

Universidad Central “Marta Abreu” de Las Villas

Facultad de Matemática, Física y Computación



**Trabajo para optar por el Título de Máster en Ciencias
de la Computación**

**Detección de anomalías en imágenes de tomografía
computarizada de cerebro**

Autor: Alexander Falcón Ruiz

Tutor: Dr. C. Alberto Taboada Crispí

Santa Clara

2008

"Año 50 de la Revolución"

Universidad Central “Marta Abreu” de Las Villas
Facultad de Matemática, Física y Computación



**Trabajo para optar por el Título de Máster en Ciencias
de la Computación**

**Detección de anomalías en imágenes de tomografía
computarizada de cerebro**

Autor: Alexander Falcón Ruiz

afalcon@uclv.edu.cu

Tutor: Dr. C. Alberto Taboada Crispí

Profesor Titular, Centro de Estudios de Electrónica y Tecnologías
de la Información, Facultad de Ingeniería Eléctrica,
ataboada@uclv.edu.cu

Santa Clara

2008

"Año 50 de la Revolución"



Hago constar que el presente trabajo de diploma fue realizado en la Universidad Central “Marta Abreu” de Las Villas como parte de la culminación de estudios de la Maestría en Ciencias de la Computación, autorizando a que el mismo sea utilizado por la Institución, para los fines que estime conveniente, tanto de forma parcial como total y que además no podrá ser presentado en eventos, ni publicados sin autorización de la Universidad.

Firma del Autor

Los abajo firmantes certificamos que el presente trabajo ha sido realizado según acuerdo de la dirección de nuestro centro y el mismo cumple con los requisitos que debe tener un trabajo de esta envergadura referido a la temática señalada.

Firma del Tutor

Firma del Jefe de Departamento
donde se defiende el trabajo

Firma del Responsable de
Información Científico-Técnica

PENSAMIENTO

“Es de importancia para quien desee alcanzar una certeza en su investigación, el saber dudar a tiempo”

Aristóteles

DEDICATORIA

A mis padres y hermana.

A mis amigos.

A toda mi familia.

AGRADECIMIENTOS

A mi tutor Alberto, por ofrecerme su valiosa guía y ayuda para llegar a concluir este trabajo.

A mis colegas Denis y Maykel, por la ayuda prestada.

A mis profesores, por transmitirme los conocimientos necesarios.

TAREA TÉCNICA

- Revisión bibliográfica sobre detección de anomalías en imágenes.
- Programación de funciones en Matlab para la extracción de rasgos, distancias y medidas basadas en histogramas para la detección de anomalías en imágenes médicas.
- Desarrollo de aplicación de software para la anotación de anomalías en imágenes médicas.
- Creación de estándar de oro para evaluación de algoritmos de detección de anomalías en imágenes de tomografía computarizada de cerebro.
- Evaluación de combinaciones de rasgos, distancias y medidas basadas en histogramas para la detección de anomalías en imágenes de tomografía computarizada de cerebro.
- Escritura del informe.

Firma del Autor

Firma del Tutor

RESUMEN

Este trabajo trata de la detección de anomalías (lesiones) en imágenes de tomografía computarizada de cerebro. Existen varios trabajos reportados donde se emplean otras modalidades de adquisición de imágenes, sin embargo en el caso de la tomografía computarizada no hay mucho, a pesar de ser una modalidad ampliamente usada y extendida en la actualidad. En esta investigación se propone un algoritmo que combina clasificadores basados en vectores de rasgos y medidas para histogramas, los cuales preclasifican los bloques en la imagen obtenidos con una ventana deslizante. Luego, el cálculo de una distancia robusta dentro de cada clase, decidió si había anomalías o no. En la evaluación del desempeño del sistema se empleó, como estándar de oro, imágenes anotadas manualmente por especialistas médicos. Adicionalmente, se propone una medida de similitud para este propósito, que maneja adecuadamente el desbalance en los datos. El detector requiere refinamientos posteriores para una mejor detección, que tomen en cuenta el tamaño, forma y composición de las regiones sólidas preclasificadas. Los resultados de este trabajo pueden emplearse para encontrar imágenes con anomalías en extensas bases de datos, para asistir a especialistas, en algoritmos de compresión de imágenes, entre otros.

TABLA DE CONTENIDOS

PENSAMIENTO	i
DEDICATORIA	ii
AGRADECIMIENTOS	iii
TAREA TÉCNICA.....	iv
RESUMEN	v
INTRODUCCIÓN	1
CAPÍTULO 1. DETECCIÓN DE ANOMALÍAS EN IMÁGENES MÉDICAS	5
1.1 Detección de anomalías.....	5
1.2 Detección de anomalías en imágenes de tomografía computarizada de cerebro	8
1.3 Métodos de detección de anomalías.....	10
1.3.1 Casos especiales	10
1.3.2 Métodos estadísticos.....	12
1.3.3 Redes neuronales	13
1.3.4 Minería de datos y <i>machine learning</i>	13
1.3.5 Sistemas híbridos.....	14
1.4 Evaluación de algoritmos de detección de anomalías.....	15
1.4.1 Estándar de oro	15
1.4.2 Medidas de desempeño.....	16

CAPÍTULO 2. ALGORITMOS PARA LA DETECCIÓN DE ANOMALÍAS EN IMÁGENES DE TOMOGRAFÍA COMPUTARIZADA DE CEREBRO	20
2.1 Empleo de Matlab	20
2.2 Recolección de datos.....	21
2.3 Anotador de anomalías.....	21
2.4 Etapa preliminar	22
2.5 Clasificación basada en rasgos.....	25
2.6 Clasificación basada en histogramas.....	28
2.7 Etapas de entrenamiento y prueba.....	29
2.8 Etapa de clasificación.....	30
2.9 Medida para la validación del sistema de detección de anomalías	33
CAPÍTULO 3. RESULTADOS Y DISCUSIÓN	35
3.1 Etapas de entrenamiento y prueba.....	35
3.2 Preclasificación	36
3.3 Etapa de clasificación.....	38
CONCLUSIONES Y RECOMENDACIONES	42
Conclusiones	42
Recomendaciones	42
REFERENCIAS BIBLIOGRÁFICAS	44
ANEXOS	49
Anexo I Rasgos.....	49
Anexo II Distancias basadas en histogramas.....	54

INTRODUCCIÓN

La detección de anomalías (DA) puede ser vista como el proceso de encontrar muestras inconsistentes dentro de un conjunto de datos determinado. Los conceptos de la DA han sido aplicados en muchos estudios de análisis de imágenes médicas. En estos casos, se trata de encontrar anomalías asociadas a varias partes del cuerpo como pueden ser hemorragias, fracturas, sangramientos, tumores, etc. En dichos estudios suelen emplearse varias modalidades de adquisición de imágenes como son la resonancia magnética (RM), los rayos X, la tomografía computarizada (TC), entre otras.

Actualmente, se sigue mejorando la tecnología de adquisición y cada vez más se incorporan nuevos algoritmos a estos equipos con el propósito de asistir al personal médico. Los algoritmos DA no son la excepción, y estos pueden ayudar a los especialistas a encontrar, de una forma más rápida y segura, determinado tipo de lesiones en las imágenes obtenidas de los pacientes. Adicionalmente, estos algoritmos pueden ser empleados en la clasificación de enormes bases de datos para obtener imágenes útiles para la investigación y la impartición de docencia. En el campo de la telemedicina, pueden emplearse algoritmos DA para preservar regiones de interés clínico, y compactar el resto de la imagen a una tasa de compresión elevada y con pérdidas, lográndose así archivos más pequeños y adecuados para almacenar y transmitir por redes con anchos de banda limitados.

La presente investigación trata de la DA en imágenes de cerebro obtenidas mediante la modalidad TC. A diferencia de los abundantes trabajos existentes en la DA empleando otras modalidades de adquisición, como por ejemplo la RM, no hay mucho trabajo realizado en el caso de la modalidad TC. La modalidad RM es más adecuada que la TC para encontrar determinadas lesiones en el cerebro, sin embargo existen varios casos en que

esta no puede ser empleada. Adicionalmente, la modalidad TC se encuentra más difundida y es más barata que los sistemas RM.

En este trabajo se proponen algoritmos para detectar anomalías en imágenes de TC axiales de cerebro. Se parte de un conjunto de imágenes recolectadas en el Hospital Universitario Arnaldo Milián Castro y anotadas a mano por especialistas de este centro, empleando un software desarrollado por el autor. De las imágenes se extraen regiones de tejidos sanos las cuales se usan para crear un modelo de lo que se considera normal (modelo de fondo) obteniendo varias clases ‘normales’. Con las imágenes anotadas y las regiones sanas extraídas se conforma un estándar de oro para probar los algoritmos DA. Fue necesaria esta etapa, debido a la carencia de un estándar de oro en este dominio.

El algoritmo DA propuesto, divide la imagen de entrada en varios bloques que se solapan entre sí. Luego, se combinan dos clasificadores, uno basado en rasgos y otro basado en distancias para histogramas con el objetivo de preclasificar estos bloques. Seguidamente, en cada una de las clases obtenidas se aplica una distancia robusta y se encuentran los bloques más alejados dentro de cada clase, siendo estos considerados como anomalías. Por último, se comparan los resultados obtenidos por el sistema con las anotaciones realizadas por los especialistas, aplicando algunas medidas de similitud. Esto permite comprobar cuan efectivo son los clasificadores, los rasgos y distancias empleados, en la detección de las anomalías. En este trabajo se propone una distancia para tal fin.

Objetivo general:

Elaborar un algoritmo para la detección de anomalías en imágenes de tomografía computarizada de cerebro, combinando clasificadores basados en vectores de rasgos y medidas para histogramas, empleando Matlab.

Objetivos específicos:

- Implementar funciones en Matlab para la extracción de rasgos, distancias y medidas basadas en histogramas para la detección de anomalías en imágenes.
- Crear estándar de oro para evaluación de algoritmos de detección de anomalías en imágenes de tomografía computarizada de cerebro.

- Evaluar combinaciones de rasgos, distancias y medidas basadas en histogramas para la detección de anomalías en imágenes de tomografía computarizada de cerebro.

Hipótesis:

Si se emplean y combinan clasificadores basados en vectores de rasgos y medidas para histogramas, entonces se pueden detectar anomalías en imágenes de tomografía axial computarizada de cerebro.

Organización del informe

Este informe de investigación se estructura en introducción, capitulario, conclusiones y recomendaciones, referencias bibliográficas y anexos.

En la introducción se expone brevemente la importancia y actualidad de la detección de anomalías en imágenes médicas, la carencia de estándares de oro en esta rama, principalmente en las imágenes de tomografía computarizada de cerebro y la necesidad del desarrollo y mejora de algoritmos en esta área.

El trabajo consta de 3 capítulos y el contenido de cada uno de ellos es el siguiente:

Capítulo 1: Este se titula Detección de Anomalías en Imágenes Médicas y en él se realiza un estudio del estado del arte del tema, de los métodos y algoritmos empleados para ello y de la evaluación de los mismos.

Capítulo 2: Se llama Algoritmos para la Detección de Anomalías en Imágenes de Tomografía Computarizada de Cerebro y en él se describe la creación de un estándar de oro, a partir de imágenes médicas de tomografía axial computarizada de cerebro anotadas por un personal médico especializado. Se describe además el desarrollo de un algoritmo de detección de anomalías en las imágenes mencionadas y se propone una medida de desempeño para la evaluación del mismo.

Capítulo 3: Resultados y Discusión, este muestra los resultados alcanzados por el algoritmo propuesto en el capítulo dos y valiosas observaciones alrededor de los mismos.

En los anexos, se muestran tablas con las expresiones matemáticas para el cálculo de rasgos y medidas basadas en histogramas, empleadas en la detección de anomalías en imágenes médicas.

CAPÍTULO 1. DETECCIÓN DE ANOMALÍAS EN IMÁGENES MÉDICAS

En este capítulo se define en qué consiste la detección de anomalías, haciendo énfasis en el dominio de las imágenes de tomografía computarizada de cerebro. Se presenta una revisión bibliográfica acerca de las técnicas empleadas actualmente en la detección de anomalías. Por último, se explica el uso de un estándar de oro y de medidas de desempeño para la evaluación de los algoritmos de detección de anomalías.

1.1 Detección de anomalías

En términos generales, la detección de anomalías (DA) puede ser considerada como el proceso de detectar una pequeña fracción de un conjunto de datos, la cual difiere, en algún sentido, de la tendencia global o el patrón definido en dicho conjunto. El objetivo de un detector de anomalías puede ser visto como la identificación de las muestras más inusuales en los datos, sin tener información a priori de sus propiedades, a no ser que son raras y que tienen baja probabilidad de ocurrencia.

El problema DA es usualmente referido, dependiendo del escenario (dominio), como detección de eventos raros, detección de novedad, detección de *outliers*, detección de intrusión o detección de defectos, entre otros. De estos, detección de anomalías y detección de *outliers* son los términos más comúnmente usados, algunas veces intercambiables.

La importancia de la DA se debe al hecho de que las anomalías en los datos se traducen en información significativa (con frecuencia crítica) procesable en una amplia variedad de dominios de aplicación. Por ejemplo, un patrón de tráfico anómalo en una red de

computadoras podría significar que una computadora está enviando información sensible a un destino no autorizado (Ertoz *et al.*, 2004). Una imagen de resonancia magnética (RM) anómala puede indicar la presencia de tumores malignos (Spence *et al.*, 2001).

La DA es ampliamente usada en una gran variedad de aplicaciones tales como detección de fraudes con tarjetas de crédito, detección de intrusión en redes de computadoras, detección de fallas en sistemas de seguridad críticos, vigilancia militar de las actividades enemigas, etc. Además, los conceptos de la DA han sido aplicados a numerosos estudios de la imaginología médica para encontrar anomalías asociadas a varias partes del cuerpo, empleando para ello varias modalidades de adquisición (ver Figura 1.1). La mayor parte de los estudios reportados han tratado de la detección de tumores en mamografías digitales (De Santo, 2003; Astley, 2004; Huang, 2004; Selvi, 2005; Wei, 2005; Peng, 2006; Chiracharit, 2007; Ikedo, 2007; Karnan, 2007), imágenes de tomografía computarizada (TC) de pulmón (Minhas, 2005; Sluimer, 2006), e imágenes RM de cerebro (Chen, 1999; Gering, 2003; Prastawa, 2004; Lee, 2005; Benamrane, 2006; Menze, 2006; Shinkareva, 2006; Bouix, 2007; Ekin, 2007), aunque pueden mencionarse muchos más.

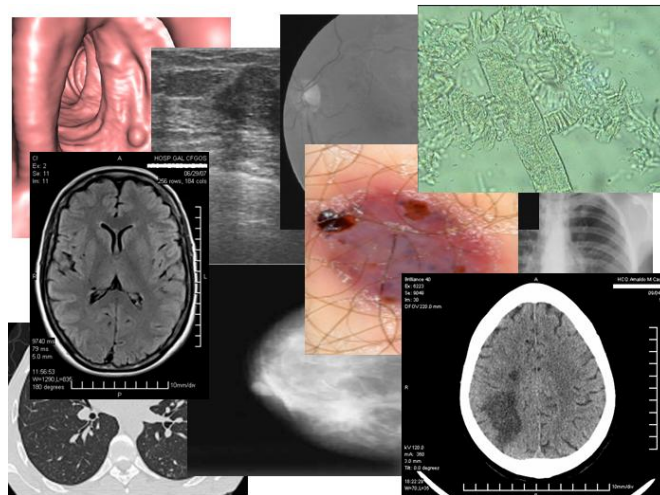


Figura 1.1. Imágenes médicas usadas en la detección de anomalías.

