



BENEMERITA UNIVERSIDAD AUTONOMA DE PUEBLA.

Facultad de ciencias de la computación.



“INSERCIÓN DE MARCA DE AGUA MEDIANTE UNA RED NEURONAL”

TESIS

QUE PARA OBTENER EL GRADO DE
LICENCIADO EN CIENCIAS DE LA COMPUTACIÓN

Presenta:

LIRA BARROETA DAVID

ASESOR:

Dr. JOSÉ ARTURO OLVERA LOPEZ

Puebla, Pue 2016

Resumen

Hoy en día vivimos en un mundo lleno de imágenes tanto en medio digital como de manera impresa, lo cual es información importante para su creador en este trabajo se buscó presentar e implementar técnicas de esteganografía en imágenes digitales, utilizando el método de la transformada discreta del coseno, el cual trabaja en el dominio de las frecuencias y el método de redes neuronales multicapa, el cual es un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso. La inserción de marca de agua utilizando redes neuronales consiste en insertar información en la imagen de manera digital mediante la realización de modificaciones sobre la misma. La ejecución de este proceso sobre la imagen es imperceptible para el ojo humano, pero no afectando su calidad.

Contenido

Capítulo I. Introducción	7
1.1 Motivación.	10
1.2 Objetivos generales y específicos del proyecto.	10
1.3 Inserción de marca de agua.	10
1.4 Requerimientos y aplicaciones de marca de agua digital.	11
1.5 Tipos de ataques.	13
1.6 Los objetivos fundamentales de las marcas de agua.	14
1.6.1 Evaluación de esquemas de marca de agua.	16
1.7 Estructura del documento.	
Capítulo II. Marco Teórico	
2 Introducción	19
2.1 Imagen digital	19
2.2 Esteganografía	20
2.3 Redes neuronales artificiales	22
2.3.1. El prototipo biológico	22
2.3.2. La neurona artificial	23
2.3.3. Arquitectura de una red neuronal	24
2.3.4. Redes Multicapa	25
2.3.5. El algoritmo retroalimentación	27
2.3.6. Redes neuronales de retroalimentación	27
2.3.7. Métodos de gradiente descendente	27
2.4 Transformada discreta del coseno	30
Capítulo III. Estado del campo o del arte	
3.1 Trabajos de investigación	33
3.2 Bit menos significativo (BMS)	34
3.2 Software propietario	36
3.2.1 Photo Helper 1.0	36
3.2.2 PicGhost	36
3.2.3 Icemark	36
3.2.4 AiS Watermarker ActiveX	37

Capítulo IV. Algoritmo de inserción de marcas de agua con redes neuronales

4.1	Algoritmo de inserción con redes neuronales	39
-----	---	----

Capítulo V. Resultados experimentales

5.1	Resultados experimentales	45
5.2	Marca de agua utilizando el método BMS	47
5.3	Marca de agua utilizando el método RN	47
5.4	Marca de agua extraída mediante RN y BMS después de utilizar el filtro gaussiano	50
5.5	Marca de agua extraída mediante RN y BMS después de utilizar el filtro negativo	51
5.6	Marca de agua extraída mediante RN y BMS después de aplicar ruido sal y pimienta	52
5.7	Experimentos con parámetros en la red neuronal	54
5.8	Marca de agua extraída mediante RN cambiando el número de épocas	55
5.9	Marca de agua extraída mediante RN cambiando el parámetro delta	55
5.10	Resultados obtenidos de proporción correcta de bits(PCR)	57

Capítulo VI. Conclusiones y trabajo a futuro

6.1	Conclusiones	61
6.2	Trabajo a futuro	61

Bibliografía	62
--------------	----

Agradecimientos

A mi madre María Yolanda Brígida Barroeta Hernández, Padre Jesus Hipólito Lira López y Hermana Carolina Lira Barroeta gracias por el apoyo acompañarme en esta carrera que hoy concluimos con mucho éxito. Me han dado todo lo que soy en este momento como persona, valores, determinación y coraje para conseguir mis objetivos.

Amigo Heber Alain Santos de la Rosa gracias por toda tu amistad es un regalo de dios para mi vida, gracias por ser mi amigo en las buenas y las malas.

A mi asesor Dr. José Arturo Olvera López por todo aquel apoyo que me ha brindado en mi carrera además de ser un gran profesor ha sido mi amigo, grandes consejos para lograr llegar al final de mi camino en esta facultad gracia.



Capítulo I. Introducción

1.1. Motivación

El desarrollo de las tecnologías multimedia hace de las imágenes datos de gran importancia en la actualidad. La facilidad para copiar y transmitir esta información ha impulsado el desarrollo de esquemas de marcas de agua, junto con algoritmos de redes neuronales, como estrategia para evitar o detectar algunas infracciones, por ejemplo, la distribución de copias ilegales.

Por esto mismo, una cuestión importante es la protección de los derechos de autor de todos los realizadores de dicha obra ya sea una imagen, audio o video, entre otras. Se ha reconocido, desde mucho tiempo, que las leyes actuales de derecho de autor son deficientes para hacerle frente. Para ello es necesario el desarrollo de nuevos mecanismos de disuasión, copia y protección.

Los esquemas de marca de agua describen técnicas, las cuales se utilizan para insertar información en materiales de forma imperceptible a los sentidos humanos. La facilidad para poder copiar y transmitir información digital ha impulsado el desarrollo de estos esquemas con marcas de agua como estrategia para evitar o detectar algunas infracciones.

Las marcas de agua algunas aplicaciones de las cuales se puede mencionar son las siguientes: transmitir claves de seguridad de algún *software*, es decir, transmitir una clave de usuario a través de una imagen ya sea del logotipo de la empresa o alguna imagen con poca característica; derechos de autor insertados dentro de una obra de arte como marca electrónica del autor (EMEA), la cual no degrada la imagen y protege al autor de cualquier falsificación y copia de su obra [1].

Una marca de agua se considera robusta si luego de muchas modificaciones o ataques como lo son la pérdida de compresión o transformaciones geométricas se extrae y es legible para el ojo humano. La marca de agua debe estar presente en las imágenes y a su vez después de las distorsiones. Por todo lo anterior, esta tesis propone la utilización de marcas de agua, las cuales se han desarrollado como una de las últimas tecnologías para la protección de los medios digitales.

Las técnicas basadas en dominio espacial han progresado en los últimos años y son utilizadas para marcar una imagen de forma eficaz. Sin embargo, siempre hay un compromiso entre la robustez y la imperceptibilidad. Por lo cual, esta investigación ofrece una técnica basado en la retropropagación para formar una imagen, así al entrenar los pesos con ciertos puntos de la imagen, esta retorne los más idóneos e inserte la marca de agua [2].

Para incrustar las marcas de agua en las imágenes contenedoras son imágenes de 8 bits por canal respecto al modelo de color RGB (figura 1).

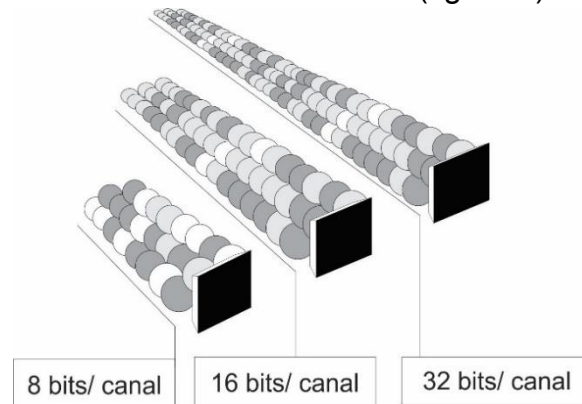


Figura 1.1: cada cuadrado negro representa un pixel con tres canales y cada uno expresa diferentes niveles de profundidad: 8, 16 y 32 bits por canal adaptado de [3].

La imagen a insertar se procesa en el dominio de las frecuencias mediante la transformación discreta del coseno (TDC), donde la forma de onda del coseno se utiliza para representar los datos originales transformando un dominio de amplitudes al dominio de la frecuencia. Los componentes frecuencia más altos son susceptibles de ser eliminados (figura 1.2). La TDC puede concentrar la energía de una imagen en la esquina superior izquierda de la imagen original. En el esquema de redes neuronales de retro propagación se ocupa para aprender la relación entre los coeficientes TDC.

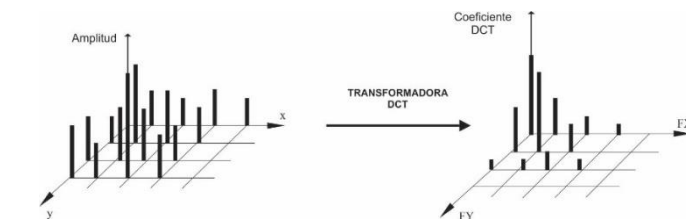


Figura 1.2: transformada discreta del coseno (DTC) (adaptado de [4]).

La marca de agua utiliza una red neuronal multicapa para insertarse, a la cual ingresan una matriz de píxeles 8x8 de la imagen a incluir en la marca de agua. Ésta es obtenida mediante un recorrido *zigzag* (figura 1.3) de los valores resultantes con la transformación discreta del coseno y junto con una función de activación tipo sigmoides¹. Se suele utilizar en el entrenamiento de la red, el cual emite un rango de número real de 0 a 1, la cual después de ocupar la función de activación obtiene el píxel óptimo para lograr hacer la inserción de una parte de la marca de agua.

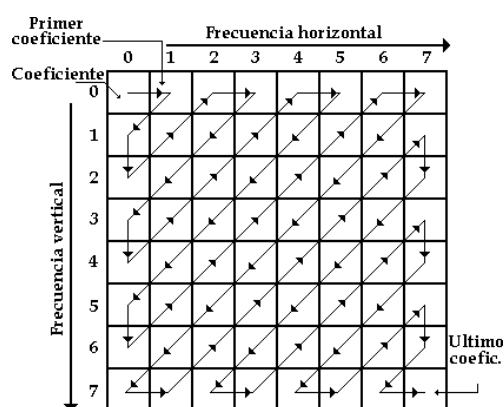


Figura 1.3: recorrido en zigzag para una matriz de 8x8 (adaptado de [6]).

Una tarea común, después de insertar una marca de agua, es la extracción, o aislamiento de la marca de agua para fines de verificación de la autenticación de una imagen. Los coeficientes originales se deben extraer para la marca de agua insertada. También se deben recuperar aproximados utilizando pesos y umbrales, mediante los primeros ajustados en el entrenamiento de la red neuronal.

El procedimiento de extracción de la marca de agua en la imagen es similar al procedimiento de incrustación, pero llevado a cabo de manera inversa.

Después de aplicar el método de inserción de marcas de agua, se lleva a cabo un análisis visual del resultado, así como el uso de una métrica de *bíts* por píxel modificado para cuantificar el resultado de la inserción de la marca de agua. Con esto se busca la inserción de una marca de agua y la alteración visual mínima de la

¹ Se trata de una función continua no lineal la cual posee un rango comprendido entre 0 y 1. Esto, aplicado a las unidades de proceso de una red neuronal artificial significa que, sea cual sea la entrada, la salida estará comprendida entre 0 y 1.

misma después de la inserción, con el objetivo de hacer a la imagen con marca de agua, siendo esta imperceptible al ojo del usuario [7].

1.2. Objetivos generales y específicos del proyecto.

Objetivo general

Desarrollar una herramienta, la cual permita insertar marcas de agua en imágenes digitales mediante redes neuronales de retroalimentación.

Objetivos particulares

1. Analizar y diseñar una topología de una red neuronal multicapa para la inserción de marcas de agua en una imagen digital.
2. Transformar contenido de imágenes al espacio de las frecuencias utilizando la TDC.
3. Generar en este espacio los índices a insertar auxiliándose del entrenamiento de una red neuronal multicapa.
4. Analizar resultados obtenidos con la marca de agua insertada mediante métrica de error cuadrático respecto a la diferencia de imágenes.

1.3 Inserción de marca de agua

Una marca de agua digital consta de datos incrustados en fotografías, películas, canciones y otros contenidos digitales. Hay software especializado para inserta la marca de agua esta puede ser texto o imagen sutilmente oculto de forma tal que la imagen que contiene la marca no sufre cambios sobre los datos originales y permitiendo de esta manera que no sea detectada y posteriormente ser extraída. Esta operación puede facilitar transacciones comerciales, autenticación de materiales impresos, acceso a contenidos multimedia, etc. Los mecanismos o aplicaciones que reproducen o acceden a dicho contenido pueden leer los datos incrustados cuando disponen del software de detección [8].

1.4 Requerimientos y aplicaciones de marca de agua digital

Entre los principales requerimientos de marca de agua se encuentran.

La robustez significa resistencia frente a aquellas modificaciones producidas por las operaciones comunes a las que están expuestos los archivos multimedia. Los archivos digitales de imágenes, audio y video, están expuestos a muchos tipos de modificaciones o ataques como son las pérdidas por compresión, los cambios producidos por el mejoramiento de imágenes, transformaciones geométricas, (recorte, rotación, escalamiento, traslación), etc. Una marca de agua se considera robusta si se conserva después de esos ataques. Esto quiere decir que la marca debe estar presente en los archivos y debe ser detectada después de las distorsiones.

La Imperceptibilidad o transparencia de la marca de agua en imágenes tiene como base el comportamiento del sistema visual humano. Una marca de agua es imperceptible, si la degradación que causa en los archivos donde se ha insertado es muy difícil de apreciar. En algunas aplicaciones se puede aceptar una pequeña degradación de los datos, a cambio de lograr mayor robustez o menor costo del sistema.

La carga útil de marca de agua hace referencia a la cantidad de información que puede contener una marca de agua, y del tipo de aplicación.

La no ambigüedad en la mayoría de los sistemas de marcas de agua es muy importante distinguir entre los archivos que contienen una marca y los que no. La probabilidad de error al detectar una marca debe ser muy pequeña. Se denomina probabilidad de falso negativo a la probabilidad de que, habiendo estado presente una marca en determinado archivo, el detector asuma que no hay tal marca.

Por otro lado, la probabilidad de falso positivo es la probabilidad de que, no estando la marca presente en un archivo, el detector asuma que la marca está presente.

