



## **“METODOLOGÍA BASADA EN INTELIGENCIA ARTIFICIAL PARA DETECCIÓN DE ANORMALIDADES ASOCIADAS AL CÁNCER DE MAMA EN IMÁGENES TERMOGRÁFICAS INFRARROJAS”**

Estudiante:  
Ing. Dulce Sofía Marín Chávez

Director:  
Dr. Luis Alberto Morales  
Hernández

Secretario:  
Dr. Jesús Alberto Basurto  
Hurtado

Vocal:  
Saúl Tovar Arriaga

Suplente:  
Andras Takacs Suplente

Suplente:  
Irving Armando Cruz Albarrán

La presente obra está bajo la licencia:  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>



CC BY-NC-ND 4.0 DEED

Atribución-NoComercial-SinDerivadas 4.0 Internacional

### Usted es libre de:

**Compartir** — copiar y redistribuir el material en cualquier medio o formato

La licenciante no puede revocar estas libertades en tanto usted siga los términos de la licencia

### Bajo los siguientes términos:



**Atribución** — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante.



**NoComercial** — Usted no puede hacer uso del material con [propósitos comerciales](#).



**SinDerivadas** — Si [remezcla, transforma o crea a partir](#) del material, no podrá distribuir el material modificado.

**No hay restricciones adicionales** — No puede aplicar términos legales ni [medidas tecnológicas](#) que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia.

### Avisos:

No tiene que cumplir con la licencia para elementos del material en el dominio público o cuando su uso esté permitido por una [excepción o limitación](#) aplicable.

No se dan garantías. La licencia podría no darle todos los permisos que necesita para el uso que tenga previsto. Por ejemplo, otros derechos como [publicidad, privacidad, o derechos morales](#) pueden limitar la forma en que utilice el material.



## Dedicatoria

A mis padres por su amor incondicional, su apoyo y la educación que me brindaron, mis hermanos y mis abuelos por su amor y paciencia.  
A mi familia, mi compañero de vida y padre de mi hija Isidoro, a mi hija Renata por el inmenso amor que le tengo, demostrarle que ella también puede y que son mi motivación y mi fuerza para seguir adelante.

## Agradecimientos

Agradezco el apoyo por SECIHTI al brindarme beca para la realización de esta investigación y mi aporte en el avance de la ciencia en México. A la facultad de ingeniería de la Universidad Autónoma de Querétaro por mi trayecto como estudiante.

A mi director de tesis Dr. Luis Morales y coordinador de la maestría el Dr. Saúl Tovar por su paciencia y apoyo en el trabajo de investigación, a mi comité de tesis por su dirección y aportación.

A mi familia por su apoyo incondicional, paciencia y amor, porque sin ellos este trabajo no sería posible, por último a mis amigos y compañeros de la maestría.



## ÍNDICE

I.	RESUMEN	7
II.	INTRODUCCION	7
III.	ANTECEDENTES	8
	Mastografía	8
	Ultrasonido	9
	Muestras de biopsia	9
	Termografía	9
IV.	JUSTIFICACIÓN	15
V.	DESCRIPCIÓN DEL PROBLEMA	15
VI.	FUNDAMENTACIÓN TEÓRICA	16
	Termografía	16
	Segmentación de imágenes	17
	Filtrado	17
	Ruido	18
	Métodos de umbralización	19
	Método Otsu	19
	Operaciones morfológicas	21
	Erosión	21
	Dilatación	22
	Transformaciones morfológicas	23
	Apertura	23
	Cerradura	24
	Operaciones morfológicas por reconstrucción	25
	Detección de bordes	26
	Características de una imagen	26
	Contraste	26
	Homogeneidad	26
	Cluster Shade	26
	Cluster prominence	27
	Clasificadores	27
	Redes Neuronales Artificiales (RNA)	28



Red Neuronal Convolucional (CNN)	31
Redes Neuronales Profundas	31
Red VGG16	31
Autoencoders	32
VII. HIPÓTESIS	33
VIII. OBJETIVOS	33
a) Objetivo general	33
b) Objetivos específicos	34
IX. METODOLOGÍA	35
Base de datos	35
Generación del ROI	35
Procesamiento de imágenes	35
Aumento de datos	35
Extracción de características	36
Clasificación	36
Métricas	36
X. RESULTADOS	37
Generación de ROI	37
Procesamiento de imágenes	38
Aumento de datos	39
Extracción de características y clasificación	39
Métricas	41
XI. CONCLUSIONES	45
XII. REFERENCIAS BIBLIOGRÁFICAS	46

### ÍNDICE DE TABLAS

Tabla 1 Antecedentes de imágenes de termografía estática en cáncer de mama	8
Tabla 2 Antecedentes de imágenes de termografía dinámica en cáncer de mama	10

### ÍNDICE DE FIGURAS

Figura 1 Formas básicas de elementos estructurantes	18
Figura 2 Erosión de X por el elemento estructurante Y	18
Figura 3 Dilatación de X por el elemento estructurante Y.	19
Figura 4 Apertura morfológica del conjunto X por el elemento estructurante Y.	20
Figura 5 Cerradura morfológica del conjunto X por el elemento estructurante Y.	21
Figura 6 Clasificador	24
Figura 7 Representación de entradas (input) en red neuronal	25
Figura 8 Representación de red neuronal	26
Figura 9 Red neuronal convolucionada	27



Figura 10 Autoencoder

28

Figura 11 Diagrama de la metodología

31



## DATOS GENERALES

- Título del proyecto de Tesis: Metodología basada en inteligencia artificial para detección de anomalías asociadas al cáncer de mama en imágenes termográficas infrarrojas
- Nombre del alumno: Ing. Dulce Sofía Marín Chávez
- Número de expediente: 272464
- Programa de Estudios a realizar: Maestría en Ciencias en Inteligencia Artificial
- Director de Tesis: Dr. Luis Alberto Morales Hernández
- Codirector: Dr. Jesús Alberto Basurto Hurtado
- Secretario: Dr. Irving Armando Cruz Albarrán
- Vocal: Dr. Saúl Tovar Arriaga
- Lugar donde se realizará la investigación: Universidad Autónoma de Querétaro, Campus Aeropuerto.
- Línea de investigación: Tecnológica
- Horario de trabajo: lunes a viernes de 7:00 am – 15:00 pm.

## I. RESUMEN

El cáncer de mama es una de las causas más comunes de muerte en mujeres a nivel mundial, con más de 2.3 millones de casos al año y se espera un aumento a 3.2 millones en 2040. Existen diferentes técnicas de diagnóstico incluyendo la mastografía, el ultrasonido, muestras de biopsia y termografía.

La termografía registra la variación de temperatura en la superficie del cuerpo, es rápida, no invasiva, sin radiación y de bajo costo. Esta técnica se divide en estática y dinámica, siendo esta última menos explorada. Varios trabajos han utilizado termografía estática con éxito, pero surge la necesidad de llevar a cabo más investigación en la versión dinámica, al igual de métodos efectivos para la detección temprana del cáncer de mama, por lo que este trabajo propone un sistema basado en inteligencia artificial para la clasificación de imágenes termográficas buscando mejorar la precisión y sensibilidad de efectividad del diagnóstico, siendo una respuesta innovadora a la demanda social y médica, por lo tanto, el escaso uso de la termografía dinámica resalta un área de enfoque de futuras investigaciones.

## II. INTRODUCCION

El cáncer de mama es la segunda causa más común de mortalidad por cáncer en mujeres. Según la Organización Mundial de la Salud (OMS), hay más de 2.3 millones de casos al año y se espera que esta cifra supere los 3.2 millones para 2024. Las técnicas más comunes para obtener imágenes son la mamografía, el ultrasonido y la resonancia magnética.

En mamografía, la imagen se produce utilizando radiación, lo que permite detectar calcificaciones y masas en el tejido con una sensibilidad del 85%. Según Jochelson et al. (2012), el uso de agentes de contraste aumenta la calidad de las imágenes para el diagnóstico. Sait (2024) utilizó un modelo de red neuronal convolucional (CNN) con los pesos de EfficientNet B7 para extraer características de las imágenes y un modelo LightGBM para la clasificación de imágenes, logrando una precisión promedio del 98.7% y un índice kappa promedio del 95.8%. Hassan (2024) introdujo un marco CAD completamente automatizado que aprovecha la red YOLOv4 y los Vision Transformers (ViT) para detectar y clasificar masas en imágenes de mamografía espectral mejorada por contraste (CESM). La CESM es un tipo avanzado de mamografía digital de campo completo (FFDM) que ofrece una mejor visualización del tejido mamario. El modelo logró puntuaciones promedio de precisión (mAP) del 98.69%, 81.52% y 71.65% en detección de masas, y precisiones de clasificación de masas del 95.65%, 97.61% y 80% en los conjuntos de datos INbreast, CE-CESM y DM-CESM, respectivamente.

