

INSTITUTO TECNOLÓGICO DE CHIHUAHUA
DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN

***“MEDICIÓN DE LA COMPLEJIDAD DE
IMÁGENES MEDIANTE UN MÉTODO EN EL
DOMINIO ESPACIAL”***

TESIS

QUE PARA OBTENER EL GRADO DE

***MAESTRO EN CIENCIAS
EN INGENIERÍA ELECTRÓNICA***

PRESENTA:

ING. DANIEL ALBERTO POSADA URRUTIA

**DIRECTOR DE LA TESIS:
DR. MARIO IGNACIO CHACÓN MURGUÍA**



SEP
SECRETARÍA DE
EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO



CHIHUAHUA, CHIH., SEPTIEMBRE 2019



EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Instituto Tecnológico de Chihuahua

"2019, Año del Caudillo del Sur, Emiliano Zapata"

Chihuahua, Chih. **17 de septiembre de 2019**

**C. DANIEL ALBERTO POSADA URRUTIA
PRESENTE**

Por este conducto le comunico que a propuesta del Jurado de Examen, la División de Estudios de Posgrado e Investigación ha concedido autorización para la impresión de su tesis para obtener el grado de en Ciencias en Ingeniería Electrónica, cuyo título es:

"Medición de la complejidad de imágenes mediante un método en el dominio espacial"

La tesis presenta el siguiente contenido de capítulos:

- I Antecedentes
- II Análisis de bases de datos
- III Características para la complejidad objetiva
- IV La resolución y sus efectos en la complejidad
- V Sistemas para determinar la complejidad
- VI Resultados y conclusiones

ATENTAMENTE

Excelencia en Educación Tecnológica

"La técnica por el Engrandecimiento de México"

**MTRO. LUIS CARDONA CHACÓN
JEFE DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN**



SECRETARÍA DE
EDUCACIÓN PÚBLICA
INSTITUTO TECNOLÓGICO
DE CHIHUAHUA



Ave Tecnológico No. 2909 Col. 10 de Mayo C.P. 31310, Chihuahua, Chih. México
Tel. 01 (614) 201 2000, (614) 413 5187, Ext. 2150 e-mail: dir_chihuahua@tecnm.mx

www.tecnm.mx | www.itchihuahua.edu.mx





Chihuahua, Chih. **17 de septiembre de 2019**

MTRO. LUIS CARDONA CHACÓN
JEFE DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN
PRESENTE

Por medio de la presente notificamos a usted que en cumplimiento de los requerimientos para la obtención del grado de en Ciencias en Ingeniería Electrónica, el documento de tesis del **C. DANIEL ALBERTO POSADA URRUTIA**, ha sido aprobado y aceptado para su impresión. El título de la tesis es:

"Medición de la complejidad de imágenes mediante un método en el dominio espacial"

Por lo que proponemos, le sea concedida la autorización de impresión correspondiente.

Agradeciendo la atención a la presente, quedamos de usted:

ATENTAMENTE

*Excelencia en Educación Tecnológica®
"La técnica por el Engrandecimiento de México"*

DR. MARIO IGNACIO CHACÓN MURGUÍA
DIRECTOR DE TESIS

Juan Ramirez

DR. JUAN ALBERTO RAMÍREZ QUINTANA
MIEMBRO DEL JURADO DE EXAMEN

MTRA. ALMA DELIA CORRAL SÁENZ
MIEMBRO DEL JURADO DE EXAMEN



SECRETARÍA DE
EDUCACIÓN PÚBLICA
INSTITUTO TECNOLÓGICO
DE CHIHUAHUA

DR. JAVIER VEGA PINEDA
MIEMBRO DEL JURADO DE EXAMEN





CARTA CESIÓN DE DERECHOS

En la ciudad de Chihuahua el día , el que suscribe **C. DANIEL ALBERTO POSADA URRUTIA** de la Maestría en Ciencias en Ingeniería Electrónica, con número de control **G17061501**, adscrito a la **DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN**, del Instituto Tecnológico de Chihuahua, manifiesta que es autor intelectual del presente trabajo de Tesis bajo la dirección del **Dr. Mario Ignacio Chacón Murguía** y cede los derechos del trabajo titulado ***"Medición de la complejidad de imágenes mediante un método en el dominio espacial"***, al Tecnológico Nacional de México y/o Instituto Tecnológico de Chihuahua para su difusión, divulgación, transmisión, reproducción, así como su digitalización con fines académicos y de investigación.

C. DANIEL ALBERTO POSADA URRUTIA



Chihuahua, Chihuahua, a 27 de septiembre de 2019

Dra. María Elena Álvarez-Buylla Roces
Directora del CONACYT

Por este conducto aprovecho la ocasión para saludarla e informarle que a la fecha he obtenido el grado de **Maestría en Ciencias de Ingeniería Electrónica** en la División de Estudios de Posgrado e Investigación del Instituto Tecnológico de Chihuahua. Agradezco todo el apoyo brindado por CONACYT ya que el otorgamiento de la beca para estudios de Posgrado, permitió dedicarme de tiempo completo al programa y con ello lograr el cumplimiento del objetivo principal del convenio establecido.

Sin otro particular por el momento, me es grato quedar de usted como su seguro servidor, no sin antes reiterar nuevamente todo mi agradecimiento.

A t e n t a m e n t e



Daniel Alberto Posada Urrutia
Exbecario CONACYT

AGRADECIMIENTOS

A Dios

Por bendecirme con tanto, brindarme lo necesario para terminar esta etapa de mi vida y haber colocado en mi camino a todas esas personas tan especiales que me acompañaron y apoyaron.

A mis padres y hermanos

Por su amor, trabajo y sacrificio al darme una formación personal y profesional. Sin ustedes simplemente no estaría aquí.

A mi novia

Por ser mi apoyo incondicional en mi vida, que con su amor y respaldo me ayuda a alcanzar mis objetivos y sueños.

A mi familia

Abuelos, tíos y primos por formar parte en mi vida y un muy especial agradecimiento a mi tía Rita y Alejandro por darme un lugar que puedo llamar mi segundo hogar.

A mis amigos y compañeros

Mi primo Eduardo, Adrián, Raúl, Luis, Abimael, José, Jorge, Óscar, Manuel, Soto, David, Xavier y Javier que hicieron mi estadía en el tecnológico más agradable.

A mi comité de tesis

Mi sincero agradecimiento al Dr. Mario Chacón, principal colaborador durante todo este proceso, quien con su dirección, conocimiento y colaboración permitió el desarrollo de este trabajo. Como también a la Mtra. Alma Corral, Dr. Juan Ramírez y Dr. Javier Vega quienes fueron parte de mi formación educativa.

Al Tecnológico Nacional de México

Gracias al Tecnológico Nacional de México / I. T Chihuahua por el apoyo recibido para el proyecto “Algoritmo de detección de movimiento en secuencias de video basado en arquitectura de redes neuronales tradicionales o de aprendizaje profundo y lógica difusa”, clave 5162.19-P.

RESUMEN

MEDICIÓN DE LA COMPLEJIDAD DE IMÁGENES MEDIANTE UN MÉTODO EN EL DOMINIO ESPACIAL

Ing. Daniel Alberto Posada Urrutia
Maestría en Ciencias en Ingeniería Electrónica
División de Estudios de Posgrado e Investigación del
Instituto Tecnológico de Chihuahua
Chihuahua, Chih. 2019
Director de Tesis: Dr. Mario Ignacio Chacón Murguía

La complejidad de la imagen ha sido utilizada en áreas del procesamiento digital de imágenes tales como detección de bordes, compresores, segmentación, etc. También se ha utilizado en áreas de psicología donde se analiza cómo ésta afecta en la memoria de trabajo e incluso en el análisis de los anuncios comerciales. Sin embargo, la complejidad es una medida subjetiva que se basa en la percepción visual del humano y puede variar debido a conocimientos, experiencias, costumbres, etc.

En esta investigación, la complejidad de una imagen se define como una medida de la cantidad de detalles contenidos en ella y que la hacen difícil de evaluar. Además, se propone un análisis y un método para disminuir la subjetividad en esta medida mediante la selección de características computables que tengan correlación con la complejidad de imagen que un humano estima como: densidad de bordes, relación de compresión, color y entropía del patrón. El primer método se trata de sistemas de inferencia de lógica difusa en el que se generó un conjunto de reglas de inferencia y se construyeron los valores difusos con algoritmos de agrupamiento. El segundo método consiste en el uso de redes neuronales artificiales, en el que se plantea un método iterativo que consiste en buscar la mejor cantidad de neuronas y la mejor red que modele la complejidad de imágenes. Los sistemas de inferencia de lógica difusa, lograron un desempeño promedio de 0.80 en correlación al criterio humano, lo cual es aceptable al estar en el promedio de la literatura investigada. Los sistemas con redes neuronales artificiales lograron un desempeño promedio de 0.89, el cual es mayor al de lógica difusa y más alto que la mayoría de los métodos de la literatura investigada.

CONTENIDO

LISTA DE FIGURAS.....	xi
-----------------------	----

LISTA DE TABLAS.....	xvi
----------------------	-----

CAPÍTULO I. ANTECEDENTES	1
---------------------------------------	----------

1.1 Introducción	1
------------------------	---

1.2 Percepción visual	2
-----------------------------	---

1.2.1 Psicología de la percepción visual	3
--	---

1.2.1.1 Leyes de percepción visual	4
--	---

1.3 Complejidad de imágenes.....	5
----------------------------------	---

1.3.1 La complejidad de una imagen en psicología	8
--	---

1.3.2 La complejidad de una imagen en ciencias de la computación	11
--	----

1.4 Métodos para determinar la complejidad de una imagen	12
--	----

1.4.1 Criterio humano	13
-----------------------------	----

1.4.2 Detección de bordes	14
---------------------------------	----

1.4.3 Compresión de imágenes	16
------------------------------------	----

1.4.4 Distribución de colores	19
-------------------------------------	----

1.4.5 Combinado	20
-----------------------	----

1.4.6 Otros métodos	22
---------------------------	----

1.5 Conclusión	23
----------------------	----

CAPÍTULO II. ANÁLISIS DE BASES DE DATOS	25
--	-----------

2.1 Selección de las bases de datos.....	26
--	----

2.2 Análisis de las bases de datos RS1 y RS2	27
--	----

2.2.1 Determinación de la complejidad subjetiva	29
---	----

2.2.2 Filtrado de las bases de datos RS1 y RS2	30
2.3 Análisis de la base de datos SAVOIAS	34
2.3.1 Determinación de la complejidad	35
2.3.2 Métodos para completar la matriz.....	37
2.3.3 Puntajes finales	38
2.4 Conclusión	39
 CAPÍTULO III. CARACTERÍSTICAS PARA LA COMPLEJIDAD OBJETIVA	40
3.1 Selección de características	40
3.1.1 Criterios para la selección de características	40
3.2 Extracción de características	45
3.2.1 Bordes	46
3.2.2 Compresión	48
3.2.3 Niveles de intensidad	50
3.2.4 Distribución.....	52
3.3 Análisis de características.....	54
3.3.1 Densidad de bordes	54
3.3.1.1 Fortalezas en densidad de bordes	54
3.3.1.2 Debilidades en densidad de bordes	54
3.3.2 Relación de compresión.....	55
3.3.2.1 Fortalezas en relación de compresión	55
3.3.2.2 Debilidades en relación de compresión.....	56
3.3.3 Medida de complejidad del color.....	57
3.3.3.1 Fortalezas en medida de complejidad del color	57

3.3.3.2 Debilidades en medida de complejidad del color.....	58
3.3.4 Entropía del patrón.....	59
3.3.4.1 Fortalezas en entropía del patrón	59
3.3.4.2 Debilidades en entropía del patrón.....	59
3.4 Conclusión	60
CAPÍTULO IV. LA RESOLUCIÓN Y SUS EFECTOS EN LA COMPLEJIDAD	61
4.1 Las características de complejidad objetiva y la resolución.....	62
4.2 Correlación entre complejidad subjetiva y las características	65
4.3 La complejidad subjetiva y la resolución	67
4.4 Conclusión	68
CAPÍTULO V. SISTEMAS PARA DETERMINAR LA COMPLEJIDAD	70
5.1 Datos de diseño y validación	70
5.2 Preprocesamiento de características	71
5.3 Complejidad por lógica difusa.....	72
5.3.1 Variables de entrada del sistema difuso.....	74
5.3.1.1 Generación de funciones de membresía.....	74
5.3.2 Determinación de las reglas de inferencia.	77
5.3.3 Ejemplo de diseño de un sistema difuso	78
5.3.4 Salida del sistema difuso.....	82
5.4 Complejidad con redes neuronales artificiales	83
5.4.1 Diseño de la RNA para determinar la complejidad de imágenes	83
CAPÍTULO VI. RESULTADOS Y CONCLUSIONES	90

6.1 Selección del algoritmo de agrupamiento para el diseño del FIS.....	90
6.2 Correlación entre características propuestas.....	91
6.3 Resultados de comparación contra la literatura	92
6.4 Tiempos de procesamiento de los sistemas	98
6.5 Conclusiones.....	100
REFERENCIAS.....	102
ANEXOS.....	109
APÉNDICES.....	113

LISTA DE FIGURAS

CAPÍTULO I. ANTECEDENTES	1
Figura 1.1. Ejemplos de principios de percepción visual de Gestalt, a) emergencia, b) reificación, c) multiestabilidad, d) invarianza.	3
Figura 1.2. Ejemplos de leyes de visión de Gestalt, a) proximidad, b) semejanza, c) cerramiento, d) buena continuidad, d) pregnancia.	5
Figura 1.3. Esquema general de la complejidad de imágenes.	6
Figura 1.4. a) Imagen muy compleja, b) imagen poco compleja.	7
Figura 1.5. Imagen usada para demostrar la diferencia entre ver y visión [14].	9
Figura 1.6. Imagen de la cabeza de la vaca remarcada [14].	9
Figura 1.7. Proceso general para determinar la CI por medio de un algoritmo.....	11
Figura 1.8. Proceso general para determinar la CI utilizando metodología del criterio humano.	14
Figura 1.9. Detección de bordes con Matlab haciendo uso de técnicas de Sobel y Canny.	15
Figura 1.10. Proceso general para determinar la complejidad de la imagen con compresión.....	19
Figura 1.11. Dos imágenes con el mismo histograma y entropía pero diferente complejidad [11].	20
Figura 1.12. Proceso general de aplicación del método combinado.	21
 CAPÍTULO II. ANÁLISIS DE BASES DE DATOS	 25
Figura 2.1. Imágenes de la base de datos RS1 de [31].	28
Figura 2.2. Imágenes de la base de datos RS2 de [31].	29

Figura 2.3. Interfaz de usuario para establecer la complejidad subjetiva de las imágenes de las bases de datos RS1 y RS2.	30
Figura 2.4. Evaluaciones de RS1 por a) sujeto 2, b) sujeto 26.	31
Figura 2.5. a) Imágenes 1 y 2 con complejidad superior a 80, b) imágenes 31 y 32 con complejidad inferior a 10, según sujeto 26.	32
Figura 2.6. a) Gráfica evaluación del sujeto 8, b) gráfica evaluación del sujeto 29 para RS2.	33
Figura 2.7. a) Imagen con complejidad menor a 5, b) imagen con complejidad mayor a 95 según sujeto 29 en RS2.	33
Figura 2.8. Algunos ejemplos de imágenes de la base de datos SAVOIAS, ordenadas de menor a mayor complejidad.	35
Figura 2.9. Ejemplo de cómo lucía la interfaz de usuario que se utilizó para realizar la decisión forzosa de la imagen más compleja.	36
Figura 2.10. Ejemplo de la matriz R	37

CAPÍTULO III. CARACTERÍSTICAS PARA LA COMPLEJIDAD OBJETIVA40

Figura 3.1. Gráfica de puntuaciones alcanzadas por cada característica y resaltadas las seleccionadas.	44
Figura 3.2. a) Bordes marcados, b) bordes tenues.	47
Figura 3.3. Estructura del bloque DCT [27].	49
Figura 3.4. Ejemplo de generación de la matriz de coocurrencia.	53
Figura 3.5. a) Imagen con muchos bordes, b) imagen con pocos bordes.	54
Figura 3.6. Dos imágenes con diferentes complejidades y densidades de bordes.	55
Figura 3.7. Dos imágenes que generan misma relación de compresión.	56
Figura 3.8. Dos imágenes con diferente relación de compresión, pero mismo contenido.	57

Figura 3.9. Dos imágenes con la misma cantidad de bordes, pero diferentes niveles de intensidad.....	58
Figura 3.10. Ejemplo de cuatro imágenes con dos niveles de intensidad distintos.	58
Figura 3.11. Dos imágenes con diferente patrón.	59
Figura 3.12. Dos imágenes con un patrón similar.	60

CAPÍTULO IV. LA RESOLUCIÓN Y SUS EFECTOS EN LA COMPLEJIDAD61

Figura 4.1. Imagen de letra ‘D’ con resolución de 200 x 200 pixeles, a) original, b) escalada al 0.5, c) escalada a 0.25, d) escalada a 0.125 y d) escalada al 0.01.	61
Figura 4.2. Gráficas de complejidad contra escala de imágenes RS1, RS2 y anuncios.	63
Figura 4.3. Gráficas de complejidad contra escala de imágenes arte y diseño de interiores.	63
Figura 4.4. Gráficas de complejidad contra escala de imágenes objetos y escenarios.	63
Figura 4.5. Gráficas de complejidad contra escala de imágenes suprematismo y visualizaciones.	64
Figura 4.6. Imagen 168 de escenas de la base de datos SAVOIAS y su complejidad para diferentes escalas.	64
Figura 4.7. Imagen 148 de escenas de la base de datos SAVOIAS y su complejidad para diferentes escalas.	65
Figura 4.8. Gráficas de correlación entre complejidad subjetiva y características contra escala de imágenes RS1, RS2 y anuncios.....	66
Figura 4.9. Gráficas de correlación entre complejidad subjetiva y características contra escala de imágenes de arte y diseño de interiores.	66
Figura 4.10. Gráficas de correlación entre complejidad subjetiva y características contra escala de imágenes de objetos y escenas.	66

Figura 4.11. Gráficas de correlación entre complejidad subjetiva y características contra escala de imágenes de suprematismo y visualizaciones.	67
Figura 4.12. Ejemplo de una reducción en la escala de la imagen al tener que ser mostrada en un monitor.	68
Figura 4.13. Ejemplo de un aumento en la escala de la imagen al tener que ser mostrada en un monitor.	68

CAPÍTULO V. SISTEMAS PARA DETERMINAR LA COMPLEJIDAD70

Figura 5.1. Esquema de validación cruzada.	70
Figura 5.2. Esquema general de un sistema difuso tipo Sugeno para determinar la complejidad de imágenes.....	73
Figura 5.3. Diagrama de flujo del diseño del FIS propuesto.	73
Figura 5.4. Ejemplo de los resultados obtenidos de las funciones de membresía.	76
Figura 5.5. Imágenes para el ejemplo de diseño del FIS.	78
Figura 5.6. Agrupamiento de cantidad de objetos.	79
Figura 5.7. Agrupamiento de cantidad de colores.	80
Figura 5.8. Posibles consecuentes de las reglas.....	80
Figura 5.9. Diagrama de una regla de un sistema Sugeno.....	82
Figura 5.10. Sistema difuso propuesto para determinar la complejidad de imágenes.....	83
Figura 5.11. Esquema general de la red neuronal propuesta.	84
Figura 5.12. Diagrama de flujo para buscar la mejor RNA.....	86
Figura 5.13. Ejemplo de inicializaciones diferentes y su repercusión en la minimización del error de la RNA.	87
Figura 5.14. Correlación de diferentes réplicas de la RNA de 77 neuronas en capa oculta para el grupo 5.	88

Figura 5.15. Desempeño promedio de validación y entrenamiento para redes neuronales con diferente cantidad de neuronas en la capa oculta.	89
--	----

CAPÍTULO VI. RESULTADOS Y CONCLUSIONES90

Figura 6.1. Tiempo de procesamiento de los sistemas propuestos y PSO de [31].	99
--	----

LISTA DE TABLAS

CAPÍTULO I. ANTECEDENTES	1
CAPÍTULO II. ANÁLISIS DE BASES DE DATOS	25
Tabla 2.1. Bases de datos.....	25
Tabla 2.2. Distintos tipos de imágenes en bases de datos.	26
CAPÍTULO III. CARACTERÍSTICAS PARA LA COMPLEJIDAD OBJETIVA	40
Tabla 3.1. Puntuación según el nivel de detalle.	42
Tabla 3.2. Pesos de los criterios para elegir características.	42
Tabla 3.3. Lista de abreviaturas utilizadas para describir la editorial.	43
Tabla 3.4. Características y criterios para su elección.....	44
Tabla 3.5. Características y criterios ponderados para su elección.	45
Tabla 3.6. Nivel de borde según el cambio en el tono de gris.....	47
CAPÍTULO IV. LA RESOLUCIÓN Y SUS EFECTOS EN LA COMPLEJIDAD	61
CAPÍTULO V. SISTEMAS PARA DETERMINAR LA COMPLEJIDAD	70
Tabla 5.1. Complejidad subjetiva para el ejemplo del diseño del FIS.	78
Tabla 5.2. Salida deseada para el ejemplo del diseño del FIS.....	78
Tabla 5.3. Características para el ejemplo del diseño del FIS.	79
Tabla 5.4. Características objetivas escaladas para el ejemplo del diseño del FIS.	79
Tabla 5.5. Imágenes que cumplen con las reglas seleccionadas.....	81
CAPÍTULO VI. RESULTADOS Y CONCLUSIONES	90

Tabla 6.1. Desempeños de los sistemas difusos al utilizar diferentes algoritmos de agrupamiento en las bases de datos RS1 y RS2.	90
Tabla 6.2. Correlación entre las características seleccionadas, utilizando las bases de datos RS1, RS2 y SAVOIAS.....	91
Tabla 6.3. Correlación promedio para las bases de datos RS1 y RS2.....	93
Tabla 6.4. Correlación de las características seleccionadas contra diferentes tipos de imágenes de la base de datos SAVOIAS.....	94
Tabla 6.5. Correlación de las características del estado del arte según [40] en su base de datos.	95
Tabla 6.6. Comparación entre resultados de correlación con los métodos propuesto y con los mejores resultados de [40] en la base de datos SAVOIAS.	95
Tabla 6.7. Comparación entre sistemas propuestos y algunos de la literatura.	97
Tabla 6.8. Tiempos de extracción de características y propagación de sistemas propuestos.....	98
Tabla 6.9. Tiempos de extracción de las características y propagación del sistema de [31].....	99

NOTACIÓN

Notación	Descripción
$I(x,y)$	Imagen
x	Índice renglón de la imagen
y	Índice columna de la imagen
CI	Complejidad de la imagen
s	Complejidad subjetiva de la imagen
i, j, l	Índices
I_{bi}	Imagen de bordes de nivel i
ρ_b	Densidad de bordes
N	Cantidad total de pixeles de la imagen
N_b	Cantidad total de pixeles bordes de la imagen
N_{bl}	Cantidad total de pixeles bordes del nivel l de la imagen
RC	Relación de compresión
T_B	Tamaño de la imagen en bytes
T_b	Tamaño de la imagen en bits
TC_B	Tamaño de la imagen comprimida en bytes
bpp	Bits por pixel
p	Pixel
$p_{bi(x,y)}$	Pixel borde de la imagen de bordes de nivel i
PC	Puntaje de la característica
CC	Correlación de la característica con el criterio humano
CR	Criterios de evaluación de las características
wc	Pesos de los criterios para elegir características
s_{ij}	Complejidad subjetiva de la imagen i por el sujeto j
s_i	Complejidad subjetiva de la imagen i
zS	z-score de la complejidad subjetiva
R	Matriz de proporciones para determinar complejidad subjetiva
$r_{i,j}$	Proporción en que la imagen i ha sido elegida como más compleja que j
$P(I_i > I_j)$	Probabilidad de elegir la imagen i sobre la imagen j como más compleja
$F(\Delta_{i,j})$	Función sigmoidea de la diferencia de puntajes subjetivos
$\Delta_{i,j}$	Diferencia entre complejidades subjetivas de las imágenes i y j
MAP	Máximo a posteriori
$\log$	Logaritmo
Ψ	Complejidad objetiva de la imagen
$\psi(x,y)$	Complejidad del pixel (x,y)
Ψ_{Bl}	Complejidad de la imagen según densidad de bordes del nivel l

Ψ_C	Complejidad de la imagen según la relación de compresión
Ψ_L	Complejidad de la imagen según los niveles de intensidad
Ψ_D	Complejidad de la imagen según la distribución del patrón
G_α	Función de ponderación gaussiana
G	Filtro gaussiano
∇p	Cambios en el tono de gris
λ_{1i}	Umbral inferior del ∇p para ser clasificado como borde nivel i
λ_{2i}	Umbral superior del ∇p para ser clasificado como borde nivel i
D	Distancia
E	Distancia euclidiana
CCM	Medida de complejidad de color
$\Omega(i, j)$	Ventana local del pixel (i, j)
τ	Número de píxeles en la ventana local Ω
γ	Factor de normalización
L	Componente L del espacio de color CIELab
a	Componente a del espacio de color CIELab
b	Componente b del espacio de color CIELab
$\Theta(u)$	Transformada discreta cosenoidal
β	Factor de la transformada discreta cosenoidal
H_x	Renglones totales de la imagen I
H_y	Columnas totales de la imagen I
O	Matriz de coocurrencia
q	Cantidad de conjuntos para diseño
d	Cantidad de imágenes de los conjuntos de diseño
gr	Grupo de validación cruzada
g	Índice de característica $g=1$, densidad de bordes nivel 1 $g=2$, densidad de bordes nivel 2 $g=3$, relación de compresión $g=4$, medida de complejidad del color $g=5$, entropía del patrón
$\Psi_{\hat{g}}$	Complejidad normalizada según la característica g
argmin	Valor mínimo
argmax	Valor máximo
t	Objetivos (<i>targets</i>) para el diseño de sistemas
μ	Función de membresía
FCM	Algoritmo de agrupamiento Fuzzy C-Means

KM	Algoritmo de agrupamiento K-Means++
GG	Algoritmo de agrupamiento Gath-Geva
GK	Algoritmo de agrupamiento Gustafson-Kessel
A	Grupo generado por un algoritmo de agrupamiento
R_n	Conjunto de reglas del sistema difuso
ω	Peso de la regla del FIS
k	Cantidad de grupos buscados por el algoritmo de agrupamiento para los valores difusos
C	Conjunto de centroides
c_i	Centroide i
x_d	Dato de diseño
X_d	Conjunto de datos de diseño
x_d'	Dato de diseño, posible nuevo centroide
σ	Desviación estándar
h	Índice de valor difuso
κ	Cantidad de grupos buscados por el algoritmo de agrupamiento para los posibles consecuentes de las reglas
B_m	Conjunto de imágenes con características que disparan a la regla m con un nivel distinto de cero
M	Número de reglas del sistema difuso
η	Constantes
z	Consecuente de la regla
w	Pesos sinápticos de las neuronas
W_h	Pesos sinápticos de la capa oculta de la RNA
W_s	Pesos sinápticos de la capa de salida de la RNA
S_{hi}	Suma de las entradas ponderadas a la neurona i de la capa oculta
S_{si}	Suma de las entradas ponderadas a la neurona i de la capa de salida
o	Número de neuronas de la capa oculta de la RNA
Φ	Función de activación
α	Coefficiente del exponencial

CAPÍTULO I. ANTECEDENTES

1.1 Introducción

La visión es uno de los sentidos más desarrollados para el ser humano [1], y por medio de éste puede conocer el medio que lo rodea, por lo que tiene un papel fundamental en la vida diaria. Las personas reciben gran cantidad de información con tan solo echar un vistazo, sin embargo, la visión no solo se limita a obtener la imagen del medio, también involucra a la percepción visual, la cual es un proceso constructivo que interpreta los estímulos visuales con los conocimientos previos del individuo. El propósito de la visión es proveer una imagen con un significado coherente con el conocimiento a priori sobre el mundo [2], lo cual le da sentido a la imagen que se ve. No es de sorprender que las imágenes jueguen un rol importante en la vida diaria del humano tanto que, en la actualidad, el mundo se encuentra inmerso en un sinnúmero de imágenes que son utilizadas para transmitir información, por ejemplo, publicidad con imágenes, señales visuales de tráfico, fotografías, pinturas, aplicaciones móviles, entre muchos ejemplos más.

Las imágenes cuentan con diferentes características como colores, resolución, abstracción, el tipo o clase de imagen (fotografía, dibujo, pintura, etc.). Estas pueden ser utilizadas para una mejor interpretación de la idea que el autor quiere transmitir. La complejidad de la imagen es una característica que permite una interpretación rápida de la imagen si ésta es baja; por otro lado, si ésta es alta la imagen puede contener más detalles y su interpretación es más lenta. Consecuentemente, la complejidad puede ser útil en diferentes formas de análisis, por ejemplo, si lo que se busca es que la imagen sea entendida rápidamente como las señales de tráfico, la complejidad debe ser baja.

La complejidad de imágenes es un punto importante en el campo de procesamiento de imágenes y reconocimiento de patrones [5] ya que el conocer o estimar dicha medida puede ser de gran utilidad para diversas aplicaciones como detección de bordes, detección de saliencia visual y área de interés [6], ocultamiento de mensajes en imágenes, e inclusive en seguridad vial, ya que estudios demuestran que una visión muy compleja puede incrementar el riesgo de accidentes de autos, debido a que la atención de las personas puede ser fácilmente perdida cuando la visión es muy compleja [7].

La complejidad en imágenes ha sido definida de diferentes maneras, por ejemplo, en [3] la definen como una medida subjetiva percibida por personas respecto al grado de atención requerido para analizar la imagen, mientras que [4] la definen como la cantidad de líneas y detalles contenidos en la imagen. Tomado en cuenta las distintas definiciones encontradas en la literatura, se optó por definir a la complejidad de imágenes como una medida de la cantidad de detalles e información contenida que dificulta el análisis de la imagen.

Dentro de este capítulo se abordan conceptos para describir lo que es la complejidad en imágenes, y también se presentan algunas metodologías utilizadas en la literatura para evaluarla.

1.2 Percepción visual

La vista es uno de los sentidos más desarrollados y tiene un rol fundamental en la percepción del humano ya que a través de ella se recibe gran parte de la información del medio en que una persona se mueve [8]. Aunque parezca lógico creer que se ve con los ojos, es más correcto decir que el cerebro es quien ve, ya que la mayoría de los procesos de visión se producen en él, mientras que los ojos pueden considerarse como los receptores de este proceso [9].

Los modelos de visión actuales consideran por lo menos tres aspectos: fisiológico, neurológico y psicológico. El aspecto fisiológico abarca las partes físicas que intervienen en la vista (partes del ojo), lo neurológico va orientado hacia el sistema nervioso que interviene en la visión (corteza cerebral, tálamo, etc), mientras que el aspecto psicológico abarca el estudio del conocimiento, la percepción y los procesos mentales [9]. No es de interés en esta investigación considerar los aspectos fisiológicos y neurológicos debido a que se parte del hecho de que a las personas que participan en la estimación de la complejidad tienen una vista saludable. Lo que sí es de interés para esta investigación es la parte psicológica, ya que se ha demostrado que conocimientos y experiencias pueden impactar en la estimación de la complejidad de la imagen [4].

1.2.1 Psicología de la percepción visual

Las teorías y leyes de Gestalt son un referente importante ya que describen y predicen algunos principios de la psicología de la percepción del humano. Los autores de Gestalt llegaron a la conclusión de que la información que llega al cerebro no es interpretada por partes aisladas, sino que se tiende a realizar agrupamientos y percibir objetos como un todo organizado [10].

A continuación, se enlistan cuatro principios claves utilizados por los autores de Gestalt para describir procesos de percepción visual.

- Emergencia: reconocimiento global, es el surgimiento de objetos a partir de la información recibida.
- Reificación: construcción de nuevas formas partiendo de las existentes gracias a la propia experiencia visual.
- Multiestabilidad: percepción ambigua entre fondo y figura que se da en algunas imágenes.
- Invariancia: prioridad en el reconocimiento y percepción de las formas y contornos de los elementos por encima de otras cualidades: color, textura, estilo, etc.

En la Figura 1.1 se muestran ejemplos de estos cuatro principios.

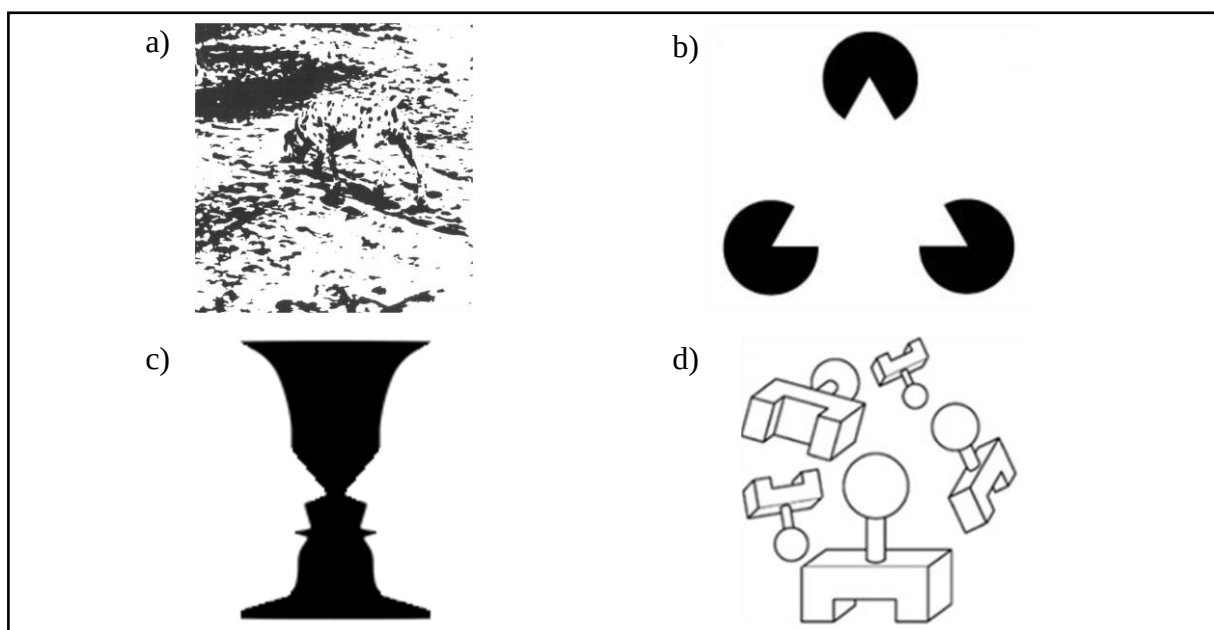


Figura 1.1. Ejemplos de principios de percepción visual de Gestalt, a) emergencia, b) reificación, c) multiestabilidad, d) invarianza.

La Figura 1.1a intenta demostrar el principio de emergencia, la cual trata de un perro olfateando cerca de la sombra de un árbol, esto emerge como un todo y no como sus partes individuales (cabeza, patas, cola, etc). La Figura 1.1b ilustra el principio de reificación, aquí se percibe la figura de un triángulo aun y cuando esta no existe, esto se debe a experiencias visuales. En la Figura 1.1c se muestra un ejemplo de multiestabilidad, se puede estar observando dos caras o bien una copa. Por último, la Figura 1.1d es un ejemplo de invarianza, aquí es posible percibir que se trata del mismo objeto aun y cuando éste ha sido trasladado y rotado.

1.2.1.1 Leyes de percepción visual

Los autores de Gestalt determinaron una serie de leyes sobre la percepción visual [10], algunas de ellas que se consideran importantes se describen brevemente a continuación:

- Ley de proximidad: establece que los objetos cercanos tienden a ser vistos como una unidad. En la Figura 1.2a se observa este caso, donde los puntos aparentan formar filas, ya que los puntos están más cercanos de forma horizontal que vertical.
- Ley de semejanza: establece que los objetos similares tienden a ser percibidos como una unidad. En la Figura 1.2b se observa cómo se alternan las filas de puntos por agrupación por color.
- Ley de cerramiento: establece que se tienden a cerrar las brechas existentes en una posible figura con líneas incompletas, las líneas que delimitan una superficie son percibidas más fácilmente como unidad que aquellas que no llegan a cerrarse. La Figura 1.2c muestra cómo percibimos el círculo aun y cuando no logra cerrarse.
- Ley de buena continuidad: establece que aquellas figuras que tengan un destino común se constituyen en unidad más fácilmente. La Figura 1.2d muestra cómo se percibe dos líneas continuas que se intersectan, en lugar de dos ángulos que se encuentran en un punto.
- Ley de la pregnancia: establece que la interpretación de algo que se observa es aquella que resulta más eficiente y requiere menos esfuerzo perceptual, esto es, la más simple que requiere menos atención. En la Figura 1.2e se muestra un ejemplo,

en el cual lo más simple es interpretar que se trata de tres cuadrados sobre puestos en vez de ver dos cuadrados que tienen una esquina inferior cortada.

- Ley de la experiencia: establece que la experiencia previa del sujeto receptor interviene con los aspectos descritos anteriormente para la constitución de las formas percibidas. El ejemplo de esta ley se puede apreciar más fácilmente en la Figura 1.1a mostrada anteriormente, donde una vez que el sujeto ya ha visto que se trata de un perro dálмата olfateando el suelo cerca de una sombra de un árbol, fácilmente lo volverá a relacionar cuando vuelva a ver la imagen.

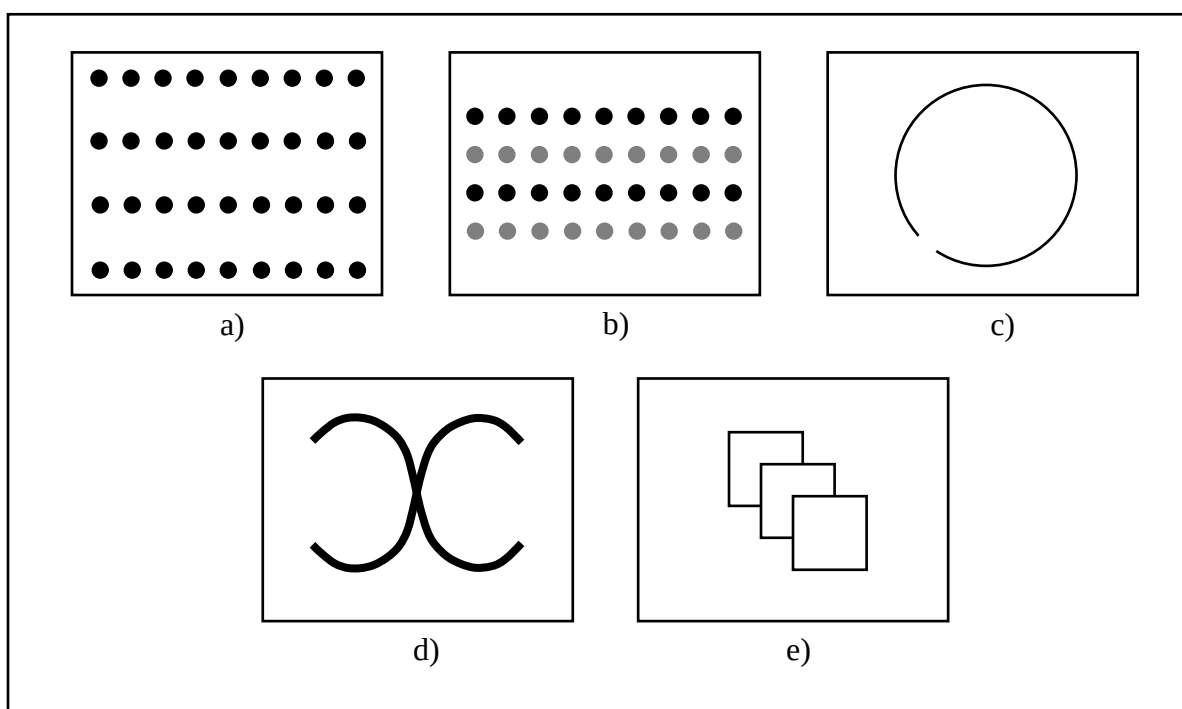


Figura 1.2. Ejemplos de leyes de visión de Gestalt, a) proximidad, b) semejanza, c) cerramiento, d) buena continuidad, e) pregnancia.

1.3 Complejidad de imágenes

Una imagen puede ser definida como una función de dos dimensiones, $I(x,y)$, donde x y y son coordenadas espaciales, e I la amplitud, intensidad o nivel de gris de cada par de coordenadas (x,y) para imágenes en tonos de grises. Cuando los valores de x , y y la amplitud son todos finitos (cantidades discretas), se tratará de una imagen digital [1].

Complejidad de imagen (CI) se ha definido en esta investigación como una medición del grado de detalles e información que dificulta su análisis, esto es, entre más detalles contenga la imagen mayor atención será requerida para analizarla y evaluarla. Por lo tanto, la complejidad será mayor que en comparación con una imagen con pocos detalles que fácilmente es analizada y evaluada. La complejidad se subdivide en dos grandes clases: complejidad subjetiva y complejidad objetiva, en donde la primera es la complejidad que estima una persona a través de su percepción visual, mientras que la complejidad objetiva es aquella que se obtiene a través de características que puedan ser estimadas y medidas a través de una computadora. En la Figura 1.3 se muestra el esquema general de la complejidad de imágenes.

La complejidad dice muchos aspectos importantes del contenido de una imagen dado que está directamente relacionada con los detalles. La CI es factor importante en la selección de material fuente para probar varios métodos de procesamiento de imágenes [11], o bien de procesos de percepción visual, de hecho una manera simple de estimar la similitud entre dos imágenes puede ser la diferencia entre sus complejidades.

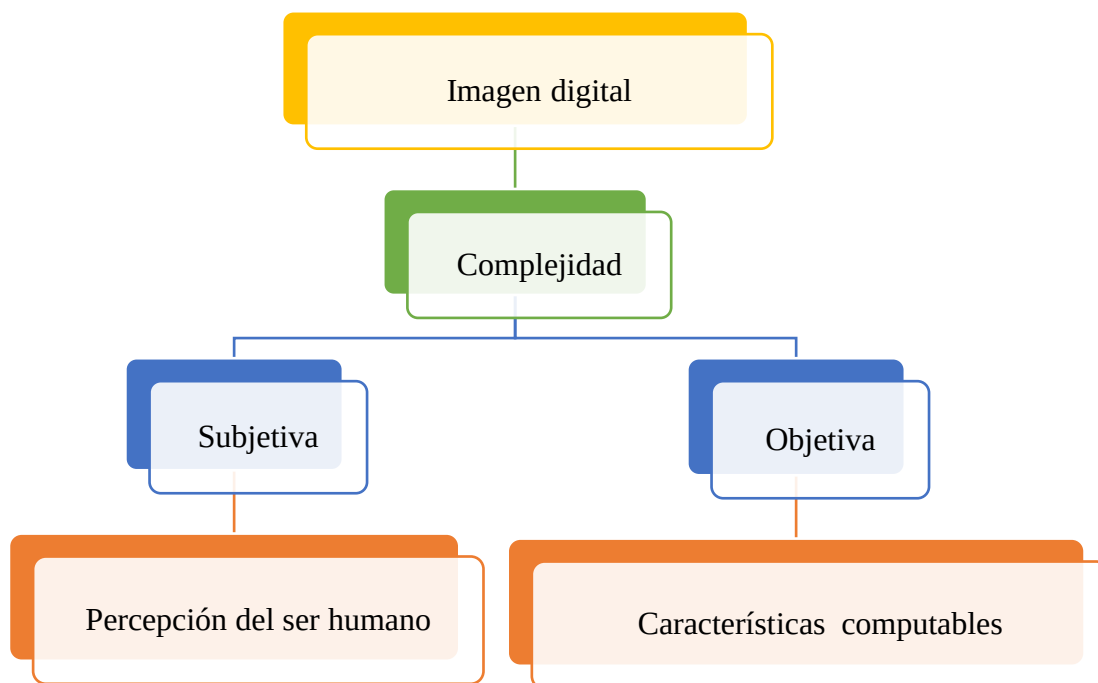
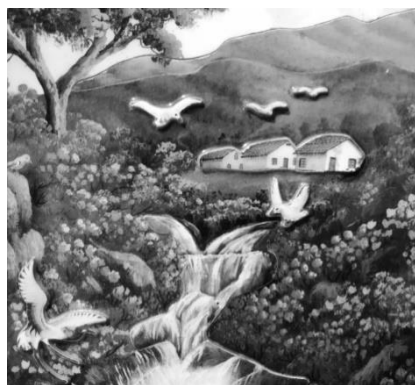


Figura 1.3. Esquema general de la complejidad de imágenes.