Este requisito debe ser tomado muy en cuenta, si se quiere que el sistema pueda ser utilizado en disputas legales. La dependencia mutua entre los requerimientos básicos de marca de agua digital, se muestra en la (figura 1.4) [9].

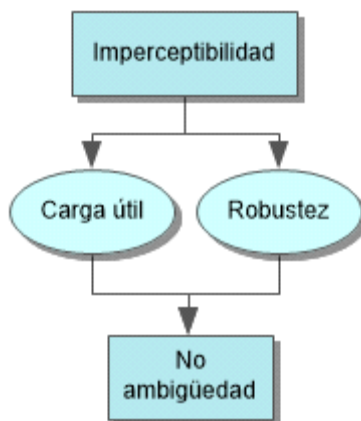


Figura 1.4: Dependencia mutua entre requerimientos básicos de marca de agua.

Entre las principales aplicaciones de marca de agua digitales se encuentran las que a continuación se describen. La protección: la protección de derechos de autor consiste en insertar marcas de agua que contengan información sobre el autor o propietario dentro de imágenes digitales. De esta manera se podrían resolver litigios legales relacionados con infracciones a los derechos de autor. Para esta aplicación se requiere una excelente robustez contra una gran cantidad de ataques, intencionales y no intencionales. Además, es necesario que los esquemas de este tipo utilicen medidas de seguridad para evitar inserciones no autorizadas en imágenes protegidas.

El rastreo de copias es un tipo de aplicación que contempla el hecho de insertar una marca de agua diferente en cada una de las copias de una determinada imagen, antes de distribuirlas entre un reducido círculo de personas. Si alguna persona redistribuye la imagen sin autorización, es posible rastrear al responsable de la infracción detectando la marca de agua de alguna de las copias ilegales, conocido como toma de huellas dactilares. Puesto que la persona que va a compartir la imagen podría saber de la existencia de la marca de agua en su imagen, este tipo

de sistemas deben evitar inserciones no autorizadas, remoción de las marcas de agua utilizando otras copias de la imagen conocido como ataque de colusión [10].

La protección de copias en algunos sistemas multimedia necesita algún mecanismo de seguridad para evitar la reproducción de copias no autorizadas. Es viable utilizar marcas de agua para ofrecer esta característica dentro de los mismos dispositivos electrónicos, sobre todo en sistemas cerrados o propietarios. Las marcas de agua, insertadas dentro de los datos multimedia, indican a los reproductores el número de copias que están autorizadas de un determinado volumen. Por ejemplo, las marcas de agua utilizadas en los Discos Versátiles Digitales (DVD), indican a los dispositivos electrónicos si se permite o no la copia de ese volumen. Generalmente, se requiere que este tipo de esquemas sean robustos contra cambios geométricos y de formato, adición de ruido, etc.

En la autenticación de imágenes, el objetivo es detectar modificaciones en los datos. Consiste en diseñar marcas de agua que se destruyan o modifiquen cuando las imágenes protegidas sufren cualquier tipo de distorsión. Los esquemas diseñados con esta aplicación son conocidos como de marcas de agua frágiles.

Algunas aplicaciones requieren marcas de agua que se destruyan con distorsiones que afecten severamente la imagen protegida, pero que sobrevivan a ciertos ataques accidentales como compresión, cambio de formato, etc. Los esquemas que se han diseñado con estas características se conocen como de marcas de agua semi-frágiles [7].

1.5 Tipos de ataques

La marca de agua ideal debería ser capaz de resistir cualquier tipo de distorsión, antes de que la calidad de la imagen marcada se degrade hasta perder su utilidad. Por la complejidad que requiere el diseño de tal esquema, usualmente se diseñan marcas de agua robustas a distorsiones relacionadas con aplicaciones específicas. Por ejemplo, si se desea utilizar marcas de agua en datos que se van a transmitir, sería necesario que dichas marcas fueran robustas a las distorsiones causadas por

ruido aditivo y multiplicativo, corrección de color y contraste, conversiones, entre otros.

Algunas de las distorsiones más comunes son:

1. Mejoramiento de la imagen (corrección de nitidez, color, contraste y gamma).
2. Ruido aditivo y multiplicativo.
3. Compresión (como compresión JPEG).
4. Transformaciones afines globales y locales (traslación, rotación, escalamiento, recorte de imagen).
5. Reducción de la imagen (recorte, modificación del histograma).
6. Composición de la imagen (inserción de logo, composición de la escena).
7. Cambio de formato (JPEG – BMP).
8. Inserción de varias marcas de agua.

Una vez que la marca de agua ha sido insertada en una imagen digital, es susceptible a una gran variedad de ataques que la distorsionan, así como degradar la calidad del material marcado. Según la causa y objetivo que los origina, éstos se pueden agrupar en ataques no intencionados e intencionados.

Las distorsiones que se aplican a una imagen protegida sin la intención de alterar la marca de agua se conocen como ataques no intencionados. Dependiendo de la aplicación, los ataques no intencionales incluyen la compresión, cambio de formato, corrección de la imagen, etc. Los ataques intencionados son las manipulaciones que realiza un intruso sobre la imagen marcada con el fin de eliminar la protección de derechos de autor [11].

1.6 Los objetivos fundamentales de las marcas de agua

Primero manejar información libre de cualquier marca de agua. Esto con el objetivo manipular la imagen para distorsionar la marca de agua insertada en la imagen, de manera que está sea indetectable. Por lo tanto, la marca de agua debe ser más resistente que la misma imagen, esto es, que la degradación de la imagen debe

ocurrir mucho antes que la marca de agua se distorsione y como consecuencia la marca sea indetectable.

El segundo objetivo es el de insertar su propia marca de agua. Este punto tiene como objetivo confundir en la determinación de la propiedad de los derechos de autor. La finalidad del agresor es insertar su propia marca de agua para reclamar que él es propietario de la imagen.

Por lo que se comprueba que la robustez no es suficiente para garantizar una adecuada protección de los derechos de autor.

En los últimos años se ha incrementado el interés de la comunidad científica por el establecimiento de las definiciones preliminares de los requisitos que deben cumplir los esquemas de marca de agua eficaces con vistas a su estandarización. Sin embargo, aún queda mucho por hacer antes de que pueda hablarse de la adopción definitiva de un estándar para algunas aplicaciones, en el caso de las imágenes, la música y el video se han experimentado diversas iniciativas con resultados positivos. Es aún largo el camino por recorrer dentro del campo, sin embargo, se continúa trabajando al respecto en investigaciones como las presentadas en esta tesis, teniendo en mente como idea principal el presentar nuevas propuestas que ayuden al desarrollo del campo.

Cuando se inserta archivos gráficos (imágenes) se aprovecha los bits menos significativos de los colores RGB para introducir en ellos la información, con lo que se hace una reducción de colores respecto a la imagen original si es necesario.

En este trabajo se propone un esquema de marcas de agua digitales que incorpora la marca de agua en el dominio de la frecuencia.

En nuestro esquema, se utiliza una red neuronal de retropropagación mejor ilustrada en la (figura 1.5). Para mejorar la seguridad y la solidez de la inserción de la imagen con marcas de agua

Es importante que la imagen que contenga la marca de agua no abarque características para su detección, ya que se modificará por la inserción de una

imagen o mensaje, de esta forma pueden ser detectadas por vigilancia electrónica utilizando técnicas de estegoanálisis para detectar las imágenes que contienen mensajes ocultos [12].

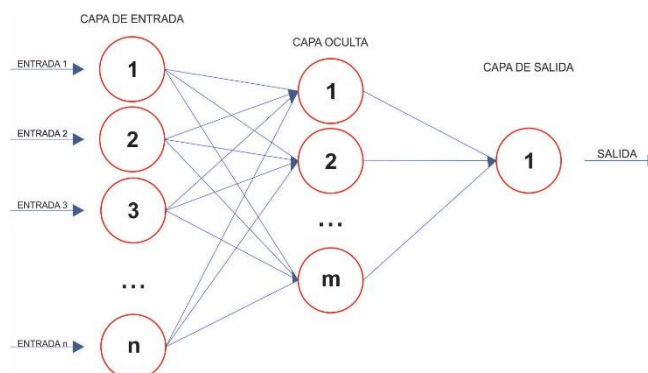


Figura 1.5: Red neuronal retro propagación (adaptado de [13]).

El algoritmo de inserción de marcas de agua utilizando redes neuronales establece la relación entre los píxeles de la imagen, seleccionados a través de números primos grandes, la transformada discreta del coseno y el entrenamiento a través de una red neuronal multicapa de retro alimentación. Para realizar el proceso de recuperación de la marca de agua se establece nuevamente la relación entre los píxeles con los números primos ya seleccionados anteriormente para poder formar los bloques de 7 píxeles además de la transformación discreta del coseno y así poder eliminar aquellos píxeles repetidos. Con el entrenamiento de la red neuronal a partir de los pesos sinápticos de las capas ocultas anteriores se obtienen nuevamente los píxeles óptimos para la recuperación de la marca de agua.

1.6.1. Evaluación de esquemas de marca de agua

Para evaluar apropiadamente los sistemas de marcas de agua se sugiere que todos los esquemas se evalúen bajo condiciones equiparables (figura 1.6) Como respuesta a la necesidad de establecer puntos de comparación para evaluar la robustez de los esquemas de marca de agua, se han propuesto métodos estándar que aplican diversas distorsiones en las imágenes marcadas. Las metodologías de evaluación aplican ataques que van desde: recortes, rotación, escalamiento, contaminación de imagen con ruido, distorsiones geométricas aleatorias, eliminación de líneas, compresión JPEG, entre otros.

En la actualidad, las técnicas de marcas de agua presentan robustez ante diversos ataques. Dando como resultado una recuperación de la marca de agua. Sin embargo, las técnicas de marca de agua deben mejorar la robustez ante los ataques. En el caso de las imágenes estos ataques pueden realizarse mediante rotación, traslación de un número entero a píxeles y cambios de escala.

En los últimos años se ha incrementado el interés por establecer la inserción de marcas de agua en los medios digitales. Sin embargo, aún queda mucho por hacer antes de que pueda hablarse de la adopción definitiva de una estándar para algunas aplicaciones en el caso de las imágenes [14].

1.7 Estructura del documento.

El resto del presente trabajo está organizado de la siguiente manera, el capítulo 2 presenta la teoría acerca de las imágenes digitales, redes neuronales y marcas de agua digital y como están relacionadas de tal manera que se puede obtener una imagen digital con marca de agua insertada.

En el capítulo 3 estado campo del arte se describen las diferentes investigaciones realizadas con diferentes métodos con la inserción de marcas de agua y el software propietario que de igual forma parte de las investigaciones realizadas.

En el capítulo 4 se explica el método que conjunta la teoría de los capítulos 2 y 3.

En el capítulo 5 se muestran los resultados obtenidos con luego de aplicar algunos filtros conceptuales.



Capítulo II.

Marco Teórico

2.1 Introducción

Existe software propietario para la inserción de marcas de agua digital visible, pero no para protección de documentos, el cual abarca desde protecciones contra edición, hasta la adición de marcas de agua visible. Sin embargo, estos esquemas, o documentos, al ser visibles le indican al usuario que este se encuentra protegido.

Una leyenda en color puede resultar en el caso de las imágenes, o verse sobre las mismas, con una sublimación. Esto permite buscar rutas, las cuales admitan una copia del documento ya sin protección, obteniendo así desde una simple selección del texto y pegado en un nuevo documento, hasta técnicas más sofisticadas para eliminar la protección. En este trabajo de tesis se aplicarán esquemas, los cuales permiten proteger imágenes, ocupando una red neuronal con retroalimentación con retro propagación, utilizando este esquema para ocultar la existencia de la protección del documento ante el ojo humano.

Por otro lado, existe una amplia gama de trabajos sobre marca de agua digital aplicada a documentos ya sean texto, imágenes o audio. La mayoría de estas investigaciones están basadas en la transformación de los documentos de texto a imágenes, para ser aplicadas a nuevas y a diversos esquemas, los cuales buscan implementar esquemas de protección, pero siempre sobre las primeras.

2.2 Imagen digital

El término se refiere a una función bidimensional de la luz y la intensidad, a la que indicamos por $f(x, y)$ donde el valor o amplitud de f en las coordenadas espaciales (x, y) da la intensidad (iluminación) de la imagen en este punto.

Para ser utilizable el procesamiento por medio de computadora, una imagen $f(x, y)$ debe ser digitalizada de manera especial como en su amplitud. La digitalización de las coordenadas espaciales (x, y) se denomina muestreo de la imagen y la digitalización de la amplitud se conoce bajo el nombre de cuantificación.

Una imagen continua $f(x, y)$ se describe de forma aproximada por una serie de muestras iguales especiales representadas mediante una matriz $M \times N$ donde M representa el número de filas N el número de columna, así como lo indica la notación siguiente:

$$f(x, y) = \begin{pmatrix} f(0, 0) & f(0, 1) & \dots & f(0, N) \\ f(1, 0) & f(1, 1) & \dots & f(1, N) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1, 0) & \dots & \dots & f(M-1, N-1) \end{pmatrix}. \quad (2.1)$$

Esto se puede explicar mejor de la siguiente manera. A cada el elemento de la matriz se le denomina elemento de la imagen o pixel. Los términos imagen y pixeles van a emplearse para indicar una imagen digital y sus elementos. Cuando se tiene una imagen digital representada en el modelo RGB (*Red, Green, Blue*) (figura 2.7), cada pixel puede adoptar algún valor dependiendo de la profundidad de la imagen en bits. Si la profundidad es de 24 bits, cada canal de color tiene 8 bits para controlar cada color, por lo tanto, los colores que podemos obtener son $256 \times 256 \times 256 = 16,777,216$ con esta cantidad de colores es suficiente para que el ojo humano percibía los colores de las imágenes definidos en el rango visible del espectro electromagnético [9].



Figura 2.7 Modelo rgb de la imagen de Lena (adaptado de [15]).

Una imagen a niveles de gris es una función bidimensional de la intensidad de la luz y cuyos valores se han obtenido muestreando la intensidad en la que el valor de cada pixel posee un valor equivalente [16].

2.3 Esteganografía

Proveniente del griego *steganos* (oculto) y *graphos* (escritura). Se puede definir, por lo anterior, como la ocultación de información en un canal encubierto con el propósito de prevenir la detección de un mensaje oculto.