Aguerchi (2024) presentó un nuevo enfoque de aprendizaje profundo que utiliza redes neuronales convolucionales (CNNs) para la detección de cáncer de mama mediante imágenes de mamografía. Aunque las CNNs son efectivas para la clasificación de imágenes, la selección de hiperparámetros y arquitecturas óptimas sigue siendo un desafío. Para abordar esto, los autores emplearon el algoritmo de optimización por enjambre de partículas (Particle Swarm Optimization, PSO) para



determinar automáticamente los hiperparámetros y la arquitectura adecuada del modelo CNN. La CNN optimizada alcanzó altas tasas de precisión del 98.23% en el conjunto de datos DDSM y del 97.98% en el conjunto MIAS.

Otro trabajo es el de Cai et al. (2021), quienes trabajaron con imágenes de mamografía e implementaron segmentación mediante una matriz de simultaneidad de combinación de niveles de gris, la transformada de onda discreta y clasificación usando una red convolucional optimizada con un intercambio térmico avanzado, logrando una precisión del 93.79%. Abdelrahman et al. (2021) aplicaron técnicas para clasificar la densidad y simetría de las imágenes, explorando la predicción de tipos de masas, como las precancerosas y cancerosas, y mejoraron la clasificación en tejido mamario denso.

Aunque se han realizado varios estudios en imágenes de mamografía, es importante mencionar que una de las desventajas de esta técnica es que está contraindicada para personas menores de 40 años y para mujeres con tejido mamario muy denso. Además, el uso de radiación para obtener imágenes es una técnica dolorosa y desagradable, lo que resalta la necesidad de buscar otras alternativas para combatir esta enfermedad.

### III. ANTECEDENTES

El cáncer de mama es una de las causas de muerte más comunes en mujeres alrededor del mundo, de acuerdo con la Organización Mundial de la Salud (OMS, 2023) existen más de 2.3 millones de casos al año, a lo cual se espera que tenga un aumento a 3.2 millones en 2040. Además, de acuerdo con el Instituto Nacional de Estadística y Geografía (INEGI, 2022) se tiene una tasa de 0.88 muertes por cada 10 mil mujeres con rango de 30 a 59 años siendo la principal causa de defunción y 4.81 casos por cada 10 mil mujeres mayores a 60 años. A continuación, se hará un análisis de los diferentes trabajos que se encargan del diagnóstico del cáncer de mama, utilizando diferentes técnicas como mastografía, ultrasonido, muestras de biopsia y termografía.

#### Mastografía

Las técnicas más comunes para la obtención de imágenes son la mastografía, ultrasonido y resonancia magnética. En la mastografía se utiliza radiación para la formación de la imagen, esta detecta calcificaciones y masas en el tejido, con una sensibilidad del 85%, de acuerdo con Jochelson et al. (2012) al hacer uso del medio de contraste este incrementa la calidad de la imagen para su diagnóstico. Otro trabajo es el Cai et al. (2021) quien trabajó con imágenes de mastografía e implementó una segmentación utilizando una matriz de concurrencia de combinación de niveles de grises, la discreta de Wavelet y una clasificación haciendo uso de una red convolucionada optimizada con un Advanced Thermal Exchange, obteniendo una exactitud del 93.79%. También, se encuentra el trabajo de Abdelrahman et al. (2021) quien aplicó técnicas para clasificar la densidad y simetría de las imágenes, también explorar la predicción de tipos de masas como precancerígenas y cancerígenas, así como mejorar la clasificación en tejido mamario denso. A pesar de que se han hecho diversos trabajos de imágenes de mastografía es importante mencionar que una de las desventajas es que esta

contraindicada para personas menores de 40 años, así como en mujeres de tejido mamario muy denso, además del uso de radiación para la obtención de imágenes y es una técnica dolorosa e incómoda, por lo tanto, es necesario buscar otras alternativas para la detección de esta enfermedad.

### Ultrasonido

De acuerdo con Ayana et al. (2021) el ultrasonido es una técnica no invasiva y sin uso de radiación, la cual no produce efectos negativos en la salud de los pacientes, es un estudio más económico a comparación de la mastografía y es un aparato portable. El trabajo de Zhang et al. (2021) propuso el modelo de Multi-task learning (SHA-ML), el cual es una red convolucional densa codificadora basada en un modelo de atención suave y fuerte para imágenes de ultrasonido mamario, haciendo uso de una segmentación y una clasificación binaria. Otro trabajo Jia et al. (2023) comparó diferentes modelos de aprendizaje profundo y encontró que DenseNet121 presento una precisión de 79.5%, sensibilidad de 90.7% y especificidad de 65.9%. El trabajo de Du et al. (2022) utilizó el modelo de Efficient-Det para identificar el área de los tumores comparando la precisión de forma manual 95.3%, el método tradicional de ingeniería 90% y Efficient-Det 92.6%. Sin embargo, el ultrasonido necesita de una validación de los resultados mediante una mastografía o un estudio patológico complementario.

### Muestras de biopsia

El análisis de imágenes de biopsia del tejido mamario es una técnica de diagnóstico la cual brinda un resultado más certero de la enfermedad dando las características del tumor y estructura de las células cancerígenas (Gurcan et al., 2009). En este mismo sentido, se encuentra el trabajo de Hamilton et al. (1994), uno de los pioneros de la técnica de análisis de imágenes patológicas, quien propuso un modelo de sistema experto para el diagnóstico de citología, el cual consistió en una red Bayesiana para la clasificación de imágenes. Otro trabajo de Fu y Dong (2022) propuso un método de red neuronal de picos (Spiking neural network) para la detección de objetos en imágenes de ultrasonido y utilizó el convertir una red neuronal profunda en una red neuronal de picos, la cual obtuvo precisión de 90.6% en ubicación y 92.8% en clasificación del tumor. También se encuentra el trabajo de Alanazi et al. (2021) en donde se realizó una comparación de modelos de CNN en el cual el que presentó mejor resultado fue de 5 capas con una precisión del 87%. Sin embargo, esta técnica tiene como desventaja el uso de equipo patológico costoso, mayor incluso que el ultrasonido y mastografía.

### Termografía

La termografía es una técnica de imágenes médicas que registran la variación de la temperatura de la superficie del cuerpo humano en función de la radiación infrarroja emitida por la superficie del cuerpo. En aplicaciones médicas, se aplica a los cambios de temperatura en el cuerpo humano las cuales se observan en las imágenes y pueden ser interpretadas por especialistas. Las células cancerígenas generan calor debido a la liberación de óxido nítrico en la sangre, la cual causa alteración en la circulación, vasodilatación al aumentar la circulación



sanguínea, neoangiogénesis, la creación de nuevos vasos sanguíneos para suministrar nutrientes al tumor y aumento de actividad metabólica de las células cancerígenas (Kakileti et al., 2017).

El uso de la termografía para toma de imágenes funciona como una alternativa a cualquier otro método de toma de imágenes ya que es una técnica rápida, no invasiva, sin uso de radiación, sin contacto y de bajo costo, la cual consta de una toma de mapeo de la zona mamaria haciendo uso de una cámara termográfica y construye la imagen en la que se puede observar el cambio de temperatura entre los tejidos, en donde se identifican las zonas con tumor cancerígeno con una temperatura mayor al resto del tejido (Ekici y Jawzal 2020a).

Existe dos versiones de la termografía: estática y dinámica, en donde la estática es la imagen con un mapa corporal de la temperatura, la cual no considera su variación en el tiempo, al igual que se requiere de condiciones ambientales rígidas y más tiempo en que el paciente se aclimate a la sala de exploración, por el contrario la dinámica utiliza un estrés térmico frío en el área de interés (De Weerd et al., 2011; Shada et al., 2013), en el cual se monitorea la respuesta de la temperatura de la piel después del estrés dinámico donde el flujo de aire frío se dirige al seno y se puede observar que en los vasos sanguíneos producidos por el tumor no suelen tener terminaciones nerviosa por lo tanto, no responden al estímulo por cambio de temperatura y el tumor permanece prácticamente sin cambios cuanto este se enfría (Resmini et al., 2021).

Entre los trabajos de termografía estática se encuentra el de Etehadtavakol et al. (2013) en donde la base de datos consistió en 9 de tumor maligno, 12 benignas y 11 normales, las cuales registró utilizando sus límites inferiores por detección de bordes, segmentó usando fuzzy c-means clustering para la extracción de características y un clasificador Adaboost, obteniendo una precisión del 95% en tumores malignos. Otro trabajo es el de Sathish et al. (2017) en donde realizó una segmentación automática de las mamas utilizando las características de forma de la mama y la textura basada en matrices de coocurrencia de niveles de gris y el ajuste de curva polinomial, obteniendo 90% de precisión, 87.5% de sensibilidad y 92.5% de especificidad. El trabajo de Singh et al. (2023) utilizó un modelo Non-Dominanted Sorting Genetic Algorithm NSGA-II y una optimización Bayesiana obteniendo una precisión de 96.16%, mencionó que el estudio consideró los pesos de los clasificadores de forma individual, a lo que se puede mejorar al tratar los pesos de los clasificadores individualmente como hiper parámetros ajustables. Mammoottil et al. (2022) realizó un modelo de aprendizaje automático basado en redes neuronales convolucionales que utilizan múltiples vistas térmicas de la mama, utilizando un conjunto de datos visuales y verificándolos con los datos clínicos, obteniendo una precisión de 93.8%. Otro trabajo es el de Yadav y Jadhav (2022) donde se utilizaron el modelo InceptionV3 en el cual aumentaron los datos de entrenamiento y prueba obteniendo 92.3% de precisión, mencionaron que se podría aplicar una técnica de agregación de características como la ultra agregación de multicontexto para mejorar las semejanzas y aumentar la base de datos. En el trabajo de Krishna y George (2021) se mencionó que uno de los principales problemas de clasificación en imágenes de termografía es la mala relación señal a ruido y la segmentación ineficiente en la región mamaria, para lo cual propuso una técnica de preservación de bordes basado en filtros de coexistencia para el

preprocesamiento del módulo, utilizando operaciones morfológicas y evolución del conjunto de niveles de regularizador de distancia, utilizó un módulo de extracción de características basado en transformadas de wavelet para la segmentación de la región mamaria. El trabajo de Sánchez-Cauce et al. (2021) realizó una red neuronal convolucionada (CNN) en la cual incluyó tres imágenes por cada paciente, plano frontal, 90° lateral derecho y 90° lateral izquierdo, donde menciona que al incluir a la CNN la información personal y clínica de cada paciente la precisión aumenta a 97%.