En la Figura 1.4 se muestra un ejemplo entre dos niveles de complejidad. La Figura 1.4a muestra una imagen con gran cantidad de detalles por lo tanto es muy compleja. Por otro lado, la Figura 1.4b muestra una imagen simple con un objeto y pocos detalles, por lo tanto es poco compleja.



a)



b)

Figura 1.4. a) Imagen muy compleja, b) imagen poco compleja.

Desafortunadamente, la complejidad en imágenes ha recibido relativamente poca atención en la investigación y esto es debido en parte a la ausencia de una método aceptable para estimarla [12]. De acuerdo con [13], una forma es mediante la complejidad de Kolmogorov, la cual se define como la longitud del programa más corto requerido para reproducir un resultado, en este caso los píxeles en una imagen [14]. Sin embargo, la complejidad de Kolmogorov no es computable y aun cuando lo fuera, el resultado sería cuestionable, ya que la similitud no está siempre relacionada a la complejidad, esto debido a que el contexto en la imagen juega un rol importante, por ejemplo una imagen de un animal y otra de ruido correlacionado pueden tener ambas la misma información de KCS (Kolmogorov-Chaitin-Solomonoff), pero solo la imagen del animal tiene sentido.

El problema de evaluar la CI es de cierta relevancia tanto en ciencia de la computación como en psicología cognitiva. Científicos de la computación usualmente analizan las dimensiones espaciales con el objetivo de automatizar problemas de visión como la extracción de características. Por otro lado, los psicólogos se ven más interesados en la dimensión temporal de la complejidad como un medio para explorar modelos de atención [15].

1.3.1 La complejidad de una imagen en psicología

En psicología, la evaluación de la complejidad de la imagen va más orientada hacia el análisis y exploración de los procesos del cerebro, como lo son la memoria de trabajo, modelos de atención e incluso en el ámbito comercial [16], entre otros más.

Uno de los métodos más utilizados para evaluar la complejidad de la imagen tanto en psicología como en ciencias de la computación es el criterio humano, el cual es un proceso que genera una evaluación subjetiva debido a que está relacionada directamente con lo que las personas estimen. En 1974, E. Berlyne enfatizó que la complejidad es en realidad subjetiva ya que ésta depende de la relación entre propiedades de estímulos físicos y estadísticos de los objetos y el proceso con el humano [2]. Estudios han demostrado que la subjetividad de la complejidad está relacionada con familiaridad, por lo que se sugiere a los investigadores considerar la relación de la complejidad con la familiaridad o bien, saber qué es lo que realmente se quiere de la medida de la complejidad visual y en caso de que la familiaridad afecte en los resultados esperados, considerar remover este factor de la ecuación. Sin embargo, de quererse esto surge la duda de saber si realmente se garantiza que se ha removido, lo cual da lugar a un gran campo de investigación [17].

La estimación de la CI dada por personas no es meramente directa con la complejidad inherente de la imagen. La cantidad de información contenida en una imagen puede variar sin cambiar la complejidad de ésta [2]. La Figura 1.5 es un experimento que tiene el propósito de demostrar lo descrito anteriormente. Para esto se pide al observador que mire la imagen por pocos segundos. Tener en cuenta que esto debe ser una evaluación rápida, ya que si se le da mucho tiempo el observador puede llegar interpretar la imagen lo cual afecta en el experimento. Después, se le indica que describa que es lo que vio y cuál es la complejidad que estima de la imagen. En la mayoría de los casos los observadores al ver la imagen por pocos segundos solo perciben una imagen de tonos de grises aleatorios y esto provoca que la evalúen con una complejidad más alta que aquellas personas que logran interpretar el contenido de la imagen.

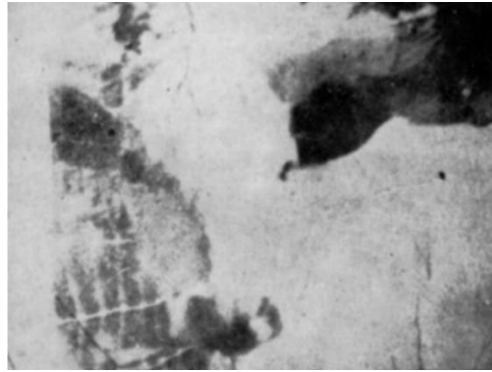


Figura 1.5. Imagen usada para demostrar la diferencia entre ver y visión [14].

Inicialmente, la imagen aparenta ser píxeles con valores de grises aleatorios, sin embargo, esto solo ocurre la primera vez que se vea, o al menos hasta que emerja el contexto de la imagen, el cual se trata de la cabeza de una vaca, donde las manchas negras que se encuentran por el centro superior de la imagen son las orejas, y la mancha inferior central es la nariz. En la Figura 1.6 se marca el contorno de la cabeza de la vaca para facilitar la visualización. Esta imagen de la Figura 1.6 es ampliamente utilizada por el *optometric extension program foundation* para demostrar las diferencias entre ver y visión [14], donde, si al inicio no se presta atención y solo se ve la imagen, difícilmente se logrará realizar una interpretación, por otro lado la visión es cuando la imagen pasa por el proceso de interpretación del contexto de ésta.



Figura 1.6. Imagen de la cabeza de la vaca remarcada [14].

Una vez que el observador ha logrado ver el contexto de la imagen difícilmente volverá a ver la imagen como solo píxeles con valores de grises aleatorios, y esto es debido al principio de percepción visual de Gestalt (emergencia). Después que el observador ha interpretado el

contexto de la imagen, se le vuelve a preguntar sobre lo que ve y cuál es la complejidad con la que evalúa a la imagen. Es común que el observador decida evaluar la imagen con una complejidad más baja que antes de interpretar su contenido, ya que la CI se ve influida por la familiaridad y conocimientos previos que se tiene con el contenido de la imagen. De acuerdo con [2] un patrón o una imagen puede ser más complejo o ambiguo para algunas personas que para otras, incluso puede cambiar para una misma persona en diferentes tiempos, lo cual es lo que se trataba de demostrar con el experimento.

En [18] se estudia cómo la ambigüedad de una imagen influye subjetivamente al ser juzgada o clasificada de acuerdo a su complejidad, aun y cuando los análisis de parámetros cuantitativos muestran que no existen diferencias entre las imágenes. Llegando a discusiones de cómo imágenes ambiguas fueron calificadas significativamente más complejas que las imágenes menos ambiguas, donde los valores objetivos de la complejidad mostraron no tener diferencias significativas entre ambos grupos.

Por otro lado, ignorando la subjetividad de la medida de la CI se ha demostrado que la complejidad puede influir en procesos del cerebro como la atención, donde estudios demuestran que una visión de una imagen muy compleja puede incrementar el riesgo de accidentes automovilísticos, esto debido a que la atención de las personas puede ser fácilmente distraída cuando la imagen que ve el piloto es demasiado compleja [7].

También existe una relación cercana entre la CI y la memoria visual de trabajo, ya que al incrementar la complejidad se reduce la eficiencia de la memoria de trabajo, dado que la prioridad de la atención juega un rol importante para nuestras habilidades de enumerar y recordar pequeños grupos de estímulos [19].

Además de estos estudios, la CI ha sido utilizada en el comercio, o más concretamente en el mercado en línea, donde se ha analizado cómo la complejidad de los anuncios puede impactar en la efectividad de estos, en muchas de las ocasiones, el vendedor añade demasiada información al anuncio para promover su producto haciendo que al final se pierda la información más importante que el vendedor quiere hacer llegar al comprador [16].

Como ya se ha mencionado, la evaluación de la CI puede verse afectada por el mecanismo de percepción humana ya que es sumamente complejo. En la etapa de análisis de lo observado, intervienen gran cantidad de factores subjetivos relacionados con los criterios y experiencias

particulares del observador, por lo que la CI es una medida altamente subjetiva cuando un humano la estima.

1.3.2 La complejidad de una imagen en ciencias de la computación

En ciencias de la computación, generalmente se estudian las dimensiones espaciales de la imagen para extraer características que puedan ser utilizadas en diferentes aplicaciones [15]. De hecho la CI indica muchos aspectos del contenido de la imagen, por lo que puede ser un factor clave para la selección del material fuente para probar algoritmos de procesamiento de imágenes [11], por ejemplo en detección de áreas de interés, exaltamiento de regiones, detección de bordes, compresión de imágenes, entre otras.

En la literatura se encuentran diversos métodos para determinar la CI por medio de algoritmos. Generalmente, el resultado del algoritmo se correlaciona con el criterio humano, el cual es una medida subjetiva que puede variar de acuerdo al criterio de la persona que está evaluando la imagen, por lo que estos métodos generalmente ofrecen una aproximación hacia la complejidad subjetiva. Aunado a esto, los algoritmos para medir la CI tienen la ventaja de que al ser métodos de computación, arrojarán un valor exacto que no se ve afectado por el tiempo o algunos otros factores como familiaridad de la imagen.

Dentro de la literatura consultada no existe un estándar o métrica universal para determinar la CI, lo cual da lugar a un campo de investigación donde se estudian y evalúan diferentes métodos para estimarla y/o aproximarla al criterio humano. En la Figura 1.7 se muestra el proceso general con el que se calcula la complejidad de una imagen de manera automática a través de un algoritmo.

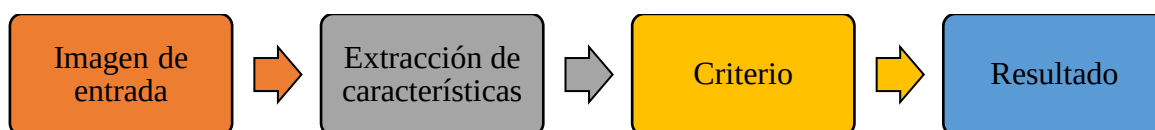


Figura 1.7. Proceso general para determinar la CI por medio de un algoritmo.

A continuación, se describe brevemente lo que engloba cada punto del proceso general para determinar la CI por medio de un algoritmo.

- **Imagen de entrada.** La imagen es la señal o estímulo a la cual se le estimará la complejidad, debe estar libre de ruido. Si éste no es eliminado de la imagen será tomado como parte de ella, lo cual afectará en el resultado de la medición de la complejidad, por lo que esta sección podría subdividirse en entrada y pre procesamiento de la imagen.
- **Extracción de características.** Conjunto de técnicas para encontrar las características relevantes de la CI, este punto puede tratarse de extracción de bordes, compresión de imagen, distribución de los colores, etc.
- **Criterio.** Determina la complejidad a partir de las características obtenidas. Esta sección puede constar de la aplicación de alguna ecuación, o ponderación de las características para llegar a un resultado final que será el de la medición de la CI.
- **Resultado.** Es lo que se obtiene después de aplicar el criterio para la evaluación de CI. Generalmente, se correlaciona con el criterio humano para que el algoritmo se comporte como una persona que evaluaría las imágenes. Este resultado puede ser utilizado para más cosas, por ejemplo, se puede relacionar con el umbral que se le debe colocar a algún algoritmo de procesamiento de imágenes, lo cual brindará un proceso más independiente del humano al automatizar esa parte de selección del parámetro.

Como ya se ha mencionado, la complejidad es una medida subjetiva cuando una persona la evalúa, debido a que depende de la percepción, la cual es un mecanismo humano sumamente complejo en el que intervienen muchos procesos subjetivos relacionados con los criterios y experiencias particulares del observador, por lo que en este sentido los algoritmos computacionales son una opción más viable [3]. Además de que éstos también tienen la fiabilidad de que la medición no va a cambiar de un momento a otro, generando resultados constantes que pueden ser utilizados en diferentes aplicaciones más autónomas.

1.4 Métodos para determinar la complejidad de una imagen

Para determinar o estimar la complejidad de una imagen existen diferentes métodos en la literatura. A continuación, se enlistan algunos de éstos que se encontraron e investigaron. Estos

fueron clasificados con base en sus características principales para la determinación de la complejidad de la imagen.

1.4.1 Criterio humano

El método de criterio humano es uno de los procesos más comunes para categorizar las imágenes según su complejidad (s), que consiste en preguntar directamente a un grupo de personas con cuanta complejidad evalúan un conjunto determinado de imágenes. La metodología consiste en establecer previamente una base de datos de imágenes. De esta base de datos se debe tener un conocimiento previo, es decir, el encargado del experimento deberá elegir imágenes realmente representativas para lograr el objetivo, que en este caso es la complejidad. Una vez que se cuenta con la base de datos de imágenes, se reúne al grupo de personas dispuestas a participar, éstas deben de estar saludables, esto es, gozar de una buena salud tanto psicológica como visual, o bien que su vista este corregida. Es importante establecer los criterios con los que se estimará la complejidad, esto consiste en un entrenamiento previo donde se dan indicaciones e instrucciones a las personas que evaluarán la base de datos de imágenes, sobre cómo determinar la complejidad y también se da un ejemplo que les sirva de guía. Las indicaciones para los participantes pueden ser concentrarse solamente en los detalles de la imagen, analizar la cantidad de líneas, determinar la complejidad con base en el tiempo que les tomo evaluar la imagen, etc. Estas indicaciones deben ser lo suficientemente claras para evitar cualquier mala interpretación de lo que se espera evaluar, recordando que se ha demostrado que la complejidad se puede ver afectada por la familiaridad, lo cual puede ser algo que no se busca cuando se trata de evaluar solamente la CI.

En [20] se realizó un estudio con 330 estudiantes griegos y una base de datos de 260 imágenes. El experimento consistió en ir mostrando imagen por imagen, donde existía un lapso entre las secuencias para que los participantes calificaran la imagen con base en la concordancia del nombre, año de adquisición y complejidad visual, donde en esta última se les indicó que está relacionada con el número de líneas y detalles en la imagen. En [4] realizaron el mismo experimento que en [20], pero en este caso, el grupo de participantes en el experimento eran personas de habla inglesa, lo cual al tratarse de diferentes culturas, pueden existir diferencias en las evaluaciones, por lo que es importante siempre aclarar de dónde era ese grupo de personas

con el que se realizó el experimento, por ejemplo, en [21] muestran tablas de diferentes grupos de personas y cómo la complejidad se vio afectada entre estos grupos, en este caso con personas de Estados Unidos, España y Francia. Sin embargo, la investigación está más orientada hacia investigadores franceses interesados en la memoria, lenguaje y otros procesos cognitivos.

En los artículos [21] y [22] definen la complejidad como la cantidad de líneas y detalles en un dibujo. Estos estudios han sido extremadamente útiles en muchos campos de psicología incluyendo memoria y lenguaje. La complejidad visual requirió que los participantes clasificaran la complejidad de cada dibujo en un rango entre 1 y 5, donde 1 es para poco complejo y 5 es para muy complejo.

En la Figura 1.8 se muestra el proceso general para determinar la complejidad subjetiva de la imagen según el criterio humano.

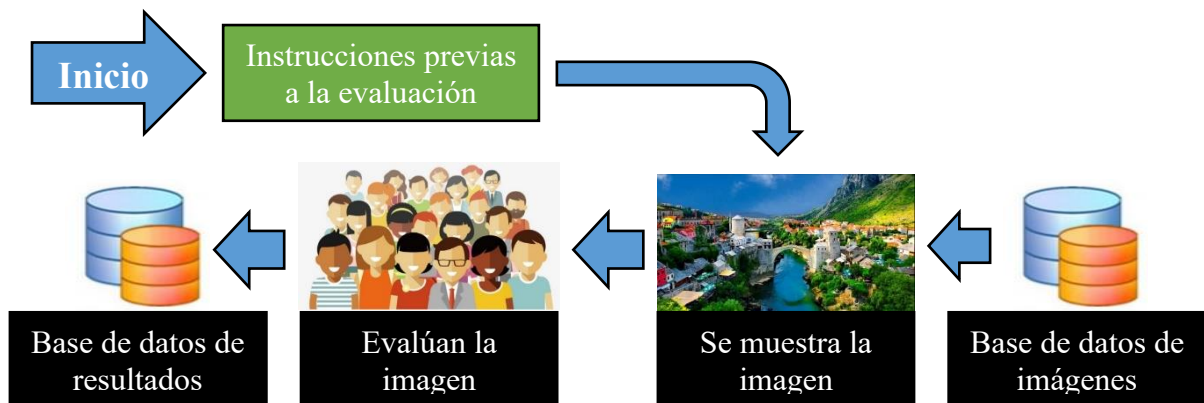


Figura 1.8. Proceso general para determinar la CI utilizando metodología del criterio humano.

1.4.2 Detección de bordes

En procesamiento de imágenes y visión por computadora, la detección de bordes es una operación de bajo nivel. El objetivo principal de la detección de bordes es localizar e identificar discontinuidades abruptas [23], las cuales delimitan regiones dentro de la imagen. Los bordes son detalles importantes ya que con solo verlos puede ser bastante descriptivo como para darse una idea clara del contexto de la imagen.

Las delimitaciones de las regiones generadas por detección de bordes pueden ser utilizadas en diferentes aplicaciones, como lo son, extracción de características, detección de objetos, compresión de imágenes, entre otros.

Existen muchos métodos para la detección de bordes en imágenes, sin embargo, los más populares y clásicos son Sobel y Canny [2]. Artículos como [24] y [25] realizan mejoras a los métodos respectivamente para incrementar la precisión de éstos y lograr mejores resultados en la detección de bordes de la imagen. La Figura 1.9 muestra un ejemplo de dos imágenes de bordes, uno utilizando el método de Sobel y otro Canny.

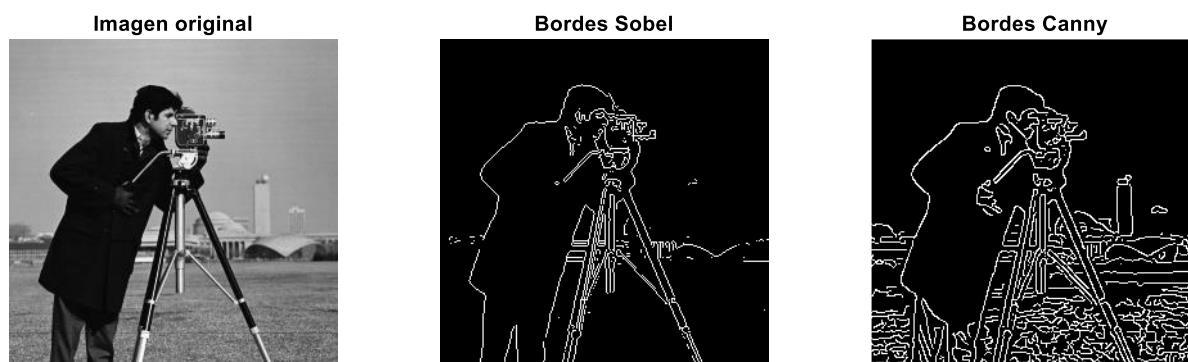


Figura 1.9. Detección de bordes con Matlab haciendo uso de técnicas de Sobel y Canny.

Dado que estos métodos de detección de bordes identifican aquellas regiones de la imagen donde se encuentran cambios en los niveles de intensidad de los píxeles, resulta fácil de imaginar que imágenes con un color uniforme a lo largo de todo su espacio no contendrá bordes, por lo tanto, tampoco tendrá muchos detalles, lo cual se traduce en que se trata de una imagen poco compleja; caso contrario, cuando en una imagen se encuentra gran cantidad de objetos que provocan muchos bordes, la CI será alta.

En la literatura se encuentran diversos artículos que determinan la CI por medio de los bordes, ya sea directamente o bien como una característica complementaria. Los métodos de [3] y [26] se basan en el análisis de los porcentajes de los bordes según su nivel, donde los niveles se distribuyen desde los bordes que son fácilmente detectados, hasta los bordes más difíciles de detectar por una persona. Dichos niveles se generan mediante diferentes umbrales a la hora de detectar los bordes, esto es, un umbral o cambio pequeño en la intensidad de los píxeles serán aquellos bordes que son difíciles de detectar por las personas, puesto que es más complicado poder ver un borde cuando el cambio de la intensidad del píxel es muy pequeño, caso contrario, el umbral será mayor para aquellos bordes fáciles de detectar por las personas (regiones que presentan un cambio muy grande en la intensidad de los píxeles). Mediante un

análisis difuso de porcentajes de bordes en los diferentes niveles, determinan la complejidad de una imagen brindando una ventaja ante el criterio humano al eliminar en cierto grado la subjetividad de las personas.

De manera similar, [15] hace uso de la lógica difusa para determinar la complejidad de la imagen ya que propone que ésta es más apropiada cuando la situación binaria es muy estricta. Sin embargo, no solo se basan en la característica de los bordes, también en características globales y locales. Donde las propiedades globales son generales de la imagen, y las locales se centran en el procesamiento por regiones como lo es detección de bordes y simetría.

La secuencia general que sigue este tipo de metodología es la siguiente:

- Adquisición de la imagen.
- Preprocesamiento. Consiste en la eliminación de ruido en la imagen para evitar desvíos de los resultados. En algunos casos también se aplica un filtro pasa bajas a la imagen para un mejor resultado con los detectores de bordes.
- Detección de los bordes.
- Análisis de la imagen de bordes.
- Aplicación de criterio. Una forma de estimar la densidad de los bordes puede ser determinando la relación entre el número de píxeles de los bordes extraídos y el número entero de píxeles de la imagen. La densidad de bordes puede ser determinada de la siguiente manera,

$$\rho_b = \frac{N_b}{N} \quad (1.1)$$

donde N es el número total de píxeles de la imagen, N_b es el número total de píxeles bordes.

- Cálculo del resultado de la medición de la complejidad.

1.4.3 Compresión de imágenes

En sistemas digitales, la compresión permite hacer una representación de la información de manera que contenga una cantidad menor de bits pero sin perder lo esencial de la información.

Por lo que es una codificación eficiente de la información que permite un manejo más eficaz tanto en almacenamiento como en transmisión.

La compresión se divide en dos grandes clases [27].

- Sin pérdidas. Aquella compresión que permite regresar a la información original. Esto con el costo de una relación de compresión más baja.
- Con pérdidas. Aquella compresión que permite pérdidas de información, por lo que al descomprimir la información puede sufrir distorsión. Sin embargo, puede lograr una relación de compresión mayor a la de sin pérdidas.

Además, las imágenes pueden contener información que se clasifica en tres tipos:

- Redundante: es la parte determinista, es decir, aquella parte que se puede predecir o bien que mantiene una relación con lo demás, como regiones homogéneas o lisas en las imágenes.
- Irrelevante: es aquella que tiene detalles que en ocasiones pueden considerarse innecesarios para cierto fin, por ejemplo, una imagen donde la calidad no es tan importante, si en dicha imagen cada pixel es representado por 16 bits, y si a cada pixel se le quita el bit menos significativo el humano difícilmente será capaz de notar una diferencia tan pequeña por lo que para este ejemplo el bit menos significativo será irrelevante.
- Útil: es aquella que es importante para mantener ya que contiene la esencia de la información, se obtiene mediante la resta de la información redundante y la irrelevante de la total [27].

Con base en los tipos de información que contiene una imagen se puede utilizar la compresión para obtener una relación de cuanta información en la imagen es redundante, esto si se utiliza un método de compresión sin pérdidas, o bien cuanta información en la imagen es redundante y/o irrelevante si se utiliza un método de compresión que permita pérdidas. La relación de compresión es determinada de acuerdo a la Ec. (1.2),

$$RC = \frac{T_B}{TC_B} \quad (1.2)$$

donde T_B es el tamaño o el espacio que ocupa en bytes la imagen sin ser comprimida y TC_B es el espacio que ocupa en bytes la imagen al ser comprimida.

Otra terminología de compresión de imagen es el número de bits por pixel (bpp). Para una imagen, esta relación es determinada de la siguiente manera,

$$bpp = \frac{T_b}{N} \quad (1.3)$$

En la literatura, se encuentran diferentes métodos para determinar la CI que se basan en estas relaciones. Puesto que los detalles de la misma son los que describen el grado de complejidad y marcan una diferencia en la relación de compresión. Entre mayor cantidad de detalles y poca redundancia, se tendrá una relación de compresión más baja que cuando se tengan muy pocos detalles y alta redundancia, lo cual está directamente correlacionado con la CI.

En [11] utilizan métodos de compresión donde se hace una relación de compresión entre la imagen original y la comprimida. Por medio de esta relación determinan la cantidad de información redundante e irrelevante para determinar el grado de complejidad, por lo que concluyen que entre más información redundante e irrelevante tenga la imagen, se tratará de una con complejidad baja. Caso contrario, si se tiene poca información redundante e irrelevante, es decir, se puede comprimir muy poco, indica que se trata de una imagen con complejidad alta.

Se pueden utilizar diferentes algoritmos de compresión para determinar la CI, por ejemplo [14] utiliza la compresión por PNG, que toma ventajas sobre las redundancias presentes en una imagen para producir mejores compresiones sin pérdidas. PNG funciona muy bien en comprimir cuadros sólidos, pero hace un trabajo relativamente pobre comprimiendo cuadros con una porción pequeña de ruido.

Por otro lado JPEG es un método de compresión de bajo nivel y de compresión local ya que usa ventanas, por lo que no toma ventaja con las regularidades globales de la imagen, mientras que compresión fractal es un esquema global, que toma ventajas de las similitudes presentes [2]. Aun así, JPEG es uno de los métodos más utilizados cuando se busca la relación de compresión como una característica para determinar la CI ya que se ha demostrado que tiene buena correlación con el criterio humano.

En la Figura 1.10 se ilustra el proceso general para determinar la complejidad de una imagen con base en la compresión de la misma.

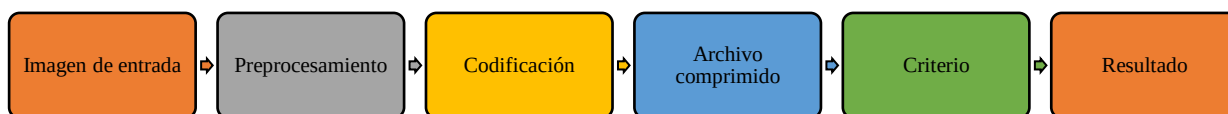


Figura 1.10. Proceso general para determinar la complejidad de la imagen con compresión.

La compresión de imágenes ha demostrado tener buena relación con la CI, aunque se ha encontrado que está correlacionada con su resolución, es decir, que existe una tendencia relacionada entre resolución y complejidad. Generalmente la complejidad incrementa conforme la resolución decrementa, excepto para imágenes con larga porción de texturas finas [11].

1.4.4 Distribución de colores

El color es una característica básica de la imagen, por lo que puede ser utilizado para describirla y representarla. La distribución de colores se puede medir local o globalmente, localmente mide la variación de los colores en regiones y globalmente mide la distribución de los colores de toda la imagen.

La característica de color en una imagen incluye la variedad y la distribución de color, en [28] definen la CI como la riqueza de sus componentes internos. El color es una fuente importante para obtener información, así pues, hacen uso de dos características para determinar la CI: variedad y distribución espacial de color.

Por otro lado, la entropía en informática mide el grado de incertidumbre de la fuente de información, donde aquellos eventos con menor probabilidad de suceder son los que contienen mayor información, mientras los más probables son los que contienen menor información. De tal modo que en [29] se hace uso de la distribución del color de la imagen de manera similar a la entropía para poder evaluar parcialmente la complejidad de la imagen, logrando un método eficiente en cuanto a costo computacional. Sin embargo, como se muestra en la Figura 1.11, la CI no puede ser definida solo por el histograma o su entropía, ya que dos imágenes con el mismo histograma pueden tener muy diferente complejidad debido a la distribución espacial de los colores [28].

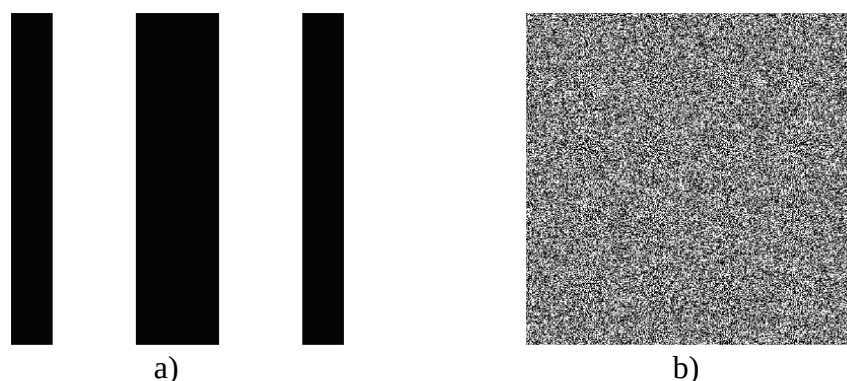


Figura 1.11. Dos imágenes con el mismo histograma y entropía pero diferente complejidad [11].

Una aproximación práctica de entropía puede ser un modelo que aproxime la función de distribución de probabilidad logarítmica por medio del uso de técnicas como análisis de componentes independientes la cual incorpora la distribución de los datos en la medida de la complejidad, diciendo que la distribución de los datos o bien la propiedad estadística prominente de las imágenes, puede ser mejor capturada con este método [13].

Por otro lado, [6] determina la CI a través de la entropía de un mapa de error, obtenido al pasar por un filtro los canales de colores de la imagen. Después la CI es utilizada para encontrar las áreas de la imagen con regiones prominentes visualmente, es decir, áreas que sobresalen al ser más llamativas para las personas, llegando a mejores resultados que otras técnicas de detección de regiones llamativas.

1.4.5 Combinado

Se le ha dado el nombre de método combinado a aquellos que hacen uso de al menos dos o más características diferentes para la determinación de la complejidad de las imágenes.

Este método consta en la extracción de diferentes características, como relación de compresión, cantidad de objetos, histograma, densidad de píxeles bordes, etc. Una vez que se reúnen dichas características, se le da cierta ponderación a cada una para la determinación de la medida de la complejidad. Generalmente, se hace uso de técnicas de reconocimiento de patrones para encontrar los pesos de cada característica que permitan tener una mejor aproximación hacia el criterio humano, la cual es la medida más común para correlacionar los

resultados de un método de medición de complejidad de imágenes. Por lo tanto, es necesario que exista un entrenamiento previo al uso de esta metodología para poder establecer los pesos que se le darán a cada característica y así lograr el resultado más preciso. Es importante tener una salida deseada que realmente esté bien clasificada, ya que el entrenamiento aprende a partir de esta salida deseada y de existir error en ésta, el algoritmo estará aprendiendo de igual forma sobre este, lo cual al final se verá reflejado en los resultados.

El proceso para realizar el entrenamiento es el siguiente:

- Tener una base de datos de imágenes y sus respectivas medidas de complejidad (salida deseada).
- A cada imagen se le extraen las características necesarias para determinar la CI.
- Se aplica alguna técnica de entrenamiento para la obtención de los pesos que mejor se ajusten para obtener la salida deseada con esas características.

Una vez que se ha entrenado al algoritmo, es posible pasar a la etapa de prueba o aplicación en la cual se hace uso del entrenamiento previo para determinar la complejidad de otras imágenes. En general el proceso de aplicación del método combinado sigue el diagrama de flujo de la Figura 1.12.

Con base en la literatura investigada, esta metodología es de las mejores correlacionadas con el criterio humano, debido a que se hace uso de diferentes características y un entrenamiento previo para lograr el resultado más parecido al esperado; sin embargo, también es de las metodologías más costosas computacionalmente, por lo que es un factor a tomar en cuenta cuando se busca algo eficiente en cuanto a procesamiento.

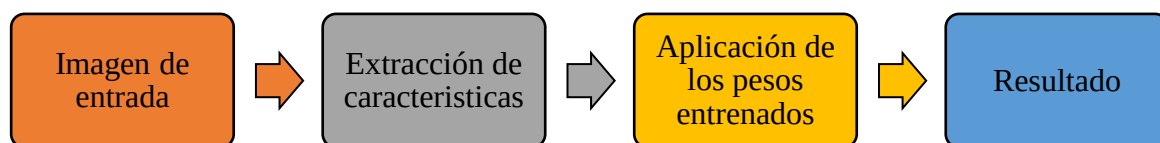


Figura 1.12. Proceso general de aplicación del método combinado.

La cantidad de características que se utilizan para la determinación de la CI no está definida o restringida, inclusive existen trabajos que hacen uso de más de 20 características para determinarla, tal es el caso de [30], donde identificaron tres factores principales que afectan la

percepción de la complejidad visual humana; distribución de composiciones, colores y contenidos. Después, mediante un estudio de conceptos psicológicos y teoría de arte, diseñaron 29 características globales, locales y regiones prominentes, con las cuales obtuvieron resultados aceptables que muestran qué modelos de regresión tienen una buena habilidad para medir calidad estética, belleza y unión de los colores.

Por otro lado [12], solo hace uso de tres características las cuales son bordes Canny, compresión JPEG y compresión GIF, las cuales se correlacionan bien con el criterio subjetivo de una persona.

Otro ejemplo de este método se encuentra en [31], donde proponen una medida de la complejidad basada en diferentes características de espacio, frecuencia y propiedades de los colores. Para encontrar la combinación lineal óptima de los pesos, se aplicó el método optimización de enjambre de partículas (PSO por sus siglas en inglés), el cual evalúa iterativamente las soluciones, recordando la locación en donde se encontró el mejor resultado. Ellos proponen realizar en un futuro modelos no lineales para combinar las medidas simples como por ejemplo máquinas de vectores de soporte.

También existen otros métodos que utilizan diferentes características a las descritas anteriormente como [32], donde demuestran que es posible evaluar objetivamente la complejidad en pinturas abstractas mediante la determinación de los valores de los números *Betti* asociados con la imagen. Los números *Betti* son cantidades que pueden ser usadas para distinguir espacios topológicos n dimensional. Estos números esencialmente miden la conectividad de tales espacios y se han utilizado para medir complejidad en otras áreas científicas.

1.4.6 Otros métodos

En la sección de otros se han clasificado aquellos métodos que no entran directamente a los descritos en los puntos anteriores, es decir, en esta clasificación no se hace un procesamiento sobre la imagen para extraer las características que se correlacionan con la complejidad de la imagen, ni tampoco se hace un cuestionario para que las personas clasifiquen las imágenes de acuerdo a su complejidad subjetiva, sino que son procesos más indirectos como el medir la

actividad cerebral de una persona mientras mira la imagen, o bien observar el movimiento ocular del observador de la imagen.

Un método utilizado para medir la complejidad de la imagen mediante la actividad cerebral es el análisis de señales electroencefalográficas (EEG) [33]. La electroencefalografía se basa en las corrientes de naturaleza iónica presentes en la corteza cerebral, que son el producto de la actividad cerebral y pueden ser registradas con unos electrodos colocados en el exterior del cráneo, lo cual lo vuelve un método poco invasivo y ampliamente utilizado.

Trabajos como [33] discuten la solución posible de aplicar ingeniería en reversa mediante un análisis en la actividad cerebral del humano a través de EEG para identificar el espacio característico utilizado por los humanos para la clasificación visual, ya que se ha reconocido que las grabaciones de actividad cerebral contienen información sobre categorías de objetos visuales. Siguiendo un proceso similar, en [34] asumen que observar las huellas de las acciones utilizadas en la creación de obras de arte como las marcas de los pinceles, se asocian con una simulación de esas acciones sensoriomotoras del observador. Por otro lado, también los resultados devueltos por las señales EEG pueden ser utilizados como otra medida para comparar los resultados obtenidos y darle mayor validez a nuevas técnicas o algoritmos para la determinación de la CI [19].

Otra manera de determinar la complejidad de la imagen de manera indirecta, es por medio del movimiento que realizan los ojos al estar mirando una imagen, de tal manera que se realiza un análisis en el patrón del movimiento ocular para determinar la complejidad de la imagen tal y como [35], donde su principal propósito es el de generar un novedoso método para la determinación de la complejidad.

1.5 Conclusión

La complejidad de la imagen es relevante tanto en ciencias de la computación como en psicología. Sin embargo, es una medida no estandarizada porque no existe una métrica única universal para determinarla. También, la CI es una medida subjetiva ya que la principal medida con la cual se correlacionan los resultados de esta característica de la imagen es por medio del criterio humano basado en el mecanismo de percepción visual humana, el cual es sumamente complejo ya que se basa no solo en la adquisición de la imagen a través de los ojos, sino que

también pasa por un proceso de análisis de lo observado en el que intervienen muchos factores como el criterio y las experiencias del observador.

Las principales ventajas de obtener la complejidad de una imagen por medio de un algoritmo son la facilidad de procesarlas de manera automática y que el algoritmo obtendrá una medida que no se basa en conocimientos a priori, o bien, regresa una medida no subjetiva. Sin embargo, estos algoritmos o métodos para determinar la complejidad de la imagen de manera automática, requieren de los resultados obtenidos con el criterio humano para realizar un entrenamiento y poder correlacionar sus resultados a éste, por lo que en cierto grado es subjetivo.

No obstante, la CI, aun y cuando no sea cien por ciento precisa o los métodos no regresen exactamente los mismos resultados, es posible utilizarla para analizar la imagen y poder establecer parámetros en diferentes procesamiento de imágenes como detección de bordes, reconocimiento de áreas de interés, o bien, en el área de psicología para explorar procesos del cerebro, como memoria de trabajo, modelos de atención, e incluso puede ser utilizada en publicidad.

CAPÍTULO II. ANÁLISIS DE BASES DE DATOS

En este capítulo se describen brevemente algunas de las bases de datos de complejidad de imágenes que se encontraron en la literatura para seleccionar aquellas que cumplan con lo necesario para ser utilizadas en esta investigación. Entre éstas, se analizaron solamente las bases de datos que proporcionan algún medio para su adquisición.

En la Tabla 2.1 se muestra un resumen de las bases de datos reportadas en la literatura donde:

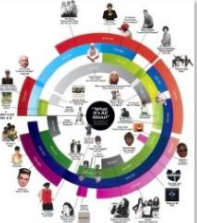
- Artículo: indica la referencia del artículo donde se reporta la base de datos.
- Cantidad de imágenes: indica el número de imágenes que contiene la base de datos.
- BD completa: indica si en la base de datos que se consiguió están incluidas todas las imágenes utilizadas por el artículo.
- Tabla de resultados: se indica si se tienen los resultados para todas las imágenes de la base de datos, donde ‘Completa’ es que se tienen todos los resultados de todas las imágenes, ‘Parcial’ indica que solo se tienen algunos resultados y ‘No’ indica que no se proporciona nada sobre resultados.
- Cantidad: se muestra cuántas bases de datos se utilizaron.

Tabla 2.1. Bases de datos.

Artículo	Cantidad de imágenes	BD completa	Tabla de resultados	Cantidad
[8]	118	Completa	Parcial	1
[18]	22	Parcial	Parcial	1
[21]	299	Completa	Completa	1
[28]	12	Parcial	Parcial	1
[31]	98	Completa	Completa	2
[32]	52	Parcial	Parcial	1
[36]	168	Completa	Parcial	1
[37]	12	Parcial	Completa	1
[38]	17,125	Completa	No	1
[39]	255,508	Completa	No	1
[40]	1,420	Completa	Completa	7
TOTAL	274,834	--	--	18

En total, se encontraron 18 bases de datos, que suman un total de 274,834 imágenes de diferentes categorías que van desde imágenes de dibujos en blanco y negro hasta imágenes de entornos reales a color. En la Tabla 2.2 se muestran algunos ejemplos de imágenes de diferentes categorías de las bases de datos adquiridas.

Tabla 2.2. Distintos tipos de imágenes en bases de datos.

Categoría	Ejemplos de imágenes		
Dibujos en blanco y negro			
Entornos reales a color			
Escalas de grises			
Pinturas abstractas			
Pinturas ambiguas			
Visualizaciones			

2.1 Selección de las bases de datos

Para seleccionar que bases de datos utilizar en el desarrollo de esta investigación se consideraron los siguientes puntos:

- La base de datos debe estar completa. En ocasiones se encuentran bases de datos parciales, donde faltan imágenes o evaluaciones, por lo que éstas no son de interés.

- La cantidad de imágenes. Se debe contar con una cantidad suficiente de imágenes (en el orden de cientos) para que los sistemas diseñados sean capaces de calcular la CI de manera más correlacionada con la estimación del humano.
- Confiabilidad de la base de datos. Se refiere a la metodología para la adquisición de los resultados de las evaluaciones. La confianza se mide con base en la forma en que se describe en detalle la evaluación de la CI. Es decir, aquellos artículos que muestren cada paso a seguir para la adquisición de los resultados de las evaluaciones, tienen mayor grado de confianza que aquellos que solo muestran los resultados.
- Base de datos a color. Se buscan bases de datos de imágenes a color, ya que algunas características importantes para estimar la complejidad se basan en color.
- Categoría de la base de datos. Se hizo una búsqueda de bases de datos con contenido común, como entornos reales, y no solo con contenido ambiguo y/o abstracto ya que estos son generalmente categorizados como más complejos de acuerdo con [18]. Esto es en parte por el principio de multiestabilidad de Gestalt, por lo que es más probable que exista mayores diferencias entre sujetos al estimar la complejidad de imágenes con este tipo de contenido (resultados menos confiables), caso que en esta investigación no se busca.

De acuerdo a los puntos descritos para la selección de la base de datos, se eligieron como las bases de datos a utilizar [31] y [40]. De aquí en adelante se referirá a estas bases de datos como los artículos las nombraron. Las dos bases de datos de [31] serán RS1 y RS2, mientras que la base de datos de [40] será SAVOIAS, la cual tiene 7 categorías y cada letra del nombre representa a una. Más adelante se describirá a detalle estas bases de datos seleccionadas.

2.2 Análisis de las bases de datos RS1 y RS2

Los autores de las bases de datos RS1 y RS2 [31] adquirieron las imágenes con alta calidad utilizando cámaras profesionales o semiprofesionales.

Para generar la base de datos RS1 se utilizaron 49 imágenes de entornos reales con diferentes grados de complejidad subjetiva aparente, en la Figura 2.1 se muestran estas imágenes. Para determinar el grado de complejidad subjetiva de esta base de datos, se realizó un experimento

que consistió en que un grupo de sujetos evaluara cada imagen y estimara la complejidad en un rango de 0 a 100. El grupo de sujetos para la evaluación de dicha base de datos RS1 consistió en un total de 26 personas, de las cuales 15 eran mujeres y 11 eran hombres con una edad promedio de 37 en un rango de 18 a 53 años.