La esteganografía de la imagen define cómo ocultar un mensaje secreto dentro de una manera tal que otros no puedan discernir la presencia o el contenido del mensaje oculto. El concepto de esteganografía se usa de manera muy frecuente, sin embargo, el término más amplio es “información oculta”, en la (figura 2.9) se muestra diagrama que permite dar una idea general de la relación entre las esteganografía y las marcas de agua [16].

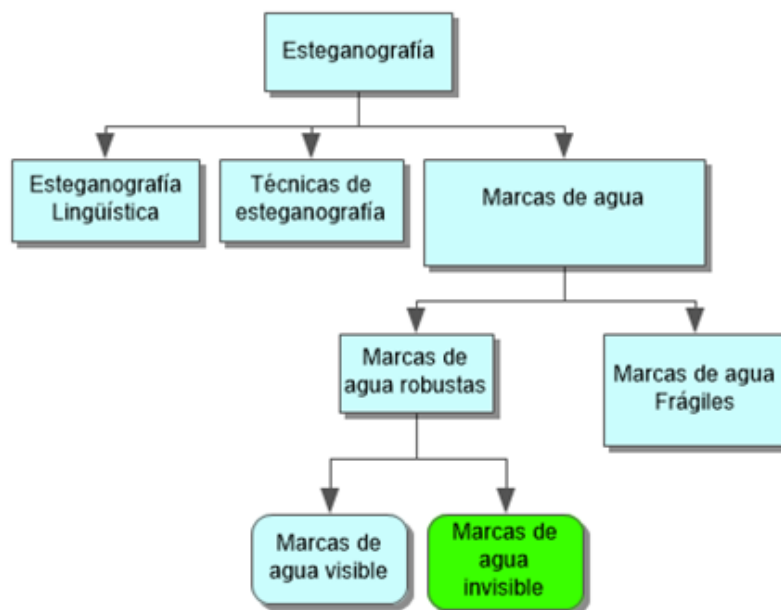


Figura 2.8 Técnicas estenográficas, inserción de marcas de agua (adaptado de [18]).

2.4 Redes neuronales artificiales

La inteligencia artificial es la disciplina que estudia la forma de diseñar procesos que exhiban características que comúnmente se asocian con el comportamiento humano inteligente [19]. La inteligencia artificial sintetiza y automatiza tareas intelectuales y es, por lo tanto, potencialmente relevante para cualquier ámbito de

la actividad intelectual humana [20]. En la actualidad esta ciencia está comprendida por varios subcampos, los cuales van desde áreas de propósito general, aprendizaje y la percepción, a otras más específicas como la demostración de teoremas matemáticos, diagnóstico de enfermedades, entre otros. Unos de los modelos que ha surgido para emular el proceso de aprendizaje es la red neuronal artificial. Las redes neuronales son modelos que intentan reproducir el comportamiento del cerebro humano. Una red neuronal consiste en un conjunto de elementos de procesamiento, llamados neuronas, los cuales se conectan entre si [21].

La organización y disposición de las neuronas dentro de una red neuronal se denomina topología, y viene dada por el número de capas, la cantidad de neuronas por capa, el grado de conectividad, y el tipo de conexión entre neuronas.

Una vez determinada la topología de la red neuronal es necesario entrenarla. En la etapa de entrenamiento la red es capaz de aprender relaciones complejas entre entradas y salidas mediante el ajuste de los pesos de las conexiones entre neuronas. Identifican una cantidad significativa de algoritmos de entrenamiento [22].

Por otro lado, dice que la mayoría de éstos utilizan información del gradiente² de una función de error para ajustar los pesos de las conexiones, y se llaman algoritmos de gradiente descendente [23].

Sin embargo, los algoritmos de gradiente descendente poseen dos características: primero, suelen quedar atrapados en mínimos locales, generándose de esta manera estimaciones subóptimas de los pesos. Segundo, suelen ser muy lentos por utilizar pasos infinitesimales para alcanzar la solución.

²El método del gradiente descendente se basa en buscar la dirección en la que una pequeña variación del vector de pesos hace que el error decrezca más rápidamente.

2.4.1. El prototipo biológico

Las diferentes configuraciones y algoritmos diseñados para las redes neuronales artificiales están inspirados en la organización del sistema neuronal del cerebro humano. No obstante, se debe aclarar que esta inspiración no supone que las *redes neuronales artificiales (RNA)* (figura 2.9) lleguen a emular al cerebro como algunos optimistas lo desean, pues entre otras limitaciones el conocimiento sobre el modo de funcionamiento y comportamiento del cerebro es bastante simple y reducido. De hecho, los diseñadores de redes artificiales van más lejos del conocimiento biológico actual y prueban nuevas estructuras que presentan un comportamiento adecuado y útil. El sistema nervioso humano constituido por células llamadas neuronas presenta una estructura muy compleja. El número estimado de neuronas es de 10^{11} y las interconexiones entre ellas son del orden de 10^{15} .

Cada neurona comparte muchas características con otras células del cuerpo humano, pero tiene propiedades particulares y especiales para recibir, procesar y transmitir señales electroquímicas a través de todas las interconexiones del sistema de comunicación del cerebro [24].

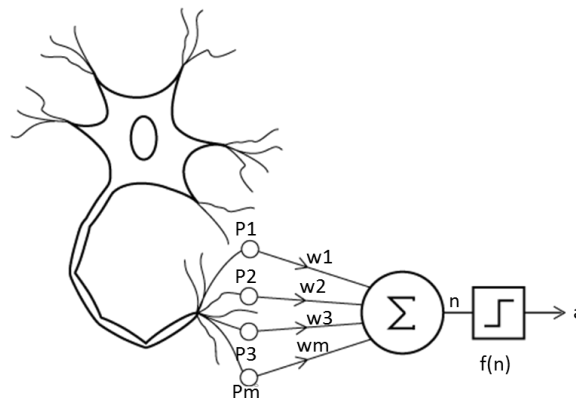


Figura 2.9: neurona prototipo biológico y su representación en la red neuronal artificial (adaptado de [25]).

2.4.2. La neurona artificial

El término tiene su origen en los intentos para encontrar representaciones matemáticas de procesamiento de la información en los sistemas biológicos (figura 2.3). Es decir, la neurona artificial fue diseñada para ‘emular’ las características del

funcionamiento básico de la neurona biológica. En esencia, se aplica un conjunto de entradas a la neurona, donde una representa una salida de otra neurona. Cada entrada se multiplica por su ‘peso’ o ponderación correspondiente análoga al grado de conexión de la sinapsis. Todas las entradas ponderadas se suman y se determina el nivel de excitación o activación de la neurona. Una representación vectorial del funcionamiento básico de una neurona artificial se indica según la siguiente expresión de la (ecuación 1.2) [27]:

donde $NET \in \mathbb{R}$ y expresa el resultado final, $X \in \mathbb{R}$ e identifica al vector de entrada y $W \in \mathbb{R}$ es el vector de pesos.

Regularmente la señal NET suele procesar una función de activación F la cual es una salida de la neurona. La función F puede ser una función lineal, una función umbral o una función no lineal que aparenta con mayor exactitud las características de transferencia no lineales de las neuronas biológicas.

En las redes neuronales artificiales, la unidad análoga a la neurona biológica es el elemento procesador (EP). Un elemento procesador tiene varias entradas y las combina, normalmente con una suma básica. La suma de las entradas es

$$NET = X \times W, \quad (2.2)$$

modificada por una función de transferencia y el valor de la salida de esta función de transferencia se pasa directamente a la salida del elemento procesador. La salida del EP se puede conectar a las entradas de otras neuronas artificiales EP mediante conexiones contenidas correspondientes [28].

2.4.3. Arquitectura de una red neuronal

Una vez definida el tipo de neurona a ser utilizadas en un modelo de redes neuronales artificiales es necesario definir la topología de la misma. A la organización y disposición de las neuronas dentro de una red neuronal se denomina topología peso sináptico, y son direccionales.

La topología de la red viene dada por el número de capas, la cantidad de neuronas por capa, el grado de conectividad, y el tipo de conexión entre neuronas. Suelen agruparse en unidades funcionales denominadas capas. Ahora, de manera análoga, se denomina capa oculta y capa de salida a aquellas que están compuestas por neuronas ocultas y de salida respectivamente. Una red neuronal artificial está compuesta por una o más capas, las cuales se encuentran interconectadas entre sí. Entre un par de neuronas de la red neuronal artificial pueden existir conexiones, estas conexiones son las sinapsis, tienen asociadas [29].

Se debe aclarar que cuando la conexión se establece entre dos neuronas de una misma capa, se habla de conexiones laterales o conexiones intra capa.

Por el contrario, si la conexión se establece entre neuronas de distintas capas se la denomina conexión inter-capas. Si la conexión se produce en el sentido inverso al de entrada-salida la conexión se llama recurrente o realimentada.

Una red puede estar formada por una única capa de neuronas. En este caso hablamos de redes mono capa, y las neuronas que conforman dicha capa cumplen la función de neuronas de entrada y salida de forma simultánea. Cuando la red está compuesta por dos o más capas hablamos de redes multicapa. Las redes de retroalimentación (figura 2.10) pueden tener señales que viajan en ambas direcciones mediante la introducción de los bucles en la red. Las redes de realimentación son muy potentes, permanecen en el punto de equilibrio hasta que cambie la entrada y un nuevo equilibrio tiene que ser encontrado. Las arquitecturas de realimentación también se conocen como interactivas o recurrentes, aunque este último término a menudo se utiliza para denotar conexiones de realimentación en las organizaciones de una sola capa [30].

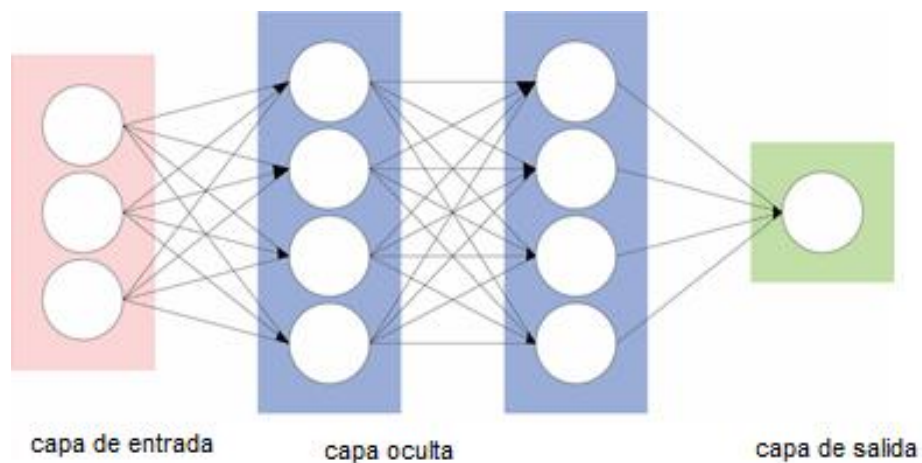


Figura 2.10: red neuronal de retro alimentación.

2.4.4.Redes Multicapa

Las redes multicapa son aquellas que disponen de un conjunto de neuronas jerarquizadas en distintos niveles o capas y son especialmente útiles en aplicaciones de reconocimiento o clasificación de patrones [31].

Las redes multicapa tienen una capa de entrada y otra de salida como se muestra en la (figura 2.11). Las redes multicapa están formadas por varias capas de neuronas. Estas redes se pueden clasificar atendiendo a la manera en que se conectan. Usualmente, las capas están ordenadas por la disposición en que reciben la señal desde la entrada hasta la salida este tipo de conexiones se denominan conexiones hacia adelante, o retroalimentación, aunque pueden existir conexiones a neuronas de cualquiera de las capas anteriores como conexiones recurrentes de alimentación hacia atrás. Dentro de estos modelos los más conocidos son: la red de memoria adaptativa lineal o memoria lineal adaptativo (MLA); la red refuerzo unidad y la red perceptrón multicapa o también llamada red propagación hacia atrás.

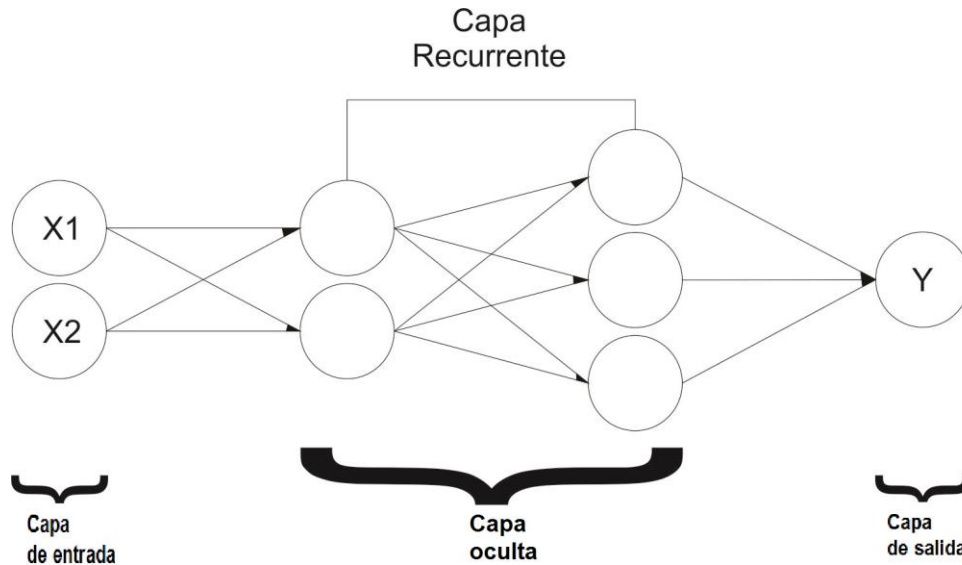


Figura 2.11: Red neuronal con tres capas ocultas y conexiones de alimentación hacia adelante (adaptado de [30]).

Dentro de los principales modelos con conexiones alimentación hacia adelante y recurrentes, se encuentran el cognitron modelo matemático hipotético del sistema perceptivo humano, neocognitron se orienta hacia el modelado del sistema visual humano, la red *TRA* (*Teoría de la resonancia adaptativa*) mostrada en la red *MAB* (*memoria asociativa bidireccional*) y las máquinas multicapa de Boltzman y Cauchy [31] como se muestra en la figura 2.12.