Por su naturaleza, las anomalías no permiten una definición positiva de sus propiedades, haciendo difícil una formulación general del problema. La mayor parte de los enfoques, por el contrario, usan una definición negativa: las anomalías son muestras que no se ajustan a la regla o modelo de lo que es normal. Esto ha conducido a un gran grupo de detectores de anomalías que se basan en algún tipo de modelo matemático o descripción que caracteriza

los datos de interés. Los datos que se ajustan a la descripción, de acuerdo con algún criterio, son considerados normales; aquellos que no se ajustan al modelo son considerados anómalos. Este enfoque de DA se adapta de manera natural a la forma en que los especialistas reconocen las anomalías en las imágenes médicas. Estos están familiarizados con la anatomía sana y la forma en que esta es representada en una modalidad de imágenes específica; por lo tanto, pueden reconocer fácilmente cualquier lesión mostrada en la imagen, comparándola con las zonas sanas que la rodean.

A pesar de la aparente simplicidad del enfoque descrito con anterioridad, existen muchos factores que lo hacen muy desafiante (Chandola *et al.*, 2007):

- Definir una región normal que englobe todos los posibles comportamientos normales puede resultar difícil.
- Las fronteras entre las regiones normales y anómalas con frecuencia son imprecisas. Así una observación anómala puede estar muy pegada al límite y ser confundida con una región normal y viceversa.
- La noción exacta de anomalía es diferente para distintos dominios de aplicación y cada uno de estos impone requerimientos y restricciones.
- La disponibilidad de datos etiquetados (marcados como que pertenecen a una clase) para el entrenamiento y prueba es con frecuencia una de las cuestiones más difíciles de lograr.
- Los datos pueden contener ruido con un comportamiento similar a las anomalías y este puede ser difícil de distinguir y eliminar.

Debido a los retos anteriormente expuestos, el problema DA, en su forma más general, no es fácil de resolver. De hecho, la mayoría de las técnicas existentes simplifican el problema enfocándose en una formulación específica. La formulación está influenciada por varios factores como la naturaleza de los datos, la disponibilidad de datos etiquetados, tipo de anomalías a detectar, etc.

1.2 Detección de anomalías en imágenes de tomografía computarizada de cerebro

A diferencia de las abundantes investigaciones en la DA en imágenes de RM de cerebro (Chen, 1999; Gering, 2003; Prastawa, 2004; Lee, 2005; Benamrane, 2006; Menze, 2006; Shinkareva, 2006; Bouix, 2007; Ekin, 2007), no hay mucho trabajo realizado en el caso de la DA en imágenes TC de esta parte del cuerpo humano. Actualmente, la modalidad RM es más sensible que la modalidad TC para diferenciar los tejidos blandos. En general, la RM supera a la TC en la detección de lesiones en el cerebro tales como los tumores, sin embargo la TC es mejor para la detección de calcificaciones, hemorragias y detalles de partes óseas. La TC es muy útil en el diagnóstico de accidentes cerebro-vasculares, hemorragias intracraneales, y para la evaluación de fracturas faciales y de cráneo. La TC puede remplazar a la RM cuando esta última está contraindicada: pacientes semiconscientes, incapaces de permanecer inmóviles, pacientes con claustrofobia, o pacientes que portan implantes metálicos. La TC es más barata y está más extendida que los sistemas de imágenes de RM. Por consiguiente, el esfuerzo en detectar anomalías en imágenes de TC de cerebro queda plenamente justificado.

Los sistemas modernos de TC proporcionan imágenes con 4096 niveles de gris, que representan diferentes niveles de densidad (dados en Unidades Hounsfield, UH), establecido a 0 UH para el agua y -1000 UH para el aire. En la Figura 1.2 se muestra la representación de tonos de gris de diferentes tejidos en una imagen DICOM (*Digital Imaging and Communication in Medicine*) de 12 bits y sus correspondientes densidades (ajustes ventana/nivel, en UH). La imágenes TC de cerebro insertadas muestran los parámetros de escala por defecto para TC ($0 \pm 2000/2$ UH (nivel $\pm$ ventana/2), arriba) y ajustes ventana/nivel reescalados (a los 256 niveles del monitor) para materia cerebral ($35 \pm 70/2$ UH, derecha) y para análisis de huesos ($800 \pm 2000/2$ HU, izquierda) en un escáner TC multicorte, Philips 'Brilliance 40'.

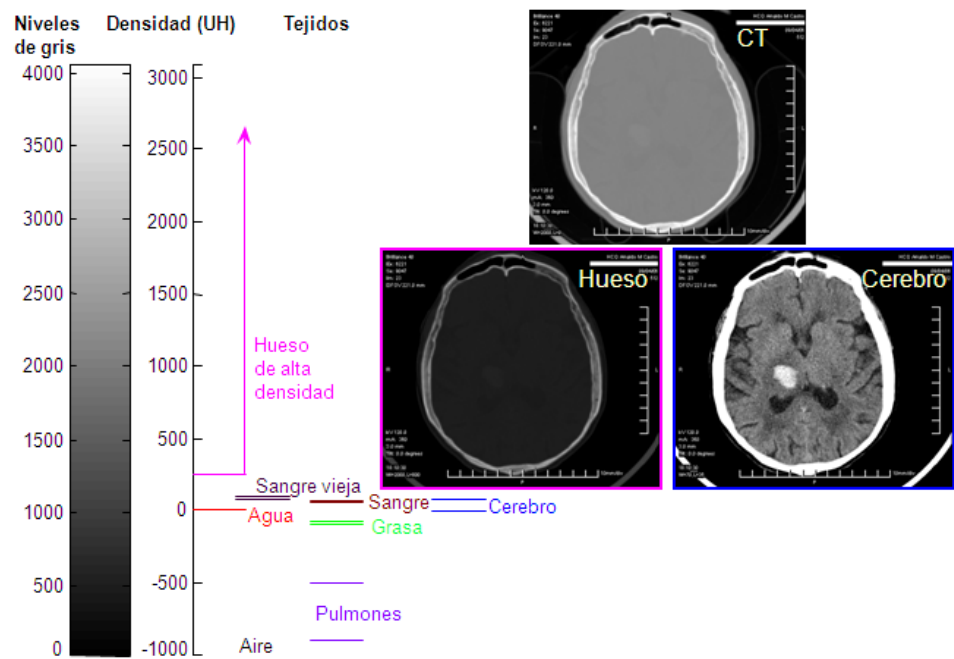


Figura 1.2. Representación de tonos de gris de diferentes tejidos en una imagen DICOM de 12 bits.

Las distintas modalidades producen imágenes muy diferentes para la misma estructura. Aun cuando tengan una apariencia similar, las imágenes de rayos X/TC y de RM, tienen diferentes niveles de grises para representar un tejido determinado, como se observa en la Tabla 1.1. Imágenes de rayos X/TC brillantes se asocian a tejidos de alta densidad (alta atenuación de los rayos X), y las imágenes oscuras se asocian a tejidos de baja densidad (Müller, 2005; Messmer, 2006; Sluimer, 2006).

Tabla 1.1. Representación de tejidos normales y anomalías en diferentes modalidades de adquisición de imágenes.

Estructura			Modalidad de adquisición de imágenes		
Normal/ anomalía	Tejido	Densidad aproximada (UH)	Rayos X /TC	RM T1	RM T2
Normal	Aire	-1000	Muy oscuro	Muy oscuro	Muy oscuro

	Hueso de alta densidad	> 250	Muy claro	Muy oscuro	Muy oscuro
	Grasa	-90 ± 10	Oscuro	Claro	Claro
	Agua	0 ± 5	Oscuro	Oscuro	Claro
	Materia cerebral	35 ± 35	Intermedio	Intermedio	Intermedio
Anomalía	Sangramiento o hemorragia	55 ± 5 (fresca) 80 ± 10 (vieja)	Claro	Claro	Claro
	Infarto		Oscuro	Oscuro	Claro
	Tumores		Oscuro (si no está calcificado)	Oscuro	Claro

1.3 Métodos de detección de anomalías

1.3.1 Casos especiales

En algunos casos la estrategia DA no es conscientemente aplicada y aparece asociada a otra técnica. La DA en el caso de las imágenes médicas puede ser vista como un caso especial de:

- Clasificación (generalmente con solo dos clases: normal y anomalía) (De Santo, 2003; Pokrajac, 2005; Selvi, 2005; Spinosa, 2005; Menze, 2006; Shinkareva, 2006; Strzelecki, 2006; Bouix, 2007). La clasificación generalmente incluye una etapa de preprocesamiento, extracción de rasgos, y el clasificador en sí. Debe resaltarse que los algoritmos típicos de clasificación no son adecuados para tratar con el desbalance de los datos (la clase normal es mucho más abundante que la poco frecuente clase anomalías) en la mayor parte de los problemas DA, pero pueden asistir en alguna etapa para preclasificar los tejidos normales. La aplicación de DA, requiere que el clasificador actúe como un detector en lugar de un clasificador.

- Segmentación (segmentación de regiones anómalas en una imagen) (Van Leemput, 2001; Gering, 2003; Zhou, 2003; Huang, 2004; Prastawa, 2004; Lee, 2005; Ségonne, 2005; Bertelli, 2006; Strzelecki, 2006; Bouix, 2007; Huang, 2007; Montgomery, 2007). La segmentación de imágenes combina los métodos de reconocimiento y de delineación. Las computadoras son más adecuadas para la delineación, superando a los humanos en especificar la extensión espacial de una región anómala. Por ejemplo, los sistemas modernos de imaginología médica pueden manipular imágenes con 4096 niveles diferentes, pero los monitores pueden representar 256 tonos y el ojo humano solo discrimina 20 de estos (Hofer, 2001). Sin embargo, los humanos pueden reconocer las regiones anómalas de manera natural, debido al conocimiento global previo, mientras que las computadoras fallan con frecuencia (Gering, 2003).
- Recuperación de imágenes basada en el contenido (recuperación de imágenes anómalas de una base de datos, y, al mismo tiempo, recuperación de los bloques locales que muestran signos anómalos del conjunto completo de bloques de la imagen) (Müller, 2004; Deselaers, 2005; Lehmann, 2005; Hersh, 2006; Xu, 2006). La creación de bases de datos conteniendo casos normales podría ser usada para comparar la distancia de un nuevo caso con los existentes, haciendo una recuperación por disimilitud en lugar de una recuperación por similitud (Müller, 2004).

La minería de imágenes médicas y el descubrimiento de conocimiento en bases de datos (KDD) (Lehmann, 2005; Margineantu, 2005; Peng, 2006), así como la detección de *outliers* (Van Leemput, 2001; Prastawa, 2004; Rousseeuw, 2006) pueden ser aplicadas para detectar anomalías. La DA puede complementar o ser complementada por otras técnicas. No obstante, la DA tiene sus propia fundamentación, basada en la estadística (Markou y Singh, 2003a), redes neuronales (Markou y Singh, 2003b), *machine learning*, minería de imágenes y KDD (Hodge, 2004; Margineantu, 2005).

1.3.2 Métodos estadísticos

Los métodos estadísticos, que modelan los datos basados en sus propiedades estadísticas, abarcan las técnicas paramétricas y no paramétricas. Los métodos paramétricos asumen que las distribuciones de los datos son gaussianas, caracterizadas por su media y covarianza. Estos métodos comprenden la mezcla de gaussianas (*Gaussian mixture modeling*, GMM), los modelos ocultos de Markov (*Hidden Markov models*, HMM), los campos aleatorios de Markov (*Markov random fields*, MRF) (Van Leemput, 2001; Gering, 2003; Selvi, 2005), los campos aleatorios condicionales (*Conditional random fields*, CRF) (Lee, 2005) y la prueba de hipótesis (*hypothesis testing*). Empleando la estadística robusta tenemos la estimación de volumen mínimo, análisis de componentes principales (*Principal component analysis*, PCA), y maximización de la estimación (*Estimation maximization*, EM) más recocido determinista (Webb, 2002; Markou, 2003a; Hodge, 2004; Rousseeuw, 2006; Daszykowski, 2007; Ekin, 2007). Los modelos no paramétricos incluyen el análisis de histograma, algoritmo de estimación de densidad basado en los k vecinos más cercanos (k-NN) empleando distancia euclidiana o la distancia de Mahalanobis (Webb, 2002; Markou, 2003a), modelos de *kernels* locales o enventanado de Parzen (*Parzen windowing*) (Hodge, 2004; Prastawa, 2004; Rousson, 2005), y coincidencia de vectores de rasgos entre los datos de entrenamiento y los datos de prueba empleando alguna medida de disimilitud.

Pokrajac *et al.* (2005) propuso un método que usa medidas de disimilitud (distancia de Mahalanobis y divergencia Kullback-Leibler) para calcular la diferencia entre distribuciones de probabilidad espaciales de regiones de interés (tales como lesiones, tumores, áreas de actividad en el cerebro, etc.) en una imagen de un nuevo sujeto y en cada una de las clases representadas por los datos históricos (o sea, clase normal vs enfermedad).

Selvi *et al.* (2005) usan campos aleatorios de Markov multiresolución (MMRF) basados en la transformada discreta wavelet para detectar tumores en mamografías. Esta técnica combina: análisis fractal, MMRF, el algoritmo ‘*dogs and rabbits*’ y una red neuronal para la clasificación.

1.3.3 Redes neuronales

Las redes neuronales demandan más recursos computacionales que los modelos estadísticos, y puede que no sean fáciles de reentrenar. Sin embargo, estas pueden ser entrenadas con número pequeño de parámetros y no necesitan hacer suposiciones a priori sobre las propiedades de los datos. Tanto redes supervisadas (como *multi-layer perceptrons*, MLP) como no supervisadas (como *self-organizing maps*, SOM, o Kohonen) (Huang, 2004) pueden ser útiles en la DA en imágenes médicas. Se pueden emplear *Radial Basis Functions* (RBF) (Minhas, 2005), *Learning Vector Quantisation* (LVQ), Redes neuronales probabilísticas (*Probabilistic Neural Networks*, PNN), redes de Hopfield, *Support Vector Machines* (SVM) (Lee, 2005; Wei, 2005; Menze, 2006; Xu, 2006; Chiracharit, 2007), *Synchronized Oscillator Network* (Strzelecki, 2006), y la teoría de Resonancia Adaptativa (*Adaptive Resonance Theory*, ART).

Gering (2003), en su tesis de doctorado, reconoció desviaciones de la normalidad usando redes de dependencia contextual (*contextual dependency network*) que incorporaban el contexto (inmediato o amplio) para detectar tumores en imágenes RM. Huang y Chen (2004) integran redes neuronales SOM con segmentación en el análisis de textura para extraer los contornos de tumores de pecho de imágenes de ultrasonido.

1.3.4 Minería de datos y *machine learning*

A diferencia de los métodos estadísticos y las redes neuronales, los enfoques de *machine learning* pueden trabajar con vectores multitypo y atributos simbólicos (Sebe, 2005). Estos métodos pueden usar sistemas basados en reglas y árboles de decisión, como la técnica C4.5. En el escenario DA, interesa las reglas que describen los casos infrecuentes, mientras los algoritmos de minería de reglas de asociación pueden generar ciertas reglas redundantes. Basado en las regiones normales, las técnicas de minería de datos pueden generar modelos de la normalidad de manera automática, y reconocer desviaciones de esta, consideradas como anomalías.