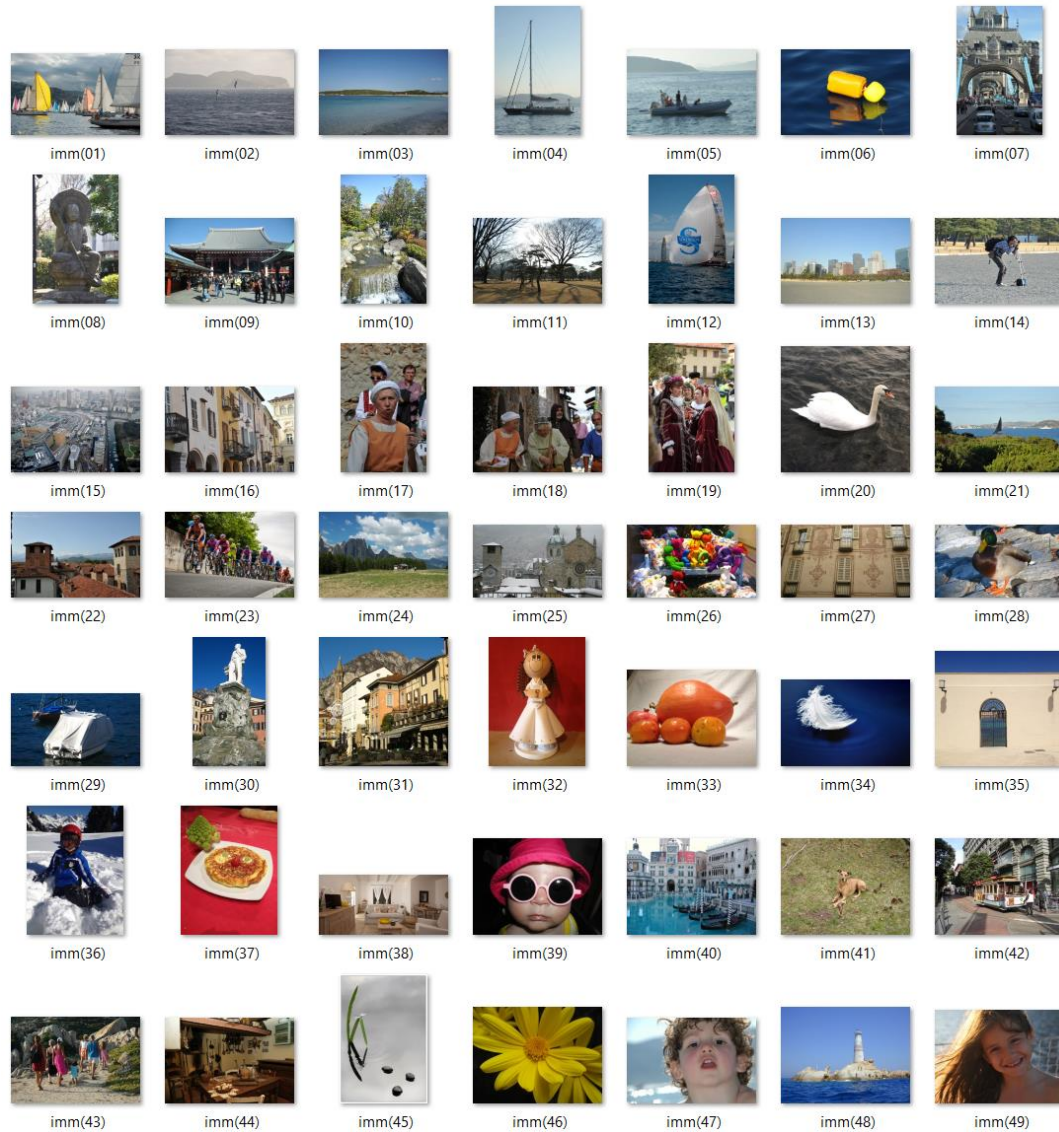


Figura 2.1. Imágenes de la base de datos RS1 de [31].

En la Figura 2.2 se muestran las imágenes de la base de datos RS2, la cual consta de un total de 49 imágenes de entornos reales con diferentes grados de complejidad subjetiva. Se determinó la complejidad de las imágenes de la misma forma que en RS1, pero esta vez el grupo de sujetos

que evaluaron la base de datos consistió en un total de 36, de los cuales eran 16 mujeres y 20 hombres con una edad promedio de 33 en un rango de 18 a 53 años.

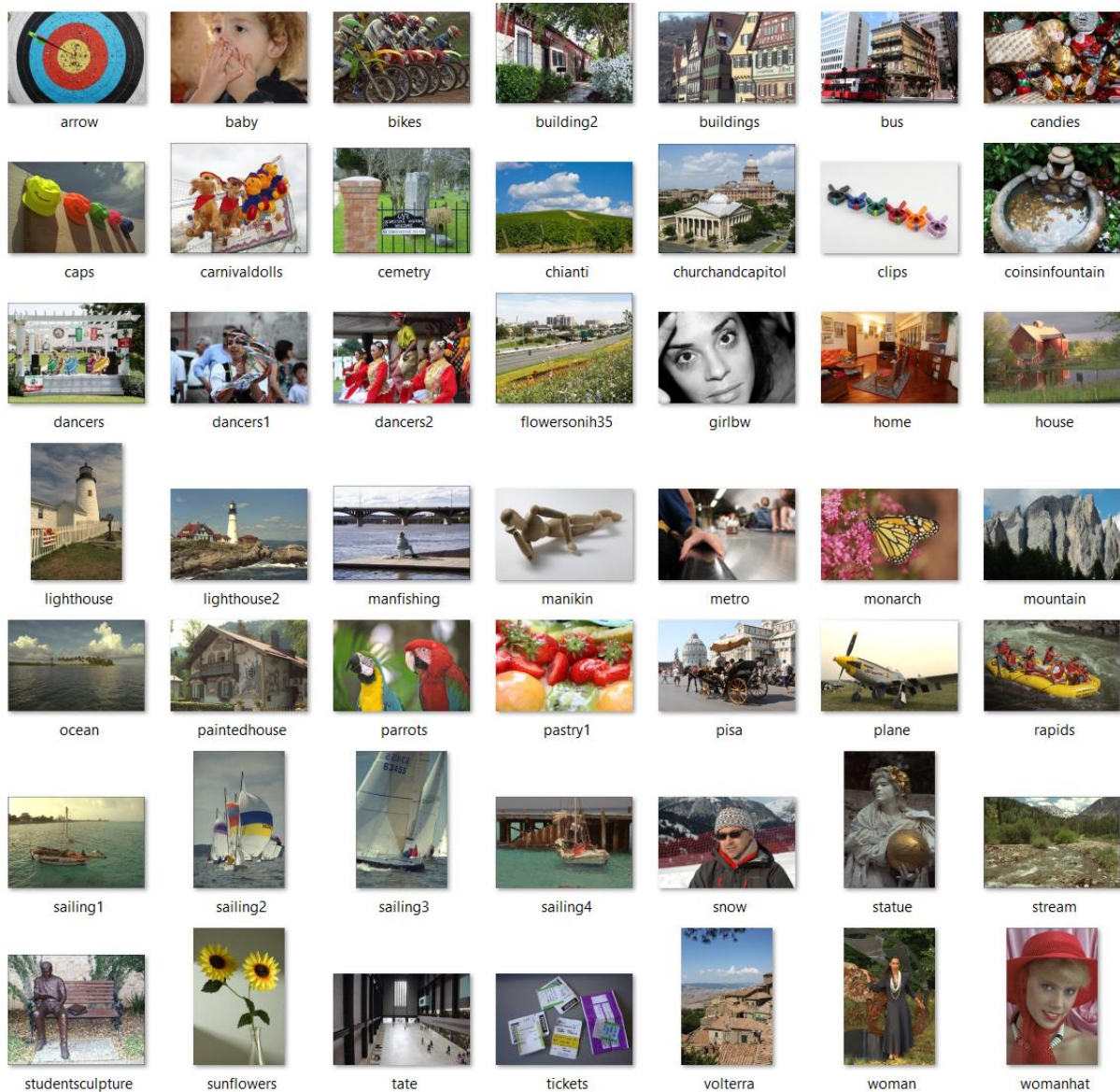


Figura 2.2. Imágenes de la base de datos RS2 de [31].

2.2.1 Determinación de la complejidad subjetiva

Para determinar la complejidad subjetiva de las bases de datos RS1 y RS2 se siguió una metodología similar a la presentada en la sección 1.4.1, donde se reúne un grupo de personas saludables, se les explica y da ejemplos de cómo evaluar la CI y por último se les va mostrando

las imágenes mediante una interfaz para que ellos vayan calificando a cada imagen con su complejidad aparente.

Hay que tener en cuenta que la muestra de imágenes se realizaba de manera aleatoria; es decir, no seguía un orden de aparición en el experimento. También es importante remarcar que aquí la puntuación de la complejidad se realizaba mediante una barra deslizante en un rango del 0 al 100 de complejidad, que de acuerdo con [40], no siempre es la mejor metodología en este tipo de casos, más adelante en el análisis de la base de datos SAVOIAS se explicará el por qué. Un ejemplo de cómo lucía la interfaz de usuario se muestra en la Figura 2.3.



Figura 2.3. Interfaz de usuario para establecer la complejidad subjetiva de las imágenes de las bases de datos RS1 y RS2.

2.2.2 Filtrado de las bases de datos RS1 y RS2

Para RS1 y RS2, las personas que evaluaron las imágenes ya habían pasado por un primer filtro, el cual consistió en descartar aquellos sujetos que no pasaran la prueba de 6 cartas Ishihara, donde se evalúa la capacidad de percibir los colores. Sin embargo, esta prueba no es

suficiente para asegurar que los resultados sean consistentes con la definición dada de complejidad.

Dado que la CI es una característica subjetiva, es posible que entre las evaluaciones de los sujetos existan diferencias, por lo que fue necesario hacer un análisis de los resultados. Esto con el fin de realizar un filtrado de la base de datos, y descartar aquellos que están lejos de la media de la población.

El primer paso para realizar el filtrado de la base de datos consiste en descartar sujetos que no estén bien correlacionados con el promedio de la población. El umbral que se utilizó fue de 0.43 [41], ya que era el umbral donde se descartaban aquellos sujetos que realmente aparentaban tener un comportamiento atípico a la población. Consecuentemente, aquellos sujetos cuya correlación con la media de la población no supere este umbral serán eliminados de la base de datos y no serán considerados en el nuevo valor promedio del puntaje subjetivo. En la Figura 2.4 se muestra una comparación entre dos sujetos, donde la Figura 2.4a es del sujeto 2 que muestra un comportamiento bien correlacionado con la media, mientras que en la Figura 2.4b muestra al sujeto 26 que no se correlaciona bien con la media (por debajo del umbral).

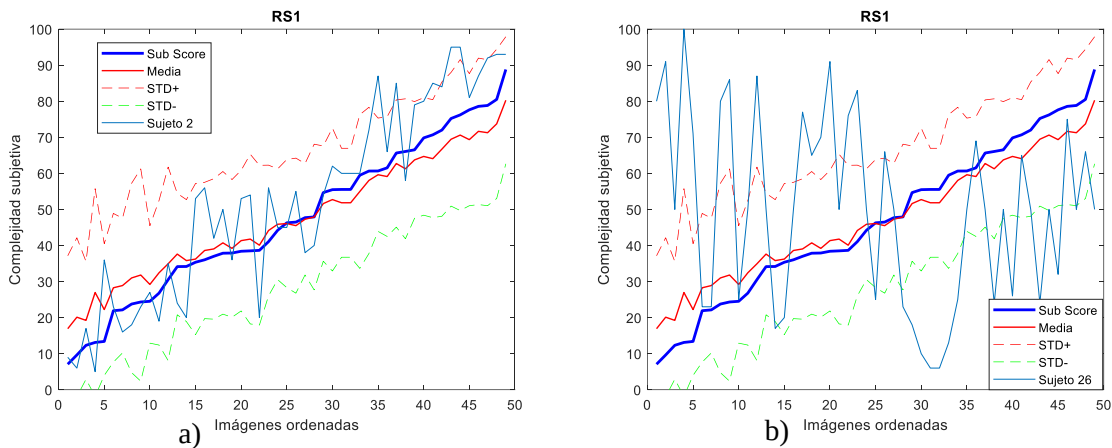


Figura 2.4. Evaluaciones de RS1 por a) sujeto 2, b) sujeto 26.

De la Figura 2.4, Sub Score es el valor medio de puntaje subjetivo con filtrado, Media es el valor medio sin filtrado, promedio que se proporciona por la base de datos, STD+ es una desviación estándar hacia arriba de la media, y STD- es una desviación estándar hacia debajo de la media.

Si bien, es verdad que la complejidad es una medida subjetiva y por lo tanto el sujeto 26 puede estar en lo correcto según su criterio, no cumple con la definición que se ha dado a la CI en esta investigación, lo cual puede ser demostrado mediante un análisis sobre sus resultados. Se puede corroborar que este sujeto sí debía ser descartado, ya que existen muchos casos fuera del promedio, por ejemplo, el sujeto otorgó un puntaje arriba de 80 a las primeras dos imágenes que aparecen de manera ordenada como se muestra en la gráfica de la Figura 2.4, mientras que a las imágenes 31 y 32 le dio un puntaje menor a 10, lo cual va muy en contra del promedio de la población. En la Figura 2.5 se muestran dichas imágenes con la complejidad subjetiva con la que el sujeto 26 las evaluó.



Figura 2.5. a) Imágenes 1 y 2 con complejidad superior a 80, b) imágenes 31 y 32 con complejidad inferior a 10, según sujeto 26.

De la base de datos RS1, se descartó un total de 3 sujetos por estar debajo del umbral y mostrar comportamientos lejanos a lo esperado.

Siguiendo el mismo procedimiento de filtrado para RS2, se encontraron problemas similares en algunos sujetos, como el caso de la Figura 2.6 donde a) muestra al sujeto 8 que se correlaciona bien con la media, mientras que b) muestra al sujeto 29 que no se correlaciona bien con la media.

Analizando los resultados del sujeto 29 para corroborar que debía ser descartado, se encontró que estaba dando valores de complejidad completamente contrarios para imágenes donde la mayoría de los sujetos decían que tenían complejidades similares. En la Figura 2.7 se muestra un ejemplo de dos imágenes donde el sujeto 29 le dio dos complejidades completamente distintas cuando el promedio de la población las calificó con una complejidad muy similar.

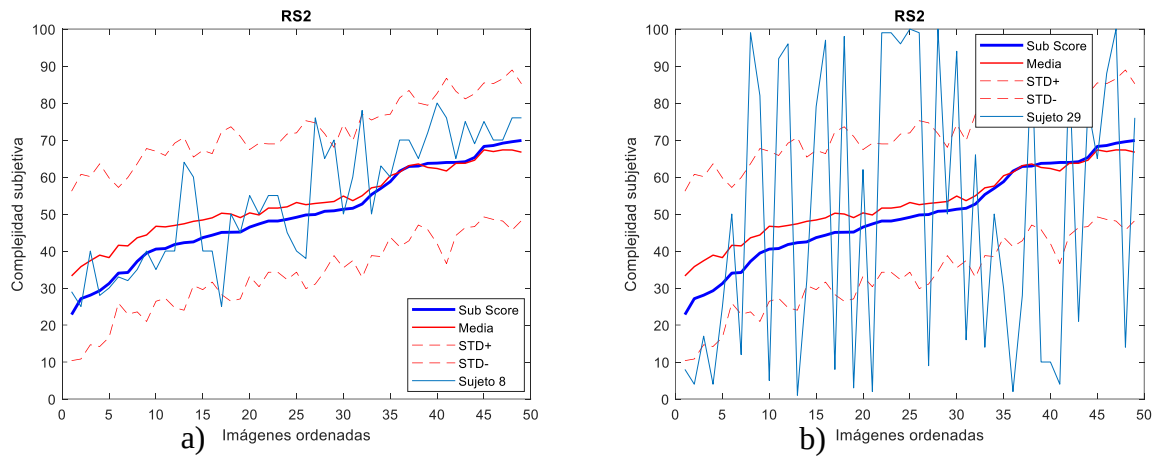


Figura 2.6. a) Gráfica evaluación del sujeto 8, b) gráfica evaluación del sujeto 29 para RS2.



a)



b)

Figura 2.7. a) Imagen con complejidad menor a 5, b) imagen con complejidad mayor a 95 según sujeto 29 en RS2.

Al final del filtrado para RS2 se descartaron un total de 11 sujetos que presentaban comportamientos atípicos a la media de la población.

Una vez que se retiraron aquellos sujetos que no se correlacionaban bien con la media de la población, se llevó a cabo un tercer filtrado de los datos que consiste en eliminar aquellas muestras de valores subjetivos que se encuentren por fuera de dos desviaciones estándar. Esto se hizo ya que se considera que aun y cuando el sujeto realizó bien la prueba, pudo existir algún error al seleccionar la complejidad para la imagen, por ejemplo, dejar el puntaje por defecto, o bien no haber ajustado bien la barra deslizante. Para esto se calculó el “z-score” el cual transforma los datos a una media cero y una desviación estándar unitaria por medio de la Ec. (2.1),

$$zs_{ij} = \frac{s_{ij} - \bar{s}_i}{\sigma_i} \quad (2.1)$$

donde zs_{ij} es el “z-score” para la imagen i del sujeto j , s_{ij} es el valor subjetivo de cada imagen por cada sujeto, $\bar{s}_i$ y σ_i son la media y la desviación estándar del valor subjetivo de todos los sujetos para cada imagen i . Por lo que aquellas muestras que se encontraran fuera de dos desviaciones estándar fueron descartadas.

Hasta este punto se tiene el conjunto de sujetos y datos aceptados de las bases de datos RS1 y RS2, solo falta obtener el valor promedio de complejidad subjetiva de cada imagen, ya que lo que se busca es modelar la complejidad a la media del criterio humano. Este valor se calcula mediante la Ec. (2.2),

$$\bar{s}_i = \frac{1}{|L|} \sum_{j \in L} s_{ij} \quad (2.2)$$

donde $\bar{s}_i$ es el valor de complejidad subjetiva final de cada imagen i , L es el conjunto de “z-scores” aceptados de cada imagen, $|L|$ denota cardinalidad, y s_{ij} es la complejidad dada a la imagen i por el sujeto j . Los valores regresados por Ec. (2.2) son lo que serán utilizados para el diseño y validación de los sistemas.

2.3 Análisis de la base de datos SAVOIAS

La segunda base de datos que se seleccionó fue SAVOIAS [40], la cual tiene 1420 imágenes en 7 categorías según su contexto, estas son: escenas (200), anuncios (200), visualización e infografía (200), objetos (200), diseño de interiores (100), arte (420) y suprematismo (100). Además, la base de datos proporciona las complejidades subjetivas de cada imagen según el experimento que realizaron para la obtención de ésta, del cual se hablará más adelante. En la Figura 2.8 se muestra un ejemplo de algunas imágenes de las diferentes categorías ordenadas de menor a mayor complejidad.

La CI de [40] de la base de datos está definida como el nivel de intrincado y detalles de una imagen o el nivel de dificultad para describir la imagen la cual concuerda con la definición dada en esta investigación.

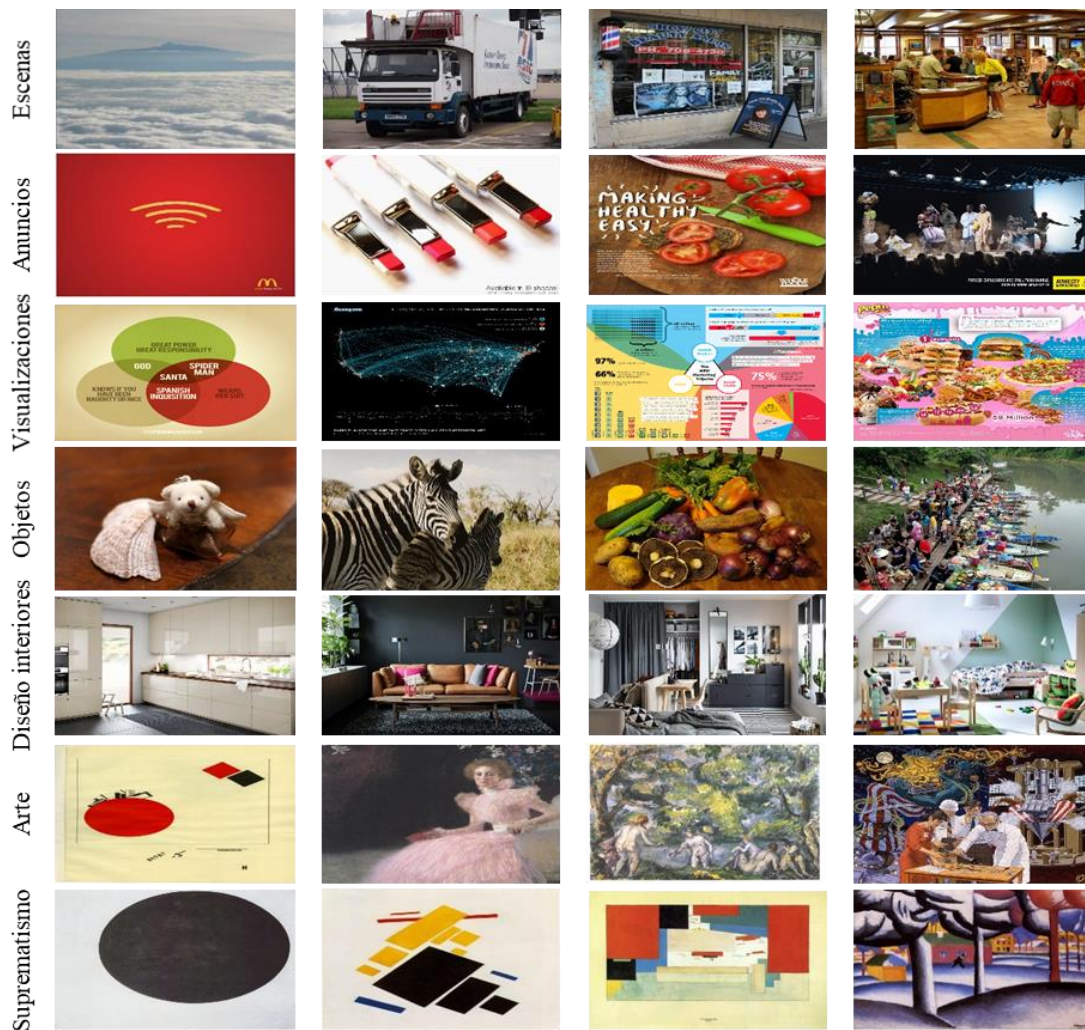


Figura 2.8. Algunos ejemplos de imágenes de la base de datos SAVOIAS, ordenadas de menor a mayor complejidad.

2.3.1 Determinación de la complejidad

En [40] proponen una metodología de decisiones forzadas que consiste en hacer comparaciones entre pares de imágenes y los participantes deben decidir entre cuál de las dos imágenes es más compleja. De acuerdo con los autores de SAVOIAS, esta metodología de decisión forzosa minimiza el sesgo potencial de los participantes, así como también el sesgo de las escalas limitadas. Los resultados se obtuvieron de 1,687 participantes en más de 37,000 pares de comparaciones de imágenes. Las puntuaciones por pares se convirtieron luego en

puntuaciones absolutas utilizando el algoritmo de Bradley-Terry y la terminación de la matriz de relaciones R .

La Figura 2.9 muestra un ejemplo de una interfaz de usuario similar a la que se utilizó para determinar la complejidad de las imágenes, la cual consiste mostrando dos imágenes y darles las opciones para que decidan cuál es la más compleja.

Por favor observe cuidadosamente las siguientes dos imágenes y decida cual es visualmente más compleja.



Por favor seleccione uno de los siguientes botones: (requerido)

- ☐ A es más compleja visualmente que B
- ☐ B es más compleja visualmente que A

Figura 2.9. Ejemplo de cómo lucía la interfaz de usuario que se utilizó para realizar la decisión forzosa de la imagen más compleja.

Preguntas de prueba, orientadas hacia el control de calidad se distribuyeron aleatoriamente a los participantes a lo largo de todo el trabajo para asegurarse de que estén atentos a la tarea. Si bien se pagaron todas las tareas de comparación (\$0.1 dólares por diez comparaciones), solo las respuestas que mantienen una calificación de aprobación del 90% o más en las preguntas de la prueba se mantuvieron.

Dado que resultaría muy costoso realizar todas las comparaciones para todas las categorías, se asignaron diferentes cantidades de comparaciones, a excepción para la categoría de suprematismo donde sí se realizaron todas las comparaciones (4,950 dado que son 100 imágenes), con esta misma se realizó un análisis para evaluar si era fiable utilizar menos comparaciones.

Por cada comparación se iba completando la matriz $R = \{r_{i,j}\}$, donde $r_{i,j}$ es la relación del número de veces que los contribuyentes seleccionaron la imagen i como más visualmente compleja comparada con la imagen j , sobre el total de número de veces comparadas. Por lo tanto $r_{i,j} + r_{j,i} = 1$. La Figura 2.10 muestra un ejemplo de la matriz R .

	1	2	3	...	n
1	-	a	?	...	
2	1-a	-	b	...	
3	1-?	1-b	-	...	
⋮	⋮	⋮	⋮	-	
n					-

Figura 2.10. Ejemplo de la matriz R .

Dado que se trata de más de 1400 imágenes sería muy complicado realizar todas las comparaciones necesarias para llenar la matriz R , por lo que se propuso solo realizar una fracción de las comparaciones e implementar algún método que permita completar la matriz, como lo son el método de Bradley-Terrey o el de terminación de matriz.

2.3.2 Métodos para completar la matriz

El método de Bradley-Terry describe la probabilidad de elegir la imagen I_i sobre la imagen I_j como una función sigmoidea de la diferencia del puntaje entre dos imágenes,

$$P(I_i > I_j) = F(\Delta_{i,j}) = \frac{e^{\Delta_{i,j}}}{1 + e^{\Delta_{i,j}}} \quad (2.3)$$

donde $\Delta_{i,j} = s_i - s_j$. El puntaje s puede ser estimado por máximo a posteriori (MAP).

$$\log \Pr(R | s) = \sum_{i,j} r_{i,j} F(\Delta_{i,j}) \quad (2.4)$$

donde lo anterior es una distribución uniforme. Este problema de optimización puede ser resuelto usando gradiente descendiente.

El método de terminación de matriz asume que si $r_{i,j}$ es más grande que 0.5 (imagen i es más compleja que la imagen j), $r_{j,l}$ es más grande que 0.5 y el par de comparación entre i y l no existe, se usa la imagen j como punto en común, por lo que se puede inferir que $r_{i,l}$ es también más grande que 0.5. Ahora se puede crear la matriz $\hat{R}$ rellenando los elementos de la matriz R :

$$\hat{r}_{i,l} = \begin{cases} r_{i,l} & \text{if } r_{i,l} \in R \\ \frac{1}{m} \sum_{j=1}^m \frac{r_{i,j} + r_{j,l}}{2} & \text{else if } r_{i,j} \in R, r_{j,l} \in R \\ & \text{and } r_{i,j} > 0.5, r_{j,l} > 0.5 \end{cases} \quad (2.5)$$

donde m es el número de pares existentes $r_{i,j}$ y $r_{j,l}$ en la matriz R . Para completar la matriz note los siguientes puntos:

- Como se indica en la ecuación, solo se considera $r_{i,j}$ y $r_{j,l}$ existe en R si son más grandes que 0.5. Por otro lado, si $r_{i,j} > 0.5$ y $r_{j,l} < 0.5$, no se realizará ningún juicio sobre el par faltante $r_{i,l}$.
- Para los pares que se tienen los resultados en una dirección, se puede llenar la matriz en la otra dirección mediante: $1 - r_{i,l}$.
- Para el caso raro que un par no esté conectado en ninguna dirección, se coloca el valor arbitrario de 0.5.

2.3.3 Puntajes finales

Cuando ya se tiene la matriz $\hat{R}$ (matriz completada), el puntaje absoluto para cada imagen es la media de los puntajes de los pares para la imagen,

$$\bar{s}_i = \frac{1}{n} \sum_{j=1, \hat{r} \in \hat{R}}^n \hat{r}_{i,j} \quad (2.6)$$

Para comparar los resultados de ambos métodos se obtuvo la correlación entre ambos resultados, donde se obtuvieron correlaciones mayores a 0.98 en todas las categorías de las imágenes. Sin embargo, los resultados reportados (base de datos) son los obtenidos con el método de Bradley-Terry.

Al tratarse de una metodología en la adquisición de los puntajes subjetivos distinta a la de RS1 y RS2 ya no se realizó un filtrado como tal (eliminar sujetos o puntajes atípicos), por lo

que aquí solo se optó por un análisis visual de los resultados. Esto es, tomar grupos de muestras de imágenes y ver si se les ha dado un puntaje subjetivo acorde a la definición que se ha dado en esta investigación a la CI, llegando a la conclusión de que los resultados de la base de datos SAVOIAS si cumple.

2.4 Conclusión

Las bases de datos que se seleccionaron para el desarrollo de esta investigación fueron RS1, RS2 de [31] y SAVOIAS de [40]. Estas fueron seleccionadas ya que cumplen los criterios establecidos en esta investigación y además son bases de datos de acceso libre de imágenes a color de entornos reales, por lo que la familiaridad no debió impactar tanto en la estimación de la complejidad subjetiva.

Si bien, en la base de datos de SAVOIAS no se realizó un filtrado de los datos como en RS1 y RS2, los autores de SAVOIAS afirman que una decisión forzosa es menos susceptible a errores como el marcar equivocadamente la complejidad de una imagen, o bien el sesgo potencial que puede ocurrir si se quiere dar una complejidad mayor o menor a lo que la escala permite, por ejemplo, si el sujeto da la máxima puntuación a la primera imagen, pero después se da cuenta de que hay imágenes aún más complejas, ya no podrá darles una mayor puntuación. Además, una decisión forzosa es más simple al solo tener que comparar dos imágenes.

Una vez que se analizaron las bases de datos se optó por utilizar ambas, pero de manera independiente, ya que se tratan de metodologías distintas e imágenes de diferentes categorías.

CAPÍTULO III. CARACTERÍSTICAS PARA LA COMPLEJIDAD OBJETIVA

La complejidad de la imagen es una medida subjetiva cuando una persona la evalúa, ya que el proceso de percepción del humano trata de darle sentido a lo que se está viendo con base en los conocimientos y experiencias personales [8]. Por ello, en ocasiones es preferible determinar la complejidad objetiva, la cual se basa en características de la imagen que se extraen y cuantifican por una computadora y se correlacionen con la complejidad subjetiva que percibe una persona. De esta manera, con estas características se puede determinar la complejidad de la imagen de una forma indirecta. En este capítulo, se describirá el procedimiento de selección de características que serán utilizadas para la determinación de la complejidad objetiva. También, se detallarán los métodos empleados para la extracción de las características, así como el análisis que se realizó a cada una de éstas.

3.1 Selección de características

En la literatura existen muchos métodos para determinar la complejidad de la imagen, los cuales hacen uso de diferentes características como bordes, número de regiones, colores, entre otras. Para determinar cuáles características utilizar se proponen una serie de criterios a considerar.

3.1.1 Criterios para la selección de características

Los criterios son detalles que se consideraron importantes para realizar una buena selección de las características. A continuación, se enlistan estos criterios con una breve descripción:

- Correlación de la característica contra el puntaje subjetivo dado por una persona. En este criterio se toma en cuenta el valor que reporta el método de la literatura de correlación entre la o las características y el puntaje dado por una persona. Sin embargo, en muchos de los artículos no se realiza esta medición o no es reportada, por lo que para esos casos se les dio un valor arbitrario de 0.5, esto con motivo de no favorecerlos más que a los que sí lo proporcionan, pero tampoco quitarle completamente toda la puntuación a esa característica. Aquellas características que han sido marcadas con ‘-’ en correlación es porque no son utilizadas en ese método.

La correlación contra el criterio subjetivo es importante ya que se están buscando las características que regresen resultados lo más similar posible al criterio subjetivo del humano al evaluar la CI.

- Citas. Este criterio marca que tanto ha sido citado el artículo, se espera que un artículo que ha sido citado por muchos autores tenga una investigación y resultados muy bien fundamentados, por lo que es un criterio a tomar en cuenta cuando se buscan las características en la literatura.
- Vistas. De igual manera que el criterio de citas, se espera que un artículo que ha recibido muchas vistas tenga características adecuadas para determinar la complejidad. Sin embargo, más adelante se muestra que se le dio un peso menor, ya que no necesariamente esta correlacionado con la calidad de las características.
- Referencias. Este criterio toma en cuenta a cuantos otros artículos investigo el autor, se espera que aquel que ha realizado una investigación más amplia tenga buenos fundamentos para seleccionar sus características.
- Número de imágenes en el experimento. Aquí se toma en cuenta con cuantas imágenes se realiza los entrenamientos y pruebas de los algoritmos.
- Editorial. La editorial fue marcada como un criterio solamente para ver en cuales revistas se encontraron mayor cantidad de artículos referentes al tema. Sin embargo, este criterio no tomó ningún peso para impactar en la selección de las características.
- Proporciona la base de dato. Este es un punto importante que se consideró, ya que se espera que si el autor proporciona la base de datos será más fácil la reproducción del experimento lo cual da más certeza de que los resultados reportados no fueron manipulados para favorecerlos.
- Nivel de detalle en la documentación. Este criterio fue al que mayor peso se le dio. En la Tabla 3.1 se muestra la puntuación que se les da a los artículos con base en ciertos puntos importantes considerados en esta investigación. Tener en cuenta que este criterio es subjetivo por lo que puede variar dependiendo de la persona que realice la evaluación.

Tabla 3.1. Puntuación según el nivel de detalle.

Puntuación	Detalles en cuenta
5	El artículo está suficientemente detallado y proporciona lo necesario para poder reproducir los resultados.
4	El artículo esta detallado, pero no lo suficiente para reproducir los resultados.
3	El artículo es claro en la metodología para determinar la complejidad, pero no en sus resultados.
2	El artículo es claro en resultados, pero no en la metodología para determinar complejidad.
1	El artículo no es del todo comprensible.

Para determinar qué características se deben utilizar en el algoritmo que determina la CI, primero se aplica una ponderación a los criterios para darle mayor peso a aquellos que se consideran más relevantes para la selección correcta y luego se utiliza la Ec. (3.1) la cual da a conocer el puntaje final otorgado a cada característica. Aquellas con los puntajes más altos serán las seleccionadas.

$$PC = CC \sum_{i \in CR} wc_i CR_i \quad (3.1)$$

donde, PC es el puntaje de la característica, CC es la correlación de la característica con el criterio humano, CR es el conjunto de criterios para elegir las características, wc_i es el peso del criterio y CR_i es el valor del criterio i . El valor de CR_i fue normalizado a uno, dividiendo cada valor con el máximo de su categoría. En la Tabla 3.2 se muestran los pesos que se le dio a cada criterio.

Tabla 3.2. Pesos de los criterios para elegir características.

Criterio	Peso wc
Citas	0.2
Vistas	0.05
Referencias	0.1
No. de imágenes en el experimento	0.1
Proporciona la base de datos	0.2
Nivel de detalle	0.35

Considere la Tabla 3.3 donde se enlistan las abreviaturas de las editoriales de los artículos analizados.

Tabla 3.3. Lista de abreviaturas utilizadas para describir la editorial.

Abreviatura de editorial	Editorial
IE	IEEE
EL	ELSEVIER
SP	SPRINGER
AR	ARVOJOURNALS
PL	PLoS ONE
AP	APA PsycNET

En la Tabla 3.4 se muestran los puntajes otorgados a cada método de la literatura investigada. Donde la primera columna indica la referencia de donde fue obtenido el método. Al lado derecho le siguen las características, en esta sección se da a conocer el puntaje de la correlación contra el valor subjetivo. Es importante recordar que en los artículos no necesariamente viene la característica como lo muestra la tabla, sino que puede estar nombrada de forma distinta, sin embargo, se buscó realizar un agrupamiento para disminuir la cantidad de diferentes tipos de características. Por último, están las columnas de los criterios considerados para la selección, los puntajes otorgados son propios de esta investigación.

De igual manera que en la Tabla 3.4, la Tabla 3.5 contiene la misma información solo que en ésta los valores ya han sido normalizados y procesados por la Ec. (3.1). Como se puede observar, la mayoría de las características alcanzan valores menores a 0.3, esto se debe a que el valor máximo que se puede alcanzar es el de 1.0. Sin embargo, para que esto se logre tendría que ser el artículo con mayor cantidad de citas, vistas, referencias y haber reportado un 100% de correlación con el criterio humano. Tomando esto en cuenta, valores superiores a 0.2 son considerados como buenos.

Luego de realizar todo el análisis descrito anteriormente, se seleccionaron las cuatro características con las mejores puntuaciones y que no fueran muy similares entre ellas, quedando las siguientes:

1. Niveles de intensidad.
2. Compresión JPEG.
3. Bordes.
4. Distribución.

Tabla 3.4. Características y criterios para su elección.

	Características															Criterios para elegir características						
Referencia.	Bordes Sobel	Bordes Canny	Bordes Otros	Compresión JPEG	Compresión GIF	Compresión PNG	Compresión Otros	Niveles de intensidad	Distribución	Energía	EEG	Contraste	Correlación	Betti Numbers	Simetría	Citas	Vistas	Referencias	No. Imágenes en experimento	Editorial	Proporciona BD	Nivel de detalle
[2]	0.5	0.5	-	0.77	-	-	0.5	0.5	-	-	-	-	-	-	-	8	59	93	800	EL	0	4
[3]	-	-	0.5	-	-	-	-	-	-	-	-	-	-	-	-	12	629	11	150	IE	0.25	4
[5]	0.5	-	-	-	-	-	-	-	0.5	0.5	-	0.5	0.5	-	-	1	57	10	200	IE	0.5	4
[6]	-	-	-	-	-	-	-	0.5	-	-	-	-	-	-	-	1	359	30	0	IE	0.5	3
[11]	-	-	0.5	0.5	-	-	0.5	-	-	0.5	-	-	-	-	-	18	422	20	30	IE	0	4
[12]	-	0.61	0.73	0.72	0.75	-	-	-	-	-	-	-	-	-	-	1	673	62	550	SP	0.5	4
[13]	-	-	-	-	-	-	-	0.5	0.5	-	-	-	-	-	-	8	1800	8	0	SP	0	2
[15]	-	-	0.5	-	-	-	-	-	-	-	-	-	-	-	0.5	29	27	32	0	EL	0	4
[17]	-	-	-	0.5	0.5	-	-	-	-	-	-	-	-	-	-	12	2400	38	0	SP	0	2
[18]	-	-	-	0.5	0.5	-	0.5	-	-	-	-	-	-	-	-	0	1	42	50	AP	0.5	3
[19]	-	-	-	-	-	-	-	-	-	0.5	0.5	-	-	-	-	1	119	28	0	IE	0	2
[26]	-	-	0.5	-	-	-	-	-	-	-	-	-	-	-	-	1	53	16	0	IE	0	4
[28]	-	-	-	-	-	-	-	0.5	0.5	-	-	-	-	-	-	0	48	9	0	IE	0.5	3
[31]	-	0.67	-	0.65	-	-	-	0.56	0.57	0.72	-	0.74	0.52	-	-	5	2000	59	98	PL	1	5
[32]	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5	1	15	34	58	EL	1	4
[36]	-	-	-	-	-	-	-	0.5	0.5	-	-	-	-	-	-	0	51	12	28	IE	1	5
[42]	-	-	-	-	-	0.5	-	-	-	-	-	-	-	-	-	0	86	30	0	IE	0	2
[43]	-	-	-	-	-	-	-	-	-	0.5	0.5	-	-	-	-	1	37	19	250	IE	0	2
[44]	-	-	-	-	-	-	-	-	-	-	0.5	-	-	-	-	0	139	16	20	IE	0	2
[45]	-	-	-	-	0.64	-	-	-	-	-	-	-	-	-	-	16	2600	58	1500	AR	0	4
[46]	-	-	0.5	-	-	-	-	0.5	-	-	-	0.5	-	-	-	14	2100	34	2000	SP	0	4
[47]	-	-	-	-	-	-	-	-	-	0.5	0.5	-	-	-	-	0	446	33	250	SP	0	2
[48]	-	-	0.5	-	0.5	-	-	0.5	-	-	-	-	-	-	-	2	19	36	150	EL	0	3
[49]	-	0.5	-	-	-	-	-	0.5	0.5	-	-	0.5	-	-	0.5	0	5	48	500	EL	0	4.5
[50]	-	0.5	0.5	0.5	-	-	-	-	-	-	-	-	-	-	-	3	1135	126	912	PL	0	4
[51]	-	-	-	-	-	-	-	0.5	-	-	-	-	-	-	-	0	47	23	0	IE	0	3
[52]	-	-	-	-	-	-	-	0.5	-	-	-	-	-	-	-	1	95	13	1000	IE	0	2
[53]	-	0.5	-	0.5	-	-	-	0.5	0.5	-	-	-	0.5	-	-	1	751	26	29	SP	1	5
[54]	-	0.5	0.5	0.5	-	0.5	0.5	0.5	-	-	-	0.5	-	-	-	25	6100	162	96	PL	0	4
[55]	-	-	0.5	-	-	-	-	0.5	0.5	-	-	-	-	-	-	1	2300	21	0	SP	0	3
Σ	1	3.78	5.23	5.14	2.89	1	2	7.06	4.07	3.22	2	2.74	1.52	0.5	1	162	24.5k	1.1k	8.7k		6.75	101.5

La Figura 3.1 muestra la gráfica de las puntuaciones alcanzadas por cada característica una vez que se les aplicó la Ec. (3.1) para determinar su puntuación.

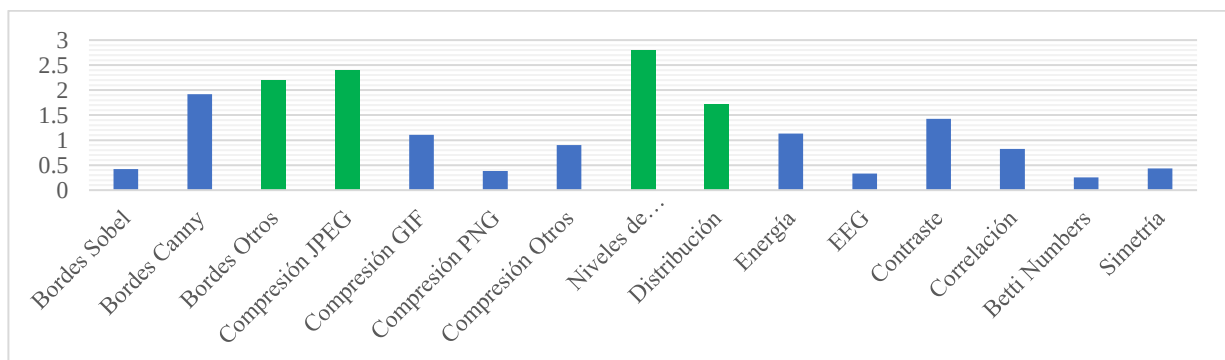


Figura 3.1. Gráfica de puntuaciones alcanzadas por cada característica y resaltadas las seleccionadas.

Tabla 3.5. Características y criterios ponderados para su elección.