A continuación, se muestra en la tabla 1 trabajos de imágenes de termografía estática para detección de cáncer de mama.

*Tabla 1 Antecedentes de imágenes de termografía estática en cáncer de mama*

Autor	Aportación	Base de datos	Segmentación	Clasificación de Algoritmo	Métricas
<b>Alfayez et al 2019</b>	Diseño de	Pública DMR-IR 1345	Geometrical and textural features	Extreme Learning Machine (ELM) and Multilayer Perceptron (MLP)	Accuracy- 82.2%
<b>(Ekici &amp; Jawzal, 2020)</b>	Uso del algoritmo de Bayer en CNN para optimización.	140 pacientes (48 sanos y 32 cáncer)	Translation Centering, eliminación de ruido Salt and noise, Estructura de objetos y firma espectral.	Convolution neural network(CNN) optimized by Bayes Algorithm	Accuracy- 98.95%
<b>(Rastghalam &amp; Pourghassem, 2016)</b>	Los patrones normales y anormales se separaron entre sí usando características de textura.	65 imágenes	Textura de imágenes: Hidden Markov Model (HMM) Local Binary Pattern (LBP) Markov Random Field (MRF-based)	Random Field (MRF-based)	Falsos negativos 8.3% Falsos positivos 5%
<b>(Singh et al., 2023)</b>	Uso del modelo NSGA2-v2-CXL en la CNN	Base de datos BreakHis 7909 imágenes		Convolutional neural network (CNN) Ensemble of gradient-boosting Algorithm Inception-ResNet-v2, modelo	Accuracy 94.40% Precision 95.77% Recall 99.29%



				NSGA2-IRv2-CXL	
<b>(Recinella et al., 2020)</b>	Detección de cáncer de mama al generar modelo 3D de las mamas	-	-	-	-
<b>(Krishna &amp; George, 2021)</b>	Forma accesible y portátil del control remoto de la salud mamaria.	71 pacientes de los cuales 34 son sanos y 37 cáncer de mama.	Transformada de Wavelet, Operaciones morfológicas y Distance Regularized Level Set Evolution (DRLSE)	K-Nearest Neighbor (KNN), Support Vector Machine (SVN) and Naives Bayes Classifiers	Accuracy 96.46%
<b>(Yadav &amp; Jadhav, 2022)</b>	Uso de Machine Learning como técnica estadística para programas de software sin ser codificador directamente.	Base de datos PROENG 67 pacientes de los cuales 43 son sanos y 24 presentan cáncer de mama.	Contrasts-enhancement, se le cambio el tamaño y una normalización	Convolutional neural network (CNN) usando el modelo VGG16 y InceptionV3-93.1%	GG16 Accuracy 87.3% InceptionV3
<b>(Fernandez-Ovies et al., 2019)</b>	Compararon Resnet como Resnet18, Resnet34, Resnet34 y Resnet50 para obtener la mejor exactitud	216 pacientes de los cuales 175 sanos y 41 presentan cáncer de mama.	-	Convolutional neural network (CNN, Resnet152	

Por otro lado, entre los trabajos de termografía dinámica se encuentra el de Cary et al. (1975) el cual utilizó enfriamiento local para diferenciar entre tumor maligno y benigno, si la diferencia de temperatura entre ambas regiones bajo enfriamiento era superior a 0.9°C se clasificó como maligno y benigno si la diferencia era inferior a 0.9°C, obtuvo como resultados de clasificación 74% de precisión en tumor maligno y 88% entre pacientes sanos y tumores benignos. Francis et al. (2014) utilizó imágenes de 24 pacientes sin cáncer de mama y 12 con cáncer, extrajo 17 diferentes características como estadísticas de primer y segundo orden,



así como de textura, posteriormente utilizó una máquina de vectores para clasificar, obteniendo una especificidad del 83.35% y sensibilidad de 83.3%. Gerasimova et al. (2014) clasificó 33 pacientes con cáncer de mama y 14 pacientes sin cáncer, en donde encontró que la fluctuación temporal cerca del tumor cancerígeno es mayor en una mama sana, hizo uso de análisis multifractal Wavelet para clasificar entre pacientes con cáncer de mama y pacientes sanos. Otro trabajo es el de Ali et al. (2015) en donde hizo uso del método estático y dinámico, posteriormente aisló las mamas y sacó su histograma para mejorar las imágenes, extrajo sus características de las cuales 6 eran estadísticas de primer orden y 15 de segundo orden, finalmente utilizó el método de Support Vector Machine (SVM) como clasificador obteniendo una precisión del 100%. Como puede apreciarse son pocos los trabajos que han utilizado termografía dinámica por lo tanto es necesario desarrollar nuevas metodologías que utilicen esta versión dinámica para la detección de anomalías que puedan estar asociadas al cáncer de mama.

A continuación, se muestra en la tabla 2 trabajos de imágenes de termografía dinámica para detección de cáncer de mama.

*Tabla 2 Antecedentes de imágenes de termografía dinámica en cáncer de mama*

Autor	Aportación	Base de datos	Segmentación	Clasificación de Algoritmo	Métricas
Mammoot til et al. (2022)	Modelo de aprendizaje supervisado en redes neuronales convolucionales utilizando múltiples vistas térmicas de la mama	293 pacientes con imágenes frontales y laterales	-	CNN con optimizador Adam	Especificidad 96.7%, sensibilidad 88.9%
<b>Sarigoz &amp; Ertan (2020)</b>	El uso de termografía dinámica fue superior al ultrasonido, mamografía y resonancia magnética en detección de ganglios linfáticos metastásicos	26 pacientes (2 imágenes estáticas y dinámicas cada 5 segundos por 4 minutos)	-	-	Sensibilidad- 83%
da Silva et al. (2020)	Evalúo un método de análisis de imágenes para detección automática de tumores benignos y malignos	64 pacientes y 1280 imágenes totales	Clustering, histograma y geometría fractal	Clasificador Support Vector Machine (SVM)	Precisión - 100%



Abdel-Nasser et al. (2019)	Utilizó una técnica de aprendizaje de clasificación y análisis de textura	37 pacientes con cáncer de mama y 19 sanos	histograma	Perceptrón multicapa con método de análisis de textura	AUC-98.9%
<b>Silva et al. (2016)</b>	Segmentación manual de la región de las mamas	40 pacientes con cáncer de mama y 40 pacientes sanos	Algoritmo k-means	Red bayesiana	Especificidad del 100%.
Saniei et al. (2015)	Detección mediante el cálculo de matching score y extracción de patrones vasculares	25 pacientes con cáncer de mama y 25 pacientes sanos	Filtro de difusión anisotrópico y el método black top-hat	-	Sensibilidad- 86% y especificidad 61%

Adicionalmente hay otros trabajos adicionales en termografía; los cuales son:

Gomathi (2023) presentó un enfoque llamado DBC-4D U-Net-DITI, que utiliza segmentación 4D U-Net con sistemas de imágenes térmicas digitales infrarrojas (IR) para el diagnóstico de cáncer de mama. Este método demostró mejoras en el rendimiento, alcanzando tasas de precisión del 39.01%, 28.34% y 37.45%, y tasas de precisión mejoradas del 17.12%, 24.12% y 32.07%.

Chebbah (2023) utilizó el modelo U-Net, logrando una intersección sobre unión (Intersection over Union, IoU) del 89.03%. Después de la segmentación, se realizaron evaluaciones texturales y análisis de la red vascular en los termogramas para extraer características relevantes. Estas características se utilizaron en clasificadores basados en algoritmos de aprendizaje supervisado para distinguir entre termogramas normales y anormales. Al aplicar el enfoque desarrollado con una máquina de soporte vectorial (Support Vector Machine, SVM), se alcanzaron una precisión del 94.4%, una exactitud del 96.2%, un recall del 86.7%, una puntuación F1 del 91.2% y una tasa de verdaderos negativos del 98.3%.