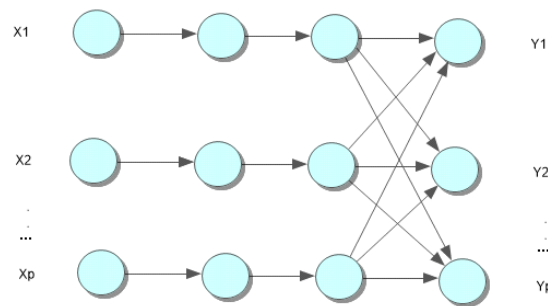


Figura 2.12: Red multicapa con diferentes tipos de conexiones (adaptado de [31]).

2.4.5. El algoritmo retroalimentación

En [33] describe al algoritmo retroalimentación, el cual es el método de entrenamiento más utilizado en redes con conexión hacia adelante. Es un método de aprendizaje supervisado de gradiente descendente, en el que se distinguen dos

fases: primero se aplica un patrón de entrada, el cual se propaga por las distintas capas que componen la red hasta producir la salida de la misma. Segundo, la salida se compara con la salida deseada y se calcula el error cometido por cada neurona de salida. Estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de las capas intermedias.

2.4.6. Redes neuronales de retroalimentación

Las redes neuronales fueron introducidas durante los primeros años del desarrollo de la tecnología, [33] aplicó una red neuronal retroalimentación con el fin de clasificar la litología mediante registros de pozos. La configuración de la red fue de tres neuronas en la capa de entrada, correspondientes a los registros de rayos *gamma*, además de la densidad y porosidad de neutrón. La capa oculta constaba de tres y cuatro neuronas y la capa de salida consta de cuatro neuronas que corresponden a la clasificación de cuatro litofacies: caliza, dolomita, lutita y arenisca.

Una red de retroalimentación es una red que consiste en varias capas o unidades todos los cuales son de adaptación. La red no tiene ciclos de ahí el nombre de retroalimentación. La red se compone de algunos nodos de entrada, algunos nodos de salida y un conjunto de nodos ocultos. Cada nodo oculto toma entradas de cada uno de los nodos de entrada y se introduce en cada uno de los nodos de salida, siguiendo la notación de Russell y Norving.

2.4.7. Métodos de gradiente descendente

Una vez seleccionada el tipo de neurona artificial que se utilizará en una red neuronal y determinada su topología es necesario entrenarla para que la red pueda ser utilizada. Partiendo de un conjunto de pesos sinápticos aleatorios, el proceso de aprendizaje busca un conjunto de pesos que permitan a la red desarrollar correctamente una determinada tarea. Durante el proceso de aprendizaje se va refinando iterativamente la solución hasta alcanzar un nivel de operación suficientemente bueno [35].

El método de entrenamiento más utilizado es el método del gradiente descendente; este método define una función $E(W)$ que proporciona el error que comete la red en función del conjunto de pesos sinápticos W . El objetivo del aprendizaje es encontrar la configuración de pesos que corresponda al mínimo global de la función de error, aunque en muchos casos es suficiente encontrar un mínimo local lo suficientemente bueno [32].

El principio general del método es el siguiente: dada un conjunto de pesos $W(0)$ para el instante de tiempo $t=0$, se calcula la dirección de máxima variación del error. La dirección de máximo crecimiento de la función $E(W)$ en $W(0)$ viene dado por el gradiente $\nabla E(W)$, luego, se actualizan los pesos siguiendo el sentido contrario al indicado por el gradiente $\nabla E(W)$, dirección que indica el sentido de máximo decrecimiento [36]. De este modo se va produciendo un descenso por la superficie de error hasta alcanzar un mínimo local dado por la (ecuación 2.3).

$$W(t + 1) = W(t) - \alpha \nabla E(W), \quad (2.3)$$

en la ecuación 2.3 $\alpha \in \mathbb{R}$ y expresa el momentum. Esta constante hace más estable la actualización de cada peso al tomar en cuenta el valor del gradiente de la iteración anterior, $t \in \mathbb{R}$.

El momentum es una manera de convergencia del error hacia el mínimo de la función. Consiste en sumar al cambio de pesos que se llevan a cabo en la fase de actualización, el valor de variación, más una fracción del valor.

Si se tiene dos gradientes consecutivos con el mismo signo entonces los pesos serán incrementados (según sea el caso) más rápidamente, si se tienen dos gradientes consecutivos con diferente signo entonces los pesos serán incrementados o decrecidos muy poco.

La inclusión de α hace que en las regiones donde el signo del gradiente oscila, los pesos se actualicen poco en magnitud con el fin de evitar una inestabilidad mayor, y en las regiones donde el signo es el mismo, los pesos se actualicen más rápido en una misma dirección que puede ser la correcta.

Puede haber más de una capa oculta, de acuerdo con la necesidad práctica. Dos unidades de cada capa adyacente están conectadas directamente entre sí, el cual es un enlace. El enlace tiene un valor ponderado, lo que representa el grado de relación entre las dos unidades. Cada vector de entrada tiene su propio vector de salida deseado. Los detalles del algoritmo de entrenamiento se describen mediante la siguiente ecuación:

$$net_j(t) = \sum_{i=0, j=0}^t w_{ij} O_i(t) - \theta_j, \quad O_j(t+1) = f_{act}(net_j(t)), \quad (2.4)$$

dónde $j \in \mathbb{R}$ expresa el índice para la unidad actual, $i \in \mathbb{R}$ es el índice de un precursor de la unidad j , $f_{act}()$ es la función de activación, $net_j(t)$ identifica la activación de la unidad j , $\theta_j \in \mathbb{R}$ es el umbral o sesgo de la unidad j , $w_{ij} \in \mathbb{R}$ es el peso del enlace desde la unidad i hasta la unidad j y $O_j(t) \in \mathbb{R}$ expresa la salida de la unidad j en t iterado [37].

Queremos encontrar w con el fin de minimizar el error y seleccionar $E(w)$

$$\nabla E(w) = \left[\frac{\partial E}{\partial w_0}, \frac{\partial E}{\partial w_1}, \dots, \frac{\partial E}{\partial w_n} \right]. \quad (2.5)$$

Un vector derivado se le llama “gradiente”: $\nabla E(w)$ cambiando cada peso:

$$W_j = W_j + \Delta W_j \quad (2.6)$$

donde:

$$\Delta W_k = -\eta \frac{\partial E}{\partial w_k}. \quad (2.7)$$

y η es la tasa de aprendizaje.

La función de error es:

$$E(W) = \frac{1}{2} \sum_{k=0}^w (t^k - o^k)^2. \quad (2.8)$$

2.5 Transformada discreta del coseno

La transformada discreta del coseno TDC es una transformada basada en la Transformada Discreta de Fourier TDF, pero utilizando únicamente coeficientes reales. Para una imagen de tamaño $M \times N$, la TDC para dos dimensiones está dada por:

$$F(u, v) = \left(\frac{2}{M}\right)^{\frac{1}{2}} \left(\frac{2}{N}\right)^{\frac{1}{2}} C_u C_v \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i, j) \cos\left(\frac{(2i+1)u\pi}{2M}\right) \cos\left(\frac{(2j+1)v\pi}{2N}\right), \quad (2.9)$$

donde $i, j \in \mathbb{R}$ son los índices de la imagen de tamaño $m \times n$ y, la función $F(u, v)$: identifica el valor del coeficiente TDC de la fila u , columna v y $f(i, j)$ es el valor de intensidad del píxel en la fila i y columna j .

Los términos C_u y $C_v \in \mathbb{R}$ se definen como:

$$C_u = \frac{1}{\sqrt{2}} \text{ para } u = 0 \quad (2.10)$$

y

$$C_v = 1 \text{ para cualquier otro caso.} \quad (2.11)$$

La transformada discreta del coseno expresa una secuencia finita de varios puntos como resultado de la suma de distintas señales sinusoidales (con distintas frecuencias y amplitudes). Como la transformada discreta de Fourier (TDF) la TDC trabaja con una serie de números finitos, pero mientras la TDC trabaja con cósenos, la TDF lo hace con exponenciales complejas [38].

Esta transformada es utilizada como base en la compresión de tipo JPEG y el tamaño de bloque utilizado en la compresión tiene dimensiones de 8×8 . Al realizar la transformada, se busca convertir la mayor cantidad posible de valores menos significativos en ceros, entendiendo por estos valores como aquellos de los cuales una imagen puede prescindir sin perder el significado y el tamaño del archivo se logra reducir, aunque de cierta manera se llega a sacrificar la calidad de la imagen.

Por otro lado, la TDC permite separar la imagen en bandas de frecuencia, conocidas como frecuencias altas, medias y bajas. En una imagen, las frecuencias bajas tienen

una mayor probabilidad de ocurrencia en comparación con las altas frecuencias, así, al realizar la transformación de una imagen al dominio de la frecuencia, se pueden quitar una gran cantidad de coeficientes de alta frecuencia con la finalidad de reducir la cantidad de datos necesarios para describir la imagen sin que ésta pierda su significado y sin degradar demasiado la calidad de la imagen.

El primer coeficiente de la ecuación 2.9 es llamado el coeficiente DC, el cual contiene el valor de luminancia promedio del bloque, siendo el coeficiente de mayor magnitud del bloque y a los demás componentes se les llama coeficientes de AC.

Se describió de una forma general la transformada discreta del coseno, la cual es utilizada para llevar a cabo la inserción de la marca de agua en el esquema propuesto en el capítulo 4, así como la clasificación de texturas utilizada en dicho esquema.

Por último, menciona que las redes neuronales artificiales han sido aplicadas con éxito en gran cantidad de problemas como por ejemplo reconocimiento de patrones, clasificación, visión, control, predicción [39]



Capítulo III.

Estado del campo o del arte

3.1. Trabajos de investigación

A continuación, una descripción de los trabajos relacionados, artículos, autores, que han abordado el tema que es el objeto de estudio de esta tesis. Dicho análisis, permite resumir el conocimiento y las conclusiones hasta ahora obtenidas, y remplazar los resultados.

En [40] se define un esquema de marcas de agua adaptativo basado en retroalimentación, este sistema reacciona a los cambios en su entorno o modificaciones ya sean traslaciones o rotaciones. La red neuronal aplica un proceso de incrustación de la marca de agua digital, de tal manera que para determinar la intensidad máxima ocupa los coeficientes de frecuencia, mediante bloques de 8×8 utilizando también la TDC (transformada discreta del coseno). El esquema de marca de agua en este trabajo se probó en imágenes médicas con un tamaño de 512×512 pixeles.

En [41] propone un esquema de incrustación y detección cuando la marca de agua está ligada con un descriptor de contenido definido por los puntos más destacados. Se estudia tres tipos diferentes puntos de función y su robustez a las transformaciones geométricas de la imagen. La incorporación de la firma se realiza mediante la extracción de puntos característicos de la imagen y realiza una función de teselación de delaunay la cual consiste en que, dado una nube de puntos en el plano, hallar una triangulación en la que los puntos más próximos entre si estén conectados por una arista o, dicho de otra forma, en la que los triángulos resultantes sean los más regulares posibles. La marca se incrusta utilizando un esquema aditivo clásico dentro de cada triángulo de la teselación. La detección se realiza utilizando propiedades de correlación en los diferentes triángulos. El rendimiento del esquema presentado se evalúa después de la compresión JPEG.

En el trabajo reportado en [42], se presenta un método de marca de agua basado en puntos característicos de la imagen. Dichos puntos son extraídos para construir un mosaico que contiene regiones triangulares y así poder insertar la información de marca de agua en el dominio espacial. En la etapa de detección, el método extrae

los puntos característicos a partir de la imagen marcada, la cual ha sido o no distorsionada. Usando estos puntos característicos, el mosaico con regiones triangulares es reconstruido y la marca de agua es detectada por una operación de correlación. Este esquema es robusto a pequeñas rotaciones y escalamiento. Sin embargo, la extracción de puntos característicos no es del todo estable en este método. Como resultado de este hecho, el conjunto de regiones triangulares obtenidas en las etapas de inserción y detección son diferentes.

En [43] se propone una marca de agua digital adaptativa de esquema con redes neuronales *FBR (fusión de base radial)* en las que visualmente la marca de agua es una imagen binaria que se incrusta en el dominio de la transformación discreta del coseno de la imagen original. La marca de agua se cifra en una serie caótica la cual es sensible respecto a las condiciones iniciales, se inserta en los coeficientes de frecuencia media de la imagen original bloqueado por la DTC basada en la transformación del dominio. Con el fin de hacer que la marca de agua sea más fuerte para resistir diferentes tipos de ataques es importante adaptar la cantidad máxima de incrustaciones. En este trabajo se utiliza las redes neuronales RBF para lograr una marca de agua de máxima fuerza es decir que la marca de agua no sea detectada al ojo humano y donde los pesos obtenidos del entrenamiento de la red sean los óptimos cuidando también que la imagen visualmente no se vea distorsionada, de acuerdo a la característica de los componentes de frecuencia de la imagen original.

3.2.Bit menos significativo (BMS)

Este método es uno de los más comunes por su fácil implementación, ya que trabaja directamente con los bits de la imagen, que consiste en modificar el bit menos significativo de cada píxel, insertando el valor que se desea almacenar. Sin ninguna alteración previa del píxel. En nuestro caso el mensaje que se inserta tiene un proceso previo de encriptación, si el usuario así lo requiere.

Por tal motivo se denominan métodos del dominio espacial, a diferencia de los métodos del dominio de frecuencias, que si tienen un proceso previo de los píxeles

antes de implantar el valor. Por tal motivo les proporcionan una seguridad más alta. Los métodos de domino espacial no son muy robustos en seguridad, por lo cual se consideran de primera generación en la esteganografía, dichos métodos tienen diferentes formas de implementación.

La gran diferencia que existe entre los métodos de inserción de marcas de agua es que para el LSB podemos deducir que la imagen llega a tener una poca discreción a la inserción de una imagen dentro de otra, mientras que utilizando el método redes neuronales (*Neural Network NN*) tenemos la seguridad que podemos elegir los pixeles óptimos de la imagen para realizar la inserción de la maca de agua [44].