Xiaoxue *et al.* (2004) usó un enfoque de minería de aprendizaje multi-instancia (MIL) para detectar exudados en imágenes de retina de pacientes diabéticos. En este caso, las imágenes de entrenamiento solo tienen etiquetas de clase ambiguas (normal, anomalía), no

disponiéndose de otra información. De ahí, que el enfoque de minería es adecuado al problema de aprender las características de las anomalías. Se extraen los rasgos relevantes de las imágenes, se descubren las anomalías y se trabaja con los errores en los datos de entrenamiento.

Wei *et al.* (2005) propuso el uso de máquina de vector de relevancia (*relevance vector machine*, RVM), como un clasificador, para detectar microcalcificaciones en mamografías digitales como un problema de aprendizaje supervisado. RVM es un procedimiento *machine learning* basado en la teoría de estimación Bayesiana, con una función de decisión esparcida definida por un número muy pequeño de vectores de relevancia. RVM no es solo exacta sino también computacionalmente eficiente.

1.3.5 Sistemas híbridos

Últimamente, los métodos estadísticos, redes neuronales y *machine learning* han sido combinados en sistemas híbridos para obtener un mejor desempeño que con los algoritmos individuales para la DA. Sin embargo, combinar algoritmos es diferente a combinar clasificadores, ya que en estos últimos trabajan combinando etiquetas de clase discretas (usando alguna técnica de votación), mientras en la combinación de algoritmos DA se combinan los rankings de los algoritmos.

Los sistemas híbridos proliferan en trabajos recientes en el contexto de la DA en imágenes médicas. Se propuso un Sistema de Múltiples Expertos para clasificar microcalcificaciones simples en imágenes de mamografía (De Santo, 2003). Se ha combinado un algoritmo genético para buscar manchas claras en mamografías con un mecanismo de descubrimiento de conocimiento para mejorar su desempeño (Peng, 2006)

Las redes neuronales, lógica difusa y algoritmos genéticos han sido combinados en sistemas híbridos para detectar anomalías presentes y especificar su naturaleza (tumores malignos o benignos) en imágenes de RM de cerebro (Benamrane, 2006). Un clasificador basado en redes neuronales, un clasificador Bayesiano, y un clasificador basado en cadenas ocultas de Markov fueron unidos para un proceso de clasificación semiautomático (Bertelli, 2006).

1.4 Evaluación de algoritmos de detección de anomalías

Para evaluar el desempeño de los sistemas DA, se requiere de (i) un adecuado estándar de oro (*ground truth*), y de (ii) medidas de similitud apropiadas para una evaluación cuantitativa en la comparación de los resultados de la detección con el estándar de oro. Sin embargo, obtener un estándar de oro aceptable es una tarea difícil, y las medidas estándares de evaluación del desempeño tales como el índice de error en la clasificación son irrelevantes en el contexto de la DA.

1.4.1 Estándar de oro

El enfoque más común para enfrentarse a este problema es comparar los resultados de las técnicas DA bajo evaluación con los de un grupo de expertos (anotados o segmentados a mano) para un conjunto de imágenes (Chen, 1999). Los expertos humanos tienen excelentes capacidades de reconocimiento debido al conocimiento previo global; sin embargo, sus habilidades de delineación son pobres. Los coeficientes de variación intraexpertos se ha reportado alrededor de 6,5%, y los coeficientes interexpertos en el orden de 22,1% de acuerdo con (Ashton, 2006), o alrededor del 15% tanto en los coeficientes intra como interexpertos según (Gering, 2003). Los *Phantoms* (imágenes sintéticas (Gering, 2003; Pokrajac, 2005) u objetos físicos conocidos, incluyendo cadáveres) también pueden ser usados en la evaluación de los algoritmos DA, al menos en la etapa preliminar, pero esto no es una tarea trivial.

A diferencia de los cientos de bases de datos de imágenes médicas disponibles para otras tareas de análisis de imágenes, las bases de datos anotadas para la evaluación de los algoritmos DA son muy poco comunes. Esta carencia de un estándar de oro para validar los algoritmos DA, forzaron a Bouix et al. (2007) a enfocar su atención en un principio de acuerdo común para validar algoritmos para la clasificación de tejidos de cerebro sin un estándar de oro. El halló que esta técnica era adecuada para la DA, pero no suficiente para una evaluación precisa del desempeño de los clasificadores.

1.4.2 Medidas de desempeño

Disponiendo de un estándar de oro adecuado (donde se muestren las regiones anómalas), la validación del algoritmo requiere la definición de medidas de similitud para comparar los datos de referencia con la salida del sistema DA. Se han empleado varias medidas, desde la diferencia en el tamaño de las anomalías detectadas (Chen, 1999; Van Leemput, 2001; Strzelecki, 2006) hasta numerosas distancias entre los límites de las anomalías segmentadas, como la distancia Hausdorff (Prastawa, 2004), o medidas de solapamiento espacial (Bouix, 2007).

En este contexto, el estándar de oro puede ser visto como una imagen binaria, X , con 0's (negro) en las regiones normales y 1's (blanco) en las regiones anómalas. La salida del sistema DA bajo consideración también puede ser visto como una imagen binaria, Y , con una asignación de valores similar. Tanto X como Y deben coincidir en tamaño. La Figura 1.3 muestra un razonamiento esquemático, basado en la teoría de la información, para calcular medidas de similitud entre Y (rojo) y X (azul). Las siguientes notaciones son consideradas:

- a = número de ocurrencias de $x_i = 1$ y $y_i = 1$ (anomalías detectadas como anomalías), conocido como verdaderos positivos.
- b = número de ocurrencias de $x_i = 0$ y $y_i = 1$ (zonas normales detectadas como anomalías), conocido como falsos positivos.
- c = número de ocurrencias de $x_i = 1$ y $y_i = 0$ (anomalía detectada como zonas normales), conocido como falsos negativos.
- d = número de ocurrencias de $x_i = 0$ y $y_i = 0$ (zonas normales detectadas como zonas normales), conocido como verdaderos negativos.

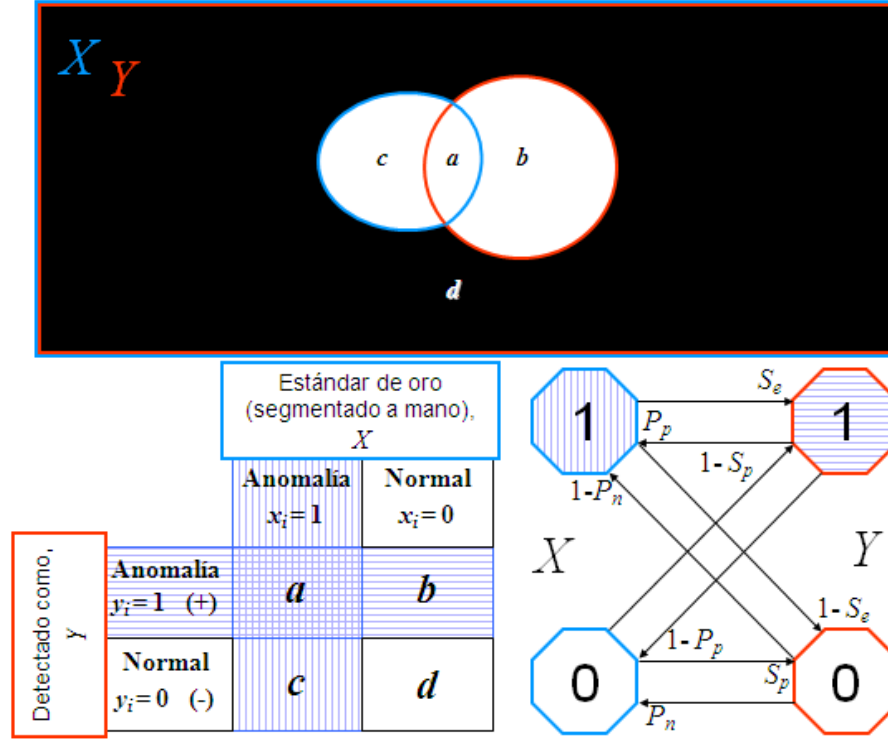


Figura 1.3. Similitud entre X y Y basado en la teoría de la información.

Se pueden calcular varios coeficientes de similitud que sirven como simples medidas para la evaluación. En general, las medidas más conocidas (esquemáticamente representada en la parte inferior derecha de la Figura 1.3) son:

- S_e = sensibilidad, índice de detección, *recall*, el cual es la probabilidad de detectar las anomalías correctamente, definido como

$$S_e = \text{Recall} = p(\text{Anomalia} = 1 | X = 1) = 1 - p(\text{Normal} = 0 | X = 1) = \frac{a}{a+c} = 1 - \frac{c}{a+c}. \quad (1.1)$$

- S_p = especificidad, el cual es la probabilidad de detectar lo que es normal correctamente, calculada como

$$S_p = p(\text{Normal} = 0 | X = 0) = 1 - p(\text{Anomalia} = 1 | X = 0) = \frac{d}{b+d} = 1 - \frac{b}{b+d}. \quad (1.2)$$

- P_p = Predictividad positiva, precisión, o índice de detección Bayesiano, el cual es la probabilidad de tener positivos correctamente detectados, esto es

$$P_p = \text{Precision} = p(X=1|Y=1) = 1 - p(X=0|Y=1) = \frac{a}{a+b} = 1 - \frac{b}{a+b}. \quad (1.3)$$

- P_n = predictividad negativa, índice negativo Bayesiano, el cual es la probabilidad de tener negativos correctamente detectados, de acuerdo con

$$P_n = p(X=0|Y=0) = 1 - p(X=1|Y=0) = \frac{d}{c+d} = 1 - \frac{c}{c+d}. \quad (1.4)$$

- R_c = índice de clasificación, exactitud, el cual es el índice de clasificaciones correctas, calculado como

$$R_c = D_{sm} = p(X=Y) = 1 - p(X \neq Y) = \frac{a+d}{a+b+c+d} = 1 - \frac{b+c}{a+b+c+d}. \quad (1.5)$$

Estas medidas por sí mismas usualmente no son empleadas en el contexto DA debido al gran número de ceros (áreas negras) en las imágenes binarias X y Y , $(d+b) \gg (a+c)$. Debe recordarse que las anomalías son eventos muy poco comunes. Para tener una mejor idea del desbalance en los datos, se puede usar el índice de anomalías, R_B , el cual es normalmente $\ll 1$ en la mayor parte de los escenarios DA, y que se expresa

$$R_B = p(X=1) = 1 - p(X=0) = \frac{a+c}{a+b+c+d} = 1 - \frac{b+d}{a+b+c+d}. \quad (1.6)$$

Para lidiar con el desbalance crítico de los datos, algunos autores han usado los siguientes coeficientes en su lugar:

- D_T = distancia de Tanimoto, o distancia de Rogers, similar al índice de clasificación pero dando el doble de peso a $(b+c)$ en el denominador

$$D_T = \frac{a+d}{a+d+2(b+c)}; \quad (1.7)$$

- D_R = distancia de Russel y Rao, la cual no involucra el término d en el numerador,

$$D_R = \frac{a}{a+b+c+d}; \quad (1.8)$$

- D_J = distancia de Jaccard (Bouix, 2007), la cual descarta completamente el término d ,

$$D_J = \frac{a}{a+b+c}; \quad (1.9)$$

- D_C = distancia de Czekanowski, Dice, Sorensen, la cual es similar a la distancia de Jaccard, pero dando doble peso a a ,

$$D_C = F_{measure} = \frac{2S_e P_p}{S_e + P_p} = \frac{2a}{2a+b+c}. \quad (1.10)$$

Las curvas ROC (*Receiver Operating Characteristic*) (S_e vs. $1-S_p$), área bajo la curva ROC (AUC), las curvas precisión/*recall* (S_e vs. $1-P_p$) y algunas otras medidas, han sido también usadas para validar los algoritmos DA. Las curvas ROC pueden ser usadas (De Santo, 2003; Wei, 2005; Gu, 2006; Peng, 2006) para validar sistemas DA, pero estas por sí mismas no puede ser usadas para determinar el mejor punto de operación. Similarmente, la AUC (De Santo, 2003; Menze, 2006; Kelm, 2007) no es apropiada para la validación debido a que esta considera todos los posibles puntos de operación para el sistema DA bajo evaluación. Las curva precisión/*recall* parecen ser más adecuadas para la evaluación DA, por estar asociadas a los detectores.

CAPÍTULO 2. ALGORITMOS PARA LA DETECCIÓN DE ANOMALÍAS EN IMÁGENES DE TOMOGRAFÍA COMPUTARIZADA DE CEREBRO

En este capítulo se describe y fundamenta el uso de herramientas como el Matlab y el Anomaly Annotator (Anotador de Anomalías) en la realización de la investigación. Luego, se describen las distintas etapas para el desarrollo de un sistema de anotación de anomalías en imágenes de tomografía computarizada de cerebro. Por último, se propone una medida para evaluar el desempeño del sistema diseñado.

2.1 Empleo de Matlab

Todos los algoritmos desarrollados para esta investigación fueron escritos empleando el programa Matlab en su versión 7.3.0 (R2006b). Se ha elegido Matlab por ser un entorno de cálculo técnico con un lenguaje de alto nivel, que integra cálculo, visualización, programación y depuración en un ambiente muy cómodo, donde los problemas y las soluciones se expresan en una notación matemática familiar. Además, este programa incorpora un extenso conjunto de paquetes especializados de funciones (*toolboxes*), como los de procesamiento digital de imágenes, procesamiento digital de señales, estadística, redes neuronales, lógica *fuzzy*, entre otros. También existen numerosos recursos de ayuda y una comunidad activa que desarrolla funciones y otros paquetes que se encuentran disponibles en Internet, como el Spider (*machine learning*), Libra (métodos estadísticos robustos) y el Tomcat (calibración multivariada). En las siguientes secciones se propone el uso de determinadas funciones de Matlab.

2.2 Recolección de datos

Se han recolectado los datos de 27 estudios de TC de cerebro, con alrededor de 33 cortes axiales (total: 890 cortes de tamaño 512×512 voxels y resolución de 12 bits (4096 niveles de gris)), empleando un escáner de TC multicorte, modelo Philips ‘Brilliance 40’, en el Hospital Universitario “Arnaldo Milián Castro”. Además, se han recolectado datos de otros estudios realizados en el mismo hospital con el equipo Siemens ‘Somaton AR.Star’, solo para la verificación final.

2.3 Anotador de anomalías

Como valor agregado del escáner Philips ‘Brilliance 40’ existe una aplicación de software llamada MxLiteView que permite la visualización de los estudios realizados. Este programa incorpora, además, facilidades de anotación, solo que estas son muy básicas e insuficientes para la creación de un estándar de oro. Por consiguiente, como parte de esta investigación, se ha creado una herramienta (Figura 2.1) (Falcón *et al.*, 2008) que permite anotar de una manera muy flexible anomalías y otras regiones de interés en imágenes médicas (principalmente en formato DICOM). El programa permite cargar imágenes y realizar operaciones como hacer y deshacer, dibujar líneas, rectángulos, elipses, dibujar a mano alzada, rellenar, borrar, ampliar/reducir, usar varios colores y niveles de transparencia, ajuste de ventana/nivel, salvar anotaciones, entre otras. El software ha sido íntegramente programado en C# 2.0 para la plataforma .Net de Microsoft, y además puede correr en varias plataformas de hardware y software, al ser compatible con Mono.

