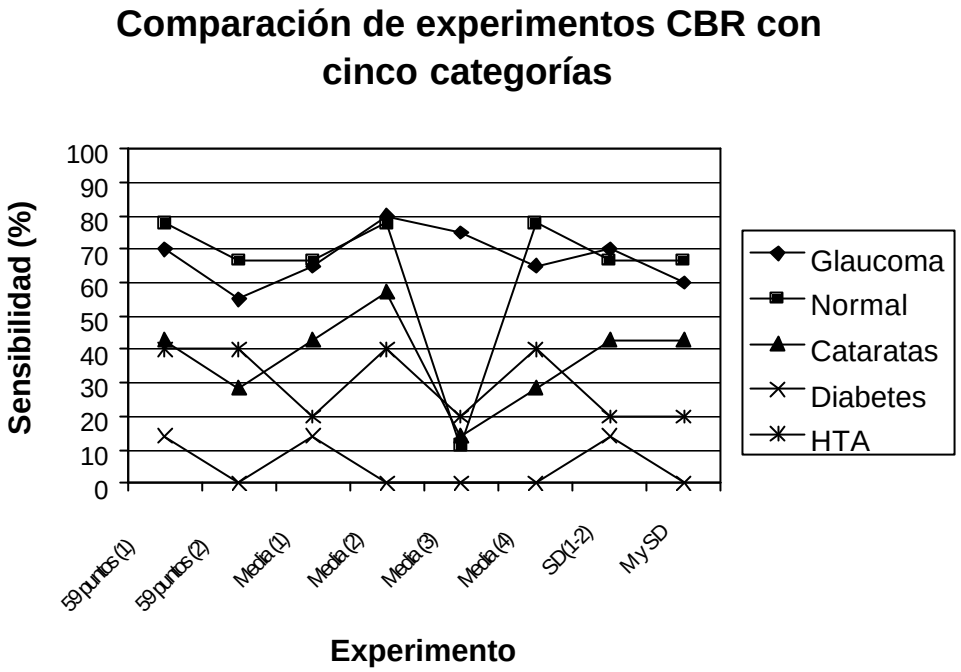


Figura 4.13: Comparación de los experimentos CB R para clasificar en cinco categorías.

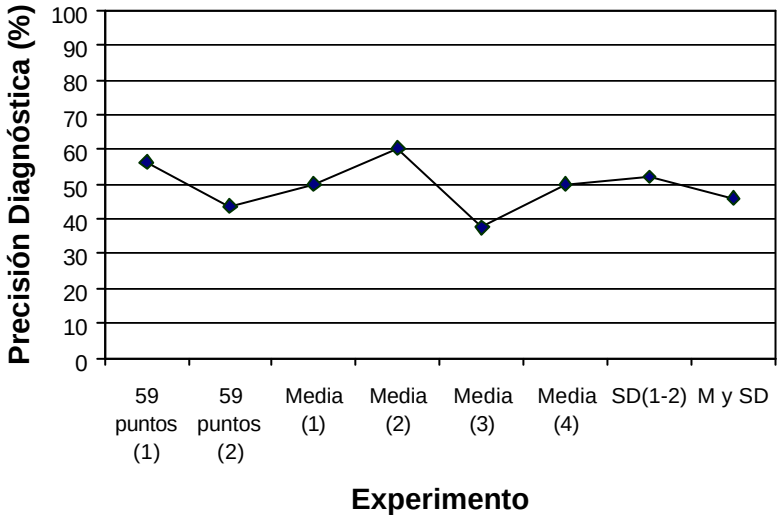


Figura 4.14: Comparación de la Precisión Diagnóstica en los experimentos CBR realizados para clasificar en cinco categorías.

En el segundo estudio para clasificar el conjunto de prueba en glaucoma, normal y otras patologías, los resultados fueron mejores que los anteriores. Los porcentajes de sensibilidad y especificidad mejores fueron los obtenidos utilizando la representación de los casos según las siete zonas del campo visual y el método de recuperación de casos por el vecino más próximo: glaucoma 80% y 78%, normal 77.8% y 89.74%, y otras patologías 52.63% y 82.76%, siendo la precisión diagnóstica 68.75% (Tabla 4.23).

En las tablas siguientes se muestran los resultados obtenidos según el tipo de representación utilizada para clasificar en tres categorías.

Clase	**Casos prueba**	**S (%)**	**E (%)**
Glaucoma	20	75	64.29
Normal	9	77.8	94.87
Otras patologías	19	52.63	86.21
Precisión Diagnóstica (%)		66.7	

Tabla 4.20: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método de recuperación de casos por inducción. Clasificación en tres categorías.

Clase	**Casos prueba**	**S (%)**	**E (%)**
Glaucoma	20	55	71.43
Normal	9	66.7	74.36
Otras patologías	19	36.84	79.31
Precisión Diagnóstica (%)		50	

Tabla 4.21: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en tres categorías.

Clase	**Casos prueba**	**S (%)**	**E (%)**
Glaucoma	20	70	85.71
Normal	9	66.7	87.2
Otras patologías	19	63.15	75.86
Precisión Diagnóstica (%)		66.7	

Tabla 4.22: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método de recuperación de casos por inducción. Clasificación en tres categorías.

Clase	Casos prueba	S (%)	E (%)
Glaucoma	20	80	78
Normal	9	77.8	89.74
Otras patologías	19	52.63	82.76
Precisión Diagnóstica (%)		68.75	

Tabla 4.23: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en tres categorías.

Clase	Casos prueba (48)	59 puntos (1)	59 puntos (2)	Media (1)	Media (2)
Glaucoma	20	75.00	55.00	70.00	80.00
Normal	9	77.80	66.70	66.70	77.80
Otras patologías	19	52.63	36.84	63.15	52.63
Precisión Diagnóstica		66.70	50.00	66.70	68.75

Tabla 4.24: Comparación de la sensibilidad y precisión diagnóstica de los experimentos CBR realizados para clasificar en tres categorías.

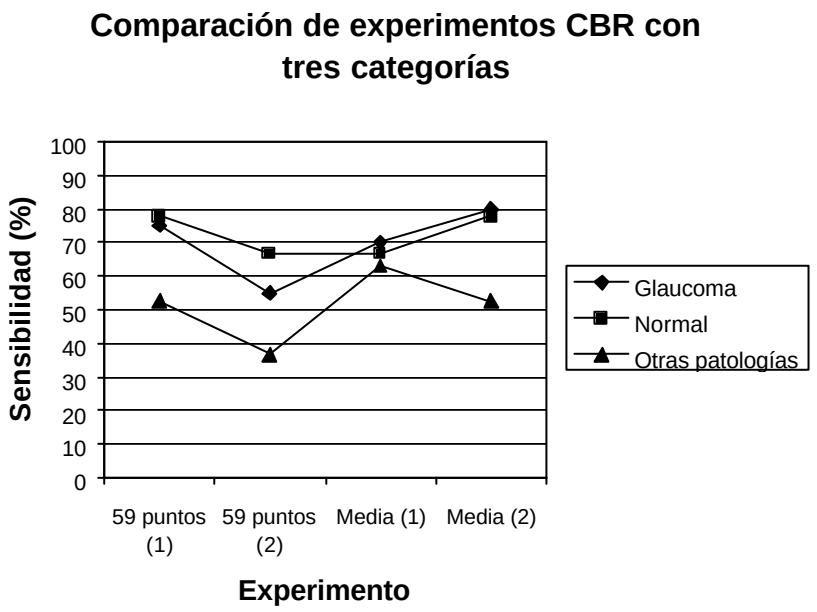


Figura 4.15: Comparación de los experimentos CBR para clasificar en tres categorías.

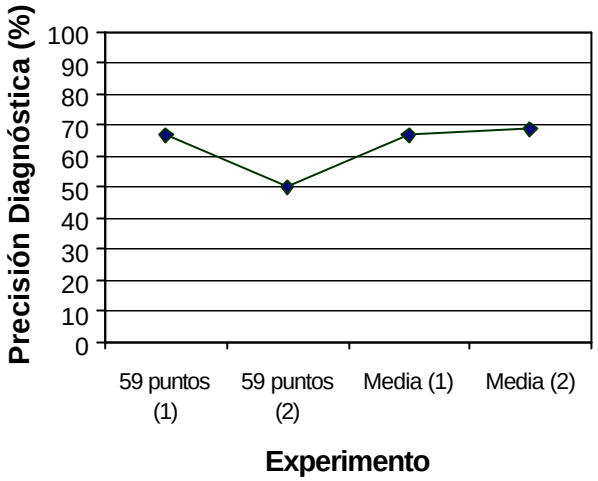


Figura 4.16: Comparación de la Precisión Diagnóstica en los experimentos CBR realizados para clasificar en tres categorías.

El mejor experimento realizado para clasificar solamente en dos categorías diferentes, glaucoma y no glaucoma, fue el de la representación de los casos según las medias de los defectos de las siete zonas del CV y el método de recuperación por el vecino más próximo. Proporcionó los siguientes porcentajes de sensibilidad y especificidad: glaucoma 80% y 78%, no glaucoma 78% y 80%. La precisión diagnóstica fue, por tanto, 79.17% (Tabla 4.28).

En las tablas siguientes se muestran los resultados obtenidos según el tipo de representación utilizada para clasificar en dos categorías.

Clase	Casos prueba	S (%)	E (%)
Glaucoma	20	90	53.57
No glaucoma	28	53.57	90
Precisión Diagnóstica (%)		68.75	

Tabla 4.25: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método de recuperación de casos por inducción. Clasificación en dos categorías.

Clase	Casos prueba	S (%)	E (%)
Glaucoma	20	55	71.43
No glaucoma	28	71.43	55
Precisión Diagnóstica (%)		64.58	

Tabla 4.26: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en dos categorías.

Clase	Casos prueba	S (%)	E (%)
Glaucoma	20	70	78
No glaucoma	28	78	70
Precisión Diagnóstica (%)		75	

Tabla 4.27: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método de recuperación de casos por inducción. Clasificación en dos categorías.

Clase	Casos prueba	S (%)	E (%)
Glaucoma	20	80	78
No glaucoma	28	78	80
Precisión Diagnóstica (%)		79.17	

Tabla 4.28: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en dos categorías.

Clase	Casos prueba (48)	59 puntos (1)	59 puntos (2)	Media (1)	Media (2)
Glaucoma	20	90.00	55.00	70.00	80.00
No glaucoma	28	53.57	71.43	78.00	78.00
Precisión Diagnóstica		68.75	64.58	75.00	79.17

Tabla 4.29: Comparación de la sensibilidad y precisión diagnóstica de los experimentos CBR realizados para clasificar en dos categorías.

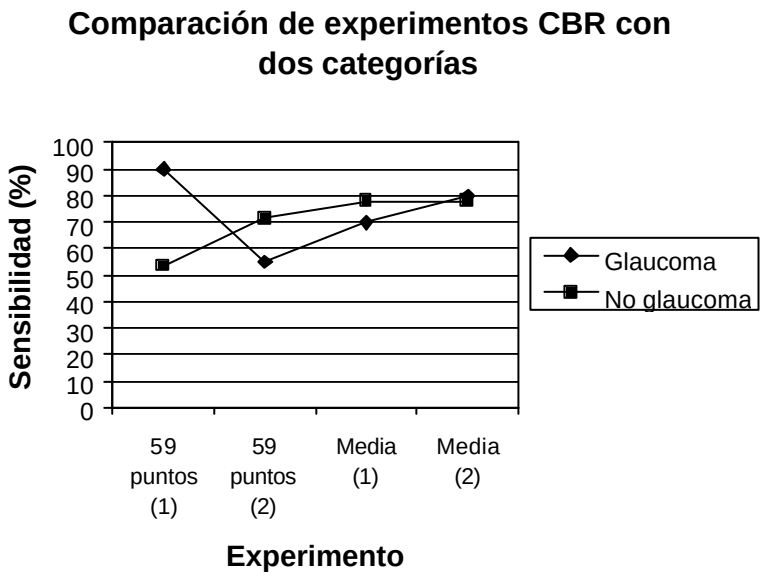


Figura 4.17: Comparación de los experimentos CBR para clasificar en dos categorías.

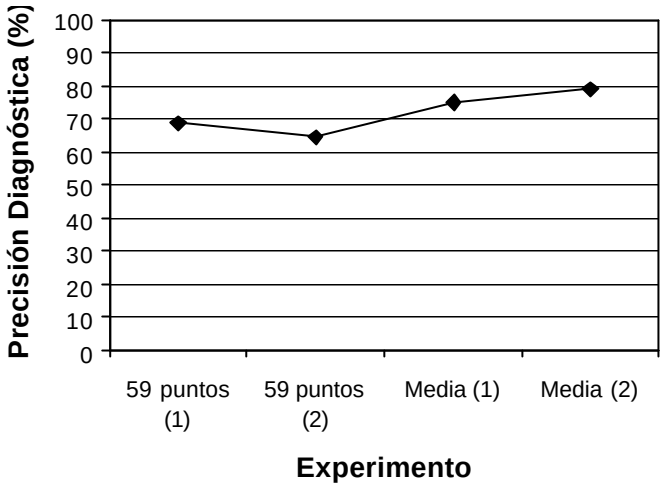


Figura 4.18: Comparación de la Precisión Diagnóstica en los experimentos CBR realizados para clasificar en dos categorías.

A continuación se muestran algunas gráficas que permiten comparar los experimentos realizados para clasificar en dos, tres y cinco categorías o clases, según el número de características que se hayan tomado para representar los casos: los defectos de los 59 puntos del campo visual o las medias de los defectos de las siete zonas en que se ha dividido el campo visual según la experiencia de los oftalmólogos. También se comparan estos experimentos según el método de recuperación de casos que se haya utilizado: recuperación por inducción o recuperación por el vecino más próximo asignando la misma importancia a todas las características del caso.

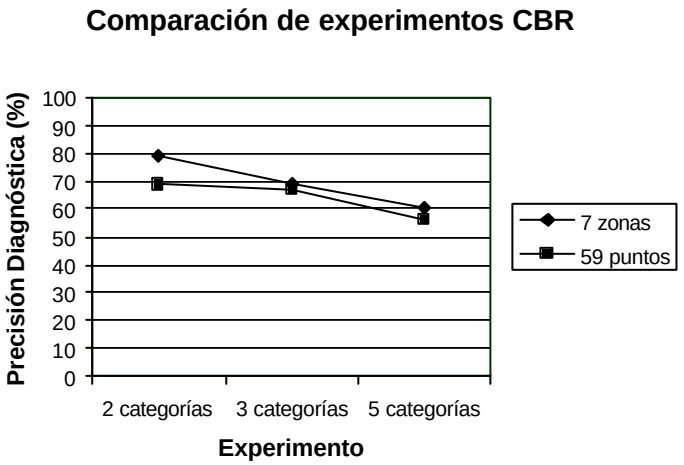


Figura 4.19: Comparación de los mejores experimentos CBR, realizados con los 59 puntos del campo visual o con las medias de los defectos de las siete zonas del campo visual, mediante la Precisión Diagnóstica.

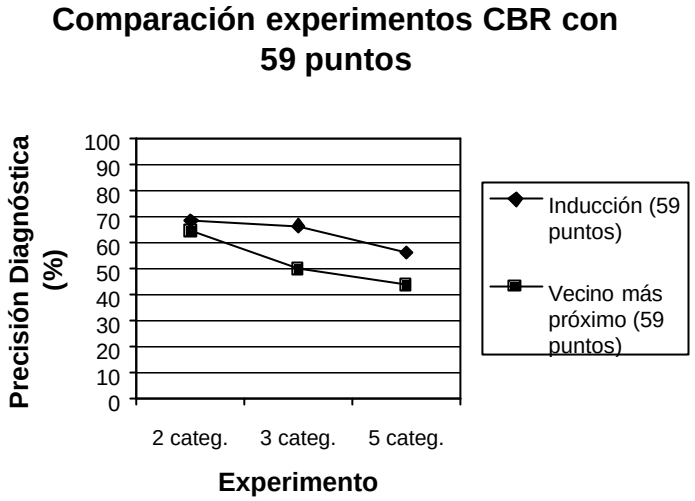


Figura 4.20: Comparación de los mejores experimentos CBR, realizados con los 59 puntos del campo visual mediante la Precisión Diagnóstica, utilizando recuperación de casos por inducción o por el vecino más próximo.

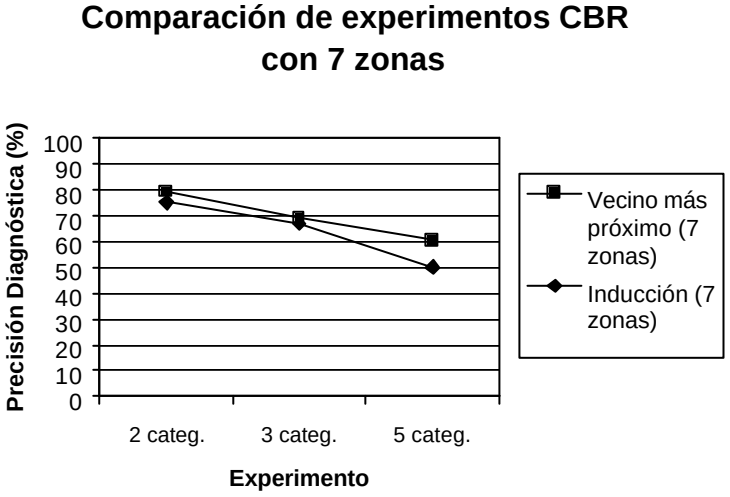


Figura 4.21: Comparación de los mejores experimentos CBR, realizados con las medias de los defectos de las siete zonas del campo visual mediante la Precisión Diagnóstica, utilizando recuperación de casos por inducción o por el vecino más próximo.

4.5.4 Conclusiones

La interpretación del CV implica obtener conclusiones clínicas a partir de información cuantitativa. Como ya se ha indicado en esta Tesis, este análisis se desarrolla generalmente de forma subjetiva por los oftalmólogos de acuerdo a su experiencia y conocimiento. Hemos intentado hacer esta evaluación más objetiva aplicando Razonamiento Basado en Casos. Los sistemas de CBR aprenden de la experiencia, y tienen la capacidad de catalogar un nuevo caso según otros anteriores almacenados en una base de casos. Todas estas características les hacen potencialmente adecuados para la interpretación del CV.

El mejor sistema CBR fue el que clasificó entre glaucomatoso, y no glaucomatoso (se obtuvieron sensibilidades del 80% y especificidades del 78% para glaucoma, 78% y 80% para no glaucoma, siendo la precisión diagnóstica 79.17%). Se utilizó la representación de las medias de los defectos de sensibilidad de las siete zonas del CV y el método de recuperación por el vecino más próximo. Los resultados son peores que los obtenidos mediante mapas SOM, no obstante, son mejores que los proporcionados por los expertos en sus consultas [56].

Por el momento, esta técnica no nos ha permitido extraer conocimiento sobre la influencia de la localización de los puntos dañados dentro del CV en el diagnóstico de glaucoma. Además, la importancia de los puntos o zonas en las que se divide el CV, dada por los expertos, ha sido introducida de forma manual en el sistema CBR cuando se ha usado para la recuperación de los casos *el vecino más próximo*, no consiguiendo mejorar los resultados de los experimentos.

En todos los grupos de experimentos realizados para clasificar en dos, tres o cinco categorías, los mejores experimentos siempre han sido aquellos en los que se han utilizado las medias de los defectos de las siete zonas del campo visual (Figura 4.19). Al disminuir el número de características de los casos hemos logrado mejores clasificaciones.

Se ha observado que los mejores resultados obtenidos con los defectos de los 59 puntos del campo visual siempre fueron realizados con el método de recuperación de casos por inducción (Figura 4.20), y sin embargo los mejores resultados obtenidos con las medias de los defectos de las siete zonas del campo visual siempre fueron realizados con el método de recuperación de casos mediante el vecino más próximo dando a todas las características de los casos la misma importancia (Figura 4.21).

180 ejemplos almacenados en la base de casos no son suficientes para un diagnóstico correcto o clasificación usando una recuperación con inducción. La experiencia nos muestra que con más casos por clase (glaucoma, normal, cataratas, retinopatía diabética, retinopatía hipertensiva) podríamos obtener un porcentaje más alto de éxito. De hecho, tenemos 96 pacientes con glaucoma y el diagnóstico de los casos de prueba es bastante bueno, no obstante, el número de casos de los otros grupos es menor de 37. Con esa muestra tan pequeña la clasificación correcta es improbable.

El solapamiento de clases es frecuente en clínica y no puede ser resuelto por sistemas CBR. Ésa podría ser la razón por la que hayamos obtenido porcentajes bajos de acierto en la interpretación de casos con diagnóstico normal, cataratas, retinopatía diabética y retinopatía hipertensiva.

Sin embargo, se ha logrado mejorar los resultados disminuyendo el número de categorías de clasificación (Figura 4.19). Los mejores resultados, como ya se ha indicado, se han obtenido al clasificar en dos categorías: glaucoma y no glaucoma.

4.6 Comparación de los experimentos SOM y CBR

El objetivo fundamental de la realización de estos experimentos mediante diversas técnicas es lograr obtener un buen clasificador que evalúe de forma automática el campo visual, de manera que se pueda integrar en el sistema de diagnóstico global del glaucoma que utiliza además otros parámetros para obtener el diagnóstico del paciente.

Los mejores resultados para clasificar en dos, tres y cinco categorías se han logrado utilizando mapas autoorganizados de Kohonen. Por ello en el capítulo siguiente se utilizará el mejor mapa SOM y se integrará en el Sistema Experto global.

Sin embargo, se ha observado una respuesta similar ante algunas características de los experimentos. Tanto en los experimentos SOM como en los CBR, los mejores resultados se han obtenido utilizando la reducción de parámetros de entrada, es decir, tomando las medias de los defectos de las siete zonas del campo visual, en lugar de los defectos de los 59 puntos que se obtienen directamente del perímetro Octopus.

En los dos tipos de experimentos -SOM y CBR- hemos logrado mejorar los resultados de la clasificación a medida que hemos disminuido el número de clases que queríamos diferenciar. Lógicamente, como se han usado los mismos conjuntos de ejemplos en ambos, y distribuidos por clases de la misma forma (Tabla 4.2), las dificultades para diferenciar las categorías con pocos ejemplos, cataratas, retinopatía diabética y retinopatía hipertensiva, son las mismas utilizando SOM que CBR.

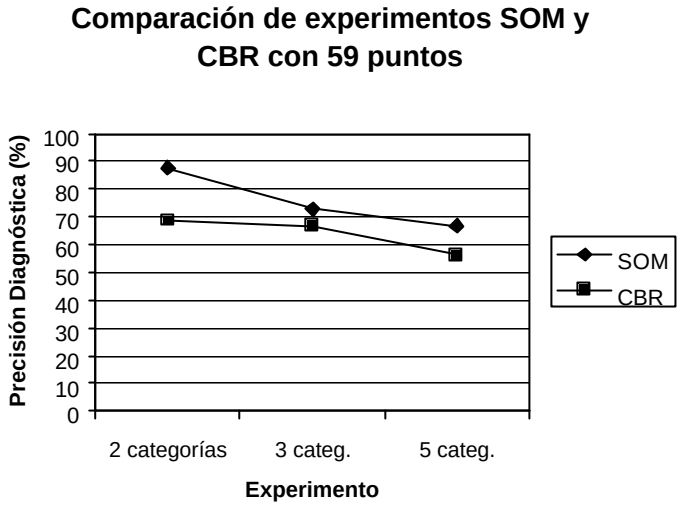


Figura 4.22: Comparación de los mejores experimentos SOM y CBR, realizados con los 59 puntos del campo visual mediante la Precisión Diagnóstica.

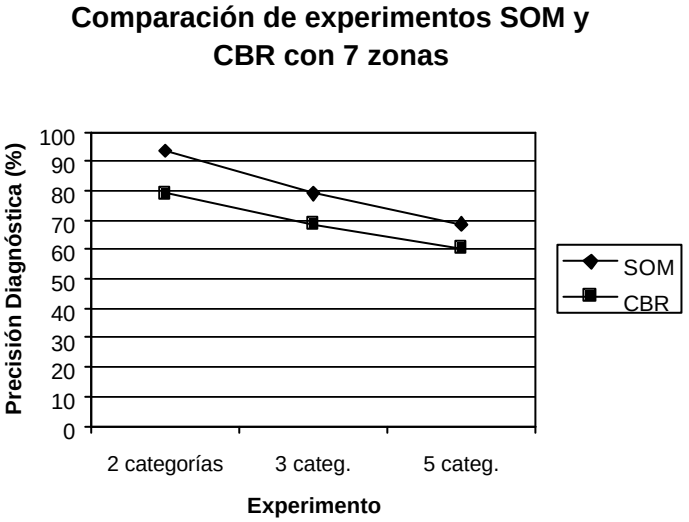


Figura 4.23: Comparación de los mejores experimentos SOM y CBR, realizados con las medias de los defectos de las siete zonas del campo visual mediante la Precisión Diagnóstica.

4.7 Comparación de los resultados obtenidos en la clasificación de campos visuales con los de otros trabajos de investigación

En el capítulo 2 de esta Tesis se expusieron los trabajos de investigación que se consideraron más significativos en la línea del diagnóstico del campo visual en el glaucoma. Llegados a este punto de nuestra investigación pensamos que es interesante resumir los resultados más relevantes para poder tener una perspectiva más completa. Para mostrar este resumen nos basamos, además de en los artículos referenciados en el capítulo 2, en el trabajo desarrollado por Reyes en su Tesis Doctoral [128].

Técnica	**Grupos**	**Grado de Glaucoma**	**S (%)**	**E (%)**
Octosmart (Hirsbrunner, Bebie [14, 76, 77])	Glaucoma vs. Normal	Cualquiera	?	?
Análisis Discriminante (Zeyen [177])	Glaucoma vs. Normal	MD < 5dB CLV 4-35 dB	80-97	80-99
Índice de Patrón de Glaucoma (GPI) (Mandava [102])	Glaucoma vs. Normal	Defectos focales	88	88
Comparación Hemicampos (Sommer [145])	Glaucoma vs. Normal	Escotomas paracentrales o arciformes	96	86
Test del Hemicampo de Glaucoma (Åsman [10])	Glaucoma vs. Normal	No avanzado	75-86	80-93
Análisis de clusters arciformes (Åsman [11])	Glaucoma vs. Normal	No avanzado	83-92	80-93
Redes neuronales MLP con BP (Goldbaum [56])	Glaucoma vs. Normal	Cualquiera	65	72
Redes neuronales MLP con BP (red jerárquica) (Zahlmann [171, 172])	Glaucoma vs. Normal vs. Otras Patologías	?	80-89	?

Redes neuronales RBF (red jerárquica) (Zahlmann [173, 176])	Glaucoma vs. Normal vs. Otras Patologías	?	72-98	70-95
Discriminación logística (7 zonas, MD y LV) (Antón [6])	Glaucoma vs. Normal vs. Otras patologías	MD ≤ 6dB menos de 10 puntos con defecto	80	92
Redes neuronales MLP con BP (7 zonas, red jerárquica) (Reyes [128])	Glaucoma vs. No glaucoma	MD ≤ 6dB menos de 10 puntos con defecto	100	82
Redes de Kohonen (SOM) (7 zonas)	Glaucoma vs. No glaucoma	MD ≤ 6dB menos de 10 puntos con defecto	95	93
CBR (7 zonas)	Glaucoma vs. No Glaucoma	MD ≤ 6dB menos de 10 puntos con defecto	80	78

Tabla 4.30: Resumen de los resultados de algunos trabajos, incluidos los de esta Tesis, sobre clasificación de campos visuales.