	Características															Criterios para elegir características						
Referencia.	Bordes Sobel	Bordes Canny	Bordes Otros	Compresión JPEG	Compresión GIF	Compresión PNG	Compresión Otros	Niveles de intensidad	Distribución	Energía	EEG	Contraste	Correlación	Betti Numbers	Simetría	Citas	Vistas	Referencias	No. Imágenes en experimento	Editorial	Proporciona BD	Nivel de detalle
[2]	0.2	0.2	-	0.3	-	-	0.2	0.2	-	-	-	-	-	-	-	0.28	0.01	0.57	0.4	EL	0	0.8
[3]	-	-	0.2	-	-	-	-	-	-	-	-	-	-	-	-	0.41	0.1	0.07	0.08	IE	0.25	0.8
[5]	0.2	-	-	-	-	-	-	-	0.2	0.2	-	0.2	0.2	-	-	0.03	0.01	0.06	0.1	IE	0.5	0.8
[6]	-	-	-	-	-	-	-	0.2	-	-	-	-	-	-	-	0.03	0.06	0.19	0	IE	0.5	0.6
[11]	-	-	0.2	0.2	-	-	0.2	-	-	0.2	-	-	-	-	-	0.62	0.07	0.12	0.02	IE	0	0.8
[12]	-	0.3	0.3	0.3	0.3	-	-	-	-	-	-	-	-	-	-	0.03	0.11	0.38	0.28	SP	0.5	0.8
[13]	-	-	-	-	-	-	-	0.1	0.1	-	-	-	-	-	-	0.28	0.3	0.05	0	SP	0	0.4
[15]	-	-	0.2	-	-	-	-	-	-	-	-	-	-	-	0.2	1	0	0.2	0	EL	0	0.8
[17]	-	-	-	0.1	0.1	-	-	-	-	-	-	-	-	-	-	0.41	0.39	0.23	0	SP	0	0.4
[18]	-	-	-	0.2	0.2	-	0.2	-	-	-	-	-	-	-	-	0	0	0.26	0.03	AP	0.5	0.6
[19]	-	-	-	-	-	-	-	-	-	0.1	0.1	-	-	-	-	0.03	0.02	0.17	0	IE	0	0.4
[26]	-	-	0.1	-	-	-	-	-	-	-	-	-	-	-	-	0.03	0.01	0.1	0	IE	0	0.8
[28]	-	-	-	-	-	-	-	0.2	0.2	-	-	-	-	-	-	0	0.01	0.06	0	IE	0.5	0.6
[31]	-	0.4	-	0.4	-	-	-	0.4	0.4	0.5	-	0.5	0.3	-	-	0.17	0.33	0.36	0.05	PL	1	1
[32]	-	-	-	-	-	-	-	-	-	-	-	-	-	0.3	-	0.03	0	0.21	0.03	EL	1	0.8
[36]	-	-	-	-	-	-	-	0.3	0.3	-	-	-	-	-	-	0	0.01	0.07	0.01	IE	1	1
[42]	-	-	-	-	-	0.1	-	-	-	-	-	-	-	-	-	0	0.01	0.19	0	IE	0	0.4
[43]	-	-	-	-	-	-	-	-	-	0.1	0.1	-	-	-	-	0.03	0.01	0.12	0.13	IE	0	0.4
[44]	-	-	-	-	-	-	-	-	-	-	0.1	-	-	-	-	0	0.02	0.1	0.01	IE	0	0.4
[45]	-	-	-	-	0.3	-	-	-	-	-	-	-	-	-	-	0.55	0.43	0.36	0.75	AR	0	0.8
[46]	-	-	0.3	-	-	-	-	0.3	-	-	-	0.3	-	-	-	0.48	0.34	0.21	1	SP	0	0.8
[47]	-	-	-	-	-	-	-	-	-	0.1	0.1	-	-	-	-	0	0.07	0.2	0.13	SP	0	0.4
[48]	-	-	0.1	-	0.1	-	-	0.1	-	-	-	-	-	-	-	0.07	0	0.22	0.08	EL	0	0.6
[49]	-	0.2	-	-	-	-	-	0.2	0.2	-	-	0.2	-	-	0.2	0	0	0.3	0.25	EL	0	0.9
[50]	-	0.2	0.2	0.2	-	-	-	-	-	-	-	-	-	-	-	0.1	0.19	0.78	0.46	PL	0	0.8
[51]	-	-	-	-	-	-	-	0.1	-	-	-	-	-	-	-	0	0.01	0.14	0	IE	0	0.6
[52]	-	-	-	-	-	-	-	0.1	-	-	-	-	-	-	-	0.03	0.02	0.08	0.5	IE	0	0.4
[53]	-	0.3	-	0.3	-	-	-	0.3	0.3	-	-	-	0.3	-	-	0.03	0.12	0.16	0.01	SP	1	1
[54]	-	0.3	0.3	0.3	-	0.3	0.3	0.3	-	-	-	0.3	-	-	-	0.86	1	1	0.05	PL	0	0.8
[55]	-	-	0.1	-	-	-	-	0.1	0.1	-	-	-	-	-	-	0.03	0.38	0.13	0	SP	0	0.6
Σ	0.4	1.9	2.2	2.4	1.1	0.4	0.9	2.8	1.7	1.1	0.3	1.4	0.8	0.3	0.4	5.59	4.03	7.09	4.34		6.75	20.3

Es importante mencionar que la fecha en que se realizó la toma de resultados fue en agosto de 2018, por lo que el número de vistas o citas del artículo pudo haber cambiado con el tiempo. Además, estas cifras fueron tomadas directamente de la página de la revista del artículo.

3.2 Extracción de características

Una vez que se seleccionaron aquellas características que lograron mejores puntuaciones se implementaron las técnicas necesarias para la extracción de éstas.

3.2.1 Bordes

Un borde o contorno se define como el lugar donde existe un cambio en la luminosidad, brillantez o color. Sin bordes no hay visión aunque el ojo se encuentre en perfectas condiciones e incida energía luminosa en él. Es decir, si un estímulo no tiene bordes es imposible ver algo [8].

Una imagen con poca cantidad de bordes ofrece poca información que se deba analizar se puede interpretar en un lapso de tiempo menor que una imagen que posee muchos bordes (la persona debe poner atención en mayor cantidad de detalles).

Durante el mecanismo de percepción visual, las personas fijan su atención en los objetos o regiones que constituyen una imagen y captan más rápidamente aquellos cuyos bordes se encuentran más marcados, es decir, aquellos en los que se observa mayor contraste entre el objeto y el fondo [8]. La percepción de objetos con bordes marcados se facilita gracias a que las diferencias de luminosidad, brillantez o color entre píxeles vecinos comunican al cerebro un estímulo con fuerza proporcional al contraste que representa y entonces se origina una respuesta inversamente proporcional en tiempo a la magnitud del estímulo. De esta manera, si los bordes son marcados se requiere un tiempo menor para percibirlos a que si fueran muy tenues.

El mecanismo de percepción visualiza primero los objetos que cuestan menos trabajo, es decir, aquellos en los que el contraste es mayor entre el objeto y el fondo. Luego, es necesario prestar más atención a los de menor contraste porque sus bordes parecen menos marcados al momento de verlos. La Figura 3.2 muestra dos imágenes con el mismo objeto, pero con diferente nivel de gris. Como la diferencia entre el fondo y el tono de gris en la Figura 3.2a es mayor que la diferencia en la Figura 3.2b se observa que es más sencillo detectar la Figura 3.2a que la Figura 3.2b, aun y cuando se trata de la misma forma geométrica.

En resumen, el mecanismo de percepción visual detecta los bordes en niveles, dependiendo de cuánta atención haya que prestárseles. Bajo este enfoque se desarrolló un algoritmo en [56] que clasifica los bordes de una imagen en cinco niveles diferentes.

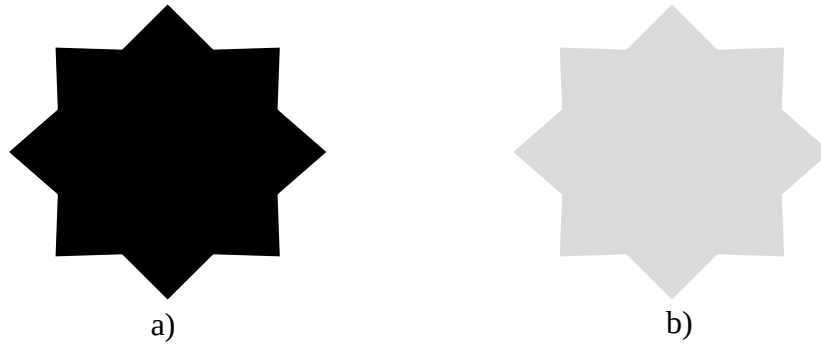


Figura 3.2. a) Bordes marcados, b) bordes tenues.

En [8] se genera una tabla para representar los niveles de bordes con base en los cambios en tonos de grises para agilizar el proceso de [56]. En la Tabla 3.6 se condensa la información y se muestran los intervalos de cambios en los tonos de grises que abarcan los niveles de bordes.

Tabla 3.6. Nivel de borde según el cambio en el tono de gris.

Nivel de bordes	Cambios en el tono de gris (∇p)
5	$1 \leq \nabla p \leq 1$
4	$2 \leq \nabla p \leq 3$
3	$4 \leq \nabla p \leq 9$
2	$10 \leq \nabla p \leq 20$
1	$21 \leq \nabla p \leq 255$

El cambio en el tono de gris del pixel (x,y) se obtiene mediante la Ec. (3.2),

$$\nabla p_{(x,y)} = \max(|p(x,y) - p(x,y+1)|, |p(x,y) - p(x+1,y)|) \quad (3.2)$$

donde p es el valor del pixel en escala de grises, y $||$ denota valor absoluto.

Para obtener la imagen de bordes de nivel i (I_{bi}) se hace uso de la Ec. (3.3),

$$p_{bi(x,y)} = \lambda_{1i} \leq \nabla p_{(x,y)} \leq \lambda_{2i} \quad (3.3)$$

donde $p_{bi(x,y)}$ es el pixel (x,y) de la imagen de bordes de nivel i (I_{bi}), λ_{1i} y λ_{2i} son los umbrales inferior y superior respectivamente del cambio de tono de gris del borde de nivel i .

De los cinco niveles en que se clasifican los bordes, se tomaron solo en cuenta los niveles 1 y 2. Los niveles 3, 4 y 5 se descartan porque se considera que los cambios en tonos de grises que representan son escasos o nulamente perceptibles para el ojo humano.

A partir de la imagen de bordes, se obtiene la densidad de bordes, la cual es la característica para medir la complejidad. Para determinar la densidad de bordes se debe dividir la cantidad de pixeles de bordes del nivel i , entre la cantidad de pixeles totales de la imagen mediante la Ec. (3.4),

$$\Psi_{Bi} = \frac{N_{bi}}{N} \quad (3.4)$$

donde N_{bi} es la cantidad de pixeles bordes del nivel i . De aquí en adelante llamaremos densidad de bordes a esta característica o bien solamente con su notación Ψ_{Bi} .

3.2.2 Compresión

Compresión de imágenes es el proceso de reducción del volumen de los datos de la imagen aplicando técnicas que permiten un trato eficiente de la información. El algoritmo de compresión JPEG permiten obtener una relación de compresión que se correlaciona positivamente a la CI subjetiva dada por personas, ya que el algoritmo es capaz de comprimir más aquellas imágenes que contienen pocos detalles, que aquellas con muchos detalles. Además, el algoritmo JPEG permite pérdidas en la compresión para descartar aquellos detalles de la imagen que el ojo humano no puede percibir fácilmente.

La base de la compresión JPEG es la transformada discreta del coseno (DCT). Se trata de una transformada ortogonal, separable y real que traslada la información del dominio espacial al dominio de la frecuencia para ser representada en una forma más compacta [27]. Por estándar en la compresión JPEG la DCT se lleva a cabo en bloques de 8x8 pixeles. La ecuación para la DCT bidimensional de una imagen $I(x,y)$ se define como,

$$\Theta(u, v) = \beta(u)\beta(v) \sum_{x=0}^{H_x-1} \sum_{y=0}^{H_y-1} I(x, y) \cos\left(\frac{(2x+1)u\pi}{2H_x}\right) \cos\left(\frac{(2y+1)v\pi}{2H_y}\right) \quad (3.5)$$

$$\beta(u) = \begin{cases} \sqrt{\frac{1}{H_x}} & \text{para } u = 0 \\ \sqrt{\frac{2}{H_x}} & \text{para } u = 1, \dots, H_x - 1 \end{cases} \quad (3.6)$$

$$\beta(v) = \begin{cases} \sqrt{\frac{1}{H_y}} & \text{para } v = 0 \\ \sqrt{\frac{2}{H_y}} & \text{para } v = 1, \dots, H_y - 1 \end{cases} \quad (3.7)$$

donde H_x son la cantidad de renglones, H_y es la cantidad de columnas de la imagen, $\theta(u,v)$ denota $H_x \times H_y$ puntos DCT del bloque $I(x,y)$, $u=0,1,\dots,H_x-1$ y $v=0,1,\dots,H_y-1$.

El coeficiente de la matriz DCT con frecuencia cero ($\theta(0,0)$) se denomina "coeficiente DC", mientras que los coeficientes restantes se denominan "coeficientes AC", que reflejan variaciones en los valores de nivel de gris en cierta dirección a cierta velocidad como se muestra en la Figura 3.3.

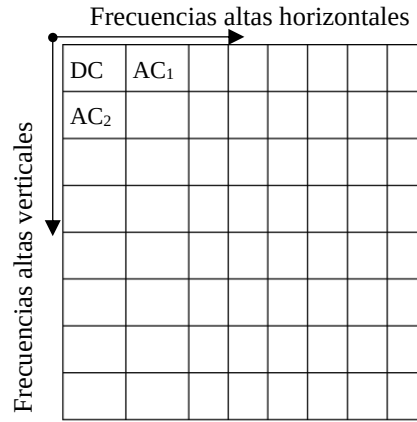


Figura 3.3. Estructura del bloque DCT [27].

Convenientemente, en la mayoría de los subbloques de imágenes, la energía se concentra en las bandas de frecuencia de los subbloques cercanos al DC. Es decir, la información de la imagen (energía) generalmente se concentra en la región de baja frecuencia (esquina superior izquierda). La región de alta frecuencia contiene poca energía y se encuentra en la esquina inferior derecha. Si se suprimen las bandas de frecuencia de AC más altas se pierden los detalles más pequeños que en ocasiones son insignificantes, sin embargo, esto permite que el bloque se represente con menos coeficientes en su estado transformado que en su estado original. Una observación simple es que cada coeficiente DCT $\theta(u,v)$ es una combinación lineal de todos los valores de píxeles dentro del bloque [27].

Para generar la característica de compresión, se utilizó la función `imwrite(I)` de Matlab. Esta función permite comprimir imágenes con diferentes formatos, sin embargo, se ha seleccionado el formato JPEG.

Como ya se ha mencionado, la forma más común de utilizar la información de un algoritmo de compresión de imágenes para calcular la complejidad es mediante la relación de compresión, la cual es una medida entre el tamaño del archivo comprimido y el tamaño del archivo sin comprimir Ec. (3.8),

$$\Psi_C = \frac{TC_B}{T_B} \quad (3.8)$$

donde TC_B es el tamaño en bytes del archivo de la imagen comprimida y T_B es el tamaño del archivo de la imagen no comprimida. De aquí en adelante llamaremos relación de compresión a esta característica o bien solamente por su notación Ψ_C .

Para obtener TC_B se comprime la imagen y se estima el tamaño de la imagen comprimida, mientras que por otro lado T_B se calcula a partir del espacio que ocuparía la imagen original en memoria. Esto se obtiene de multiplicar la cantidad de píxeles N de la imagen y los bits que se utilizan por cada píxel, por ejemplo, para imágenes RGB por lo general se usan 8 bits por cada canal ($N \times 8 \times 8 \times 8$).

3.2.3 Niveles de intensidad

El nivel de intensidad se define en esta investigación como el valor RGB para imágenes a color o la intensidad de gris para imágenes en escala de grises.

Algunas investigaciones (vea Tabla 3.4) han demostrado que en varios casos los niveles de intensidad que se encuentran en la imagen pueden influir en la percepción de la CI subjetiva.

El color es una característica inmediata y distinguida en la percepción visual humana, permite que un observador pueda distinguir las diferencias entre dos campos sin estructura del mismo tamaño y misma forma. Además, se ha identificado que las composiciones de colores influyen incluso en el precio de pinturas, por lo que el color es la primera y más importante característica en las pinturas artísticas de entre todas las características globales [30].

Para medir la complejidad de los colores en una imagen, se emplea el método de medición de la complejidad del color (CCM) [57], CCM es definido por la Ec. (3.9),

$$\psi(i, j) = \sum_{x, y \in \Omega(i, j)} G_{\alpha} (\|p(x, y) - \bar{p}(i, j)\|) \quad (3.9)$$

donde $p(x, y)$ es el valor del píxel que pertenece a la ventana local $\Omega(i, j)$, G_{α} denota una función de ponderación gaussiana y $\bar{p}$ es un valor de color promedio dentro de un tamaño de ventana local $\Omega(i, j)$ centrado en (i, j) y $\| \cdot \|$ denota la medida de diferencia de color.

El promedio del valor del color de una ventana es calculado por,

$$\bar{p}(i, j) = \frac{1}{\tau} \sum_{x, y \in \Omega(i, j)} p(x, y) \quad (3.10)$$

donde τ es el número de píxeles en la ventana local $\Omega(i, j)$. La medida de la diferencia de color es importante para representar la percepción visual humana de los colores. En general, la diferencia de color se evalúa utilizando alguna medida de distancia que depende del espacio de color en el que se esté trabajando.

Se hace uso del espacio de color CIELab ya que describe todos los colores visibles para el ojo humano. Un espacio de color puede ser descrito como un método para expresar el color de un objeto usando algún tipo de anotación. La Comisión Internacional de Iluminación (*Commission Internationale de l'Éclairage* o *CIE*) es una organización sin fines de lucro que desarrolla espacios de color. Definió el espacio de color CIELab que correlaciona los valores numéricos de color consistentemente con la percepción visual humana. Este espacio de color fue modelado con base en una teoría de color oponente que establece que dos colores no pueden ser rojo y verde o amarillo y azul al mismo tiempo. CIELab tiene tres componentes L, a y b, donde 'L' es luminosidad, 'a' se refiere a las coordenadas rojo/verde (+a indica rojo, -a indica verde) y 'b' indica las coordenadas amarillo/azul (+b indica amarillo, -b indica azul). Para medir la diferencia de color en este espacio se emplea la Ec. (3.11) propuesta por [57],

$$D(p(i, j), p(x, y)) = 1 - \exp \left[- \frac{E(p(i, j), p(x, y))}{\gamma} \right] \quad (3.11)$$

donde γ es el factor normalizado y $E(p(i, j), p(x, y))$ es la distancia euclidiana en el espacio de color CIELab. Una pequeña distancia euclidiana entre dos puntos de color en CIELab es

proporcional a la diferencia que percibe el sistema visual humano. En [57] se identifica que una distancia euclidiana más grande no tiene significado, sino solo una gran diferencia en el sistema visual humano.

$$E(p(i, j), p(x, y)) = \sqrt{(L_{ij} - L_{xy})^2 + (a_{ij} - a_{xy})^2 + (b_{ij} - b_{xy})^2} \quad (3.12)$$

Mediante la Ec. (3.9) se calcula la complejidad del color del píxel (i, j) con base en sus vecinos de la ventana local. De manera que para calcular la medida de complejidad del color total de la imagen se utiliza la Ec. (3.13),

$$\Psi_L = \frac{1}{N} \sum_{i, j \in I(x, y)} \psi(i, j) \quad (3.13)$$

donde, $\psi(i, j)$ es la complejidad del píxel que pertenece a la imagen $I(x, y)$, por lo que esto es obtener el promedio de la complejidad de todos los píxeles de la imagen. De aquí en adelante se hará referencia a esta característica como medida de complejidad del color (CCM) o bien por su notación Ψ_L .

3.2.4 Distribución

La distribución de la imagen es definida en esta investigación como una medida de la aleatoriedad de la repartición de cierta característica en el espacio de la imagen. La frecuencia con la que un patrón aparece en la escena podría explicar cómo percibimos la aleatoriedad.

Antes de realizar cualquier operación sobre la imagen para calcular la distribución de ésta, primero se aplica un filtrado de mediana, el cual es particularmente efectivo en la presencia de ruido impulso, también llamado sal y pimienta [1]. La razón de utilizar este filtro es para eliminar aquellos valores de los píxeles atípicos de la imagen. Suponga que la ventana en que se está aplicando el filtrado contiene los siguientes valores (1,1,2,2,1,2,3,3,12), estos valores son ordenados de forma ascendente de la siguiente manera (1,1,1,2,2,2,3,3,12), por lo que la mediana será el 2, como podemos observar el valor 12 el cual es un valor no esperado en esa ventana será ignorado, y ya no afectará. Con esto se busca tener patrones en la imagen más consistentes y que el posible ruido no afecte tanto.

Para determinar el patrón de la imagen ya filtrada se construye una matriz de coocurrencia con la función de Matlab `graycomatrix(I)` la cual calcula la frecuencia que un pixel con valor η_1 de nivel de gris aparece horizontalmente adyacente a un pixel con valor η_2 . De forma predeterminada, si I es una imagen binaria, la función `graycomatrix` escala la imagen a dos niveles de grises. Si I es una imagen de intensidad, `graycomatrix` escala la imagen a ocho niveles de gris.

En la Figura 3.4 se muestra cómo `graycomatrix` calcula varios valores en la matriz de coocurrencia de la imagen $I(x,y)$ de 4 por 5. El elemento (1,1) de la matriz de coocurrencia contiene el valor 1 porque sólo hay una instancia en la imagen donde dos píxeles adyacentes horizontales tienen los valores 1 y 1. El elemento (1,2) de la matriz de coocurrencia contiene el valor 2 porque hay dos instancias en la imagen donde dos píxeles adyacentes horizontales tienen los valores 1 y 2, y así de forma similar para todos los demás casos.

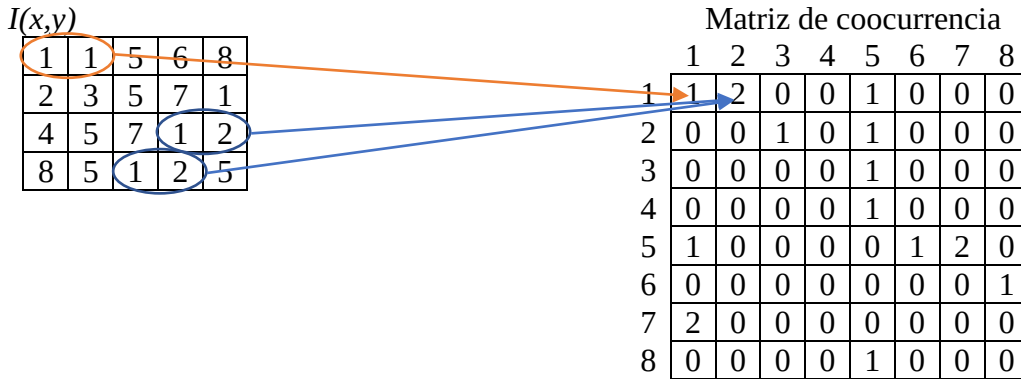


Figura 3.4. Ejemplo de generación de la matriz de coocurrencia.

Hasta este punto se ha construido una matriz con el patrón de la imagen, sin embargo, aún hay que estimar que tan aleatorio es éste, y para esto se puede emplear la entropía, la cual es una medida estadística de aleatoriedad, por lo que se utiliza para calcular la aleatoriedad de la matriz o bien del patrón mediante la Ec. (3.14),

$$\Psi_D = -\sum_{i,j} O_{i,j} \log(O_{i,j}) \quad (3.14)$$

donde O es la matriz de coocurrencia. De aquí en adelante se le llamará entropía del patrón o bien por su notación Ψ_D a esta característica para determinar la complejidad de las imágenes.

3.3 Análisis de características

Una vez seleccionadas las características que se utilizarán, se realizó un análisis más a fondo sobre cada una de éstas, de tal forma que se encontraron algunas fortalezas y debilidades. A continuación, se describen brevemente.

3.3.1 Densidad de bordes

3.3.1.1 Fortalezas en densidad de bordes

Una imagen con mayor cantidad de bordes generalmente indica mayor complejidad ya que los bordes contienen información importante de la imagen, como lo pueden ser objetos y/o áreas de interés. En la Figura 3.5 se muestra un ejemplo, en el cual se observa que los bordes de la Figura 3.5a la hacen ser más compleja (por lo general para el ser humano) que la Figura 3.5b aun y cuando esta otra imagen tiene más colores diferentes.

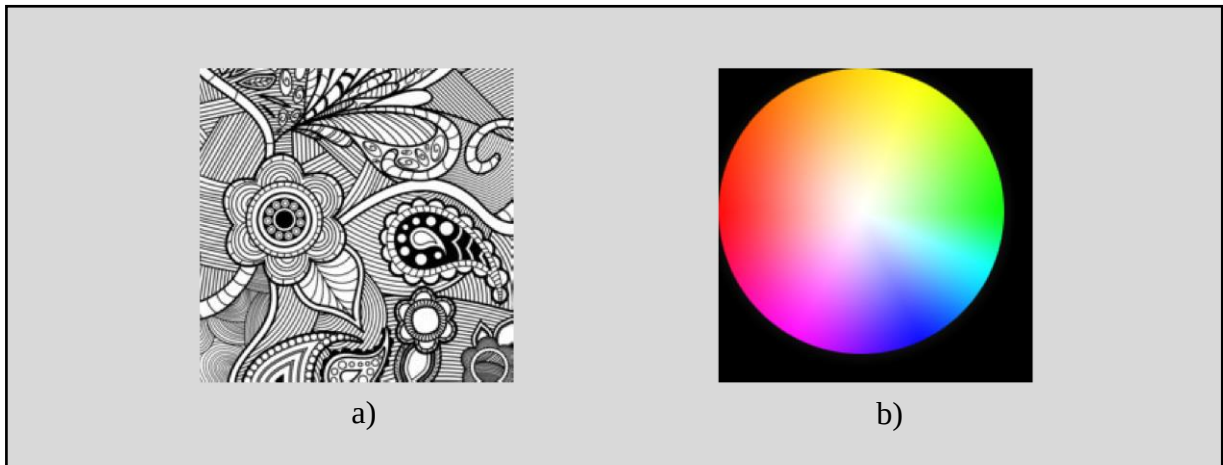


Figura 3.5. a) Imagen con muchos bordes, b) imagen con pocos bordes.

Al calcular la densidad de bordes de ambas imágenes se obtiene que para la imagen a) una densidad de bordes de 0.4928 mientras que para b) se tiene 0.0305.

3.3.1.2 Debilidades en densidad de bordes

Mediante el análisis de bases de datos se ha encontrado cómo imágenes que presentan patrones simples en los bordes tienden a ser catalogadas menos complejas que aquellas imágenes con patrones más irregulares. La Figura 3.6 muestra dos imágenes tomadas de [21],

donde por lo general las personas clasificaron menos compleja a la Figura 3.6a con una complejidad subjetiva promedio de 3, mientras que la Figura 3.6b obtuvo una complejidad promedio de 4. Sin embargo, se puede observar que la Figura 3.6a contiene una mayor cantidad de pixeles de bordes. De hecho se extrajo la densidad de bordes para ambas imágenes y efectivamente la Figura 3.6a tiene una densidad de bordes de 0.1358 mientras que la Figura 3.6b tiene 0.0803, lo cual es contrario a valor promedio subjetivo que se le dio de complejidad en [21].

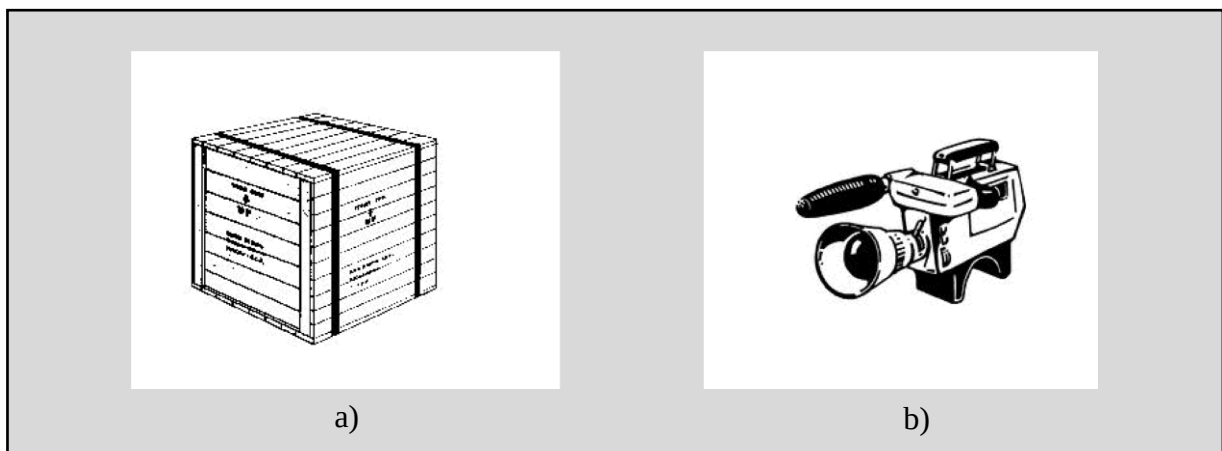


Figura 3.6. Dos imágenes con diferentes complejidades y densidades de bordes.

3.3.2 Relación de compresión

3.3.2.1 Fortalezas en relación de compresión

La compresión de imágenes tiene un comportamiento similar a la forma de exclusión de información del ser humano, ya que la compresión que permite pérdidas como lo es el caso de compresión JPEG puede descartar información excesiva, por ejemplo, variar en uno el bit menos significativo de un pixel puede ser imperceptible al ojo humano al igual que para el compresor de imágenes. También la compresión de imágenes permite manejar de manera más eficiente la información redundante, esto es, en los lugares uniformes de la imagen se comprime mejor ya que el algoritmo los puede representar con menos datos que si se manejan pixel por pixel, de manera similar el humano asocia regiones lisas o uniformes con regiones simples o bien poco complejas.

En la Figura 3.7 se muestra un ejemplo que cumple con la ventaja de ignorar detalles muy pequeños, donde la Figura 3.7a muestra un cuadro liso gris, y la Figura 3.7b muestra un cuadro que existe textura en él o más concretamente otros cuadrados más pequeños por toda la superficie. Estos cuadros más pequeños tienen un valor de gris muy similar al del cuadro grande, por eso son escasamente o nulamente perceptibles por el ojo humano. A ambas imágenes se les calculo la relación de compresión y se obtuvo un resultado de 0.063, lo cual era de esperarse ya que la compresión ignora esos detalles tan pequeños y produjo una relación de compresión igual para ambos casos.



Figura 3.7. Dos imágenes que generan misma relación de compresión.

3.3.2.2 Debilidades en relación de compresión

Algunos algoritmos de compresión como el JPEG se pueden ver afectados por el escalamiento y/o la rotación de la imagen, de manera que generen diferentes resultados si la misma imagen es procesada por este algoritmo, pero se ha rotado o escalado, por ejemplo, en la Figura 3.8, se observa la misma escena pero con diferente rotación y diferente relación de compresión. Para la Figura 3.8a se obtuvo una relación de compresión de 0.10, mientras que la Figura 3.8b fue de 0.09, cuando un humano podría identificar que se trata de la misma imagen y la misma complejidad.



Figura 3.8. Dos imágenes con diferente relación de compresión, pero mismo contenido.

Otro punto a tener en cuenta con el algoritmo de compresión JPEG es que utiliza máscaras de 8x8 por lo que características globales de la imagen no serán captadas por completo. Esto es, si se tiene un patrón muy repetido como el caso de las imágenes fractales, la compresión JPEG no será capaz de identificarlo, por lo que no cumple con el principio de invarianza de Gestalt explicado en el capítulo 1.

3.3.3 Medida de complejidad del color

3.3.3.1 Fortalezas en medida de complejidad del color

Los niveles de intensidad pueden ser útiles para diferenciar en complejidad cuando otras características no lo hacen. Tal es el caso de la Figura 3.9 donde los bordes no serían capaces de diferenciar entre ellas, es decir, determinar la complejidad de ambas imágenes a través de bordes regresaría el mismo resultado. Sin embargo, determinar la complejidad de ambas imágenes a través de la cantidad de niveles de intensidad es mejor ya que esta característica sí es capaz de diferenciar entre ambos casos, se puede apreciar que la Figura 3.9a tiene más colores o niveles de intensidad diferentes que la Figura 3.9b. Calculando CCM para la Figura 3.9a se obtiene 0.2875 mientras que para la Figura 3.9b se obtiene 0.0654.

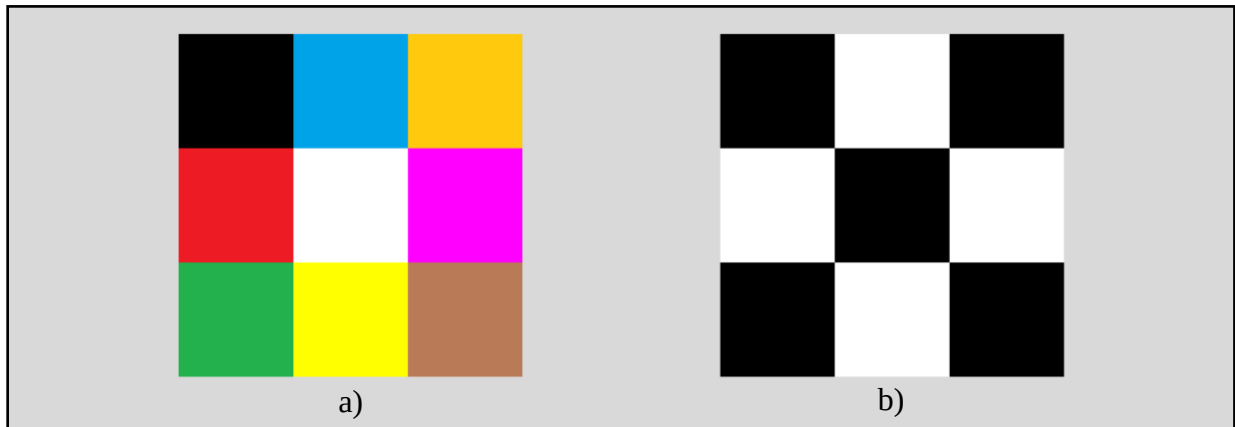


Figura 3.9. Dos imágenes con la misma cantidad de bordes, pero diferentes niveles de intensidad.

3.3.3.2 Debilidades en medida de complejidad del color

Un mayor número de niveles de intensidad no siempre está relacionado a una mayor complejidad de la imagen tal y como se vio en la fortaleza de densidad de bordes. Sin embargo, en el caso donde se tengan imágenes con la misma cantidad de colores, la medida de complejidad de color puede dar resultados distintos, ya que la CCM se construye a partir de la diferencia en el espacio de color entre píxeles, lo cual no siempre implica que una mayor diferencia en el espacio de color sea una mayor complejidad. Este caso se puede observar en la Figura 3.10 la cual consta de cuatro imágenes, cada una de ellas se trata de dos rectángulos sólidos de diferentes colores, la Figura 3.10a tiene una CCM de 0.0645, Figura 3.10b de 0.0430, Figura 3.10c de 0.0430 y Figura 3.10d de 0.0543.

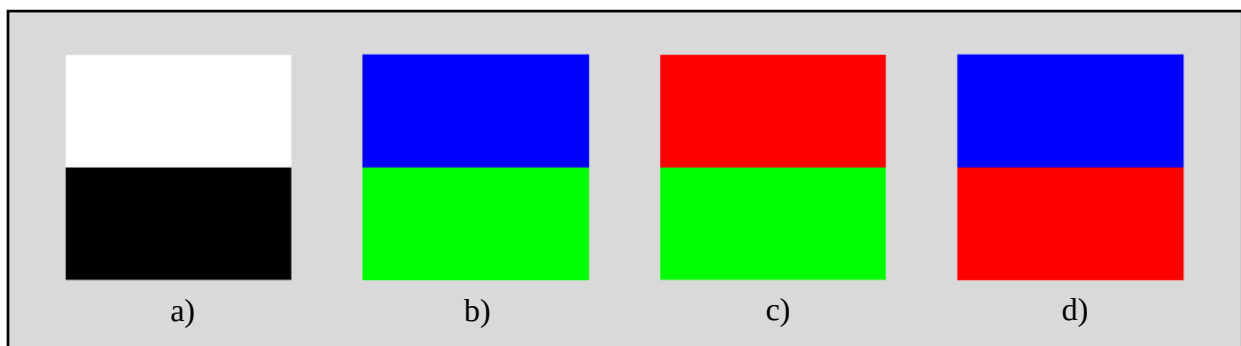


Figura 3.10. Ejemplo de cuatro imágenes con dos niveles de intensidad distintos.

3.3.4 Entropía del patrón

3.3.4.1 Fortalezas en entropía del patrón

Como ya se ha mencionado anteriormente, la aleatoriedad del patrón de la imagen es un factor importante cuando se trata de estimar la complejidad de ésta, tal es el caso de la Figura 1.11 donde se muestran dos imágenes, una de ruido y otra de regiones lisas, ambas imágenes tienen la misma cantidad de colores y la misma cantidad de píxeles para cada uno de estos colores, sin embargo la imagen de ruido es más compleja al presentar un desorden.

Otro ejemplo similar se encuentra en la Figura 3.11 la cual muestra dos imágenes en blanco y negro. La Figura 3.11a tiene el patrón ordenado de un tablero de ajedrez, mientras que la Figura 3.11b muestra un patrón aleatorio. Al calcular la entropía del patrón para la Figura 3.11a se obtuvo un 0.0571 mientras que para la Figura 3.11b se obtuvo un 0.0754.

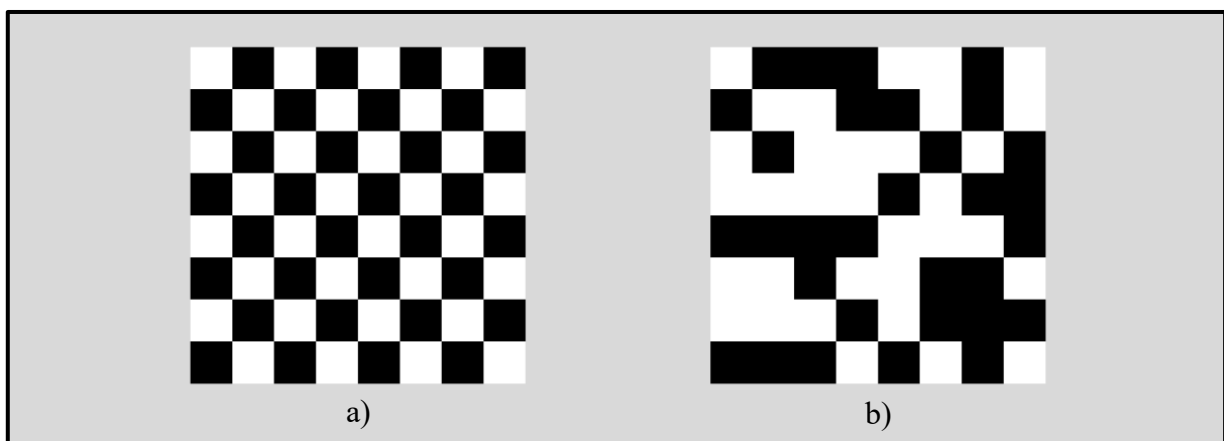


Figura 3.11. Dos imágenes con diferente patrón.

3.3.4.2 Debilidades en entropía del patrón

La creación de la matriz de coocurrencia es de forma horizontal a la derecha, lo cual influye en el resultado al estimar la entropía del patrón. Tal es el caso de la Figura 3.12 donde se muestran dos imágenes cuadradas “iguales”, solo que la Figura 3.12b ha sido rotada en 90 grados. Cuando se calcula la entropía del patrón de la Figura 3.12a se obtiene un resultado de 0.0552, mientras que para Figura 3.12b se obtiene 0.0414, el cual es diferente y esto es por la rotación que se realizó.

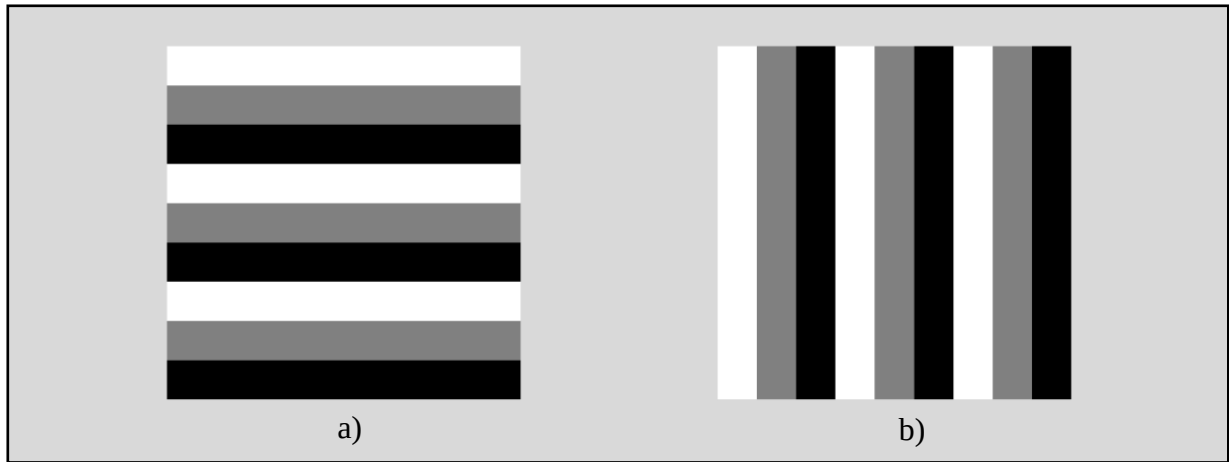


Figura 3.12. Dos imágenes con un patrón similar.

3.4 Conclusión

Mediante el análisis de la literatura realizado se seleccionaron las características de densidad de bordes, relación de compresión, medida de complejidad de color y entropía del patrón, ya que estas fueron las que lograron las mejores puntuaciones según los criterios establecidos en esta investigación. Es importante recordar que se busca una buena correlación entre estas características y el criterio subjetivo del humano, por lo que estas no pueden ser las mejores para ciertos grupos de personas, por ejemplo, es posible que para personas del oriente se correlacione mejor relación de compresión GIF que JPEG.

También se realizó un análisis de fortalezas y debilidades de las características seleccionadas, y se observó que en muchos casos lo que puede representar una debilidad para una característica para otra puede ser una fortaleza. Tal es el caso de la Figura 3.9, donde la densidad de bordes regresaría el mismo valor para ambas imágenes, sin embargo, la CCM puede distinguir cuál de las dos es más compleja. Por lo que las características seleccionadas se complementan, lo cual es deseado para los sistemas que se diseñaran más adelante.

CAPÍTULO IV. LA RESOLUCIÓN Y SUS EFECTOS EN LA COMPLEJIDAD

La resolución de una imagen indica la cantidad de detalles que puede observarse en ésta. Imágenes con resolución alta se traduce en imágenes que puede tener una mejor calidad visual, ya que una resolución alta permite más detalles que una baja. El escalar una imagen implica modificar su resolución, ya que si la imagen es escalada con un factor menor a 1, ésta debe perder pixeles, caso contrario si es escalada a un factor mayor a 1.

Para modificar la resolución espacial de las imágenes se usó la función `imresize(im,escala)` de Matlab la cual puede regresar una nueva imagen con diferente resolución. Esta función hace interpolación para obtener el nuevo valor del pixel, lo cual consiste en un promedio ponderado en el vecindario más cercano al pixel en análisis.

En la Figura 4.1 se muestra una serie de imágenes con la letra 'D', donde la Figura 4.1a tiene una resolución de 200x200 pixeles, la Figura 4.1b se redujo a una escala de 0.5 quedando una resolución de 100x100 pixeles, la Figura 4.1c se bajó a una escala de 0.25 quedando con una resolución de 50x50 pixeles, la Figura 4.1d se redujo la escala a 0.125 quedando una resolución de 25x25 pixeles y por último la Figura 4.1e se ha reducido a una escala de 0.01 quedando a una resolución de 2x2 pixeles.