El método de ocultación de la información BMS (Bit Menos Significativo) consiste en aprovechar el bit menos significativo de cada byte para guardar información. Para poder amplificar más el concepto ocuparemos el siguiente ejemplo basándonos en una imagen (llamada "A") en la que queremos ocultar un documento de texto (llamado "B") creando la imagen "A+B".

Para utilizar esta técnica, lo que se debe hacer es almacenar todos los bits del documento B en los bits menos significativos de cada uno de los colores que componen los pixeles de la imagen A, sustituyendo los de la propia imagen por los del documento B y añadiendo 0's una vez se ha completado el documento B. Al ser los bits que proporcionan menos información de color al pixel, los cambios realizados en los colores de la imagen no serán apreciables por el ojo humano, con esto podemos ver que un requisito a este método (aplicado a imágenes) es que el archivo a ocultar tiene que ser 3 veces más pequeño que el número total de pixeles de la imagen portadora (en la que ocultaremos el archivo). Esto debido a que para cada 3 bits del documento B, necesitaremos un pixel de la imagen (suponiendo que el pixel está formado por los tres bytes de RGB).

Una vez se tiene la imagen A+B, para poder recuperar el archivo oculto, únicamente deberemos acceder a cada uno de los pixeles de la imagen obtener el contenido del bit menos significativo de cada uno de ellos en un nuevo archivo, obteniendo así el documento B tal y como estaba antes de ser ocultado [43].

3.3. Software propietario

También existe software programado para la inserción de marca de agua actualmente y se nombran algunos ejemplos a continuación.

3.3.1. Photo Helper 1.0

Este software permite rotar imágenes y añadir una marca de agua de copyright con solo unos clics del ratón. Crea marcas de agua, ya sea con texto o un archivo gráfico, establece la opacidad marca de agua y colócalos en las imágenes el problema que se encuentra con esta aplicación es que las marcas de agua son visibles [44].

3.3.1. PicGhost

Pic Ghost soporta todos los tipos de extensiones de imágenes además de que permite un máximo de 40 imágenes cada imagen con un máximo de 10MB lo que quiere decir que permite subir hasta 400Mb a la vez. El problema que se llega a tener con lo referente a las marcas de agua es que son visibles y para imágenes médicas son alteradas lo que para un diagnostico no serviría. Se puede elegir el color y la tipografía y la posición de la marca de agua [45].

3.3.2. Icemark


Es un sistema de tecnología de última generación para la protección de imágenes digitales contra el uso ilegal. Te permite incrustar marcas de agua digital o (watermarking) invisibles con información sobre el propietario en una imagen. Esto puede ser un poco de texto o un identificador. Uno de los problemas que se encuentra con este tipo de software es que es de licencia dado que el enfoque que se propone es para imágenes médicas si es aplicado a una clínica esto tendría un costo muy alto [46].

3.3.3. AiS Watermarker ActiveX

Es un software que lee una imagen de cualquier tipo de extensión, se suma a la imagen de color marca de agua transparente y guarda la imagen resultante a un nuevo archivo

- Algunas de las funciones que tiene incluida este software son:
- Personalizar la configuración de marcas de agua
- Guardar y ajustar las cagas desde archivos
- Añadir marcas de agua de color transparente a las imágenes
- Hacer imágenes miniaturas.

Cabe mencionar que el software antes mencionado no cuenta con ningún algoritmo de redes neuronales para la incrustación de marcas de agua [47].



Capítulo IV. Algoritmo de inserción de marcas de agua con redes neuronales

4.1 Algoritmo de inserción con redes neuronales

La red neuronal multicapa es una extensión del perceptron a la cual añade una serie de capas que, básicamente, hacen una transformación sobre las variables de entrada. Esta red multicapa se define con una capa de entrada y una capa de salida y al menos una capa oculta. Dichas capas se unen de forma total hacia adelante, esto es la capa de entrada se une con la primera capa oculta; cada capa oculta se conecta hacia adelante de manera sucesiva y la última capa se une con la capa de salida. A este tipo de topología se le conoce como retro alimentación y comúnmente el error se retroalimentación para garantizar la mejor convergencia en la salida de cada capa en el entrenamiento de la red neuronal.

Para lograr el objetivo planteado en este trabajo se sigue la metodología de [47], las imágenes a procesar en este trabajo de tesis, son imágenes de 8 bits por canal respecto al modelo de color RGB. Se procesa una imagen de 24 bits para realizar la incrustación de la marca de agua digital. La imagen a insertar como marca de agua debe sea resistente a transformaciones geométricas. Debido a que si la imagen puede ser sensible a estas transformaciones se dice que no es robusta porque se altera su descripción geométrica. La imagen robusta la representaremos por θ donde:

$$\theta = \begin{pmatrix} \theta(0, 0) & \theta(0, 1) & \dots & \theta(0, 0_w - 1) \\ \theta(1, 0) & \theta(1, 1) & \dots & \theta(\theta_h, 0_w) \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \theta(\theta_h - 1, 0) & \theta(\theta_h - 1, 1) & \dots & \theta(\theta_h - 1, 0_w - 1) \end{pmatrix}, \quad (4.1),$$

donde $\theta(i, j)$ es un número entero entre $0 < \theta(i, j) < 255$ y θ_h y θ_w son la altura de la imagen original y ancho, respectivamente.

Una marca de agua W se utiliza en nuestro esquema de inserción. La imagen de marca de agua la definimos como:

$$W = \begin{pmatrix} W(0, 0) & W(0, 1) & \dots & W(0, W_w - 1) \\ W(1, 0) & W(1, 1) & \dots & W(W_h, W_w) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ W(W_h - 1, 0) & W(W_h - 1, 1) & \dots & W(W_h - 1, W_w - 1) \end{pmatrix}, \quad (4.2),$$

donde es un número entero entre $0 < W(i, j) < 255$ y W_h y W_w son la altura de la imagen original y ancho, respectivamente.

La marca de agua se representa como una matriz. Con el fin de ocultar la marca de agua en la imagen original, se convierte los datos en un vector de la manera siguiente:

$$W = (W_0, W_1, \dots, W_k, \dots, W_{h \times w} - 1). \quad (4.3),$$

donde: $W_k = W(i, j)$ es la conversión de la matriz, $k = j + (i \times W_w)$, $0 \leq i < W_h$ y $0 \leq j < W_w$. Inspirado por el sistema de cifrado Rabin's ³ de llaves públicas, donde se eligen dos números primos P y Q , el producto de $P \times Q = n$, dos claves secretas k_1 y k_2 son elegidas por el usuario utilizados para decidir los lugares (x_i, y_j) donde se oculta la marca de agua (x_i, y_j) .

Para ubicar las posiciones (x_i, y_j) de donde se ocultará la marca de agua se debe seguir los siguientes cálculos: para calcular la posición inicial (x_0, y_0) se lleva a cabo las siguientes operaciones:

$$X_0 = K_1^2 \mod n \quad (4.4)$$

$$Y_0 = K_2^2 \mod n \quad (4.5)$$

$$x_i = X_0 \mod \theta_w \quad (4.6)$$

$$y_i = Y_0 \mod \theta_h \quad (4.7)$$

calcular $W_h \times W_w - 1$ de la siguiente manera:

³ El criptosistema de Rabin es una técnica criptográfica asimétrica cuya seguridad, al igual que RSA, se basa en la complejidad de la factorización de los números primos. Sin embargo, la ventaja del criptosistema de Rabin es que se ha demostrado que la complejidad del problema en el que se basa es tan duro como la factorización de enteros

$$X_I = X_{I-1}^2 \bmod n \quad (4.8),$$

$$Y_I = Y_{I-1}^2 \bmod n \quad (4.9)$$

$$x_i = X_I \bmod \theta_w \quad (4.10),$$

$$y_i = Y_I \bmod \theta_h \quad (4.11),$$

donde K_1^2 y K_2^2 son dos claves secretas elegidas por el usuario para iniciar el calculo de las posiciones de iniciar la obtención de bloques para la TDC $i = 1, 2, \dots$ Si (x_i, y_i) posiciones en la imagen para la inserción de la marca de agua se duplican estos valores deben ser eliminados. Luego de obtener las posiciones (x_i, y_i) debemos obtener los lugares W_h y $W_w - 1$.

Estos lugares son diferentes de (x_0, y_0) y entre sí. Por lo tanto, utilizando el procedimiento de decisión de localización, se puede producir W_h y W_w ubicaciones disponibles en total. Teniendo en cuenta que si (x_i, y_i) tienen las mismas coordenadas esto producirá un ciclo infinito en las operaciones de búsqueda de la posición.

Las imágenes a insertar se procesan en el dominio de las frecuencias mediante la transformación discreta del coseno, donde utilizando el procedimiento anterior de localización se pueden producir sub matrices correspondientes a la imagen original la cual denotaremos por:

$$M_i = \text{submatriz}(\theta, x_i, x_i + 7, y_i, y_i + 7). \quad (4.12),$$

Aquí la función M_i denota la adquisición de la TDC en bloques de 7 la cual concentra la energía de una imagen en la esquina superior izquierda de la imagen original. Estos bloques pueden solaparse parcialmente entre sí mostrado en la (figura 4.14)

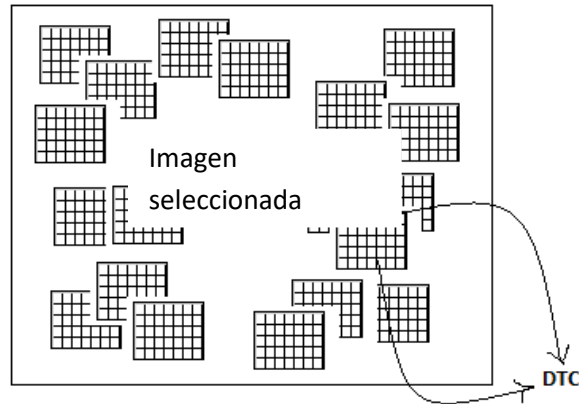


Figura 4.14: Bloques seleccionados la TDC para insertar (adaptada de [50]).

Para insertar la marca de agua, se utiliza una red neuronal multicapa cuya topología se muestra en la (figura 4.14) con nueve unidades de entrada a la cual ingresan una sub-matriz de píxeles de 1×9 cuyos valores son obtenidos de la imagen de marca de agua esta sub-matriz es obtenida mediante un recorrido zigzag de los valores resultantes con la TDC como se muestra en la figura 4.15 y se obtiene un vector el cual es representado por $AC12'_i$.

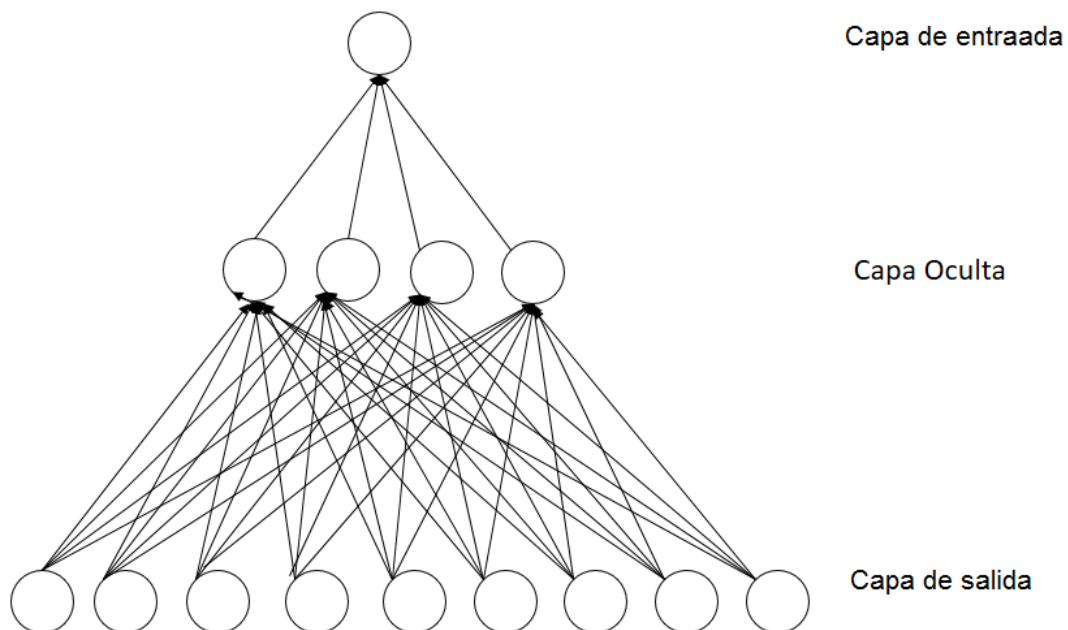


Figura 4.15 Modelo de red neuronal propuesto

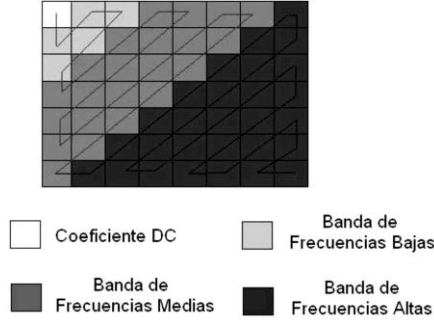


Figura 4.10. Bandas de frecuencias en un bloque TDC de tamaño 8×8 (adaptado de [51]).

donde la función $AC12'_i$ es la salida del entrenamiento de la red neuronal, para insertar la marca de agua esta incrustado mediante la sustitución con la función $AC12''_i$ donde $AC12''_i$ se calcula de acuerdo con $AC12'_i$ y W_i como sigue:

$$AC12''_i = \begin{cases} AC12'_i - \delta, & \text{si } W_i = 0 \\ AC12'_i + \delta, & \text{si } W_i = 1 \end{cases} \quad (4.13)$$

donde δ es un parámetro el cual entre mayor es el valor: la marca de agua será más robusta, pero la distorsión de la imagen contenedora de la marca de agua será más visible. El valor de δ puede ser determinado por las necesidades del usuario.