Observando la tabla anterior podemos decir que los resultados obtenidos en este trabajo de investigación son bastante significativos respecto a algunos de los anteriores trabajos, sobre todo los que se han conseguido utilizando mapas autoorganizados de Kohonen. Hemos logrado clasificar en dos categorías, glaucoma y no glaucoma, con sensibilidad y especificidad de 95 y 93% respectivamente. Estos valores son similares a los obtenidos en la Tesis Doctoral de Reyes [128] con los mismos datos de campos visuales que en esta Tesis.

Es importante destacar que en este trabajo se han utilizado campos visuales glaucomatosos con **defectos muy incipientes** para poder realizar un diagnóstico precoz, al igual que hicieron Reyes y Antón anteriormente con los mismos datos. Además, aunque en los experimentos de esta Tesis en que se diferencia en dos clases se han obtenido resultados similares a otros anteriores utilizando otras técnicas, las clases que se han diferenciado aquí han sido Glaucoma y No glaucoma, es decir, entre los campos que no son de pacientes glaucomatosos hay algunos pertenecientes a pacientes sin patología y otros con patologías distintas de glaucoma. Este hecho dificulta la separación de clases mucho más que si sólo hubiera campos con glaucoma y normales.

Podemos decir, por tanto, que las redes neuronales artificiales y en particular los mapas autoorganizados de Kohonen son una buena herramienta en la ayuda al diagnóstico precoz del glaucoma.

5 Sistema híbrido de ayuda al diagnóstico del glaucoma

Este capítulo, como ya se ha indicado, presenta el sistema de diagnóstico final que se ha implementado utilizando lógica borrosa y en el que se reúnen los dos enfoques de la IA: *simbólico*, al implementar un sistema basado en reglas de producción, y *conexionista*, al utilizar el mejor método de clasificación del CV mediante redes neuronales artificiales obtenido en nuestros experimentos mostrados en el capítulo anterior, e integrarlo en el sistema.

El objetivo de este trabajo es conseguir un sistema que, integrando todas las exploraciones utilizadas habitualmente para diagnosticar el Glaucoma Crónico Simple, logre mejores resultados en la clasificación de los pacientes que utilizando únicamente los resultados del campo visual. En el capítulo 2 se mostraron algunos sistemas que se basan sólo en los datos del campo visual, por ser una de las exploraciones fundamentales. Sin embargo, en ese mismo capítulo se recogieron algunas referencias de autores que indican la necesidad de utilizar también los resultados de las demás exploraciones.

Para abordar este estudio, primeramente vamos a exponer los fundamentos sobre lógica borrosa y su aplicación al desarrollo de sistemas expertos.

5.1 Introducción. Lógica Borrosa

Los problemas del mundo real se caracterizan por la necesidad de procesar información incompleta, imprecisa, vaga o con incertidumbre. La lógica clásica es incapaz de plasmar ese conocimiento en reglas de inferencia, por muy complejas que éstas sean. En esa lógica convencional se reduce todo ese conocimiento a un simple sí o no, es blanco o negro, es verdadero o falso, tiene la presión intraocular alta o baja, etc. Sin embargo, el hombre se mueve en un mundo de conocimiento que fluctúa habitualmente entre gradaciones. Las cosas no suelen ser blancas o negras sino que hay una escala de grises. Por ejemplo, cómo es posible que una

persona sea alta si mide 1,75 m. y otra no lo es si mide 1,74 m, si sólo se diferencian 1 cm en su estatura. Era necesario establecer una teoría que fuera capaz de reflejar el modo de pensar, razonar y expresarse del hombre para poder resolver problemas reales. Poder utilizar términos como alta, casi alta, muy alta, etc., así como poder definir acciones del tipo: "si cae *mucha* agua, cerrar *un poco* el grifo", era algo importante para afrontar problemas al estilo del hombre.

Después de la lógica clásica surgieron variaciones de ésta, como es la lógica multivaluada. Más tarde, en 1965, Lotfi Zadeh, profesor de la Universidad de California en Berkeley, publicó un artículo donde enunció las bases teóricas de la *lógica borrosa* o *fuzzy logic* [140, 170].

La teoría de conjuntos borrosos fue la base formal para desarrollar los sistemas de inferencia borrosos, que permiten utilizar fácilmente el conocimiento de los expertos en un determinado tema al formalizar éste, normalmente ambiguo, de una forma realizable.

5.1.1 Conjuntos borrosos

Como ya hemos señalado, en los conjuntos clásicos algo está incluido completamente en él o no lo está en absoluto. Esto puede describirse asignando un 1 a todos los elementos incluidos en el conjunto, y un 0 a los no incluidos. La teoría de conjuntos borrosos parte de la teoría clásica de conjuntos, añadiendo una función de pertenencia al conjunto definida como un número real entre 0 y 1, es decir, se incluyen todos los valores intermedios. (Se puede decir que la teoría clásica de conjuntos es un caso particular de la borrosa). Así, se introduce el concepto de **conjunto o subconjunto borroso** asociado a una determinada variable lingüística, definido por una palabra, adjetivo, valor o **etiqueta lingüística** A. Se llama **variable lingüística** a aquella variable que se va a calificar de forma borrosa y es capaz de tomar como valores palabras del lenguaje natural. Para cada conjunto o subconjunto borroso se define una **función de pertenencia** $m_A(t)$, que indica el grado en que la variable t está incluida en el concepto representado por la etiqueta A [111]. Por ejemplo, para la variable lingüística *estatura de una persona*, se pueden definir tres subconjuntos borrosos, cada uno identificado por una etiqueta, {Bajo, Medio, Alto}, y con una función de pertenencia $\{m_{Bajo}(t), m_{Medio}(t), m_{Alto}(t)\}$.

Sea U un conjunto de objetos, por ejemplo, $U=\Re^n$, que se denominará **universo de discurso**. En términos matemáticos, un **conjunto borroso** F en U queda caracterizado por una función de pertenencia m_F que toma valores en el rango [0,1], es decir,

$$m_F : U \rightarrow [0,1]$$

donde $m_F(u)$ representa el grado en el que $u \in U$ pertenece al conjunto borroso F.

Dado un cierto conjunto borroso F, se definen los siguientes términos (Figura 5.1):

- El **conjunto soporte** es el conjunto (clásico) de todos los valores de U para los que $m_F(u)>0$.
- Los **puntos de cruce** son aquellos valores para los que $m_F(u)=0.5$.
- Un conjunto borroso es de tipo **singleton** si su conjunto soportado es de un solo valor.
- El **conjunto** a **-corte** $F\alpha$ de un conjunto borroso F es el conjunto clásico de todos los puntos de u de U para los que se cumple $m_F(u)>\alpha$.

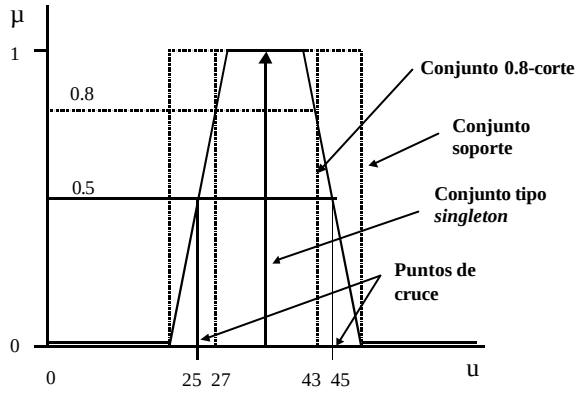


Figura 5.1: Términos relativos a los conjuntos borrosos [111].

5.1.2 Funciones de pertenencia de conjuntos borrosos

La **función de pertenencia** de un conjunto borroso consiste en un conjunto de pares ordenados $F=\{(u, m_F(u)) / u \in U\}$ si la variable es discreta, o una función continua si no lo es. Para la definición de estas funciones de pertenencia se utilizan convencionalmente ciertas familias de formas estándar, por coincidir con el significado lingüístico de las etiquetas más utilizadas. Las más frecuentes son de tipo trapezoidal, singleton, triangular, S, exponencial y tipo π.

Función trapezoidal:

Se define por cuatro puntos a, b, c y d, tal que a > b > c > d. Toma valor 0 para los valores menores que a y para los mayores que d, valor 1 para los valores comprendidos entre b y c, y valores intermedios para los restantes puntos. Se utiliza en sistemas borrosos sencillos, ya que permite definir un conjunto borroso con pocos datos y calcular fácilmente la función de pertenencia de cada punto. Se emplea especialmente en sistemas basados en microprocesador, ya que su formato es muy similar al de las funciones S, π, triangular y singleton. Formalmente, se define así:

$$S(u;a,b,c,d)=\begin{cases}0 & u<a\\ \left(\dfrac{u-a}{b-a}\right) & a\le u\le b\\ 1 & b\le u\le c\\ \left(\dfrac{d-u}{d-c}\right) & c\le u\le d\\ 0 & u>d\end{cases}$$

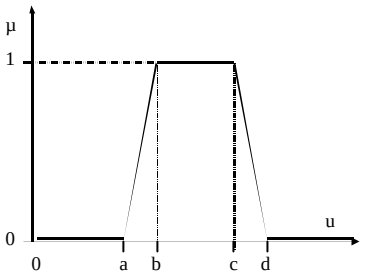


Figura 5.2: Función trapezoidal.

Función *singleton*:

Toma valor 1 para uno y solo uno de los puntos del universo de discurso, y 0 para el resto. Es utilizado con frecuencia en sistemas borrosos sencillos, para definir los conjuntos borrosos de las particiones de las variables de salida, pues permite simplificar los cálculos y requiere menos memoria para almacenar la base de reglas.

$$S(u;a)=\begin{cases}1 & u=a\\ 0 & u\neq a\end{cases}$$

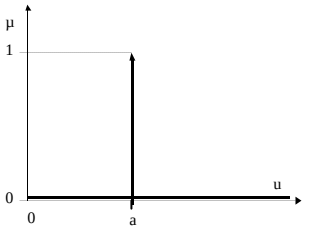


Figura 5.3: Función de tipo singleton.

Función de tipo T o triangular:

Puede verse como un caso especial de la función trapezoidal, en la que b=c, y por tanto toma valor 1 únicamente en un punto. Esta función es adecuada para modelar propiedades con un valor de pertenencia distinto de cero para un rango de valores estrecho en torno a un punto b.

$$T(u;a,b,d)=\begin{cases}0 & u<a\\ \left(\dfrac{u-a}{b-a}\right) & a\le u\le b\\ \left(\dfrac{d-u}{d-b}\right) & b\le u\le d\\ 0 & u>d\end{cases}$$

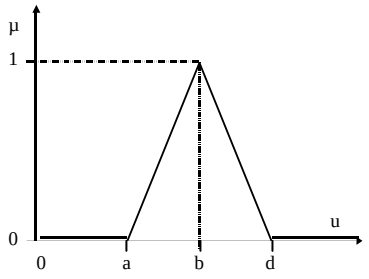


Figura 5.4: Función triangular.

Función tipo S:

Es una función adecuada para modelar propiedades como *grande, mucho, positivo*, etc. Toma valor 0 para los puntos menores que un cierto valor, a, valor 1 para valores mayores que otro valor, c, y valores intermedios para los puntos comprendidos entre ambos valores, de forma que el punto de cruce (valor 0.5)

será b=(a+c)/2. Para definir la variación entre dichos puntos, se han utilizado varias familias de funciones, como la cuadrática y la exponencial:

Usando una función cuadrática:

$$S(u;a,b,c)=\begin{cases}0 & u<a\\ 2\left(\dfrac{u-a}{c-a}\right)^2 & a\le u\le b\\ 1-2\left(\dfrac{u-c}{c-a}\right)^2 & b\le u\le c\\ 1 & u>c\end{cases}$$

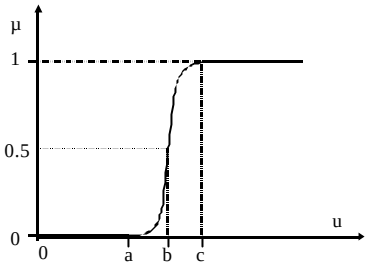


Figura 5.5: Función de tipo S.

Usando una función exponencial:

$$S(u;k,c)=\frac{1}{1+\exp(-k(u-b))}$$

Funcion tipo p **:**

Es la típica funcion gaussiana, y es la más adecuada para los conjuntos definidos en torno a un valor c, como *medio, normal,* etc. Se define a partir de la función S, utilizando para ello cualquiera de las dos familias de funciones citadas, si bien la más habitual es la exponencial:

$$p(u;b,c)=\begin{cases}S(u;c-b,c-b/2,c) & u\le c\\ 1-S(u;c-b,c-b/2,c) & u\ge c\end{cases}$$

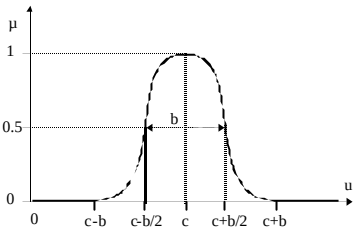


Figura 5.6: Función de tipo p**.**

5.1.3 Particiones borrosas

Dada una variable lingüística A, definida en un rango entre u_1 y u_2, es posible establecer en ella diversas particiones. Se llama **partición** a un grupo de los conjuntos borrosos que se han definido para la variable A. Por ejemplo, para la variable lingüística *estatura*, una posible partición sería la de la Figura 5.7, con tres subconjuntos borrosos, cada uno identificado por una etiqueta, {Bajo, Medio, Alto}, y con una función de pertenencia $\{m_{Bajo}(t), m_{Medio}(t), m_{Alto}(t)\}$. Se dice que una partición es **completa** si para todos los valores posibles de U, existe en la partición un conjunto con pertenencia no nula, es decir, los conjuntos definidos

cubren todo U. Se dice que dos conjuntos borrosos están **solapados** si su intersección es no nula [111].

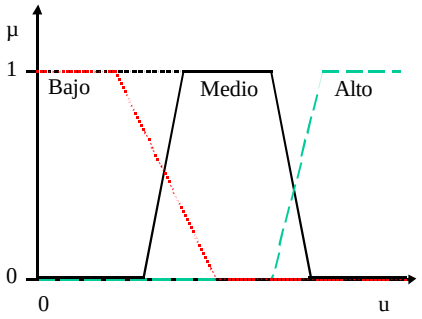


Figura 5.7: Ejemplo de conjuntos borrosos para la variable Estatura.

Normalmente se recomienda que las particiones sean completas, y que se solapen [111]. El solapamiento entre las distintas funciones de pertenencia de los conjuntos borrosos aumenta la robustez del sistema [140].

5.1.4 Operaciones borrosas

A los conjuntos borrosos se les puede aplicar determinados operadores, o bien pueden realizarse operaciones entre ellos. Al aplicar un operador sobre un solo conjunto borroso se obtiene otro conjunto borroso; de la misma manera, al combinar dos o más subconjuntos mediante alguna operación se obtendrá otro conjunto.

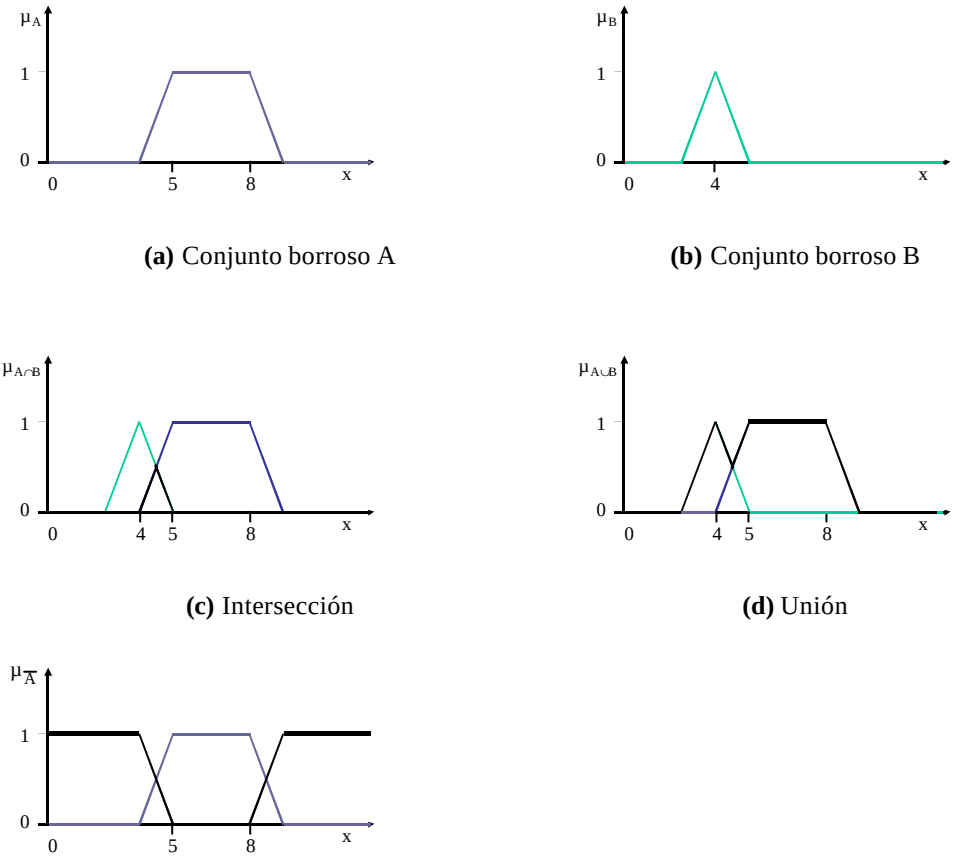
Sean los subconjuntos borrosos identificados por las etiquetas A y B, asociados a una variable lingüística x, podemos definir tres operaciones básicas: intersección, unión y complemento.

- **Unión** $m_{A\cup B}(x) = max\left[m_A(x), m_B(x)\right] \quad \forall x \in U$
- **Intersección** $m_{A\cap B}(x) = min\left[m_A(x), m(x)_B\right] \quad \forall x \in U$
- **Complemento** $m_{\overline{A}}(x) = 1 - m_A(x) \quad x \in U$

El funcionamiento de estas operaciones básicas coincide con el de las correspondientes a las de la teoría clásica de conjuntos. Las funciones correspondientes a la definición de la unión y la intersección pueden generalizarse, a condición de cumplir ciertas restricciones. Las funciones que cumplen estas condiciones se conocen respectivamente como **Conorma Triangular (T-Conorma)** y **Norma Triangular (T-Norma)**. Algunas de las más usadas son las siguientes [111]:

Conormas	**Normas**
$MAX(a,b)$	$MIN(a,b)$
$(a+b-ab)$	(ab)
$a \dot{+} b = MIN(1, a+b)$	$a * b = MAX(0, a+b-1)$

Ejemplo: Sean los conjuntos borrosos A y B, la Figura 5.8 muestra esos conjuntos borrosos, su intersección, unión y complemento de A.



(a) Conjunto borroso A

(b) Conjunto borroso B

(c) Intersección

(d) Unión

(e) Complemento

Figura 5.8: Ejemplo de aplicación de operaciones básicas sobre conjuntos borrosos.

Se pueden definir además otras operaciones:

- **Norma** $m_{Norma(A)}(x) = \dfrac{m_A(x)}{max[m_A(x)]} \qquad x \in U$
- **Concentración** $m_{Conc(A)}(x) = (m_A(x))^2 \qquad x \in U$
- **Dilatación** $m_{Dilat(A)}(x) = (m_A(x))^{0.5} \qquad x \in U$

5.1.5 Inferencia borrosa

Al igual que la lógica clásica, la borrosa también se ocupa del razonamiento formal con proposiciones, pero en ésta las proposiciones pueden tomar valores intermedios entre verdadero y falso. Se puede establecer un isomorfismo entre la lógica y la teoría de conjuntos borrosos, y de éstos, a su vez, con un álgebra de Boole. El objeto de la lógica borrosa es proporcionar un soporte formal al razonamiento basado en el lenguaje natural, que se caracteriza por ser de tipo aproximado, y que usa proposiciones que expresan información de carácter impreciso.

5.1.5.1 Principio de extensión

El principio de extensión [81, 111] permite convertir conceptos no borrosos en borrosos, y es además la base de la inferencia en los sistemas borrosos. Sean U y V dos universos de discurso, y f una función de U a V. En general, para un conjunto borroso A en U **el principio de extensión** define un conjunto borroso B en V dado por

$$m_B(v) = \max_{u \in f^{-1}(v)} \left[m_A(u) \right]$$

es decir, $m_B(v)$ es el máximo de $m_A(u)$ para todos los $u \in$ U que cumplen que $f(u)=v$, donde $v \in$ V y suponemos que $f^{-1}(v)$ no es vacío. Si $f^{-1}(v)$ es vacío para algún $v \in$ V, definiremos $m_B(v)=0$.

5.1.5.2 Relación borrosa

Sean U y V dos universos de discurso, una **relación binaria borrosa** se define como un conjunto borroso R en el espacio UxV, cuya función de pertenencia se denota como $m_R(u,v)$, con $u \in$ U y $v \in$ V.

Puede definirse también la **composición *Max-Star*** $R \circ S$ para dos relaciones borrosas R y S en UxV y VxW, respectivamente, como otra relación borrosa con la siguiente función de pertenencia

$$m_{R \circ S}(u, w) = \max_{v \in V} \left[m_R(u, v) * m_S(v, w) \right]$$

donde $u \in$ U, $v \in$ V, $w \in$ W, y el operador $*$ puede indicar la operación intersección definida anteriormente (entonces tendríamos la relación borrosa **max-min**) u otras definiciones (T-normas [81, 111]). Por tanto, la composición *max-star* de dos relaciones borrosas $R \circ S$ es un conjunto borroso en UxW, y uno de los principios básicos para el desarrollo de un sistema de inferencia borrosa.

5.1.5.3 Modus ponens generalizado y modus tolens generalizado

Las reglas borrosas son básicamente del tipo IF-THEN y expresan una relación borrosa o proposición borrosa. En lógica borrosa, el razonamiento es

aproximado, por tanto, se puede inferir de una regla una conclusión aunque el antecedente (premisa) no se cumpla plenamente. Existen dos métodos básicos de inferencia entre reglas o **leyes de inferencia**, el **modus ponens generalizado (GMP)** y el **modus tolens generalizado (GMT)**, que son generalizaciones del razonamiento clásico. Al GMP se le llama razonamiento directo y puede resumirse de la siguiente forma:

Conocimiento: Si x es A entonces y es B
Hecho: x es A'

Consecuencia: y es B'

donde A, A', B y B' son conjuntos borrosos. Esta relación se expresa también como B'=A'◦R.

El GMT se conoce como razonamiento inverso y puede resumirse como sigue:

Conocimiento: Si x es A entonces y es B
Hecho: y es B'

Consecuencia: x es A'

Lo que se expresa como A'=B'◦R.

5.1.5.4 Implicación borrosa

Si se definen dos conjuntos borrosos A y B en U y V, respectivamente, una implicación borrosa de A en B, que se indica con A $\rightarrow$ B, es una relación borrosa en UxV, que puede interpretarse de dos formas: A **asociado con** B, o A **implica** B [81].

Si adoptamos la primera interpretación, resultan varias relaciones diferentes dependiendo de la T-norma empleada. Las más típicas son las propuestas por Mamdani y Larsen:

- $m_{A\rightarrow B}(u,v) = \min\left[m_A(u), m_B(v)\right]; \quad \forall u \in U, \forall v \in V$ (Mamdani)

- $m_{A\rightarrow B}(u,v) = m_A(u)\, m_B(v); \quad \forall u \in U, \forall v \in V$ (Larsen)

Cuando adoptamos la segunda interpretación como significado de A $\rightarrow$ B, tenemos otras cuantas funciones de implicación borrosas. Un ejemplo de ellas es una de las propuestas por Zadeh:

- $m_{A\rightarrow B}(u,v) = \max(\min(\ m_A(u), m_B(v)), 1 - m_A(u)); \quad \forall u \in U, \forall v \in V$

Un ejemplo de interpretación gráfica de GMP usando la implicación borrosa de Mamdani y la composición max-min puede verse en la Figura 5.9.

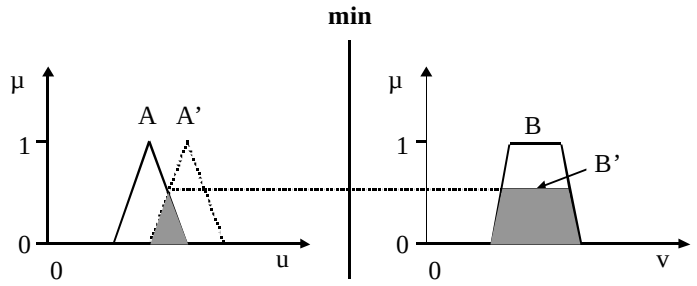


Figura 5.9: Interpretación gráfica del GMP usando la implicación borrosa de Mamdani y la composición max-min.

5.1.6 Reglas borrosas

Las reglas borrosas permiten expresar el conocimiento del que se dispone sobre la relación entre uno o varios conjuntos borrosos (antecedentes) y otro de ellos (consecuente). Para expresar este conocimiento de forma completa normalmente se precisa de varias reglas, que se agrupan formando lo que se conoce como **base de reglas**, es decir, el conjunto de reglas que expresan las relaciones conocidas entre antecedentes y consecuentes.

Formalmente, una base de reglas borrosa es una colección de reglas de implicación $R^{(l)}$, que puede tener los siguientes formatos:

- **Tipo Mamdani,** o borroso puro:
  $R^{(l)}$ *: IF* x_1 *is* F_1^l *and ... and* x_n *is* F^l_n *THEN y is* G^l

- **Tipo Sugeno**, donde la salida es una función de las variables de entrada:
  $R^{(l)}$ *: IF* x_1 *is* F_1^l *and ... and* x_n *is* F^l_n *THEN* $y^l = f^l(\mathbf{x})$

donde F_i^l y G_i^l son conjuntos borrosos en $U_i \subset \Re$ y $V \subset \Re$, respectivamente, y $\mathbf{x}=(x_1,...,x_n)^T \in U_1 * ... * U_n$ e $y \in V$ son variables lingüísticas.

A los sistemas borrosos descritos con una sola salida, se les conoce como **MISO** *(Multiple Input Single Output)*, mientras que a los que tienen varias salidas son conocidos como **MIMO** *(Multiple Input Multiple Output)*; éstos últimos pueden descomponerse en varios sistemas MISO.

5.1.7 Sistemas de inferencia borrosa

Los Sistemas de Inferencia Borrosa permiten el uso de conjuntos borrosos para extraer conclusiones y tomar decisiones. Esencialmente la ventaja de un enfoque desde la teoría de conjuntos borrosos es el poder describir información imprecisa, incompleta, o con incertidumbre. No obstante, eso sería poco práctico a menos que pudiéramos utilizarlo para realizar inferencia. El método aceptado de

la aplicación de conjuntos borrosos es análogo, aunque diferente, al modo en que se organizan los Sistemas Basados en Conocimiento convencionales.

Se llaman sistemas de inferencia borrosa a aquellos que interpretan las reglas de tipo IF-THEN de una base de reglas, para obtener los valores de salida a partir de los actuales valores de las variables lingüísticas de entrada al sistema.

La salida final de un dispositivo de inferencia borrosa puede ser:

- *M* conjuntos borrosos B^l, con *l*=1,2,...*M*, cada uno de los cuales es el resultado de aplicar una entrada A' a cada una de las *M* reglas de la base de reglas del dispositivo.
- Un único conjunto borroso B', que es la unión de los *M* conjuntos borrosos B^l del punto anterior.
- *M* escalares y^l, con *l*=1, 2,....*M*, si las reglas son de tipo Sugeno, cada uno de los cuales es el resultado de aplicar una entrada A' a cada una de las *M* reglas de la base de reglas.