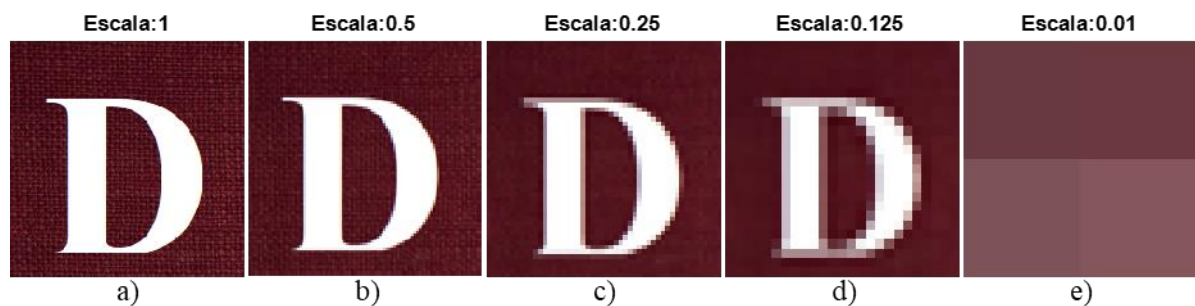


Figura 4.1. Imagen de letra 'D' con resolución de 200 x 200 pixeles, a) original, b) escalada al 0.5, c) escalada a 0.25, d) escalada a 0.125 y d) escalada al 0.01.

En la Figura 4.1 se puede observar que al bajar la resolución los detalles se van perdiendo y la calidad visual se degrada. Además al disminuirla causa que los detalles finos como las texturas (para este caso de la Figura 4.1 el fondo) se pierdan, llegando incluso a aparentar que la textura son regiones uniformes y solamente se puedan ver las características más generales como lo es el color predominante y las formas. El disminuir demasiado la resolución de la

imagen puede causar que información importante de la imagen se pierda, tal es el caso de la Figura 4.1e donde la letra ‘D’ se perdió por completo.

En este capítulo se muestran algunos experimentos realizados para observar si variar la resolución de las imágenes impacta en las características seleccionadas para determinar la complejidad objetiva.

4.1 Las características de complejidad objetiva y la resolución

Con el fin de observar si existen cambios en las características seleccionadas al variar la resolución de las imágenes se realizó un experimento similar al expuesto en [11], el cual consiste en extraer las características en diferentes escalas de las imágenes.

Para este análisis se tomaron las imágenes de las bases de datos RS1, RS2 y SAVOIAS, las cuales suman un total de 1,518. Estas fueron clasificadas en 8 categorías. La categoría 1 corresponde a RS1 y RS2 ya que se juntaron estas bases de datos que contienen imágenes de entornos reales. Las categorías 2 al 8 se formaron de acuerdo a la base de datos SAVOIAS Escenas, Anuncios, Visualizaciones, Objetos, Diseño de interiores, Arte y Suprematismo.

Para variar la resolución de las imágenes éstas fueron escaladas de 0.05 a 1.0 en intervalos de 0.05 (20 escalas distintas), por lo que suman un total de 30,360 imágenes distintas y se les extrajo las siguientes características Ψ_i :

- 1- Densidad de bordes nivel 1 (DBordes1).
- 2- Densidad de bordes nivel 2 (DBordes2).
- 3- Relación de compresión JPEG (JPG).
- 4- Medida de complejidad del color (CCM).
- 5- Entropía del patrón (Dist).

Ya con las características se calculó el valor promedio por categoría y escala. De tal forma que se generaron las gráficas mostradas de la Figura 4.2 a la 4.5, donde el eje de Complejidad representa el valor promedio de las características para una escala específica que se indica en el eje Escala.

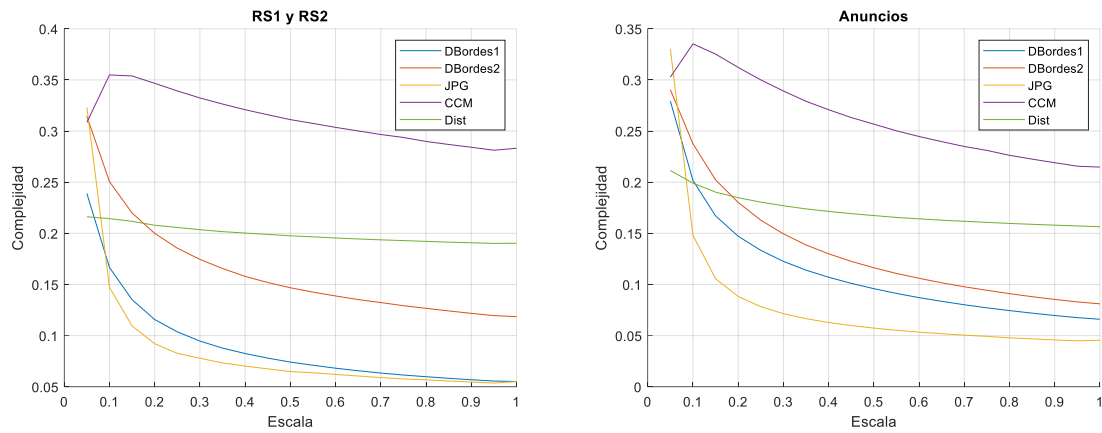


Figura 4.2. Gráficas de complejidad contra escala de imágenes RS1, RS2 y anuncios.

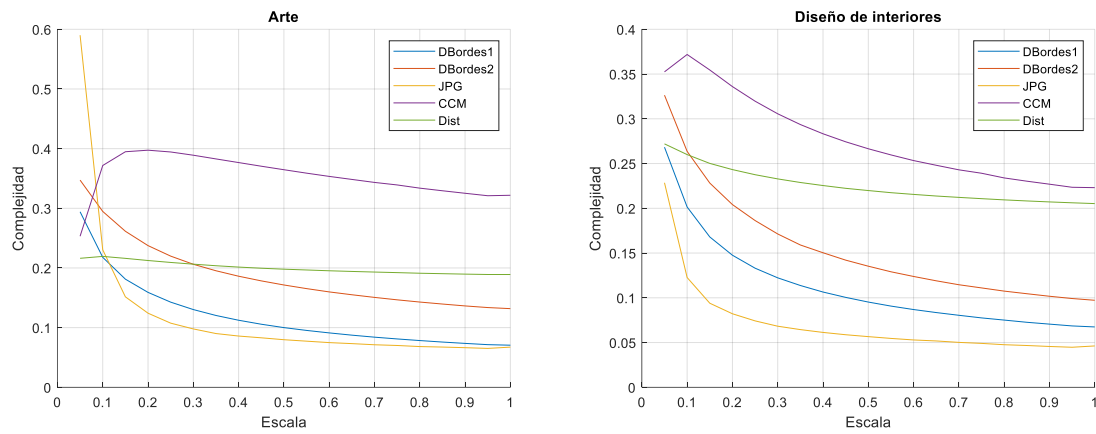


Figura 4.3. Gráficas de complejidad contra escala de imágenes arte y diseño de interiores.

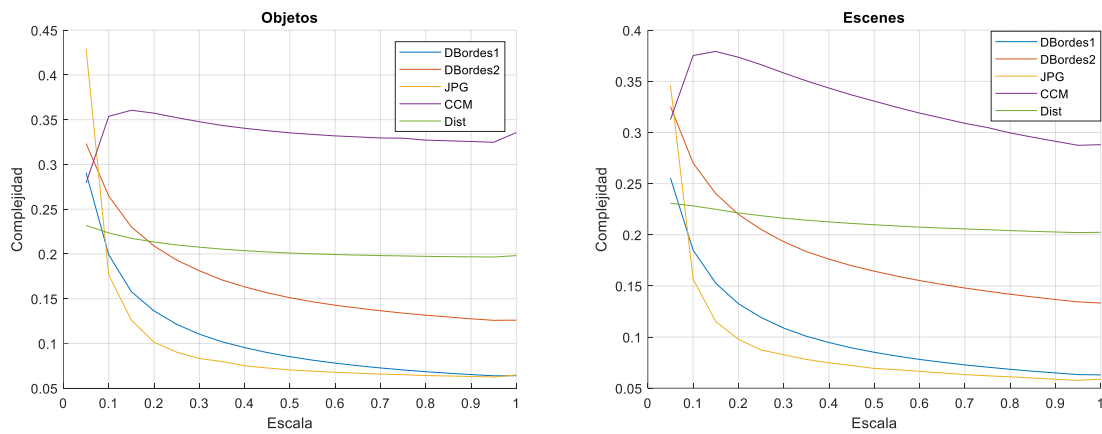


Figura 4.4. Gráficas de complejidad contra escala de imágenes objetos y escenarios.

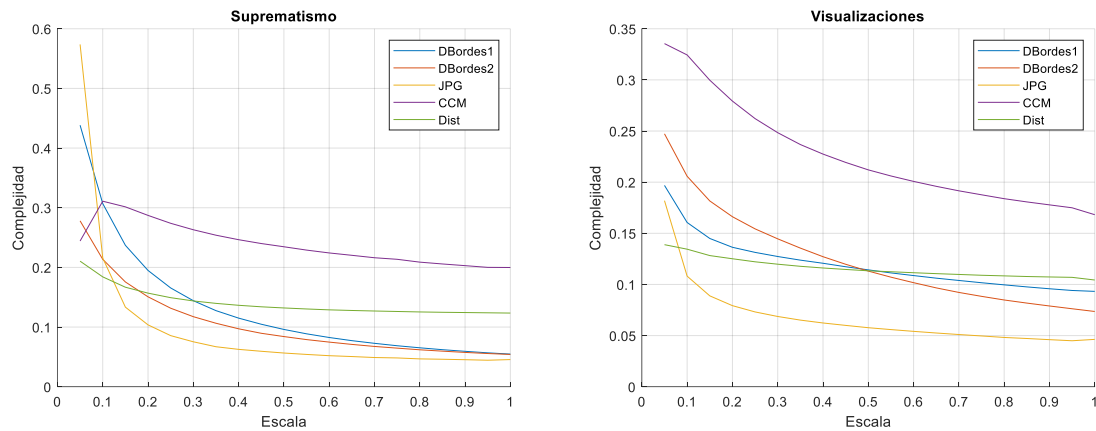


Figura 4.5. Gráficas de complejidad contra escala de imágenes suprematismo y visualizaciones.

De la Figura 4.2 a la 4.5 se puede observar que el comportamiento promedio de las Ψ_i es incrementar conforme la resolución de las imágenes disminuye. Sin embargo, este es un comportamiento promedio por lo que se realizó un análisis individual de distintas imágenes. Con el análisis individual de imágenes se encontraron dos casos comunes: el primero que es el más frecuente es aquel donde las Ψ_i se incrementan conforme la resolución disminuye, un ejemplo de esto se muestra en la Figura 4.6. El segundo caso se da cuando la escala disminuye y las Ψ_i también disminuyen, un ejemplo de esto se muestra en la Figura 4.7.

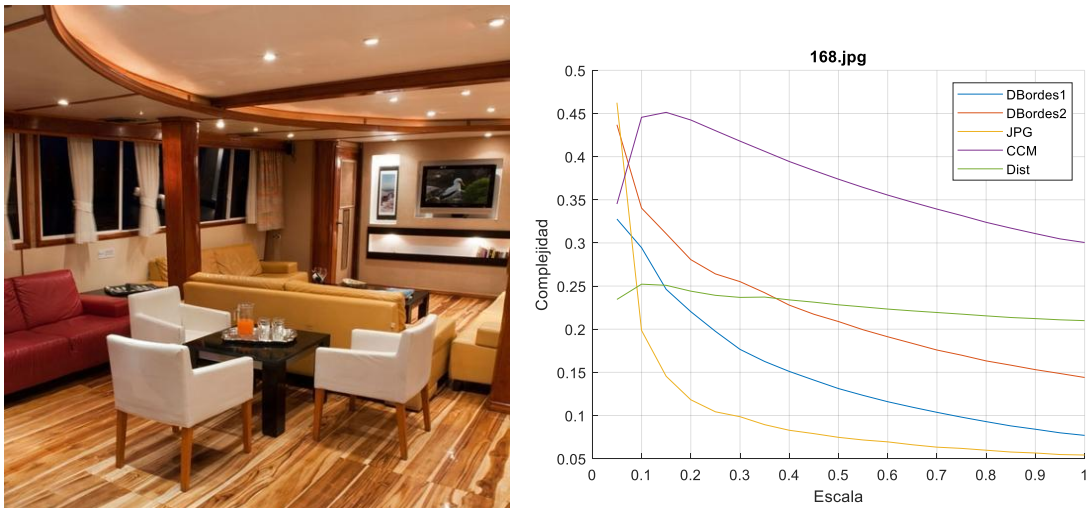


Figura 4.6. Imagen 168 de escenas de la base de datos SAVOIAS y su complejidad para diferentes escalas.

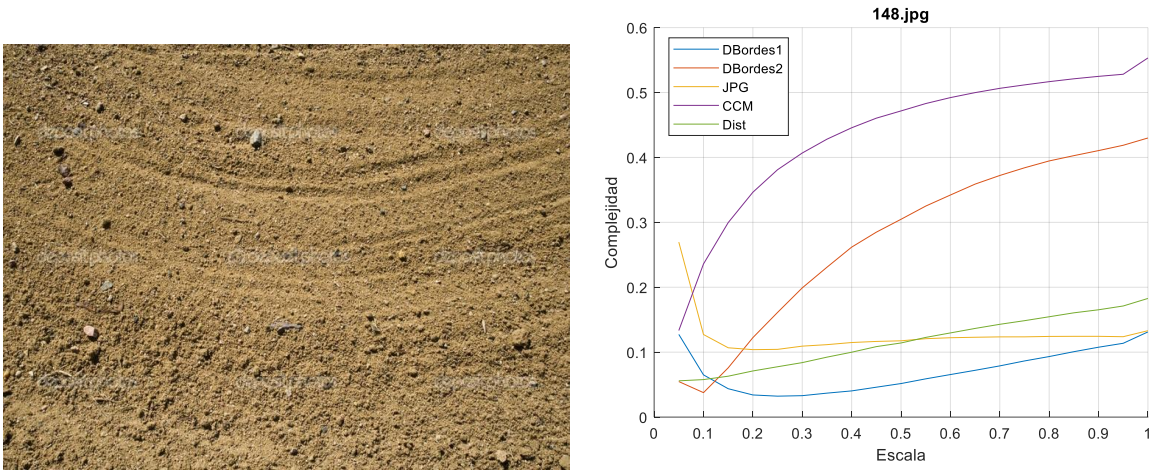


Figura 4.7. Imagen 148 de escenas de la base de datos SAVOIAS y su complejidad para diferentes escalas.

Las imágenes que contienen en su mayoría detalles finos como texturas las Ψ_i se ven atenuadas conforme la resolución baja, ya que de acuerdo con [11], los detalles finos se pierden cuando la imagen es submuestreada, y la creación de nuevas características localizadas no puede compensar esta pérdida de información. Caso contrario, cuando las imágenes casi no contienen detalles finos como texturas las Ψ_i se ven incrementadas conforme la resolución decrementa, ya que las regiones largas se convierten en pequeñas y las regiones pequeñas se convierten en características localizadas, como líneas, bordes o esquinas.

4.2 Correlación entre complejidad subjetiva y las características

Como ya se ha demostrado, para las bases de datos RS1, RS2 y SAVOIAS las Ψ_i de las imágenes se puede ver afectadas por el cambio de resolución, por lo que surge la duda sobre qué ocurre con la correlación entre la complejidad subjetiva y las Ψ_i al variar la resolución de las imágenes. Para esto, se propone un experimento similar al de la sección 4.1, donde se separaron las imágenes de las bases de datos en 8 categorías. Después cada imagen es escalada en un rango de 0.05 a 1.0 para variar la resolución, y a cada imagen se le extraen las características Ψ_i . Una vez que se tienen todos los valores se obtiene la correlación entre complejidad subjetiva y las características. Por último, se graficó la correlación por categoría y escala como se muestra en las Figura 4.8 a 4.11.

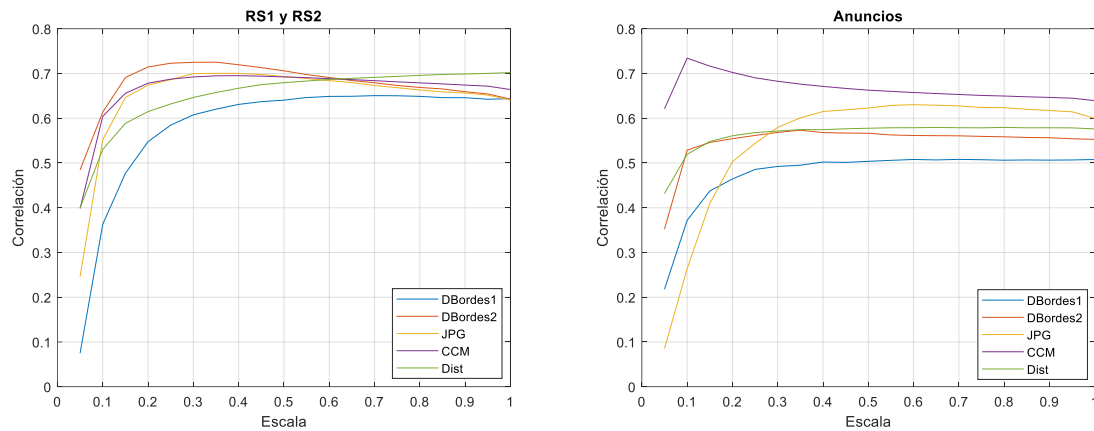


Figura 4.8. Gráficas de correlación entre complejidad subjetiva y características contra escala de imágenes RS1, RS2 y anuncios.

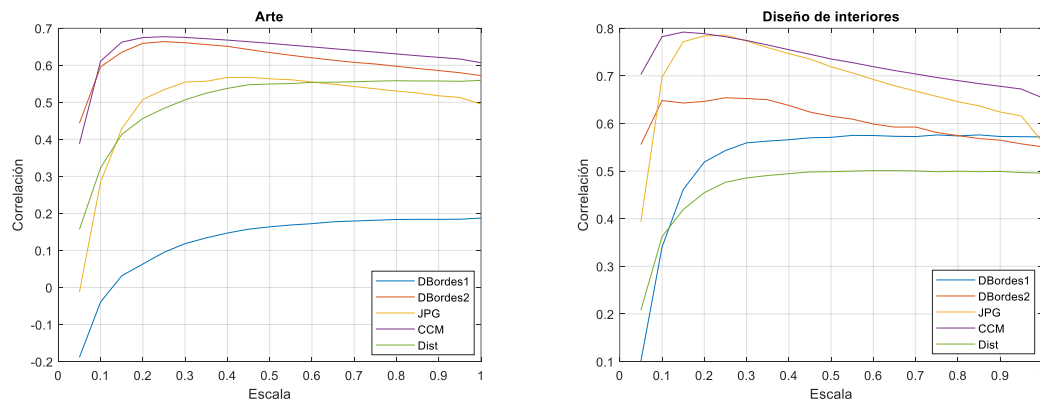


Figura 4.9. Gráficas de correlación entre complejidad subjetiva y características contra escala de imágenes de arte y diseño de interiores.

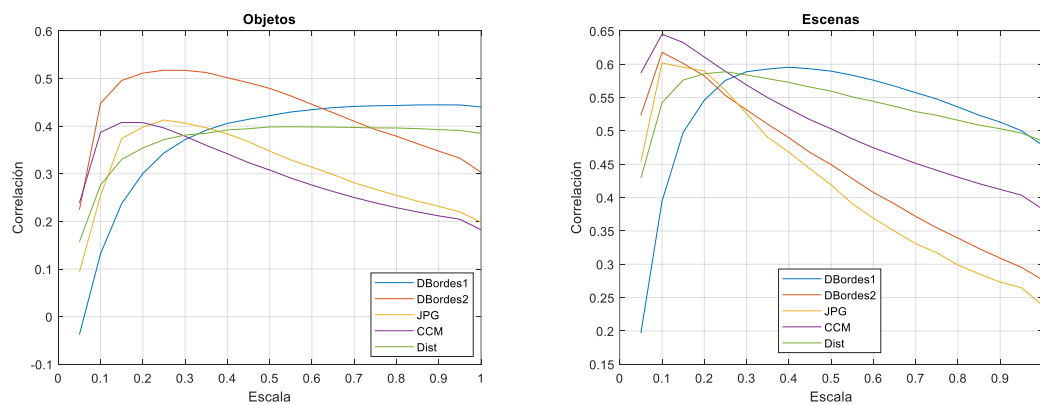


Figura 4.10. Gráficas de correlación entre complejidad subjetiva y características contra escala de imágenes de objetos y escenas.

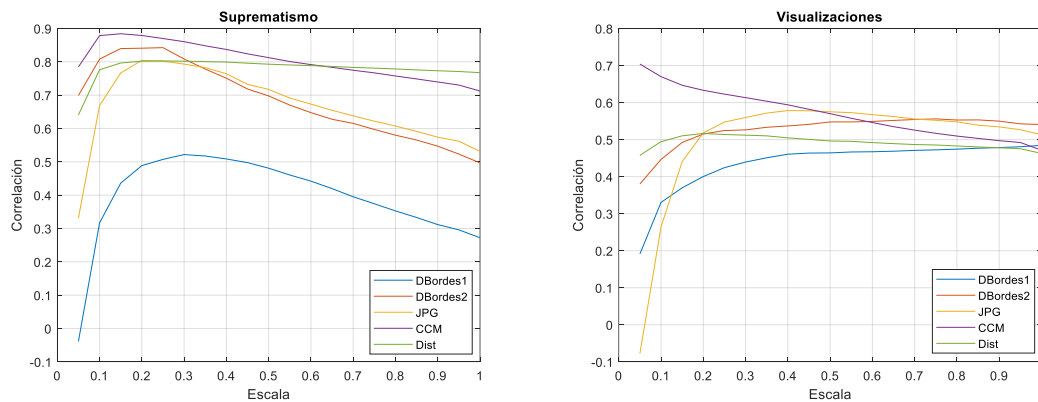


Figura 4.11. Gráficas de correlación entre complejidad subjetiva y características contra escala de imágenes de suprematismo y visualizaciones.

De la Figura 4.8 a la 4.11 se puede observar que por lo general las características de densidad de bordes nivel 2, relación de compresión JPEG y medida de complejidad de color, mejoran su correlación con la complejidad subjetiva conforme la resolución de la imagen disminuye. También se puede apreciar que la correlación cambia drásticamente para las imágenes de la categoría Escena, mientras que para la categoría de imágenes de Anuncios la correlación se mantiene más constante. De acuerdo con [11], esto se debe a que las imágenes de Escenas tienen en su mayoría detalles finos como texturas, los cuales los humanos tienden a ignorar cuando se les pide que estimen la complejidad, mientras que la computadora no. Por lo tanto, al reducir la escala mejora la correlación para este caso, ya que al submuestrear se eliminan detalles finos y la complejidad objetiva de la imagen se ajusta más a la que un humano estima.

4.3 La complejidad subjetiva y la resolución

Cuando se realiza el experimento para la estimación de la complejidad subjetiva es común que las imágenes a evaluar sean desplegadas en un monitor. Existen casos donde no se considera el efecto que puede tener dicho monitor en los resultados, ya que la imagen que se está viendo puede estar escalada, que se traduce a que hubo un aumento de datos o una reducción de estos.

En la Figura 4.12 se muestra un ejemplo de una imagen con una resolución alta (5100x3300) que debe ser mostrada en un área determinada de un monitor Full HD, el cual escala la imagen

a un factor de 0.19, por lo tanto esta imagen está perdiendo información y no llegará completa a la persona que este evaluando a la imagen en ese monitor.



Figura 4.12. Ejemplo de una reducción en la escala de la imagen al tener que ser mostrada en un monitor.

Por otro lado, también se puede tener el caso contrario, donde la imagen sea de una baja resolución y se deba de mostrar en una área de un monitor con una resolución más grande, como el ejemplo de la Figura 4.13 donde se tiene una imagen de baja resolución (286x368) y ésta es escalada en un factor de 2.6. Para este ejemplo a la imagen se le debe añadir información mediante algún método de escalamiento, por lo tanto, la información que le está llegando al humano ya ha sido modificada.



Figura 4.13. Ejemplo de un aumento en la escala de la imagen al tener que ser mostrada en un monitor.

4.4 Conclusión

La resolución de la imagen no se toma en cuenta en la mayoría de trabajos sobre complejidad, sin embargo, se ha demostrado en [11] y en los experimentos realizados en este

capítulo que la resolución afecta a las características para determinar la complejidad objetiva. Por lo tanto, es importante considerar la resolución de las imágenes en los sistemas que determinan la complejidad objetiva.

Por otro lado, no se reportó en la literatura estudios o investigaciones sobre cómo puede afectar la resolución de la imagen en la percepción de la complejidad de un humano. Consecuentemente, se planteó el análisis de la sección 4.3, llegando a la conclusión de que la resolución si afecta, dado que influye en que tantos detalles pueda ser capaz de ver el humano, por lo que se propone cuidar este aspecto en futuros experimentos que se realicen para estimar la complejidad subjetiva de imágenes.

La falta de evidencia concreta acerca de que tanto afecta la resolución de la imagen cuando un humano estima la CI y la falta de bases de datos donde si se tome en cuenta este aspecto se ha optado por el diseño de sistemas que no contemplen la resolución. Por lo que se ha dejado como un posible trabajo futuro puesto que se ha demostrado que si puede impactar en las características para determinar la complejidad objetiva.

CAPÍTULO V. SISTEMAS PARA DETERMINAR LA COMPLEJIDAD

En el capítulo III se analizaron y seleccionaron algunas características computables para determinar la complejidad de imágenes de una forma objetiva. Sin embargo, estas características pueden ser más robustas si se utilizan en conjunto, debido a que algunas de éstas pueden ser mejores donde otras presentan debilidades. Por esto, se ha propuesto el diseño de dos sistemas para determinar la CI, el primero que hace uso de lógica difusa y el segundo de redes neuronales artificiales.

En este capítulo se presenta el preprocesamiento que se les da a las características que alimentan a los sistemas, así como también el diseño de éstos.

5.1 Datos de diseño y validación

Debido a que los sistemas propuestos para determinar la complejidad de imágenes deben ser diseñados o entrenados para que tengan un comportamiento similar a la percepción del humano, se realiza una validación cruzada, la cual es útil para garantizar que los datos utilizados para diseñar o entrenar no estén sesgados y obtener un desempeño confiable. En la Figura 5.1 se muestra el esquema que sigue la validación cruzada, la cual consiste en particionar en η grupos, donde cada uno de los grupos contiene la misma cantidad de datos o una muy similar. Cada uno de estos grupos es utilizado $\eta-1$ veces como diseño y una vez como validación, pero nunca al mismo tiempo como diseño y validación.

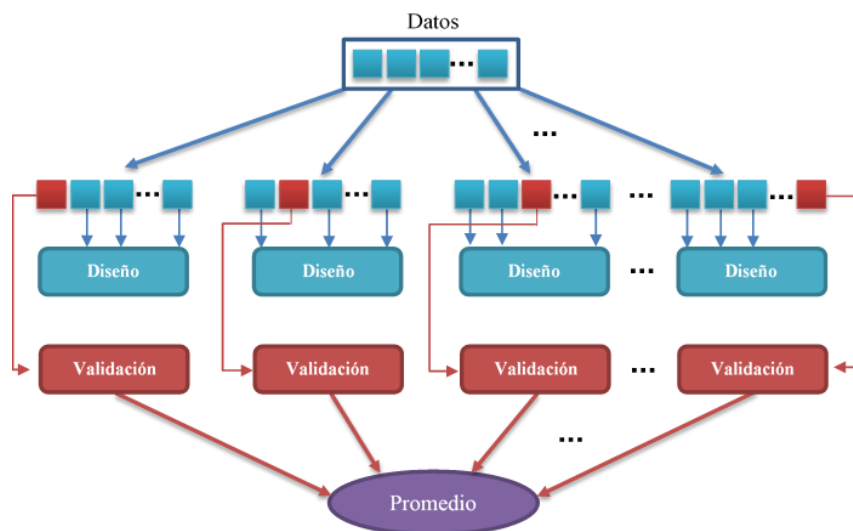


Figura 5.1. Esquema de validación cruzada.

La selección de los grupos de diseño y validación se realiza de la siguiente manera:

1. Las imágenes son ordenadas de menor a mayor complejidad subjetiva (s).
2. Se forman η grupos, donde cada grupo contiene imágenes de todo el rango de complejidad subjetiva. Para lograr esto, se puede ir tomando una imagen cada determinado número de imágenes ya que las imágenes fueron ordenadas de acuerdo con su s en el paso 1, por ejemplo, se busca generar dos grupos, por lo que para el primer grupo se toman las imágenes 1,3,5,... y para el grupo dos se toman las imágenes 2,4,6,...
3. Todos los grupos deben ser utilizados una sola vez como validación, cuando el grupo gr sea utilizado como validación, el resto de los grupos generados se unen para formar el grupo de diseño.

5.2 Preprocesamiento de características

Las características para determinar la complejidad objetiva que serán usadas como alimentación de los sistemas fueron diseñadas para estar dentro de un rango entre 0 y 1, sin embargo, es posible que entre características se encuentren en distintos rangos para las imágenes que se usen para diseño, por ejemplo, si se tiene un grupo de imágenes de diseño con densidad de bordes nivel 1 en un rango entre 0.01 y 0.2, mientras que para densidad de bordes nivel 2 se tiene un rango entre 0.2 y 0.4, esto podría sesgar el diseño del sistema, ya que ambas características son importantes pero una está regresando valores más altos por lo tanto tendrá mayor peso, es por esto que se realiza un escalamiento de 0 a 1 de las características con la Ec. (5.1),

$$\Psi_{\hat{g}} = \frac{\Psi_g - \arg \min(\Psi_g)}{\arg \max(\Psi_g) - \arg \min(\Psi_g)} \quad (5.1)$$

donde $\Psi_{\hat{g}}$ representa a la característica normalizada, g es el índice de la característica, $\arg \min(\Psi_g)$ es el mínimo valor de la característica en el grupo de diseño y $\arg \max(\Psi_g)$ es el máximo valor de la característica en el grupo de diseño. Las características de las imágenes del grupo de validación o de cualquier nueva imagen deben ser procesadas con los mismos parámetros que se utilizaron al normalizar las características de diseño.

Por último, la complejidad subjetiva será utilizada como la salida deseada t_i para el diseño de los sistemas, pero dividida entre 100 para obtener un rango de 0 a 1. Por lo tanto, los objetivos se definen con la Ec. (5.2).

$$t_i = \frac{s_i}{100} \quad (5.2)$$

5.3 Complejidad por lógica difusa

La lógica difusa ha logrado difundirse en diversas áreas de investigación gracias a la flexibilidad que ofrece para modelar y procesar conocimiento humano bajo un esquema matemático formal. El esquema de la lógica difusa permite un mejor modelado que la lógica bivalente del mundo real y sobre todo de sistemas que contengan manejo de incertidumbre y subjetividad [58]. Además esta reduce significativamente el tiempo de desarrollo y la alta complejidad en modelos de sistemas no lineales [8].

La lógica difusa se conforma generalmente por tres etapas: fusificación (se asignan grados de pertenencia a cada variable de entrada según su valores difusos), operaciones en el plano de membresías (aplicación de reglas de inferencia) y defusificación (traducir el resultado del dominio difuso a un valor cuantificable en lógica clásica), este último para sistemas Mamdani [8]. Por otro lado, para sistemas Sugeno la salida es un valor rígido que surge de aplicar cierta función, generalmente mediante un promedio ponderado.

En la Figura 5.2 se muestra un esquema general de un sistema difuso tipo Sugeno para determinar la complejidad de la imagen. La entrada recibe las variables lingüísticas que pueden ser características, después se aplican las reglas de inferencia y por último a la salida se obtiene la CI.

Para el diseño del sistema difuso que determine la CI se propone el diagrama de flujo que se muestra en Figura 5.3, el cual es un proceso donde se realizan varios sistemas difusos, después se selecciona y guarda el que logró el mejor desempeño para ser utilizado.

En las siguientes secciones se describirá más a detalle las partes del proceso propuesto para el diseño del sistema difuso que mide la CI.

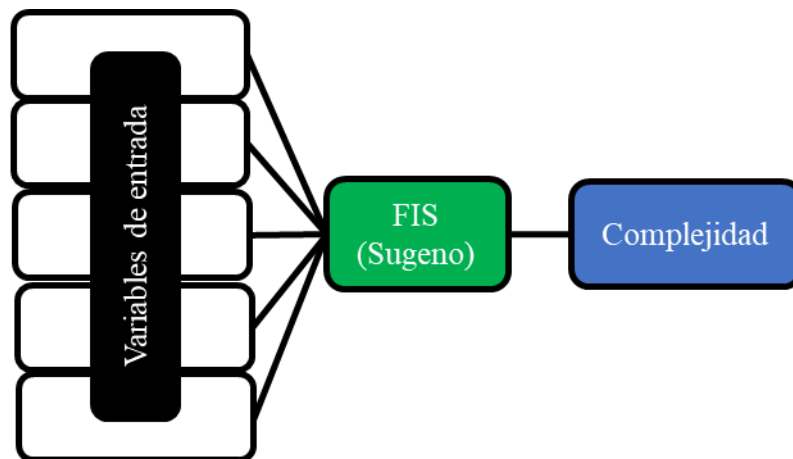


Figura 5.2. Esquema general de un sistema difuso tipo Sugeno para determinar la complejidad de imágenes.

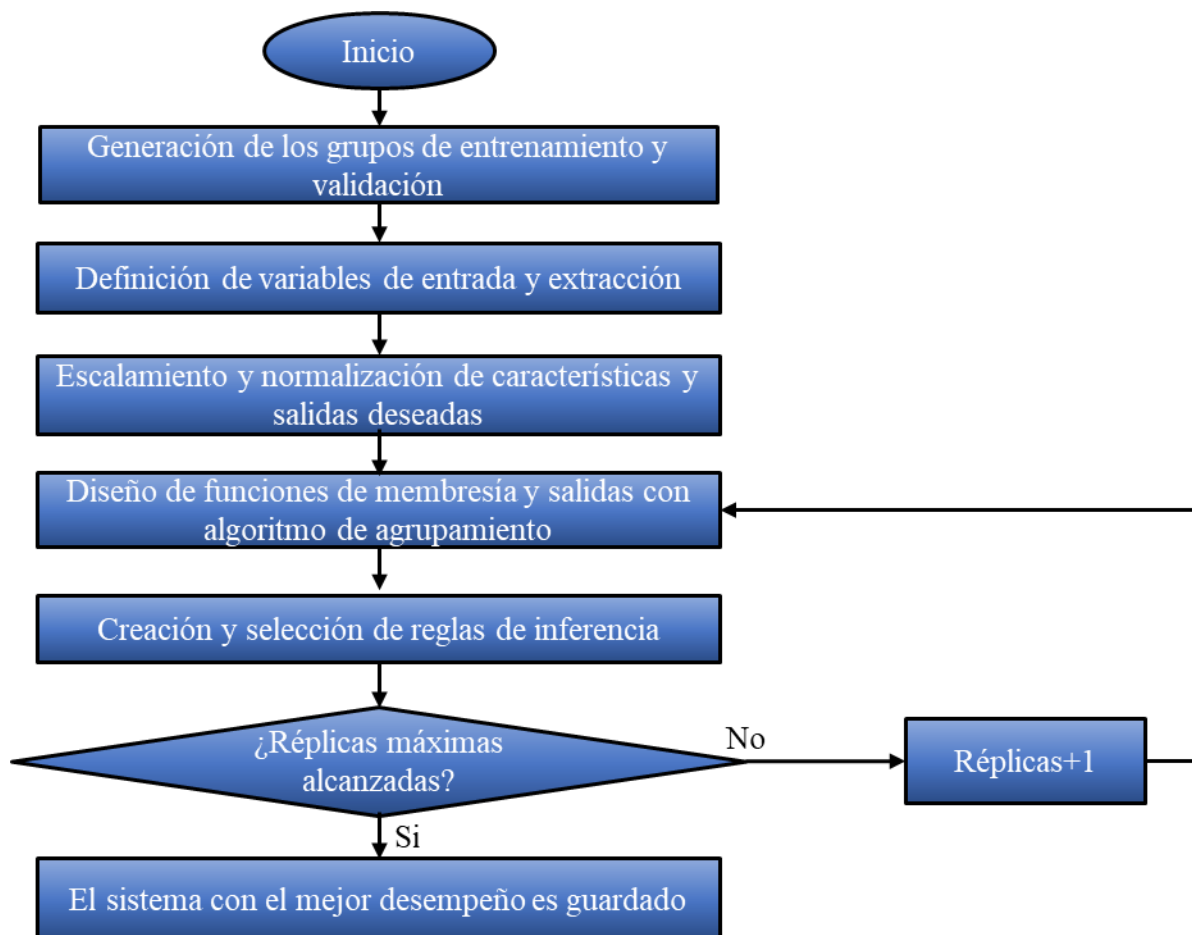


Figura 5.3. Diagrama de flujo del diseño del FIS propuesto.

5.3.1 Variables de entrada del sistema difuso

Los sistemas difusos propuestos reciben como entradas variables lingüísticas obtenidas a través de la Ec. (5.1), como se muestra en la Figura 5.2, la cual normaliza las características vistas en el capítulo III. Dichas variables lingüísticas toman valores difusos, por ejemplo ‘Poco’, ‘Algo’, ‘Mucho’, etc.

5.3.1.1 Generación de funciones de membresía

Las funciones de membresía describen matemáticamente los valores difusos. Dichas funciones pueden ser diseñadas por un experto, una red neuronal, etc. Aquí se propone el uso de un algoritmo de agrupamiento para el diseño de éstas. Los algoritmos seleccionados para el agrupamiento fueron k-means++ (KM), fuzzy c-means (FCM), Gath-Geva (GG) y Gustafson–Kessel (GK). El objetivo del agrupamiento es la clasificación de objetos de acuerdo a similitudes entre ellos de tal forma que cada grupo sea homogéneo y distinto de los demás. Un criterio de homogeneidad simple está basado en la distancia: se espera que la distancia entre los datos de un grupo a su centroide sea significativamente menor que la distancia entre los datos de ese grupo a otros centroides.

Entre los algoritmos de agrupamiento, se seleccionó k-means++ por tener mejores desempeños en el diseño de los sistemas difusos. Estos desempeños se muestran en la Tabla 6.1 del capítulo VI. Por lo que en esta sección solo se describirá el método k-means++, para más información de los otros algoritmos de agrupamiento dirigirse a la sección de anexos.

El método k-means++ se basa en el algoritmo k-means por lo que se describirá primeramente éste. K-means es un algoritmo simple y rápido que intenta mejorar localmente un agrupamiento y funciona de la siguiente manera [59]:

1. Seleccionar k centroides iniciales $C = \{c_1, \dots, c_k\}$ de forma arbitraria.
2. Para cada $i \in \{1, \dots, k\}$, se establece la agrupación A_i para que sea el subconjunto de datos del conjunto de diseño X_d que estén más cerca de c_i que de c_j para todo $j \neq i$.
3. Para cada $i \in \{1, \dots, k\}$, se establece c_i para ser el centro de masa de todos los datos en A_i , donde,

$$c_i = \frac{1}{|A_i|} \sum_{x_d \in A_i} x_d \quad (5.3)$$

4. Repetir los pasos 2 y 3 hasta que C no cambie más.

K-means++ tiene una forma diferente de iniciar los centroides. En cualquier momento, suponga que $D(x_d)$ denota la distancia del dato x_d al centroide más cercano. A continuación, se enlistan los pasos del algoritmo k-means++ [59].

- 1a. Seleccione un dato uniformemente al azar del conjunto de datos, X_d . El dato elegido es el primer centroide, y se denota c_1 .
- 1b. Seleccione el siguiente centro c_i , seleccionando

$$c_i = x_d' \in X_d \quad (5.4)$$

donde x_d' es un nuevo dato de X_d y es seleccionado mediante probabilidad

$$\frac{D(x_d')^2}{\sum_{x_d \in X_d} D(x_d)^2} \quad (5.5)$$

- 1c. Repetir el paso 1b hasta que se han elegido los k centroides.
- 2-4. Proceder con el estándar del algoritmo de k-means.

Cada grupo A_i generado con k-means++ está definido por sus objetos miembros y por su centroide o centro (c). El centroide para cada grupo es el punto en el que se minimiza la suma de las distancias de todos los datos en ese grupo.

Como ya se ha mencionado, la inicialización de los centroides es aleatoria. Sin embargo, como muchos otros tipos de minimización numérica, la solución que alcanza k-means++ a menudo depende de los centroides iniciales. Es posible que k-means++ alcance un mínimo local, donde la reasignación de cualquier punto a un nuevo grupo aumentaría la suma total de la distancia de dato a centroide, pero donde existe una mejor solución (centroides que generan un sistema mejor correlacionado al criterio humano). Para tratar de solucionar este problema se puede hacer uso de réplicas, las cuales realizan de nuevo el algoritmo y como consecuente una posible mejor solución.

De esta manera, se buscan aquellos centroides que mejor representan a los datos, tratando de evitar caer en el problema de mínimos locales, por lo que se hace un proceso iterativo que consiste en realizar varias réplicas para buscar la mejor solución.

Para definir las funciones de membresía μ de los valores difusos, se hace uso del algoritmo de agrupamiento k-means++ sobre un conjunto de diseño, en el cual se busca un grupo para cada valor difuso de cada variable lingüística denotados como A_{gh} , donde g representa a la variable lingüística y h al valor difuso. Después, se analiza cada grupo para brindarles el nombre del valor difuso con base en su centroide c_{gh} , es decir, para el centroide con valor más pequeño recibirá el nombre de ‘Poco’, mientras que para el grupo con el centroide con valor más alto tomará el nombre de ‘Mucho’ y de manera similar para el resto. Cada centroide c_{gh} de cada grupo A_{gh} será la media de una función gaussiana que definirá el valor difuso, mientras que la desviación estándar σ_{gh} del grupo A_{gh} será la desviación estándar de la función de membresía que define a A_{gh} . De tal forma que cada función de membresía μ_{gh} del sistema está definida por la Ec. (5.6).

$$\mu_{gh} = e^{\frac{-(\Psi_g - c_{gh})^2}{2\sigma_{gh}^2}} \quad (5.6)$$

En la Figura 5.4 se muestra un ejemplo de las funciones de membresía que se generan para una variable lingüística, donde se ordenan de izquierda a derecha de menor a mayor, esto es, la línea en azul ilustra el valor difuso ‘Poco’, mientras que la línea verde representa el valor difuso ‘Mucho’ y de manera similar para las demás funciones.

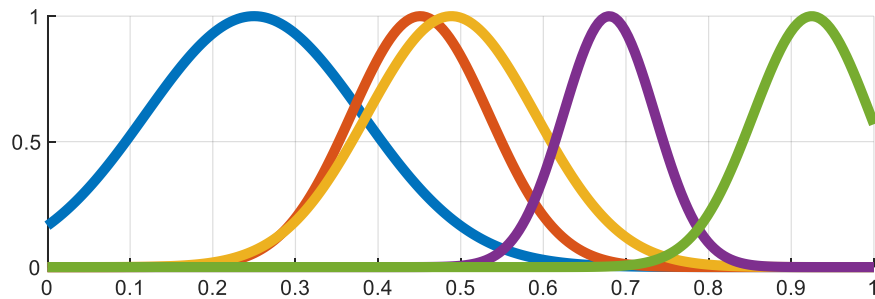


Figura 5.4. Ejemplo de los resultados obtenidos de las funciones de membresía.