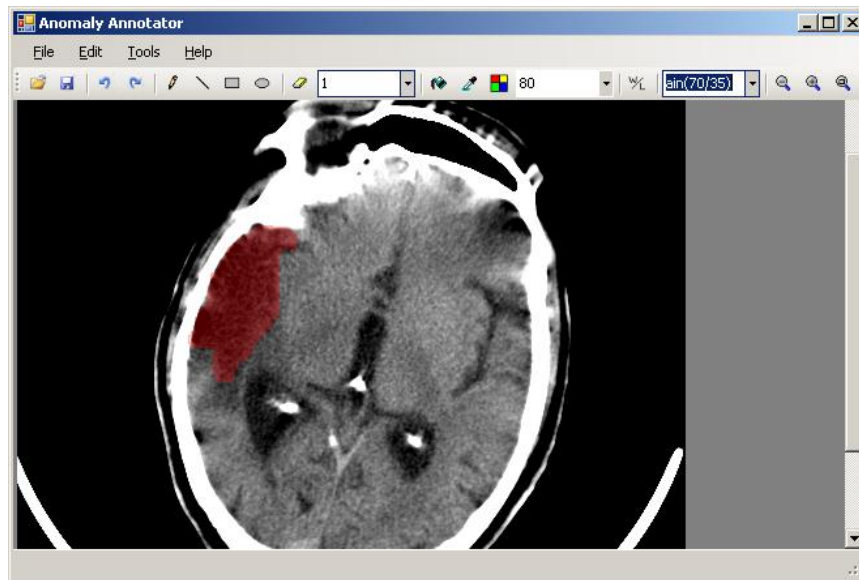


Figura 2.1. Anomaly Annotator: Software para la anotación de anomalías en imágenes médicas.

2.4 Etapa preliminar

En una etapa preliminar, con la ayuda de radiólogos y neurólogos, se seleccionaron un conjunto de regiones que representan las siguientes clases normales (Figura 2.2): aire/fondo (BG), fluido cerebroespinal (CSF), materia cerebral (BM) dividida en materia blanca y materia gris (WM y GM, respectivamente), y cráneo o hueso de alta densidad (HDB). Debe enfatizarse que solo se consideraron regiones normales en la primera etapa. Se decidió fijar la atención en anomalías de $8\text{mm} \times 8\text{mm}$, o sea, 16×16 voxels a la resolución dada, y, en consecuencia, trabajar con bloques de 16×16 . Cada uno de estos bloques incluye 256 voxels, lo cual es un número estadísticamente significativo y no implica una excesiva carga computacional.

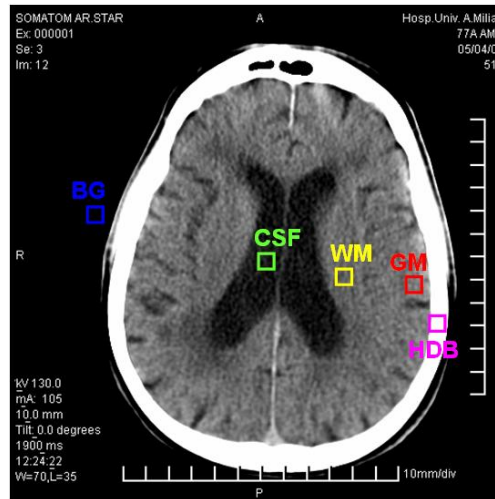


Figura 2.2. Bloques de 16×16 tomados de regiones normales en la imagen de tomografía computarizada de cerebro.

De las regiones normales segmentadas a mano (empleando el software Anomaly Annotator), se extrajeron automáticamente miles de bloques de 16×16 de cada clase para aprender de ellas. Se calcularon los histogramas de niveles de gris de los voxels (*bins* desde 0 hasta 4095 con ancho de 4) en cada bloque extraído. Luego, se promediaron, *bin a bin*, todos los histogramas de una misma clase para obtener un histograma prototipo por clase, los cuales se normalizaron dividiendo el número de ocurrencias en cada *bin* por el número de voxels por bloque (256) para obtener un estimado de la distribución de probabilidad de los niveles de gris de cada clase.

La Figura 2.3 ilustra los histogramas obtenidos para cada clase. Como puede notarse las clases BG y HDB son fácilmente reconocibles. De hecho, los bloques que pertenecen a estas clases pueden ser clasificados usando un simple rasgo, como la mediana de los niveles de gris y un umbral. Para ser específico, si la mediana de los niveles de gris es < 500 , entonces el bloque pertenece a la clase BG; si la mediana de los niveles de gris > 1250 , entonces el bloque pertenece a la clase HDB.

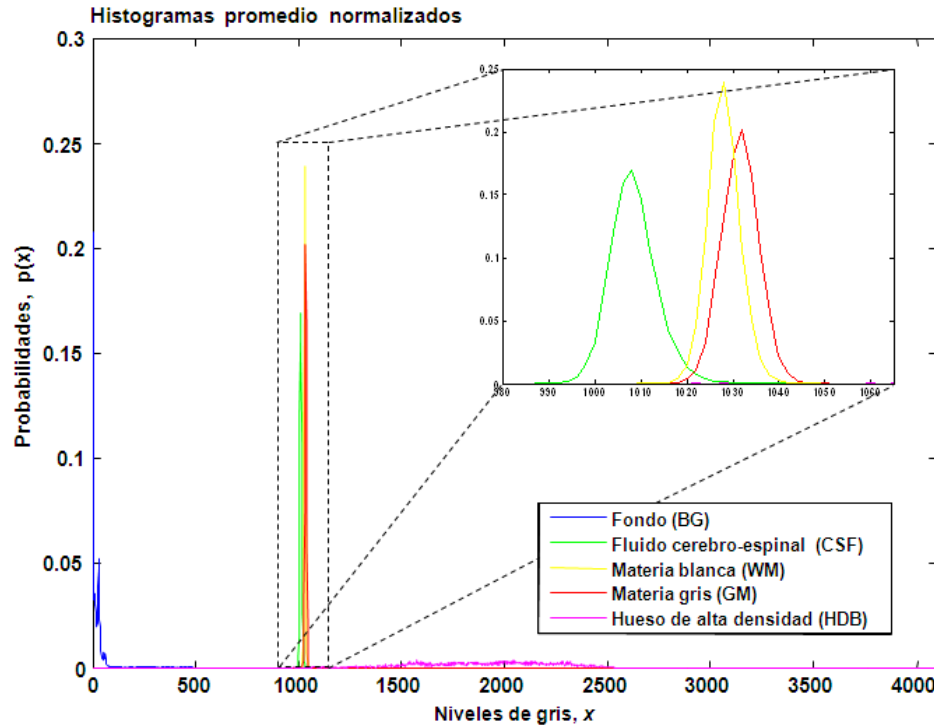


Figura 2.3. Histogramas promedio normalizados para las clases normales en las imágenes de tomografía computarizada de cerebro.

Se confirmó que las clases WM y GM se superponen en gran medida, de modo que se decidió agruparlas en una única clase: la clase materia cerebral (BM), cuyo histograma prototipo es un promedio, *bin a bin*, de los histogramas que representan las subclases WM y GM. Las anomalías en el cerebro, como las hemorragias, infartos, tumores y calcificaciones, son también consideradas tejidos blandos (con la excepción de ciertas calcificaciones). De esta manera, estos deben aparecer en la imagen de TC con los ajustes ventana/nivel (que permiten el mapeo de porciones del rango dinámico de la imagen al rango dinámico del monitor) usados para la materia cerebral, o sea, alrededor de 35 UH (nivel de gris 1035). Por consiguiente, de aquí en adelante el enfoque estará en dos clases normales: fluido cerebroespinal (CSF) y materia cerebral (BM).

Uno de los problemas más desafiantes en la DA en imágenes médicas está asociado con la gran variabilidad en la apariencia de los tejidos normales. La mayor parte de los trabajos publicados en la DA en el contexto de imágenes médicas crean un modelo estadístico de las estructuras normales y entonces buscan anomalías (Ashton, 2006). En el caso de la DA en

imágenes de TC de cerebro, se propone una combinación de técnicas de clasificación basada en rasgos con la clasificación basada en histogramas.

2.5 Clasificación basada en rasgos

Se experimentó con 40 rasgos diferentes (Tabla A1 en los anexos) , descritos en numerosos trabajos (Duncan, 2000; Webb, 2002; DeSanto, 2003; Astley, 2004; Huang, 2004; Müller, 2004; Deselaers, 2005; Lehmann, 2005; Müller, 2005; Pokrajac, 2005; Radke, 2005; Benamrane, 2006; Peng, 2006; Sluimer, 2006; Stoitsis, 2006; Strzelecki, 2006; Xu, 2006; Taboada, 2008) para preclasificar las regiones CSF y BM. Algunos de estos rasgos son medidas basadas en la entropía (1-5 en la Tabla A1), otras representan la tendencia central de los niveles de gris en los bloques bajo evaluación (6-10 en la Tabla A1), otros representan la dispersión (11-16 en la Tabla A1), o la forma de la distribución de probabilidad (17-18 en la Tabla A1), junto con la matriz de coocurrencia de niveles de gris (GLCM) medidas basadas en la textura (Xie, 2006) (19-42 en la Tabla A1). En nuestro caso, para los rasgos GLCM, se usaron solo orientación este y sur, y la distancia $d = 1$, reduciendo el número de niveles de gris (N) a 20, de 985 a 1065 con ancho 4, de acuerdo con el ajuste ventana/nivel usado (los niveles de gris por debajo de 985 o por encima de 1065 fueron forzados a 985 o 1065, respectivamente).

Se calcularon todos los rasgos para cada bloque etiquetado y se ordenaron de acuerdo a su relevancia por medio un enfoque de selección de rasgos usando la función FSV del Spider. Los resultados se muestran en la última columna de la Tabla A1 en los anexos.

En correspondencia con estos resultados, los mejores rasgos para reconocer áreas del cerebro en imágenes de TC son los rasgos basados en la entropía (1-5 en la Tabla A1), los cuales describen la complejidad de los niveles de gris en el bloque bajo evaluación, en concordancia con observaciones previas realizadas con un sistema experimental (Taboada, 2008). Las estadísticas de primer orden que representan la tendencia central (6-10 en la Tabla A1) también se desempeñan muy bien (ranking entre 4 y 8). Por otra parte, no hay una diferencia notable entre los parámetros de dispersión (11-16 en la Tabla A1) de las dos clases, ni en la forma de las distribuciones de probabilidad (Kurtosis y Skewness (Simetría)), como puede esperarse de la Figura 2.3. Sorprendentemente, ninguno de los rasgos basados en GLCM para el análisis de textura quedó en el ranking entre los 8

primeros lugares. Los mejores fueron el 9 y el 10 que se relacionan con la media en dirección sur y este. Diferentes valores de orientación (θ diferente de este y sur) y distancia ($d > 1$) podrían dar mejores resultados, pero no es trivial.

Empleando un pequeño conjunto de bloques etiquetados que representan las clases normales, en este caso: 250 de regiones CSF (bloques 1-250), y 250 de regiones BM (125 WM, bloques 251-375, y 125 GM, bloques 376-500), se verificó que los 12 primeros rasgos en el ranking fueron excelentes para reconocer los bloques CSF y BM. Los rangos dinámicos obtenidos al calcular estos rasgos son mostrados en la Tabla 2.1. Al no existir artefactos, existe una gran correspondencia entre los valores de todas las medidas de tendencia central (media aritmética, media geométrica, media armónica, media recortada y mediana). Adicionalmente, a pesar de la significativa diferencia en los rangos dinámicos, se notó una interesante coincidencia en la tendencia de todos estos valores.

Tabla 2.1. Rango dinámico (valores mínimo y máximo) de los primeros rasgos en el ranking.

Rasgo	Mínimo	Máximo
Medidas de la tendencia central: • Media aritmética (ma) • Media geométrica (mg) • Media armónica (mh) • Media recortada con $\alpha = 10\%$ ($m\alpha$) • Mediana (m50)	1 005,0	1 039,0
Entropía de Shannon (-HSh)	3 582 300 000,0	3 830 100 000,0
Entropía Logaritmo de la Energía (HlogE)	3 539,8	3 555,8
Entropía de norma $p=2$ (Hnorm)	259 068 081,0	275 747 614,0
Entropía SURE con $T2 = 1065$ (HSURE)	259 067 825,0	275 747 358,0
Media GLCM (orientación sur y este)	4,6	12,6
Entropía con umbral $T1 = 1020$ (Hthr)	0,0	256,0

Para verificar este comportamiento similar en los primeros rasgos en el ranking, se normalizó cada medida (x_i) según la media (μ) y desviación estándar (σ) del rasgo correspondiente, calculado como en (A6) y (A13), respectivamente,

$$x_n(i) = \frac{x_i - \mu}{\sigma}. \quad (2.1)$$

Luego, se graficaron estos valores de rasgos normalizados (x_n) calculado para cada bloque, como se muestra Figura 2.4. Cualquiera de estos rasgos por sí mismo es capaz de distinguir un bloque CSF (primera mitad en la gráfica, o sea bloques del 1-125, correspondiente a los valores del rasgo normalizado alrededor de -1) de un bloque BM (segunda mitad en la gráfica, o sea bloques del 251-500, correspondiente a los valores del rasgo normalizado alrededor de +1). Sin embargo, como se ha asumido, estos rasgos no pueden diferenciar tan bien las subclases WM (materia blanca, bloques 251-375) de GM (materia gris, bloques 376-500). De estos 12 rasgos, solo la entropía con umbral (con umbral $T_1 = 1020$, o sea el nivel de gris de intersección entre los histogramas prototipos de CSF y BM) se desempeña ligeramente diferente.

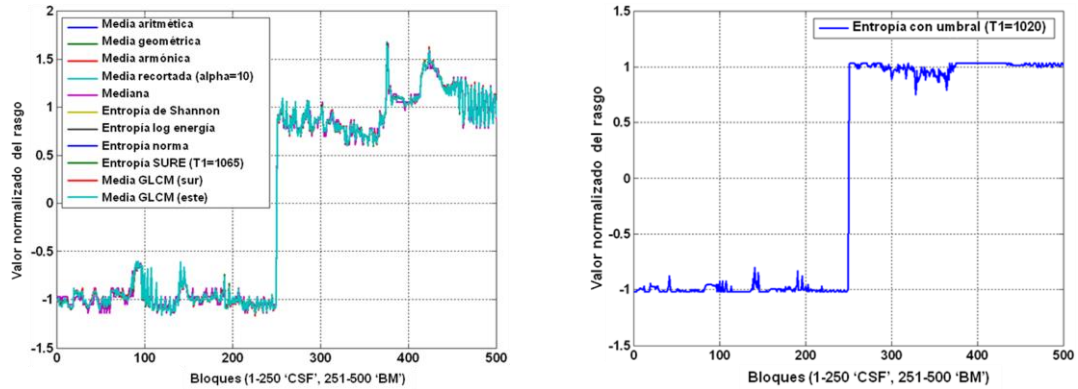


Figura 2.4. Comportamiento de los mejores rasgos al preclasificar los bloques normales CSF y BM.