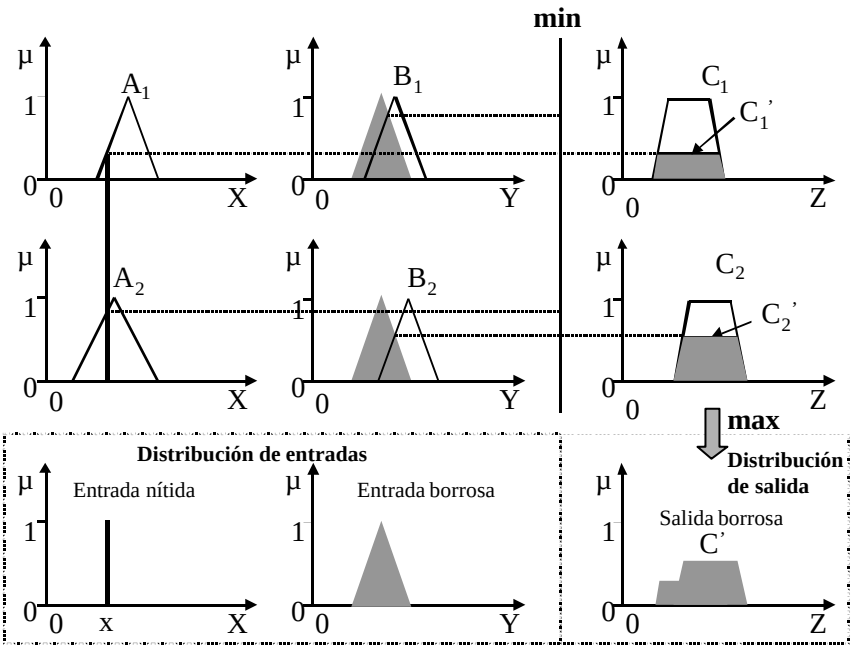


Figura 5.10: Un sistema de inferencia borrosa tipo Mamdani con dos entradas, una nítida y otra borrosa, dos reglas con dos antecedentes cada una y una salida borrosa, usando min y max como T-norma y T-conorma respectivamente.

Podemos decir que un Sistema de Inferencia Borrosa está constituido por:

- La base de conjuntos borrosos que se van a usar, con sus respectivas funciones de pertenencia.
- Las reglas que combinan los conjuntos borrosos.

- La composición borrosa de las reglas.
- El aclarado del conjunto borroso solución.

Cada componente presenta su complejidad y se han de tomar muchas decisiones para definirlas.

5.1.8 Determinación de las funciones de pertenencia

El primer problema que se ha de afrontar es cómo determinar las funciones de pertenencia [82].

Normalmente tendrán que determinarse mediante la adquisición del conocimiento a partir de un experto o un grupo de expertos. Una vez que los nombres de los conjuntos borrosos se han establecido, hay que considerar sus funciones de pertenencia asociadas.

El enfoque adoptado para determinar la forma de una particular función de pertenencia suele depender de la aplicación. Para la mayor parte de los problemas de control borroso se suele suponer que la función de pertenencia es lineal, normalmente con forma triangular. No obstante, para muchas otras aplicaciones esas funciones no son adecuadas y han de elicitarse directamente a partir del experto, mediante un enfoque manual de tipo "estadístico" o generándolas automáticamente.

5.1.8.1 Métodos manuales para determinar funciones de pertenencia

Hay varias técnicas estadísticas para determinar las funciones de pertenencia. Watanabe señala que se pueden clasificar en dos categorías amplias: el **uso de frecuencias** y la **estimación directa**. El método de frecuencias obtiene las funciones midiendo el porcentaje de gente de un grupo (normalmente expertos del dominio concreto) que responden afirmativamente a la pregunta sobre si un objeto pertenece a un conjunto particular. Los métodos de estimación directa se basan en pedir a los expertos que gradúen un evento según una escala. Watanabe llegó a la conclusión después de unos experimentos de que son preferibles los métodos de estimación directa.

Turksen cubre estos dos métodos considerando cuatro enfoques: estimación directa, votación, estadística de conjuntos valuados y estimación inversa. Para explicar estos enfoques vamos a hacer las siguientes definiciones:

- El conjunto de elementos $u \in U$.
- A, una variable lingüística que es una etiqueta para $u \in U$.
- B, un término lingüístico o conjunto borroso de una variable lingüística.
- Un intervalo de medida numérico $X \in [-\infty, \infty]$.
- $m_B(u)$, el valor de la función de pertenencia. Representa el grado de pertenencia de u al conjunto de elementos determinado por el término lingüístico B.

1. La **estimación directa** presenta $u \in U$ aleatoriamente seleccionados, con valores $A(u) \in X$ a personas que deben responder a la pregunta ¿Cómo de B es $u(A(u))$? Es decir, se le hace la pregunta al experto y él debe responder según una escala. Después, usando la opinión de los expertos sobre el rango de la variable A, se calcula $m_B(u)$.

2. El **método de votación** hace una pregunta diferente: "¿Estás de acuerdo con que u es B?" y la respuesta debe ser sí o no. La proporción de respuestas afirmativas respecto del total se usa para calcular la función de pertenencia.

3. La **estadística de conjuntos valuados** se basa en la idea de combinar conjuntos ordinarios usando un enfoque de frecuencia basada en la observación.

4. La **estimación inversa** usa un enfoque distinto, preguntando al experto "¿Está de acuerdo con que $u(A(u))$ tiene $m_B(u) = y$?"

Hay otros métodos para determinar funciones de pertenencia basados en varias técnicas de entrevistas para adquirir conocimiento. La desventaja que pueden tener estos métodos es que se basan en una interpretación subjetiva de los conceptos, además de todos los problemas que lleva consigo la adquisición del conocimiento [57].

5.1.8.2 Métodos automáticos para la determinación de funciones de pertenencia

La generación automática de funciones de pertenencia abarca una amplia variedad de enfoques. Esencialmente lo que diferencia estos métodos de los manuales es que el experto está ausente completamente del proceso o bien interviene inicialmente para establecer las formas de las funciones que luego se ajustarán mediante métodos automáticos.

Las técnicas que vamos a describir brevemente, aplicadas al problema que nos atañe, son:

- Redes Neuronales Artificiales
- Algoritmos Genéticos
- Prototipos Deformables
- Búsqueda del Gradiente
- Razonamiento Inductivo

Redes Neuronales Artificiales

Takagi y Hayashi señalan que el razonamiento borroso presenta varios problemas:

- La falta de un método definido para determinar funciones de pertenencia.
- La falta de una función de aprendizaje.

Para solucionar esos problemas describen un enfoque basado en RNA. Este método consiste en analizar las reglas if-then usando RNA para determinar las funciones de pertenencia del antecedente y luego determinar la componente del consecuente como la salida de cada regla. El enfoque usado consiste en tomar los datos, aplicar un algoritmo de clustering convencional para agrupar los datos en clusters y aplicar RNA a estos datos agrupados para determinar la función de pertenencia de un patrón dentro de unos conjuntos borrosos concretos.

Esta combinación de redes neuronales y razonamiento borroso no sólo permite generar funciones de pertenencia sino que aporta un enfoque potencialmente muy potente.

Como ya hemos descrito, utilizar el conocimiento de los expertos es el modo más común de determinar funciones de pertenencia. Wang de apoya en la experiencia aportada por un experto y luego usa RNA para refinar las funciones. Es decir, los pares (*X*, *y*) que describen la relación entre X e y se presentan a la red neuronal, la cual ajusta una función a los puntos. Utilizan una versión del algoritmo *back propagation* estándar para conseguir mejorar la interpolación entre los puntos.

Algoritmos Genéticos

Esta técnica adopta algunas de las ideas de la genética y representa los datos como cromosomas y genes, desarrollando operaciones tales como cruce y mutación.

Meredith aplicó algoritmos genéticos para refinar funciones de pertenencia en un controlador borroso para un helicóptero. Hace unas estimaciones iniciales de las funciones de pertenencia mediante ingeniería de control, y luego los algoritmos genéticos ajustan los parámetros que las definen. Utilizan funciones triangulares.

Otros autores como Lee y Takagi determinaron, utilizando algoritmos genéticos, el número óptimo de reglas así como las funciones de pertenencia, que fueron también triangulares. Obtuvieron buenos resultados, pero hay que resaltar que su sistema sólo tenía dos variables de entrada y una de salida.

Los algoritmos genéticos ofrecen un modo adecuado de modelar funciones de pertenencia, especialmente si son de tipo triangular.

Prototipos Deformables

Un enfoque relativamente reciente aportado por Bremermann es el de los prototipos deformables. Este método aconsejado inicialmente para reconocimiento

de patrones, parece ser potencialmente útil para determinación automática de funciones de pertenencia. Está basado en la idea de tomar un objeto que queremos identificar y deformarlo para hacerlo coincidir con un prototipo. La cantidad de coincidencia y distorsión se miden mediante sus funciones respectivas, que combinadas dan lugar a un *funcional cibernético*. Usando la notación de Bremermann el funcional de coincidencia sería

$$<M(\Phi_i(p_1, ..., p_n)), f>$$

donde f es el objeto que va a ser identificado, Φ_i es el prototipo i-ésimo y p_1, ..., p_n son los parámetros que controlan la distorsión del prototipo. Sea

$$D(\Phi_i(p_1, ..., p_n))$$

la función distorsión, entonces el funcional cibernético para el prototipo i-ésimo puede definirse por

$$< (f_i, f) >= Min_{p_i,\ldots,p_n} < F_i(p_1,\ldots,p_n), f >$$

donde

$$<F_i(p_1, ..., p_n), f> = <M(\Phi_i(p_1, ..., p_n)), f> + c\, D(\Phi_i(p_1, ..., p_n))$$

y c es una constante. Bremermann explica que este funcional cibernético puede usarse para generar una función de pertenencia

$$X(f) = 1 - \frac{< f, f >}{Max}$$

donde *Max* es el menor límite superior de f. Bremermann obtuvo buenos resultados en sus implementaciones en reconocimiento de patrones y en particular para conjuntos borrosos en interpretación de ECG.

Búsqueda del Gradiente

Basándose en el trabajo de Procyck y Mamdani, Burhardt y Bonissone usaron métodos de búsqueda del gradiente para refinar bases de conocimiento borrosas. Afinaron las reglas y también las funciones de pertenencia para una aplicación de control. Suponen primeramente, como con la mayor parte del control borroso, que las funciones de pertenencia son triangulares y usan la búsqueda del gradiente para determinar el factor de ajuste óptimo para la base de los triángulos.

Razonamiento Inductivo

El enfoque adoptado por Kim y Russell es utilizar razonamiento inductivo para generar las funciones de pertenencia y las reglas. Suponen que no tienen más información que un conjunto de datos. El método consiste en hacer particiones en

el conjunto de datos en clases minimizando la entropía. Ésta, considerando que solamente un resultado es verdadero, es el valor esperado de la información contenida en el conjunto de datos y viene dado por

$$S = k\sum_{i=1}^{N}\left[P_i \ln P_i + (1-P_i)\ln(1-P_i)\right]$$

donde la probabilidad de que el ejemplo i-ésimo sea verdadero es P_i y N es el número de ejemplos del conjunto de datos. En su trabajo muestran un ejemplo de dos clases donde, haciendo particiones en el espacio iterativamente, calculan una estimación de la entropía que les permite determinar las funciones de pertenencia triangulares. Este método tiene el inconveniente de que no hay modo de saber si estas funciones son realistas y además los conjuntos borrosos obtenidos son triangulares.

5.1.9 Emborronador

El **emborronador** establece una relación entre puntos de entrada no borrosos y sus correspondientes conjuntos borrosos. Es decir, emborrona los valores procedentes del exterior. Para esto se puede seguir varios métodos:

- **Emborronador singleton:** Es el método más utilizado, y consiste en considerar los propios valores discretos como conjuntos borrosos, es decir, se definiría un conjunto borroso A con la siguiente función de pertenencia:

$$m_A(x') = \begin{cases} 1 & x = x' \\ 0 & x \neq x' \end{cases}$$

- **Emborronador no singleton:** en este método se utiliza una función exponencial con forma de campana, centrada en el valor de entrada x, de anchura s y de amplitud a:

$$m_A(x') = a \cdot \exp\left[-\left(\frac{x'-x}{s}\right)^2\right]$$

5.1.10 Aclarador

Es una función que transforma un conjunto borroso (normalmente la salida de un dispositivo de inferencia borrosa) en un valor no borroso. Al igual que para emborronar, también existen varios métodos para aclarar:

- **Dispositivos de inferencia borrosa con reglas tipo Mamdani:**

  - **Aclarador por máximo:**

$$y = \arg\max_{y \in V}(m_{B'}(y))$$

- **Aclarador por media de centros**, donde $\bar{y}^l$ representa el punto en el que la función de pertenencia del conjunto G^l alcanza su máximo valor:

$$y = \frac{\sum_{l=1}^{n} \bar{y}^l (m_{B'}(\bar{y}^l))}{\sum_{l=1}^{n} (m_{B'}(\bar{y}^l))}$$

- **Aclarador por centro de área**, donde M^l es el momento de la función de pertenencia del conjunto G^l y A^l es el área:

$$y = \frac{\sum_{l=1}^{n} M^l (m_{B'}(\bar{y}^l))}{\sum_{l=1}^{n} A^l (m_{B'}(\bar{y}^l))} = \frac{\sum_{l=1}^{n} \int_V m_{B'}(\bar{y}^l)^2 dy \cdot M^l (m_{B'}(\bar{y}^l))}{\sum_{l=1}^{n} \int_V m_{B'}(\bar{y}^l) dy}$$

- **Dispositivos de inferencia borrosa con reglas tipo Sugeno:** la salida se obtiene como media ponderada de las salidas de cada regla de la base de reglas según la siguiente fórmula, donde y^l es la salida de la regla *l*, calculada, como ya vimos anteriormente, como función del vector de entradas:

$$y = \frac{\sum_{l=1}^{n} y^l (m_{A^l}(x))}{\sum_{l=1}^{n} (m_{A^l}(x))}$$

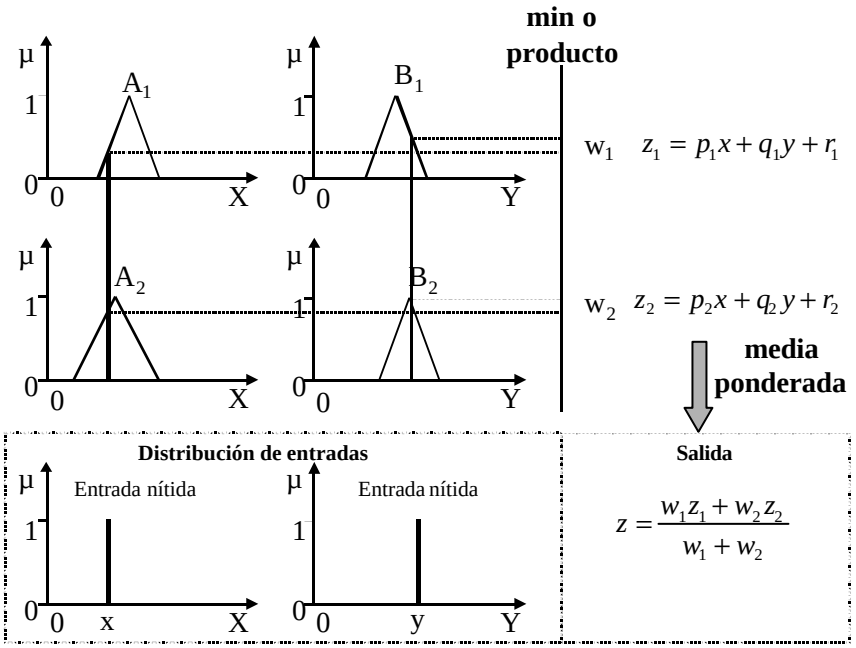


Figura 5.11: Un sistema de inferencia borrosa tipo Sugeno con dos entradas nítidas, y dos reglas con dos antecedentes cada una, usando min o producto como T-norma y calculando la media ponderada de la salida de cada regla para obtener la salida final.

5.2 Lenguaje de desarrollo: FuzzyClips

Para desarrollar este primer prototipo de Sistema Borroso se eligió la herramienta FuzzyCLIPS.

FuzzyCLIPS en una extensión de CLIPS, desarrollado también por la NASA bajo la misma filosofía y con la misma sintaxis, y surge como soporte para los sistemas expertos difusos. El sistema integra de forma ponderada los consecuentes para dar una respuesta global afectada por la confianza o nivel de disparo de cada una de las reglas [23]. FuzzyCLIPS permite trabajar con razonamiento exacto, borroso o con una mezcla de ambos, permitiendo mezclar en las reglas y hechos de un SE términos normales y borrosos. El sistema usa dos conceptos básicos inexactos: borrosidad e incertidumbre [123].

Para hacer llamadas al motor de inferencia de FuzzyCLIPS desde nuestra aplicación, desarrollada en Visual C++, se ha utilizado una biblioteca de vínculos dinámicos dll, desarrollada también por la NASA. Sin embargo, en la interfaz usada para comunicarse con dicha biblioteca, se utilizaron los archivos de cabecera e implementación de la interfaz estándar de CLIPS, y no una propia de FuzzyCLIPS. Esto es debido a la falta de documentación existente sobre el uso dicha biblioteca. Dado que FuzzyCLIPS no es más que una extensión del propio lenguaje CLIPS, esto no ha supuesto ningún problema, ya que la sintaxis es idéntica y sólo el funcionamiento interno del motor de inferencia varía de uno a otro.

5.2.1 Técnica de inferencia de FuzzyCLIPS

La evaluación de una regla depende de varios factores, como son si las variables, borrosas o no, forman parte del antecedente o consecuente de la regla, si ésta contiene múltiples antecedentes o consecuentes, si se añaden además factores de certeza, etc. Vamos a ver cómo resuelve todo ello FuzzyCLIPS [123].

5.2.1.1 Reglas simples

Son reglas del tipo:

si A entonces C	CF_r
A’	CF_h
---	---
C’	CF_c

donde A es el antecedente de la regla, A’ es el hecho de la base de hechos que confronta, C es el consecuente de la regla, C’ es el consecuente calculado, CF_r es el factor de certeza de la regla, CF_h es el factor de certeza del hecho y CF_c es el factor de certeza de la conclusión.

Existen tres tipos de reglas simples: CRISP_, FUZZY_CRISP, y FUZZY_FUZZY.

Regla simple CRISP_

El antecedente de la regla no contiene ningún objeto borroso independientemente de cómo sea el consecuente.

Para que este tipo de reglas se disparen A' debe ser igual a A. Sería una regla clásica de CLIPS. En ese caso la conclusión C' es igual a C, y el coeficiente de certeza es el resultado de multiplicar los coeficientes de certeza de la regla y el hecho

$$CF_c = CF_r * CF_h$$

Algunas de las reglas de nuestro sistema son de este tipo: tanto el antecedente como el consecuente son nítidos.

Regla simple FUZZY_CRISP

Sólo el antecedente contiene un objeto borroso.

A' debe ser un hecho borroso con la misma variable borrosa que la especificada en A para que puedan confrontar y la regla se pueda disparar. Los valores de las variables borrosas A y A' representados por los conjuntos borrosos F_α y F'_α no tienen que ser iguales pero sí deben solaparse.

En este tipo de reglas la conclusión C' es igual a C, y el coeficiente de certeza es igual a

$$CF_c = CF_r * CF_h * S$$

donde S es una medida de similitud entre los conjuntos borrosos F_α y F'_α. Esta medida S se basa en la medida de posibilidad P y la de necesidad N:

$S = P(F_a \mid F_a')$ si $N(F_a \mid F_a') > 0.5$

$S = (N(F_a \mid F_a') + 0.5) * P(F_a \mid F_a')$ en otro caso

donde

$$P(F_a \mid F_a') = \max(\min(\ m_{F_a}(u), m_{F_a'}(u))) ; \quad \forall u \in U$$

$$N(F_a \mid F_a') = 1 - P(\overline{F_a} \mid F_a'), \quad \overline{F_a} \text{ es el complemento de } F_a .$$

Por tanto, si la similitud entre los conjuntos borrosos del patrón A y del hecho que confronta A' es alta, el factor de certeza de la conclusión es muy cercano a $CF_r * CF_h$ ya que S será cercano a 1. Si la confrontación es pobre entonces el factor de certeza de la conclusión es más pequeño. Si los conjuntos borrosos no solapan, la similitud sería cero y el factor de certeza de la conclusión también, y, por tanto, la regla no se dispararía.

Este tipo de reglas de FuzzyCLIPS vamos a usarlas habitualmente en nuestro sistema. La mayor parte de las reglas contienen antecedentes borrosos y el consecuente es nítido.

Regla simple FUZZY_FUZZY

Tanto el antecedente como el consecuente contienen objetos borrosos.

En este tipo de reglas el antecedente y el consecuente están conectados mediante la siguiente relación borrosa

$$R = F_a * F_c$$

donde F_α es un conjunto borroso que representa el valor del antecedente borroso y F_c el del consecuente borroso.

La función de pertenencia de la relación R se calcula mediante la siguiente fórmula:

$$m_R(u,v) = \min(m_{F_a}(u), m_{F_C}(v)), \quad \forall (u,v) \in U \times V$$

El cálculo de la conclusión se basa en la regla de composición siguiente:

$$F_c' = F_a' \circ R$$

donde F_c' es un conjunto borroso que representa el valor del objeto borroso del consecuente. La función de pertenencia de F_c' se calcula como sigue:

$$m_{F_C'}(v) = \max_{u \in U}(\min(m_{F_a'}(u), m_R(u,v)))$$

que puede simplificarse a

$$m_{F_C'}(v) = \min(z, m_{F_c}(v))$$

donde

$$z = \max(\min(m_{F_a'}(u), m_{F_a}(u)))$$

El factor de certeza de la conclusión se calcula de la siguiente forma: $CF_c = CF_r * CF_h$.

Ninguna de las reglas de nuestra base de reglas del sistema borroso son de este tipo, ya que no hay ninguna con consecuente borroso.

5.2.1.2 Reglas complejas

Múltiples consecuentes

En CLIPS, el consecuente puede contener múltiples patrones (C_1, C_2, ..., C_n) con la conjunción **y** implícita entre ellos. Se les trata como múltiples reglas con un solo consecuente. Así la siguiente regla:

si Antecedentes entonces C_1 y C_2 y ... y C_n

es equivalente a las siguientes reglas:

si Antecedentes entonces C_1
si Antecedentes entonces C_2
...
si Antecedentes entonces C_n

Múltiples antecedentes

Si el consecuente no es un hecho borroso, no es necesario un tratamiento especial ya que la conclusión será el hecho nítido. Si el consecuente es un hecho borroso, el valor se calcula de la siguiente forma.

Si usamos **y** lógico, tenemos:

si A_1 **y** A_2 entonces C	CF_r
A_1'	CF_{h1}
A_2'	CF_{h2}

C'	CF_c

donde A_1' y A_2' son hechos (nítidos o borrosos) que confrontan con los antecedentes A_1 y A_2 respectivamente. En este caso el conjunto borroso que describa el valor de la declaración borrosa de la conclusión se calcula según la siguiente fórmula

$$F_c' = F_{c1}' \cap F_{c2}'$$

donde $\cap$ es la intersección de conjuntos borrosos que, en este software, viene dada por la siguiente expresión $m_c(x) = \min(m_A(x), m_B(x)), \forall x \in U$, F_{c1}' es el resultado de la inferencia borrosa para el hecho A_1' y la regla simple si A_1 entonces C, y F_{c2}' es el resultado de la inferencia borrosa para el hecho A_2' y la regla simple si A_2 entonces C.

Por tanto, si tanto A_1 como A_2 son hechos nítidos, la conclusión es el hecho borroso C, ya que serían como dos reglas simples CRISP_FUZZY. Si uno de los patrones es nítido, por ejemplo A_1, y el otro borroso, la conclusión es F_{c2}' ya que la regla simple CRISP_FUZZY concluiría C y la FUZZY_FUZZY concluiría F_{c2}'. La intersección de las conclusiones es justamente F_{c2}'.

El factor de certeza de la conclusión se calcula de acuerdo al modelo MYCIN

$$CF_c = \min(CF_{h1}', CF_{h2}') * CF_r$$

donde CF_{h1}' es el CF de la regla simple si A_1 entonces C dado el hecho que confronta A_1', y CF_{h2}' es el CF de la regla si A_2 entences C dado el hecho que confronta A_2'.

Todo lo anterior que hemos visto para dos antecedentes se puede generalizar a n antecedentes

$$F_c' = F_{c1}' \cap F_{c2}' \ldots \cap F_{cn}'$$
$$CF_c = \min(CF_{h1}', CF_{h2}', \ldots, CF_{hn}') * CF_r$$

5.2.1.3 Inserción de varios hechos

En CLIPS estándar, cuando insertamos un hecho y éste ya existe en la lista de hechos, puede o no ser insertado en dicha lista dependiendo de la opción que permite la duplicación de hechos. En FuzzyCLIPS la inserción de un hecho borroso que ya existe en la lista de hechos modifica su función de pertenencia. Por supuesto deben pertenecer a la misma variable. El nuevo conjunto se calcula de la siguiente manera:

$$F_g = F_h \cup F_c'$$

donde F_g es el nuevo conjunto borroso, F_h es el conjunto actual y F_c' es el hecho borroso que estamos insertando. El coeficiente de certeza del hecho modificado se calcula como:

$$CF_g = \max(CF_h, CF_c)$$

Un ejemplo de esto se puede ver en la Figura 5.10.

5.3 Diseño de las funciones de pertenencia

La especificación del Sistema Experto vendrá determinada por las variables lingüísticas de entrada y salida. En nuestro caso, al igual que en el Sistema Experto Glaucom-Easy (apartado 3.2), las variables de entrada son la presión intraocular (PIO), la relación excavación/papila (E/P) del nervio óptico, la asimetría de esta relación entre ambos ojos, el campo visual, los factores de riesgo (edad, antecedentes glaucomatosos, miopía elevada, diabetes, enfermedades cardiovasculares) y las características de la papila del nervio óptico (ovalización vertical de la excavación, excentricidad, alguna muesca en el borde, cambios bruscos en la trayectoria de los vasos, hemorragias, excavación con bordes netos). Los valores de estas variables son introducidos al sistema a excepción de la asimetría entre las E/P de ambos ojos y la edad que las calcula él mismo. La salida del sistema es también la misma que en Glaucom-Easy: un diagnóstico o consejo sobre la pauta a seguir (periodo de revisión, repetición de alguna exploración, conveniencia de instaurar un tratamiento o de acudir al especialista en glaucoma).



Figura 5.12: Entradas y salidas del Sistema Experto Borroso.

En la Figura 5.12 se pueden observar las variables de entrada y salida. Una de las diferencias fundamentales respecto al Sistema Experto Glaucom-Easy es que en éste, como ya se ha indicado anteriormente, el campo visual no se introduce al sistema como un valor consecuencia de una decisión humana por parte del experto, sino que es clasificado de forma automática.

El siguiente paso es definir los conjuntos borrosos asociados a cada variable lingüística. Esto puede hacerse utilizando métodos manuales o automáticos. Nosotros los hemos definido utilizando los manuales, basándonos en la experiencia de los expertos.

La primera decisión que tomamos ha sido el número de conjuntos borrosos asociados a cada variable lingüística. Nos apoyamos para ello en el conocimiento experto adquirido en las primeras versiones de sistemas expertos desarrolladas, explicadas en el capítulo 3, y en las decisiones implementadas mediante reglas de producción en esos sistemas, especialmente en el Glaucom-Easy. Las variables lingüísticas y los términos elegidos para sus conjuntos borrosos se muestran en la Tabla 5.1.

PIO	**E/P**	**Diferencia E/P**	**CV**
Normal	Normal	Normal	Normal
Dudosa	Dudosa	Patológica	Dudoso
Límite	Límite		Patológico
Más límite	Patológica		
Patológica			

Tabla 5.1: Conjuntos borrosos. Variables y términos lingüísticos.