Otro trabajo es el de Mishra et al. (2020), que utilizó termogramas mamarios de 56 sujetos. Las características de textura se extrajeron de estas imágenes utilizando métodos como la matriz de longitud de carrera de nivel de gris (Gray Level Run Length Matrix, GLRLM) y la matriz de concurrencia de nivel de gris (Gray Level Co-occurrence Matrix, GLCM). Al analizar la correlación de estas características, el estudio estableció relaciones lineales entre variables, lo que ayudó a evaluar cuantitativamente los datos. Se aplicaron técnicas de reducción de características no supervisadas, como el análisis de componentes principales (Principal Component Analysis, PCA) y Autoencoder (AE), para seleccionar las características más relevantes en la detección de anomalías entre tejido mamario sano y no sano. Entre los clasificadores probados, el bosque aleatorio (Random Forest) combinado con PCA alcanzó la mayor precisión, logrando un 95.45% en la distinción entre tumores benignos y malignos.

#### IV. JUSTIFICACIÓN

La necesidad de desarrollar métodos efectivos y tempranos para la detección del cáncer de mama se deriva de su alta mortalidad a nivel global. Este proyecto, centrado en la clasificación de imágenes termográficas de cáncer de mama mediante el uso de Inteligencia Artificial (IA), emerge como una respuesta innovadora ante esta demanda social. La implementación de técnicas avanzadas de procesamiento de imágenes y aprendizaje automático no sólo será una herramienta para un diagnóstico más acertado y rápido, sino que también promete aligerar la carga operacional en los profesionales de la salud, lo cual podría ayudar en una mejora en la calidad de la atención médica, lo cual es una forma para avanzar en la lucha contra el cáncer de mama.

Como se vio en la revisión en el estado del arte, existen diferentes técnicas para la detección del cáncer de mama, sin embargo como se mencionó la mastografía es una técnica dolorosa, requiere contacto, uso de radiación, costosa y sólo se puede realizar en personas mayores de 40 años, el ultrasonido necesita de una validación complementaria del estudio mediante una mastografía o biopsia y el uso de esta última es un estudio invasivo y costoso.

El presente trabajo pretende promover el bienestar en todas las edades ya que para este estudio no es necesario que el paciente tenga una edad a partir de los 40 para poder realizarse como lo es en la mastografía, al igual que proporcionar una estrategia de diagnóstico anticipado y más accesible para el cáncer de mama, este proyecto pretende detectar anomalías que puedan relacionarse al cáncer de mama.

#### V. DESCRIPCIÓN DEL PROBLEMA

De acuerdo con el INEGI (2023) en México se registraron 23790 nuevos casos de cáncer de mama en 2022 en población mayor a 20 años, de los cuales se registró 27.64 casos por cada 100 mil personas, donde la cantidad en mujeres es de 51.92 y 1.25 en hombres, catalogándolo como uno de los tipos de cáncer más frecuentes y causa de muerte en mujeres alrededor del mundo.

Los métodos convencionales de detección enfrentan desafíos como los altos costos de un estudio de mastografía, una biopsia que necesita de equipo especializado de alto costo, al igual que la incapacidad para identificar tumores en etapas tempranas de manera eficaz y los riesgos o incomodidades asociadas con procedimientos invasivos o la exposición a la radiación. Lo cual es una problemática para la detección temprana del cáncer de mama por lo que es un elemento crucial para mejorar las tasas de supervivencia y asegurar un tratamiento efectivo y menos costoso.

La termografía es una opción prometedora debido a su carácter no invasivo, sin radiación, sin contacto, sin dolor y con la posibilidad de identificar cambios fisiológicos asociados con la presencia de células cancerígenas. No obstante, la interpretación precisa de estas imágenes puede ser una tarea desafiante que requiere un alto grado de especialización y experiencia.

Al revisar los trabajos en el estado del arte, se puede determinar la falta de aplicación de una red convolucional con una base de datos utilizando diferentes





ángulos por paciente, el uso de termografía dinámica ayudaría a mejorar el contraste de las imágenes para una clasificación con mayor precisión, proponer el desarrollo de un algoritmo para la detección del área donde se encuentra la anomalía de forma automática, al igual que proponer una caracterización basada en texturas y la geometría de las imágenes para reducir el tiempo de ejecución computacional y aumentar la precisión de clasificación de las imágenes.

El propósito es superar las limitaciones actuales en la detección del cáncer de mama mediante la creación de un sistema robusto que pueda identificar indicativos tempranos de cáncer de manera más eficaz y precisa, proporcionando así una herramienta valiosa para los profesionales médicos y mejorando la accesibilidad y efectividad de la detección del cáncer de mama. Esta problemática se encuentra como una necesidad social y médica apremiante, y aborda una oportunidad tecnológica emergente para mejorar significativamente la atención sanitaria en el ámbito de la oncología mamaria

## VI. FUNDAMENTACIÓN TEÓRICA

### Termografía

El calor es una forma de energía que es emitida por el movimiento molecular dentro de un sistema, la termografía es una técnica la cual permite visualizar la distribución de la temperatura en superficies, objetos y cuerpos. La intensidad de radiación emitida en función de la temperatura se determina de la siguiente manera:

$$I(\lambda, T) = \frac{2\pi hc^2}{\lambda^4} \left( \frac{1}{e^{\frac{hc}{\lambda kT}} - 1} \right) \quad (1)$$

Dónde:

$\lambda$  es la longitud de onda.

$h$  es la constante de Planck.

$c$  es la velocidad de la luz en el vacío.

$k$  es la constante de Boltzman.

La potencia de radiación emitida de acuerdo con la ecuación de Stefan-Boltzman es la siguiente:

$$E = \epsilon \sigma T^4 \quad (2)$$

Dónde:

$E$  es la energía total emitida.

$\sigma$  es la constante de Boltzman.

$T$  es la temperatura absoluta.

$\epsilon$  es la emisividad de las superficies de radiación.

La ecuación nos afirma que la potencia de la onda emitida se relaciona directamente con la temperatura de la piel.

La termografía es una técnica de imágenes médicas que registran la variación de la temperatura de la superficie del cuerpo humano en función de la radiación infrarroja emitida por la superficie del cuerpo.

En aplicaciones médicas, se aplica a los cambios de temperatura en el cuerpo humano las cuales se observan en las imágenes y pueden ser interpretadas por especialistas.

Las células cancerígenas generan calor debido a la liberación de óxido nítrico en la sangre, la cual causa alteración en la circulación, vasodilatación al aumentar la circulación sanguínea, neoangiogénesis, la creación de nuevos vasos sanguíneos para suministrar nutrientes al tumor y aumento de actividad metabólica de las células cancerígenas (Kakileti et al., 2017).

Como se mencionó anteriormente, la termografía estática se refiere a la captura de imágenes es un momento específico y sin cambios, por otro lado, la dinámica se refiere a imágenes en movimiento para observar el cambio en la temperatura a lo largo del tiempo (De Weerd et al., 2011). Para poder hacer uso de la termografía estática es necesario estabilizar la temperatura corporal en un lapso de 10 a 15 min para poder hacer la toma de las imágenes. Al contrario de la termografía dinámica, esta es necesario realizar un enfriamiento del tejido mamario hasta determinada temperatura y realizar la toma de imágenes (Gonzalez-Hernandez et al., 2019).

### Segmentación de imágenes

Es el proceso de dividir una imagen en regiones para cambiar la representación de esta en algo que sea más fácil de analizar para obtener la información en la región de interés, esto es necesario para identificar el contenido de la imagen, para ello se usan diferentes métodos para clasificar dentro de una imagen recorriendo píxel por píxel (Liu et al., 2021; Saroha et al., 2013).

Este proceso se puede alcanzar combinando las diversas técnicas, por tal motivo, hay varias metodologías que ayudan a lograr la segmentación, ya que no todos los métodos son adecuados para un tipo particular de imagen. La segmentación se clasifica de acuerdo con la región, bordes, umbral y agrupación por ciertas características (Kuruvilla et al., 2016).

### Filtrado

En 1980 Grossman y Morlet desarrollaron una función cuadrada integrable y Mallat encontró una relación entre los filtros de cuadratura, los algoritmos piramidales y las bases ortonormales, así es como Daubichies realiza una base ortogonal la más utilizada hoy en día, lo que facilita la transformada directa e inversa para una señal o imagen (Palomares et al., 2016).

El ruido presente en las imágenes, que puede alterar y distorsionar su contenido, suele originarse por los diferentes medios de transmisión o almacenamiento. El ruido gaussiano se utiliza como un modelo para representar la aleatoriedad de los valores de los píxeles. Su función de densidad de probabilidad  $p_q(x)$  se describe a partir del promedio ( $\mu$ ) y la varianza ( $\sigma^2$ ) de una variable aleatoria, tal como se expresa en la ecuación 3.

(3)



$$p_q(x) = (2\pi\sigma^2)^{-1/2} e^{-(x-\mu)^2/2\sigma^2}$$

Existen diversos tipos de filtros. Entre ellos, los filtros espaciales, que pueden ser lineales, como los de promedio y gaussiano, o no lineales, como los de mediana y sigma.