De acuerdo con este resultado, se seleccionó un vector de rasgos formado por solo uno de los rasgos en la gráfica a la izquierda de la Figura 2.4 (concretamente se eligió la mediana por ser la más adecuada en la presencia de artefactos) y la entropía con umbral (la cual se desempeña ligeramente diferente al resto, como se observa en el gráfico de la derecha) para preclasificar los bloques. Por consiguiente, la clasificación basada en rasgos, usará dos

vectores prototipos de dos rasgos cada uno (mediana y entropía con umbral), denotados como V_{CSF} y V_{BM} para las clases CSF y BM, respectivamente. Estos vectores se caracterizarán por sus centros (m_{CSF} y m_{BM} , para las clases CSF y BM, respectivamente) y sus matrices de covarianza (C_{CSF} y C_{BM} , para CSF y BM, respectivamente).

Las distancias de Mahalanobis (Webb, 2002; Rousseeuw, 2006; Daszykowski, 2007; Taboada, 2008), D_M , entre el vector de rasgos del bloque i -ésimo, V_i , y los vectores prototipos, V_{CSF} and V_{BM} , son calculadas como:

$$\begin{aligned} D_M(V_i, V_{CSF}) &= \sqrt{(V_i - m_{CSF})^T C_{CSF}^{-1} (V_i - m_{CSF})} \\ D_M(V_i, V_{BM}) &= \sqrt{(V_i - m_{BM})^T C_{BM}^{-1} (V_i - m_{BM})} \end{aligned} \quad (2.2)$$

Después de calcular estas distancias (2.2) para todos los bloques a clasificar, el bloque i -ésimo es preasignado a la clase para la cual la distancia D_M tenga el valor mínimo. Una técnica similar se ha usado en (Gerin, 2003).

2.6 Clasificación basada en histogramas

A pesar de su simplicidad, las técnicas basadas en histogramas han probado su valor como un enfoque de bajo nivel y bajo costo computacional. Estas son invariantes a la traslación y rotación, son insensibles a la distribución espacial exacta de los voxels. Por lo tanto, para complementar la clasificación basada en rasgos, se propone un esquema basado en histogramas con un tamaño de *bin* variable.

El tamaño variable de *bin* está basado en el tradicional ajuste ventana/nivel usado para analizar visualmente la materia cerebral y el cráneo en las imágenes de TC. En el primer *bin*, se incluyeron los niveles de gris por debajo de 985, correspondiente a la clase fondo/aire, y el último *bin* los que están por encima de 1250, correspondiente al hueso de alta densidad del cráneo. El resto de los *bins* quedaron como en el ajuste ventana/nivel para cerebro, entre los valores 985 y 1065 con un ancho de 4. Luego, los bordes de los *bins* de los histogramas a calcular son 0, 985, 989, 993, ..., 1057, 1061, 1065, 1250, y 4095; arrojando solo 23 *bins* en total. Después del cálculo de los histogramas con *bin* variable para cada bloque (como se describe arriba), se estimó el histograma promedio (prototipo) normalizado para cada clase.

Sean, hCSF y hBM estos histogramas de referencia, para las clases fluido cerebroespinal (CSF) y materia cerebral (BM), respectivamente. Luego, se evaluaron varias distancias basadas en histogramas, D_h , reportadas en la literatura (Webb, 2002; Deselaers, 2005; Lehmann, 2005; Rousseeuw, 2006; Stoitsis, 2006; Taboada, 2008), y listadas en la Tabla A2 de los anexos. Se hizo un análisis similar al de la clasificación basada en rasgos, y se encontró una relativa coincidencia entre algunas de las medidas.

Para la evaluación de las diferentes medidas, se calculó el índice de clasificación, R_C , de acuerdo con (1.5). También, se propuso una medida, S , para cuantificar la habilidad de separación entre clases de la medida bajo consideración. Se calcularon la suma de las distancias bajo comprobación, después de una normalización de amplitud, con respecto a cada histograma prototipo menos dos veces la distancia con respecto al histograma prototipo de la clase esperada. Por lo tanto, las distancias ‘no adecuadas’ contribuyen positivamente a esta medida, mientras las distancias ‘adecuadas’ contribuyen negativamente, dando un valor mayor a las distancias con mejores propiedades de separación. Por ejemplo, si la clase esperada es BM, entonces la medida (S) es calculada para el i -ésimo bloque usando la medida D_h como

$$S = D_h(h_i, h_{CSF}) - D_h(h_i, h_{BM}) , \quad (2.3)$$

donde h_i es el histograma del i -ésimo bloque, y hCSF and hBM son los histogramas prototipos para las clases CSF y BM, respectivamente.

El índice de clasificación y la medida de separación (S) permite ordenar los desempeños de las distancias D_h . Para la etapa de clasificación se consideran las 2 primeras distancias para histogramas en el ranking, denotadas por D_{h1} y D_{h2} . Después de calcular estas dos distancias para todos los bloques a clasificar, el i -ésimo bloque es preasignado a la clase para la cual la distancia D_h tiene valor mínimo.

2.7 Etapas de entrenamiento y prueba

Para cada uno de los enfoques propuestos, el basado en rasgos y el basado en histogramas (Figura 2.5), se considera una etapa de entrenamiento durante la cual se estima el conjunto de rasgos y de histogramas que representan las clases normales, llamadas CSF y BM.

Además, se aplica una etapa de prueba, para verificar la preclasificación de los bloques usando los prototipos obtenidos en la etapa de entrenamiento. Aleatoriamente se tomaron cientos de bloques de cada clase de 14 de los 27 estudios (conjunto de entrenamiento) y entonces se probaron los resultados con un número similar de bloques tomados de los restantes 13 estudios (conjunto de prueba).

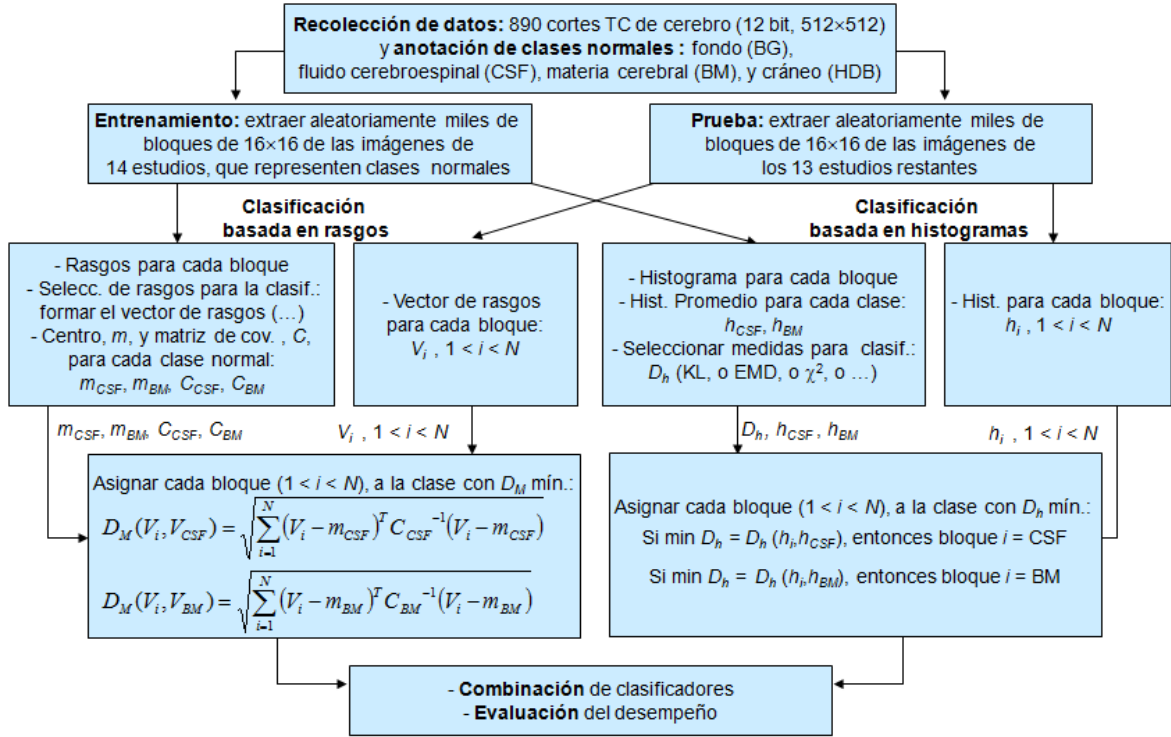


Figura 2.5. Sistema de detección de anomalías propuesto: diagrama de bloques para etapas de entrenamiento y prueba.

2.8 Etapa de clasificación

El diagrama de bloques de la etapa de clasificación se muestra en la Figura 2.6. Como se observa, después de cargar la imagen o el conjunto de imágenes (de tamaño 512×512 cada una) para cada estudio, una ventana deslizante de 16×16 voxels, en correspondencia con el tamaño de la menor anomalía a detectar, genera un conjunto de bloques con un solapamiento del 50% (en total, $(2 \cdot (512 \div 16) - 1)^2 = 3969$ bloques por imagen) (Taboada, 2008). Una parte importante de estos bloques (N_e) puede ser inmediatamente preclasificada como fondo (BG), o hueso de alta densidad del cráneo (HDB), por tener una

mediana de niveles de gris por debajo de 500 (< -500 HU), o por encima de 1250 ($> +250$ UH), respectivamente, como se aprendió en la primera etapa. El resto de los bloques ($N_r = 3969 - N_e$) tienen que ser clasificados.

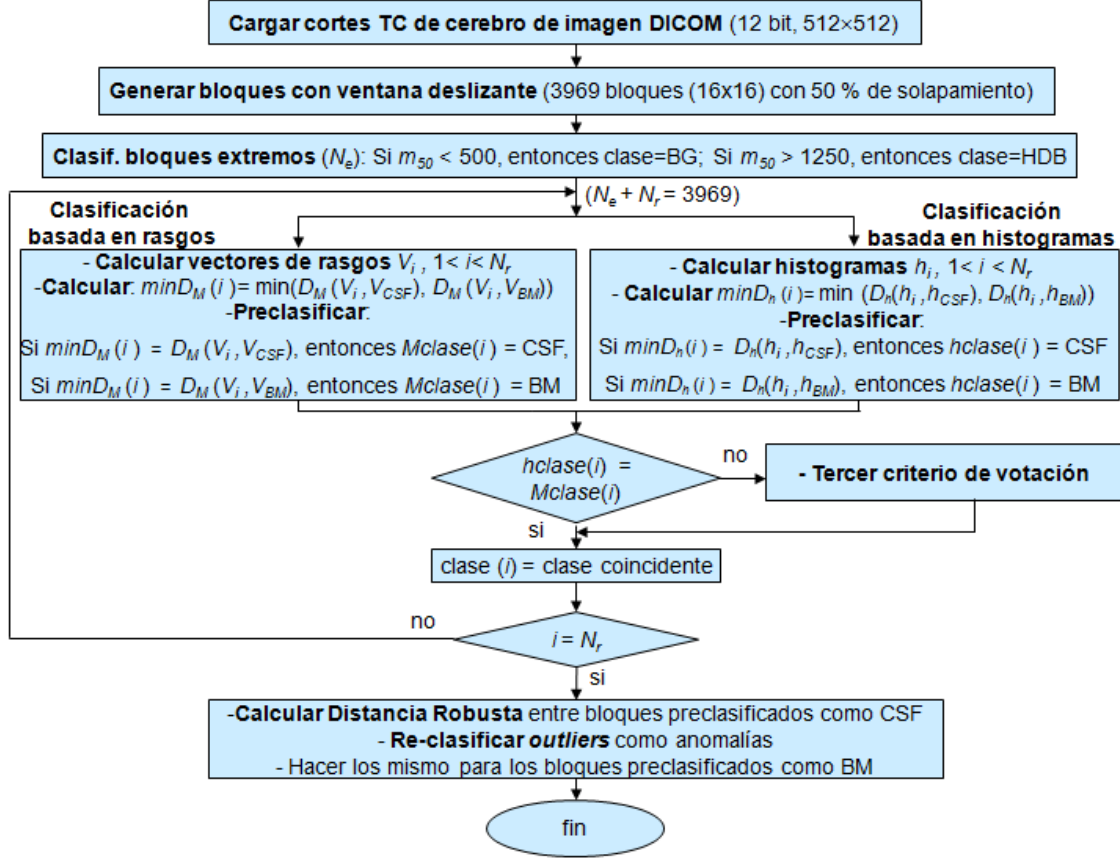


Figura 2.6. Sistema de detección de anomalías propuesto: diagrama de bloques para la etapa de clasificación.

Para la clasificación basada en rasgos tenemos:

$$\begin{aligned}
 \min DM(i) &= \min(DM(V_i, V_{CSF}), DM(V_i, V_{BM})), & 1 < i < N_r, \\
 \text{Si } \min DM(i) &= DM(V_i, V_{CSF}), \text{ entonces } Mclase(i) = CSF, & 1 < i < N_r, \\
 \text{Si } \min DM(i) &= DM(V_i, V_{BM}), \text{ entonces } Mclase(i) = BM, & 1 < i < N_r.
 \end{aligned}
 \tag{2.4}$$

Para la clasificación basada en histogramas, se emplearon las dos primeras distancias en el ranking Dh_1 y Dh_2 :

$$\begin{aligned}
 \min Dh(i) &= \min (Dh1(h_i, h_{CSF}), Dh1(h_i, h_{BM})) , & 1 < i < N_r, \\
 \text{Si } \min Dh(i) &= Dh1(h_i, h_{CSF}), \text{ entonces } h_{clase}(i) = CSF, & 1 < i < N_r, \quad (2.5) \\
 \text{Si } \min Dh(i) &= Dh1(h_i, h_{BM}), \text{ entonces } h_{clase}(i) = BM, & 1 < i < N_r.
 \end{aligned}$$

$$\begin{aligned}
 \min Dh2(i) &= \min (Dh2(h_i, h_{CSF}), Dh2(h_i, h_{BM})) , & 1 < i < N_r, \\
 \text{Si } \min Dh2(i) &= Dh2(h_i, h_{CSF}), \text{ entonces } h2_{clase}(i) = CSF, & 1 < i < N_r, \quad (2.6) \\
 \text{Si } \min Dh2(i) &= Dh2(h_i, h_{BM}), \text{ entonces } h2_{clase}(i) = BM, & 1 < i < N_r.
 \end{aligned}$$

Como hay solo dos posibles clases normales, la clasificación final es como sigue:

$$\begin{aligned}
 \text{Si } h_{clase}(i) &= M_{clase}(i), \text{ entonces } clase(i) = h_{clase}(i) , \\
 \text{sino } clase(i) &= h2_{clase}(i) .
 \end{aligned} \quad (2.7)$$

Después de estas etapas, algunas anomalías son preclasificadas como CSF y BM, o sea hay algunos falsos negativos. Para mejorar el desempeño del sistema de DA, se calculó una distancia robusta (Rousseeuw, 2006), D_{rob} , entre todos los vectores de rasgos (V) que caracterizan los bloques preclasificados en la misma clase normal,

$$D_{rob}(V) = \sqrt{(V - m_{rob})^T S_{rob}^{-1} (V - m_{rob})} \quad (2.8)$$

Se pueden detectar bloques que contienen anomalías por sus grandes valores de D_{rob} . Esta distancia robusta D_{rob} es muy similar a la D_M en (2.2), pero a diferencia del clásico centro (media), la distancia robusta usa el determinante de mínima covarianza (MCD), m_{rob} . Adicionalmente, D_{rob} usa una matriz esparcida, S_{rob} , en lugar de la matriz de covarianza en (2.2) y esta no es afectada por el efecto de enmascaramiento (Webb, 2002). El MCD es un método robusto en el sentido de que el estimado no es inapropiadamente influenciado por los datos anómalos, aun si hubiese muchos de ellos. Se usa el método MCD de Rousseeuw (Rousseeuw, 2006), implementado en la función de Matlab FASTMCD (disponible en <http://www.agoras.ua.ac.be>).