Para definir las funciones de pertenencia de los conjuntos borrosos hemos usado los siguientes métodos:

- **Estimación directa.** El experto ha ido indicando por ejemplo cómo de normal, dudoso, límite o patológico es un determinado valor de presión intraocular según una escala ordinal, y así con las demás variables.
- **Estadística de conjuntos valuados.** Se han realizado estudios epidemiológicos sobre la prevalencia de determinados valores de E/P y

diferencias de E/P en la población de Castilla y León que nos han ayudado a fijar los valores de pertenencia [8, 9].

- **Estimación inversa.** Hemos elaborado unas funciones de pertenencia para los conjuntos borrosos y hemos preguntado al experto si estaba de acuerdo con los valores fijados.

Los estudios estadísticos de la frecuencia de la población de Castilla y León con unos determinados valores de E/P o de diferencias de E/P entre ambos ojos, han sido realizados en el IOBA (Instituto de Oftalmobiología Aplicada de la Universidad de Valladolid) [8, 9] y vienen recogidos en las figuras siguientes (Figuras 5.13, 5.14 y 5.15):

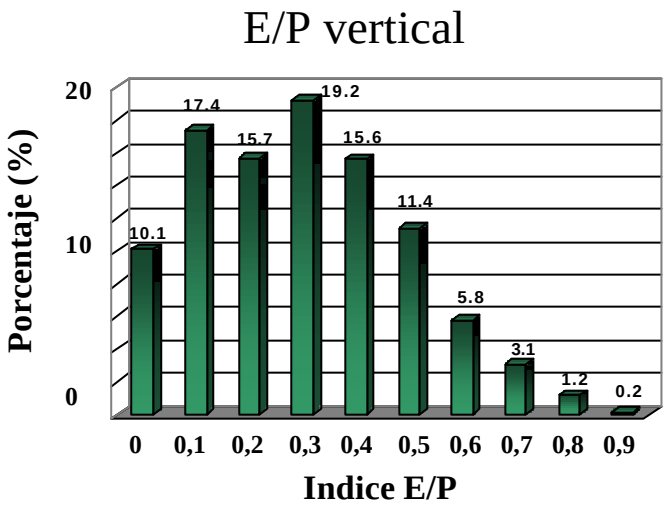


Figura 5.13: Frecuencia de los valores de la E/P en la población normal.

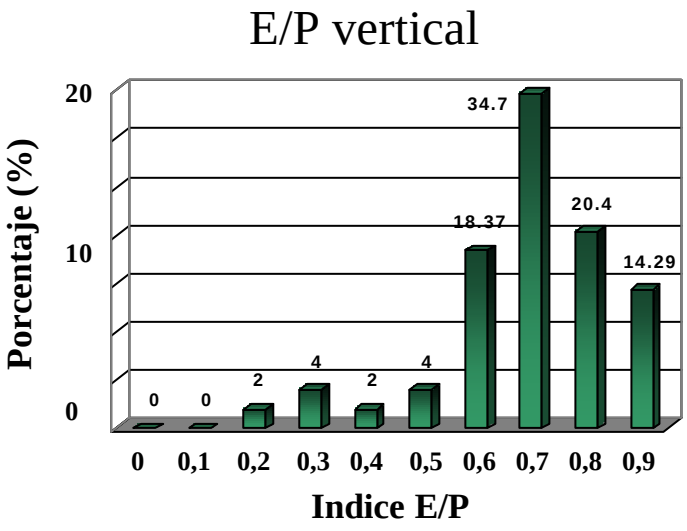


Figura 5.14: Frecuencia de los valores de la E/P en la población glaucomatosa.

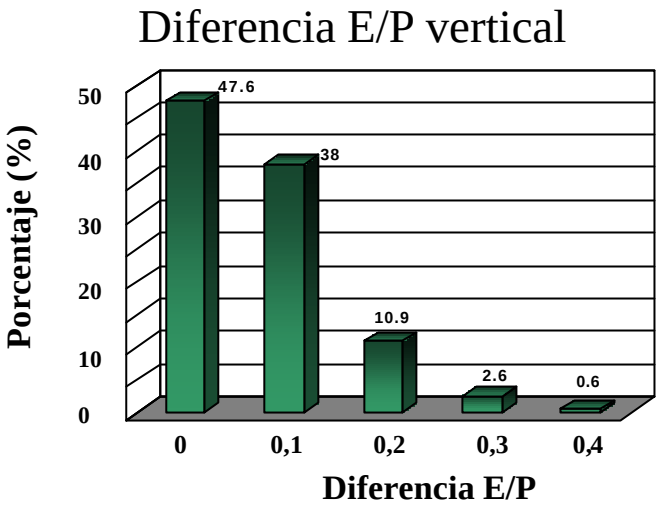


Figura 5.15: Frecuencia de los valores de la diferencia de la E/P en la población normal.

Las funciones de pertenencia definidas para el sistema borroso se muestran en la Tabla 5.2. La notación usada es la misma que se utilizó cuando se explicó al comienzo de este capítulo la teoría sobre la forma que podían adquirir estas funciones (apartado 5.1.2).

Podemos observar que las particiones borrosas diseñadas para las tres primeras variables lingüísticas (PIO, E/P y Diferencia E/P) son completas, es decir, para todos los posibles valores del universo de discurso existe en la partición un conjunto borroso con pertenencia no nula, y, además, existe solapamiento entre los distintos conjuntos borrosos. Todo esto da robustez al sistema [111, 140].

El campo visual está definido mediante tres conjuntos "s*ingleton*" debido a la dificultad que entraña la evaluación de éste. A diferencia de las anteriores magnitudes, no poseemos una medida fija del campo visual que nos indique cuándo éste es normal, dudoso o patológico, es decir, gráficamente no dispondríamos de ninguna escala que dibujar en el eje de abscisas.

La solución propuesta a este problema es evaluar el campo visual mediante un sistema de diagnóstico, que se explicará en siguientes apartados, y pasar a FuzzyCLIPS los valores "normal", "dudoso" o "patológico" en lugar de los valores numéricos de los que carecemos, utilizando conjuntos singleton como método para emborronarlos. Así pues, consideramos el campo visual como una variable discreta que puede tomar uno de tres valores, cada uno perteneciente a un solo conjunto borroso, siendo su grado de pertenencia a éste igual a uno, y a los otros dos conjuntos igual a cero. Por lo tanto, no hay solapamiento entre los conjuntos borrosos.

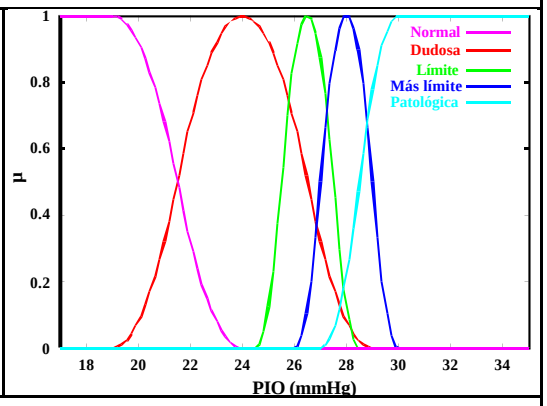
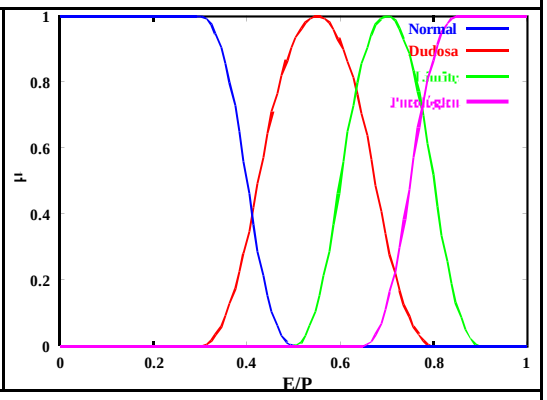
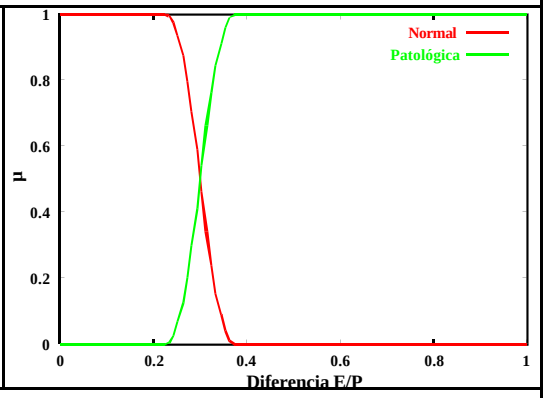
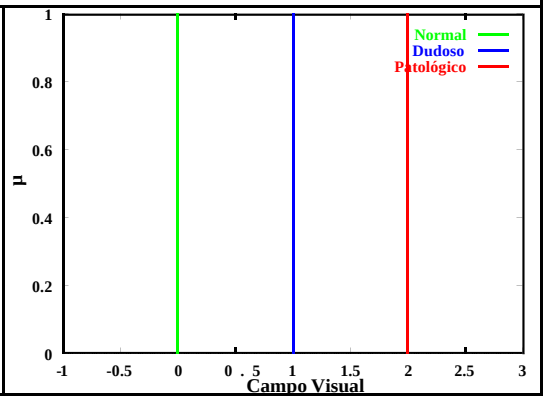
PIO		
• Normal • Dudosa • Límite • Más límite • Patológica	1 – S(x; 19, 21.5, 24) π(x; 5, 24) π(x; 2, 26.5) π(x; 2, 28) S(x; 27, 28.5, 30)	
E/P		
• Normal • Dudosa • Límite • Patológica	1 – S(x; 0.3, 0.4, 0.5) π(x; 0.25, 0.55) π(x; 0.20, 0.7) S(x; 0.65, 0.75, 0.85)	
Diferencia E/P		
• Normal • Patológica	1 – S(x; 0.225, 0.3, 0.375) S(x; 0.225, 0.30, 0.375)	
CV		
• Normal • Dudoso • Patológico	$\begin{cases}1 & x=0\\0 & x\neq 0\end{cases}$ $\begin{cases}1 & x=1\\0 & x\neq 1\end{cases}$ $\begin{cases}1 & x=2\\0 & x\neq 2\end{cases}$	

Tabla 5.2: Funciones de pertenencia diseñadas para las variables lingüísticas del sistema borroso.

Para el resto de los datos de entrada se definen también conjuntos singleton, pero tomando como base el eje de abscisas utilizado para definir sus correspondientes conjuntos borrosos. De esta manera, el solapamiento del conjunto singleton definido para un valor de entrada (numérico) con los diversos conjuntos borrosos que componen la partición de su correspondiente variable lingüística nos indicará claramente el grado de pertenencia a cada uno de dichos conjuntos.

5.4 Arquitectura del sistema híbrido de ayuda al diagnóstico del glaucoma

La arquitectura del sistema es de tipo *filtro-tubería* y básicamente consta de los módulos que aparecen en la Figura 5.16:

- **Sistema Experto Borroso.** Obtiene los datos a través de la interfaz de usuario, del clasificador del campo visual, y del gestor de históricos, y proporciona su salida a la interfaz. Tiene la arquitectura básica de un SE. Consta de dos partes claramente diferenciadas, que se complementan para obtener el diagnóstico:
  - Base de conocimiento, y motor de inferencia de FuzzyCLIPS. Se han introducido los hechos y reglas de la base según su sintaxis.
  - Base de conocimiento y motor de inferencia programados en Visual C++.

  Se han omitido algunos componentes opcionales (apartado 3.1.2).
- **Clasificador del Campo Visual.** De este módulo, como su nombre indica, depende la clasificación del campo visual. Toma su entrada comunicándose con la interfaz de usuario, al igual que el Sistema Experto, y su salida constituye una de las entradas de éste.
- **Interfaz de usuario.** Es el módulo que permite al SE comunicarse con el usuario. Se ha diseñado especialmente para que la herramienta pueda ser empleada por usuarios poco familiarizados con los ordenadores.
- **Ayuda.** El sistema proporciona dos tipos de ayuda a petición del usuario:
  - Ayuda general de la aplicación.
  - Ayuda contextual.
- **Histórico.** Base de datos de los pacientes que almacena sus datos personales y todos los datos recogidos en cada consulta realizada. Se han utilizado para ello tablas de MS Access.
- **Gestor de históricos.** Es un gestor de base de datos. Es el módulo que se encarga de gestionar los datos de los pacientes.
- **Validación de datos.** Este módulo se comunica con la interfaz de usuario para validar los datos que introduce el usuario, y con el gestor de

históricos para comprobar si los datos introducidos pertenecen a algún paciente dado ya de alta en la base de datos de pacientes.

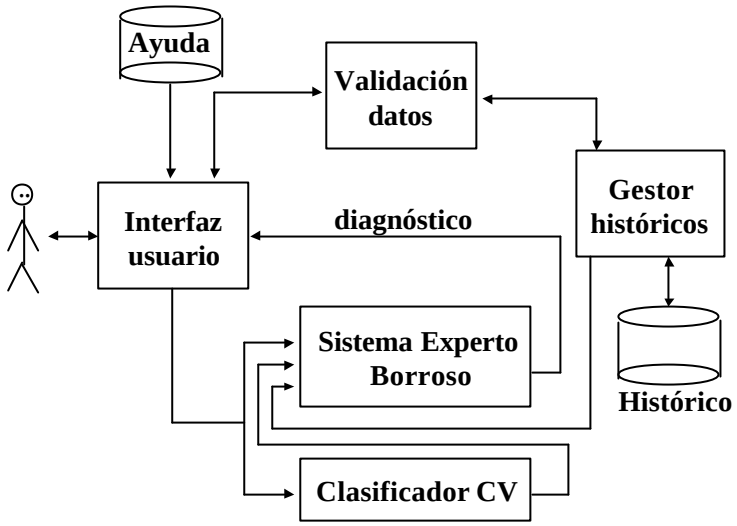


Figura 5.16: Diseño arquitectónico del Sistema híbrido de ayuda al diagnóstico del glaucoma.

5.5 Inferencia a través de la base de conocimiento

El proceso que se sigue al realizar una consulta es básicamente el mismo que el del Glaucom-Easy, aunque facilitado por una interfaz más flexible y cómoda.

Inicialmente se deben introducir los datos necesarios para obtener un diagnóstico. El modelo de ficha de datos que usa este Sistema Experto es el de las Figuras 5.17 y 5.18.

Mediante la ventana principal se recogen los factores de riesgo del paciente y se pasa a las ventanas de cada ojo en las que se introducen los demás parámetros, que luego se mostrarán también en la principal para tener una visión global del paciente.

Observando la Figura 5.18 podemos apreciar la diferencia esencial entre este sistema y el Glaucom-Easy. El usuario no introduce la clasificación del campo visual, hecha por él siguiendo determinadas pautas o consejos, sino que introduce los datos obtenidos directamente del perímetro para que sea el sistema el que ayude a clasificar el campo.

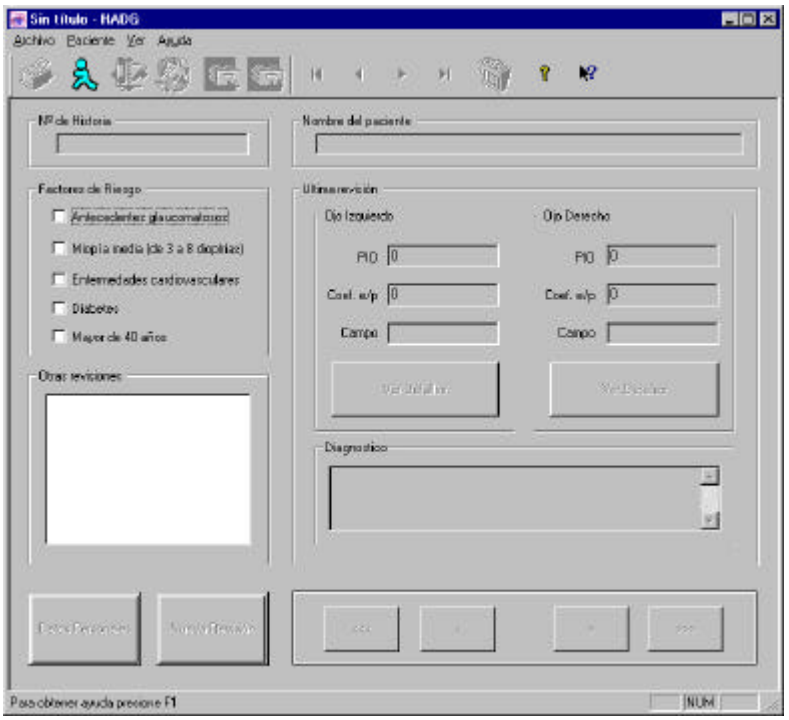


Figura 5.17: Ventana principal de la aplicación [74].

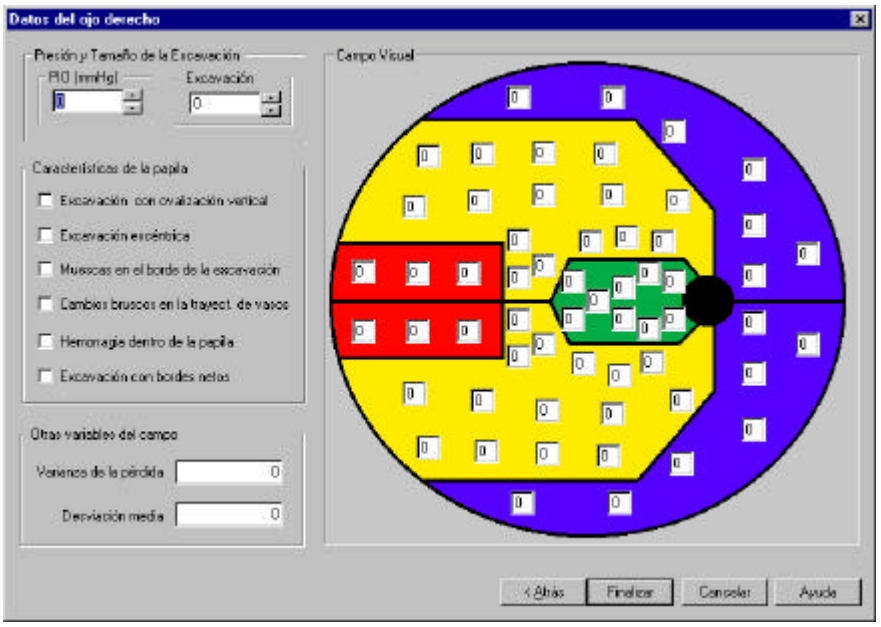


Figura 5.18: Ventana en la que se recogen todos los datos del ojo derecho [74].

Los grafos de decisión, que se han implementado en forma de reglas de producción para que el sistema experto pueda efectuar el nuevo diagnóstico teniendo en cuenta los anteriores, son los mismos que los del Glaucom-Easy (Figuras 3.10 y 3.11).

En este sistema también usamos dos tipos de reglas para obtener un diagnóstico:

- **Reglas instantáneas**: sólo requieren la información que se introduce cuando se realiza la consulta. Se disparan en caso de haber alto riesgo de padecer glaucoma. Estas reglas evalúan tanto valores extremos de PIO como excavaciones demasiado grandes en la papila del nervio óptico, una fuerte asimetría en el tamaño de las mismas entre ambos ojos o un campo visual claramente glaucomatoso.
- **Reglas progresivas**: utilizan históricos del paciente para determinar el diagnóstico actual, es decir, realizan un seguimiento del paciente. Mediante estas reglas implementamos los grafos de decisión.

Alguna de las reglas instantáneas se muestra a continuación en pseudocódigo:

```
si PIO_derecha patológica
o PIO_izquierda patológica
entonces
    diagnóstico=Muy alto riesgo. Conviene instaurar un
    tratamiento o enviarlo al especialista en glaucoma
fin si

si campo_visual_derecho patológico
o campo_visual_izquierdo patológico
entonces
    diagnóstico=Muy alto riesgo. Conviene instaurar un
    tratamiento o enviarlo al especialista en glaucoma
fin si
```

Un ejemplo de las reglas progresivas es el siguiente:

```
si (diagnóstico_anterior es “Bajo riesgo. Puede volver
al año si lo desea.”
y PIO_derecha dudosa
y campo_visual_derecho dudoso)
entonces
    diagnóstico=Presión dudosa. Tomarla de nuevo en un
    plazo de 7 días, preferiblemente por la mañana.
fin si
```

Como ya se indicó al hablar del diseño arquitectónico del sistema, la parte del mismo que produce el diagnóstico consta de dos módulos: el clasificador del campo visual y el sistema experto borroso. Para implementarlos se ha utilizado FuzzyCLIPS y Visual C++ .

El clasificador de campos visuales a su vez realiza la clasificación utilizando dos subsistemas: un sistema experto para el diagnóstico del campo visual y una red neuronal. Las salidas de ambos subsistemas se integran mediante una serie de reglas de producción. Todo esto ha sido implementado en Visual C++.

El hecho de unir en un sistema de clasificación para el diagnóstico del campo visual los dos enfoques, simbólico y conexionista, pensamos que puede mejorar la precisión diagnóstica del sistema, ya que introducimos la experiencia del experto en un sistema experto y entrenamos también una red neuronal que a su vez dará un diagnóstico. Contrastando los dos diagnósticos podríamos lograr un sistema más robusto.

A continuación explicamos con detalle cómo se realiza esta clasificación.

5.6 Clasificador del campo visual

El diseño arquitectónico de este clasificador del campo visual se muestra en la figura siguiente (Figura 5.19):

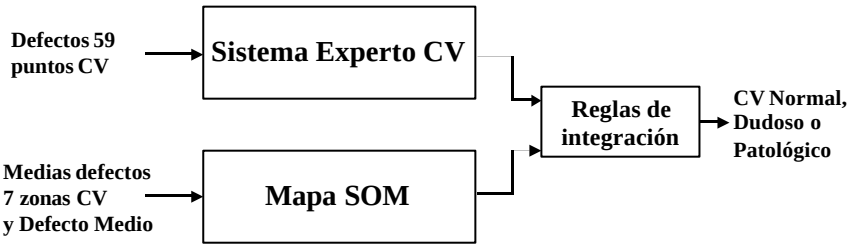


Figura 5.19: Diseño arquitectónico del Clasificador del Campo Visual.

En los siguientes subapartados vamos a ir explicando cada uno de sus módulos:

- Sistema Experto del campo visual
- Mapa SOM clasificador del campo visual
- Integración de los dos anteriores mediante reglas

5.6.1 Desarrollo del Sistema Experto para el diagnóstico del campo visual

El objetivo de este sistema experto es ayudar en la clasificación de campos visuales para detectar glaucoma incipiente. Tener en cuenta esto es fundamental para entender las decisiones que se han tomado al elaborar las reglas de producción.

Para desarrollar este sistema experto nos hemos inspirado en un trabajo de investigación de Gaasterland y col. [54]. Éstos desarrollaron un método cuantitativo para medir la confiabilidad de la prueba y la gravedad de los defectos del campo visual glaucomatoso realizado con el analizador Humphrey. Sobre todo nos hemos fijado en su segundo objetivo. Para medir el grado de defectos glaucomatosos se basan en el número y profundidad de los clusters de puntos con defecto de la exploración en el hemicampo superior e inferior y en el área nasal, tomando esas zonas tal como indica la Figura 5.20. Toman un rango que va desde 0 (ningún defecto) hasta 20 (todos los puntos con defectos profundos). Partiendo del valor inicial 0, mediante una serie de reglas relativas a los defectos que aparecen en las zonas van aumentando ese valor.

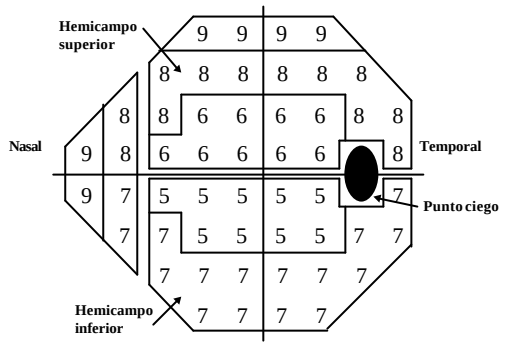


Figura 5.20: Cantidad mínima de depresión, en decibelios, que identifica localizaciones del campo como defectuosas en la prueba umbral 24-2 del Analizador de campo visual Humphrey.

Nuestro sistema experto para el campo visual valora los defectos de sensibilidad que muestra la prueba campimétrica, y va aumentando o disminuyendo una marca global dependiendo de si los defectos hallados son característicos de un campo con glaucoma incipiente o pueden ser debidos a otras patologías.

Por tanto, como datos de entrada al sistema experto se han tomado los 59 puntos evaluados por el perímetro computerizado Octopus 500.

Para diseñar este sistema consideramos importante recordar cómo se va afectando el campo visual en el glaucoma (apartado 2.1) [104].

A la hora de realizar un diagnóstico del campo visual, especialmente en fases incipientes, es posible dividir los 30 grados centrales del mismo (la zona evaluada por la mayoría de los perímetros computerizados existentes en el mercado) en varias zonas según la evolución de las lesiones, teniendo en cuenta la distribución de las fibras nerviosas en la retina, la experiencia clínica de la Unidad de Glaucoma del Hospital Clínico Universitario de Valladolid y las ya empleadas en el sistema experto Glaucom-Easy [6, 7, 142]. Esas zonas se muestran en la Figura 5.21. El punto negro representa la mancha ciega, más cercana al lado temporal del campo en oposición a la zona nasal.

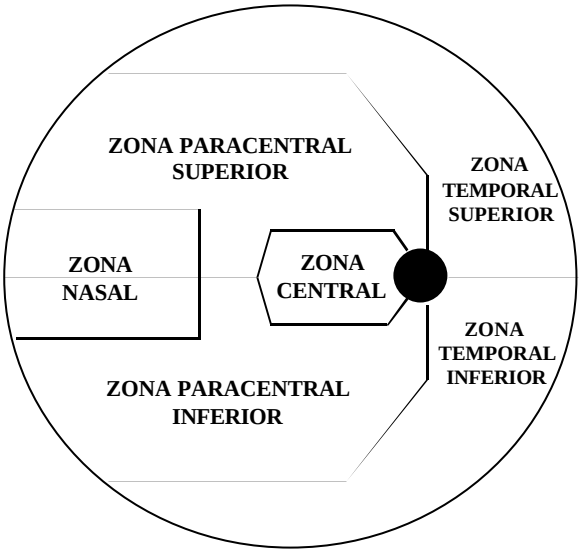


Figura 5.21: Zonas del campo visual.

Cada una de las zonas tendrá una importancia determinada a la hora de diagnosticar un glaucoma:

- **Zonas 1 y 6 (nasal superior e inferior):** las más importantes, ya que en los seis puntos que las componen es donde se forman los escalones nasales, defectos típicamente glaucomatosos.
- **Zonas 2 y 5 (paracentral superior e inferior):** también poseen una gran importancia en el diagnóstico precoz de glaucoma, ya que será en ellas donde se formen los primeros escotomas aislados.
- **Zonas 3 y 4 (temporal superior e inferior):** tienen cierto interés diagnóstico los defectos localizados en la zona superior.
- **Zona 7 (central):** sus defectos guardan escasa relación con el glaucoma, pues como ya se dijo en el capítulo 2 (apartado 2.1), la agudeza visual se mantiene intacta hasta estadios terminales de la enfermedad.

Los puntos del campo visual se numeran tal como se mostró en el capítulo 4 (Figura 5.22). Se considerará que dos puntos forman un cluster si están próximos en la figura y están aproximadamente a la misma distancia del centro del campo visual (punto 36). La justificación a esta restricción de adyacencia se debe a que típicamente, los defectos glaucomatosos del campo visual tienden a formar anillos alrededor de dicho punto y no a seguir una trayectoria radial.