Luego del entrenamiento de los pesos W_{ij} se puede utilizar para la recuperación de la marca de agua. El vector de salida correspondiente $AC12'_i$ se adquiere a partir de $AC12_i$ al final. $AC12'_i$ a continuación, se calcula:

$$AC12'_i = N(AC1_i, AC2_i, \dots, AC9_i), \quad (4.14)$$

Una tarea común después de insertar una marca de agua es la extracción de la marca de agua para verificar la propiedad del autor respecto a la imagen. El procedimiento de extracción de la marca de agua es similar al procedimiento de incrustación; excepto que para el procedimiento de la TDC este se efectúa de forma inversa. Cuando las claves secretas correctas (K_1, K_2) y los pesos se introducen, correspondientemente $AC12$ y $AC12'$ la marca de agua se obtiene. Esta marca de agua recuperada se procesa utilizando la relación $AC12$ y $AC12'$ como:

$$w_i = \begin{cases} 0, & AC12_i < AC12'_i \\ 1, & AC12_i > AC12'_i \end{cases} \quad (4.15)$$

luego de insertar la marca de agua, se lleva a cabo un análisis visual del resultado, así como el uso de una métrica de bits por pixel modificado para cuantificar el resultado de la inserción de la marca de agua. Recordar que se busca con la inserción de una marca de agua es la alteración visual mínima de la imagen después de la inserción, con el objetivo de que la marca de agua sea casi imperceptible al ojo del usuario [52].



Capítulo V. Resultados experimentales

5.1 Resultados experimentales

Este capítulo describe las pruebas realizadas para comparar el desempeño de la inserción de marcas de agua ocupando el software Matlab R2010a para la implementación, experimentación y comparación utilizando los algoritmos del bit menos significativo BMS (Bit menos significativo) y el método de redes neuronales RN.

La interfaz gráfica realizada que se muestra en la figura 5.11 cuenta con carga de imagen con la cual se puede aplicar algún filtro de la lista desplegable, entre los que están escala gris, negativo, componente y (rojo, verde, azul, regional y alfa). También cuenta con una opción de carga de marca de agua que a su vez esta debe ser una imagen para poder ser insertada utilizando el método BMS o RN. Cuando es seleccionado el método RN se solicita al usuario 2 llaves públicas para su inserción o extracción de la marca de agua. La imagen inicialmente es a color luego se convierte a escala de grises utilizando las rutinas de Matlab.

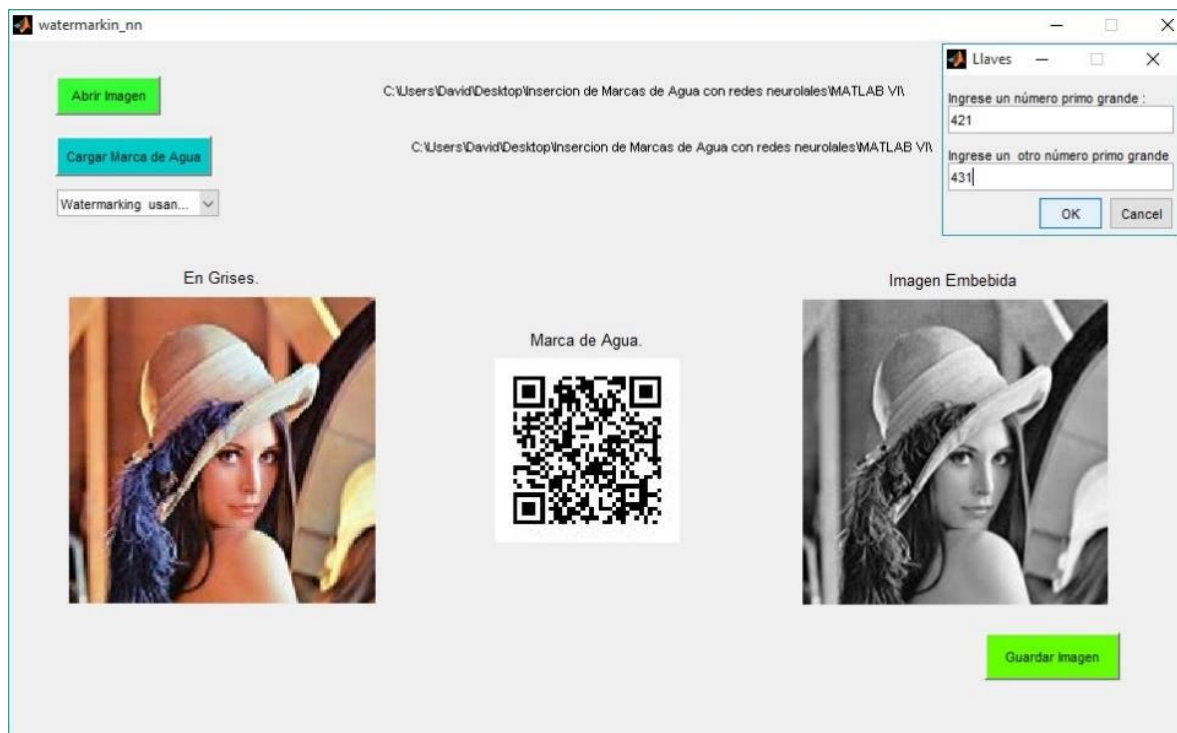


Figura 5.11: Interfaz gráfica de usuario para la inserción de marcas de agua.

En nuestro experimento, se compara dos métodos BMS contra RN para lo cual se evalúan las imágenes obtenidas utilizando distintos filtros del procesamiento digital de imágenes, así como el cambio de topología para la red neuronal.

En la figura 5.12 se muestra las imágenes que ocupamos para realizar el proceso de experimentación:

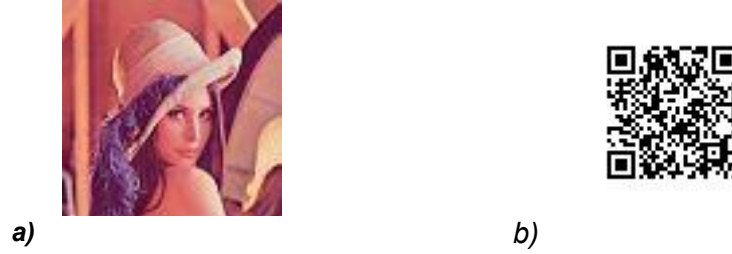


Figura 5.12: a) Imagen Lena y b) Marca de agua a insertar

Con los algoritmos de BMS y RN, se encontró que las diferencias visuales entre la imagen original y la correspondiente imagen cargada con la marca de agua son casi apenas detectadas visualmente. Una manera de cuantificar el nivel exacto de distorsión causada a la marca de agua recuperada, está dada por la función de la proporción correcta de bits (*PCB*).

$$BCR = \frac{\sum_{i=1}^{W_h} \sum_{j=1}^{W_w} \overline{w(i,j) \oplus w'(i,j)}}{W_h \times W_w} \times 100\%. \quad (1.1)$$

Donde: $W(i, j) \in \mathbb{R}$ es la marca de agua original, $W'(i, j) \in \mathbb{R}$ es el elemento de la marca de agua recuperada, \oplus denota el operador de exclusividad *OR* entre las marcas de agua recuperadas, $W_h \in \mathbb{R}$ Es la altura de la imagen y $W_w \in \mathbb{R}$ Es el ancho de la imagen.

El resultado de esta operación entre $(W(i, j) \oplus W'(i, j))$ es ajustada a través de la función piso⁴. Entre mayor sea el valor PCB, la marca de agua recuperada es de mejor calidad en caso contrario dicha marca no se ha recuperado satisfactoriamente.

⁴ La función piso se aplica a un número $x \in \mathbb{R}$ y devuelve el máximo número entero que no superior a x

5.2 Marca de agua utilizando el método BMS

En el siguiente experimento se inserta la marca de agua de la figura 5.12b en la figura 5.12a utilizando el BMS. Se aprovecha el hecho de que el nivel de precisión en muchos formatos de imagen es muy superior a la perceptibilidad por el ojo humano. Por lo tanto, una imagen alterada con ligeras variaciones será indistinguible de la original. BMS requiere ocho bits de pixeles para almacenar 1 byte de datos; estos divididos en grupos de cuatro donde se almacena ya sea texto o una imagen como podemos ver en la figura 5.13.



Figura 5.13: Imagen de resultado luego de ocupar el método BMS.

5.3 Marca de agua utilizando el método RN

En este trabajo de tesis, se insertó una marca de agua utilizando un sistema basado en una red neuronal con una topología de retroalimentación. El primer método de inserción ocupa las figuras 5.12a y 5.12b utilizando una red neuronal de retro propagación para mejorar la seguridad y calidad.



Figura 5.14: Imagen resultante luego de ocupar el método RN de la figura 12a) y b).

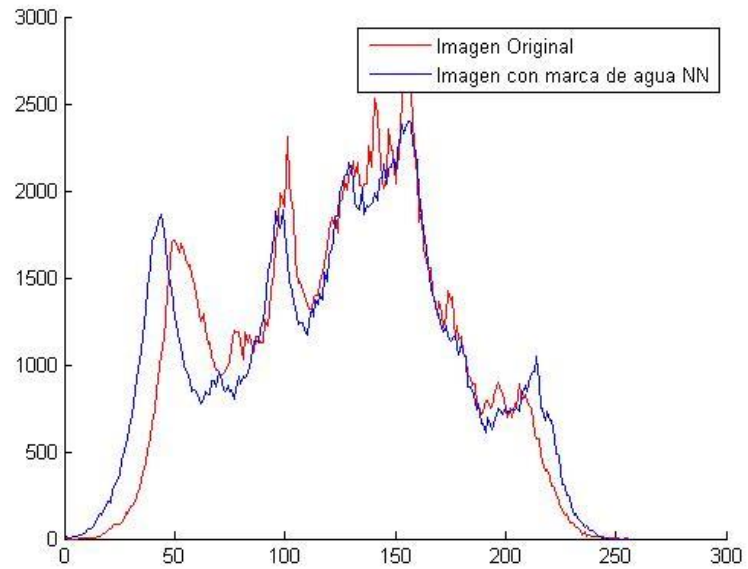
Se puede visualizar en la figura 5.14 que la imagen que contienen la marca de agua utilizando el método RN se nota menos distorsionada en comparación con el método BMS.

Se aplicaron filtro gaussiano, negativo aplicando ruido sal y pimienta, modificando parámetros en la red neuronal, cambiando el número de épocas y variando el parámetro delta sobre la figura 5.13 y 5.14 además de la configuración de activación de salida para la red neuronal, así como de diferentes topologías de la red neuronal. Con esto se demostrando que, aunque la imagen sea expuesta a diferentes cambios, la marca de agua es recuperada exitosamente observando las ventajas de utilizar el método de redes neuronales.

En la figura 5.15a y 5.15d se muestran los histogramas de las figuras 5.15a y 5.15d, se puede observar que la figura 5.15a en comparación con la figura 5.15d las crestas del histograma son casi parecidas a diferencia de la figura 5.15e la cual la separación de las crestas es considerable. El valor PCB de las marcas de agua de la figura 15a es de 42.12 % ya que la marca de agua no ha sido alterada en su totalidad, en la figura 15d podemos observar que su valor algo más 32.10% por lo que en un principio podemos ver que el método de redes neuronales empieza a ser mejor que el método del bit menos significativo.



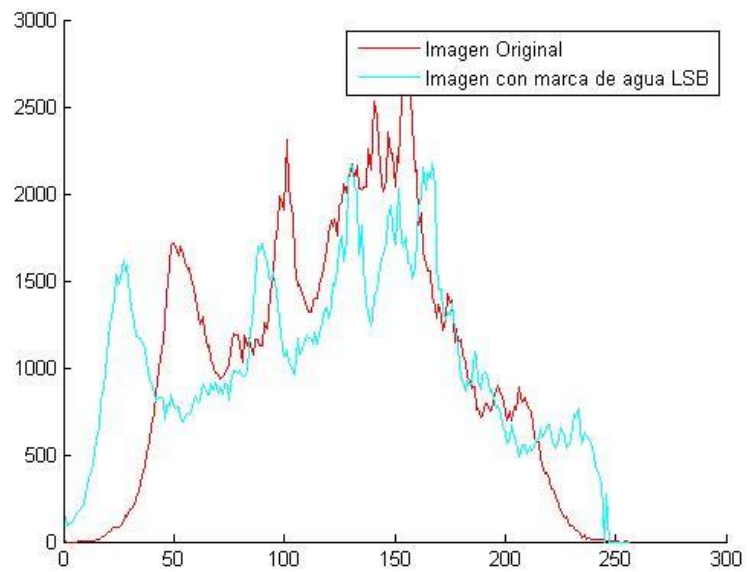
a)



b)



d)



e)

Figura 5.15 a) imagen con marca de agua utilizando redes neuronales, b) histograma imagen con marca de agua utilizando redes neuronales, d) imagen con marca de agua utilizando bit menos significativo y e) histograma imagen original junto con imagen cargada utilizando bit menos significativo.

En las secciones siguientes se distorsiona las imágenes 5.14 y 5.15 aplicando filtro gaussiano, negativo aplicando ruido sal y pimienta, modificando parámetros en la red neuronal, cambiando el número de épocas y variando el parámetro delta esto con el fin de poder comprobar la eficiencia de la inserción de marca de agua utilizando el método de RN sobre el BMS.

5.4 Marca de agua extraída mediante RN y BMS después de utilizar el filtro gaussiano

El filtro gaussiano un objetivo es el de eliminar el ruido de una imagen de manera tal que cuando una imagen tiene gran cantidad de ruido este sea discernido. También pueden utilizarse, para resaltar la información correspondiente a una determinada variabilidad o distorsión de una imagen.