2.9 Medida para la validación del sistema de detección de anomalías

En este trabajo, se propone adaptar una nueva medida del trabajo de Gu *et al.* (2006) en la detección de intrusión, para validar los algoritmos de DA en el escenario del análisis de imágenes médicas. Esta medida, conocida como índice o capacidad de detección de anomalías, R_{AD} , puede ser vista como la relación de la información mutua, $I(X; Y)$, entre el estándar de oro, X , y la imagen de salida del sistema de DA, Y , y la entropía del estándar de oro, $H(x)$,

$$R_{AD} = \frac{I(X; Y)}{H(X)} = \frac{H(X) - H(X|Y)}{H(X)} = 1 - \frac{-\sum_x \sum_y p(x, y) \log \frac{p(x, y|x)}{p(y)}}{-\sum_x p(x) \log p(x)}, \quad (2.9)$$

donde,

$$\begin{aligned} H(X|Y) = & -R_B S_e \log \frac{R_B S_e}{R_B S_e + (-R_B)(-S_p)} - R_B (-S_e) \log \frac{R_B (-S_e)}{R_B (-S_e) + S_p (-R_B)} \\ & - (-R_B) S_p \log \frac{(-R_B) S_p}{(-R_B) S_p + R_B (-S_e)} - (-R_B) (-S_p) \log \frac{(-R_B) (-S_p)}{(-R_B) (-S_p) + S_e R_B}, \end{aligned} \quad (2.10)$$

y

$$H(X) = -R_B \log R_B - (-R_B) \log (-R_B). \quad (2.11)$$

Por lo tanto, R_{AD} ($0 \leq R_{AD} \leq 1$) toma en cuenta todos los aspectos importantes de la capacidad de detección, de acuerdo a

$$R_{AD} = 1 - \frac{a \log P_p + b \log (-P_p) + c \log (-P_n) + d \log P_n}{(-R_B + c) \log R_B + (-R_B + d) \log (-R_B)}. \quad (2.12)$$

El índice de detección de anomalías propuesto proporciona una medida intrínseca de la capacidad de DA y es sensible a los parámetros de operación del sistema (Gu, 2006). Hemos verificado todas estas ventajas sobre las medidas tradicionales. De ahí, aunque no se haya reportado antes, proponemos a R_{AD} como una medida para el ajuste fino y la validación de los algoritmos DA en el escenario del análisis de imágenes médicas.

CAPÍTULO 3. RESULTADOS Y DISCUSIÓN

En este capítulo se muestran y discuten los resultados alcanzados en las distintas etapas del sistema de detección de anomalías propuesto en el capítulo anterior.

3.1 Etapas de entrenamiento y prueba

Para las etapas de entrenamiento y prueba usando clasificación basada en histogramas, los resultados fueron casi perfectos con todas las distancias evaluadas (ver Tabla 3.1). La media armónica, la distancia χ^2 , y las normas L2 y L_∞ alcanzan 100% de clasificación correcta.

Tabla 3.1. Distancias basadas en histogramas – propiedades de clasificación y separación.

Medida	Casos de entrenamiento		Casos de prueba	
	Índice de clasificación (%)	Medida de separación (S)	Índice de clasificación (%)	Medida de separación (S)
Norma L1	100,00	0,78	99,99	0,78
Norma L2	100,00	0,76	100,00	0,76
Norma L_∞	100,00	0,77	100,00	0,77
Distancia Matusita	99,94	0,79	99,99	0,79
Distancia Bhattacharyya	99,94	0,46	99,99	0,44

Intersección de histogramas	100,00	0,78	99,99	0,78
Distancia Earth Mover's	99,94	0,90	99,98	0,90
Media armónica	100,00	0,45	100,00	0,43
Distancia χ^2	100,00	0,88	100,00	0,89
Divergencia de Kullback-Leibler	99,88	0,65	99,85	0,63

La Tabla 3.1 muestra además los resultados de la medida de separación, S , calculada como en (2.3), para evaluar las medidas basadas en histogramas en esta aplicación. En este caso, la medida ha sido calculada para ambos conjuntos, entrenamiento y prueba, y ha sido normalizada entre 0 y 1. El mejor desempeño respecto a esta propiedad es para la distancia Earth Mover's y nuevamente para la distancia χ^2 . Por consiguiente, el mejor compromiso entre ambas medidas de evaluación es para la distancia χ^2 , que será la primera opción para la clasificación basada en histogramas en las restantes etapas.

3.2 Preclasificación

Para las etapas de entrenamiento y prueba (Figura 2.5), los resultados coinciden con los valores esperados. Algunas veces, los bloques CSF fueron clasificados como BM equivocadamente. Sin embargo, el otro caso (BM clasificado como CSF) no ocurrió en los experimentos.

Combinando el clasificador basado en rasgos usando el vector (H_{thr}, m_{50}) , y el basado en histogramas con las distancias χ^2 y L^∞ como D_{h1} y D_{h2} , respectivamente, el índice de clasificación obtenido fue del 100%.

Nótese que la preclasificación realizada aquí (BM o CSF) es solo una etapa intermedia que tiene solo en cuenta los tejidos normales y que las anomalías pueden aparecer

enmascaradas. El propósito final es detectar las anomalías o lesiones que se encuentran en las imágenes de TC.

Para todos los bloques preclasificados como normales en una clase particular (BM o CSF), se ha calculado una distancia robusta como en (2.8) para reclasificar aquellos bloques cuya D_{rob} asociada está por encima de un umbral calculado con la función FASTMCD. La Figura 3.1 muestra un ejemplo donde tres bloques fueron intencionalmente adicionados en las posiciones 602, 814 y 873, de un conjunto de más de 1000 bloques.

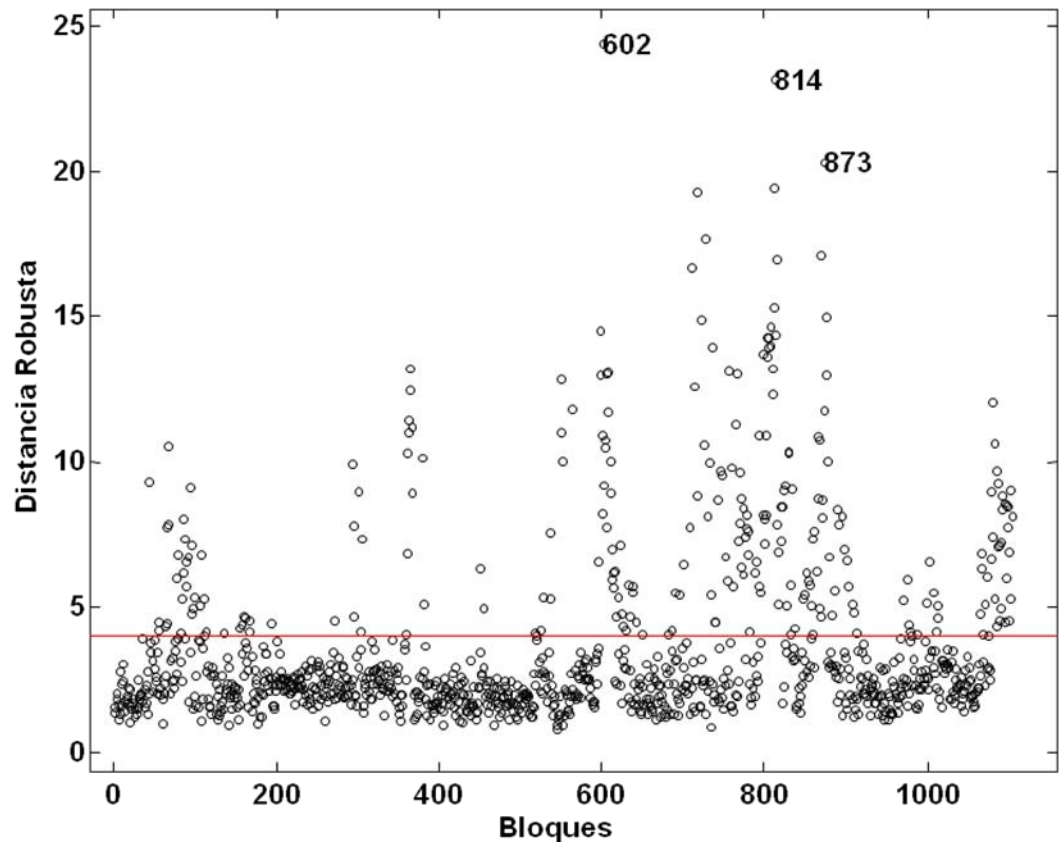


Figura 3.1. Distancia robusta vs bloques (anomalías incluidas en las posiciones 602, 814 y 873).

Nótese que aplicando (2.8) con la localización y matriz de covarianza estimada en la etapa de entrenamiento (m_{CSF} y C_{CSF} para la clase CSF, o m_{BM} y C_{BM} para la clase BM), donde fueron usados solo bloques normales, se pueden obtener resultados similares. A pesar de la

alta carga computacional, se propone una distancia robusta con el estimado MCD en su lugar debido a que este enfoque robusto puede evadir cualquier ligero problema de calibración o desempeño. Se verificó esta equivalencia, entre la distancia robusta y el enfoque más simple, con los datos disponibles del mismo equipo (Philips ‘Brilliance 40’), pero se encontró diferencias algo mayores empleando datos extras adquiridos con un sistema de TC de un modelo diferente (Siemens ‘SOMATOM AR.STAR’).

3.3 Etapa de clasificación

Se realizó la evaluación del algoritmo de DA de la Figura 2.6 para la etapa de clasificación empleando 10 cortes de nuestra base de datos, los cuales contenían algunas lesiones representativas: tumores, hemorragias y áreas de infarto (la Figura 3.2 muestra 4 de ellas). Obsérvese que algunas regiones son hiperdensas (más brillantes que el tejido que las rodea) y algunas hipodensas (más oscuras que los alrededores). Después de los últimos refinamientos propuestos, este algoritmo debe ser evaluado con la base de datos completa recolectada en este trabajo, que luego puede ser incrementada, lo que demandaría una cooperación más elevada de los especialistas para formar un estándar de oro totalmente operativo.

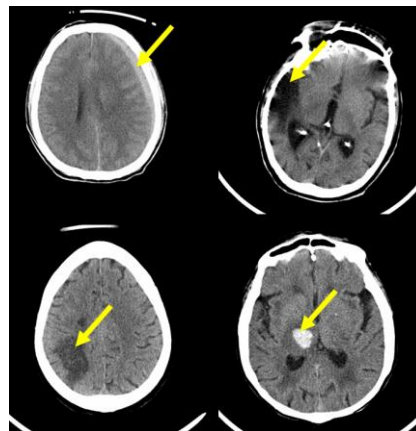


Figura 3.2. Ejemplos de lesiones a detectar con el sistema de detección de anomalías implementado.

El algoritmo de DA detecta las lesiones automáticamente. Las lesiones hiperdensas han sido detectadas casi siempre, o sea los sangramientos, hemorragias, calcificaciones o

tumores calcificados, como en el ejemplo de la Figura 3.3, aunque algunas veces estas son confundidas con la materia gris, la cual es parte de la clase normal BM.

Los bloques usados para detectar las anomalías son relativamente grandes puesto que el objetivo principal es realizar el reconocimiento, no la delineación. Adicionalmente, se permite cierta tolerancia; en este caso se admitieron diferencias de ± 1 (o sea ± 16 voxels) en las posiciones espaciales de las anomalías con respecto al estándar de oro. Si se requiriera una mejor precisión para cierta aplicación, pudiera emplearse cualquier algoritmo de delineación probado para segmentar la anomalía (Pratikakis, 2005).

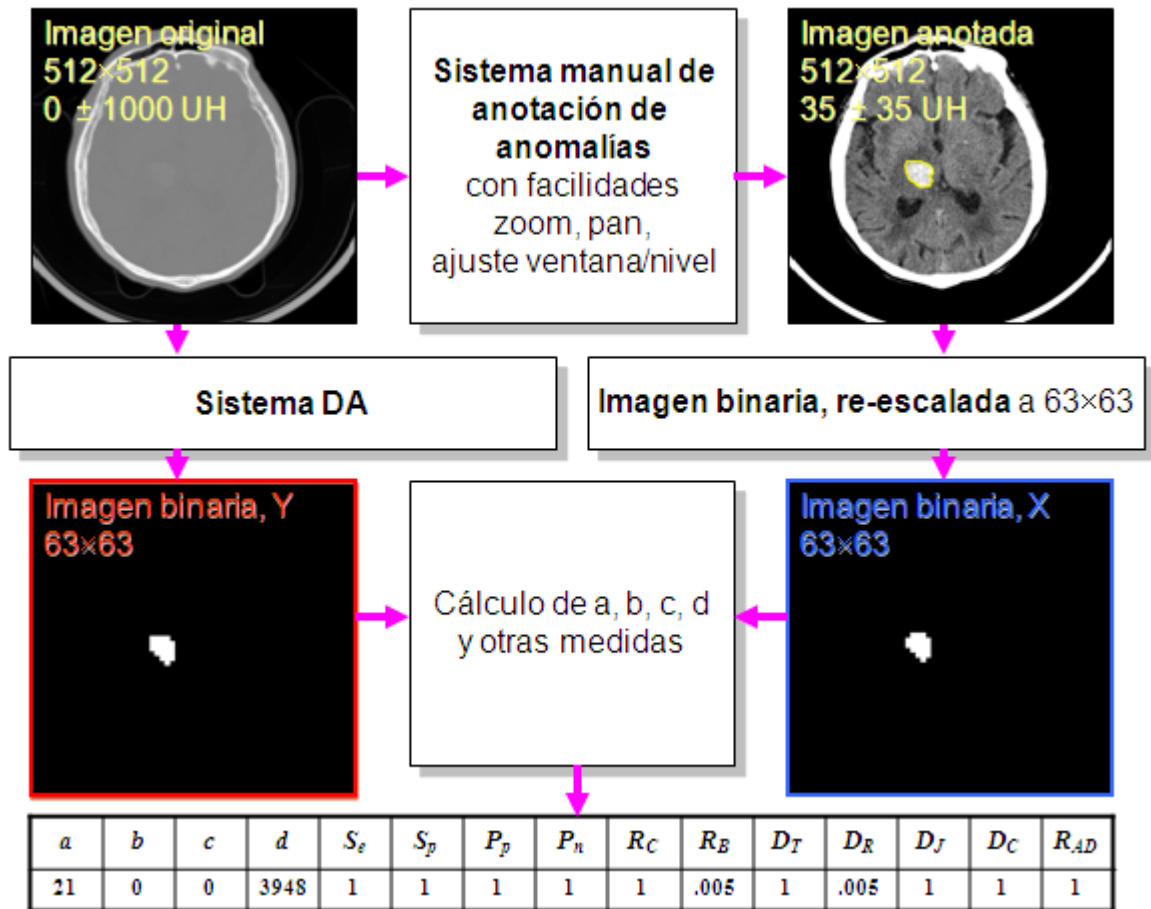


Figura 3.3. Ejemplo de evaluación del desempeño del sistema AD.