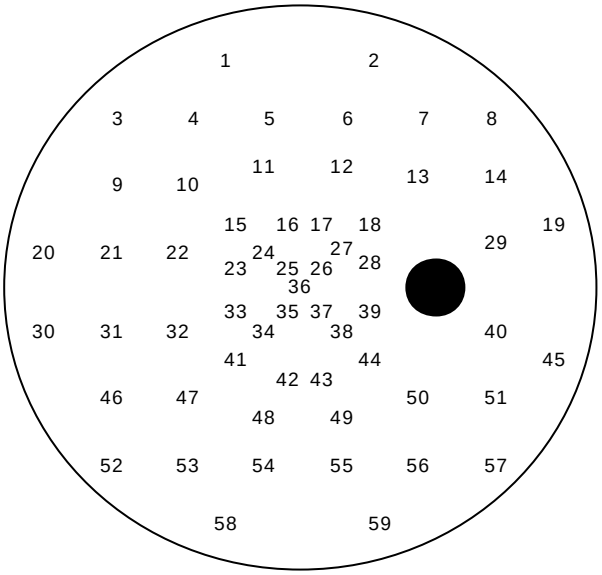


Figura 5.22: Número asignado a cada punto del campo visual.

Solamente se ha tenido en cuenta la distribución de los defectos, y no su profundidad, ya que en los estados incipientes de glaucoma, ésta es irrelevante, mientras que su forma y situación son fundamentales para el diagnóstico.

El modo de funcionamiento de este sistema experto se basa en puntuar los defectos que se vayan encontrando en las diversas zonas. Esta puntuación será positiva en caso de encontrar un defecto típicamente glaucomatoso, y negativa en caso contrario.

Las puntuaciones totales de cada zona se multiplicarán por un factor proporcional a su importancia en el diagnóstico de glaucoma y se sumarán, obteniendo así la valoración total del sistema.

Si dicha valoración es claramente positiva, el campo es considerado glaucomatoso; si es claramente negativa, no glaucomatoso, y si es cercana a cero, se considerará dudoso.

Los defectos que se han puntuado son los siguientes:

- **Zona Nasal (zonas 1 y 6):**
  - Escalón nasal: defecto típicamente glaucomatoso en el que existen varios puntos con pérdida alineados horizontalmente en la parte nasal del ecuador del campo visual. Probabilidad alta de glaucoma.
  - Defecto nasal: varios puntos unidos con pérdida a ambos lados del ecuador. Defecto demasiado grande para los primeros estadios de un glaucoma.
  - Algún punto con pérdida. Posible formación de un escalón nasal.

- **Zona Paracentral (zonas 2 y 5):**
  - Más puntos dañados en el hemisferio inferior que en el superior hace sospechar que no se trate de un glaucoma en gran parte de los casos. La asimetría de las lesiones entre los hemicampos superior e inferior es típica de glaucoma: en un 70% de los casos está más dañado el hemicampo superior y en un 30% hay más lesiones en el hemicampo inferior.
  - Número de clusters de puntos con pérdida: un número excesivo de puntos con pérdida unidos no hacen pensar en un glaucoma incipiente sino posiblemente en otra patología, ya que en el glaucoma inicialmente, los defectos suelen ser más localizados.
  - Número de puntos aislados: por el contrario, si la gran mayoría de los puntos están aislados, tampoco podremos pensar en un glaucoma, ya que en éstos se suelen formar escotomas arciformes, es decir, defectos en forma de anillo alrededor del punto central de visión.

- **Zona Temporal (zonas 3 y 4):**
  - Más puntos dañados en el hemisferio inferior que en el superior hace sospechar que no se trate de un glaucoma en gran parte de los casos. La asimetría de las lesiones entre los hemicampos superior e inferior es típica de glaucoma: en un 70% de los casos está más dañado el hemicampo superior y en un 30% hay más lesiones en el hemicampo inferior.
  - Número de clusters de puntos con pérdida: un número excesivo de puntos con pérdida unidos no hacen pensar en un glaucoma incipiente sino posiblemente en otra patología, ya que en el glaucoma inicialmente, los defectos suelen ser más localizados.
  - Número de puntos aislados: por el contrario, si la gran mayoría de los puntos están aislados, tampoco podremos pensar en un glaucoma, ya que en éstos se suelen formar escotomas arciformes, es decir, defectos en forma de anillo alrededor del punto central de visión.

- **Zona Central (zona 7):**
  - Número de puntos con pérdida: la zona central es la última en verse afectada por un glaucoma, por lo que demasiados puntos con perdida serán indicativos de otra patología.

Se ha tenido en cuenta que el campo era normal si no tenía ningún punto con defecto en el campo o sólo uno.

Resultados

El sistema experto se probó con los 48 ejemplos de campos visuales de la muestra de prueba cuya clasificación es conocida y con los que se han realizado los experimentos en el capítulo anterior (Tabla 5.3).

	Muestra de prueba
Glaucoma	20
Normal	9
Cataratas	7
Diabetes	7
HTA	5
TOTAL	48

Tabla 5.3: Distribución de casos según clases.

Una vez ajustados los niveles dentro de los cuales el campo será considerado dudoso, el número de casos de cada clase (glaucoma o no glaucoma) que no están claros se muestran en la Tabla 5.4.

	Muestra de prueba
Dudosos (Glaucomas)	10 (50%)
Dudosos (No glaucoma)	9 (32.14%)
Total Dudosos	19 (39.58%)

Tabla 5.4: Número de ejemplos de cada clase que han sido clasificados como dudosos por el SE del CV, y su porcentaje.

Estos casos se han eliminado de la muestra a la hora de calcular la sensibilidad (S), especificidad (E) y precisión diagnóstica (PD). Los resultados obtenidos para la clase glaucoma se muestran en la Tabla 5.5.

	S (%)	**E (%)**	**PD (%)**	**FP**	**FN**	**Dudosos (%)**
Glaucoma	100	84.21	89.66	3	0	39.58

Tabla 5.5: Sensibilidad, especificidad y precisión diagnóstica del sistema experto para el diagnóstico del campo visual.

Podemos observar que la sensibilidad es del 100%, y la especificidad del 84.21% en la muestra de prueba. Normalmente se considera que el sistema tiene una capacidad diagnóstica aceptable cuando los valores de sensibilidad y especificidad están por encima del 80% y están preferiblemente equilibrados. De todas formas, es recomendable una especificidad al menos del 90% para evitar en la mayor parte de los casos que pacientes sin glaucoma sean diagnosticados y tratados como tales.

En nuestro sistema, si bien tanto la sensibilidad como la especificidad están dentro de los niveles aceptados (por encima del 80%), están bastante desequilibradas, y la especificidad no alcanza los límites recomendados para este caso. Es más, la totalidad de los 3 errores cometidos por el sistema son falsos positivos.

Otra cuestión a tener en cuenta es la cantidad elevada de campos diagnosticados como dudosos por el sistema, un 39.58%.

Para intentar reducir el número de campos dudosos y solucionar el sesgo de nuestro sistema hacia el diagnóstico de glaucoma positivo (tratar de aumentar la especificidad sin dañar excesivamente la alta sensibilidad alcanzada), utilizaremos como apoyo una red neuronal.

5.6.2 Elección y modificación de la RNA para clasificación de campos visuales

La RNA elegida para clasificar el campo visual en nuestro sistema borroso junto con el sistema experto anterior es el mapa SOM que mejores resultados produjo al diferenciar los ejemplos en dos categorías. Recordamos sus características (Tabla 5.6):

Topología*	a **1**	a **2**	**Variables**	**S(%)**	**E(%)**	**PD (%)**
8x3	0.03	0.01	7 zonas y MD	95	92.86	93.75

Tabla 5.6: El mejor mapa SOM que clasifica en glaucoma y no glaucoma entrenado con las medias de los defectos de sensibilidad de las siete zonas del campo visual.

Recordemos también que las reglas que se aplicaron en el etiquetado que realiza la herramienta GESL [51] ayudaron a refinar la clasificación y, por tanto, a mejorar nuestros resultados. Dos de esas decisiones que se tomaron en el etiquetado son las siguientes:

1. Cada neurona se etiqueta con la clase que más ejemplos han caído en esa neurona.
2. Si hay dos clases con el mismo número de ejemplos que han activado la neurona correspondiente, se etiqueta con la clase cuya suma de distancias a esa neurona es menor.

Esto nos llevó a estudiar qué ejemplos concretos de la muestra de entrenamiento caían en cada una de las neuronas del mapa SOM elegido. De esta forma podemos saber qué ejemplos de la muestra de entrenamiento habrían sido mal clasificados con este mapa, y, por tanto, qué ejemplos nuevos serían mal clasificados si fueran muy parecidos a éstos.

Hemos podido observar cómo algunas neuronas (0-2, 0-3, 0-7, 1-2, 1-3, 1-5, 2-1, 2-6, 2-7) provocan casi tantos errores como aciertos puesto que son activadas por ejemplos muy similares pertenecientes a clases distintas. Estas neuronas que no permiten definir con nitidez su etiqueta, además de la 1-1, que no resulta ganadora para ninguno de los 180 casos de entrenamiento, van a ser etiquetadas como dudosas.

La topología original de la red se puede ver en la Figura 5.23, donde se representan en blanco las neuronas con la etiqueta de "Otras patologías" y en negro las neuronas con etiqueta "Glaucoma":

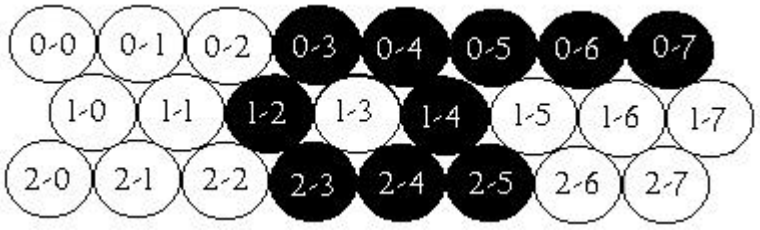


Figura 5.23: Mapa SOM que clasifica en dos categorías: Glaucoma y no glaucoma.

Se puede observar que el SOM ha actuado como un mapa de rasgos, gracias a la función de vecindad (de tipo hexagonal) definida en su entrenamiento, y vectores de datos de entrada con rasgos similares aparecen próximos en la red.

Podemos ver ahora la Figura 5.24, en la cual se ha cambiado la etiqueta de las neuronas con gran porcentaje de error o duda, y observamos cómo éstas se agrupan en torno a la frontera entre neuronas "glaucomatosas" y "no glaucomatosas", como era de esperar.

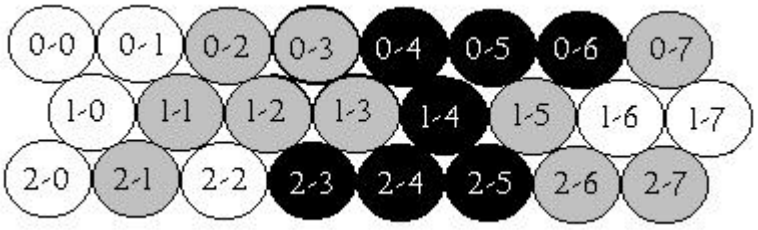


Figura 5.24: Mapa SOM en el que se han modificado algunas etiquetas para clasificar campos dudosos además de glaucomatosos y no glaucomatosos.

Hay que indicar que en el entrenamiento de la red se considera que la primera columna de neuronas es vecina de la última al actualizar los pesos (formando así una especie de anillo), por lo que las neuronas 1-6 y 1-7, aparentemente aisladas del resto de neuronas "no glaucomatosas" no lo están, sino que son adyacentes a la 0-0, la 1-0 y la 2-0.

Dicha solución aprovecha el hecho de que nuestro sistema borroso, que será el destinatario de la información proporcionada por la red junto con el sistema experto del campo visual, dispondrá de más datos a la hora de realizar un diagnóstico final, por lo que es preferible dar un diagnóstico dudoso a dar uno erróneo. Marcando el campo como dudoso, la base de reglas de nuestro sistema borroso nos dará, con gran probabilidad, un diagnóstico adecuado en función del resto de los parámetros del ojo en cuestión.

Resultados

Una vez seleccionadas las neuronas que se etiquetarán como dudosas, el número de ejemplos de cada clase (glaucoma o no glaucoma) que caen en ellas se muestran en la Tabla 5.7.

	Muestra de prueba
Dudosos (Glaucomas)	5 (25%)
Dudosos (No glaucoma)	12 (42.86%)
Total Dudosos	17 (35.42%)

Tabla 5.7: Número de ejemplos de cada clase que han sido clasificados como dudosos por el mapa SOM, y su porcentaje.

Estos casos se han eliminado de la muestra al calcular la sensibilidad, especificidad y precisión diagnóstica. Los resultados obtenidos para la clase glaucoma se muestran en la Tabla 5.8.

	S (%)	**E (%)**	**PD (%)**	**FP**	**FN**	**Dudosos (%)**
Glaucoma	100	87.5	93.55	2	0	35.42

Tabla 5.8: Sensibilidad, especificidad y precisión diagnóstica del mapa SOM para el diagnóstico del campo visual.

Podemos observar que la sensibilidad es del 100%, y la especificidad del 87.5% en la muestra de prueba. Como ya se ha indicado al analizar los resultados del sistema experto del campo visual, normalmente se considera que el sistema tiene una capacidad diagnóstica aceptable cuando los valores de sensibilidad y especificidad están por encima del 80% y están preferiblemente equilibrados. De todas formas, es recomendable una especificidad al menos del 90% para evitar en la mayor parte de los casos que pacientes sin glaucoma sean diagnosticados y tratados como tales.

Mediante esta red neuronal conseguimos que tanto la sensibilidad como la especificidad estén dentro de los niveles aceptados (por encima del 80%), además están menos desequilibradas que en el sistema experto, y aunque la especificidad no alcanza los límites recomendados para este caso está bastante cercana al 90%.

Por otra parte, hemos disminuido la cantidad de campos diagnosticados como dudosos por el SE del CV: con este sistema tenemos un 35.42%.

Vamos a ver si la integración de ambos sistemas: el sistema experto del campo visual y la red neuronal, mejora los resultados.

5.6.3 Sistema híbrido simbólico-conexionista para el diagnóstico del campo visual

Como recordatorio, se muestran en la Tabla 5.9 los resultados obtenidos por ambos sistemas individualmente:

	S (%)	E (%)	PD (%)	FP	FN	Dudosos (%)
Sistema Experto	100	84.21	89.66	3	0	39.58
Red Neuronal	100	87.5	93.55	2	0	35.42

Tabla 5.9: Resultados de sensibilidad, especificidad, precisión diagnóstica, falsos positivos y negativos, y porcentaje de dudosos obtenidos con el sistema experto del campo visual y con el mapa SOM.

Para lograr una integración eficaz de ambos sistemas hemos diseñado una pequeña base de reglas teniendo en cuenta los resultados de la Tabla 5.9. Dichas reglas de producción han sido programadas en Visual C++ y son las siguientes:

- **El sistema experto diagnostica "No glaucoma":** el diagnóstico final será "No glaucoma" independientemente de lo que diga la red neuronal, debido a la menor especificidad del sistema experto.
- **El sistema experto diagnostica "Dudoso":** a fin de reducir el nivel de duda, el diagnóstico final será el que efectúe la red (que posee menos campos dudosos).
- **El sistema experto diagnostica "Glaucoma":**
  - **Si la red también diagnostica "Glaucoma",** el resultado final será ése.
  - **Si la red diagnostica "No glaucoma",** el resultado será "Dudoso", debido al conflicto entre ambos sistemas. Hay que mencionar que esta posibilidad es bastante rara de encontrar, de hecho no nos hemos encontrado ningún caso de los 48 de la muestra de prueba, sin embargo hay que considerar la posibilidad.
  - **Si la red diagnostica "Dudoso",** el diagnostico final será "Dudoso", para compensar el sesgo que el sistema experto tiene a favor del glaucoma.

Resultados

Integrados ambos sistemas y unificado el diagnóstico, obtenemos los siguientes resultados (Tabla 5.11):

	Muestra de prueba
Dudosos (Glaucomas)	5 (25%)
Dudosos (No glaucoma)	10 (35.71%)
Total Dudosos	15 (31.25%)

Tabla 5.10: Número de ejemplos de cada clase que han sido clasificados como dudosos por el sistema integrado, y su porcentaje.

	S (%)	**E (%)**	**PD (%)**	**FP**	**FN**	**Dudosos (%)**
Glaucoma	100	94.44	96.96	1	0	31.25

Tabla 5.11: Sensibilidad, especificidad y precisión diagnóstica del sistema integrado para el diagnóstico del campo visual.

En la Tabla 5.12 mostramos de nuevo los resultados obtenidos con el sistema experto del campo visual, con el mapa SOM, y con el sistema integrado.

	S (%)	**E (%)**	**PD (%)**	**FP**	**FN**	**Dudosos (%)**
Sistema Experto	100	84.21	89.66	3	0	39.58
Red Neuronal	100	87.5	93.55	2	0	35.42
Sistema integrado	100	94.44	96.96	1	0	31.25

Tabla 5.12: Resultados de sensibilidad, especificidad, precisión diagnóstica, falsos positivos y negativos, y porcentaje de dudosos obtenidos con el sistema experto del campo visual, con el mapa SOM, y con el sistema integrado.

Podemos observar que:

- La sensibilidad sigue siendo la máxima, al igual que en cada sistema por separado.
- La especificidad es superior tanto a la de la red como a la del sistema experto.
- La precisión diagnóstica es también superior a la de ambos sistemas por separado.
- El porcentaje de duda es inferior al de ambos sistemas por separado.
- El número de falsos positivos se ha reducido.

Sensibilidad y especificidad son ambas superiores a 90%, por tanto, valores no sólo aceptables, sino por encima de lo recomendable.

Mediante la integración de ambos sistemas se ha conseguido, manteniendo la sensibilidad, aumentar la especificidad, y, por tanto, la precisión diagnóstica, y reducir el porcentaje de campos considerados dudosos, lo que demuestra las ventajas de un sistema híbrido frente a uno convencional.

5.7 Integración del clasificador del campo visual en el sistema de diagnóstico global

Una vez diseñado el clasificador del campo visual, el siguiente paso es integrarlo en el sistema de diagnóstico global, es decir, su salida deberá ser una de las entradas del sistema experto borroso implementado en FuzzyCLIPS. Como ya se ha explicado, la salida del clasificador no es borrosa sino que da uno de tres posibles valores: normal, dudoso o patológico, y luego esos valores se le pasan a FuzzyCLIPS utilizando conjuntos singleton. Podríamos considerar que en cierta medida el haber pasado de clasificar el campo visual en dos categorías, glaucoma y no glaucoma, a tres introduciendo los campos dudosos, ha sido una forma de emborronar el resultado de esa clasificación, y de hacerla más parecida a la que realmente efectúa el experto en su consulta.

La salida del sistema borroso implementado en FuzzyCLIPS es un consejo o pauta a seguir con el paciente y un coeficiente de certeza que da la confianza, expresada en tanto por ciento, que tiene el sistema en que ese consejo es adecuado para el paciente según la presión intraocular, la relación excavación/papila y el campo visual. Pero aún quedan una serie de datos que se recogen en la historia clínica del paciente y que pueden intervenir en el diagnóstico final. La influencia va a ser no sobre el consejo, que ya no va a variar, sino sobre el factor de certeza.

5.8 Modificación del factor de certeza del Sistema Experto Borroso

Los diagnósticos resultantes de la ejecución del sistema borroso se han refinado a partir de datos no borrosos que no se han introducido al sistema de FuzzyCLIPS, sino que se han valorado aparte mediante reglas de producción implementadas en Visual C++. Estos datos, tal y como ya se hizo en el sistema experto Glaucom-Easy, pueden agruparse en dos tipos:

- Factores de riesgo: antecedentes familiares de glaucoma, miopía media, enfermedades cardiovasculares, diabetes, ser mayor de 40 años.
- Características de la papila del nervio óptico: si la excavación tiene una ovalización vertical, es excéntrica, presenta su borde alguna muesca, hay cambios bruscos en la trayectoria de los vasos, o hemorragias dentro de la papila, si la excavación tiene bordes netos.

Cada respuesta afirmativa a las preguntas sobre los factores o características anteriores modificará el coeficiente de certeza de los diagnósticos dados, ajustándose así a la nueva información y personalizando de algún modo la respuesta del sistema a las particularidades de cada paciente concreto. Así, por ejemplo, a igualdad en el tamaño de excavación y la presión intraocular, un paciente con antecedentes familiares de glaucoma y con enfermedades cardiovasculares será diagnosticado de glaucoma con mayor certeza que otro paciente sin ese tipo de antecedentes y con bordes netos en la excavación de su papila.

Los factores de riesgo aumentan, obviamente, la probabilidad de padecer glaucoma. Asimismo, todas las características de la papila influyen a favor de un diagnóstico de glaucoma, a excepción de la última: si la excavación presenta bordes netos, que es signo bastante probable de una excavación fisiológica no glaucomatosa.

Por tanto, finalmente, el diseño arquitectónico del Sistema Experto Borroso es el que se muestra a continuación en la Figura 5.25.

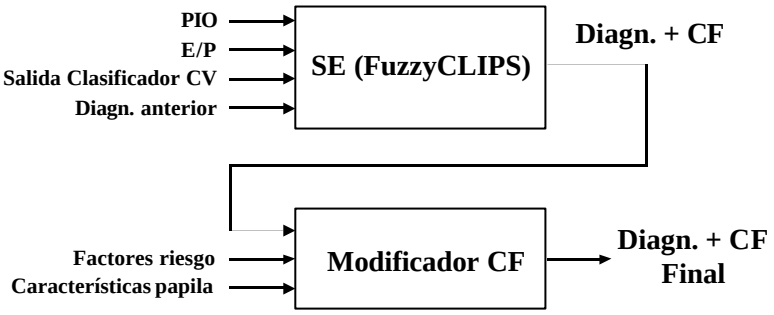


Figura 5.25: Diseño arquitectónico del Sistema Experto Borroso.

5.9 Validación del sistema híbrido de ayuda al diagnóstico del glaucoma

Se incluyeron en el estudio 77 pacientes elegidos aleatoriamente de entre los 228 ejemplos con los que hemos trabajado en los experimentos de esta Tesis y conservando aproximadamente el porcentaje de las cinco clases de la muestra total. Disponíamos de los datos del campo visual obtenidos con el perímetro Octopus y de la relación excavación/papila del nervio óptico de los 77 pacientes elegidos. La distribución de clases se muestra en la Tabla 5.13:

	Muestra
Glaucoma	39
Normal	14
Cataratas	10
Diabetes	8
HTA	6
TOTAL	77

Tabla 5.13: Distribución de casos según clases.

No disponíamos de ningún dato sobre la presión intraocular de dichos pacientes por lo que se tomó la decisión de añadir unas presiones simuladas y que el conjunto de parámetros fuera valorado por un experto. Se le pidió que hiciera dos valoraciones, una sobre el campo visual y otra sobre el paciente globalmente, dando un consejo o pauta a seguir dentro de la lista de consejos que puede emitir

el sistema. Estas pautas son las mismas que las del sistema experto Glaucom-Easy y aconsejan sobre el periodo de revisión, repetición de alguna exploración, necesidad de tratamiento o envío al especialista en glaucoma.

Los 77 pacientes fueron valorados por el experto la primera mitad con PIO normal y la otra mitad con PIO fuera de lo normal. Pasados unos días volvió a valorar los mismos casos pero ahora la primera mitad con PIO fuera de lo normal y el resto con PIO normal. Por tanto, el número de pacientes de la muestra se duplicó.

Se introdujeron los datos de las historias clínicas en la base de datos del sistema. Al introducir la presión intraocular, que debe expresarse en mmHg, se consideró una presión normal de 17 y una presión fuera de lo normal de 25 mmHg. La excavación papilar se debe introducir en décimas. Se introdujeron los datos obtenidos a partir de la prueba campimétrica: los defectos de los 59 puntos del campo visual y el defecto medio.

Lo primero que se pudo observar al comparar los casos valorados por el experto fue que la mayor parte de los campos correspondientes a otras patologías distintas de glaucoma fueron valorados como glaucomas. Eso es lógico ya que al experto le faltaban otros datos sobre el paciente que le indicasen si tenía cataratas, diabetes o hipertensión, y sabemos que el campo glaucomatoso puede adquirir todas las formas de defectos que se dan en otras patologías. Esto nos llevó a eliminar de la muestra todos los casos correspondientes a esas tres patologías. De otro modo no hubiera sido posible comparar los diagnósticos del experto con los del sistema.

Por tanto, finalmente los casos que se utilizaron para la validación del sistema siguen la distribución que se muestra en la Tabla 5.14:

	Muestra
Glaucoma	39
Normal	14
TOTAL	53

Tabla 5.14: Distribución de casos que se usaron para validar el sistema.

A la hora de establecer conclusiones resulta interesante distinguir entre las dos revisiones que el experto hizo: llamaremos **grupo 1** a la primera revisión y **grupo 2** a la segunda. Además vamos a comparar las dos valoraciones que se le pidieron al experto: la clasificación del campo visual, y el consejo o pauta a seguir con el paciente.

5.9.1 Validación de la clasificación del campo visual

Como se ha hecho en estudios anteriores, vamos a analizar primeramente el porcentaje de campos dudosos (Tabla 5.15). Eliminando éstos de la muestra

calcularemos la sensibilidad, especificidad, precisión diagnóstica, falsos positivos y negativos.

	Grupo 1	**Grupo 2**	**Clasificador CV**
Dudosos (Glaucomas)	27 (69.23%)	18 (46.15%)	12 (22.64%)
Dudosos (No glaucoma)	3 (21.43%)	5 (35.71%)	5 (35.71%)
Total Dudosos	30 (56.6%)	23 (43.4%)	17 (32.08%)

Tabla 5.15: Número de ejemplos de cada clase que han sido clasificados como dudosos por el experto en la primera y en la segunda revisión de los datos, y por el clasificador de campos visuales, y su porcentaje.

Podemos observar que el porcentaje global de dudosos en la primera revisión de los datos por parte del experto es bastante alto, y es mucho mayor en la clasificación de campos glaucomatosos que en los normales. Lo que indica que el diagnóstico no es fácil en casos incipientes.

El porcentaje global de dudosos en la segunda revisión del experto ha disminuido respecto a la primera revisión de los datos. Se sigue apreciando un mayor porcentaje de dudosos al clasificar los campos glaucomatosos. El porcentaje de dudosos de la clase normal ha aumentado respecto a la primera revisión.

El porcentaje global de dudosos cuando se usa el sistema clasificador del campo visual disminuye respecto a cualquiera de las dos revisiones del experto, especialmente en la clasificación de campos glaucomatosos, que son detectados con mucha más facilidad. El porcentaje de dudosos en normales es el mismo que el del experto en la segunda revisión y mayor que el del experto en la primera revisión de los datos.

También nos ha parecido interesante calcular el porcentaje de dudosos que coinciden entre las diferentes revisiones (Tabla 5.16).

		Grupo 1	**Grupo 2**
Clasificador CV	**Glaucomas**	83.33%	41.67%
	Normales	40%	80%
	Total	70.59%	52.94%

Tabla 5.16: Porcentaje de dudosos de cada clase que coincide comparando el clasificador del campo visual con cada una de las revisiones del experto.

Por lo que se ve, las coincidencias en dudosos son bastante variables, aunque en ambos casos coinciden más de la mitad de los campos analizados.

Vamos a ver qué coincidencias encontramos entre las dos revisiones del experto (Tabla 5.17).

		Grupo 1
Grupo 2	**Glaucomas**	88.89%
	Normales	40%
	Total	78.26%

Tabla 5.17: Porcentaje de dudosos de cada clase que coincide comparando la segunda revisión del experto con la primera.