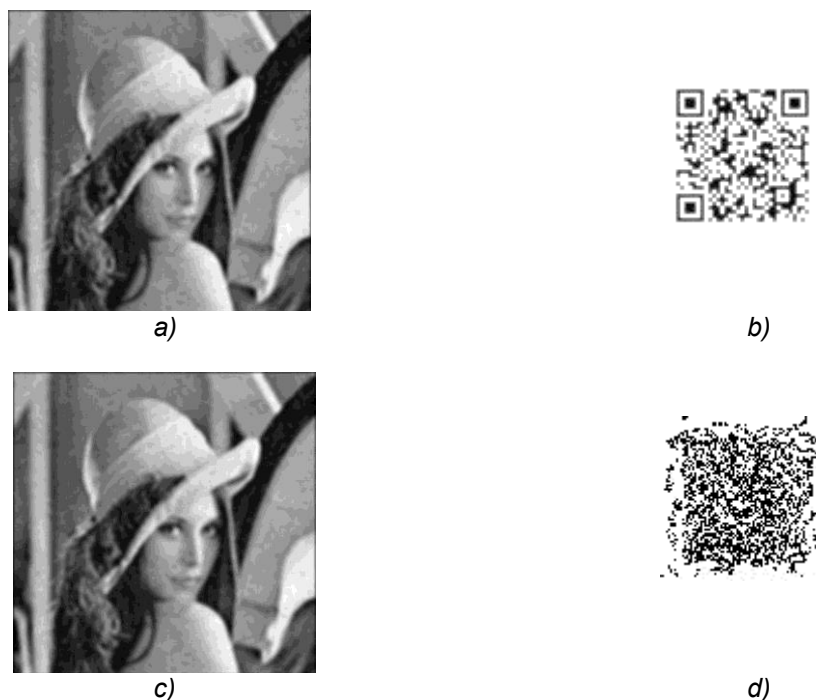


Figura 5.16: a) imagen con marca de agua utilizando RN con filtro gaussiano, b) marca de agua con RN con filtro gaussiano, c) imagen con marca de agua utilizando BMS con filtro gaussiano y d) marca de agua con BMS usando filtro gaussiano.

Se aplica este filtro a la figuras 5.16a y 5.16c luego de extraer la marcas de agua figuras 5.16b y 5.16d utilizando RN y BMS respectivamente se puede observar que para la marca de agua (figura 5.16d) ocupando el método RN para su recuperación, sufre una alteración no tan agresiva sin embargo con el método BMS la marca de

agua figura 5.16d no se obtiene un resultado bueno y esto se puede determinar mediante el PCB de la figura 16b el cual tiene un valor de 89.69% en comparación al de la (figura 5.16d) el cual es 32.12% lo que se concluye que este filtro es agresivo para las marcas de agua insertadas.

5.5 Marca de agua extraída mediante RN y BMS después de utilizar el filtro negativo

El filtro negativo convierte los colores de un pixel a su contrario en la matriz de colores RGB, para obtener cierta visualización de una imagen, el cual ante el ojo humano no es posible distinguir detalles de color referidos en dicha matriz.



Figura 5.17: a) imagen con marca de agua utilizando RN con filtro negativo, b) marca de agua RN después de usar filtro negativo, c) imagen con marca de agua utilizando BMS con filtro negativo y d) marca de agua BMS después de con filtro negativo.

Las marcas de agua (figuras 5.17b y 5.17d) sufren alteraciones mínimas y esto se puede determinar mediante el PCB la (figura 5.17b), tiene un valor de 88.63%, en comparación al valor de la figura 5.17d cuyo valor es 81.54%.

5.6 Marca de agua extraída mediante RN y BMS después de aplicar ruido sal y pimienta

El ruido es la información no deseada que contamina la imagen. Este puede estar en el proceso de adquisición de la imagen (errores en el medio de la comunicación), como en el de transmisión (debido a interferencias en el canal de transmisión) en la inserción de marcas de agua. El ruido puede generarse, luego de ser insertada la imagen.

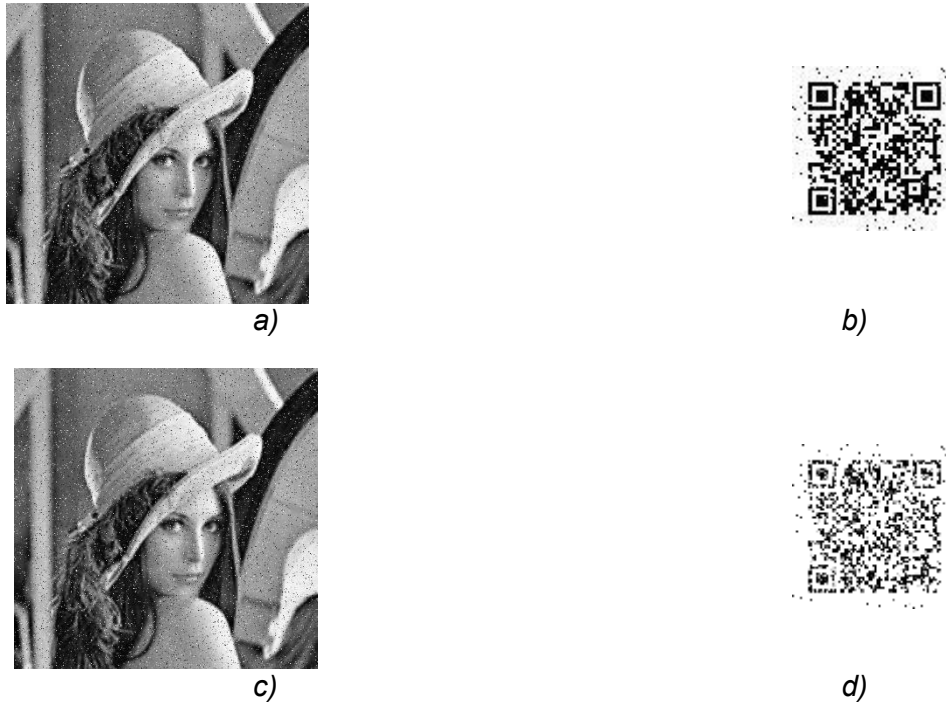


Figura 5.18: a) imagen con marca de agua utilizando RN con ruidos sal y pimienta, b) marca de agua extraída RN con ruido sal y pimienta, c) imagen con marca de agua utilizando BMS con ruidos sal y pimienta y d) marca de agua extraída con ruido sal y pimienta.

En la figura 5.18 podemos observar que después de añadir ruido “sal y pimienta” en ambas imágenes las marcas de agua (figuras: 5.18b y 5.18d) sufren alteraciones mínimas y esto se puede determinar mediante el PCB, la figura 18b tiene un valor 82.61% en comparación al valor de la figura 5.18d el cual su valor es de 71.20%.

5.7 Experimentos con parámetros en la red neuronal

La función de activación calcula la aceleración de la unidad de la red neuronal en función de la entrada total y la activación previa, aunque en la mayor parte de los casos es simplemente una función no decreciente de la entrada total. Los tipos de

función más empleados son: la función escalón, función lineal, función Gaussiana y la función sigmoïdal. La función de salida empleada usualmente es la función identidad y así la salida de la unidad de procesamiento es idéntica a su nivel de activación.



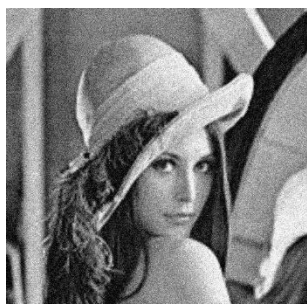
Figura 5.19 : a) imagen cargada activación lineal, b) marca extraída activación lineal, c) imagen cargada activación escalón, d) marca extraída activación escalón, e) imagen cargada activación sigmoïdal y f) marca extraída activación sigmoïdal

Como podemos observar en la figuras: 5.19a, 5.19c y 5.19e con respecto al cambio de configuración de activación : (lineal, escalon, gaussiana y sigmoïdal respectivamente). Las imágenes que contienen las marcas de agua recuperadas (figuras 5.19a, 5.19d y 5.19f) llegan a ser borrosas. Las configuraciones de activación que permiten que la marca de agua se vea reconocible son

“retropropagación” y “lineal”. Las alteración que tiene las marcas de agua las podemos determinar mediante su PCB el cual podemos ver en la figuras: 5.19c, 5.19b, 5.19d y 5.19f.

5.8 Marca de agua extraída mediante RN cambiando el número de épocas

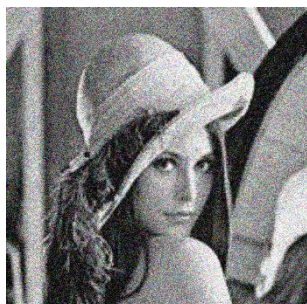
El aprendizaje en las redes se realiza por épocas. La red se hace evolucionando durante un tiempo (época) en el que se le ha introducido una secuencia. Una vez llegando al final de una época se reinicializa la red para que el estado iniciado no dependa del estado final de la época anterior y se entra en una nueva época. La variación de los pesos solo se realiza al acabar por lotes (calcular todos los incrementos de los patrones-épocas) o de forma incremental (después de cada patrón época).



a)



b)



c)



d)

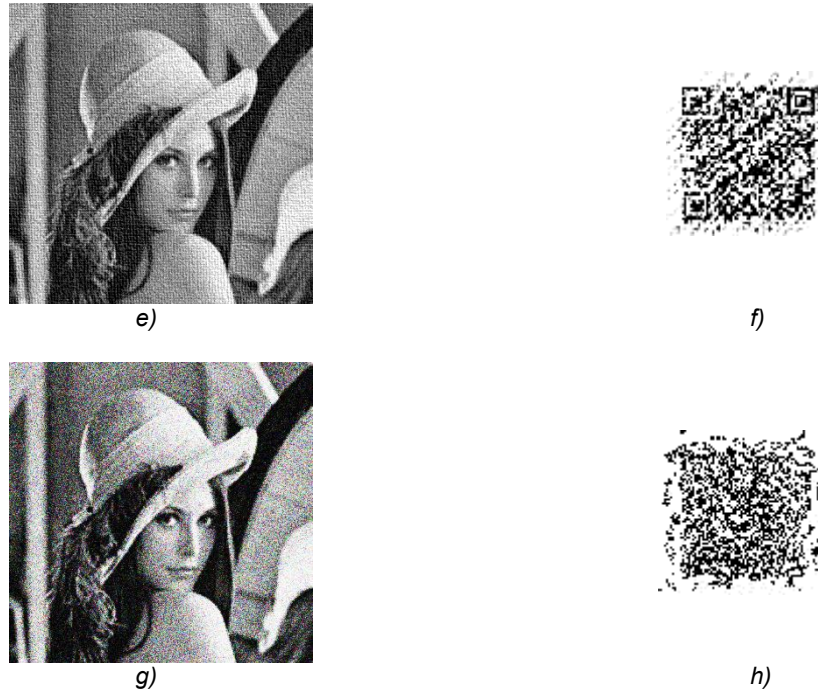


Figura 5.20 : a) imagen cargada 50 épocas, b) marca de agua extraída, c) imagen cargada 70 épocas, d) marca de agua extraída, e) imagen cargada 300 épocas, f) marca de agua extraída, g) imagen cargada 500 épocas y h) marca de agua extraída.

Como podemos apreciar en las figuras: 5.20b, 5, 21e y 5.21h respecto al cambio de épocas en la recuperación de la marca de agua, la imagen llega a embeberse un poco borrosa y se llega a perder por completo en las 500 épocas, pero en general podemos observar que durante todas las modificaciones hechas la marca de agua ha sufrido, leves alteraciones, las cuales no afectan el mensaje que se intenta transmitir.

5.9 Marca de agua extraída mediante RN cambiando el parámetro delta

El parámetro δ es un valor que se le asigna al resultado del entrenamiento de la red neuronal (ecuación 4.25 Capítulo IV); el cual nos permite que la inserción de la marca de agua sea más robusta. Entré más grande sea su valor sin embargo la distorsión de la imagen original se incrementará también.

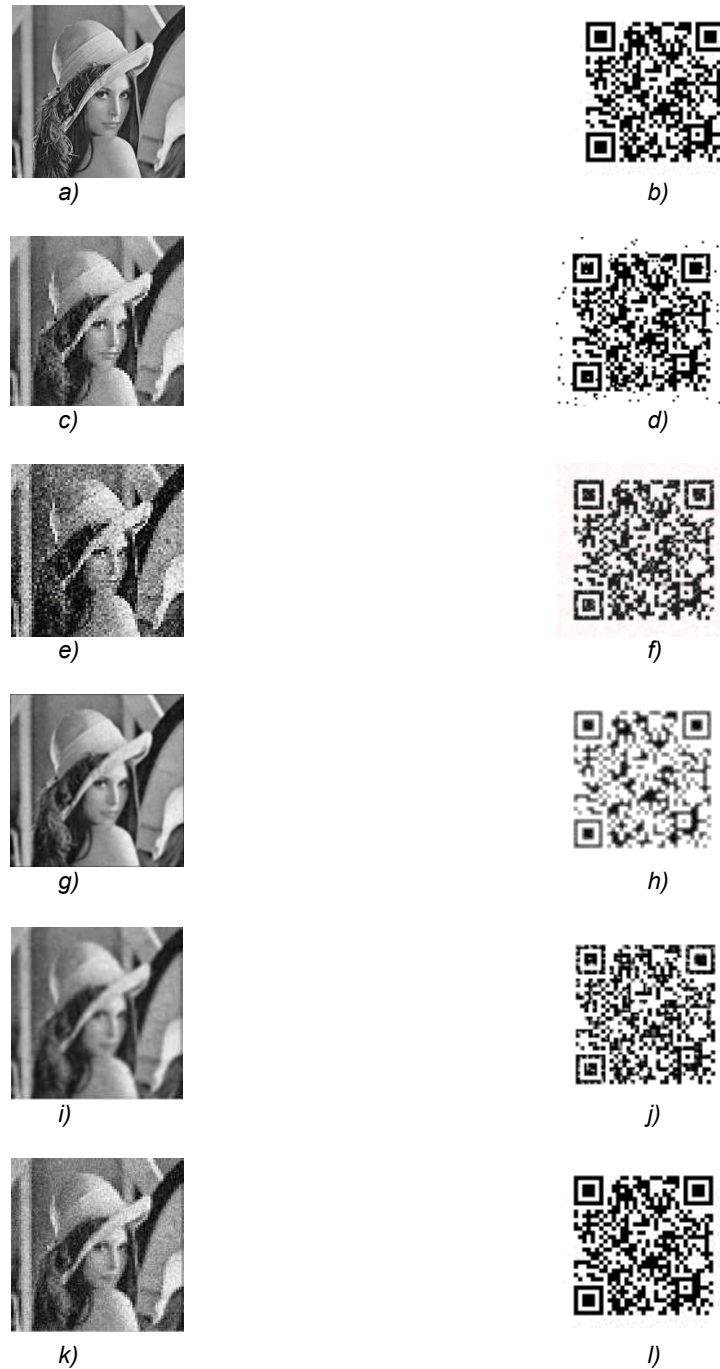


Figura 5.21: a) imagen cargada con marca de agua utilizando configuracion $\delta = 20$, b) Marca de agua extraida utilizando RN luego de el valor $\delta=20$, c) Imagen cargada con marca de agua utilizando configuracion $\delta = 75$, d) Marca de agua extraida utilizando RN luego de el valor $\delta=75$, e) Imagen cargada con marca de agua utilizando configuracion $\delta = 105$, f) Marca de agua extraida utilizando RN luego de el valor $\delta=105$, g) Imagen cargada con marca de agua utilizando configuracion $\delta = 20$ y filtro gaussiano, h) Marca de agua extraida utilizando RN luego de el valor $\delta=20$ de una imagen con filtro gaussiano, i) Imagen cargada con marca de agua utilizando configuracion $\delta = 75$ y filtro gaussiano, j) Marca de agua extraida utilizando RN luego de el valor $\delta=75$ de una imagen con filtro gaussiano, k) Imagen cargada con marca de agua utilizando configuracion $\delta = 105$ y filtro gaussiano y l) Marca de agua extraida utilizando RN luego de el valor $\delta=105$ de una imagen con filtro.