Por otro lado, los filtros frecuenciales se emplean principalmente para minimizar el error cuadrático medio y mejorar la calidad de la imagen. Su aplicación requiere calcular el espectro de energía de la imagen, tal como se muestra en la ecuación 4.

$$H(u, v) = \frac{D * (u, v)}{D * (u, v)D(u, v) + \frac{S_w(u, v)}{S_f(u, v)}} \quad (4)$$

Donde:

$S_w(u, v)$  es el espectro de energía de la imagen ideal.

$S_f(u, v)$  es el ruido.

$D(u, v)$  es la estimación de una función de degradación y su conjugado  $D * (u, v)$ .

El filtro sigma es muy efectivo para la eliminación del ruido gaussiano, como su principal característica, esta conserva los bordes por técnicas difusas.

Se puede resolver el problema del ruido mediante el uso de redes neuronales, brindando eficientes resultados en la segmentación, clasificación, identificación de imágenes, etc. (Ortiz Rangel et al., 2016).

## Ruido

Conocido como “datos sin significado o que no se utilizan y se producen como un subproducto no deseado”. Las características del ruido dependen de su fuente, así como del operador que mejor reduce sus efectos. “El ruido de impulso representa picos aleatorios de energía que ocurren durante la transferencia de datos de una imagen” (Kumar & Kumar, 2015). Para generar ruido, se daña un porcentaje de la imagen cambiando un punto de canal seleccionado aleatoriamente a un valor de píxel entre 0-255.

El modelo de ruido  $I_n$  es representado como se observa en la ecuación 5.

$$I_n(i, j) = \left\{ \begin{aligned} &I(i, j) - x \geq p(I_r(i, j), I_g(i, j), z) \text{ y } x < \frac{1}{3}x \\ &< p(I_r(i, j), z, I_b(i, j)) (z, I_g(i, j), I_b(i, j)) \frac{1}{3} \leq y \leq \frac{2}{3} \frac{2}{3} \\ &\leq y \text{ } x < p \text{ } x < p \end{aligned} \right\} \quad (5)$$

Donde:

$I$  es la imagen original.



$I_r$  Componentes del color rojo en la imagen original.  
 $I_g$  Componentes del color verde en la imagen original.  
 $I_b$  Componentes del color azul en la imagen original.  
 $x, y = [0,1]$  son valores aleatorios entre  $z = [0,255]$ .  
 $p = [0,1]$  son parámetros que representan la probabilidad de ruido en la imagen.

## Métodos de umbralización

La umbralización tiene como propósito transformar una imagen en escala de grises en una imagen binaria, es decir, de dos niveles, con el fin de distinguir el objeto del fondo. Este proceso permite segmentar los píxeles de acuerdo con su intensidad en escala de grises, y puede realizarse mediante diferentes enfoques: umbral global, local o adaptativo. En particular, el umbral global genera una imagen cuyos niveles de gris se encuentran entre 0 y 1, produciendo finalmente una imagen binaria. La imagen se debe fragmentar en subregiones para poder alcanzar un umbral, la asignación del umbral local no considera el tamaño o forma del algoritmo y se le da un umbral a cada subregión. (Ochoa González et al., 2019)

A continuación, se describirá el método de Otsu el cual trabaja con un umbral óptimo como se mostrará a continuación.

## Método Otsu

El método de Otsu se basa en determinar un umbral óptimo que permite dividir la imagen en dos clases: una compuesta por los píxeles con valores de intensidad menores al umbral, y otra con aquellos mayores o iguales al mismo. Cuando la varianza entre los niveles de gris de ambas clases es máxima, se considera que se ha encontrado el umbral más adecuado. La probabilidad de ocurrencia del nivel de gris  $i$  en una imagen se expresa según la ecuación 6.

$$p_i = \frac{f_i}{N} \quad (6)$$

Donde:

$f_i$  es el número de píxeles

$N$  denota el nivel de gris en una imagen entre 1 y  $L$ .

$p_i$  probabilidad de ocurrencia.

Cuando se da una umbralización de dos niveles en una sola imagen, los píxeles se dividen en  $C_1$  con niveles de gris de  $[1, \dots, t]$  y  $C_2$  con  $[t + 1, \dots, L]$ . Completando la distribución de probabilidad de los niveles de gris para  $C_1$  y  $C_2$  como se muestra en las Ecuaciones 6- 17.

$$C_1: \frac{p_1}{\omega_1(t)}, \dots, \frac{p_t}{\omega_1(t)} \quad (7)$$

$$C_2: \frac{p_{t+1}}{\omega_2(t)}, \frac{p_{t+2}}{\omega_2(t)}, \dots, \frac{p_L}{\omega_2(t)} \quad (8)$$



Donde:

$f_i$  es el número de píxeles

$N$  denota el nivel de gris en una imagen entre 1 y  $L$ .

$p_i$  probabilidad de ocurrencia.

$$\omega_1(t) = \sum_{i=1}^t p_i \quad (9)$$

$$\omega_2(t) = \sum_{i=t+1}^L p_i \quad (10)$$

La media para la clase  $C_1$  y  $C_2$  está dada por:

$$\mu_1 = \sum_{i=1}^t \frac{ip_1}{\omega_1(t)} \quad (11)$$

$$\mu_2 = \sum_{i=t+1}^L \frac{ip_1}{\omega_2(t)} \quad (12)$$

Donde  $\mu_T$  es la intensidad media de toda la imagen, la cual se calcula de la siguiente manera:

$$\mu_T = \omega_1\mu_1 + \omega_2\mu_2 \quad (13)$$

$$\omega_1 + \omega_2 = 1 \quad (14)$$

Se define la varianza entre clases de la imagen umbralizada como:

$$\sigma_B^2 = \omega_1(\mu_1 - \mu_T)^2 + \omega_2(\mu_2 - \mu_T)^2 \quad (15)$$

Para una umbralización de dos niveles, Otsu verificó que el umbral óptimo  $t^*$  se elige de manera que  $\sigma_B^2$  sea máxima; esto es:

$$t^* = \text{Max}\{\sigma_B^2(t)\} \quad (16)$$

$$1 \leq t \leq L \quad (17)$$

### Operaciones morfológicas

Toda operación morfológica se obtiene a partir de una o más operaciones entre conjuntos, como la unión, intersección o complemento, involucrando dos conjuntos —generalmente denotados como A y B— que son subconjuntos de un mismo espacio.

El propósito de las transformaciones morfológicas es extraer estructuras geométricas de los conjuntos sobre los que se trabaja, mediante la aplicación de otro conjunto denominado elemento estructurante. El tamaño y la forma de dicho elemento se seleccionan en función de la morfología del objeto que se desea analizar y de las características que se buscan resaltar o aislar. En la figura 2.3 se presentan los elementos estructurantes básicos más utilizados.

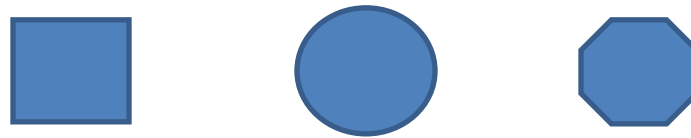


Figura 2.3 Formas básicas de elementos estructurantes

### Erosión

En la teoría de retículos se describe a la erosión como una “operación que conmuta con el infinito” (Ortiz Zamora, 2002).

La erosión consiste en verificar si el elemento estructurante B se encuentra contenido dentro del conjunto A. Cuando esta condición no se cumple, el resultado es un conjunto vacío. En términos formales, la erosión de un conjunto A se define como el conjunto de puntos x que pertenecen a A, de manera que al trasladar el elemento estructurante B a esa posición, este queda completamente incluido en A. Este proceso se expresa matemáticamente en la ecuación 18.

$$\varepsilon_Y(X) = \{x | Y_x \subseteq X\} \quad (18)$$

En la ecuación 19 se puede observar la erosión en una imagen binaria de un elemento estructurante Y de tipo cuadrado.

$$\begin{aligned} &\{00000 \ 01110 \ 000 \ 110 \ 110 \ 101000 \} \\ &\rightarrow \{00000 \ 00000 \ 000 \ 000 \ 100 \ 00000 \} \end{aligned} \quad (19)$$

En la figura 2.4 se muestra el resultado de aplicar la erosión utilizando un elemento estructurante con forma de disco circular. Este tipo de elemento provoca que

desaparezcan las estructuras o detalles de la imagen que son más pequeños que el propio elemento estructurante, conservando únicamente las formas de mayor tamaño.

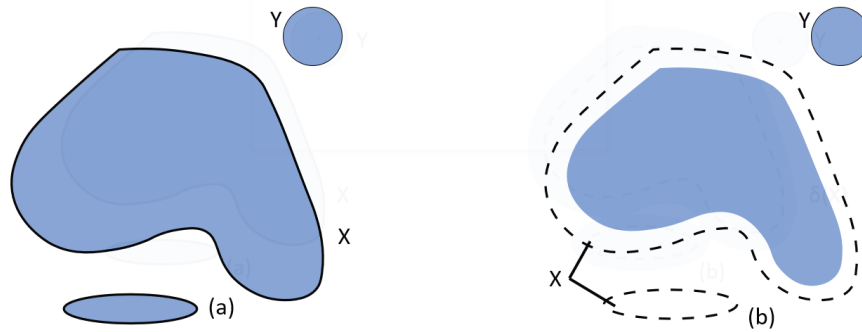


Figura 2.4 Erosión de X por el elemento estructurante Y

Se puede concluir que la erosión aplicada a señales bidimensionales en escala de grises produce una imagen con valores de intensidad menores, es decir, una imagen más oscura, ya que este proceso reduce los valores de la señal.