El porcentaje de coincidencias de dudosos que consigue el experto en la segunda revisión respecto de la primera es poco mayor que el que consigue el sistema clasificador del campo visual en comparación con la primera revisión del experto. Parece que es normal, por tanto, que haya un conjunto de casos cuyo diagnóstico no coincida con el del experto.

Los resultados obtenidos en las tres clasificaciones son los mostrados en la Tabla 5.18:

	S (%)	**E (%)**	**PD (%)**	**FP**	**FN**	**Dudosos (%)**
Grupo 1	83.33	100	91.3	2	0	56.6
Grupo 2	85.71	100	90	3	0	43.4
Clasificador CV	96.3	100	97.22	1	0	32.08

Tabla 5.18: Sensibilidad, especificidad, precisión diagnóstica, falsos positivos y negativos del experto en la primera y segunda revisión de los datos, y del sistema clasificador en el diagnóstico del campo visual.

Los resultados que obtiene el experto, lógicamente, están dentro de lo recomendable. Sin embargo el sistema clasificador aún supera estos resultados, disminuyendo el número de falsos positivos sin afectar a la especificidad, y logrando que el número de dudosos sea menor.

5.9.2 Validación del consejo o pauta a seguir con el paciente

Una vez evaluado el clasificador del campo visual respecto al experto, vamos a ver si los consejos que da el sistema de diagnóstico global concuerdan con los del experto. Es importante tener en cuenta que los grafos de decisión que se han implementado para obtener el diagnóstico y seguimiento del paciente fueron diseñados y refinados siguiendo los criterios de un experto distinto al que ha evaluado los casos de pacientes, por lo que las discrepancias entre el consejo del experto y el que ofrece el sistema pueden ser más acusadas. Recordemos que Hirsbrunner, Bebie y col. [14, 76, 77] afirmaron que **no es fácil evaluar la eficacia de un sistema experto ya que los propios expertos no se ponen de acuerdo en el diagnóstico de los casos dudosos**. De hecho, compararon la salida de su SE con el diagnóstico de tres oftalmólogos distintos encontrando importantes diferencias. Por otra parte, los resultados que obtuvo Martin [108] mostraron que la concordancia entre su programa y el experto no fue muy buena especialmente en algunos casos, pero si se ajustaba el programa al modo de actuar de cada uno, los resultados fueron muy buenos (concordancia de 92% a 100%).

De cada uno de los 53 casos tenemos dos diagnósticos distintos, uno con PIO normal y otro con PIO fuera de lo normal. Por tanto, en realidad tenemos 106 casos distintos evaluados. Sin embargo, vamos a analizar primeramente los diagnósticos de la primera revisión, luego de la segunda, y finalmente obtendremos una valoración global.

Estableceremos una división clara a la hora de analizar los diagnósticos: si el resultado de la clasificación del campo visual del experto y del sistema coinciden, o si no coinciden. Lógicamente, en el primer caso, se podrá exigir una mayor correspondencia de diagnósticos que en el segundo, ya que la prueba campimétrica es fundamental para el diagnóstico, y va a condicionarlo en gran medida. Por ejemplo, si un campo visual es para el sistema patológico el consejo va a ser enviarlo al especialista o instaurar un tratamiento, independientemente de los otros parámetros recogidos en la consulta, pero si ese mismo campo es dudoso para el experto, puede que el diagnóstico sea repetir las exploraciones dentro de algunos meses, por lo que resulta muy difícil comparar ambos diagnósticos.

Grupo 1

Coincide la clasificación del campo visual		
28 (52.83%)	18	coinciden experto y sistema
	4	experto: revisión a los 6 meses sistema: revisión a los 3 meses
	2	experto: revisión al año sistema: revisión a los 3 meses
	4	experto: acudir al especialista o instaurar tratamiento sistema: revisión a los 3 meses
No coincide la clasificación del campo visual		
25 (47.17%)	10	coinciden experto y sistema
	1	experto: revisión a los 6 meses sistema: revisión a los 3 meses
	2	experto: revisión a los 6 meses sistema: revisión al año
	2	experto: revisión al año sistema: revisión a los 3 meses
	8	experto: revisión a los 6 ó 12 meses sistema: acudir al especialista o instaurar tratamiento
	2	experto: acudir al especialista o instaurar tratamiento sistema: revisión a los 3 meses

Tabla 5.19: Análisis comparativo del diagnóstico del sistema y del experto en los datos del grupo 1.

Grupo 2

Coincide la clasificación del campo visual		
29 (54.72%)	22	coinciden experto y sistema
	4	experto: revisión a los 6 meses sistema: revisión a los 3 meses
	2	experto: revisión a los 3 meses sistema: acudir al especialista o instaurar tratamiento
	1	experto: acudir al especialista o instaurar tratamiento sistema: revisión a los 3 meses
No coincide la clasificación del campo visual		
24 (45.28%)	9	coinciden experto y sistema
	1	experto: revisión a los 6 meses sistema: revisión a los 3 meses
	12	experto: revisión a los 3 ó 6 meses sistema: acudir al especialista o instaurar tratamiento
	2	experto: acudir al especialista o instaurar tratamiento sistema: revisión a los 3 meses

Tabla 5.20: Análisis comparativo del diagnóstico del sistema y del experto en los datos del grupo 2.

Para calcular una precisión diagnóstica global del sistema vamos a considerar fallos aquellos diagnósticos o consejos que puedan ir en contra de la salud del paciente bien por retrasar la siguiente revisión más de lo que lo haría el experto o por no enviarlos al especialista o instaurar tratamiento cuando el experto lo considere necesario, siempre y cuando el experto haya clasificado bien el campo visual. Los fallos están marcados en negrita en las Tablas 5.19 y 5.20.

	Grupo 1 (53 casos)	**Grupo 2 (53 casos)**	**Total (106 casos)**
Precisión diagnóstica	84.9%	94.34%	89.62%

Tabla 5.21: Precisión diagnóstica del sistema híbrido de ayuda al diagnóstico del glaucoma respecto al experto.

Al comienzo de este apartado recordábamos que Martin [108] no obtuvo una concordancia entre su programa y el experto muy buena, pero si ajustaba el programa al modo de actuar de cada experto, los resultados fueron: concordancia de 92% a 100%. Zahlmann [173, 176] obtuvo una sensibilidad y especificidad de su sistema de 98 y 100% respectivamente cuando fue evaluado por oftalmólogos clínicos, y de 82 y 67% respectivamente cuando la evaluación fue realizada por oftalmólogos privados. Podemos decir que la precisión diagnóstica del sistema implementado en este trabajo es buena comparando el diagnóstico con el del experto: una concordancia del 89.62% (Tabla 5.21), y teniendo en cuenta que el conocimiento incluido en el sistema proviene de otro experto distinto al que lo ha evaluado.

Además, podríamos decir que aunque en algunos casos el sistema obligaría, de seguir su consejo, a revisar al paciente con más frecuencia que siguiendo el consejo del experto, también es cierto que detecta antes los casos patológicos y, por tanto, en esos casos se puede poner antes un tratamiento que pueda frenar, en la medida de lo posible, deterioros sucesivos en el campo visual.

5.10 Ejemplo del seguimiento de un paciente con el sistema híbrido desarrollado

Para ver cómo funciona el sistema ante un caso dudoso, planteamos el siguiente caso clínico.

Un paciente llega por primera vez a una consulta médica y las exploraciones realizadas proporcionan los siguientes datos: PIO 22 mmHg, relación E/P: 0.5. Se le hace también una prueba campimétrica. Todos los datos son introducidos al sistema híbrido. El clasificador del CV considera el campo dudoso. El sistema da dos posibles diagnósticos o consejos considerando todos los parámetros:

- 2': Presión dudosa. Tomarla de nuevo en el plazo de 7 días, preferiblemente por la mañana (CF 0.65).
- 3'': Repetir la tonometría y el campo visual a los 3 meses (CF 0.27).

El primero de ellos se obtiene como consecuencia de considerar la presión de 22 mmHg, además de los otros datos. El segundo es debido a considerar que una PIO de 22 mmHg tiene también un grado de pertenencia distinto de cero al conjunto borroso "normal" puesto que hay personas con esa PIO que no tienen glaucoma.

Si suponemos que el médico elige el primer consejo, volverá a tomar la presión al paciente en un plazo de 7 días. En esta segunda consulta la PIO sigue siendo de 22 mmHg. El sistema podría dar dos posibles diagnósticos:

- 3': Repetir la tonometría y el campo visual a los 3 meses (CF 0.65).
- 3'': Repetir la tonometría y el campo visual a los 3 meses (CF 0.27).

Al igual que en la consulta anterior, el primero se obtiene al considerar la PIO dudosa y el segundo al considerarla normal. Pero puesto que el consejo o pauta a seguir con el paciente es la misma, al usuario sólo se proporcionará el primero de ellos que es el de mayor factor de certeza.

Al cabo de tres meses el paciente vuelve a consulta y esta vez la PIO es de 26 mmHg, la relación E/P sigue siendo 0.5 y el CV también es considerado dudoso por el clasificador del CV. El sistema proporciona dos posibles diagnósticos:

- 6: Conviene instaurar un tratamiento o enviarlo al especialista en glaucoma. (CF 0.88).
- 4': Revisión a los 6 meses (CF 0.65).

En este caso se está considerando por una parte que el paciente sigue teniendo un CV dudoso, y por otra parte que la PIO ha aumentado. Ésta tiene ahora un grado de pertenencia al conjunto borroso "límite" distinto de cero y también al conjunto borroso "dudosa". Debido a PIO límite y CV dudoso se obtiene el diagnóstico 6, y por tener PIO dudosa y CV dudoso el sistema también aconseja 4'.

Evidentemente el médico será el que tenga la decisión final en cada una de las consultas.

Por el modo en que se realiza el diagnóstico y a la vista de los resultados se puede observar la importancia de utilizar todos los datos recogidos en las exploraciones y no solo los del CV. En este caso la PIO límite -junto con el campo dudoso- produce la advertencia sobre el alto riesgo de glaucoma: "Conviene instaurar un tratamiento o enviarlo al especialista en glaucoma".

Esta capacidad del sistema de facilitar el seguimiento del paciente y de proporcionar, por tanto, un diagnóstico teniendo en cuenta los datos de las consultas anteriores, aprovechando a la vez las ventajas de un sistema borroso, no había sido implementada en sistemas anteriores, como por ejemplo en los de Martin [108] y Zahlmann [173, 176] que son los más completos y utilizan técnicas parecidas.

6 Conclusiones y trabajo futuro

6.1 Conclusiones

El diagnóstico precoz del Glaucoma Crónico Simple no es un tema sencillo de abordar y de resolver. Como hemos podido comprobar, sobre todo a través de los trabajos destacados en el capítulo 2, muchos esfuerzos se han hecho y se siguen haciendo para lograr diseñar herramientas que utilicen técnicas que ayuden a mejorar ese diagnóstico en las fases iniciales de la enfermedad.

Del mismo modo, también se ha realizado un gran esfuerzo en este trabajo, y considero que se ha conseguido una buena aportación a la investigación en este campo.

Hemos logrado desarrollar un sistema híbrido de ayuda al diagnóstico, después de pasar por varias etapas intermedias de diseño e implementación de algunos prototipos, que logra recoger y manejar gran parte del conocimiento y experiencia del experto en el rastreo del glaucoma y diagnóstico en las fases iniciales de esta enfermedad. El sistema diseñado integra todos los datos de una consulta, no sólo el campo visual, para obtener el diagnóstico, y, por tanto, es una aproximación más real al diagnóstico que elabora el experto. Además, de esta forma se mejora el diagnóstico.

Este sistema híbrido combina técnicas simbólicas borrosas y técnicas conexionistas. Se ha probado el buen comportamiento de la mezcla de esas técnicas. Mediante la interacción entre redes neuronales y sistemas borrosos, conseguimos que el sistema sea capaz de manejar información imprecisa, más próxima a lo que ocurre en la realidad ya que es la que maneja el experto, y por otro lado que tenga la flexibilidad, velocidad y adaptabilidad de una red neuronal a la hora de manejar informaciones más complejas, como ocurre en este caso con el campo visual.

Hemos comprobado que combinar las predicciones de varios clasificadores es un modo eficaz de crear clasificadores más precisos que cualquiera de los que los componen.

El sistema híbrido desarrollado realiza un seguimiento del paciente y proporciona el diagnóstico teniendo en cuenta los datos de las consultas anteriores. Esta capacidad, que junto con las ventajas de un sistema borroso, mejora los resultados, no había sido implementada en otros sistemas de ayuda al diagnóstico del glaucoma.

En cuanto al tipo de red neuronal artificial utilizada para la clasificación del CV, mapas autoorganizados de Kohonen, podemos decir que a pesar del solapamiento de clases que conllevan los problemas médicos, y que limita cualquier método de clasificación, los resultados han sido satisfactorios. Además, mediante la modificación introducida en los algoritmos del SOM se ha conseguido mejorar los resultados que se habían obtenido utilizando directamente el software original de Kohonen.

En todos nuestros resultados la dependencia con respecto a la muestra escogida es muy grande. Esto se ha observado especialmente en la clasificación del campo visual. El resultado que proporciona una RNA de cualquier tipo depende directamente de los datos, de cómo han sido seleccionados, de la proporción de ejemplos para cada una de las clases, etc. La selección de los datos es, por tanto, de vital importancia, ya que una selección mal hecha producirá unos resultados poco satisfactorios. En el caso del glaucoma la selección de pacientes se ha realizado de una forma estricta y minuciosa, estableciendo una serie de criterios de inclusión y exclusión determinados por el Instituto de Oftalmobiología Aplicada (IOBA). Aunque los resultados obtenidos en este trabajo son buenos, del orden a los obtenidos en trabajos anteriores o superiores, habría que repetir los experimentos con unas muestras de entrenamiento y prueba más grandes. Con el número de ejemplos que hemos manejado podemos decir que las técnicas aplicadas y herramientas desarrolladas pueden ser adecuadas para la ayuda al diagnóstico precoz del glaucoma, pero lógicamente, hará falta una fase mucho más exhaustiva de validación y refinamiento. En estos momentos ya se podría disponer de un número mucho mayor de datos y mejor elegidos aún que los anteriores.

Se ha comprobado que la reducción de los parámetros de entrada en la clasificación del CV, dividiendo el CV en siete zonas según la experiencia de los oftalmólogos, y tomando la media de los defectos de sensibilidad de cada zona, mejora los resultados obtenidos en todos los experimentos realizados, tanto usando SOM como CBR.

Es importante tener en cuenta que el objetivo de la Tesis es lograr un diagnóstico precoz, y, por tanto el sistema diseñado sólo es adecuado para esa función. Todas las decisiones que se han ido tomando en su desarrollo van encaminadas a ese fin. Si quisiéramos tener una herramienta que proporcionara una ayuda en el diagnóstico del glaucoma en cualquiera de sus fases deberíamos

completar fundamentalmente el clasificador del campo visual del sistema híbrido, entrenando y probando con otros casos de glaucomas moderados y avanzados.

Es seguro que se hubieran obtenido mejores resultados en la clasificación del campo si la muestra sólo contuviera normales y glaucomatosos, y eliminásemos por tanto otras patologías; sin embargo, no se hizo así para comprobar la capacidad de los mapas SOM de encontrar características comunes en patrones sin clasificar previamente. Después, al validar el sistema híbrido de ayuda al diagnóstico sí se han eliminado los casos de patologías distintas al glaucoma.

Como conclusiones finales se puede destacar que:

- Los objetivos planteados en este trabajo han sido alcanzados: se ha desarrollado un sistema de ayuda al diagnóstico precoz del glaucoma que recoge y maneja gran parte del conocimiento y experiencia del experto, consiguiendo una concordancia con éste del 90%, y logrando detectar antes los casos patológicos.

- Hemos logrado un sistema híbrido que integra los resultados de una red neuronal artificial con aprendizaje no supervisado –mapa SOM- y de varios sistemas expertos, uno de ellos borroso. Con esta mezcla de clasificadores de distinta naturaleza logramos mejorar la precisión de los resultados de cada uno por separado.

- Las redes neuronales artificiales y en particular los mapas autoorganizados son una herramienta adecuada para clasificar campos visuales en oftalmología.

6.2 Trabajo futuro

Si bien las aportaciones que se han hecho en esta Tesis al diagnóstico precoz del Glaucoma Crónico Simple son importantes, como se ha podido constatar a lo largo de esta memoria, este campo de investigación sigue estando abierto a continuas iniciativas o mejoras de las líneas de trabajo ya iniciadas.

En particular, podemos destacar una serie de cuestiones que han ido surgiendo en el desarrollo de este trabajo, que bien por quedar fuera de los objetivos iniciales, bien por falta de tiempo o datos no han sido abordadas o se hallan en un estado de desarrollo incipiente. Entre ellas están las siguientes:

- Hemos resaltado en las conclusiones la importancia de la muestra que se utilice en los experimentos. Aunque la selección de los datos ha sido hecha de una forma bastante rigurosa, sin embargo el factor tamaño de la muestra condiciona enormemente los resultados, no sólo en cuanto a que alcancen una cota deseable o mejoren los ya obtenidos en otros trabajos, sino en cuanto a poder extraer conclusiones definitivas sobre la técnica usada. Por este motivo

deberíamos volver a repetir los experimentos realizados con una muestra mayor. En estos momentos ya dispondríamos de esa muestra, sin embargo los datos de ella correspondientes a la perimetría computerizada han sido obtenidos con el analizador Humphrey. Manejar esos datos habría supuesto por un lado bastantes cambios en las herramientas que ya habíamos desarrollado puesto que nuestros datos campimétricos provienen del Octopus, y por otro lado no habríamos podido hacer comparaciones de nuestros resultados con los anteriores realizados con los mismos datos [6, 128].

- Aunque el objetivo de esta Tesis es el diagnóstico precoz del glaucoma, y la herramienta desarrollada es adecuada para ello, si queremos que ésta sea realmente operativa, debería poder ayudar a diagnosticar el Glaucoma Crónico Simple en cualquiera de sus fases, y, por tanto, deberíamos ampliar la aplicación, y entrenar y probar con otros casos de glaucomas moderados y avanzados. Para conseguir este nuevo objetivo no sólo sería conveniente disponer de una muestra más grande, como hemos señalado en el punto anterior, sino que contuviese datos de pacientes con glaucomas en todas las fases de desarrollo de la enfermedad.

- Una muestra más grande además nos permitiría refinar, utilizando métodos automáticos, las funciones de pertenencia de los conjuntos borrosos de nuestro sistema híbrido de ayuda al diagnóstico que ahora han sido diseñadas de forma manual.

- Dentro de ese refinamiento de las funciones de pertenencia de los conjuntos borrosos, sería conveniente probar si el emborronado del campo visual aporta ayuda al usuario del sistema. En la herramienta desarrollada el diagnóstico del campo visual se realiza de forma automática pero sólo con tres posibles resultados nítidos: normal, dudoso y patológico. Podría ser una buena ayuda si obtuviésemos también el grado de normal, dudoso o patológico. De esta forma, también se podría detectar la progresión del CV con el tiempo.

- Habría que revisar la base de conocimiento de los dos sistemas expertos que están desarrollados mediante reglas de producción, el SE del CV y el SE Borroso, añadiendo detalles diagnósticos que mejorarían su funcionamiento. Por ejemplo, en el SE del CV se han desarrollado reglas que en algún caso sólo se adaptan al modo de comportamiento general de los pacientes, dejando un porcentaje de éstos sin incluir. Esto ocurre al considerar que el hemicampo superior se daña más en el glaucoma que el inferior. Aunque es cierto, esta regla podría enmascarar el diagnóstico en el 30% de casos de glaucoma donde ocurre lo contrario: que se afecta más el hemicampo inferior. Aunque se ha comprobado que los resultados del clasificador del campo visual son muy buenos, puesto que no se basa sólo en el SE del CV, sin duda este y otros refinamientos podrían mejorarlo. Esto se llevaría a cabo en la fase de validación del sistema híbrido en una unidad de glaucoma real, como se indica más adelante.

- En la clasificación del campo visual mediante redes neuronales hemos probado la capacidad de los mapas SOM, que no se habían utilizado con anterioridad para el mismo fin. Sí se habían utilizado otras redes neuronales como MLP y RBF. En estos momentos estamos desarrollando un generador de experimentos RBF e híbridos RBF-MLP que da la posibilidad de utilizar distintos algoritmos de entrenamiento para la capa oculta y de salida, así como variar otros parámetros. Queremos probar la capacidad de las RBF en el problema de la clasificación de campos visuales, utilizando estas redes de forma distinta a la mostrada en otros trabajos anteriores.

- Si bien la perimetría es una exploración fundamental para el diagnóstico del glaucoma, no lo es menos el estudio de la papila del nervio óptico. Actualmente se están utilizando en investigación datos sobre la papila más objetivos a la hora de emitir un diagnóstico que la relación Excavación/Papila, utilizada en este trabajo. Una mejora sustancial del mismo podría ser la incorporación de esos nuevos datos y de su significado en nuestro sistema híbrido.

- Por último, la validación que hemos realizado del sistema híbrido de ayuda al diagnóstico se puede considerar como un paso inicial para poder concluir que la herramienta puede ser adecuada. Sin embargo habría que validarla durante un periodo de tiempo largo y de forma continuada, y a ser posible en una unidad de glaucoma de un hospital, para asegurar que el número de pacientes es elevado.

Apéndice A: Sensibilidad, especificidad y precisión diagnóstica

La **sensibilidad** es la capacidad de cualquier prueba diagnóstica de detectar cambios en un parámetro. Si la sensibilidad es elevada podrá detectar variaciones pequeñas. Cuando se utiliza una prueba dicotómica (una cuyos resultados se puedan interpretar directamente como positivos o negativos), la sensibilidad es la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como positivo respecto a la condición de la prueba, razón por la que también es denominada fracción de verdaderos positivos (FVP) [98].

$$Sensibilidad = \frac{Número\ de\ positivos\ detectados\ como\ positivos}{Número\ total\ de\ positivos}$$

La **especificidad** es la capacidad de detectar aquello para lo que ha sido diseñado el test. Es la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como negativo. Es igual al resultado de restar a uno la fracción de falsos positivos (FFP).

$$Especificidad = \frac{Número\ de\ negativos\ que\ dieron\ negativo}{Número\ total\ de\ negativos}$$

La sensibilidad y especificidad de un prueba diagnóstica ideal sería del 100%, nivel que raramente se logra.

La **precisión diagnóstica** es la proporción de individuos positivos y negativos que fueron detectados como tales:

$$Precisión\ diagnóstica = \frac{N^o\ de\ negativos\ que\ dieron\ negativo + N^o\ de\ positivos\ que\ dieron\ positivo}{N^o\ total\ de\ negativos\ +\ N^o\ total\ de\ positivos}$$

Los **falsos positivos** (FP) son individuos negativos que aparecen como positivos según la prueba diagnóstica que se está usando.

Los **falsos negativos** (FN) son individuos positivos que aparecerían como negativos.

		Presencia de enfermedad		
		Sí	**No**	**Total**
Resultado de la prueba	**Positiva**	*VP*	*FP*	*VP + FP*
	Negativa	*FN*	*VN*	*VN + FN*
Total		*VP + FN*	*VN +FP*	*VP + FN + VN + FP*
Sensibilidad		*= VP/(VP + FN) = FVP*		
Especificidad		*= VN/(VN + FP) = FVN = 1-FFP*		
Valor predictivo positivo		*= VP/(VP + FP)*		
Valor predictivo negativo		*= VN/(VN + FN)*		

Tabla A.1: Análisis de los parámetros de filtrado.

El **valor predictivo positivo** es la probabilidad de que el individuo padezca una enfermedad cuando el test es positivo.

El **valor predictivo negativo** es la probabilidad de que el individuo no padezca una enfermedad cuando el test es negativo.

Apéndice B: Curvas ROC

Muchas veces los investigadores biomédicos se enfrentan al problema de saber la precisión de un test o prueba clínica para identificar casos de individuos que padecen una determinada enfermedad. La capacidad de un test de diferenciar casos con la enfermedad de otros normales se evalúa mediante el análisis de la curva de Característica Operativa del Receptor (Receiver Operating Characteristic, ROC). Estas curvas también pueden usarse para comparar el funcionamiento diagnóstico de dos o más tests.

Como curiosidad una reseña histórica: la procedencia del nombre “Receiver Operating Characteristic”. El análisis ROC es parte de un campo llamado “Teoría de Detección de la Señal” desarrollado durante la segunda guerra mundial para el análisis de imágenes de radar. Los operadores de radar tenían que decidir si una irregularidad momentánea en la pantalla era un destino enemigo, un amigo, o sólo ruido. La teoría de detección de la señal mide la capacidad de los operadores de los receptores de radar de hacer estas distinciones importantes. A esta capacidad se le llamó características operativas del receptor. En los años 70 se aceptó la utilidad de esa teoría de detección de la señal para interpretar resultados de las pruebas médicas.

Cuando se consideran los resultados de un test particular en dos poblaciones, una con una enfermedad y la otra sin ella, raramente se observa una separación perfecta entre los dos grupos. De hecho, la distribución de los resultados del test se solapan, como se muestra en la Figura B.1.

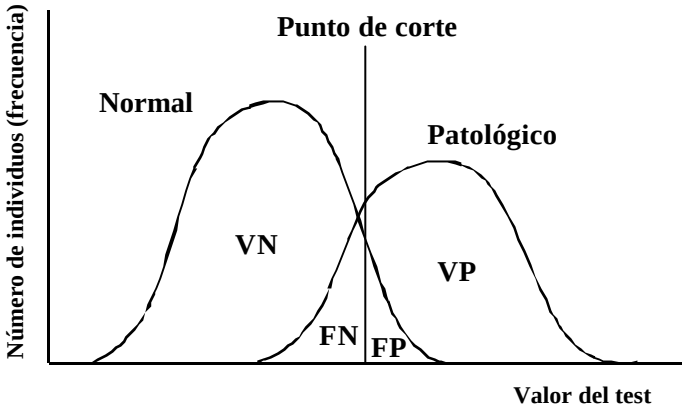


Figura B.1: Ejemplo de distribución de los resultados de un test.

En la práctica, se elige un punto de corte (indicado por una línea vertical en la Figura B.1) por encima del cual se considera que el test es anormal y por debajo que es normal. Para cada posible punto de corte que se fija para discriminar entre las dos poblaciones, habrá algunos casos patológicos correctamente clasificados como positivos (VP = verdaderos positivos), pero otros casos patológicos serán clasificados como si fueran normales (FN = falsos negativos). Por otra parte, algunos casos normales serán clasificados correctamente como negativos (VN = verdaderos negativos), pero otros casos normales serán clasificados como positivos (FP = falsos positivos).

La curva ROC se obtiene al representar esos valores obtenidos con cada punto de corte. Es una gráfica que expresa la proporción de verdaderos positivos frente a la proporción de falsos positivos para los posibles puntos de corte de una prueba diagnóstica (Figura B.2).

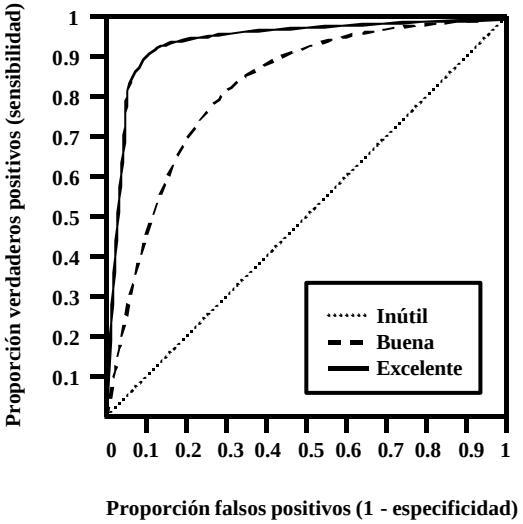


Figura B.2: Curvas ROC. Comparación.