Como podemos apreciar en las figuras 5.21a a 5.21l al incrementar el valor de δ respecto a la marca de agua empieza a deformarse sin embargo la marca de agua (figuras 5.21b – 5.21k) por el contrario llegan a ser de mejor calidad, utilizando la función PCB podemos ver las diferentes alteraciones respecto a la marca de agua (figuras 5.21c a la 5.21l).

5.10 Resultados obtenidos de proporción correcta de bits(PCR)

El valor ilustrado en la tabla 5.1 demuestra que tan buena fue la recuperación de la marca de agua respecto a los filtros utilizados.

Figura N°	Aplicando	Valor PCB luego de la marca de agua obtenida mediante el BMS	Valor PCB luego de la marca de agua obtenida mediante el RN	Margen de diferencia al PCB entre los métodos
5.16 b y 5.16d	Filtro gaussiano	89.69%	32.12%	57.57%
5.17 b y 5.17 d	Filtro negativo	88.63%	81.54%	7.09%
5.18 b y 5.18 d	Ruido sal y pimienta	82.61%	71.20%	11.41%
Tabla 5.1 Comparación de la extracción de las marcas de agua Redes neuronales contra el método bit menos significativo.				

De igual forma se realizaron experimentos variando los parámetros en la red neuronal esto con el fin de poder observar de manera cuantitativa las alteraciones de la marca de agua (tabla 5.2).

Imagen de la marca de agua obtenida	Parámetro modificado	Valor PCB luego de la marca de agua obtenida mediante el RN
5.19b	Función de activación lineal	91.51%
5.19e	Función de activación escalón	68.03%
5.19h	Función de activación sigmoideal	62.86%
5.20b	Número de épocas 50	97.51%
5.20e	Número de épocas 70	82.03%
5.20h	Número de épocas 300	78.86%
5.20k	Número de épocas 500	68.03%
5.21b	Parámetro $\delta=20$	94.51%
5.21e	Parámetro $\delta=75$	96.03%
5.21h	Parámetro $\delta=105$	97.86%
5.21k	Parámetro $\delta=20$ filtro gaussiano	80.69%
5.21n	Parámetro $\delta=75$ filtro gaussiano	89.09%
5.21q	Parámetro $\delta=105$ filtro gaussiano	94.96%
Tabla 5.2 Comparación de la extracción de las marcas de agua redes neuronales luego de realizar diferentes experimentos.		



Capítulo VI.

Conclusiones y trabajo a futuro

6.1. Conclusiones

En el presente trabajo de tesis se cumplieron los objetivos, se implementó una red neuronal con retro alimentación para la inserción de una marca de agua dentro de una imagen con el fin de transmitir información de manera segura.

Puesto que la inserción de marca de agua tiene un alto número de métodos de inserción en esta tesis se propone el método más robusto para realizar las inserciones de la marca de agua.

En el proceso de experimentación se puede observar en la tabla de resultados del PCB, luego de la recuperación de la marca de agua, que las marcas de agua insertadas y luego extraídas mediante la red neuronal tiene un valor bajo debido a que la marca de agua además de ser visible también es de buena calidad según el PCB.

El método de inserción de marcas de agua utilizando RN sobre el BMS se demuestra en los resultados presentados que el método de RN es superior sobre BMS ya que la marca insertada y luego extraída sufre una menor modificación con el primer método.

6.2. Trabajo a futuro

Como trabajo a futuro se explora el seguir optimizando con otro tipo de redes neuronales e incluso el poder utilizar algoritmos bio-inspirados para poder realizar la inserción de las marcas de agua, una adición a este proyecto se plante la colonia de hormigas esto con el fin de comparar resultados de las extracciones de las marcas de agua y poder comparar con el PCB cuál es la mejor marca de agua extraída.

Como en el trabajo realizado con el entrenamiento de redes neuronales basado en algoritmos evolutivos [53] o técnicas robustas de marca de agua basada en normalización de imágenes [54].

Bibliografía

Bibliografía

- [1] Marie, R. R. Data Hiding in Digital Images using Cryptography and Steganography Techniques (CryptSteg). *World of Computer Science & Information Technology Journal*, 4(4) (2014).
- [2] Disponible: http://images.slideplayer.es/3/1095811/slides/slide_60.jpg
- [3] Disponible: http://images.slideplayer.es/3/1095811/slides/slide_60.jpg
- [4] Johnson, N. F., and Jajodia, S. Steganalysis: The investigation of hidden information. In *Information Technology Conference*, pp. 113-116 1998.
- [5] http://1.bp.blogspot.com/-D4dBuIFN8U/UbX_zpn1fzI/AAAAAAAAASw/_GKrwPAn0nw/s1600/compre5.gif
- [6] Moskowitz, S. A., and Cooperman, M. *Method for human-assisted random key generation and application for digital watermark system*. U.S. Patent No 5,822,432, 13 October. 1998.
- [7] Huang, S., Zhang, W., Feng, W., and Yang, H. Blind watermarking scheme based on neural network. In *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on* pp. 5985-5989 2008, June.
- [8] Tefas, A. N., and Solachidis, V. A survey on watermarking application scenarios and related attacks. In *Proc. Int. Conf. Image Processing* Vol. 3, pp. 991-994 2001, October.
- [9] Moskowitz, S. A., and Cooperman, M. *Method for human-assisted random key generation and application for digital watermark system*. Patent No 5,822,432, 13 October.. 1998.
- [10] Stern, Julien P., et al. Robust object watermarking: Application to code. En *International Workshop on Information Hiding*. Springer Berlin Heidelberg, 1999. p. 368-378.
- [11] Moskowitz, S and Cooperman, M. *Method for human-assisted random key generation and application for digital watermark system*. U.S. Patent No 5,822,432, 13 October.. 1998..
- [12] Disponible: http://docs.opencv.org/2.4/_images/mlp.png
- [13] Dumitrescu, S., Wu, X., and Wang, Z. *Detection of LSB steganography via sample pair analysis*. *transactions on Signal Processing*, vol. 51, no 7, p. 1995-2007, 2003.
- [14] Disponible: <http://obsessive-coffee-disorder.com/wp-content/uploads/2014/02/lena.png>

- [15] Katzenbeisser, S., and Petitcolas, F. *Information hiding techniques for steganography and digital watermarking*. Artech house (2000).
- [16] Gonzalez, R. C., and Woods, R. *Digital image processing* 2002.
- [17] Disponible: <http://www.revista.unam.mx/vol.9/num4/art23/imagenes/Figura%201a.jpg>
- [18] Cabrera Cortés, I. A. El procesamiento humano de la información: en busca de una explicación (2003). Disponible: http://scielo.sld.cu/scielo.php?pid=S1024-94352003000600006&script=sci_arttext&lng=pt
- [19] Eskicioglu, A. M., and Fisher, P. S. Image quality measures and their performance. *Transactions on communications*, vol. 43, no 12, p. 2959-2965 1995.
- [20] Ruan, Benfang, et al. Binding of rapamycin analogs to calcium channels and FKBP52 contributes to their neuroprotective activities. *Proceedings of the National Academy of Sciences*, vol. 105, no 1, p. 33-38 2008.
- [21] Govindaraju, Rao S.; Rao, Adiseshappa Ramachandra. *Artificial neural networks in hydrology*. Springer Science and Business Media, 2013.
- [22] Ruan, Benfang, et al. Binding of rapamycin analogs to calcium channels and contributes to their neuroprotective activities. *Proceedings of the National Academy of Sciences*, vol. 105, no 1, p. 33-38 2008.
- [23] Disponible:
https://upload.wikimedia.org/wikipedia/commons/thumb/6/60/ArtificialNeuronModel_english.png/600px-ArtificialNeuronModel_english.png
- [24] Disponible:
https://upload.wikimedia.org/wikipedia/commons/thumb/6/60/ArtificialNeuronModel_english.png/600px-ArtificialNeuronModel_english.png
- [25] Matich, Damián Jorge. *Redes Neuronales: Conceptos básicos y aplicaciones*. Cátedra de Informática Aplicada a la Ingeniería de Procesos–Orientación I, 2001.
- [26] Johnson, N. F., and Jajodia, S. Steganalysis of images created using current steganography software. In *International Workshop on Information Hiding* (pp. 273-289). Springer Berlin Heidelberg (1998, April).
- [27] Yu, P. T., and Lin, J. S. Digital watermarking based on neural networks for color images. *Signal processing*, 81(3), 663-671 (2001).
- [28] Disponible: http://1.bp.blogspot.com/-D4dBuIF-8U/UbX_zpn1fzI/AAAAAAAAASw/_GKrwPA0nw/s1600/compre5.gif
- [29] I. J. Cox et al., 1999 July Digital watermarking based on neural networks for color images

- [30] Ingemar J et al., L. Watermarking as communications with side information. *Proceedings of the*, 1999, vol. 87, no 7, p. 1127-1141.
- [31] Bas, P., Chassery, J. M., and Macq, B. Geometrically invariant watermarking using feature points. *transactions on image Processing*, 11(9), 1014-1028 2002.
- [32] Bender, Walter, et al. for data hiding. *IBM systems journal*, 35(3.4), 313-336.
- [33] Lou, D. C and Liu, J. L. Adaptive digital watermarking using neural network technique. In *Security Technology*, pp. 325-332 2003, October.
- [34] Ruan B, et al. Binding of rapamycin analogs to calcium channels and FKBP52 contributes to their neuroprotective activities. *Proceedings of the National Academy of Sciences*, vol. 105, no 1, p. 33-38 , 2008.
- [35] RUAN, Benfang, et al. Binding of rapamycin analogs to calcium channels and FKBP52 contributes to their neuroprotective activities. *Proceedings of the National Academy of Sciences*, vol. 105, no 1, p. 33-38, 2008.
- [36] Lou, D. C., Liu, J. L., and Hu, M. C. (2003, October). Adaptive digital watermarking using neural network technique. In *Security Technology, 2003. Proceedings. 37th Annual International Carnahan Conference on* (pp. 325-332) 2003.
- [37] Bansal, E. A., and Bhadauria, D. S. Watermarking using neural network and hiding the trained network within the cover image. *Journal of Theoretical and Applied Information Technology*, 663-670 (2008).
- [38] Mezquida. D et al. Ocultación de imagenes mediante Esteganografía. *Novática: Revista de la Asociación de Técnicos de Informática*, (163), 52-57.
- [39] Zhi-Ming, Z., Rong-Yan, L., and Lei, W Adaptive watermark scheme with RBF neural networks. En *Neural Networks and Signal Processing, 2003. Proceedings of the International Conference on*, p. 1517-1520 2003.
- [40] Zhi-Ming, Z., Rong-Yan, L., and Lei, W. Adaptive watermark scheme with RBF neural networks. In *Neural Networks and Signal Processing. Procings of the n'Op 76 International Conference on* (Vol. 2, pp. 1517-1520 ,December 2003.
- [41] Lou, D. C and Liu, J. L. Adaptive digital watermarking using neural network technique. In *Security Technology*, pp. 325-332 2003, October.
- [42] Zhi-Ming, Z., Rong-Yan, L., and Lei, W Adaptive watermark scheme with RBF neural networks. En *Neural Networks and Signal Processing, 2003. Proceedings of the International Conference on*, p. 1517-1520 2003.
- [43] TechSpot, Inc. All Rights Reserved. 2016 Photo Helper 1.0 Disponible: <http://www.techspot.com/downloads/6091-photo-helper.html>
- [44] PicGhost TreeNodes 2015 picghost Disponible: <http://www.picghost.com/>

- [45] Icemark mark 2016 Icemark mark Disponible:
<http://www.phibit.com/images/download.png>
- [46] Watermarker.com 2010 Watermarker.com Disponible:
<http://www.watermarker.com/download.aspx>
- [47] Johnson et al., 1996 Techniques for data hiding Disponible:
<http://ieeexplore.ieee.org/document/5387237/?denied>
- [48] Disponible:
http://electronicimaging.spiedigitallibrary.org/data/Journals/ELECTIM/22345/023006_1_1.png
- [49] Disponible:
http://electronicimaging.spiedigitallibrary.org/data/Journals/ELECTIM/22345/023006_1_1.png
- [50] Disponible: http://images.books24x7.com/bookimages/id_13509/fig463_01.jpg
- [51] Huang, S., Zhang, W., Feng, W., and Yang, H. Blind watermarking scheme based on neural network. In *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on* pp. 5985-5989 2008, June.
- [52] Disponible: <http://www.iuma.ulpgc.es/users/lhdez/inves/tesis/memoria-tesis/node5.html>
- [53] Britos, M. I. P. ENTRENAMIENTO DE REDES NEURONALES BASADO EN ALGORITMOS EVOLUTIVOS (2005).
- [54] Hernández, M. C et al, Técnica robusta de marca de agua basada en normalización de imágenes. *Revista Facultad de Ingeniería Universidad de Antioquia*, (52), 147-160.