5.3.2 Determinación de las reglas de inferencia.

Para definir los valores del consecuente de las reglas, se utiliza nuevamente el algoritmo k-means++ sobre t_i , donde se eligen los κ grupos que mejor representan al conjunto de diseño. Los centroides de los grupos encontrados corresponderán a los posibles consecuentes de las reglas de inferencia.

Una vez que ya se tienen todos los valores difusos se buscan las reglas con las cuales el sistema de inferencia difuso (FIS) operará. Las reglas siguen una estructura típica de un sistema difuso de tipo Sugeno de orden cero, donde el antecedente de la regla es como en los sistemas Mamdani pero el consecuente es una salida constante,

SI *Entrada1* es A_1 **Y** *Entrada2* es $A_2, \dots$, **ENTONCES** la salida es z

El conjunto de reglas (R_n) con el cual el FIS operará se determina mediante un análisis de cada una de las posibles combinaciones, donde se evalúan y se seleccionan solo aquellas que son disparadas al menos por una imagen de diseño, esto con el propósito de reducir la cantidad de las mismas y por consiguiente disminuir el procesamiento. La Ec. (5.7) muestra el criterio de selección,

$$\text{Seleccionamos } R_m \text{ si y solo si } B_m \neq \emptyset \quad (5.7)$$

donde B_m es el conjunto de imágenes con características que hacen disparar a la de la regla m con un nivel distinto que cero.

Para determinar el consecuente z_m de cada regla R_m , se buscan los κ grupos que mejor representen la complejidad subjetiva de las imágenes seleccionadas para diseño, donde el consecuente z_m toma el valor del centroide más cercano al promedio de las complejidades subjetivas de las imágenes de diseño que disparan la regla Ec. (5.8) y Ec. (5.9).

$$\bar{S}_m = \frac{1}{|B_m|} \sum_{t_i \in B_m} t_i \quad (5.8)$$

donde $\bar{S}_m$ es el promedio de complejidades subjetivas de las imágenes que intervienen en la regla m . Seleccionamos el centroide más cercano para la salida de la regla m ,

$$z_m = c_t : c_t \left(\min \left(\bar{S}_m - c_t \right) \right) \quad (5.9)$$

donde c_t son los centroides que mejor representan a las complejidades subjetivas de las imágenes de diseño, generados por el agrupamiento en t_i .

5.3.3 Ejemplo de diseño de un sistema difuso

Para facilitar el entendimiento de este proceso de diseño del sistema difuso se muestra un ejemplo paso a paso; Se desea diseñar un sistema difuso tipo Sugeno de orden cero para calcular la complejidad de imágenes simples. Las especificaciones del sistema son las siguientes; a la entrada recibe dos variables lingüísticas, la primera se tratará de la cantidad de objetos, y la segunda de la cantidad de colores, cada una de estas variables lingüísticas pueden tomar tres valores difusos, Pocos, Más o menos, y Muchos. Por último, como salida del sistema es la complejidad.

Paso 1, adquisición de la base de datos: para este ejemplo se tiene como base de datos un total de 10 imágenes diferentes para el diseño como se muestra en Figura 5.5.

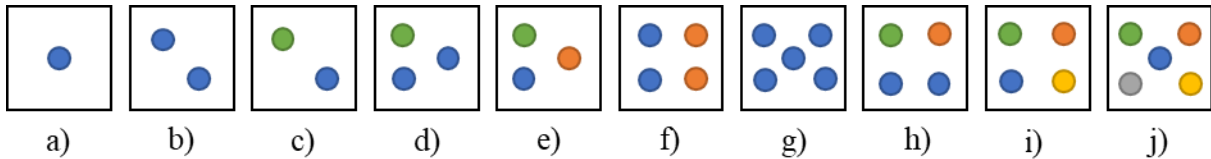


Figura 5.5. Imágenes para el ejemplo de diseño del FIS.

La base de datos de imágenes ya ha sido evaluada por un grupo de personas que estimaron la complejidad de las imágenes con base en su percepción y ya se tiene el promedio de estos como se muestra en la Tabla 5.1.

Tabla 5.1. Complejidad subjetiva para el ejemplo del diseño del FIS.

Imagen	a	b	c	d	e	f	g	h	i	j
s_i	0	12.5	25	37.5	50	50	50	62.5	75	100

Suponer que la base de datos no requiere ningún filtrado de los resultados y ya está completa, por lo que se aplica la Ec. (5.2), sobre los puntajes subjetivos de complejidad para obtener los objetivos o salidas deseadas para el sistema t_i , los resultados de hacer esto se muestran en la Tabla 5.2.

Tabla 5.2. Salida deseada para el ejemplo del diseño del FIS.

Imagen	a	b	c	d	e	f	g	h	i	j
t_i	0.0	0.125	0.25	0.375	0.5	0.5	0.5	0.625	0.75	1.0

Paso 2, extracción de características: Como ya se ha establecido para este ejemplo, las características con las que se estimará la complejidad objetiva son cantidad de objetos (Ψ_o) y cantidad de colores (Ψ_e), suponga que ya se cuenta con los extractores de estas características y los resultados regresados son los que se muestran en la Tabla 5.3.

Tabla 5.3. Características para el ejemplo del diseño del FIS.

Imagen	a	b	c	d	e	f	g	h	i	j
Ψ_o	1	2	2	3	3	4	5	4	4	5
Ψ_e	1	1	2	2	3	2	1	3	4	5

Para garantizar un mejor diseño del sistema y que no tengan mayor peso ciertas características, estas deben ser escaladas a un rango de 0 a 1 por medio de la Ec. (5.1), quedando como se muestra en la Tabla 5.4.

Tabla 5.4. Características objetivas escaladas para el ejemplo del diseño del FIS.

Imagen	a	b	c	d	e	f	g	h	i	j
$\Psi_{\hat{o}}$	0	0.25	0.25	0.5	0.5	0.75	1	0.75	0.75	1
$\Psi_{\hat{e}}$	0	0	0.25	0.25	0.5	0.25	0	0.5	0.75	1

Paso 3, diseño de funciones de membresía: Se aplica un método de agrupamiento sobre los pares $(t_i, \Psi_{\hat{o}})$ para generar las funciones de membresía de la variable lingüística cantidad de objetos, en la Figura 5.6 se muestra el resultado de la generación de los grupos de este ejemplo,

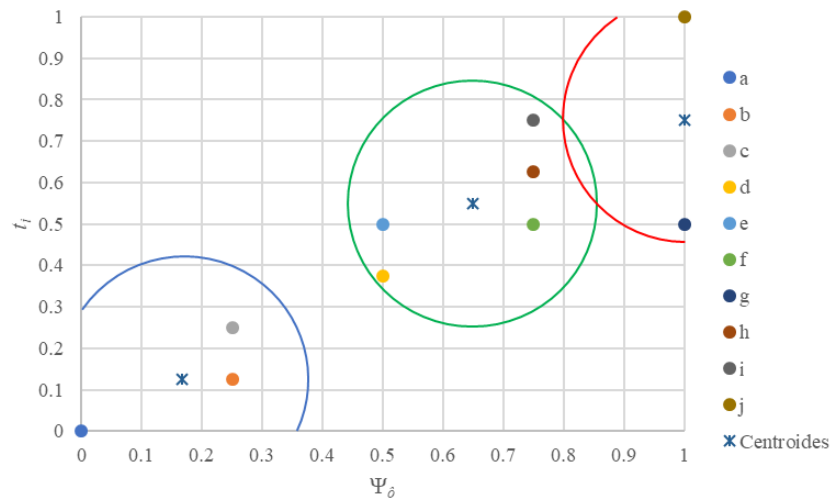


Figura 5.6. Agrupamiento de cantidad de objetos.

Los grupos adquieren el nombre del valor difuso de acuerdo a su centroide, por lo que el grupo del semicírculo azul será para el valor difuso Pocos, el círculo verde representa el valor de Más o menos y el semicírculo rojo el valor de Muchos. Ya con los grupos es posible generar las funciones de membresía gaussianas con la Ec. (5.6), tomando su centroide como la media, y la desviación estándar de las características de las imágenes que pertenecen al conjunto generado como la desviación estándar de la función. Realizamos lo mismo para la segunda variable lingüística de entrada, pero esta vez tomando los pares de (t_i, Ψ_θ) , en la Figura 5.7 se muestra el resultado de la generación de los grupos,

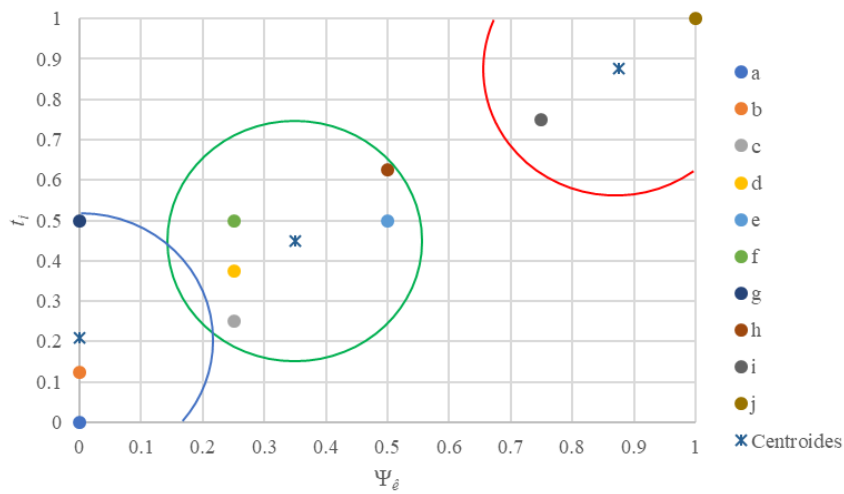


Figura 5.7. Agrupamiento de cantidad de colores.

Paso 4, definición del consecuente: Para los posibles consecuentes de las reglas se aplica algún método de agrupamiento sobre los datos de diseño, más específicamente sobre las salidas deseadas t_i ; para este ejemplo se buscaron 5 posibles consecuentes (5 centroides) como se muestra en la Figura 5.8.

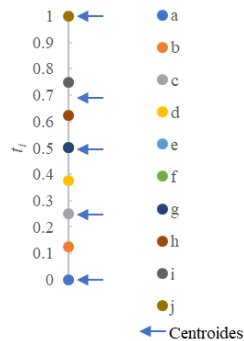


Figura 5.8. Posibles consecuentes de las reglas.

El conjunto de estos centroides corresponde a $c_t = [0 \ 0.25 \ 0.5 \ 0.6875 \ 1]$.

Paso 5, definición de reglas de inferencia: Al tratarse de dos variables lingüísticas con tres valores difusos cada una se pueden tener hasta un total de nueve reglas de inferencias, estas nueve posibles reglas se enlistan a continuación,

1. **SI** Ψ_{δ} es Pocos, **Y** Ψ_{ϵ} es Pocos **ENTONCES** $z = z_1$
2. **SI** Ψ_{δ} es Pocos, **Y** Ψ_{ϵ} es Más o menos **ENTONCES** $z = z_2$
3. **SI** Ψ_{δ} es Pocos, **Y** Ψ_{ϵ} es Muchos **ENTONCES** $z = z_3$
4. **SI** Ψ_{δ} es Más o menos, **Y** Ψ_{ϵ} es Pocos **ENTONCES** $z = z_4$
5. **SI** Ψ_{δ} es Más o menos, **Y** Ψ_{ϵ} es Más o menos **ENTONCES** $z = z_5$
6. **SI** Ψ_{δ} es Más o menos, **Y** Ψ_{ϵ} es Muchos **ENTONCES** $z = z_6$
7. **SI** Ψ_{δ} es Muchos, **Y** Ψ_{ϵ} es Pocos **ENTONCES** $z = z_7$
8. **SI** Ψ_{δ} es Muchos, **Y** Ψ_{ϵ} es Más o menos **ENTONCES** $z = z_8$
9. **SI** Ψ_{δ} es Muchos, **Y** Ψ_{ϵ} es Muchos **ENTONCES** $z = z_9$

Sin embargo, no todas las reglas son necesarias, ya que hay casos en donde ninguna imagen cumple con ellas, por lo que haciendo uso de la Ec. (5.7) seleccionamos solo las necesarias, de tal manera que solo quedan las reglas 1, 2, 5, 6, 7 y 9.

Aplicamos la Ec. (5.8) en cada regla seleccionada para obtener los valores $\bar{S}$, por ejemplo, para el caso de la regla 1, solo las imágenes a y b cumplen con ella por lo que sería $\bar{S} = (0.0+0.125)/2 = 0.0625$. Después mediante la Ec. (5.9) encontramos los valores z correspondientes para cada regla, igual para el caso de la regla 1 sería el c_t más cercano por lo tanto le corresponde 0. La Tabla 5.5 muestra los resultados de realizar todo esto a las reglas seleccionadas.

Tabla 5.5. Imágenes que cumplen con las reglas seleccionadas.

Regla	Imágenes que cumplen la regla	$\bar{S}$	z
1	a, b	0.0625	0
2	c	0.25	0.25
5	d, e, f, h	0.5	0.5
6	i	0.75	0.6875
7	g	0.5	0.5
9	j	1	1.0

En este punto, ya se tienen completas todas las reglas con las que operará el sistema difuso de este ejemplo, la cuales se enlistan a continuación,

1. **SI** Ψ_{δ} es Pocos, **Y** Ψ_{ϵ} es Pocos **ENTONCES** $z = 0$
2. **SI** Ψ_{δ} es Pocos, **Y** Ψ_{ϵ} es Más o menos **ENTONCES** $z = 0.25$
5. **SI** Ψ_{δ} es Más o menos, **Y** Ψ_{ϵ} es Más o menos **ENTONCES** $z = 0.5$
6. **SI** Ψ_{δ} es Más o menos, **Y** Ψ_{ϵ} es Muchos **ENTONCES** $z = 0.6875$
7. **SI** Ψ_{δ} es Muchos, **Y** Ψ_{ϵ} es Pocos **ENTONCES** $z = 0.5$
9. **SI** Ψ_{δ} es Muchos, **Y** Ψ_{ϵ} es Muchos **ENTONCES** $z = 1.0$

5.3.4 Salida del sistema difuso.

Para un sistema difuso tipo Sugeno una regla de inferencia funciona como se muestra en la Figura 5.9.

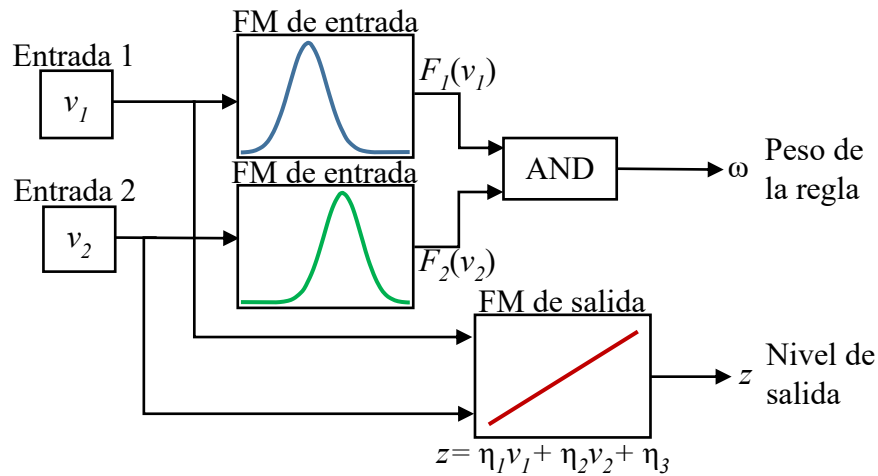


Figura 5.9. Diagrama de una regla de un sistema Sugeno.

donde $F_{1,2}(\cdot)$ son las funciones de pertenencia para las entradas 1 y 2, v_1 y v_2 son los valores de las entradas 1 y 2 respectivamente, η_i son constantes. Sin embargo, al tratarse de un sistema difuso tipo Sugeno de orden cero, los valores de z serán constantes en vez del polinomio. El peso ω de la regla, es obtenido mediante el operador *AND* con *Entrada 1* = v_1 y *Entrada 2* = v_2 , que para el sistema difuso propuesto en esta investigación se trata del operador mínimo,

$$\omega_i = \min(F_1(v_1), F_2(v_2))$$

Por último, la salida final del sistema es el promedio ponderado de todas las salidas de las reglas, Ec. (5.10).

$$\Psi = \frac{\sum_{i=1}^M \omega_i z_i}{\sum_{i=1}^M \omega_i} \quad (5.10)$$

donde M es el número total de reglas del sistema.

Por lo que el sistema difuso propuesto para determinar la complejidad de imágenes por medio de las características densidad de bordes nivel 1 y 2, relación de compresión, medida de complejidad del color y entropía del patrón, queda como se ilustra en la Figura 5.10.

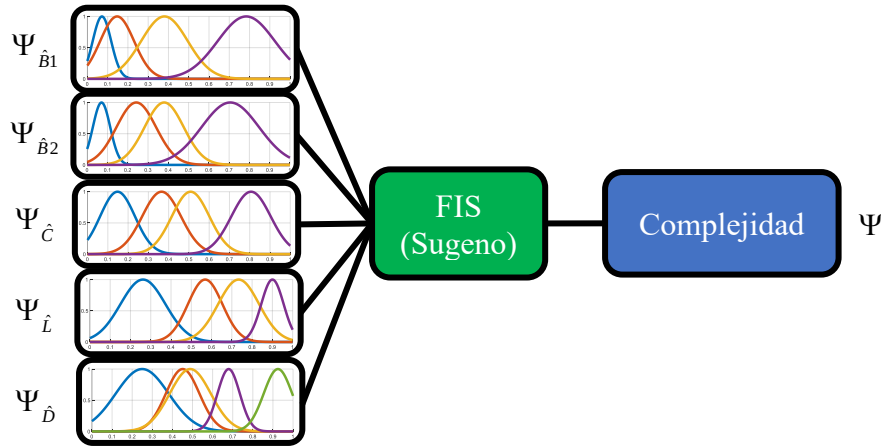


Figura 5.10. Sistema difuso propuesto para determinar la complejidad de imágenes.

5.4 Complejidad con redes neuronales artificiales

Otro sistema que se propone para determinar la CI objetiva es mediante el uso de una red neuronal artificial (RNA). En la siguiente sección se explicará cómo se diseña una para el cálculo de complejidad objetiva de imágenes.

5.4.1 Diseño de la RNA para determinar la complejidad de imágenes

La RNA que se propone en este documento tiene una arquitectura clásica, es decir está compuesta por una capa de entrada, una capa oculta, y una capa de salida, en donde no existen conexiones laterales es decir conexiones entre neuronas de la misma capa y ninguna

retroalimentación, por lo que es una red *feedforward*. La Figura 5.11 se muestra el esquema general de la red neuronal propuesta.

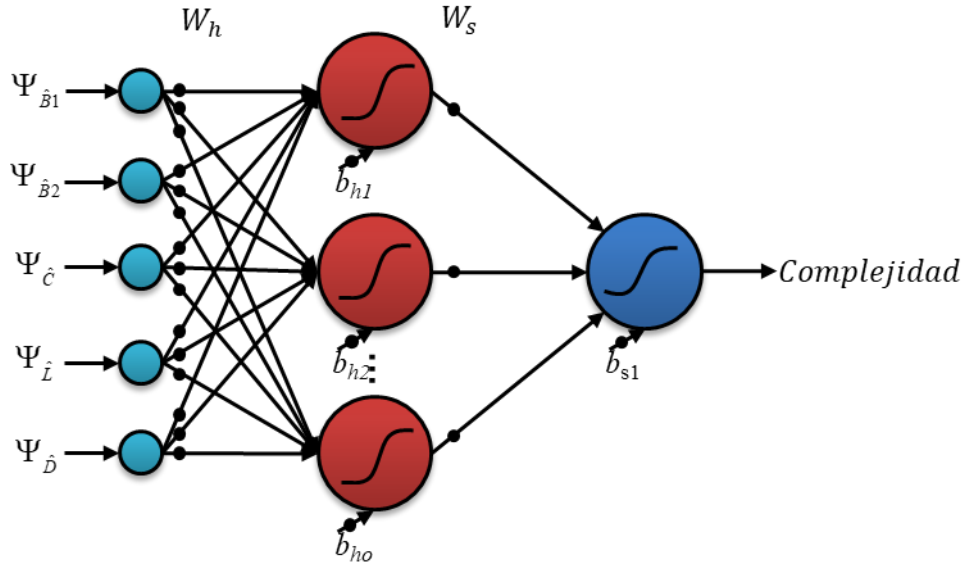


Figura 5.11. Esquema general de la red neuronal propuesta.

En la capa de entrada está el vector de características Ψ_i , para este caso se trata de las características de densidad de bordes nivel 1, densidad de bordes nivel 2, relación de compresión JPEG, medida de complejidad de color y entropía del patrón. Tener en cuenta que la RNA no recibe la imagen como entrada, sino solo las características que se han mencionado anteriormente, por lo que cualquier información que no es captada por las características se pierde.

Después en la capa oculta se encuentran las neuronas las cuales son las principales responsables de realizar la transformación de la entrada en la salida deseada [2]. Donde W_h son los pesos sinápticos, así como b_h son los *bias* de la capa oculta. Se propone el uso de funciones de activación tipo tangente hiperbólica sigmoidea para esta capa Ec. (5.11)

$$\Phi_{hi}(S_{hi}) = \frac{2}{(1 + e^{-2S_{hi}})} - 1 \quad (5.11)$$

donde S_{hi} es la suma ponderada de las entradas a la neurona i de la capa oculta.

En la capa de salida se tiene una única neurona que calcula la complejidad, para este caso con base en las salidas de todas las neuronas de la capa oculta. Donde W_s son los pesos

sinápticos y b_s el *bias* de la capa de salida. Aquí se sugiere el uso de una función de activación del tipo log-sigmoidea [5] Ec. (5.12), ya que la salida se espera que esté en un rango de 0 a 1, donde 0 es para imágenes poco complejas y 1 es para imágenes muy complejas.

$$\Phi_{si}(S_{si}) = \frac{1}{1 + e^{-S_{si}}} \quad (5.12)$$

donde S_{si} es la suma ponderada de las entradas a la neurona i de la capa de salida.

Dado que no existe una fórmula o un método que garantice el número de neuronas en la capa oculta que mejor modelará el problema que se esté tratando con la RNA, se propone un proceso iterativo, donde se varía el número de neuronas en la capa oculta. A continuación, se enlista el proceso propuesto.

1. Generación de los grupos de diseño (entrenamiento) y validación, al tratarse de validación cruzada generalmente se particiona en 10 grupos como se explicó en la sección 5.1.
2. El grupo de diseño es reordenado de manera aleatoria, ya que los grupos se formaron de forma ascendente según su complejidad subjetiva, la razón de esta organización aleatoria es tratar de evitar que el entrenamiento se sesgue.
3. Se realiza el entrenamiento de la RNA, aquí cada vez que se haga este punto es una inicialización distinta de los pesos sinápticos de la RNA.
4. Repetimos los pasos 2 y 3 según réplicas deseadas.
5. Incrementar el número de neuronas de la capa oculta de la RNA, y repetir los pasos 3 y 4. De haber alcanzado el número máximo de neuronas en la capa oculta establecido pasar a 6.
6. Pasar al siguiente grupo de diseño y volver hacer los pasos 2, 3, 4 y 5. De haber terminado toda la validación cruzada pasar al punto 7.
7. Buscar las RNAs con los mejores desempeños y guardar.

En la Figura 5.12 se muestra el diagrama de flujo general propuesto para la búsqueda de la mejor RNA.

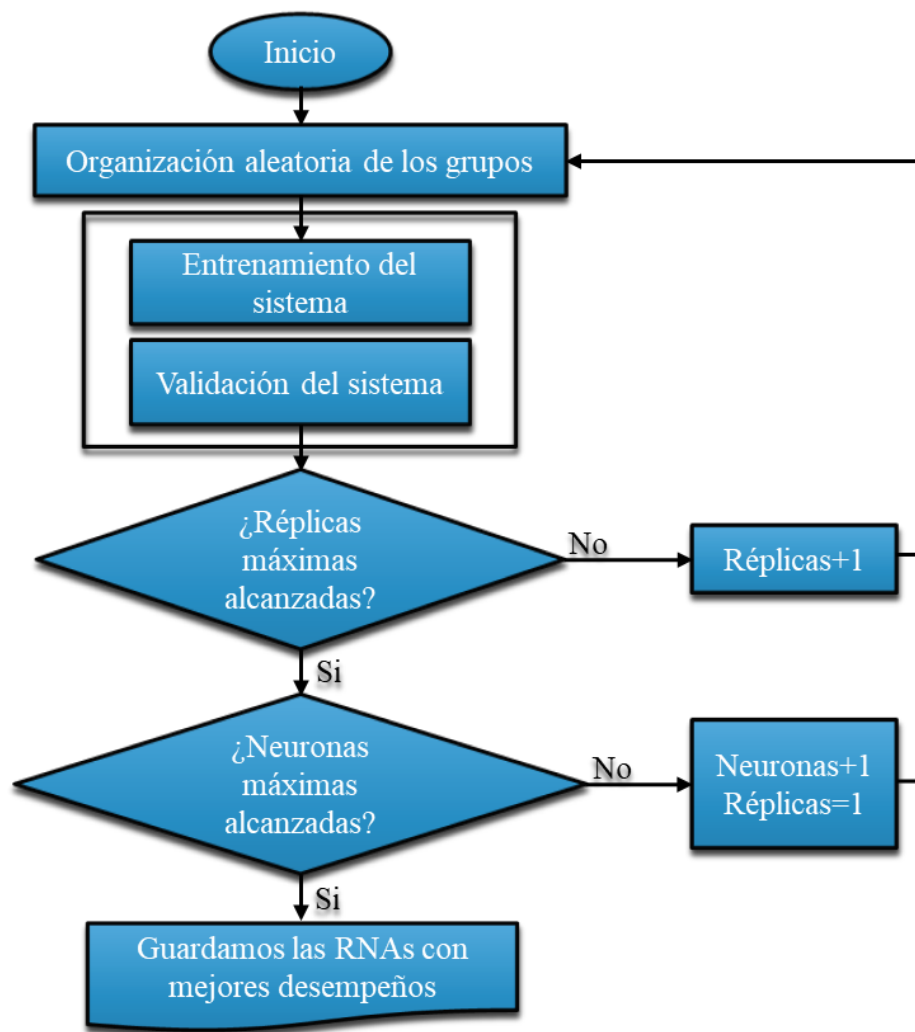


Figura 5.12. Diagrama de flujo para buscar la mejor RNA.

Como se muestra en el diagrama de flujo del proceso de búsqueda de la mejor RNA se propone el realizar réplicas, ya que el medir el desempeño de la primer RNA que se diseñe no es del todo confiable, puesto que para este caso y por lo general la inicialización de los pesos sinápticos de las neuronas es de manera aleatoria, por lo que esto supone que existe la posibilidad de que una inicialización distinta lleve a resultados diferentes.

En la Figura 5.13 se muestra un ejemplo simple de cómo una inicialización aleatoria de los pesos sinápticos puede afectar en el resultado final en la minimización del error en la RNA.

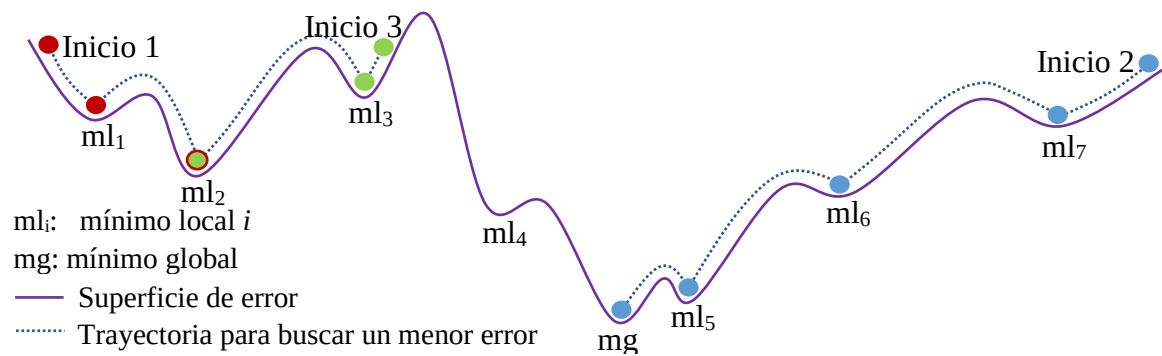


Figura 5.13. Ejemplo de inicializaciones diferentes y su repercusión en la minimización del error de la RNA.

De la Figura 5.13 se puede resaltar como la inicialización 1 y 3 llegaron al mismo mínimo local, sin embargo, en la inicialización 2 se llegó al mínimo global, es importante recordar que en la realidad no se cuenta con la superficie del error, esta es desconocida por lo que la red no es capaz de saber si está en un mínimo local o uno global, es aquí donde se pueden realizar diferentes técnicas para buscar el mínimo global, como lo puede ser el momento, el cual trata de evitar el estancamiento en estos mínimos locales. Sin embargo, como el caso del inicio 3 un momento muy grande no lo llevaría al mínimo global e incluso un momento muy grande puede causar la inestabilidad y no converger en una solución.

Para tratar con el problema descrito anteriormente se propone realizar réplicas, esto es, volver a realizar el entrenamiento desde cero para una nueva RNA, con una inicialización distinta de los pesos sinápticos, pero manteniendo la misma arquitectura.

Para el diseño y entrenamiento de la red se utiliza las funciones que ofrece Matlab `patternnet(hiddenSizes)`, la cual hace uso de un algoritmo derivado del algoritmo de propagación hacia atrás (*backpropagation*), este algoritmo está basado en la idea de la propagación del error hacia atrás. Una explicación general sobre este algoritmo es la siguiente: se presenta un patrón de entrenamiento y se calcula la salida de la RNA. Luego, esta salida se compara con la respuesta deseada, la diferencia entre estos valores es el error, el cual se propaga hacia atrás entre capas de la RNA. Los pesos de las conexiones entre las neuronas de la capa oculta y de salida se ajustan para disminuir el error. En el siguiente paso, los pesos de las conexiones entre las neuronas de la entrada y las capas ocultas también se ajustan para disminuir el error asociado a cada una de las neuronas de la capa oculta.

Durante el entrenamiento, la RNA se expone repetidamente a todos los patrones de entrenamiento hasta que se cumple con alguna de las condiciones siguientes: el máximo número de épocas es alcanzado, el desempeño de la RNA alcanza al desempeño deseado, el desempeño del gradiente está por debajo de lo mínimo o el desempeño en validación ha incrementado más veces de lo predefinido desde la última vez en que decrementó. Este proceso implica que la RNA se ajustará progresivamente con el fin de reducir el error.

Por último los parámetros utilizados fueron los siguientes: máximo número de épocas 1000, desempeño por buscar de 0.999, y como réplicas máximas un total de 100 ya que con esta cantidad se espera que la aleatoriedad de la inicialización de los pesos de las neuronas no sea un problema grave, es decir, es menos probable que la red se estanque en un mínimo local con un error alto, por lo que realizar varias réplicas puede ayudar en la búsqueda de una mejor solución. En cuanto a neuronas máximas se estableció en 100 también, ya que más de 100 neuronas supondría más de una neurona por dato para algunas bases de datos como la de RS1 y RS2 que tiene 98 imágenes en total.

En la Figura 5.14 se muestra como varía el desempeño en términos de correlación de la salida de la red contra la salida esperada al realizar varias réplicas para una RNA con 77 neuronas para el grupo de entrenamiento número 5, como se puede observar si se deja guiar por el primer desempeño se hubiera tenido un desempeño menor del 0.65 para entrenamiento, el cual es muy bajo comparado con lo deseado.

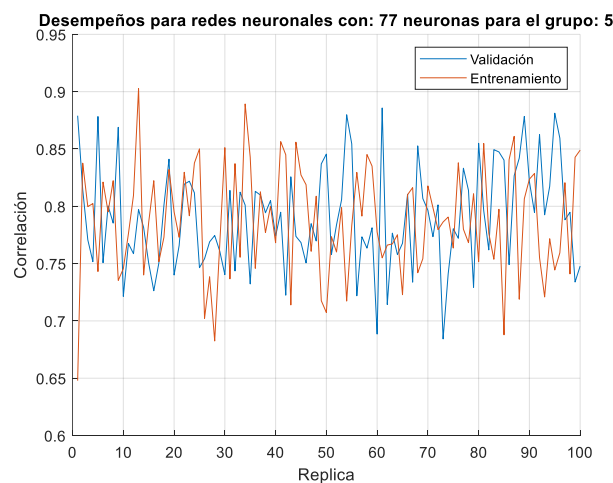


Figura 5.14. Correlación de diferentes réplicas de la RNA de 77 neuronas en capa oculta para el grupo 5.

Por otro lado, en la Figura 5.15 se muestra como varía el desempeño promedio en el grupo 5 al ir cambiando el número de neuronas en la capa oculta de la RNA. Se puede observar que con menos de 10 neuronas los desempeños son muy bajos, mientras que con más de 50 los desempeños tienen una tendencia a la baja, por lo que para este ejemplo lo mejores desempeños están en un rango de 15 a 40 neuronas o más específicamente con 24 neuronas se encontró el mejor desempeño.

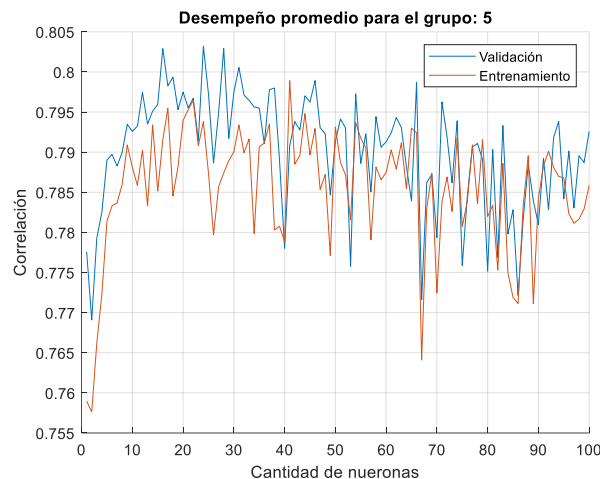


Figura 5.15. Desempeño promedio de validación y entrenamiento para redes neuronales con diferente cantidad de neuronas en la capa oculta.

Tener en cuenta que este proceso iterativo para buscar la mejor RNA se realizó para cada una de las diferentes categorías de imágenes de las bases de datos seleccionadas, ya que se trata de imágenes distintas por lo que es lógico imaginar que al tratarse de diferentes categorías las RNAs puedan necesitar diferentes números de neuronas en la capa oculta: para RS1 y RS2 el mejor desempeño se encontró con 97 neuronas, para escenas 90, para anuncios 24, para visualizaciones 77, para objetos 96, para diseño de interiores 85, para arte 96 y para suprematismo 84.

CAPÍTULO VI. RESULTADOS Y CONCLUSIONES

En este capítulo se describen los desempeños obtenidos en términos de correlación de Pearson entre los resultados de los sistemas propuestos y el criterio subjetivo del humano. Como ya se ha mencionado en el capítulo V se generaron varios grupos de diseño y validación para seguir el esquema de validación cruzada, por lo que los desempeños mostrados aquí son los promedios de dicha validación cruzada. También en esta sección se realiza la comparación contra algunos resultados reportados en la literatura. Por último, se dan a conocer las conclusiones a las que se llegó.

6.1 Selección del algoritmo de agrupamiento para el diseño del FIS

Antes de presentar los desempeños logrados por los sistemas difusos se expone por qué se eligió k-means++ sobre los demás algoritmos de agrupamiento para el diseño de las funciones de membresía del sistema difuso.

El experimento para la selección de un algoritmo de agrupamiento consistió en realizar diferentes sistemas difusos, utilizando distintos algoritmos de agrupamiento y variando las cantidades de grupos (valores difusos), en la Tabla 6.1 se muestra un resumen de los resultados de dicho experimento.

Tabla 6.1. Desempeños de los sistemas difusos al utilizar diferentes algoritmos de agrupamiento en las bases de datos RS1 y RS2.

No. grupos	KM		FCM		GG		GK	
	Validación	Diseño	Validación	Diseño	Validación	Diseño	Validación	Diseño
2	0.7433	0.7275	0.7029	0.7244	0.6462	0.7274	0.7138	0.7120
3	0.7923	0.7964	0.7531	0.7947	0.6872	0.7732	0.7824	0.7449
4	0.7895	0.8053	0.7545	0.8147	0.7444	0.7937	0.7910	0.7729
5	0.8035	0.8290	0.7320	0.8087	0.7314	0.8147	0.7996	0.7660
6	0.8230	0.8432	0.7688	0.8396	0.7012	0.8386	0.7913	0.8047
7	0.8200	0.8849	0.7782	0.8643	0.7354	0.8594	0.7262	0.8251
8	0.8145	0.8866	0.7764	0.8970	0.6960	0.8797	0.8197	0.8505
9	0.8179	0.9138	0.7633	0.9066	0.7254	0.8871	0.8222	0.8349
10	0.8271	0.9173	0.7833	0.9095	0.7127	0.8947	0.8134	0.8499
Promedio	0.8035	0.8449	0.7569	0.8399	0.7089	0.8298	0.7844	0.7957

De la Tabla 6.1 se puede resaltar que k-means++ es el que logra mejores desempeños (remarcados en negritas), también se puede apreciar que existe una tendencia de ir mejorando el desempeño conforme la cantidad de grupos va incrementando, sin embargo, solo se dejó hasta un máximo de 10 grupos, ya que el ir incrementando el número de grupos supone un aumento exponencial en el tiempo de análisis y selección de las reglas de inferencia del sistema difuso, lo cual para más de 10 grupos lo hace un proceso muy pesado que conlleva mucho tiempo de procesamiento y posiblemente un incremento muy leve en el desempeño de acuerdo a las tendencias mostradas.

6.2 Correlación entre características propuestas

Las características son un punto muy importante para determinar la CI de manera objetiva, sin embargo, puede existir información redundante entre éstas y no aportar demasiado a los sistemas. Se realizó un análisis para determinar cuáles de estas características contienen información redundante, con el fin de generar nuevos sistemas que descarten aquellas características que aporten menos y con esto mejorar el tiempo de procesamiento sin comprometer demasiado el desempeño.

Para determinar cuáles características contienen más información redundante se extrajeron éstas de las imágenes de las bases de datos seleccionadas (RS1, RS2 y SAVOIAS) y después se obtuvo la correlación entre pares de características, la Tabla 6.2 muestra la correlación entre pares de características obtenidas con el experimento previamente descrito.

Tabla 6.2. Correlación entre las características seleccionadas, utilizando las bases de datos RS1, RS2 y SAVOIAS.

	DBordes1	DBordes2	RC_JPEG	CCM
DBordes2	0.6790			
RC_JPEG	0.6971	0.8652		
CCM	0.6297	0.8645	0.9306	
Entropía	0.5661	0.6965	0.6771	0.7239

Como se puede apreciar en la Tabla 6.2, los pares de características que están más correlacionadas y que por lo tanto contienen información más redundante entre ellas son:

RC_JPEG-CCM, RC_JPEG-DBordes2, CCM-DBordes2 y CCM-Entropía (resaltadas en negritas).

El objetivo es descartar alguna o algunas de estas características para generar nuevos sistemas que hagan uso de menos características, lo cual implica que tendrá mejores tiempos de procesamiento en la extracción de características. A continuación, se enlistan los sistemas que se diseñaron tomando como criterio principal ir descartando aquellas características más correlacionadas:

5C: Sistemas que incorporan las 5 características, estos sistemas no descartan ninguna característica ya que es importante realizar la comparación entre otros sistemas con menos características y estos, con el objetivo de ver si el reducir el número de características puede mejorar el tiempo sin afectar demasiado en el desempeño.

4C: Sistemas que hacen uso de las siguientes 4 características; densidad de bordes nivel 1, densidad de bordes nivel 2, medida de complejidad de color, y entropía del patrón. Aquí se ha descartado la relación de compresión JPEG por ser de las más correlacionadas con otras características.

3C: Sistemas que hacen uso de las siguientes 3 características; densidad de bordes nivel 1, densidad de bordes nivel 2, y entropía del patrón. Al igual que para 4C aquí se descarta también relación de compresión JPEG por las mismas razones, además también se descarta medida de complejidad de color por tener una correlación elevada con otras características y además ser de las más lentas en procesamiento, más adelante en la sección 6.4 se muestra.

2C: Sistemas que hacen uso de las siguientes 2 características; densidad de bordes nivel 1 y densidad de bordes nivel 2. En estos sistemas solo se mantiene densidad de bordes por ser de las características menos correlacionadas y tratarse de dos características que se sacan en el mismo proceso.

6.3 Resultados de comparación contra la literatura

Al tratarse de diferentes trabajos con diferentes bases de datos se realizó primero una comparación directa contra los resultados que se reportan para las bases de datos que se utilizaron para el diseño e implementación de los sistemas propuestos. En la Tabla 6.3 se muestra una comparación entre el desempeño en validación de los sistemas en términos de

correlación contra los resultados reportados por el artículo [31] de la base de datos RS1 y RS2, donde hacen uso de la técnica de optimización por enjambre de partículas (PSO).

Tabla 6.3. Correlación promedio para las bases de datos RS1 y RS2.

Métodos propuestos								Artículo [31]	Métodos contra los que se evaluó el artículo [31]	
Lógica difusa				Redes neuronales				PSO	FC	SE
5C	4C	3C	2C	5C	4C	3C	2C			
0.85	0.85	0.83	0.80	0.92	0.93	0.92	0.88	0.84	0.75	0.65

De la Tabla 6.3 se resalta en negritas aquellos desempeños que se encuentran por encima del método que se propuso en [31], se puede ver que tanto con lógica difusa como con redes neuronales se lograron mejores desempeños utilizando 5 o 4 características, sin embargo al utilizar menos de 4 características con el sistema de lógica difusa este cae por debajo del método de [31], mientras que redes neuronales aún con menos de 4 características logró mantenerse por encima.

Donde:

Particle Swarm Optimization (PSO): PSO permite optimizar un problema a partir de una población de soluciones candidatas, denotadas como "partículas", moviendo éstas por todo el espacio de búsqueda según reglas matemáticas que tienen en cuenta la posición y la velocidad de las partículas. El movimiento de cada partícula se ve influido por su mejor posición local hallada hasta el momento, así como por las mejores posiciones globales encontradas por otras partículas a medida que recorren el espacio de búsqueda. El fundamento teórico de esto es hacer que la nube de partículas converja rápidamente hacia las mejores soluciones.

Feature Congestion (FC): Tres mapas de desorden para la imagen, que representan el color, la textura y la congestión de orientación, se evalúan en diferentes escalas y se combinan adecuadamente para obtener una sola medida.

Subband Entropy (SE): Se relaciona con el número de bits necesarios para la codificación de imágenes de subbanda. Después de descomponer la luminancia y los canales de crominancia en subbandas wavelet, la entropía se calcula dentro de cada banda y se propone una suma ponderada de estas entropías como medida de saturación.

Los autores de SAVOIAS [40] no proponen ningún método nuevo para determinar la CI de manera objetiva, sino simplemente hacen una comparación contra ciertas características, las cuales ellos consideran las del estado del arte para determinar la CI. Por lo que no es posible comparar directamente los sistemas, así que se realizó una comparación similar a la que ellos presentan, que es ver que tan bien se correlacionan las características para determinar la complejidad objetiva contra la complejidad subjetiva.