Una curva ROC muestra el compromiso entre sensibilidad y especificidad (cualquier aumento en sensibilidad vendrá acompañado por una disminución en especificidad).

La precisión del test depende de la capacidad que tenga para separar el grupo que está examinando en individuos normales y patológicos. El área bajo la curva es una medida de esa precisión. Un área de 1 representa un test perfecto; un área de 0.5 representa un test inútil. Una guía aproximada para clasificar la precisión de un test diagnóstico es la siguiente:

- 0.9-1 = excelente
- 0.8-0.9 = bueno
- 0.7-0.8 = moderado
- 0.6-0.7 = pobre
- 0.5-0.6 = defectuoso

Bibliografía

1. “The Field Analyzer Primer”, Editado por Mary Jean Haley, Humphrey Instruments, Inc., A company of the Carl Zeiss Group, San Leandro, California, 1987.

2. "Machine Learning. A Multistrategy Approach", Morgan Kaufmann Publishers, 1994, Editores Ryszard Michalski and Gheorghe Tecuci.

3. Aamodt A., Plaza E., “Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches”, AI Communications, 7(i), 39-59, 1994.

4. Aleshunas, J. J., Clair, D. C., Bond, W. E., “Classification Characteristics of SOM and ART2”, Proceedings of the 1994 ACM Symposium on Applied Computing, April 1994, pp. 297-302.

5. Alonso Romero, L., Simón Hurtado, A., Antón López, A., “Use of Self Organizing Methods for Visual Fields Classification”, Actas de la VII Reunión de Trabajo en Procesamiento de la Información y Control, San Juan, Argentina, 17-19 de Septiembre de 1997, Volumen I, pp. 171-173.

6. Antón, A., “Valor de las Redes Neuronales y la Discriminación Logística en el análisis de los defectos del campo visual”, Tesis doctoral. Universidad de Valladolid, 1995.

7. Antón, A., Jordano, J., Maquet, J. A., “Sistema Experto de diagnóstico de glaucoma: GLAUCOM-EASY”, Arch Soc Esp Oftalmol, 1995; 69:23-28.

8. Antón, A., “Prevalence of primary open angle glaucoma. The Segovia study”, ARVO annual meeting. Fort Lauderdale, FL, USA, Mayo 2000.

9. Antón, A., “Characteristics of the normal optic disk in a Spanish population”, ARVO annual meeting. Fort Lauderdale, FL, USA, Mayo 2001.

10. Åsman, P., Heijl, A., "Glaucoma Hemifield Test. Automated Visual Field Evaluation.", Arch Ophthalmol, Vol 110, June 1992.

11. Åsman, P., Heijl, A., Olsson, J., Rootzen, H., "Spatial analyses of glaucomatous visual fields; a comparison with traditional visual field indices", Acta Ophthalmol (Copenh) 1992 Oct; 70(5):679-86.

12. Aulhorn, E., Karmeyer, H., "Frequency distribution in early glaucomatous visual field defects", Doc Ophthalmol Proc Series 14: 75-83, 1977.

13. Bazoon, M., Stacey, D. A., Cui, C., Harauz, G., "A Hierarchical Artificial Neural Network System for the Classification of Cervical Cells", Proceedings of International Congress on Computational Intelligence ICNN'94.

14. Bebie, H., "Computer-assisted evaluation of visual fields", Graefe's Arch Clin Exp Ophthalmol 228: 242-245, 1990.

15. Birch, M. K., Wishart, P. K., O'Donnell, N. P., "Determining progressive visual field loss in serial Humphrey visual fields", Ophthalmology 1995 Aug; 102(8):1227-35.

16. Brassard, G., Bratley, P., "Fundamentos de Algoritmia", Prentice Hall, Madrid, 1997.

17. Brechner, R. J., Whalen, W. R., "Creation of the Transformed Q Statistic Probability Distribution to Aid in the Detection of Abnormal Computerized Visual Fields", Ophthalmic Surgery, Oct. 1984, Vol. 15, nº 10: 833-836.

18. Brigatti, L., Hoffman, D., Caprioli, J., "Neural Networks to Identify Glaucoma With Structural and Functional Measurements", American Journal of Ophthalmology 1996; 121:511-521.

19. Brigatti, L., Nouri-Mahdavi, K., Weitzman, M., Caprioli, J., "Automatic detection of glaucomatous visual field progression with neural networks", Arch Ophthalmol 1997 Jun; 115(6):725-8.

20. Calonge Cano, Teodoro, "Simulación de Redes Neuronales Artificiales mediante Procesamiento en Paralelo", Departamento de Informática. Universidad de Valladolid (España), 1998.

21. Calonge, T., Alonso, L., Ralha, R., Sánchez, A. L., "Parallel Implementation of an Hybrid Neural Network Used for Speech Recognition Task", EUROSPEECH'95 Madrid, septiembre 18-21 1995.

22. Caminhas, W., Tavares, H., Gomide, F., "Competitive Learning of Fuzzy, Logical Neural Network", 6th IFSA World Congress, Sao Paulo, Julio 22-28, 1995.

23. Cazorla, M. A., Escolano, F., Rizo, R., Satorre, R., “Laboratorio de Inteligencia Artificial”, Publicaciones de la Universidad de Alicante, 1998.

24. Chauhan, B. C., Drance, S. M., Douglas, G. R., "The use of visual field indices in detecting changes in the visual field in glaucoma", Invest Ophthalmol Vis Sci 1990 Mar 1; 31(3):512-20.

25. Chauhan, B. C., Drance, S. M., Lai, C., "A cluster analysis for threshold perimetry", Graefe's Arch Clin Exp Ophthalmol 227:216-220, 1989.

26. Chauhan, B. C., Drance, S. M., LeBlanc, R. P., Lieberman, M. F., Mills, R. P., Werner, E. B., "Technique for determining glaucomatous visual field progression by using animation graphics", Am J Ophthalmol 1994 Oct 15; 118(4):485-91.

27. Chauhan, B. C., Henson, D. B., Hobley, A. J., "Cluster analysis in visual field quantification", Doc Ophthalmol 69: 25-39, 1988.

28. Cheng, G., Wu, J.X., Liu, X., Fitzke, F., Hitchings, R., “Pattern discovery of motion sensitivity loss in glaucoma using unsupervised neural networks”, Invest Ophthalmol Vis Sci, 1994; 35: 2185.

29. Cho, K. W., Liu, X., Loizou, G., Wu, J. X., “An AI approach to dynamic visual field testing”, Comput Biomed Res, 1998, Jun, VOL: 31 (3),143-63.

30. Cox, M. J., Wood, I. C. J., "Measurement of Friedmann Visual Field Analyzer Tests in Primary Open-Angle Glaucoma", Optometry and Vision Science, Vol. 70, nº 5, pp. 421-432, 1993.

31. Darlington, K., “The essence of Expert Systems”, Prentice Hall, 2000.

32. Davis, R., Shrobe, H., Szolovits, P., "What is a Knowledge Representation?", AI Magazine, 14(1):17-33, 1993.

33. Doggett, K., “Glossary of Verification Terms”, NOAA Space Environment Center/Space Weather Operations, Jun. 1998.

34. Dorffner, G., “Taxonomies and Part-Whole Hierarchies in the Acquisition of Word Meaning – A Connectionist Model”, The Austrian Research Institute for Artificial Intelligence, Vienna, Austria. Technical Report oefai-tr-92-38, 1992.

35. Dorffner, G., Porenta, G., "On using feedforward neural networks for clinical diagnostic tasks", Artificial Intelligence in Medicine 6, 417-435, 1994.

36. Dorffner, G., “Classification through Hyperplane Fitting with Feedforward Neural Networks”, The Austrian Research Institute for Artificial Intelligence and Dept. of medical Cybernetics and Artificial Intelligence, Vienna, Austria. Technical Report oefai-tr-95-19, 1995.

37. Dorffner, G., “A generalized view on learning in feedforward neural networks”, The Austrian Research Institute for Artificial Intelligence and Dept. of medical Cybernetics and Artificial Intelligence, Vienna, Austria. Technical Report oefai-tr-95-20, 1995.

38. Duggan, C., Sommer, A., Auer, C., Burkhard, K., "Automated Differential Therhold Perimetry for Detecting Glaucomatous Visual Field Loss", American Journal of Ophthalmology 100: 420-423, Sep. 1985.

39. Edenbrandt, L., Devine, B., Macfarlane, P. W., "Neural networks for classification of ECG ST-T segments", J Electrocardiol 1992 Jul; 25(3):167-73.

40. Faúndez-Abans, M., Ormeño, M. I., Oliveira-Abans, M. de, “Classification of planetary nebulae by cluster analysis and artificial neural networks”, Astron. Astrophys. Suppl. Ser. 116, 395-402, 1996.

41. Feigenbaum, E., McCorduch, P., Nii, H. P., “The Rise of the Expert Company: How Visionary Companies Are Using Artificial Intelligence to Achieve Higher Productivity and Profits”, New York: Times Books, 1988.

42. Figueiredo, M., Gomide, F., Pedrycz, W., "Fuzzy Neurons and Networks: Models and Learning", Journal ECLA005, 1995.

43. Filligoi, G. C., Capitanio, L., Accornero, N., "Automatic visual field: a software diagnostic procedure", Comput Methods Programs Biomed 1992 Jun; 38(1):49-59.

44. Fitzke, F. W., Hitchings, R. A., Poinoosawmy, D., McNaught, A. I., Crabb, D. P., "Analysis of visual field progression in glaucoma", Br J Ophthalmol 1996 Jan; 80(1):40-8.

45. Flammer, J., Drance, S. M., Augustiny, L., Funkhouser, A., "Quantification of glaucomatous visual field defects with automated perimetry", Invest Ophthalmol Vis Sci 1985 Feb; 26(2):176-81.

46. Flammer, J., Jenni, F., Bebie, H., Keller, B., "The Octopus Glaucoma G1 Program", Glaucoma 1987; 9:67-72.

47. Flexer, A., "Connectionists and Statisticians, Friends or Foes?", The Austrian Research Institute for Artificial Intelligence, Vienna, Austria. Technical Report oefai-tr-95-06, 1995.

48. Flexer, A., "Limitations of self-organizing maps for vector quantization and multidimensional scaling", The Austrian Research Institute for Artificial Intelligence, Vienna, Austria. Technical Report oefai-tr-96-23, 1996.

49. Flexer, A., "On the use of self-organizing maps for clustering and visualization", The Austrian Research Institute for Artificial Intelligence, Vienna, Austria. Technical Report oefai-tr-99-04, April 1999.

50. Flexer, A., "Statistical Evaluation of Neural Network Experiments: Minimum Requirements and Current Practice", The Austrian Research Institute for Artificial Intelligence, Vienna, Austria. Technical Report oefai-tr-95-16, 1995.

51. Fuente Gallego, E. S. de la; Bahillo Pérez, D. "Clasificación de campos visuales mediante redes neuronales". Proyecto Fin de Carrera. EUP. Universidad de Valladolid, 1998.

52. Funkhouser, A. T., Fankhauser, F., "A comparison of the mean defect and mean deviation indices within the central 28 degrees of the glaucomatous visual field", Jpn J Ophthalmol 1990; 34(4):414-20.

53. Funkhouser, A. T., Fankhauser, F., Weale, R. A., Problems related to diffuse versus localized loss in the perimetry of glaucomatous visual fields", Graefe's Arch Clin Exp Ophthalmol 230: 243-247, 1992.

54. Gaasterland, D. E., Ederer, F., Sullivan, E. K., Caprioli, J., Cyrlin, M. N., "Advanced Glaucoma Intervention Study. 2. Visual Field Test Scoring and Reliability.", Ophthalmology, Vol. 101, nº 8, Aug. 1994, pp. 1445-1455.

55. Giarratano, J., Riley, G., "Expert Systems. Principles and Programming", Boston, 1998.

56. Goldbaum, M. H., Sample, P. A., White, H., Côté, B., Raphaelian, P., Fechtner, R. D., Weinreb, R. N., "Interpretation of Automated Perimetry for Glaucoma by Neural Network", Investigative Ophthalmology and Visual Science, Aug. 1994, Vol. 35, nº 9, pp. 3362-3373.

57. Gómez, A., Juristo, N., Montes, C., Pazos, J., "Ingeniería del conocimiento", Ed. Ceura, Madrid, 1997.

58. González de la Rosa, M., Martínez Piñero, A., González Hernández, M., "Reproducibility of the TOP algorithm results versus the ones obtained with the bracketing procedure", XIIIth International Perimetric Society Meeting. Gardone Riviera, 1998.

59. González Hernández, M., González de la Rosa, M., "Perimetría de movimiento", Arch. Soc. Canar. Oftal., 1999 – nº 10.

60. Greenfield, R. H., Kass, M. A., Livingston, J. P., "A Computerized Glaucoma Data Base", Arch Ophthalmol, Vol 95, Aug 1977.

61. Guo, A. K., Yang, X. Y., "Neural Network approaches to visual motion perception", SCI CHINA B 1994 Feb; 37(2):177-89.

62. Gutte, C., "Note on Free Funches and Cross Validation", Department of Mathematical Modelling, building 321. Technical University of Denmark, DK-2800 Lyngby, Denmark, 1997.

63. Hansen, B. K., “Weather Prediction Using Case-Based Reasoning and Fuzzy Set Theory”, M. Sc., Dalhousie University – Daltech, Halifax, Nova Scotia 2000.

64. Hart, W. M., Hartz, R. K., "Computer Processing of Visual Field Data. I. Recording, Storage, and Retrieval", Arch Ophthalmol, Vol 99, Jan 1981.

65. Hart, W. M., "Computer Processing of Visual Data. II. Automated Pattern Analysis of Glaucomatous Visual Fields", Arch Ophthalmol, Vol 99, Jan 1981.

66. Hart, W. M., Becker, B., “The onset and evolution of glaucomatous visual field defects”, Ophthalmology 89:268-279, 1982.

67. Haykin, S., "Neural Networks. A Comprehensive Foundation", PH, 1994.

68. Heijl, A., Lundqvist, L., “The frequency distribution of earliest glaucomatous visual field defects documented by automatic perimetry”, Acta Ophthalmol 62, pp. 658-664, 1984.

69. Heijl, A., Lindgren, G., Olsson, J., "Normal Variability of Static Perimetric Threshold Values Across the Central Visual Field", Arch Ophthalmol, Vol 105, Nov 1987, pp. 1544-1549.

70. Heijl, A., Lindgren, G., Olsson, J., Åsman, P., “Visual Field Interpretation With Empiric Probability Maps”, Arch Ophthalmol, Vol 107, Feb 1989.

71. Henson, D. B., Darling, M. N., "Detecting progressive visual field loss", Ophthalmic Physiol Opt 1995 Sep; 15(5):387-90.

72. Henson, D. B., Spenceley, S. E., Bull D. R., "Artificial neural network analysis of noisy visual field data in glaucoma", Artif Intell Med 1997 Jun; 10(2):99-113.

73. Henson, D. B., Spenceley, S. E., Bull, D. R., "Spatial Classification of Glaucomatous Visual Field Loss", British Journal of Ophthalmology, 1996; 80:526-531.

74. Hernández Merayo, D., “Sistema borroso híbrido simbólico-conexionista para el diagnóstico del glaucoma”, Proyecto Fin de Carrera. ETSI Informática. Universidad de Valladolid, 2002.

75. Hintz-Madsen, M., Hansen, L. K., Larsen, J., "Detection of Malignant Melanoma using Neural Classifiers", Proceedings of International Conference on Engineering Applications of Neural Networks, EANN'96.

76. Hirsbrunner, H. P., Fankhauser, F., Funkhouser, A. T., Jenni, A., "Evaluating human and automated interpretation of visual field data in perimetry", Jpn J Ophthalmol, Vol 34: 72-80, 1990.

77. Hirsbrunner, H. P., Fankhauser, F., Jenni, A., Funkhouser, A., "Evaluating a perimetric expert system: experience with Octosmart", Graefe's Arch Clin Exp Ophthalmol 228: 237-241, 1990.

78. Honkela, T., Pulkki, V., Kohonen, T., "Contextual Relations of Words in Grimm Tales, Analyzed by Self-Organizing Map.", ICANN, Paris, 1995, 3-7.

79. Iester, M., Swindale, N.V., Mikelberg, F.S., "Sector-based analysis of optic nerve head shape parameters and visual field indices in healthy and glaucomatous eyes", J Glaucoma, 1997, Dec; 6 (6):370-6.

80. Jain, A. K., Mao, J., Mohiuddin, K. M., "Artificial Neural Networks: A Tutorial", IEEE Computer, Marzo, 1996.

81. Jang, J.-S. R., Sun, C.-T., Mizutani, E., "Neuro-Fuzzy and Soft Computing. A Computational Approach to Learning and Machine Intelligence.", Matlab Curriculum Series, Prentice Hall, 1997.

82. John, R. I., "Fuzzy Inferencing Systems. Problems and Some Solutions", Computing Sciences Research Working Paper No. 62, School of Computing Sciences, De Montfort University, The Gateway, Leicester. December 1995.

83. Kaski, S., Honkela, T., Lagus, K., Kohonen, T., "Creating an Order in Digital Libraries with Self Organizing Maps", World Congress on Neural Networks, Laurence Erlbaum and INNS Press, Mahwah, MJ, 1996, 814-817.

84. Katz, J., Sommer , A., Gaasterland, D. E., Anderson, D. R., "Comparison of analytic algorithms for detecting glaucomatous visual field loss", Arch Ophthalmol 1991 Dec; 109(12):1684-9.

85. Kohavi, R., "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection", International Joint Conference on Artificial Intelligence, 1995.

86. Kohonen, T., "Self-Organization and Associative Memory", SV, 1989.

87. Kohonen, T., Hyaninen, J., Kangas, J., Laaksomen, J., Torkkola, K., "LVQ_PAK. The Learning Vector Quantization Program Package V3.1", LVQ Programming Team of the Helsinki University, 1995.

88. Kohonen, T., Kangas, J., Laaksomen, J., "SOM_PAK. The Self-Organizing Map Program Package V3.1", SOM Programming Team of the Helsinki University, 1995.

89. Kolodner, J., "Case-Based Reasoning", Morgan Kaufmann Publishers, 1993.

90. Krakau, C. E. T., "Artificial intelligence in computerized perimetry", Doc Ophthalmol Proc Series 49, 1987, pp. 169-174.

91. Kröse, B., Smagt, P., "An Introduction to Neural Networks", The University of Amsterdam, 1996.

92. Lawrence, S., Tsoi, A. C., Back, A. D., "Function Approximation with Neural Networks and Local Methods: Bias, Variance and Smoothness", Australian Conference on Neural Networks, ACNN 96, Bartlett, P., Burkitt, A., Williamson, R. Eds. Australian Natinal University, pp. 16-21, 1996.

93. Lietman, T., Eng, J., Katz, J., Quigley, H. A., "Neural Networks for Visual Field Analysis: How Do They Compare with Other Algorithms?", Journal of Glaucoma 8:77-80, 1999.

94. Liu, X., Cheng, G., Wu, J. X., "Identifying the measurement noise in glaucomatous testing: An artificial neural network approach", Artificial Intelligence in Medicine 6:401-416, 1994.

95. Llamas Bello, César, "Análisis Comparativo de Tres Técnicas de Reconocimiento de Patrones Aplicado a un Problema de Reconocimiento de Habla", Departamento de Informática. Universidad de Valladolid (España), 1995.

96. Leake, D. B., "CBR in Context: The present and Future.", Case-Based Reasoning: Experiences, Lessons, and Future Directions. Menlo Park: AAAI Press. 1996.

97. López de Mántaras, R., Plaza, E., "Case-Based Reasoning: An Overview", AI Communications Journal, 10(1), 21-29, 1997.

98. López de Ullibarri Galparsoro, I., Pita Fernández, S., "Metodología de la Investigación. Curvas ROC.", Cad Aten Primaria, 1998; 5 (4): 229-235.

99. Long, W. J., "The Probability of Disease", Proceedings of the Fifteenth Annual Symposium on Computer Applications in Medical Care, pp 619-623, 1991.

100. Luger, G. F., Stubblefield, W. A., "Artificial Intelligence. Structures and Strategies for Complex Problem Solving.", Addison-Wesley, Third edition, 1998.

101. Madsen, E. M., Yolton, R. L., "Demonstration of a neural network expert system for recognition of glaucomatous visual field changes", Mil Med 1994 Aug; 159(8):553-7.

102. Mandava, S., Caprioli, J., Zulauf, M., "Glaucoma Pattern Index to Quantify Glaucomatous Visual Field Loss", Journal of Glaucoma, Vol. 1, Nº 3, pp. 178-183, 1992.

103. Mandava, S., Zulauf, M., Zeyen, T., Caprioli, J., "An evaluation of clusters in the glaucomatous visual field", Am J Ophthalmol, 1993 Dec 15; 116(6):684-91.

104. Maquet, J. A., Pastor, J. C., Fuente, L. F. de la, “El glaucoma crónico simple. Guía para su comprensión.”, IOBA, Universidad de Valladolid, 1993.

105. Marraffa, M., Marchini, G., Albertini, R., Bonomi L., "Comparison of different screening methods for the detection of visual field defects in early", Int Ophthalmol 1989 Jan; 13(1-2):43-5.

106. Martin-Boglind, L. M., Wanger, P., “Computer-assisted evaluation of the results from high-pass resolution perimetry: a knowledge-based system”, Perimetry Update 1990/91, Proceedings of the IXth International Perimetric Society Meeting, edited by R. P. Mills and A. Heijl, pp. 297-301, 1991.

107. Martin-Boglind, L., “Computer-assisted interpretation of resolution visual fields. Basic studies and program development.” , Kongl Carolinska Medico Chirurgiska Institutet, Stockholm 1993.

108. Martin, L., “Computer-Assisted Management of Primary Open-Angle Glaucoma. Knowledge Acquisition and Prototype Testing.”, Computers in Nursing, September-October, 1996, Vol. 14, Nº 5, pp. 267-271.

109. Martin, L., “Share Care and Computer Assistance in Glaucoma Management”, Nursing Informatics, U. Gerdin y col. (Eds.) IOS Press 1997, pp. 288-290.

110. Martin, L., “Knowledge acquisition and evaluation of an expert system for managing disorders of the outer eye”, Computers in Nursing, 2001.

111. Martín del Brío, B., Sanz Molina, A., "Redes Neuronales y Sistemas Borrosos", Ed. Ra-Ma Textos Universitarios, Madrid 1997.

112. Martínez Piñero, A., Pareja Ríos, A., Mesa Moreno, C., González de la Rosa, M., “Reproducibility and accuracy of the top perimetry”, Arch Soc Esp Oftalmol nº 6, Jun 1999.

113. Metz, H. S., Madden, E. E., Williams, V. R., Vaughan, C. M., "Use of a Computer in a Glaucoma Clinic", Arch Ophthal, Vol 81, Feb 1969.

114. Miglior, S., Brigatti, L., Lonati, C., Rossetti, L., Pierrottet, C., Orzalesi, N., "Correlation between the progression of optic disc and visual field changes in glaucoma", Curr Eye Res 1996 Feb; 15(2):145-9.

115. Mikelberg, F. S., Drance, S. M., “The mode of progression of visual field defects in glaucoma”, American Journal of Ophthalmology 98:443-445, 1984.

116. Mizoguchi, F., Ohwada, H., Daidoji, M., Shirato, S., "Learning rules that classify ocular fundus images for glaucoma diagnosis", Inductive Logic Programming Workshop 1996: 146-159. Lecture Notes in Computer Science, vol. 1314, Springer, 1997.

117. Moriarty, D. E., Miikkulainen, R., “Hierarchical Evolution of Neural Networks”, Proceedings of the 1998 IEEE Conference on Evolutionary Computation, Piscataway, NJ.

118. Moro, Q. I., Alonso, L., "Reviewing Some Validation Methods for Classifiers", ECLA005, 1996.

119. Moro, Q. I., Alonso, L., "Una contribución a la Predicción Meteorológica basada en Técnicas Conexionistas", I Taller AIRENE sobre Reconocimiento de Patrones con Redes Neuronales, Universidad Católica del Norte, Antofagasta, Chile, Mayo 1999.

120. Moro, Q. I., “Una Aportación a la Predicción Meteorológica Basada en Técnicas Conexionistas”. Tesis Doctoral. Universidad de Valladolid, 2000.

121. Mutlukan, E., Keating, D., "Visual field interpretation with a personal computer based neural network", Eye 1994; 8 (3):321-3.

122. Nilsson, N. J., “Inteligencia Artificial. Una nueva síntesis.”, Mc Graw-Hill, Madrid, 2000.

123. Orchard, R. A., “FuzzyCLIPS Version 6.04A. User’s Guide.” Integrated Reasoning. Institute for Information Technology. National Research Council, Canadá, 1998.

124. Paycha, F., Paycha, A., "Un programme d'aide au diagnostic par identification (PADI), système-expert pour les hypertonies acquises de l'adulte", J. Fr. Ophthalmol., 1986, 9, 6-7, 485-491.

125. Peña, D., "Estadística Moderna y Métodos", Editorial Alianza, 1994.

126. Perera, R. S., “Case-Based Estimating and Design: towards a colloborative approach”, Ph. D., T.I.M.E. Research Institute, Department of Surveying, University of Salford, U.K., August 1997.

127. Pradines, F., Delbosc, B., Royer, J., "Statistical analysis of visual fields by the automatic perimeter Octopus 2000 R with the delta program", J Fr Ophthalmol 1985; 8(8-9):541-8.

128. Reyes de los Mozos, Mario, “SoftComputing & Aplicaciones Biomédicas. Estudio del Glaucoma en estado incipiente”, Tesis doctoral. Universidad Autónoma de Barcelona, 1998.

129. Riesbeck, C. K., Shank, R. C., “Inside CaseBased Reasoning”, Lawrence Erlbaum Associates, Hillsdale, NJ, US, 1989.

130. Ríos, S., "Métodos Estadísticos", Segunda Edición, Ediciones del Castillo S.A., 1983.

131. Rodríguez Martín, F.J., Martínez Piñero, A., Abraldes López-Veiga, M., "Perimetría Top-Flicker en pacientes normales, hipertensos oculares y glaucomatosos iniciales", Capítulo 3, 1998.

132. Rolston, D. W., "Principios de Inteligencia Artificial y Sistemas Expertos", Mc Graw Hill, Colombia, 1990.

133. Ronco, E., Gawthrop, P., "Modular Neural Networks: a state of the art", Centre for System and Control, University of Glasgow, UK, Technical Report: CSC-95026, May 1995.

134. Rouleau, J., Paul, C., Robert, J. L., Sole, P., "Système expert appliqué au diagnostic et au traitement du glaucome primitif à angle ouvert (GPAO)", Ophtalmologie 1991; 5: 127-128.

135. Sánchez García, M., "Modelos Estadísticos Aplicados al Tratamiento de Datos", Universidad Complutense - Madrid, 1978.

136. Sarle, W. S., "Frequently Asked Questions", ftp://ftp.sas.com/pub/neural, 2002.

137. Sarle, W. S., "Neural Networks and Statistical Models", Proceedings of Nineteenth Annual SAS Users Group International Conference, April, 1994.

138. Schwartz, B., Nagin, P., "Probability maps for evaluating automated visual fields", Doc Ophthalmol Proc Ser 1985, 42:39-48.

139. Shapiro, L. R., Johnson, C. A., "Evaluation of the diagnosis of visual fields". In Heijl A., Mills R. P. (eds): Perimetry Update 1990/1991, Kugler Publications, Amsterdam 1991, pp. 281-285.

140. Siler, W., "Building Fuzzy Expert Systems", Birmingham, AL 35217, USA, http://members.aol.com/wsiler.

141. Simón, M. A., "Desarrollo de un sistema experto para el diagnóstico del glaucoma". Trabajo de Licenciatura. Universidad de Valladolid, 1990.