Desafortunadamente, el sistema de DA implementado en correspondencia con la Figura 2.6 no siempre es tan exitoso como en el ejemplo de la Figura 3.3, el cual alcanzó la detección de anomalías ‘perfecta’ (dentro de los límites de ± 1 cuadrícula). Los resultados finales, para un conjunto de 10 cortes, se muestran en la Tabla 3.2. Obsérvese que la distancia de

Jaccard tiene un comportamiento similar al índice de detección de anomalías, pero este último es aun más sensible a cualquier variación del desempeño.

El algoritmo de DA propuesto no puede identificar las lesiones hipodensas (como en las imágenes 3 y 4 de la Tabla 3.2) con niveles de gris entre las clases normales BM y CSF, y algunas veces totalmente solapadas con esta última. Una manera de identificar este tipo de anomalía es tomando en cuenta su posición, vecinos, tamaño y forma, lo que es fácil de calcular, pero no está implementado en este trabajo.

Tabla 3.2. Desempeño del sistema de detección de anomalías para 10 cortes de tomografía computarizada de cerebro de 512×512.

No.	1	2	3	4	5	6	7	8	9	10	total
<i>a</i>	38	37	47	21	55	60	53	18	69	63	461
<i>b</i>	3	15	6	0	2	8	6	15	12	20	87
<i>c</i>	3	54	39	0	3	5	20	5	8	23	160
<i>d</i>	3925	3863	3877	3948	3909	3896	3890	3931	3880	3863	38982
<i>S_e</i>	0,9268	0,4066	0,5465	1	0,9483	0,9231	0,7260	0,7826	0,8961	0,7326	0,7424
<i>S_p</i>	0,9992	0,9961	0,9985	1	0,9995	0,9980	0,9985	0,9962	0,9969	0,9949	0,9978
<i>P_p</i>	0,9268	0,7115	0,8868	1	0,9649	0,8824	0,8983	0,5455	0,8519	0,7590	0,8412
<i>P_n</i>	0,9992	0,9862	0,9900	1	0,9992	0,9987	0,9949	0,9987	0,9979	0,9941	0,9959
<i>R_C</i>	0,9985	0,9826	0,9887	1	0,9987	0,9967	0,9935	0,9950	0,9950	0,9892	0,9938
<i>R_B</i>	0,0103	0,0229	0,0217	0,0053	0,0146	0,0164	0,0184	0,0058	0,0190	0,0217	0,0157
<i>D_T</i>	0,9970	0,9658	0,9776	1	0,9975	0,9935	0,9870	0,9900	0,9900	0,9786	0,9876
<i>D_R</i>	0,0096	0,0093	0,0118	0,0053	0,0139	0,0151	0,0134	0,0045	0,0174	0,0159	0,0116
<i>D_J</i>	0,8636	0,3491	0,5109	1	0,9167	0,8219	0,6709	0,4737	0,7753	0,5943	0,6511

D_C	0,9268	0,5175	0,6763	1	0,9565	0,9023	0,8030	0,6429	0,8734	0,7456	0,7887
R_{AD}	0,8455	0,2706	0,4277	1	0,8904	0,8103	0,6021	0,5681	0,7592	0,5496	0,5999

Aun después de la etapa de refinamiento, empleando (2.7) y (2.8), hay un grupo considerable de bloques anómalos (alrededor del 15 ó 20 % del total) erróneamente clasificados como la clase normal CSF. Por lo tanto, sería necesario incorporar en un trabajo futuro otro criterio para discriminarlos correctamente. Un buen candidato debería tener en cuenta la posición y propiedades de las regiones donde los bloques fueron preclasificados. Se pueden calcular varias propiedades de las regiones, por ejemplo usando la función REGIONPROPS de Matlab, la cual incluye: área, centroide, excentricidad, diámetro equivalente, número de Euler, extensión, etc.

Otros datos, como la edad del paciente puede ayudar en la interpretación de los resultados. Lo que es normal en un sujeto de 81 años, por ejemplo la excesiva degradación materia cerebral y la presencia de mucho más CSF, puede ser patológico en una persona de 20 años. Este dato particular, la edad del paciente, puede ser fácilmente tomado de la información DICOM que acompaña a la imagen, para complementar el diagnóstico.

Debería tomarse en cuenta algún enfoque de *machine learning* para producir reglas útiles. De acuerdo con estas reglas, parte de las ‘regiones normales’ pueden ser reclasificadas como anomalías, mejorando así los resultados finales.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

En este trabajo se han llegado a las siguientes conclusiones:

1. Se implementó un conjunto de funciones en Matlab para extraer rasgos, distancias y medidas basadas en histogramas que pueden ser de gran utilidad para la detección de anomalías en imágenes.
2. Se implementó una aplicación de software que supera a las disponibles para anotar anomalías en imágenes médicas.
3. Se creó un estándar de oro y se establecieron procedimientos para evaluar algoritmos de detección de anomalías en imágenes de tomografía computarizada de cerebro.
4. Se evaluaron combinaciones de rasgos y distancias para detectar anomalías en imágenes de TC de cerebro, arrojando los mejores resultados, la mediana y la entropía con umbral, junto con la distancia de Mahalanobis modificada (robusta).
5. El mejor desempeño de las medidas basadas en histogramas para la detección de anomalías en imágenes de TC de cerebro, fueron para las distancias χ^2 y L^∞ .

Recomendaciones

A partir de las limitaciones actuales del trabajo se recomienda:

1. Extender la base de datos de imágenes y emplear un mayor número de especialistas en la anotación de las mismas.
2. Extender las prestaciones del software Anomaly Annotator incorporando funciones para el cálculo de histogramas, rasgos, etc.

3. Refinar las técnicas empleadas tomando en cuenta el tamaño, forma y composición de las regiones sólidas preclasificadas.
4. Implementar y probar otros rasgos, distancias y medidas basadas en histogramas para detección anomalías en imágenes de TC de cerebro.

REFERENCIAS BIBLIOGRÁFICAS

- Ashton, E. (2006). Method and system for automatic identification and quantification of abnormal anatomical structures in medical images. Patente de Estados Unidos 7,103,224.
- Astley, S.M. (2004). Computer-based detection and prompting of mammographic abnormalities. *The British Journal of Radiology*, 77, 194–S20.
- Benamrane, N., & al. (2005). A Hybrid Fuzzy Neural Networks for the Detection of Tumors in Medical Images. *American Journal of Applied Sciences*, 2(4), 892-896.
- Benamrane, N., Aribi, A., & Kraoula, L. (2006). Fuzzy Neural Networks and Genetic Algorithms for Medical Images Interpretation. *Proceedings of the Geometric Modeling and Imaging— New Trends (GMAI'06)*, IEEE Computer Society.
- Bertelli, L., Cucchiara, R., Paternostro, G., & Prati, A. (2006). A semi-automatic system for segmentation of cardiac M-mode images. *Pattern Analysis and Applications*, 9(4), 293-306.
- Bouix, S., Martin-Fernandez, M., Ungar, L., Nakamura, M., Koo, M.S., McCarley, R.W., Shenton, M.E. (2007). On evaluating brain tissue classifiers without a ground truth. *NeuroImage*, 36, 1207–1224.
- Chandola, V., Arindan, V., Kumar, V. (2007), Outlier Detection – A Survey, Reporte Técnico, Departamento de Ciencias de la Computación e Ingenierías, Universidad de Minnesota.
- Chen, M., Kanade, T., Pomerleau, D., & Rowley, H.A. (1999). Anomaly Detection through registration. *Pattern Recognition*, 32, 113-128.
- Chiracharit, W., Sun, Y., Kumhom, P., Chamnongthai, K., Babbs, C.F., & Delp, E.J. (2007). Normal mammogram detection based on local probability difference transforms and support vector machines. *IEICE Transactions on Information and Systems*, E90-D(1), 258-27.
- Daszykowski, M., Kaczmarek, K., Vander Heyden, Y., & Walczak, B. (2007). Robust statistics in data analysis — A review of Basic concepts. *Chemometrics and Intelligent Laboratory Systems*, 85, 203–219
- De Santo, M., Molinara, M., Tortorella, F., & Vento, M. (2003). Automatic classification of clustered microcalcifications by a multiple expert system. *Pattern Recognition*, 36, 1467 – 1477.
- Deselaers, T., & Müller, H. (2005). Medical Image Retrieval / Visual Features. Tutorial Medical Image Retrieval, Medical Informatics Europe.
- DiMaio, S., Kapur, T., Cleary, K., Aylward, S., Kazanzides, P., Vosburgh, K., Ellis, R., Duncan, J., Farahani, K., Lemke, H., Peters, T., Lorensen, W., Gobbi, D., Haller, J.,

- Clarke, L., Pizer, S., Taylor, R., Galloway R., Fichtinger, G., Hata, N., Lawson, K., Tempany, C., Kikinis, R., & Jolesz, F. (2007). Challenges in image-guided therapy system design. *NeuroImage*, 37, 144–151.
- Duncan, J.S., & Ayache, N. (2000). Medical Image Analysis: Progress over Two Decades and the Challenges Ahead. *IEFE Transactions On Pattern Analysis And Machine Intelligence*, 22(1), 85-108.
- Ekin, A., Jasinski, R., van der Grond, J., & van Buchem, M. (2007). Improving information quality of MR brain images by fully automatic and robust image analysis methods. *Journal Of The Society For Information Display*, 15 (6), 367-376.
- Ertöz, L., Eilertson, E., Lazarevic, A., Tan, P.-N., Kumar, V., Srivastava, J., and Dokas, P. 2004. MINDS – Minnesota Intrusion Detection System. In *Data Mining - Next Generation Challenges and Future Directions*. MIT Press.
- Falcón, A., Taboada, A., Hernández, D., Orozco, M. (2008). Anomaly Annotator for Medical Images, IV Conferencia Científica de la Universidad de las Ciencias Informáticas, Ciudad de La Habana.
- Gering, D.T. (2003). Recognizing deviations from normalcy for brain tumor segmentation. Tesis de doctorado. Departamento de Ingeniería Eléctrica y Ciencias de la Computación. Instituto de Tecnología de Massachusetts.
- Ginneken, B.V., Romeny, B.M.H., & Viergever, M.A. (2001). Computer-Aided Diagnosis in Chest Radiography: A Survey. *IEEE Trans Med Imaging*, 20, 1228-1241.
- Goldberg, I.G., Allan, C., Burel, J.-M., Creager, D., Falconi, A., Hochheiser, H., Johnston, J., Mellen, J., Sorger, P.K., Swedlow, J.R. (2005). The open microscopy environment (OME) data model and XML file: Open tools for informatics and quantitative analysis in biological imaging. *Genome Biol*, 6(5), R47.
- Goode, A., Sukthankar, R., Mummert, L., Chen, M., Saltzman, J., Ross, D., Szymanski, S., Tarachandani, A., & Satyanarayanan, M. (2007). Distributed Online Anomaly Detection in High-Content Screening. CMU-CS-07-177, Universidad Carnegie Mellon.
- Gu, G., Fogla, P., Dagon, D., & Lee, W. (2006). An Information-Theoretic Measure of Intrusion Detection Capability. *ASIACCS '06*, Marzo 21-24, Taipei, Taiwan, 90-101.
- Hersh, W. (2006). Medical Image Retrieval Protocol. ImageCLEFmed 2006, disponible en <http://www.clef-campaign.org/>, última actualización - 14 de Junio de 2006.
- Hodge, V.J., & Austin, J. (2004). A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, 22, 85–126.
- Hofer, M. (2001). Manual Páctico de TC: Introducción a la TC. 3ra Edición, Editorial Médica Panamericana, ISBN 84-7903-595-1, Madrid, España.
- Huang, Y.L., & Chen. D.R. (2004). Watershed Segmentation For Breast Tumor In 2-D Sonography. *Ultrasound in Med. & Biol.*, 30(5), 625–632.
- Huang, X., Zhou, X., Jerebko, A., Krishnan, A., Guan, H., & Kubota, T. et al. (2007). System and Method for Whole Body Landmark Detection, Segmentation and Change Quantification in Digital Images. Patente WO2007044508.
- Ikedo, Y., Fukuoka, D., Hara, T., Fujita, H., Takada, E., & Endo, T. et al. (2007). Development of a fully automatic scheme for detection of masses in whole breast ultrasound images. *Medical Physics*, 34(11), 4378-4388.
- Iyatomi, H., Oka, H., Saito, M., Miyake, A., Kimoto, M., Yamagami, J., Kobayashi, S., Tanikawa, A., Hagiwara, M., Ogawa, K., Argenziano, G., Soyer, H.P., & Tanaka, M. (2006). Quantitative assessment of tumour extraction from dermoscopy images and

- evaluation of computer-based extraction methods for an automatic melanoma diagnostic system. *Melanoma Research*, 16 (2), 183-19.
- Karnan, M., & Thangavel, K. (2007). Automatic detection of the breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications. *Computer Methods and Programs in Biomedicine*, 87(1), 12-2.
- Kelm, B.M., Menze, B.H., Zechmann, C.M., Baudendistel, K.T., & Hamprecht, F.A. (2007). Automated Estimation of Tumor Probability in Prostate Magnetic Resonance Spectroscopic Imaging: Pattern Recognition vs. Quantification. *Magnetic Resonance in Medicine*, 57,150–159.
- Kraeft, S.K., Sutherland, R., Gravelin, L., Hu, G.H., Ferland, L.H., Richardson, P., Elias, A. & Chen, L.B. (2000). Detection and analysis of cancer cells in blood and bone marrow using a rare event imaging system. *Clinical Cancer Research*, PNAS, 6(2), 434–442.
- Krivacic, R.T., Ladanyi, A., Curry, D.N., Hsieh, H.B., Kuhn, P., Bergsrud, D.E., Kepros, J.F., Barbera, T., Ho, M.Y., Chen, L.B., Lerner, R.A., & Bruce, R.H. (2004). A rare-cell detector for cancer. *Proc. of the National Academy of Sciences of the United States of America*, *Clinical Cancer Research*, 101(29), 10501 – 10504.
- Lee, C.-H., Schmidt, M., Murtha, A., Bistritz, A., Sander, J., & Greiner, R. (2005). Segmenting Brain Tumors with Conditional Random Fields and Support Vector Machines. *Lecture Notes in Computer Science*, 3765, 469-478.
- Lehmann, T.M., Meinzer, H.P., & Tolxdorff, T. (2004). Advances in biomedical image analysis - Past, present and future challenges. *Methods of Information in Medicine*, 43 (4), 308-314.
- Lehmann, T.M., Güld, M.O., Deselaers, T., Keysers, D., Schubert, H., Spitzer, K., Ney, H., & Wein, B.B. (2005). Automatic categorization of medical images for content-based retrieval and data mining. *Computerized Medical Imaging and Graphics*, 29, 143–155.
- Margineantu, D., Bay, S., Chan, P., & Lane, T. (2005). Data Mining Methods for Anomaly Detection. *KDD-2005 Workshop Report*, KDD-2005, Agosto 21, Chicago, 7(2), 132-136.
- Markou, M., & Singh, S. (2003). Novelty Detection: A Review, Part 1: Statistical Approaches. *Signal Processing*, 83, 2499–2521.
- Markou, M., & Singh, S. (2003). Novelty Detection: A Review, Part 2: Neural network based approaches. *Signal Processing*, 83, 2481–2497.
- Menze, B.H., Lichy, M.P., Bachert, P., Kelm, B.M., Schlemmer, H.-P., & Hamprecht, F.A.(2006). Optimal Classification of Long Echo Time In Vivo Magnetic Resonance Spectra in the Detection of Recurrent Brain Tumors. *Multidimensional Image Processing*, IWR, Universidad de Heidelberg.
- Messmer, P., Matthews, F., Jacob, A., Kikinis, R., Regazzoni, P., & Noser, H. (2006). A CT Database for Research, Development and Education: Concept and Potential. *Journal of Digital Imaging*, 0(0), 1-6, Publicación online 07 de Agosto de 2006.
- Miller, E.L. (1999). Statistically Based Methods for Anomaly Characterization in Images from Observations of Scattered Radiation. *IEEE Transactions on Image Processing*, 8(1), 92-101.
- Minhas, A.S., & Reddy, M.R. (2005). Neural network based approach for anomaly detection in the lungs region by electrical impedance tomography. *Physiol. Meas.*, 26, 489–502