En la Tabla 6.4 se muestra la correlación que se obtuvo entre las características seleccionadas en esta investigación y la complejidad subjetiva de la base de datos de [40], donde se han resaltado en negritas aquellos promedios más altos que los de [40].

Tabla 6.4. Correlación de las características seleccionadas contra diferentes tipos de imágenes de la base de datos SAVOIAS.

	Características					
Clase de imagen	DBordes1	DBordes2	RC_JPEG	CCM	Entropía	Promedio
Escenas	0.48	0.28	0.24	0.38	0.48	0.37
Anuncios	0.51	0.55	0.60	0.64	0.58	0.58
Visualizaciones	0.48	0.54	0.51	0.47	0.46	0.49
Objetos	0.44	0.30	0.20	0.18	0.38	0.30
Diseño de interiores	0.57	0.55	0.56	0.65	0.50	0.57
Arte	0.19	0.57	0.50	0.61	0.56	0.48
Suprematismo	0.27	0.50	0.53	0.71	0.77	0.56
Promedio	0.42	0.47	0.45	0.52	0.53	0.48

En la Tabla 6.5 se muestra la correlación que obtuvo [40] con las características seleccionadas como las del estado del arte contra sus valores de complejidad subjetiva, al igual que para la Tabla 6.4 se han resaltado en negritas aquellos promedio más altos.

De la Tabla 6.4 y Tabla 6.5 se puede observar que la mayoría de los mejores promedios se lograron con las características que se seleccionaron en esta investigación, lo cual es indicativo de que se realizó una mejor selección de características, por otro lado, número de regiones fue la que mejor correlación logro en promedio, sin embargo es una característica que por lo general implica mucho procesamiento computacional, más adelante en la sección 6.4 se muestra como

ésta es muy pesada computacionalmente en comparación con las seleccionadas en esta investigación.

Tabla 6.5. Correlación de las características del estado del arte según [40] en su base de datos.

Clase de imagen	Características					
	Bordes	Compresión	N. Regiones	FC	SE	Promedio
Escenas	0.16	0.3	0.57	0.42	0.16	0.32
Anuncios	0.54	0.56	0.41	0.56	0.54	0.52
Visualizaciones	0.57	0.55	0.38	0.52	0.61	0.53
Objetos	0.28	0.16	0.29	0.30	0.10	0.23
Diseño de interiores	0.61	0.68	0.67	0.58	0.31	0.57
Arte	0.48	0.51	0.65	0.22	0.33	0.44
Suprematismo	0.18	0.60	0.84	0.48	0.39	0.50
Promedio	0.40	0.48	0.54	0.44	0.35	0.44

Como ya se ha mencionado en [40] no se propuso ningún método nuevo para determinar la complejidad de imágenes, sin embargo si presentan algunas características para determinarla, por lo que se realizó una comparación entre los sistemas propuestos en esta investigación contra los mejores resultados de las características propuestas en [40]. La Tabla 6.6 muestra dicha comparación entre desempeños logrados con los sistemas propuestos y los mejores desempeños de las características (MC) seleccionadas por [40] para la base de datos SAVOIAS.

Tabla 6.6. Comparación entre resultados de correlación con los métodos propuesto y con los mejores resultados de [40] en la base de datos SAVOIAS.

Clase de imagen	Métodos propuestos								[40]
	Lógica difusa				Redes neuronales				MC
	5C	4C	3C	2C	5C	4C	3C	2C	
Escenas	0.69	0.67	0.66	0.66	0.87	0.82	0.82	0.79	0.57
Anuncios	0.75	0.75	0.72	0.69	0.82	0.82	0.82	0.79	0.56
Visualizaciones	0.68	0.69	0.69	0.67	0.84	0.84	0.83	0.79	0.61
Objetos	0.70	0.69	0.70	0.67	0.80	0.79	0.78	0.77	0.30
Diseño de interiores	0.83	0.82	0.80	0.74	0.91	0.92	0.93	0.89	0.68
Arte	0.68	0.68	0.66	0.62	0.81	0.77	0.76	0.72	0.65
Suprematismo	0.94	0.93	0.91	0.78	0.97	0.97	0.96	0.92	0.84
Promedio	0.75	0.75	0.73	0.69	0.86	0.85	0.84	0.81	0.60

Como se puede observar en la Tabla 6.6 la mayoría de los sistemas propuestos tienen mejores desempeños. Solamente en los sistemas difusos de dos características para arte y suprematismo se está por debajo de [40], sin embargo, la diferencia es mínima de no más de 0.06, en comparación con la diferencia que se alcanza en sistemas como la RNA para objetos, donde se está por encima de [40] en un 0.5.

De las Tabla 6.3 y Tabla 6.6 se puede concluir lo siguiente: utilizar 4 características mantuvo el desempeño logrado con 5. Utilizar 3 características el desempeño casi se mantiene, no baja más de 0.02. Mientras que utilizar solamente 2 características el desempeño si se ve afectado bajando hasta en un 0.06. Sin embargo, más adelante en la sección 6.4 se muestra un análisis de tiempos de procesamiento, el cual puede ayudar a decidir si la ganancia en tiempo es mejor que lo que se pierde en desempeño.

Por último, se realizó una comparación contra toda la literatura investigada en aquellos artículos que si reportan una correlación entre la complejidad objetiva y la complejidad subjetiva, para esto se analizaron los siguientes puntos:

- 1- Correlación promedio entre complejidad objetiva y complejidad subjetiva.
- 2- Cantidad de imágenes utilizadas en el diseño y validación de los sistemas.
- 3- Cantidad de características utilizadas para determinar la complejidad objetiva.
- 4- Cantidad y tipo de categorías de las imágenes utilizadas.

La Tabla 6.7 muestra el resumen de los resultados promedios alcanzados por los sistemas propuestos (5C) y los resultados reportados por la literatura, en donde se ha resaltado en negritas los mejores tres desempeños en términos de correlación, las tres cantidades más grandes en el uso de imágenes, los tres que utilizan menos características y los tres que hacen uso de más categorías de imágenes.

Donde:

LD: Lógica difusa.

RNA: Red neuronal artificial.

PSO: Optimización por enjambre de partículas.

MC: Mejor desempeño de las características propuestas.

GIF: Algoritmo de compresión GIF (Formato de Intercambio de Gráficos)

BDM: Método de descomposición en bloque.

RF: Clasificador *Random Forest*.

VCD: Variedad de color y su distribución.

Tabla 6.7. Comparación entre sistemas propuestos y algunos de la literatura.

Método	Correlación promedio	No. de imágenes	No. de características	Categorías de las imágenes
Propuesto LD	0.80	1518	5	8 categorías a color
Propuesto RNA	0.89	1518	5	8 categorías a color
RNA [2]	0.83	800	329	5 categorías a color
LD [3]	‘Algo’	150	3	Imágenes entornos reales
RNA [5]	‘Consistente’	600	5	-
VCD [28]	‘No consistente’	-	2	Imágenes entornos reales a color
RF [30]	0.78	500	29	Imágenes de pinturas
PSO [31]	0.83	98	11	Entornos reales a color
MC [40]	0.60	1420	1	7 categorías a color
GIF [45]	0.95	180	1	Patrones en blanco y negro
RF [50]	0.91	1164	4	Patrones en blanco y negro
LD [60]	0.73	-	-	Arte a color
BDM [61]	0.53	150	-	Patrones en blanco y negro

De la Tabla 6.7 se puede resaltar que los desempeños logrados por los sistemas propuestos para determinar la complejidad de manera objetiva es competente, ya que solo se está por debajo de dos artículos [45] y [50], sin embargo, estos artículos que se encuentran por encima de los sistemas propuestos hacen uso de menos imágenes para el diseño y validación. Además, hacen uso de imágenes de patrones en blanco y negro, tipo tablero de ajedrez, por lo que podría ser de esperarse estos resultados, ya que se trata de imágenes más simples donde los sistemas se adaptan más fácilmente. También se puede observar que si se realiza una comparación por tipo de método utilizado los sistemas propuestos tienen mejores desempeños, por ejemplo para el sistema de RNA se está por encima en un 0.06 sobre [2], mientras que para el sistema de

lógica difusa se está por encima en un 0.07 sobre [60], que son los sistemas que si proporcionaron una correlación numérica.

6.4 Tiempos de procesamiento de los sistemas

El tiempo de procesamiento de los sistemas también puede ser de interés, por lo que aquí se calcularon los tiempos de procesamiento de la siguiente manera; se implementó un programa cíclico, el cual era alimentado por imágenes RGB de 1MP (1,000,000 pixeles), el programa se ciclaba un total de 100 veces y después se obtenía el tiempo promedio que le tomaba por cada una. En cada ciclo se extraen las características y se calcula la complejidad con alguno de los sistemas. Los sistemas utilizados para este experimento fueron los diseñados para RS1 y RS2 los cuales se encuentran entre el promedio ya sea en número de neuronas en la capa oculta para RNA, como en cantidad de reglas de inferencia para el sistema difuso. Se tomaron tiempos por separados, es decir tiempo en extracción de cada una de las características y los tiempos de propagación de los sistemas.

En las Tabla 6.8 se muestran los resultados en tiempo (segundos). Tener en cuenta que los resultados del tiempo se tomaron desde una computadora que cuenta con lo siguiente: procesador Intel Xeon E5-1620 @3.5 GHz, 16 GB de RAM @ 2133 MHz, GPU NVIDIA Quadro K4200 4 GB GDDR5, HDD.

Tabla 6.8. Tiempos de extracción de características y propagación de sistemas propuestos.

Características y sistemas	Tiempo (s)
DBordes 1 y 2	0.1414
RC_JPEG	0.0383
CCM	1.1990
Entropía	0.1444
FIS	0.0278
RNA	0.0149

En cuanto a los tiempos de procesamiento de los sistemas de la literatura no se reportan, por lo que solo se hizo la comparación contra [31] ya que los autores de este artículo si proporcionaron los programas necesarios para realizar el mismo experimento para la medición de tiempos lo cual se muestra en la Tabla 6.9.

Tabla 6.9. Tiempos de extracción de las características y propagación del sistema de [31].

Características y sistemas	Tiempo (s)
Contraste	0.1506
Correlación	0.1468
Energía	0.1476
Homogeneidad	0.1462
Factor de frecuencia	0.0670
Densidad de bordes	0.0554
Relación de compresión	0.0611
Número de regiones	8.2142
Colorido	0.0403
Número de colores	0.1384
Armonía del color	0.1322
PSO	0.0002

Como se puede observar de las Tabla 6.8 y Tabla 6.9 los tiempos en promedio de extracción de características tanto de las propuestas como de la literatura son muy similares, alrededor de 0.14 segundos por cada característica, a excepción de CCM que toma más de un segundo y número de regiones que se encuentra por encima de los ocho segundos, de esta última ya se había mencionado anteriormente, ya que presenta muy buena correlación con la CI subjetiva pero el procesamiento es muy pesado, y el tiempo es demasiado elevado en comparación con las demás características.

En la Figura 6.1 se muestra una gráfica para comparar los tiempos totales de los sistemas, tanto extracción de características como la propagación, la cual es un resumen de las Tabla 6.8 y Tabla 6.9.

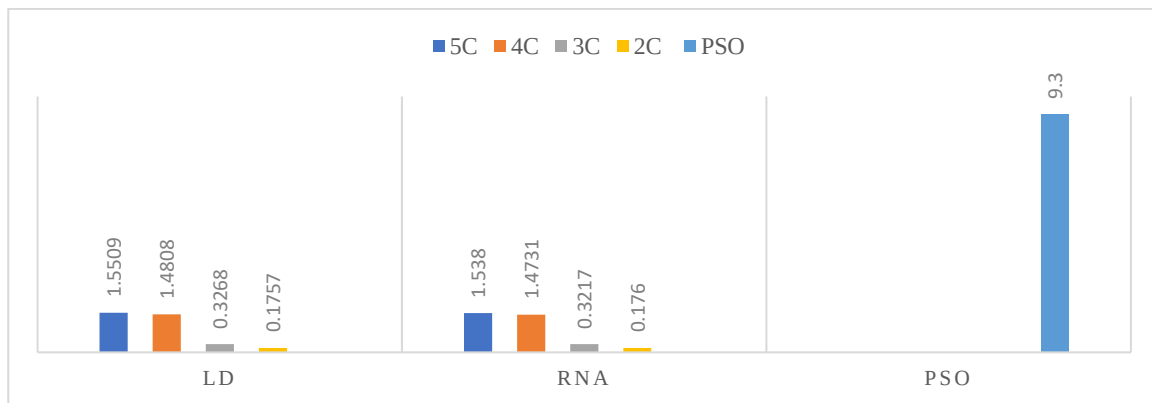


Figura 6.1. Tiempo de procesamiento de los sistemas propuestos y PSO de [31].

De la Figura 6.1 se puede resaltar que los sistemas propuestos son más rápidos, ya que el sistema más lento propuesto es el de lógica difusa para 5 características y este tiene un tiempo promedio de 1.5509 segundos mientras que el de [31] le toma en promedio 9.3 segundos lo cual es aproximadamente 6 veces más lento, o bien hasta casi 53 veces más lento que la RNA para dos características, el cual es el sistema más rápido y que mantiene un desempeño aun superior al de [31].

6.5 Conclusiones

Como se ha tratado a lo largo de esta investigación la complejidad de imágenes es una medida subjetiva cuando un humano la evalúa, ya que para esto hay que pasar por el proceso de percepción visual en el cual intervienen muchos mecanismos sumamente complejos y que pueden ser afectados por conocimientos, cultura, experiencias, etc. Sin embargo, aquí se demostró que la complejidad de las imágenes si puede ser evaluada por un sistema de manera objetiva por medio de características que se correlacionan con la complejidad subjetiva.

Las características que se seleccionaron lograron mejores desempeños promedios en correlación contra la complejidad subjetiva, que aquellas características utilizadas en otros artículos como se muestra en las Tabla 6.4 y Tabla 6.5.

Los sistemas propuestos para determinar la complejidad de imágenes tienen desempeños competentes de hasta 0.89 para redes neuronales y de 0.80 con lógica difusa (Tabla 6.7), solo por debajo de 2 artículos [45] y [50] que logran desempeños de 0.95 y 0.91 respectivamente. Sin embargo, esos dos artículos que se encuentran por encima del desempeño logrado en esta investigación fueron de sistemas que determinan la complejidad a imágenes de patrones en blanco y negro, tipo tablero de ajedrez, por lo que es de esperar al tratarse de imágenes más simples que con los que se trabajó en esta investigación.

Si se compara contra métodos que hacen uso de imágenes de categorías más similares a las utilizadas, los sistemas propuestos están por encima, para RNA se está por encima en un 0.06 contra [2] y para lógica difusa en un 0.07 contra [60], que son los artículos más similares que si reportan la correlación de forma numérica.

En cuanto a los tiempos de procesamiento se demostró que los sistemas propuesto logran tiempos promedios por debajo de [31], que fue el único que si proporcionó el código para poder

realizar una evaluación justa, bajo las misma condiciones. Logrando tiempos promedios hasta 6 veces más rápido si se utilizan los sistemas que hacen uso de todas las características o de hasta 53 veces más rápido si se compara con el sistema más veloz diseñado que mantiene un mejor desempeño en correlación que [31], como se muestra en la Figura 6.1.

Por lo tanto queda a criterio de la aplicación en la que se vaya utilizar cual sistema utilizar, puesto que los sistemas que tienen los mejores desempeños en correlación con la complejidad subjetiva son los más lentos, sin embargo, quitando algunas características se puede reducir considerablemente el tiempo de procesamiento hasta 8 veces (Figura 6.1) sin afectar mucho el desempeño en correlación, en no más de 0.06 (Tabla 6.3 y Tabla 6.6).

REFERENCIAS

- [1] R. C. Gonzalez and R. E. Woods, *Digital image processing*, 2nd ed. New Jersey: Prentice Hall, 2002.
- [2] P. Machado, J. Romero, M. Nadal, A. Santos, J. Correia, and A. Carballal, “Computerized measures of visual complexity,” *Acta Psychol. (Amst.)*, vol. 160, pp. 43–57, 2015.
- [3] M. I. Chacón and A. D. Corral, “Image complexity measure: A human criterion free approach,” in *Annual Conference of the North American Fuzzy Information Processing Society - NAFIPS*, 2005, pp. 241–246.
- [4] J. G. Snodgrass and M. Vanderwart, “A standardised set of 260 pictures: normal for name agreement, familiarity and visual complexity,” *J. Exp. Psychol. Hum. Learn. Mem.*, vol. 6, no. 2, pp. 174–215, 1980.
- [5] Y.-Q. Chen, J. Duan, Y. Zhu, X.-F. Qian, and B. Xiao, “Research on the image complexity based on neural network,” in *2015 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2015, pp. 295–300.
- [6] M. Liu, K. Gu, G. Zhai, and P. Le Callet, “Visual saliency detection via image complexity feature,” in *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 2777–2781.
- [7] B. Reimer, “Impact of Cognitive Task Complexity on Drivers’ Visual Tunneling,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2138, pp. 13–19, 2009.
- [8] A. D. Corral, “Modelado de percepción de bordes con lógica difusa,” M.C. Tesis, Instituto Tecnológico de Chihuahua, CUU, MX, 2003.
- [9] J. Alberich, D. Gómez Fontanills, and A. Ferrer Franquesa, *Percepción visual*, 1st ed. Cataluña: UOC, 2013.
- [10] J. Wagemans, J. Elder, M. Kubovy, S. Palmer, M. Peterson, and M. Singh, “A century

- of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization,” *Psychol. Bull.*, vol. 138, no. 6, pp. 1172–1217, 2012.
- [11] H. Yu and S. Winkler, “Image complexity and spatial information,” in *2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*, 2013, pp. 12–17.
 - [12] A. Forsythe, G. Mulhern, and M. Sawey, “Confounds in pictorial sets: The role of complexity and familiarity in basic-level picture processing,” *Behav. Res. Methods*, vol. 40, no. 1, pp. 116–129, 2008.
 - [13] J. Perkiö and A. Hyvärinen, “Modelling image complexity by independent component analysis , with application to content-based image retrieval,” *19th Int. Conf. Artif. Neural Networks Part II*, vol. 5769, pp. 1–11, 2009.
 - [14] W. Ewert, W. Dembski, and R. Marks, “Measuring meaningful information in images: algorithmic specified complexity,” *IET Comput. Vis.*, vol. 9, no. 6, pp. 884–894, 2015.
 - [15] M. Cardaci, V. Di Gesù, M. Petrou, and M. E. Tabacchi, “A fuzzy approach to the evaluation of image complexity,” *Fuzzy Sets Syst.*, vol. 160, no. 10, pp. 1474–1484, 2009.
 - [16] K. Wu, J. Vassileva, Y. Zhao, Z. Noorian, W. Waldner, and I. Adaji, “Complexity or simplicity? Designing product pictures for advertising in online marketplaces,” *J. Retail. Consum. Serv.*, vol. 28, pp. 17–27, 2016.
 - [17] A. Forsythe, “Visual Complexity: Is That All There Is?,” in *Engineering Psychology and Cognitive Ergonomics: 8th International Conference, EPCE 2009*, 2009, vol. 5639, pp. 158–166.
 - [18] M. Jakesch and H. Leder, “The qualitative side of complexity: Testing effects of ambiguity on complexity judgments,” *Psychol. Aesthetics, Creat. Arts*, vol. 9, no. 3, pp. 200–205, 2015.
 - [19] J. Huo, “An image complexity measurement algorithm with visual memory capacity and

- an EEG study,” in *2016 SAI Computing Conference*, 2016, pp. 264–268.
- [20] M. Dimitropoulou, J. A. Duñabeitia, P. Blitsas, and M. Carreiras, “A standardized set of 260 pictures for Modern Greek: Norms for name agreement, age of acquisition, and visual complexity,” *Behav. Res. Methods*, vol. 41, no. 2, pp. 584–589, 2009.
 - [21] P. Bonin, R. Peereman, N. Malardier, A. Méot, and M. Chalard, “A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies,” *Behav. Res. Methods, Instruments, Comput.*, vol. 35, no. 1, pp. 158–167, 2003.
 - [22] F. X. Alario and L. Ferrand, “A set of 400 pictures standardized for French: norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition,” *Behav. Res. methods, instruments, Comput.*, vol. 31, no. 3, pp. 531–552, 1999.
 - [23] S. Bhardwaj and A. Mittal, “A Survey on Various Edge Detector Techniques,” *Procedia Technol.*, vol. 4, pp. 220–226, 2012.
 - [24] Y. Zhang, X. Han, H. Zhang, and L. Zhao, “Edge detection algorithm of image fusion based on improved Sobel operator,” in *2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*, 2017, pp. 457–461.
 - [25] X. Li and H. Zhang, “An Improved CANNY Edge Detection Algorithm,” in *2009 Second International Workshop on Computer Science and Engineering*, 2009, pp. 497–500.
 - [26] M. I. Chacon, L. E. Aguilar, and A. Delgado, “Fuzzy adaptive edge definition based on the complexity of the image,” in *10th IEEE International Conference on Fuzzy Systems. (Cat. No.01CH37297)*, 2002, vol. 3, pp. 675–678.
 - [27] A. J. Hussain, A. Al-Fayadh, and N. Radi, “Image compression techniques: A survey in lossless and lossy algorithms,” *Neurocomputing*, vol. 300, pp. 44–69, 2018.
 - [28] B. Zhou, S. Xu, and X. X. Yang, “Computing the color complexity of images,” in *2015*

- 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2015, pp. 1898–1902.
- [29] M. Ivanovici and N. Richard, “A Naive Complexity Measure for Color Texture Images,” in *2017 International Symposium on Signals, Circuits and Systems (ISSCS)*, 2017, pp. 1–4.
 - [30] X. Guo, Y. Qian, L. Li, and A. Asano, “Assessment model for perceived visual complexity of painting images,” *Knowledge-Based Syst.*, vol. 159, pp. 110–119, 2018.
 - [31] S. E. Corchs, G. Ciocca, E. Bricolo, and F. Gasparini, “Predicting complexity perception of real world images,” *PLoS One*, vol. 11, no. 6, pp. 1–22, 2016.
 - [32] E. M. de la Calleja and R. Zenit, “Topological invariants can be used to quantify complexity in abstract paintings,” *Knowledge-Based Syst.*, vol. 126, pp. 48–55, 2017.
 - [33] C. Spampinato, S. Palazzo, I. Kavasidis, D. Giordano, M. Shah, and N. Souly, “Deep Learning Human Mind for Automated Visual Classification,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
 - [34] M. Hoenen, K. T. Lübke, and B. M. Pause, “Sensitivity of the human mirror neuron system for abstract traces of actions: An EEG-study,” *Biol. Psychol.*, vol. 124, pp. 57–64, 2017.
 - [35] B. Bonev, L. L. Chuang, and F. Escolano, “How do image complexity, task demands and looking biases influence human gaze behavior?,” *Pattern Recognit. Lett.*, vol. 34, no. 7, pp. 723–730, 2013.
 - [36] M. Ivanovici and N. Richard, “A Naive Complexity Measure for color texture images,” in *ISSCS 2017 - International Symposium on Signals, Circuits and Systems*, 2017, pp. 1–4.
 - [37] Y.-Q. Chen, J. Duan, Y. Zhu, X.-F. Qian, and B. Xiao, “Research on the image complexity based on neural network,” in *2015 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2015, pp. 295–300.

- [38] B. Xiao, J. Duan, X. Liu, Y. Zhu, and H. Wang, "Evaluation of Image Complexity Based on SVOR," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 32, no. 7, pp. 1–18, 2018.
- [39] D. Kundu, D. Ghadiyaram, A. C. Bovik, and B. L. Evans, "Large-Scale Crowdsourced Study for Tone-Mapped HDR Pictures," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4725–4740, 2017.
- [40] E. Saraee, M. Jalal, and M. Betke, "SAVOIAS : A Diverse , Multi-Category Visual Complexity Dataset," *arXiv*, pp. 1–10, 2018.
- [41] L. Madrid, M. I. Chacón, D. A. Posada, and J. A. Ramirez, "Human Image Complexity Analysis Using a Fuzzy Inference System," in *2019 International Conference on Fuzzy Systems*, 2019, pp. 1–6.
- [42] W. A. Dembski, W. Ewert, and R. J. Marks, "Measuring meaningful information in images: algorithmic specified complexity," *IET Comput. Vis.*, vol. 9, no. 6, pp. 884–894, 2015.
- [43] J. Huo, "A measurement method for the mismatch between the image target and salient points as a metric for image complexity," in *2015 Science and Information Conference (SAI)*, 2015, pp. 645–649.
- [44] E. Z. H. Tan, K. G. Smitha, and A. P. Vinod, "Detection of Familiar and Unfamiliar Images Using EEG-Based Brain-Computer Interface," in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 2015, pp. 3152–3157.
- [45] L. Palumbo, A. D. J. Makin, and M. Bertamini, "Examining visual complexity and its influence on perceived duration," *J. Vis.*, vol. 14, no. 14, pp. 1–18, 2014.
- [46] F. Yaghmaee and M. Jamzad, "Estimating watermarking capacity in gray scale images based on image complexity," *EURASIP J. Adv. Signal Process.*, vol. 2010, p. 9, 2010.
- [47] J. Huo, "Image Complexity and Visual Working Memory Capacity," *Emerg. Trends Adv. Technol. Comput. Intell.*, vol. 647, pp. 301–314, 2016.
- [48] N. Gauvrit, F. Soler-Toscano, and A. Guida, "A preference for some types of complexity

- comment on ‘perceived beauty of random texture patterns: A preference for complexity,’” *Acta Psychol. (Amst.)*, vol. 174, pp. 48–53, 2017.
- [49] S. H. Yoon, J. Lim, and Y. G. Ji, “Assessment model for perceived visual complexity of automotive instrument cluster,” *Appl. Ergon.*, vol. 46, no. A, pp. 76–83, 2015.
- [50] A. Gartus and H. Leder, “Predicting perceived visual complexity of abstract patterns using computational measures: The influence of mirror symmetry on complexity perception,” *PLoS One*, vol. 12, no. 11, pp. 1–30, 2017.
- [51] P. Khanzadi, B. Majidi, and E. Akhtarkavan, “A Novel Metric for Digital Image Quality Assessment using Entropy-Based Image Complexit,” in *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, 2017, pp. 440–445.
- [52] X. Li, Q. Wu, Y. Kou, L. Hou, and H. Xie, “Driver’s Eyes State Detection Based on Adaboost Algorithm and Image Complexity,” in *2015 6th International Conference on Intelligent Systems Design and Engineering Applications (ISDEA)*, 2015, pp. 349–352.
- [53] G. Ciocca, S. Corchs, F. Gasparini, E. Bricolo, and R. Tebano, “Does Color Influence Image Complexity Perception?,” *Lect. Notes Comput. Sci.*, vol. 9016, pp. 139–148, 2015.
- [54] M. M. Marin and H. Leder, “Examining Complexity across Domains: Relating Subjective and Objective Measures of Affective Environmental Scenes, Paintings and Music,” *PLoS One*, vol. 8, no. 8, pp. 1–35, 2013.
- [55] L. Sun, T. Yamasaki, and K. Aizawa, “Relationship Between Visual Complexity and Aesthetics: Application to Beauty Prediction of Photos,” in *Computer Vision - ECCV 2015 Workshops*, 2015, pp. 20–34.
- [56] L. Aguilar, “La percepción humana en la detección de niveles de borde en una imagen, basada en lógica difusa,” M.C. Tesis, Instituto Tecnológico de Chihuahua, CUU, MX, 2002.

- [57] K.-J. Yoon and I. S. Kweon, "Color image segmentation considering of human sensitivity for color pattern variations," *Proc. SPIE*, vol. 4572, pp. 269–278, 2001.
- [58] T. Law, H. Itoh, and H. Seki, "Image filtering, edge detection, and edge tracing using fuzzy reasoning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 5, pp. 481–491, 1996.
- [59] D. Arthur and S. Vassilvitskii, "K-Means++: the Advantages of Careful Seeding," *Proc ACM-SIAM Symp. Discret. algorithms.*, vol. 8, pp. 1027–35, 2007.
- [60] M. Cardaci, V. Di Gesù, M. Petrou, and M. E. Tabacchi, "A fuzzy approach to the evaluation of image complexity," *Fuzzy Sets Syst.*, vol. 160, no. 10, pp. 1474–1484, 2009.
- [61] N. Gauvrit, F. Soler-Toscano, and A. Guida, "A preference for some types of complexity comment on 'perceived beauty of random texture patterns: A preference for complexity,'" *Acta Psychol. (Amst)*, vol. 174, pp. 48–53, 2017.
- [62] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: the fuzzy c-means clustering algorithm," *Comput. Geosci.*, vol. 10, no. 2–3, pp. 191–203, 1984.
- [63] D. E. Gustafson and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," in *1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, 1978, pp. 761–766.
- [64] Gath I. and Geva A. B., "Unsupervised optimal fuzzy clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 773–781, 1989.

ANEXOS

A. ALGORITMOS DE AGRUPAMIENTO

Algoritmo FCM

De [62], FCM es uno de los algoritmos más populares en el agrupamiento difuso y fue realizado por Bezdek. FCM intenta encontrar una partición (c grupos difusos) para un grupo de datos $x_i \in \mathbb{R}^d, i = 1, \dots, n$ mientras se minimiza la función costo:

$$J(\mathbf{U}, \mathbf{M}) = \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^m D_{ij}$$

Donde:

$\mathbf{U} = [u_{ij}]_{c \times n}$ es la matriz de partición difusa y $u_{ij} \in [0,1]$ es el coeficiente de membresía del objeto i en el grupo j .

$\mathbf{M}$ es la matriz de prototipos de grupo con elementos $a_j, j = 1, \dots, c$
 m es el parámetro de fusificación y usualmente es 2.

D_{ij} es la medida de distancia entre el dato i y el prototipo j .

A continuación, se explica el algoritmo FCM, en donde la métrica utilizada es la distancia Euclidiana.

1. Seleccionar valores apropiados para m, c y un número pequeño positivo ϵ . Inicializar la matriz prototipo $\mathbf{M}$ aleatoriamente. Establecer el valor de iteraciones $t = 0$.
2. Calcular (en $t = 0$) o actualizar (para $t > 0$) la matriz de pertenencia mediante:

$$u_{ij}^{(t+1)} = \frac{1}{(\sum_{l=1}^c (D_{ij} / D_{il})^{1/(1-m)})}$$

3. Actualizar los elementos de la matriz prototipo $\mathbf{M}$ mediante:

$$a_j^{(t+1)} = \frac{(\sum_{i=1}^n (u_{ij}^{(t+1)})^m x_i)}{(\sum_{i=1}^n (u_{ij}^{(t+1)})^m)}$$

para $j = 1, \dots, c$.

4. Repetir los pasos 2) y 3) hasta que $\|M^{(t+1)} - M^{(t)}\| < e$

A pesar de su popularidad FCM sufre de la presencia de ruido, es muy sensible a los límites, y tiene dificultad para encontrar las particiones iniciales.

Algoritmo GK

De [63], el algoritmo GK es una poderosa técnica de agrupamiento con un gran número de aplicaciones en varios campos incluyendo el procesamiento de imagen y clasificación e identificación de sistemas. Su mayor característica es su capacidad de adaptación a la forma de los grupos mediante la matriz de covarianza de los agrupamientos formados. Los pasos que sigue este algoritmo son:

Dado un conjunto de datos $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$, determine el número de grupos que se formarán $1 < c < N$, el exponente de fusificación m (usualmente es igual a 2), la tolerancia de terminación $\varepsilon > 0$ (comúnmente 10^{-3}) y el volumen del grupo, ρ , (usualmente igual a 1). Inicializa la matriz de partición $\mathbf{U}$ de manera aleatoria.

Repita para $l = 1, 2, \dots$

Paso 1. Calcula la media de los grupos

$$\mathbf{v}_i^{(l)} = \frac{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m \mathbf{x}_k}{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m}, \quad 1 \leq i \leq c.$$

Paso 2. Calcula la matriz de covarianza, $\mathbf{F}$, de cada grupo

$$\mathbf{F}_i = \frac{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m (\mathbf{x}_k - \mathbf{v}_i^{(l)}) (\mathbf{x}_k - \mathbf{v}_i^{(l)})^T}{\sum_{k=1}^N (\mu_{ik}^{(l-1)})^m}, \quad 1 \leq i \leq c$$

Paso 3. Calcula la distancia de los datos con cada grupo

$$D_{i,k} = (\mathbf{x}_k - \mathbf{v}_i^{(l)})^T [\rho_i \det(\mathbf{F}_i)^{1/n} \mathbf{F}_i^{-1}] (\mathbf{x}_k - \mathbf{v}_i^{(l)}), \quad 1 \leq i \leq c, \quad 1 \leq k \leq N$$

donde $\det$ significa la operación de determinante.

Paso 4. Actualiza la matriz de partición

Para $1 \leq k \leq N$

si $D_{ikl} > 0$ para $1 \leq i \leq c$,

$$\mu_{ik}^{(l)} = \frac{1}{\sum_{j=1}^K (D_{ikl} / D_{jkl})^{2/m-1}}$$

de otra manera

$$\mu_{ik}^{(l)} = 0 \quad \text{si } D_{ikl} > 0, \quad \text{y } \mu_{ik}^{(l)} \in [0,1] \quad \text{con } \sum_{i=1}^C \mu_{ik}^{(l)} = 1$$

Hasta

$$\|\mathbf{U}^{(l)} - \mathbf{U}^{(l-1)}\| < \varepsilon$$

El algoritmo de GK tiene un problema cuando el número de muestras (en algunos grupos) es pequeño o cuando los datos dentro de un grupo están linealmente correlacionados. En este caso la matriz de covarianza de los grupos se hace singular y su inversa no puede ser calculada.

Algoritmo GG

De [64], este algoritmo utiliza como métrica de distancia la FMLE (*fuzzy maximum likelihood estimates*) propuesto por Bezdek y Dunn:

$$D_{ij}(x_i, a_j) = \frac{\sqrt{\det(F_{wj})}}{\alpha_j} \exp\left(\frac{1}{2}(x_i - a_j)^t F_{wj}^{-1} (x_i - a_j)\right)$$

Note que al contrario del algoritmo GK, la distancia incluye un término exponencial y por lo tanto disminuye más rápidamente que la norma interna, y además envuelve el término F_{wj} que denota el valor del grupo j en la matriz de covarianza difusa, y está dado por:

$$F_{wj} = \frac{\sum_{i=1}^n \mu_{ij}^m (x_i - a_j)(x_i - a_j)^t}{\sum_{i=1}^n (\mu_{ij})^m}, 1 \leq j \leq c$$

Donde $m = 1$ en el FMLE original, pero en este caso se utiliza 2 para que la partición se vuelva más difusa, para compensar el exponencial en la distancia. A_i es la probabilidad previa de seleccionar un grupo i y está dada por:

$$\alpha_j = \frac{1}{n} \sum_{i=1}^n \mu_{ij}$$

Los grados de membresía μ_{ij} son interpretados como las probabilidades posteriores de seleccionar el grupo j dado un punto x_i . Gath y Geva reportan que este algoritmo detecta grupos de varias formas, densidades y tamaños. La matriz de covarianza se utiliza conjuntamente con la distancia exponencial para que los grupos no estén forzados en volumen. Las desventajas de este algoritmo es su falta de robustez, ya que requiere una buena inicialización por la distancia exponencial.

APÉNDICE A

Manual de usuario de interfaz para medir la complejidad

En esta sección se muestra como luce la interfaz de usuario generada para determinar la complejidad de imágenes mediante los sistemas previamente descritos, así como también algunas de las características y como instalarla, en la Figura A.1 se muestra la interfaz gráfica de usuario.

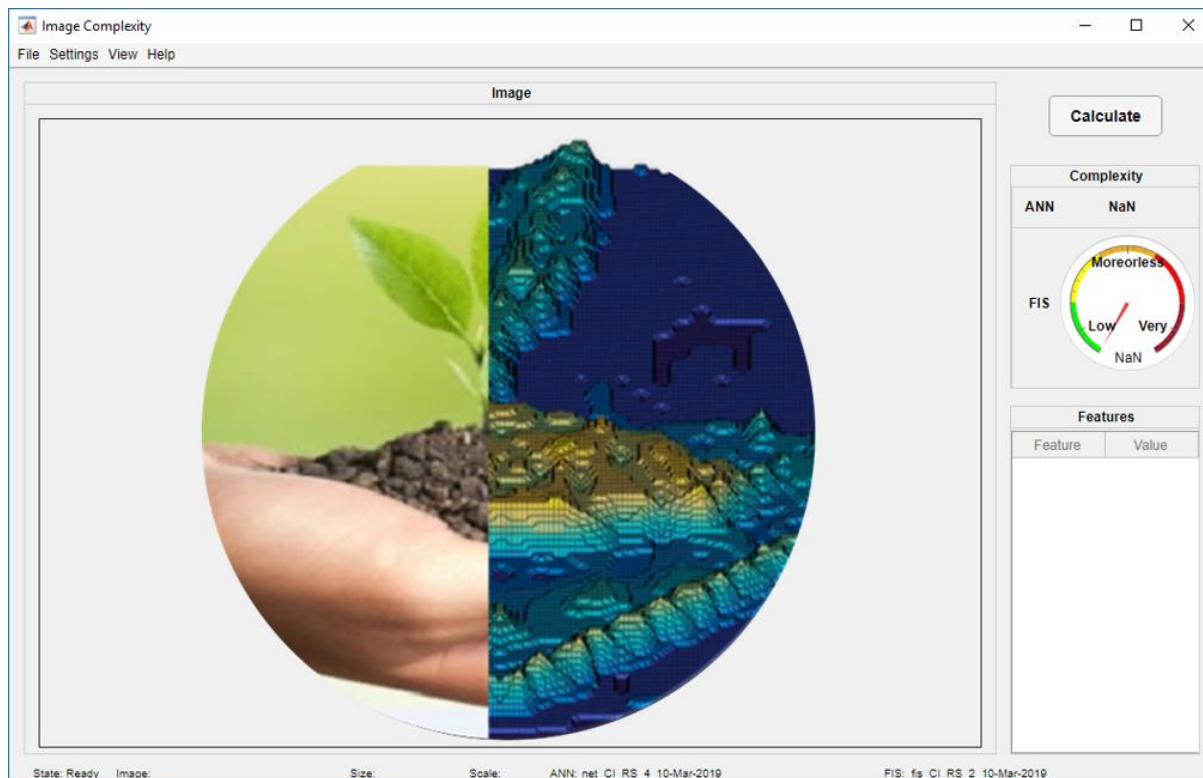


Figura A.1. Interfaz gráfica de usuario.

Instalación de la aplicación

Para instalar la aplicación hay que dirigirse al instalador, el cual se encuentra en la dirección `GUIImageComplexity\for_redistribution` bajo el nombre de `Installer_imComplexityApp`, como se muestra en la captura de pantalla de la Figura A.2, dar doble clic sobre él y después aceptar cuando salga la ventana emergente preguntando si está seguro, para esto debe tener permiso de administrador.

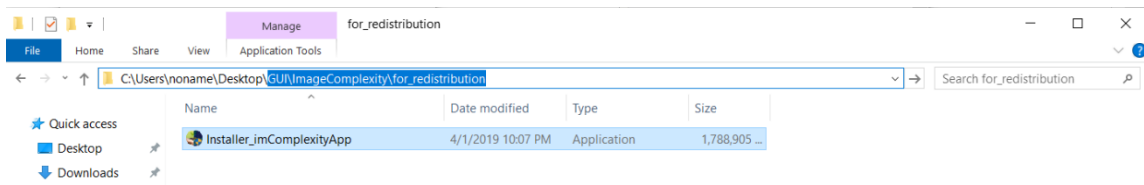


Figura A.2. Directorio de instalador de aplicación.

Una vez que se tiene el permiso del administrador la instalación comenzará, en la Figura A.3, se muestra una serie de ventanas que se irán desplegando, las cuales hay que ir siguiendo para realizar la instalación de la aplicación de forma correcta.

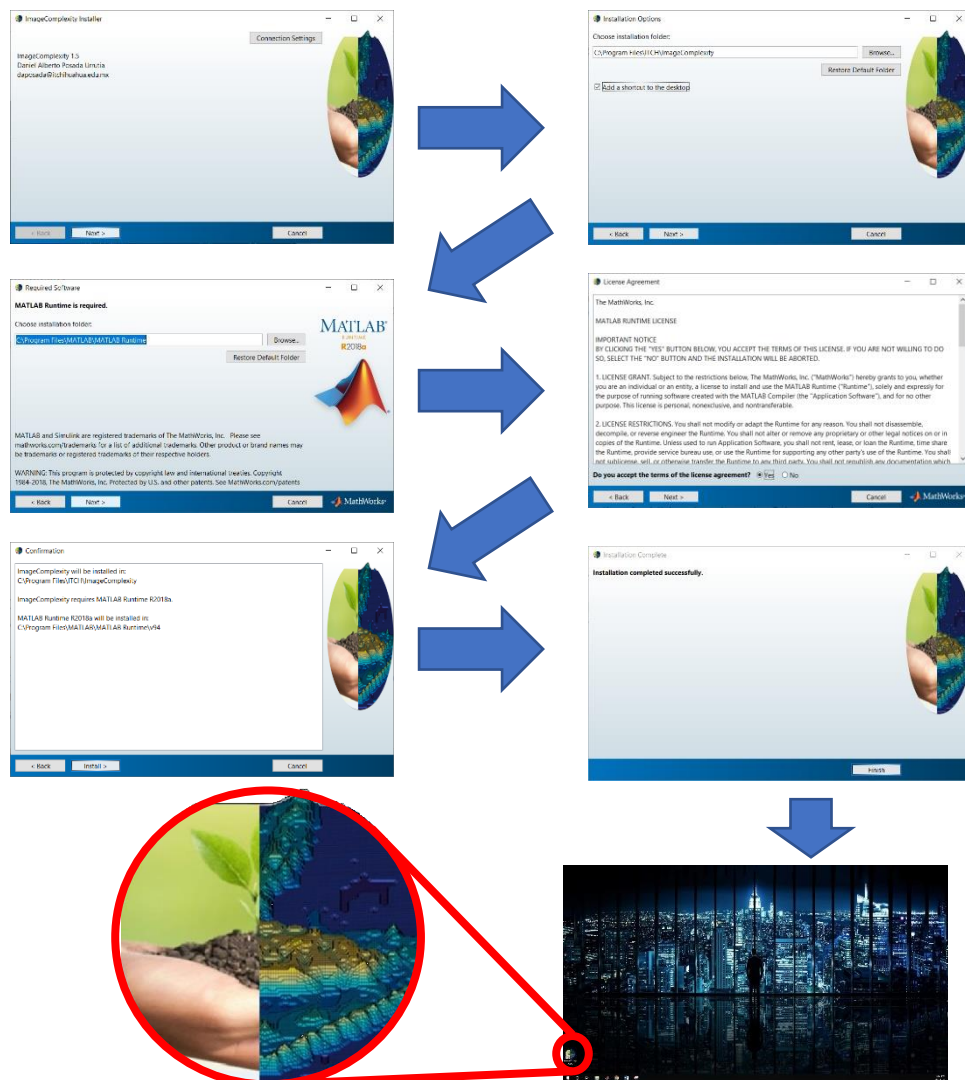


Figura A.3 Proceso de instalación de aplicación.

Ya que el instalador ha finalizado el programa está listo para usarse, sin embargo, dos características para determinar la complejidad no estarán disponibles, para poder utilizarlas dirigirse a la sección de [NumberOfColors y ColorHarmony error](#).