142. Simón, M. A., de la Fuente, L. F., "Sistema experto Glaucom-Easy v 1.0." Universidad de Valladolid, 1993.

143. Simón, A., Alonso, L., Cabana, J. J., Antón, A., Maquet, J., Pastor, J. C., "An Expert System for the Diagnosis of Glaucoma: GLAUCOM-EASY", Proceedings of the Thirteenth IASTED International Conference APPLIED INFORMATICS, Austria, February 21-23, 1995. Editor: M. H. Hamza. Publication Code: 230. pp. 133-135.

144. Simón, A., Alonso, L., Bahillo, D., de la Fuente, E., Antón, A., "Clasificación de campos visuales mediante redes neuronales", Actas del I TALLER AIRENE sobre Reconocimiento de Patrones con Redes Neuronales, Antofagasta, Chile, 24-28 de Mayo de 1999, pp. 90-94.

145. Sommer A., Enger, C., Witt, K., "Screening for glaucomatous visual field loss with automated threshold perimetry", Am J Ophthalmol 1987 May 15; 103(5):681-4.

146. Spenceley, S. E., Henson, D. B., Bull, D. R., "Visual field analysis using artificial neural networks", Ophthalmic Physiol Opt 1994 Jul; 14(3):239-48.

147. Sponsel, W. E., Hobley, A. J., Williams, A. H., Dallas, N. L., "Glaucoma Assessment by Microcomputer", Trans. Ophthalmol. Soc. U.K., 1984, 104:100-105.

148. Stevens, R. H., Najafi, K., "Can artificial neural networks provide an expert's view of medical students performances on computer based simulation?", Proc Annu Symp Comput Appl Med Care 1992:179-83.

149. Suzuki, Y, Araie, M., Ohashi, Y., "Sectorization of the central 30 degrees visual field in glaucoma", Ophthalmology 1993 Jan; 100(1):69-75.

150. Szolovits, P., Pauker, S. G., "Categorical and Probabilistic Reasoning in Medicine Revisited", Artificial Intelligence, 59:167-180, 1993.

151. Takeda, T., Kishi, K., Yamanouchi, T., Mizoe, H., Matsuoka, T., "Significance of distributed representation in the output layer of a neural network in a pattern recognition task", Med Biol Eng Comput 1994 Jan; 32(1):77-84.

152. Togneri, R., Farrokhi, D., Zhang, Y., Attikiouzel, Y., "A comparison of the LBG, LVQ, MLP, SOM and GMM algorithms for vector quantization and clustering analysis", Proceedings of The Fourth Australian International Conference on SST-92, Brisbane, 1-3 Dec 1992, pp. 173-177.

153. Tumer, K., Ramanujam, N., Ghosh, J., Richards-Kortum, R., "Ensembles of Radial Basis Function Networks for Spectroscopic Detection of Cervical Precancer", IEEE Transactions on Biomedical Engineering, Vol. 45, Nº 8, August 1998: 953-961.

154. Uchida, H., Brigatti, L., Caprioli, J., "Detection of Structural Damage From Glaucoma With Confocal Laser Image Analysis", Investigative Ophthalmology and Visual Science, Nov. 1996, Vol. 37, nº 12, pp. 2393-2401.

155. Ulieru, M., Cuzzani, O., Rubin, S. H., Ceruti, M. G., "Application of Soft Computing Methods to the Diagnosis and Prediction of Glaucoma", 2000 IEEE International Conference on Systems, Man and Cybernetics, SMC 2000, "Cybernetics Evolving to Systems, Humans, Organizations, and their Complex Interactions," October 8-11, 2000, Nashville, TN, USA, pp. 3641-3645.

156. van der Weg, M. C., "Prediction of the Water Level During Storm Situations using Neural Networks", Department of Computer Science, Leiden University, 1997.

157. Viswanathan, A. C., McNaught, A. I., Poinoosawmy, D., Fontana, L., Crabb, D. P., Fitzke, F. W., Hitchings, R. A., "Correlation of patient perception of glaucomatous visual loss and automated perimetry (1): A cross-sectional and longitudinal study", Vth Meeting of the European Glaucoma Society, Paris 1996.

158. Walter, J., Ritter, H., Schulten, K., "Non-linear Prediction with Self-organizing Maps", Proc. Int. Joint Conf. On Neural Networks, (1):587-592, San Diego, Jun. 1990.

159. Walter, J., Ritter, H., "Local PSOMs and Chebyshev PSOMs Improving the Parametrised Self-Organizing Maps", Proc Int. Conf. on Artificial Neural Networks (ICANN-95), Paris, vol. 1, pp. 95-102, 1995.

160. Walter, J. A., "PSOM: Network: Learning with Few Examples", Proc. Int. Conf. on Robotics and Automation (ICRA) IEEE 1998, pp 2054-2059.

161. Watson, I., "Applying Case-Based Reasoning: Techniques for Enterprise Systems", Morgan Kaufmann Publishers, 1997.

162. Weijters, A., van den Bosch, A., van den Herik. H.J., "Behavioural Aspects of Combining Backpropagation Learning and Self-organizing Maps", Connection Science, 9:3, pp 235-252, 1997.

163. Weijters, T., van den Bosch, A., van den Herik, H. J., "Interpretable Neural Networks with BP-SOM", Proceedings of 10th European Conference on Machine Learning, Chemnitz, Germany. Lecture Notes in Computer Science, vol. 1398. Springer, 1998.

164. Weijters, T., van den Herik, H. J., van den Bosch, A., Postma, E., "Avoiding Overfitting with BP-SOM", 15th International Joint Conference on Artificial Inteligence, 1997, 1140-1145.

165. Weiss, S. M., "A system for model-based computer-aided diagnosis and therapy", Rutgers University. The State University of New Jersey, Ph. D., 1974.

166. Weiss, S. M., Kulikowski, C. A., Amarel, S., Safir, A., "A Model-Based Method for Computer-Aided Medical Decision-Making", Artificial Intelligence 11, 1978, pp. 145-172.

167. Weiss, S., Kulikowski, C. A., Safir, A., "Glaucoma Consultation by Computer", Comput Biol Med, 1978, Vol. 8, pp. 25-40.

168. Wise, B. M., Holt, B. R., Gallagher, N. B., Lee, S, "A Comparison of Neural Networks, Non-Linear Biased Regression and Genetic Algorithm For Dynamic Model Identification", 1996.

169. Wu, D. C., Schwartz, B., Nagin, P., "Trend analyses of automated visual fields", Doc Ophthalmol Proc Series 49, 1987, pp. 175-189.

170. Zadeh, L., "Fuzzy Sets", Information and Control, vol. 8, pp. 338-353, 1965.

171. Zahlmann, G., Obermaier, M., Ritzke, C., Scherf, M., "Knowledge-based monitoring of glaucoma patients - a connectionist's approach", GSF, Medis-institute, Neuherberg, Germany, Dept. of Ophthalmology, Techn. University of Munich, 1996. Proceedings MIE96, 19-22.8.96 Kopenhagen.

172. Zahlmann, G., Schubert, M., Obermaier, M., Mann, G., "Concept of a knowledge based monitoring system for glaucoma and diabetic retinopathy using a telemedicine approach", Proc. of the 13th Annual International Conf. IEEE-EMBS, IEEE Press 1996, 111pp.

173. Zahlmann, G., Scherf, M., Wegner, A., "A neuro-fuzzy-classifier for a knowledge-based glaucoma monitor", AIME 97, Lecture Notes in Artificial Intelligence 1211, Keravnou et al. Artificial Intelligence in Medicine, Springer Verlag, 273-284.

174. Zahlmann, G., Scherf, M., Wegner, A., "NeuroFuzzy and EUBAFES as tools for Knowledge discovery in visual field data", 20th Annual Int. Conference on the IEEE EMBS, 29.10-1.11.1998 Hong Kong.

175. Zahlmann, G., Kochner, B., Ugi, I., Schuhmann, D., Liesenfeld, B., Wegner, A., Obermaier, M., Mertz, M., "Hybrid Fuzzy Image Processing for Situation Assessment. A Knowledge-Based System for Early Detection of Diabetic Retinopathy.", IEEE Engineering in Medicine and Biology, January/February 2000, pp. 76-83.

176. Zahlmann, G., Scherf, M., Wegner, A., Obermaier, M., Mertz, M., "Situation Assessment of Glaucoma Using a Hybrid Fuzzy Neural Network. A Differential Diagnostic-Decision Support System to Aid in Early Detection in Primary-Care Environments.", IEEE Engineering in Medicine and Biology, January/February 2000, pp. 84-91.

177. Zeyen, T. G., Caprioli, J., "Progression of disc and field damage in early glaucoma", Arch Ophthalmol 1993 jan; 111(1):62-5.

178. Zulauf, M., LeBlanc, R. P., Flammer, J., "Normal visual fields measured with Octopus-Program G1. I. Differential light sensitivity at individual test locations.", Graefe's Arch Clin Exp Ophthalmol, 232:509-515, 1994.

179. Zulauf, M., LeBlanc, R. P., Flammer, J., "Normal visual fields measured with Octopus-Program G1. II. Global visual field indices.", Graefe's Arch Clin Exp Ophthalmol, 232:516-522, 1994.

Índice de Figuras

Figura 2.1: Esquema en el que se muestra la forma de crecimiento más frecuente de la excavación glaucomatosa [104]. 12
Figura 2.2: Escotomas aislados en el área de Bjerrum [104]. 13
Figura 2.3: Escotomas arciformes paracentrales, asociados a estrechamientos periféricos [104]. 13
Figura 2.4: Campo biinsular [104]. 13
Figura 2.5: Lucerna temporal [104]. 13
Figura 2.6: CASNET. Los cuatro niveles de la red semántica [165, 166 y 167]. 17
Figura 2.7: Zonas del campo visual. 19
Figura 2.8: Actual (dB). Tabla numérica del programa G1 del Octopus 500 [104]. 23
Figura 2.9: Diferencia (dB). Tabla corregida según la edad del programa G1 del Octopus 500 [104]. 23
Figura 2.10: Tablas de datos obtenidas con el programa central 24-2 en el analizador de campos visuales Humphrey [1]. 24
Figura 2.11: Curva de características operativas de la regla de clasificación desarrollada a partir de los 59 puntos del CV y evaluada con la muestra independiente. 32
Figura 2.12: Curva de características operativas de la regla de clasificación desarrollada a partir de los 59 puntos del CV, MD y LV, evaluada con la muestra independiente. 33
Figura 2.13: Curva de características operativas de la regla de clasificación desarrollada a partir de las 7 zonas del CV y evaluada con la muestra independiente. 33
Figura 2.14: Curva de características operativas de la regla de clasificación desarrollada a partir de las 7 zonas del CV, el MD y la LV, evaluada con la muestra independiente. 34
Figura 2.15: Curva de características operativas de la regla de clasificación desarrollada a partir de las 7 zonas del CV, teniendo en cuenta el valor de los defectos de sensibilidad, y evaluada con la muestra independiente. 35
Figura 2.16: Siete zonas de Schwartz [138]. 36

Figura 2.17: Sectores de Asman [10]. 37
Figura 2.18: El mapa de dimensión 5x5 de una red de Kohonen entrenada con 560 campos visuales glaucomatosos y normales. Los datos de entrenamiento contienen los defectos de sensibilidad de los 25 puntos del hemicampo superior examinados por el perímetro Humphrey. Cada escala de grises representa los pesos para cada nodo sobre una escala que va de 0 a 100 organizada espacialmente para representar el campo visual [73]. 46
Figura 2.19: Red jerárquica RBF para datos de perimetría [173]. 55
Figura 3.1: Grafo de decisión 1. 63
Figura 3.2: Grafo de decisión 2. 64
Figura 3.3: Grafo de decisión 3. 65
Figura 3.4: Grafo de decisión 4. 65
Figura 3.5: Componentes de un Sistema Experto [122]. 66
Figura 3.6: Etapas en que se realiza el diagnóstico. 70
Figura 3.7: Factor de certeza según la PIO. 72
Figura 3.8: Factores de certeza según la E/P. 72
Figura 3.9: Entradas y salidas del SE GLAUCOM-EASY. 76
Figura 3.10: Grafo de decisión 1. GLAUCOM-EASY. 77
Figura 3.11: Grafo de decisión 2. GLAUCOM-EASY. 78
Figura 3.12: Diseño arquitectónico del Sistema Experto Glaucom-Easy. 80
Figura 3.13: Datos solicitados por el Glaucom-Easy 81
Figura 3.14: Criterios propuestos por el Glaucom-Easy para calificar el campo visual como normal, dudoso o patológico. 82
Figura 3.15: Factor de certeza según la PIO. Glaucom-Easy. 84
Figura 3.16: Factor de certeza según la E/P. Glaucom-Easy. 85
Figura 3.17: Variación de la sensibilidad y especificidad si se modificara el riesgo de glaucoma a partir del cual el Glaucom-Easy indica la necesidad de tratamiento o visita al especialista. 88
Figura 4.1: Número asignado a cada punto del CV. 99
Figura 4.2: Distribución del campo visual en zonas. 100
Figura 4.3: Vecindad tipo "burbuja" en un SOM. 102
Figura 4.4: Vecindad tipo "Gaussiana" en un SOM. 102
Figura 4.5: Topología rectangular (a) y hexagonal (b) de un SOM. 103
Figura 4.6: Ventana de GESL para generar una batería de entrenamientos SOM 105
Figura 4.7: Mapa SOM para clasificar en 5 categorías. La clase Diabetes no está representada. 108
Figura 4.8: Mapa SOM para clasificar en 5 categorías. Hay varias neuronas sin etiquetar. 109
Figura 4.9: Comparación de los experimentos SOM, realizados con los defectos de sensibilidad de los 59 puntos del campo visual o con las medias de los defectos de sensibilidad de las siete zonas del campo visual, mediante la Precisión Diagnóstica. 113
Figura 4.10: El ciclo CBR [3]. 114
Figura 4.11: Diagrama de flujo de CBR [126]. 115
Figura 4.12: Tareas de aplicación CBR [126]. 119

Figura 4.13: Comparación de los experimentos CBR para clasificar en cinco categorías. 125
Figura 4.14: Comparación de la Precisión Diagnóstica en los experimentos CBR realizados para clasificar en cinco categorías. 125
Figura 4.15: Comparación de los experimentos CBR para clasificar en tres categorías. 127
Figura 4.16: Comparación de la Precisión Diagnóstica en los experimentos CBR realizados para clasificar en tres categorías. 128
Figura 4.17: Comparación de los experimentos CBR para clasificar en dos categorías. 129
Figura 4.18: Comparación de la Precisión Diagnóstica en los experimentos CBR realizados para clasificar en dos categorías. 130
Figura 4.19: Comparación de los mejores experimentos CBR, realizados con los 59 puntos del campo visual o con las medias de los defectos de las siete zonas del campo visual, mediante la Precisión Diagnóstica. 130
Figura 4.20: Comparación de los mejores experimentos CBR, realizados con los 59 puntos del campo visual mediante la Precisión Diagnóstica, utilizando recuperación de casos por inducción o por el vecino más próximo. 131
Figura 4.21: Comparación de los mejores experimentos CBR, realizados con las medias de los defectos de las siete zonas del campo visual mediante la Precisión Diagnóstica, utilizando recuperación de casos por inducción o por el vecino más próximo. 131
Figura 4.22: Comparación de los mejores experimentos SOM y CBR, realizados con los 59 puntos del campo visual mediante la Precisión Diagnóstica. 134
Figura 4.23: Comparación de los mejores experimentos SOM y CBR, realizados con las medias de los defectos de las siete zonas del campo visual mediante la Precisión Diagnóstica. 134
Figura 5.1: Términos relativos a los conjuntos borrosos [111]. 141
Figura 5.2: Función trapezoidal. 142
Figura 5.3: Función de tipo singleton. 142
Figura 5.4: Función triangular. 142
Figura 5.5: Función de tipo S. 143
Figura 5.6: Función de tipo π. 143
Figura 5.7: Ejemplo de conjuntos borrosos para la variable Estatura. 144
Figura 5.8: Ejemplo de aplicación de operaciones básicas sobre conjuntos borrosos. 145
Figura 5.9: Interpretación gráfica del GMP usando la implicación borrosa de Mamdani y la composición max-min. 148
Figura 5.10: Un sistema de inferencia borrosa tipo Mamdani con dos entradas, una nítida y otra borrosa, dos reglas con dos antecedentes cada una y una salida borrosa, usando min y max como T-norma y T-conorma respectivamente. 149
Figura 5.11: Un sistema de inferencia borrosa tipo Sugeno con dos entradas nítidas, y dos reglas con dos antecedentes cada una, usando min

o producto como T-norma y calculando la media ponderada de la salida de cada regla para obtener la salida final. 155
Figura 5.12: Entradas y salidas del Sistema Experto Borroso. 161
Figura 5.13: Frecuencia de los valores de la E/P en la población normal. 162
Figura 5.14: Frecuencia de los valores de la E/P en la población glaucomatosa. 162
Figura 5.15: Frecuencia de los valores de la diferencia de la E/P en la población normal. 163
Figura 5.16: Diseño arquitectónico del Sistema híbrido de ayuda al diagnóstico del glaucoma. 166
Figura 5.17: Ventana principal de la aplicación [74]. 167
Figura 5.18: Ventana en la que se recogen todos los datos del ojo derecho [74]. 167
Figura 5.19: Diseño arquitectónico del Clasificador del Campo Visual. 169
Figura 5.20: Cantidad mínima de depresión, en decibelios, que identifica localizaciones del campo como defectuosas en la prueba umbral 24-2 del Analizador de campo visual Humphrey. 170
Figura 5.21: Zonas del campo visual. 171
Figura 5.22: Número asignado a cada punto del campo visual. 172
Figura 5.23: Mapa SOM que clasifica en dos categorías: Glaucoma y no glaucoma. 176
Figura 5.24: Mapa SOM en el que se han modificado algunas etiquetas para clasificar campos dudosos además de glaucomatosos y no glaucomatosos. 176
Figura 5.25: Diseño arquitectónico del Sistema Experto Borroso. 181
Figura B.1: Ejemplo de distribución de los resultados de un test. 198
Figura B.2: Curvas ROC. Comparación. 198

Índice de Tablas

Tabla 2.1: Estudio comparativo del programa G1 del perímetro Octopus con otras estrategias de perimetría computerizada (Marrafa [105]). 27
Tabla 2.2: Estudio comparativo de la capacidad de varios índices de discriminar campos normales de glaucomatosos incipientes (Mandava [102]). 29
Tabla 2.3: Sensibilidad (S) y especificidad (E) de las ecuaciones de regresión a la hora de clasificar los CV como glaucomatosos (GL), normales (NOR), con catarata (CAT) o diabéticos o hipertensos (D+H). 31
Tabla 2.4: Validez diagnóstica de las reglas de clasificación, obtenidas mediante regresión logística, según los distintos tipos de datos de entrada. Estos datos se consiguieron teniendo en cuenta la afectación o normalidad en la sensibilidad de cada punto. 32
Tabla 2.5: Validez diagnóstica del algoritmo obtenido mediante regresión logística, agrupando los puntos en 7 zonas, teniendo en cuenta el valor cuantitativo del defecto en cada punto. 34
Tabla 2.6: Resultados obtenidos por Goldbaum [56] en la clasificación de campos visuales. 43
Tabla 2.7: Resultados de sensibilidad y especificidad obtenidos en la clasificación de campos visuales en tres categorías utilizando la red $7 \rightarrow 4 \rightarrow 3$. 48
Tabla 2.8: Resultados de sensibilidad y especificidad obtenidos en la clasificación de campos visuales en dos categorías: glaucoma y no glaucoma. 48
Tabla 2.9: Resultados de sensibilidad y especificidad obtenidos en la clasificación de campos visuales entrenando redes de forma independiente. 49
Tabla 2.10: Resultados obtenidos por Zahlmann [171, 172] en la clasificación de campos visuales utilizando MLP con BP. 54
Tabla 2.11: Resultados de la red especializada. 56
Tabla 2.12: Resultados de la clasificación de la red RBF (entrenamiento). 57
Tabla 2.13: Resultados de la clasificación de la red RBF (prueba). 58

Tabla 2.14: Sensibilidad y especificidad de la estimación final de la situación. 58
Tabla 3.1: Distribución de los pacientes según el diagnóstico previo y el riesgo de padecer glaucoma calculado por el Glaucom-Easy. En el Grupo 1 se incluyeron los pacientes diagnosticados de glaucoma y en el Grupo 2 los sanos o con distinta patología. 86
Tabla 3.2: Distribución de los pacientes según su diagnóstico previo y el resultado obtenido con el Glaucom-Easy. 87
Tabla 3.3: Valores de sensibilidad, especificidad, falsos positivos, falsos negativos, valor predictivo positivo y valor predictivo negativo, según el valor del riesgo que el Glaucom-Easy considera indicación de tratamiento o visita al especialista. 87
Tabla 4.1: Distribución inicial de casos según clases. 95
Tabla 4.2: Distribución final de casos según clases. 96
Tabla 4.3: Resultados obtenidos en la clasificación del campo visual en cinco categorías utilizando el SOM_PACK [88]. 106
Tabla 4.4: Resultados obtenidos en la clasificación del campo visual en dos categorías utilizando el SOM_PACK [88]. 106
Tabla 4.5: Los dos mejores mapas SOM que clasifican en glaucoma y no glaucoma entrenados con los defectos de sensibilidad de los 59 puntos del campo visual. 110
Tabla 4.6: El mejor mapa SOM entrenado con los defectos de sensibilidad de los 59 puntos del campo visual que clasifica en tres categorías: Glaucoma, Normal y Otras patologías distintas de Glaucoma. 110
Tabla 4.7: El mejor mapa SOM entrenado con los defectos de sensibilidad de los 59 puntos del campo visual que clasifica en cinco categorías: Glaucoma, Normal, Cataratas, Diabetes y HTA. 111
Tabla 4.8: El mejor mapa SOM que clasifica en glaucoma y no glaucoma entrenado con las medias de los defectos de sensibilidad de las siete zonas del campo visual. 111
Tabla 4.9: El mejor mapa SOM que clasifica en tres categorías: Glaucoma, Normal y Otras patologías distintas del Glaucoma, entrenado con las medias de los defectos de sensibilidad de las siete zonas del campo visual. 111
Tabla 4.10: El mejor mapa SOM que clasifica en cinco categorías: Glaucoma, Normal, Cataratas, Diabetes y HTA, entrenado con las medias de los defectos de sensibilidad de las siete zonas del campo visual. 112
Tabla 4.11: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método de recuperación de casos por inducción. Clasificación en cinco categorías. 122
Tabla 4.12: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en cinco categorías. 123

Tabla 4.13: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método de recuperación de casos por inducción. Clasificación en cinco categorías. 123
Tabla 4.14: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en cinco categorías. 123
Tabla 4.15: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Las características tienen importancias 8, 4, 2, 0 según la zona del campo visual. Clasificación en cinco categorías. 123
Tabla 4.16: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Las características tienen importancias 16, 8, 4, 2 según la zona del campo visual. Clasificación en cinco categorías. 124
Tabla 4.17: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las desviaciones estándar de los defectos de las siete zonas del campo visual y el método de inducción tanto con búsqueda primero en anchura como primero en profundidad. Clasificación en cinco categorías. 124
Tabla 4.18: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias y las desviaciones estándar de los defectos de las siete zonas del campo visual y el método de inducción con búsqueda primero en anchura. Clasificación en cinco categorías. 124
Tabla 4.19: Comparación de la sensibilidad y precisión diagnóstica de los experimentos CBR realizados para clasificar en cinco categorías. 124
Tabla 4.20: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método de recuperación de casos por inducción. Clasificación en tres categorías. 126
Tabla 4.21: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en tres categorías. 126
Tabla 4.22: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método de recuperación de casos por inducción. Clasificación en tres categorías. 126
Tabla 4.23: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Todas las

características tienen la misma importancia. Clasificación en tres categorías. 127
Tabla 4.24: Comparación de la sensibilidad y precisión diagnóstica de los experimentos CBR realizados para clasificar en tres categorías. 127
Tabla 4.25: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método de recuperación de casos por inducción. Clasificación en dos categorías. 128
Tabla 4.26: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando los defectos de los 59 puntos del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en dos categorías. 128
Tabla 4.27: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método de recuperación de casos por inducción. Clasificación en dos categorías. 129
Tabla 4.28: Sensibilidad, Especificidad y Precisión diagnóstica obtenidas utilizando las medias de los defectos de las siete zonas del campo visual y el método del vecino más próximo. Todas las características tienen la misma importancia. Clasificación en dos categorías. 129
Tabla 4.29: Comparación de la sensibilidad y precisión diagnóstica de los experimentos CBR realizados para clasificar en dos categorías. 129
Tabla 4.30: Resumen de los resultados de algunos trabajos, incluidos los de esta Tesis, sobre clasificación de campos visuales. 136
Tabla 5.1: Conjuntos borrosos. Variables y términos lingüísticos. 161
Tabla 5.2: Funciones de pertenencia diseñadas para las variables lingüísticas del sistema borroso. 164
Tabla 5.3: Distribución de casos según clases. 174
Tabla 5.4: Número de ejemplos de cada muestra que han sido clasificados como dudosos por el SE del CV dentro de cada clase, y su porcentaje. 174
Tabla 5.5: Sensibilidad, especificidad y precisión diagnóstica del sistema experto para el diagnóstico del campo visual. 174
Tabla 5.6: El mejor mapa SOM que clasifica en glaucoma y no glaucoma entrenado con las medias de los defectos de sensibilidad de las siete zonas del campo visual. 175
Tabla 5.7: Número de ejemplos de cada clase que han sido clasificados como dudosos por el mapa SOM, y su porcentaje. 177
Tabla 5.8: Sensibilidad, especificidad y precisión diagnóstica del mapa SOM para el diagnóstico del campo visual. 177
Tabla 5.9: Resultados de sensibilidad, especificidad, precisión diagnóstica, falsos positivos y negativos, y porcentaje de dudosos obtenidos con el sistema experto del campo visual y con el mapa SOM. 178
Tabla 5.10: Número de ejemplos de cada clase que han sido clasificados como dudosos por el sistema integrado, y su porcentaje. 179
Tabla 5.11: Sensibilidad, especificidad y precisión diagnóstica del sistema integrado para el diagnóstico del campo visual. 179

Tabla 5.12: Resultados de sensibilidad, especificidad, precisión diagnóstica, falsos positivos y negativos, y porcentaje de dudosos obtenidos con el sistema experto del campo visual, con el mapa SOM, y con el sistema integrado. 179
Tabla 5.13: Distribución de casos según clases. 181
Tabla 5.14: Distribución de casos que se usaron para validar el sistema. 182
Tabla 5.15: Número de ejemplos de cada clase que han sido clasificados como dudosos por el experto en la primera y en la segunda revisión de los datos, y por el clasificador de campos visuales, y su porcentaje. 183
Tabla 5.16: Porcentaje de dudosos de cada clase que coincide comparando el clasificador del campo visual con cada una de las revisiones del experto. 183
Tabla 5.17: Porcentaje de dudosos de cada clase que coincide comparando la segunda revisión del experto con la primera. 184
Tabla 5.18: Sensibilidad, especificidad, precisión diagnóstica, falsos positivos y negativos del experto en la primera y segunda revisión de los datos, y del sistema clasificador en el diagnóstico del campo visual. 184
Tabla 5.19: Análisis comparativo del diagnóstico del sistema y del experto en los datos del grupo 1. 185
Tabla 5.20: Análisis comparativo del diagnóstico del sistema y del experto en los datos del grupo 2. 186
Tabla 5.21: Precisión diagnóstica del sistema híbrido de ayuda al diagnóstico del glaucoma respecto al experto. 186
Tabla A.1: Análisis de los parámetros de filtrado 196