De este modo, en una imagen en niveles de gris, la erosión se interpreta como una contracción o reducción del objeto original, empleada comúnmente para disminuir su tamaño o eliminar detalles pequeños. En contraste, la dilatación actúa de manera opuesta, provocando una expansión o crecimiento de las regiones de la imagen.

### Dilatación

Operación dual de la erosión, en la teoría de los retículos, donde el resultado de la dilatación es el conjunto de los puntos de origen del elemento estructurante Y tal que el elemento estructurante contiene algún elemento del conjunto X, cuando el elemento se desplaza por el espacio que contienen a ambos conjuntos.

$$\delta(\vee_{i \in I} x_i) = \vee_{i \in I} \delta(x_i) \quad (20)$$

Donde:

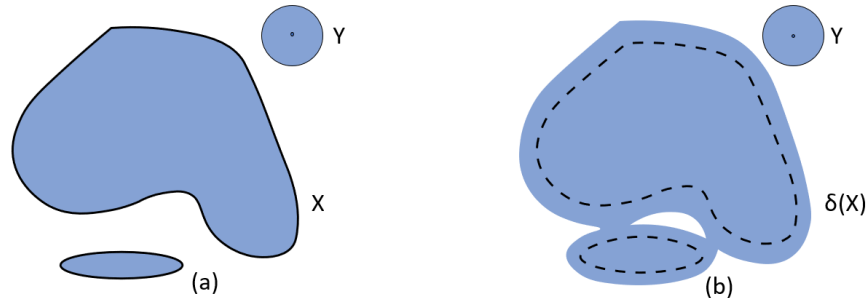
I es cualquier conjunto de índices.

$y\{x_i\}$  es una colección arbitraria de valores, de tal manera que  $x_i \in X$ .

En la ecuación 21 se puede observar la dilatación en una imagen binaria, tomando en cuenta un elemento estructural de Y.

$$\begin{aligned} &\{00000 \ 00000 \ 000 \ 000 \ 100 \ 00 \ 00 \ 00 \ 00 \} \\ \rightarrow &\{00000 \ 01110 \ 000 \ 110 \ 110 \ 10 \ 10 \ 00 \ 00 \} \end{aligned} \quad (21)$$

El efecto de la operación de dilatación se aprecia en la figura 2.5, donde se emplea un elemento estructurante con forma de disco circular. Este elemento provoca un aumento en el tamaño y la definición del objeto, permitiendo que sus bordes se expandan y que las formas se vuelvan más prominentes dentro de la imagen.



*Figura 2.5 Dilatación de X por el elemento estructurante Y.*

Como resultado de aplicar una dilatación sobre una imagen bidimensional en escala de grises, se obtiene una imagen más clara, ya que el proceso incrementa los valores de intensidad de los píxeles. Esto genera un efecto visual en el que los objetos más brillantes adquieren mayor definición y contraste en comparación con las zonas más oscuras.

En esencia, la dilatación actúa de forma opuesta a la erosión: mientras la erosión reduce o adelgaza las estructuras de la imagen, la dilatación expande y realza los objetos, aumentando su tamaño y presencia visual.

### Transformaciones morfológicas

En las transformaciones morfológicas se debe conformar y definir la imagen de partida para el proceso de segmentación y así poder tomar una correcta elección del elemento estructurante y la variable de gradiente. La elección de un gradiente ya sea por erosión o dilatación va a depender de la geometría y la luminosidad de los elementos a destacar en la imagen. (Ortiz Zamora, 2002)

### Apertura

La apertura de una señal  $f$  mediante un elemento estructurante  $B$  se representa como  $f \circ Y$  y se define como la erosión de  $f$  por  $Y$ , seguida de una dilatación con el mismo elemento estructurante. Este proceso combina ambas operaciones para suavizar los contornos, eliminar pequeñas irregularidades y preservar la forma general del objeto, tal como se expresa en la ecuación 22.

$$\gamma_Y(f) = \delta_Y(\varepsilon_Y(f)) \quad (22)$$

La apertura de una imagen es independiente del origen del elemento estructurante utilizado. Esto se debe a que, mientras la erosión puede interpretarse como una intersección de traslaciones del elemento estructurante, la dilatación posterior equivale a una unión de traslaciones en la dirección opuesta.



Por tanto, la apertura puede definirse como la unión de todos los elementos estructurantes  $Y$  que se encuentran totalmente contenidos dentro del conjunto  $X$ , tal como se describe en la ecuación 23.

$$\gamma_Y(X) = U\{Y | Y \subseteq X\} \quad (23)$$

En la figura 2.6 se muestra el resultado de una operación de apertura, en la cual se utiliza un elemento estructural con forma de disco. Durante la fase de erosión, este elemento provoca la eliminación de una estructura pequeña que, posteriormente, en la etapa de dilatación, no logra recuperarse completamente a su forma original. Este comportamiento refleja la naturaleza suavizadora de la apertura, que tiende a eliminar detalles finos o ruidos, preservando únicamente las formas más grandes y estables de la imagen.

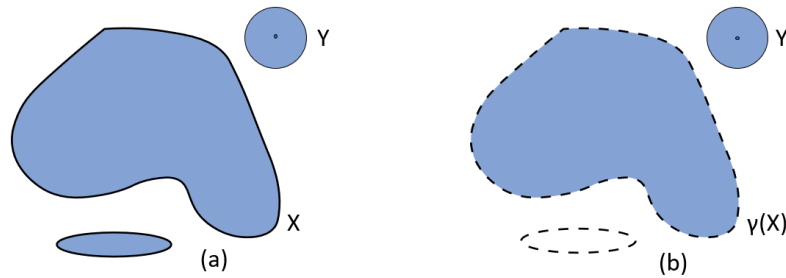


Figura 2.6 Apertura morfológica del conjunto  $X$  por el elemento estructural  $Y$ .

Se puede apreciar que, durante la apertura, se eliminan los objetos cuyo tamaño es menor al del elemento estructural utilizado. Como consecuencia, este proceso suaviza los contornos y redondea las convexidades más relevantes del objeto, contribuyendo a una representación más limpia y uniforme de las estructuras principales de la imagen.

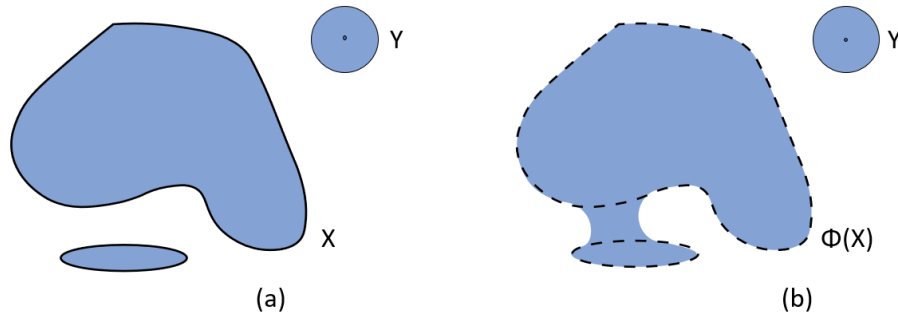
### Cerradura

La cerradura de una señal  $f$  mediante un elemento estructural  $Y$  se representa como  $f \bullet Y$  y se define como la dilatación de  $f$  por  $Y$ , seguida de una erosión con el mismo elemento estructural. Este procedimiento permite rellenar huecos, suavizar bordes y conectar regiones cercanas, tal como se expresa en la ecuación 24.

$$\phi_Y(X) = \cap \{Y^c | X \subseteq Y^c\} \quad (24)$$

Al realizar el cierre morfológico, se observa que, al igual que en la apertura, este proceso es independiente del origen del elemento estructural. En el caso de un conjunto  $A$  y un elemento estructural con forma de disco, como se muestra en

la figura 2.7, el cierre puede interpretarse como el espacio generado por el recorrido del elemento estructurante cuando este es forzado a mantenerse fuera de los límites del conjunto. Este procedimiento tiende a rellenar huecos pequeños y suavizar las concavidades, preservando la forma general del objeto.



*Figura 2.7 Cerradura morfológica del conjunto  $X$  por el elemento estructurante  $Y$ .*

Como se puede observar, en una imagen en niveles de grises, el cierre morfológico resulta útil para rellenar detalles finos, conectar objetos cercanos entre sí, suavizar los contornos, y cerrar pequeños vacíos o discontinuidades presentes en las formas. En conjunto, esta operación contribuye a obtener una imagen más continua y compacta, eliminando imperfecciones menores sin alterar significativamente la estructura principal del objeto.