Elementos básicos de la aplicación

Al inicializar la aplicación se mostrará una ventana como en la Figura A.4, en la cual se describe brevemente los elementos principales que la conforman.

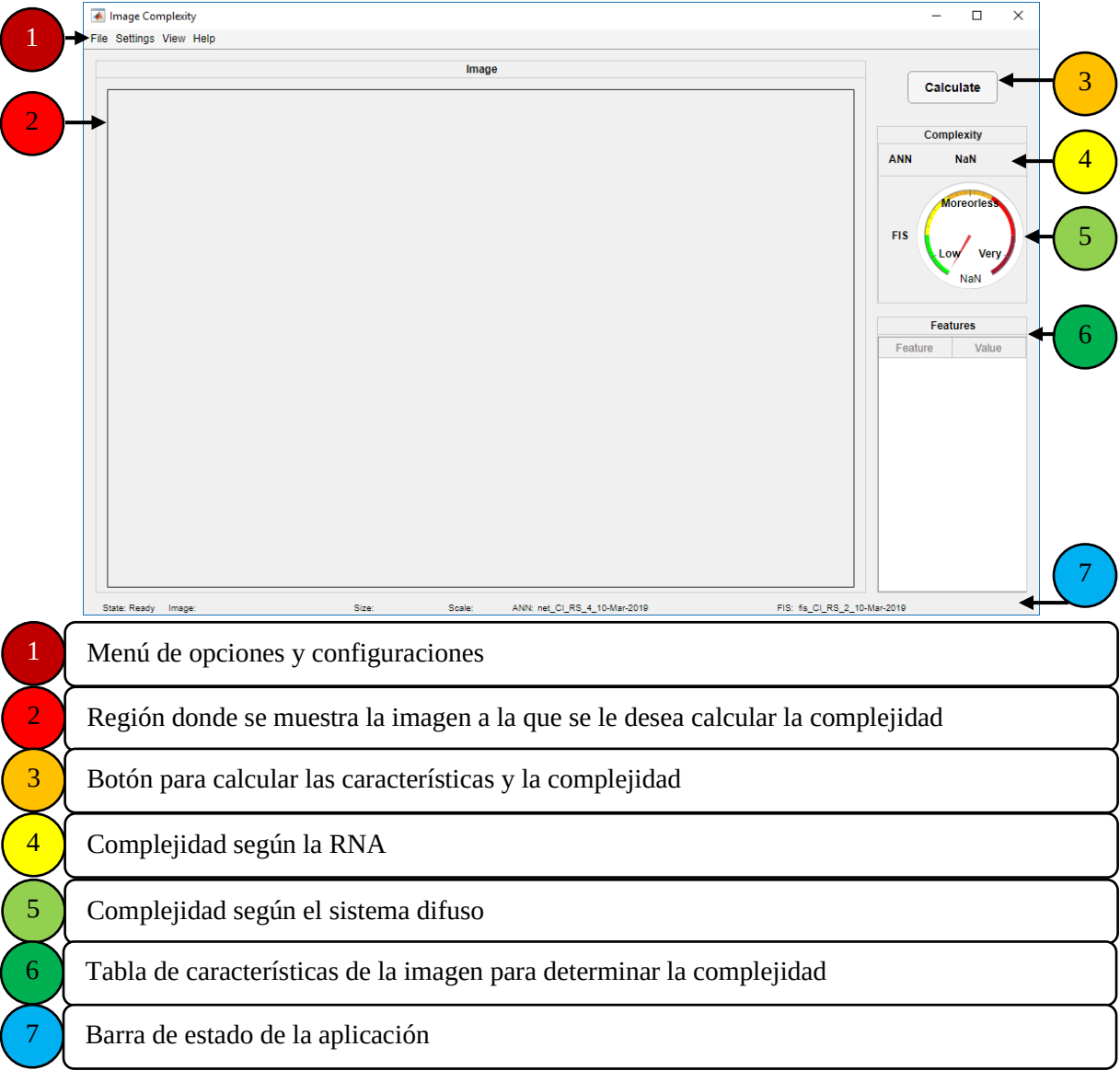


Figura A.4. Elementos básicos de la interfaz de usuario.

Menú 'File'

El menú 'File' abrirá una pestaña en la cual se dan ciertas opciones que permiten desde cargar una nueva imagen a la cual se le calculará la complejidad hasta cargar otros sistemas, en la Figura A.5 se muestra.

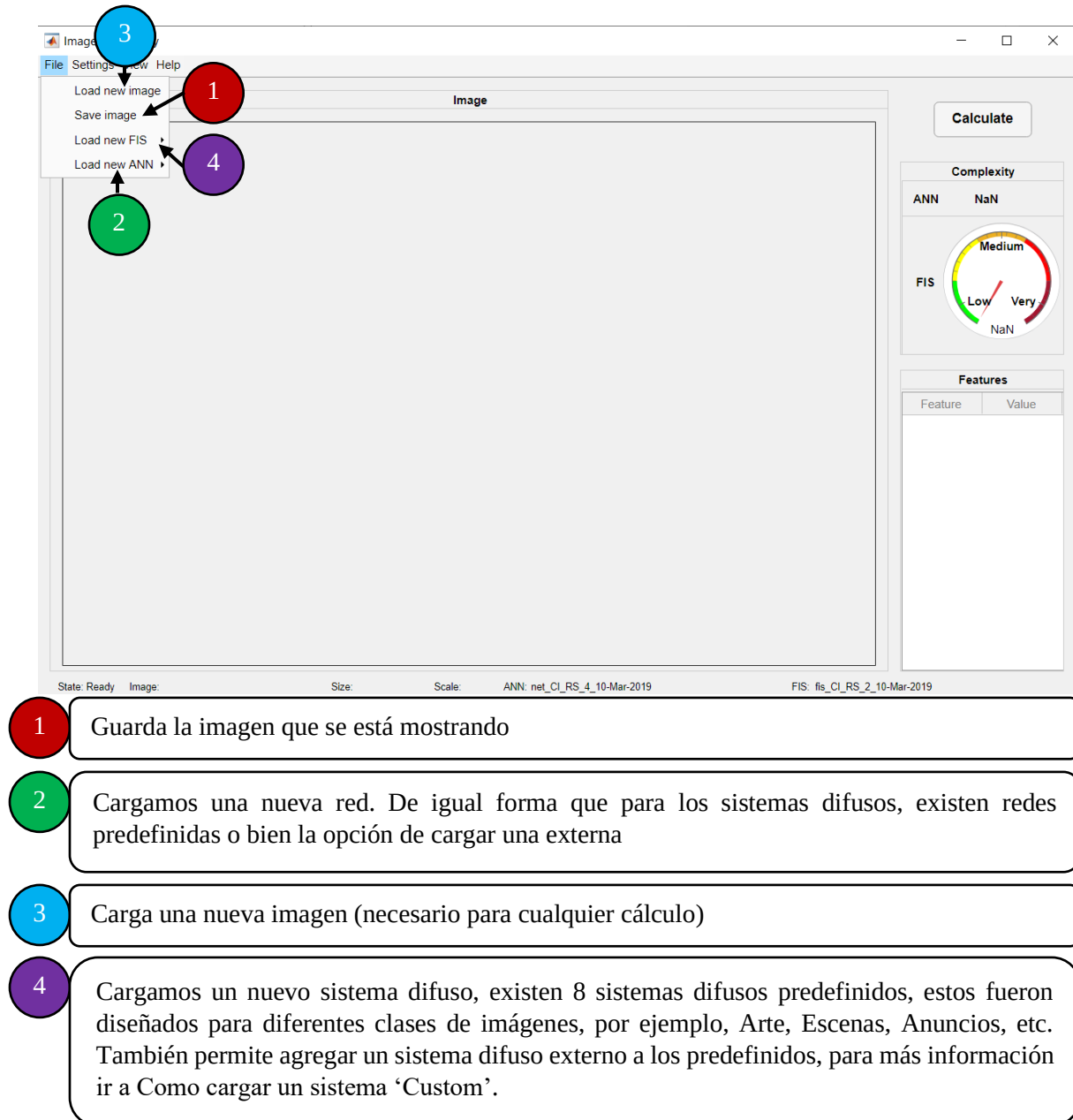
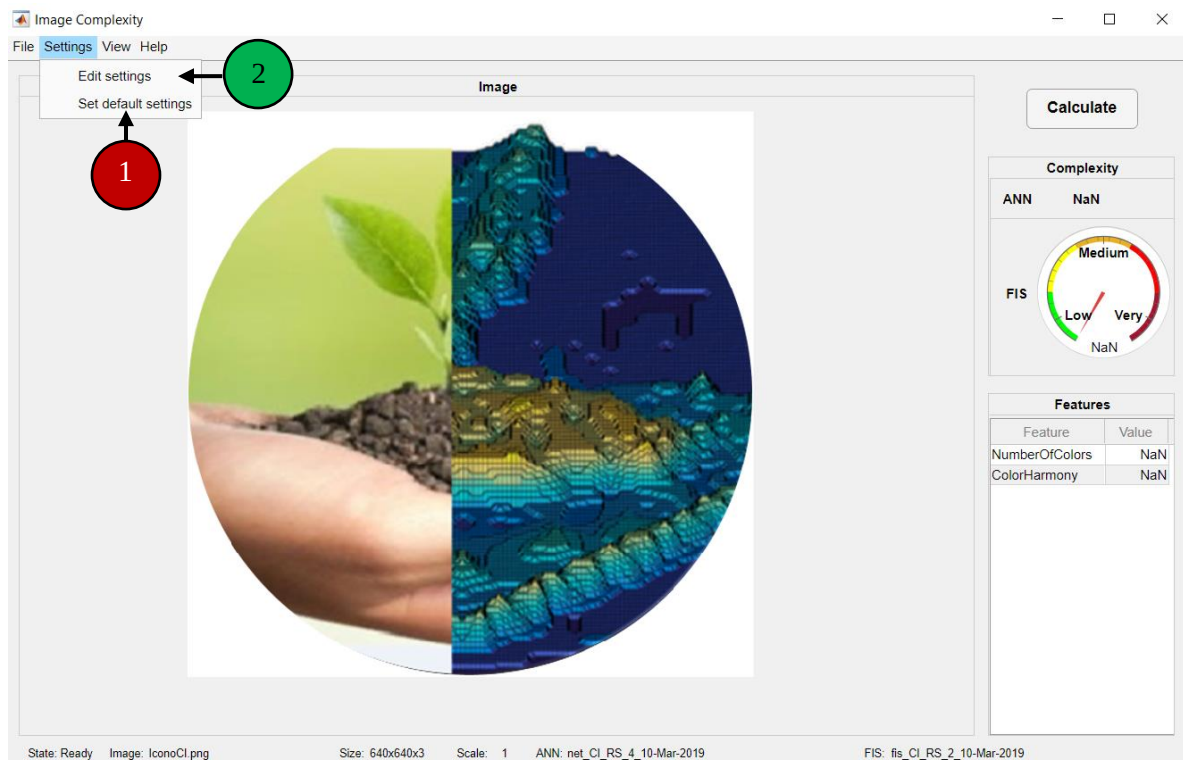


Figura A.5. Elementos del menú 'File'.

Menú 'Settings'

El menú 'Settings' permite configurar algunas características y seleccionar cuales extraer, también permite editar algunos aspectos de la imagen o bien regresar a la configuración predefinida, un ejemplo de cómo luce se encuentra en la Figura A.6.



1

Reestablece los valores por defecto de las configuraciones

2

Abre un nuevo panel de control para editar ciertas configuraciones de las características y los sistemas, así como también de la imagen

Figura A.6. Elementos del menú 'Settings'.

‘Edit settings’

El menú ‘Edit settings’, abrirá un panel de configuración como se muestra en la Figura A.7.

1 Características que se pueden extraer

2 Calidad de la compresión JPG

3 Configuraciones de la característica medida de complejidad del color

4 Selecciona si se quiere una escala o un reajuste de la imagen por pixeles

5 Marcar si se desea calcular la imagen con los sistemas

6 Guarda lo editado y cierra panel

7 Cancela lo editado y cierra panel

8 Escala a la que se ajustara la imagen

9 Tamaño en pixeles a los que se ajustara la imagen

10 Método de escalado de la imagen

Figura A.7. Panel ‘Settings’.

Menú 'View'

El menú 'View' permite visualizar ciertas características de la imagen, como lo es ver la imagen en su espacio de color RGB, ver los bordes según su nivel o bien los bordes Canny o Sobel, visualizar la medida de complejidad del color a nivel pixeles, ver la imagen en escala de grises, así como también permite abrir un nuevo panel el cual muestra algunos de los parámetros y desempeños de los sistemas (red neuronal o sistema de inferencia difuso).

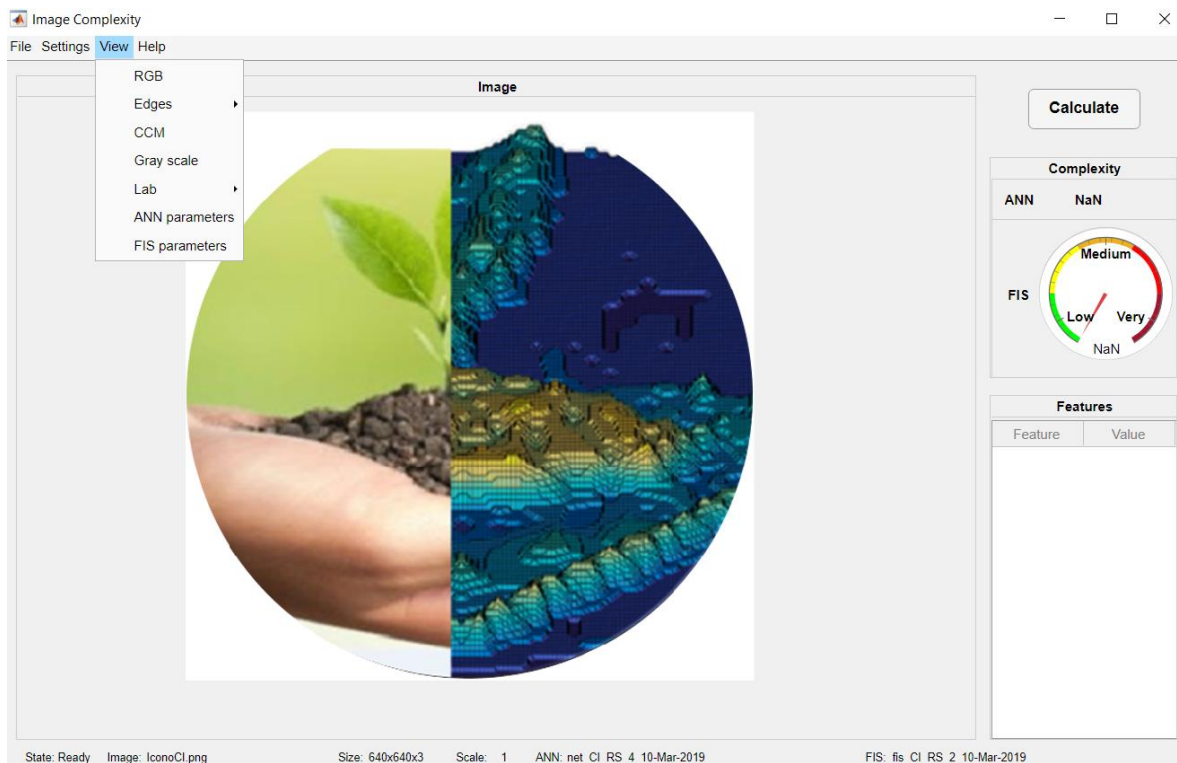


Figura A.8. Elementos del menú 'View'.

En la Figura A.9 se muestran tres ejemplos de las características de las imágenes que se pueden ver a través del menú 'View'.

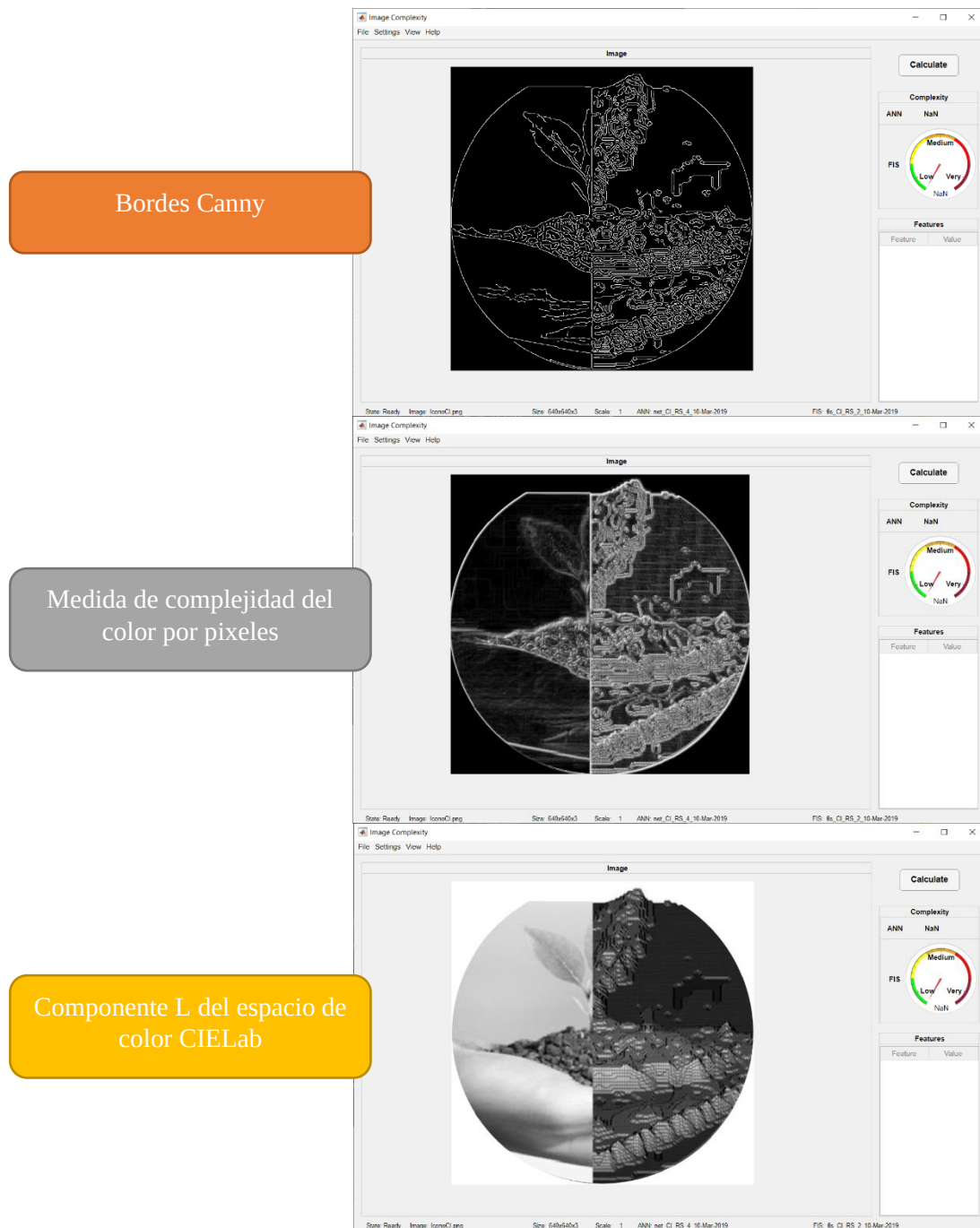


Figura A.9. Ejemplos de características del menú 'View'.

Además de las características, el menú View permite ver algunos de los parámetros de los sistemas, como se muestra en la Figura A.10.

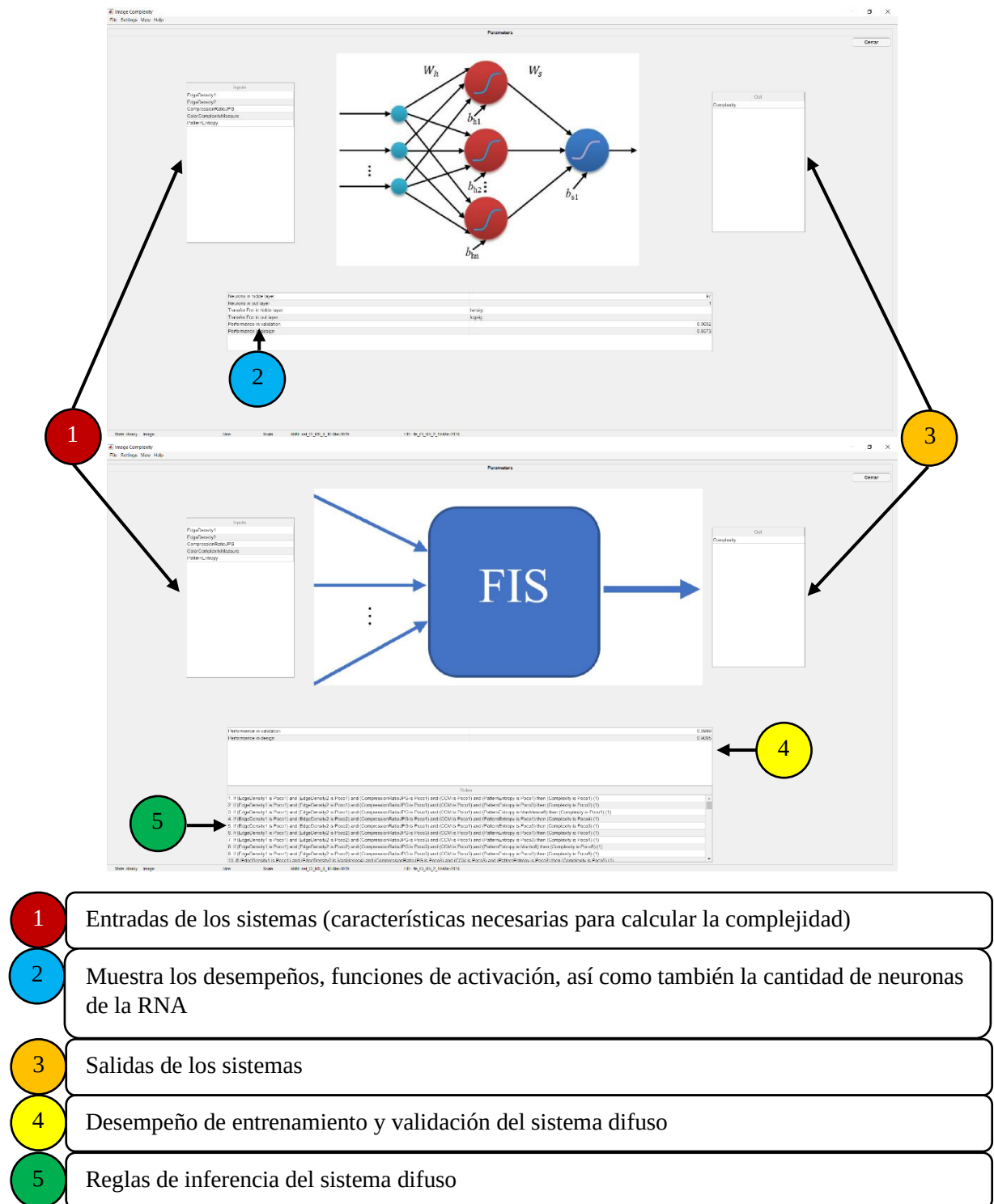


Figura A.10. Paneles de parámetros de los sistemas.

Menú 'Help'

El menú 'Help' abre el manual de usuario, ya que es el documento donde se encuentra la ayuda para el uso de la aplicación, en la Figura A.11 se muestra un ejemplo de cómo se abre el manual de usuario como un documento PDF.

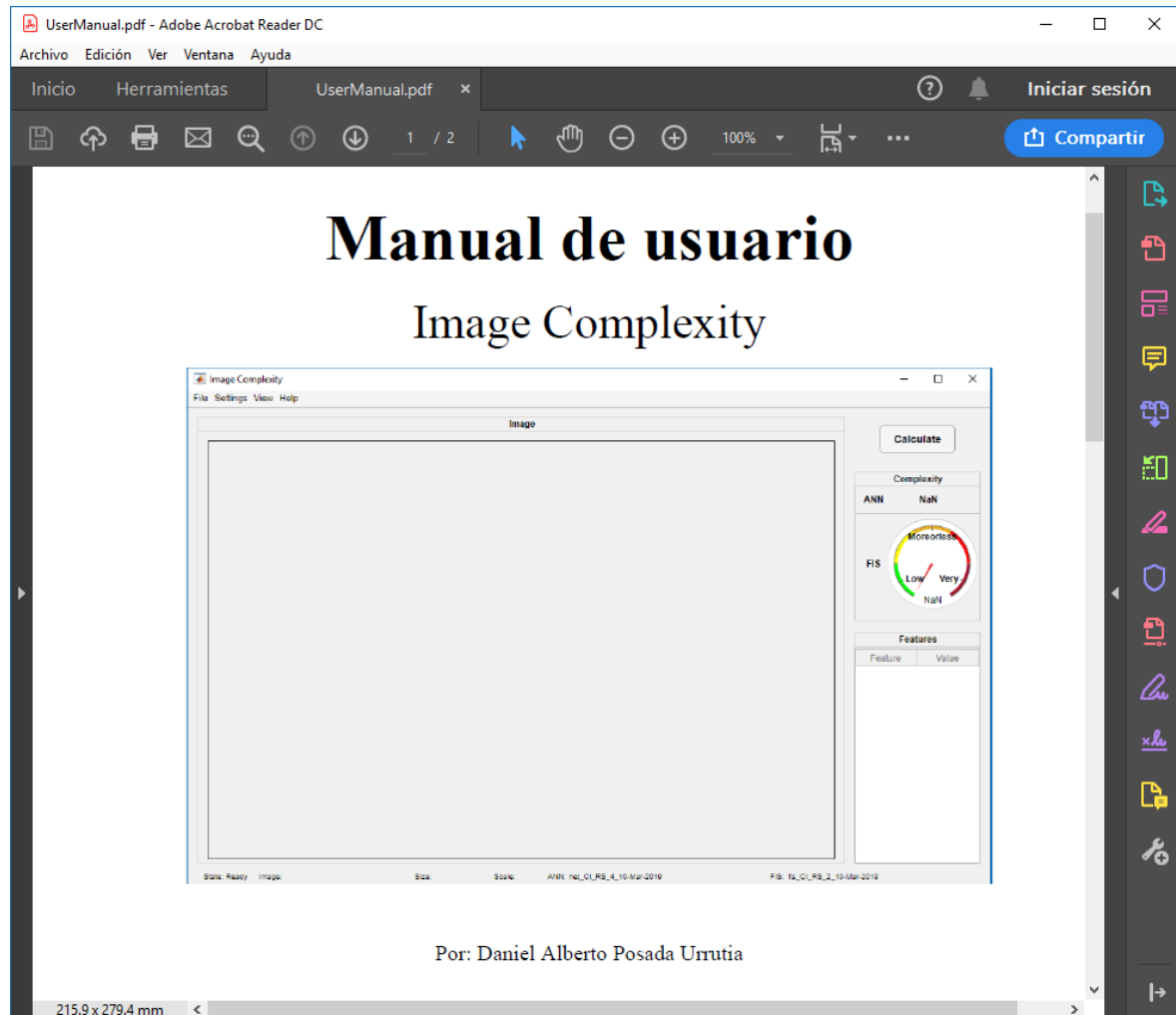


Figura A.11. Manual de usuario.

Como cargar un sistema 'Custom'

Para cargar un sistema personalizado como lo puede ser una red o bien un sistema difuso, debe irse al menú 'File', después en cargar nuevo sistema y en sistema personalizado, en este punto se abrirá una ventana exploradora de Windows, esta primer ventana será para la selección del sistema (RNA o FIS), una vez seleccionado el sistema y aceptado se abrirá una nueva ventana que pedirá un archivo '.mat' con los parámetros del sistema, este archivo es importante ya que contiene con que características fue diseñado el sistema y le proporciona la información necesaria a la aplicación para poder trabajar con él. A continuación, se enlista como está estructurado los parámetros de los sistemas para un manejo adecuado de los mismos.

- `params.Features`: Aquí se guardan en forma de celdas las características que utiliza el sistema, como: `params.Features = {'EdgeDensity1' 'PatternEntropy'}`; Es importante el orden, ya que la entrada uno corresponderá a la primer características y así sucesivamente. (Imprescindible tanto para FIS como para RNA).
- `params.Fmax`: Parámetro necesario para realizar el escalamiento de las características, aquí está el valor máximo con el que se diseñó el sistema. Es necesario para el funcionamiento ya que todas las características son escaladas.
- `params.Fmin`: Parámetro necesario para realizar el escalamiento de las características, aquí están los mínimos de las características con los que se diseñó el sistema. Es necesario para el funcionamiento, ya que todas las características deben ser escaladas de similar forma que como fueron cuando se diseñó.
- `params.InputDesign`: Contiene los valores de las características con las que fue diseñado el sistema, no necesario para funcionamiento, pero si para ver el desempeño que se logró en entrenamiento.
- `params.TargetDesign`: Contiene los valores de los objetivos, para este caso la complejidad subjetiva con los que se diseñaron los sistemas, tampoco necesario para funcionamiento, solo para ver el desempeño en entrenamiento.
- `params.InputValidation`: Contiene los valores de las características de las imágenes para validación, no necesarias para funcionamiento, solo para ver el desempeño.

- `params.TargetValidation`: Contiene los valores de los objetivos de las imágenes para validación, no necesarias para funcionamiento, solo para ver el desempeño.
- `params.TypeCluster`: Indica con qué tipo de agrupamiento se realizó el diseño del sistema difuso. Es un parámetro informativo del sistema difuso, por lo que no es necesario para el funcionamiento del FIS.
- `params.inFuzzyVars`: Indica las variables lingüísticas que se le dio como entrada al sistema difuso, de igual forma es un parámetro informativo que puede ser de ayuda pero puede trabajar sin él. Se utiliza para ver el diagrama del sistema en parámetros de la aplicación.
- `params.outFuzzyVars`: Indica la salida del sistema difuso, de igual forma que para `inFuzzyVars`.
- `params.kGroupsIn`: Indica bajo cuantos valores difusos se diseñó el sistema difuso (no necesario).
- `params.kGroupsOut`: Indica bajo cuantos valores de salida se diseñó el sistema difuso (no necesario).
- `params.maxit`: Indica cuantas réplicas del sistema se realizó para buscar el mejor (no necesario).
- `params.Nneuronas`: Indica bajo que rango de neuronas se buscó la mejor RNA, no necesario para funcionamiento.
- `params.intentos`: Indica con cuantas réplicas se buscó la mejor RNA (no necesario).
- `params.WV`: Indica el peso que se le dio a validación para la búsqueda del mejor sistema (no necesario).
- `params.WD`: Indica el peso que se le dio a diseño para la búsqueda del mejor sistema (no necesario).

Es importante que el sistema que se desea cargar sea compatible con la versión con la que se realizó la aplicación, para este caso Matlab 2018a, por ejemplo, para Matlab 2019 las estructuras de los sistemas difusos cambiaron, por lo que no se garantiza la compatibilidad de la aplicación con estos sistemas.

NumberOfColors y ColorHarmony error

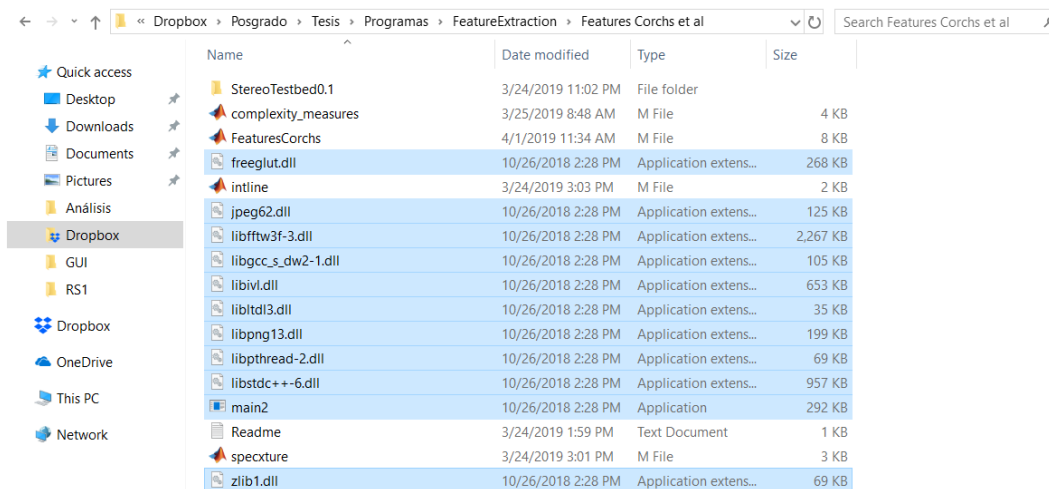
Si al calcular las características de NumberOfColor y ColorHarmony aparece el valor de NaN como se muestra en la Figura A.12.

Features	
Feature	Value
NumberOfColors	NaN
ColorHarmony	NaN

Figura A.12. Error con características.

Debe seguir los siguientes pasos:

- Copiar los archivos que se muestran en la Figura A.13, estos pueden estar en la carpeta de FeatureExtraction\Features Corchs et al, o bien en la carpeta GUI.



Name	Date modified	Type	Size
StereoTestbed0.1	3/24/2019 11:02 PM	File folder	
complexity_measures	3/25/2019 8:48 AM	M File	4 KB
FeaturesCorchs	4/1/2019 11:34 AM	M File	8 KB
freeglut.dll	10/26/2018 2:28 PM	Application extens...	268 KB
intline	3/24/2019 3:03 PM	M File	2 KB
jpeg62.dll	10/26/2018 2:28 PM	Application extens...	125 KB
libfftw3f-3.dll	10/26/2018 2:28 PM	Application extens...	2,267 KB
libgcc_s_dw2-1.dll	10/26/2018 2:28 PM	Application extens...	105 KB
libivl.dll	10/26/2018 2:28 PM	Application extens...	653 KB
libltdl3.dll	10/26/2018 2:28 PM	Application extens...	35 KB
libpng13.dll	10/26/2018 2:28 PM	Application extens...	199 KB
libpthread-2.dll	10/26/2018 2:28 PM	Application extens...	69 KB
libstdc++-6.dll	10/26/2018 2:28 PM	Application extens...	957 KB
main2	10/26/2018 2:28 PM	Application	292 KB
Readme	3/24/2019 1:59 PM	Text Document	1 KB
speccture	3/24/2019 3:01 PM	M File	3 KB
zlib1.dll	10/26/2018 2:28 PM	Application extens...	69 KB

Figura A.13. Archivos por copiar.

- Dirigirse a la carpeta donde se instaló la aplicación, por defecto será algo similar a lo que se muestra en la Figura A.14,

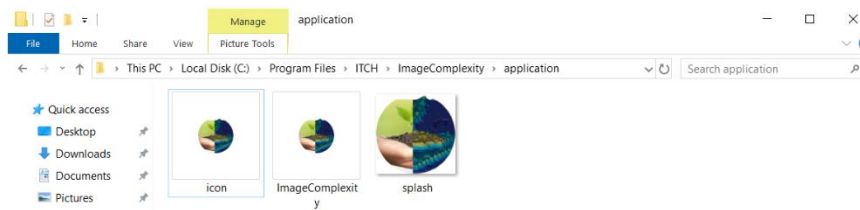


Figura A.14. Carpeta destino.

- Pegar los documentos y proporcionar permiso para esto (necesario permiso de administrador Figura A.15),

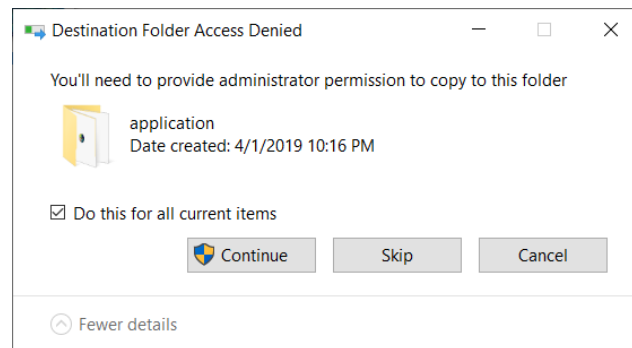


Figura A.15. Permiso para pegar.

- Debe quedar algo similar a lo que se muestra en la Figura A.16,

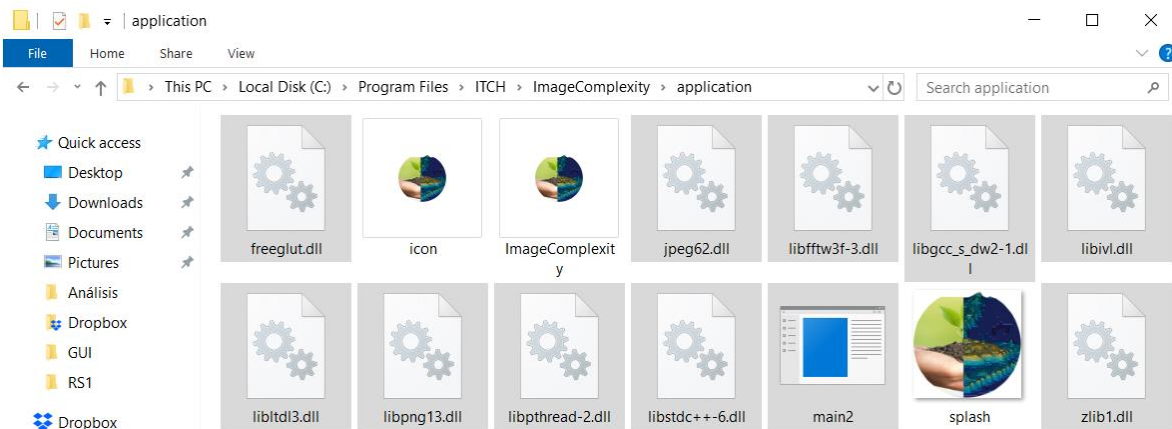


Figura A.16. Vista final de la carpeta destino.

- En este punto ya deberá ser capaz de extraer cualquier característica, en la Figura A.17 se muestra un ejemplo de la aplicación ya extrayendo todas las características,

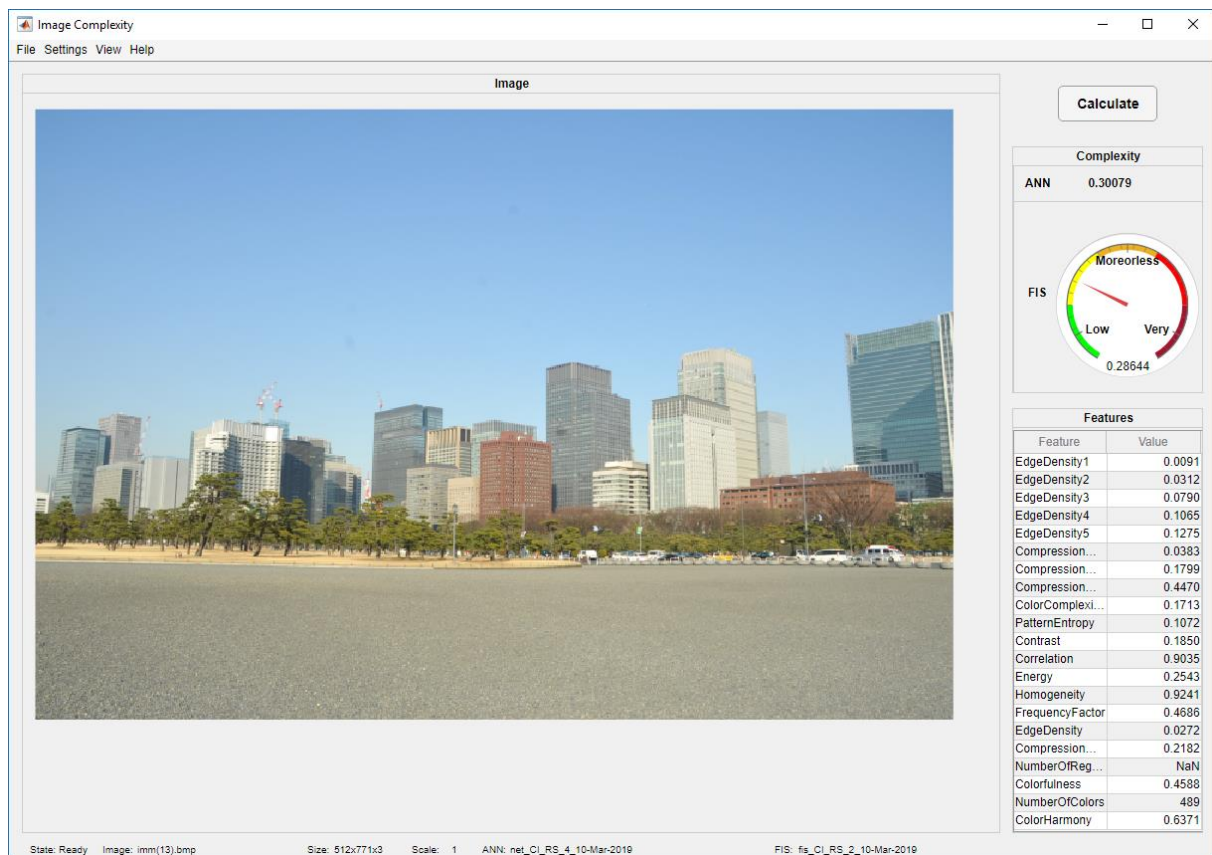


Figura A.17. Aplicación extrayendo todas las características.

Errores y soluciones

- * Error 101: No existe alguna imagen a la cual se le deban extraer las características y calcular la complejidad, para esto hay que cargar una imagen.
- * Error 201: Problema al leer el sistema difuso, puede tratarse de una incompatibilidad con el sistema que se desea cargar, revisar compatibilidad con Matlab 2018a.
- * Error 202: Problema al cargar los parámetros del sistema, revisar que estén correctos con el sistema (FIS).
- * Error 203: Problema al leer el sistema RNA, puede tratarse de una incompatibilidad con el sistema que se desea cargar, revisar compatibilidad con Matlab 2018a.
- * Error 204: Problema al cargar los parámetros del sistema, revisar que estén correctos con el sistema (RNA).

- * Error 301: La carga de algún sistema difuso predefinido no se pudo realizar, intentar cargar otro diferente o cargar uno personalizado. También se puede presentar para el caso de las redes, realizar lo mismo.
- * Error 501: Incapaz de determinar el formato de la imagen, puede tratarse de que se intentó cargar un archivo que no era imagen o bien un formato no soportado, se recomienda cargar imágenes del tipo, jpg, png, bnp, volver a cargar nueva imagen con otro formato para determinar si era el problema.
- * Error 601: No hay imagen que guardar, primero debe cargar una imagen, la imagen que se guarda será aquella que es mostrada en el panel de Image.
- * Error 701: Incapaz de extraer las características, alguna característica fallo, puede deberse a la compatibilidad, tratar con una nueva imagen y si sigue persistiendo reinstalar el programa y seguir los pasos para poder extraer todas las características.
- * Error 801: Error al calcular la complejidad de la imagen con el sistema FIS, puede deberse a no tener las características necesarias, para esto debe ir edit settings y agregar las características que utiliza el sistema o bien también puede deberse a problemas con los parámetros del sistema, revisar los parámetros del sistema según a los indicado en el manual.
- * Error 802: Error al calcular la complejidad de la imagen con la RNA, puede deberse a no tener las características necesarias, para esto debe ir edit settings y agregar las características que utiliza el sistema o bien también puede deberse a problemas con los parámetros del sistema, revisar los parámetros del sistema según a los indicado en el manual.