- Montgomery, D.W.G., Amira, A. & Zaidi, H. (2007). Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model. *Medical Physics*, 34 (2), 722-736.
- Müller, H., Michoux, N., Bandon, D., & Geissbuhler, A. (2004). A Review of Content-Based Image Retrieval Systems in Medical Applications: Clinical Benefits and Future Directions. *Int. J. Med. Inform.*, 73(1), 1-23.
- Müller, H., Rosset, A., Garcia, A., Vallee, J.P., Geissbuhler, A. (2005). Benefits of Content-based Visual Data Access in Radiology. *RadioGraphics*, 25, 849-858.
- Peng, Y., Yao, B., & Jiang, J. (2006). Knowledge-discovery incorporated evolutionary search for microcalcification detection in breast cancer diagnosis. *Artificial Intelligence in Medicine*, 37(1), 43-53.
- Pokrajac, D., Megalooikonomou, V., Lazarevic, A., Kontos, D., & Obradovic, Z., (2005). Applying spatial distribution analysis techniques to classification of 3D medical images. *Artificial Intelligence in Medicine* 33, 261—28.
- Prastawa, M., Bullit, E., Ho, S., & Gerig, G. (2004). A brain tumor segmentation framework based on outlier detection. *Medical Image Analysis*, 26(8), 275–283.
- Pratikakis, I., Sahli, H., Cornelis, J., (2005). Three-Dimensional Multiscale watershed segmentation of MR images, in *Medical Image Analysis Methods*, Editor. Lena Costaridou, CRS - Taylor & Francis, 271—314.
- Radke, R.J., Andra, S., Al-Kofahi, O., & Roysam, B. (2005). Image Change Detection Algorithms: A Systematic Survey. *IEEE Transactions On Image Processing*, 14(3), 294-307.
- Rakotomamonjy, A. (2003). Variable Selection Using SVM based Criteria. *Journal of Machine Learning Research (Special Issue on Variable and Feature Selection)*, 3, 1357-137.
- Rousseeuw, P.J., Debruyne, M., Engelen, S., & Hubert, M. (2006). Robustness and Outlier Detection in Chemometrics. *Critical Reviews in Analytical Chemistry*, 36, 221–242.
- Ségonne, F. (2005). Segmentation of Medical Images under Topological Constraints. Tesis de doctorado. Departamento de Ingeniería Eléctrica y Ciencias de la Computación. Instituto de Tecnología de Massachusetts.
- Selvi, S.T., Rejani, Y.I., Sowmy, A. (2005). Breast Cancer Detection Using MMRF. *Proceedings of the International Conference on Information and Automation*, Diciembre 15-18, Colombo, Sri Lanka, 127-132.
- Shinkareva, S.V., Ombao, H.C., Sutton, B.P., Mohanty, A., & Miller, G.A. (2006). Classification of functional brain images with a spatio-temporal dissimilarity map. *NeuroImage*, 33, 63–71.
- Sklansky, J., & Ballard, D. (1973). Tumor Detection in Radiographs. *Computers and Biomedical Research*, 6(4), 299-321.
- Sluimer, I., Schilham, A., Prokop, M., & Ginneken, B. (2006). Computer Analysis of Computed Tomography Scans of the Lung: A Survey. *IEEE Trans. on Medical Imaging*, 25(4), 385-405.
- Spence, C., Parra, L., and Sajda, P. 2001. Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model. In *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*. IEEE Computer Society, Washington, DC, USA, 3.
- Spinosa, E.J., & de Carvalho, A. (2005). Support vector machines for novel class detection in Bioinformatics. *Genet. Mol. Res.* 4 (3), 608-615.

- Stoitsis, J., Golemati, S., & Nikita, K.S. (2006). A Modular Software System to Assist Interpretation of Medical Images—Application to Vascular Ultrasound Images. *IEEE Transactions On Instrumentation And Measurement*, 55(6), 1944-1952.
- Strzelecki, M., Materka, A., Drozd, J., Krzeminska-Pakula, M., & Kasprzak, J.D. (2006). Classification and segmentation of intracardiac masses in cardiac tumor echocardiograms. *Computerized Medical Imaging and Graphics*, 30, 95–107.
- Taboada, A., Sahli, H. (2008). Experimental System for Image Anomaly Detection Based on Features and Distances. *Ingeniería Electrónica, Automática y Comunicaciones*, 29(2), ISSN 0258-5944.
- Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., & Suetens, P. (2001). Automated Segmentation of Multiple Sclerosis Lesions by Model Outlier Detection. *IEEE Transactions on Medical Imaging*, 20(8), 677-688.
- Verrelst, H., & Moreau, Y. (1999). Use of Multi-Layer Perceptron to Predict Malignancy in Ovarian Tumours. Department of Electronical Engineering, Catholic University Leuven.
- Vovk, U., Pernuš, F., & Likar, B. (2007). A Review of Methods for Correction of Intensity Inhomogeneity in MRI. *IEEE Transactions On Medical Imaging*, 26(3), 405-421.
- Webb, A.R. (2002). *Statistical Pattern Recognition*. Second Edition, John Wiley & Sons Ltd, Malvern, UK, 2002.
- Wei, L., Yang, Y., Nishikawa, R.M., Wernick, M.N., & Edwards, A. (2005). Relevance vector machine for automatic detection of clustered microcalcifications. *IEEE Transactions on Medical Imaging*, 24(10), 1278-1285.
- Xiaoxue, Y., Hsu, W., Lee, W.S., & Lozano-Perez, T. (2004). Abnormality detection in retinal images. Reporte Técnico MIT-3845, MIT, Instituto de Tecnología de Massachussets.
- Xie, X. (2006). Defect Detection in Random Colour Textures. Tesis de doctorado. Facultad de Ingeniería. Departamento de Ciencias de la Computación. Universidad de Bristol.
- Xu, R., Zhao, X., Li, X., Kwan, C., & Chang, C.-I, (2006). Target Detection with Improved Image Texture Feature Coding Method and Support Vector Machine. *International Journal Of Intelligent Technology*, 1(1), 47-56, ISSN 1305-6417.
- Zhou, J., Lim, T.-K., Chong, V., Huang, J. (2003). Segmentation and visualization of nasopharyngeal carcinoma using MRI. *Computers in Biology and Medicine*, 33, 407–424.
- Zöllei, L. (2006). A Unified Information Theoretic Framework for Pair- and Group-wise Registration of Medical Images. Tesis de doctorado. Departamento de Ingeniería Eléctrica y Ciencias de la Computación. Instituto de Tecnología de Massachussets.

ANEXOS

Anexo I Rasgos

En este anexo se presenta una tabla con los rasgos empleados en la detección de anomalías en imágenes médicas. Junto a cada rasgo aparece la expresión matemática (referenciada, ej. A1) para calcularlo, así como el lugar ocupado por este en el ranking (idoneidad para la clasificación) empleando la función FSV del paquete para Matlab Spider.

Tabla A1. Diferentes rasgos evaluados en el contexto de la imaginología de tomografía computarizada de cerebro.

No.	Rasgo	Fórmula (x es el bloque de n voxels bajo comprobación; p es la distribución de probabilidades de k bins de x . Para los rasgos basados en GLCM: N es el número de niveles de gris reescalados (aquí, $N = 20$), y P_{ij} es el número de ocurrencias del par de niveles de gris i y j)	R A N K I G
1	Entropía de Shannon	$H_{Sh} = -\sum_{i=1}^k p_i \cdot \log(p_i), \text{ con } 0 \cdot \log(0) = 0$ (A1)	1
2	Entropía de norma	$H_{norm} = \sum_{i=1}^k p_i ^p$	2

	$(p = 2)$	(A2)	
3	Entropía Log Energía	$H_{\log E} = - \sum_{i=1}^k \log \left(\frac{E_i}{E} \right), \text{ con } \log(0) = 0$ (A3)	11
4	Entropía con umbral ($T_1 = 1020$)	$H_{thr} = - \sum_{i: x_i > T_1} \log \left(\frac{x_i}{T_1} \right)$ (A4)	12
5	Entropía SURE ($T_2 = 1065$)	$H_{SURE} = n + \sum_{i=1}^n \min \left(\frac{x_i^2}{T_2^2}, 1 \right) - \sum_{i: x_i \leq T_2} \log \left(\frac{x_i}{T_2} \right)$ (A5)	3
6	Media aritmética (o simplemente, media)	$m_a = \mu = \frac{1}{n} \sum_{i=1}^n x_i$ (A6)	8
7	Media recortada ($\alpha=10\%$)	m_α = media de x_i excluyendo los valores mayores y menores que $\alpha/2$ de las observaciones	5
8	Mediana	m_{50} = valor de la variable que deja el mismo número de datos antes y después que él, una vez ordenados estos	4
9	Media armónica	$m_h = n \left(\sum_{i=1}^n \frac{1}{x_i} \right)^{-1}$ (A7)	6
10	Media geométrica	$m_g = \sqrt[n]{\prod_{i=1}^n x_i}$ (A8)	7
11	Rango	$Rango = \max(x) - \min(x)$	32

		(A9)	
12	Rango intercuartil	$IQR = x_{[0.75 \cdot n]} - x_{[0.25 \cdot n]}$ (A10)	34
13	Desviación estándar	$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}}$ (A11)	33
14	Varianza	$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}$ (A12)	36
15	Desviación media absoluta (basada en la media)	$MAD_1 = \frac{\sum_{i=1}^n x_i - \mu }{n}$ (A13)	37
16	Desviación media absoluta (basada en la mediana)	$MAD_2 = \text{median} (x_i - m_{50})$ (A14)	38

17	Skewness	$A_s = \frac{\sum_{i=1}^n \frac{(x_i - \mu)^3}{n}}{\sigma^3}$ (A15)	42
18	Kurtosis	$A_p = \frac{\sum_{i=1}^n \frac{(x_i - \mu)^4}{n}}{\sigma^4}$ (A16)	41
19, 20	Contraste GLCM (este, sur)	$Contrast = \sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2$ (A17)	22, 25
21, 22	Disimilitud GLCM (este, sur)	$Dissimilarity = \sum_{i,j=0}^{N-1} P_{i,j} i - j $ (A18)	21, 28
23, 24	Similitud GLCM (este, sur)	$Similarity = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + i - j }$ (A19)	24, 32
25, 26	Homogeneidad GLCM (este, sur)	$Homogeneity = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2}$ (A20)	23, 29
27, 28	Segundo momento angular GLCM (este, sur)	$ASM = \sum_{i,j=0}^{N-1} P_{i,j}^2$ (A21)	17, 13

29, 30	Energía GLCM (este, sur)	$Energy = \sqrt{\sum_{i,j=0}^{N-1} P_{i,j}^2}$ (A22)	18, 14
31, 32	Máxima probabilidad GLCM (este, sur)	$Maxpro = \max P_{i,j}$ (A23)	35, 16
33, 34	Entropía GLCM (este, sur)	$Entropy = \sum_{i,j=0}^{N-1} P_{i,j} \log P_{i,j}$ (A24)	19, 15
35, 36	Media GLCM (este, sur)	$\mu_i = \sum_{i,j=0}^{N-1} iP_{i,j} \quad \mu_j = \sum_{i,j=0}^{N-1} jP_{i,j}$ (A25)	10, 9
37, 38	Varianza GLCM (este, sur)	$\sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j} (i - \mu_i)^2 \quad \sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j} (j - \mu_j)^2$ (A26)	31, 27
39, 40	Desviación estándar GLCM (este, sur)	$\sigma_i = \sqrt{\sum_{i,j=0}^{N-1} P_{i,j} (i - \mu_i)^2} \quad \sigma_j = \sqrt{\sum_{i,j=0}^{N-1} P_{i,j} (j - \mu_j)^2}$ (A27)	26, 20
41, 42	Correlación GLCM (este, sur)	$Correl = \sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(i - \mu_i)^2 (j - \mu_j)^2}} \right]$ (A28)	40, 39

Anexo II Distancias basadas en histogramas

En este anexo se presenta una tabla con un conjunto de distancias basadas en histogramas, empleadas en la detección de anomalías en imágenes médicas. Aparece además la expresión matemática para el cálculo de las distancia junto con una referencia a las mismas (ej. A29).

Tabla A2. Distancias basadas en histogramas.

Distancia	Fórmula (n es el número de <i>bins</i> en los histogramas x e y)
Norma L1	$D_{L1}(x, y) = \sum_{i=1}^n x_i - y_i \quad (A29)$
Norma L2	$D_{L2}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (A30)$
Norma L_∞	$D_{L_\infty}(x, y) = \max_{1 \leq i \leq n} x_i - y_i \quad (A31)$
Distancia Matusita	$D_{Matu}(x, y) = \sqrt{\sum_{i=1}^n (\sqrt{x_i} - \sqrt{y_i})^2} \quad (A32)$
Distancia Bhattacharyya	$D_{Bhat}(x, y) = -\log \sum_{i=1}^n \sqrt{x_i \cdot y_i} \quad (A33)$
Intersección de histogramas	$D_{hi}(x, y) = 1 - \frac{\sum_{i=1}^n \min(x_i, y_i)}{\sum_{i=1}^n x_i} \quad (A34)$
Distancia χ^2	$D_{\chi^2}(x, y) = \sum_{i=1}^n \frac{(x_i - y_i)^2}{x_i + y_i} \quad (A35)$

Divergencia Kullback-Leibler	$D_{KL}(x, y) = \sum_{i=1}^n \left(x_i - y_i \right) \log \frac{x_i}{y_i} \quad (A36)$
Media armónica de Kullback-Leibler (simétrica)	$D_{KLS}(x, y) = \frac{1}{\frac{1}{\sum_{i=1}^n x_i \log \frac{x_i}{y_i}} + \frac{1}{\sum_{i=1}^n y_i \log \frac{y_i}{x_i}}} \quad (A37)$
Distancia Earth Mover's	$D_{EM}(x, y) = \frac{\sum_{i=1}^m \sum_{ij=1}^n f_{ij} d_{ij}}{\sum_{i=1}^m \sum_{i=1}^n f_{ij}}, \quad (A38)$ $work(x, y, F) = \sum_{i=1}^m \sum_{j=1}^n f_{ij} d_{i,j} \quad (A39)$