### Operaciones morfológicas por reconstrucción

Las operaciones morfológicas constituyen un conjunto de técnicas empleadas para procesar imágenes a partir de sus formas geométricas. En este tipo de operaciones, cada píxel se modifica según los valores de sus vecinos, dentro de un entorno definido por un elemento estructurante. Al seleccionar apropiadamente la forma y el tamaño de dicho entorno, es posible diseñar operaciones morfológicas sensibles a patrones o estructuras específicas presentes en la imagen original.

La reconstrucción morfológica es una operación más avanzada que involucra dos imágenes y un elemento estructurante. En ella, se utiliza una imagen marcadora, que indica el punto de partida de la transformación, y una imagen máscara, que limita la extensión de la reconstrucción. El elemento estructurante define la conectividad entre píxeles. Una de sus principales aplicaciones es filtrar o eliminar regiones de la imagen que no estén contenidas en el elemento estructurante. En particular, la apertura por reconstrucción consiste en aplicar una apertura a la imagen marcadora y luego reconstruirla mediante dilataciones geodésicas iterativas, hasta que no se produzcan más cambios en la imagen.

Las operaciones por reconstrucción permiten preservar mejor las formas originales. Así, la apertura por reconstrucción tiende a suavizar los bordes y eliminar pequeñas protuberancias, mientras que la cerradura por reconstrucción actúa de forma opuesta, rellenando huecos y uniendo componentes cercanas.

En contraste, las operaciones morfológicas básicas, como la dilatación y la erosión, únicamente expanden o reducen las formas, sin conservar necesariamente la geometría original de los objetos en la imagen.

### Detección de bordes

Se puede definir como el cambio de líneas que marcan el límite y la división entre la apariencia de la imagen, esta utiliza un enfoque en la variación de la intensidad, con el propósito de extraer información necesaria como la nitidez para la ubicación de un objeto en la imagen (Vyas et al., 2018).

### Características de una imagen

Para llevar a cabo una clasificación de imágenes en una red neuronal, se necesita encontrar características en común de las diferentes imágenes, a continuación, se presentan algunas de las características más comunes:

### Contraste

Se refiere a aumentar el rango dinámico en los niveles de gris en las imágenes, es decir a la variación de intensidades lo que da como resultado una imagen con mayor iluminación, lo que puede ayudar a resaltar objetos o detalles en la nueva imagen (Lacombe et al., 2020). Como se muestra en la ecuación 25.

$$h_2 = \sum_{n=0} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\}, |i - j| = n \quad (25)$$

Donde:

$p(i, j)$  es la probabilidad en la que un píxel con valor  $i$  se encuentre adyacente a un píxel de valor  $j$ .

### Homogeneidad

Una imagen se considera homogénea cuando los valores de la diagonal principal de su matriz de coocurrencia son altos. Esto se debe a que, según lo expresado en la ecuación 26, los valores de probabilidad en dicha matriz tienden a ser mayores en la diagonal principal, reflejando una alta similitud entre los niveles de gris de los píxeles vecinos, mientras que su influencia disminuye exponencialmente conforme se alejan de la diagonal.

$$\sum_{i,j=0}^{N-1} \frac{p_{i,j}}{1 + (i - j)^2} \quad (26)$$

Donde:

$p(i, j)$  es la probabilidad en la que un píxel con valor  $i$  se encuentre adyacente a un píxel de valor  $j$ .

### Cluster Shade

Es una medida de asimetría de la matriz que mide los conceptos perceptivos de uniformidad, en donde primero se crea una nueva imagen  $i + j$  con un rango de



intensidades entre 0 a 2. El valor  $\mu(i + j)$  se calcula y almacena para la primera imagen y se actualiza a medida que se mueve un píxel. Cuando la sombra del cúmulo es alta, la imagen es asimétrica como se observa en la ecuación 27.

$$\sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i, j) \quad (27)$$

### Cluster prominence

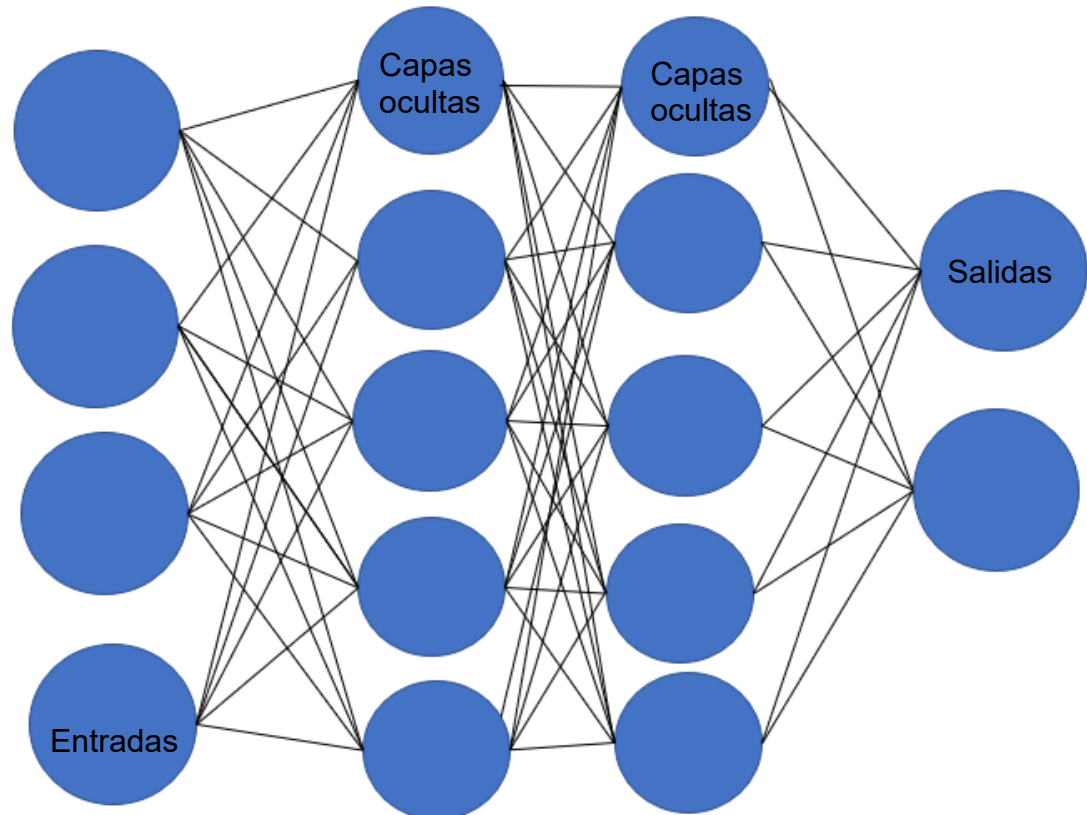
Es una medida de asimetría cuando el valor de prominencia del grupo es alto y la imagen es menos simétrica, al igual que cuando el valor de prominencia es bajo, existe un pico en la matriz de coocurrencia de nivel de gris alrededor de los valores medios (Yang et al., 2012) como se observa en la ecuación 28.

$$\sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i, j) \quad (28)$$

### Clasificadores

La clasificación digital de imágenes consiste en organizar los píxeles de una imagen dentro de un número determinado de categorías o clases, en función de los valores de sus variables asociadas. Cuando un píxel cumple con ciertos criterios o condiciones, se asigna a la clase correspondiente, lo que permite generar una nueva imagen clasificada, en la que cada píxel está representado por un nivel digital que identifica la categoría a la que pertenece.

Para realizar este proceso, es necesario emplear un método cuantitativo que permita evaluar las semejanzas entre los píxeles. Uno de los enfoques más comunes es el uso de redes neuronales, las cuales pueden reconocer patrones complejos y realizar una clasificación automatizada. En la figura 2.8 se muestra la representación esquemática de un clasificador basado en este principio.



*Figura 2.8 Clasificador*

Existen diversos tipos de clasificadores, los cuales se diferencian según el algoritmo o enfoque que emplean para realizar la clasificación. Estos pueden ser paramétricos o no paramétricos, estadísticos, basados en redes neuronales, supervisados o no supervisados, así como de tipo monoestructural, multiestructural o hiperestructural.

A continuación, se presentan y explican los principales conceptos y características de cada uno de estos tipos de clasificadores, con el fin de comprender sus ventajas, limitaciones y aplicaciones dentro del procesamiento digital de imágenes.

### **Redes Neuronales Artificiales (RNA)**

Es una herramienta de estadística de modelado en la información que no es lineal, donde esta se compone de nodos que están interconectados, los cuales pueden llegar a describir las relaciones complejas entre entradas (input) y salidas (Penm et al., 2013).

Las redes neuronales artificiales se utilizan como modelos matemáticos y computacionales inspirados en la estructura del cerebro humano, caracterizándose por ser sistemas paralelos y distribuidos de procesamiento de información. Están formadas por un conjunto de unidades simples, denominadas neuronas artificiales, las cuales se interconectan mediante enlaces que poseen valores numéricos ajustables o pesos sinápticos, responsables de modificar la intensidad y dirección de la transmisión de información entre ellas.

El proceso de entrenamiento toma la información contenida en las variables de entrada y ajusta los valores de los pesos, los cuales conectan todas las capas para hacer coincidir la entrada con cada clase, de esta forma se detecta y almacena el patrón oculto que comparte todas las entradas y su correspondiente clase, de esta manera es necesario hacer uso de una base de datos suficiente para entrenar la red neuronal (Beura et al., 2015).

Se representa cada unidad como la suma de todas las entradas mediante los pesos  $w$  donde está relacionado con la conexión entre neuronas como se muestra en la ecuación 29.

$$y_i = \sum_j w_{ij} y_j \quad (29)$$

A continuación, en la figura 2.9 se puede observar la representación de una red neuronal.

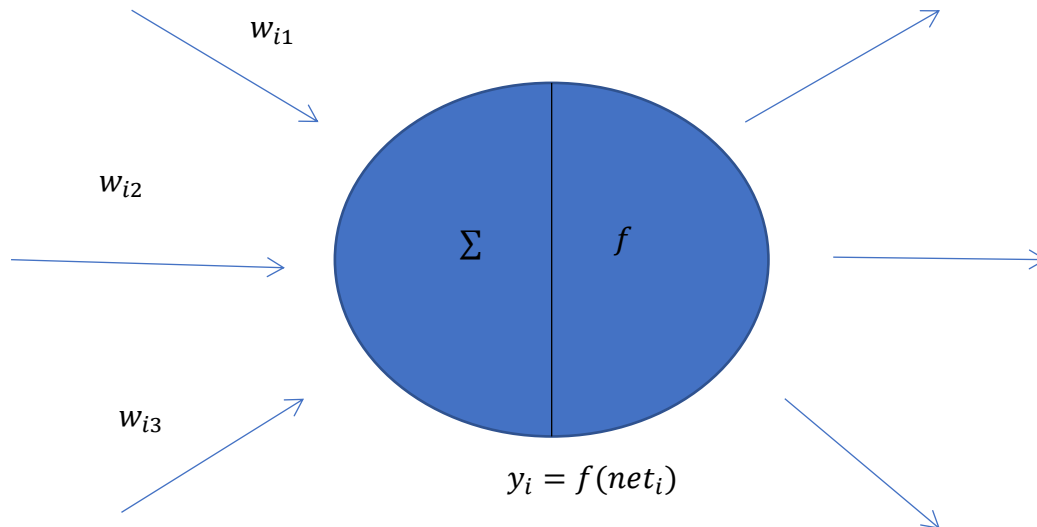


Figura 2.9 Representación de entradas (input) en red neuronal

Las redes neuronales artificiales se distinguen por varias características fundamentales. En primer lugar, presentan una autoorganización, ya que utilizan algoritmos de aprendizaje adaptativos que les permiten ajustarse y optimizar su desempeño sin intervención externa directa. Esta capacidad les confiere un procesamiento robusto, capaz de adaptarse a diferentes tipos de datos.

Además, realizan un procesamiento no lineal, lo que incrementa su habilidad para aproximar funciones complejas, reconocer patrones y resistir el ruido en los datos de entrada. También emplean un procesamiento paralelo, donde numerosos nodos trabajan simultáneamente, lo que permite una alta interconectividad y eficiencia en el manejo de información.

Estructuralmente, una red neuronal artificial está compuesta por una capa de entrada (input), una o más capas ocultas (hidden layers), generalmente entre una y tres, y una capa de salida (output). La relación matemática que describe la conexión y transmisión de información entre neuronas se presenta en la ecuación 30.

$$h_i = \sigma \left( \sum_{j=1}^N w_{ij} x_j + T_i^{hid} \right) \quad (30)$$

Donde:

$\sigma$  es la función de transferencia.

$N$  el número de entradas (input).

$w_{ij}$  el peso.

$x_j$  entradas (input) a la neurona.

$T_i^{hid}$  límite de las neuronas ocultas.

En la figura 2.10 se puede observar la representación de una red neuronal con sus diferentes componentes.

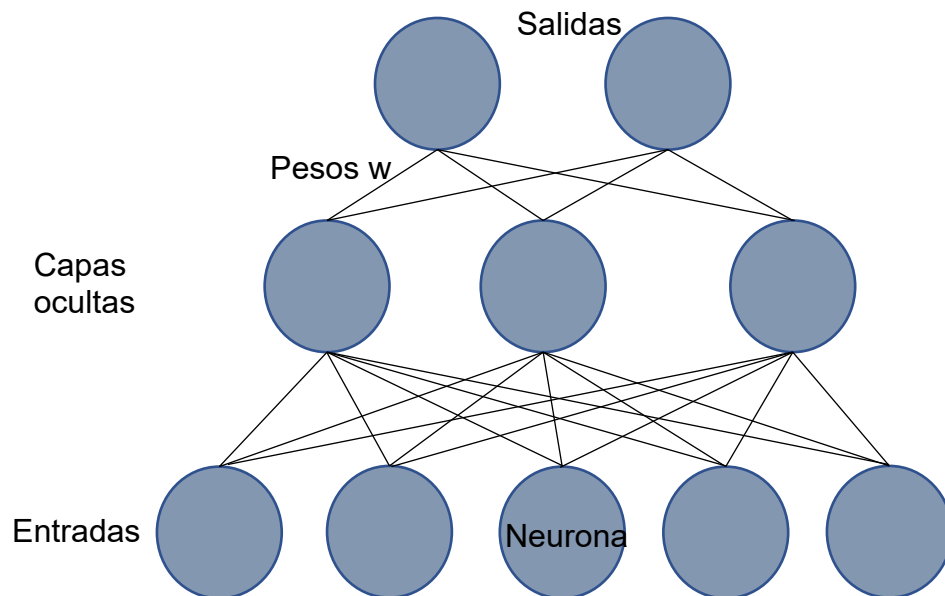


Figura 2.10 Representación de red neuronal

Cuando varias neuronas trabajan de manera paralela y coordinada, forman una capa neuronal. Por ello, una red neuronal está constituida por múltiples capas interconectadas, en las que la salida de una capa actúa como entrada para la siguiente. Esta disposición define la arquitectura o topología de la red, la cual determina su complejidad, capacidad de aprendizaje y el tipo de problemas que puede resolver.

A continuación, se describen los principales tipos de arquitecturas de redes neuronales, destacando sus características estructurales y aplicaciones más comunes.

- Red de capa única: El cual consiste en un vector de entrada que está conectado a una capa de  $N$  número de neuronas.
- Red multicapa: Este se caracteriza por tener una o más capas intermedias, las cuales se conocen como capas ocultas.



- Red recursiva: Red con retroalimentación la cual en algunas de sus salidas se conectan con las entradas.

### Red Neuronal Convolutiva (CNN)

Es un tipo de red neuronal especializada, la cual procesa datos y está dividida en cuatro etapas (Ekici & Jawzal, 2020):

1. Consiste en generar un conjunto lineal en paralelo con operaciones de convolución.
2. Se calcula un conjunto de los mapas de características pertenecientes al filtro de la capa anterior.
3. Se realiza el cálculo de la función lineal por activación lineal rectificado (RELU) o llamado también como etapa del detector.
4. Se modifica la salida para la función de agrupación de capas y el tamaño de los mapas de características.

Las redes convolucionales analizan una gran cantidad de imágenes aplicando patrones específicos, esta consta de un núcleo, una agrupación y capas conectadas, el propósito de la capa del núcleo es detectar y extraer características específicas de las imágenes, lo cual se realiza con el operador de convolución (Le et al., 2019). Las CNN se aplican para explorar patrones en una imagen, esto se hace convolucionando una imagen, una red puede detectar líneas en las capas frontales. En la figura 4 se puede observar la representación del funcionamiento de una red neuronal convolucionada.

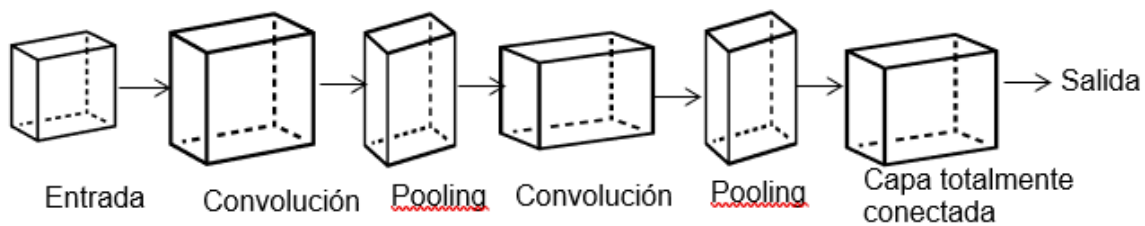


Figura 2.11 Red neuronal convolucionada

### Redes Neuronales Profundas

Es un modelo computacional que consta de muchas unidades de procesamiento simples o neuronas (Arisoy et al., 2012), las cuales funcionan en paralelo y cuentan con capas interconectadas, estas redes cuentan con una capa de entrada y una de salida, conforme las capas se apilan, estas se llaman profundas (Hinton, 2007). La principal ventaja de este modelo es que es capaz de manejar entradas inciertas que pueden eliminar la incertidumbre de los datos, haciéndolo útil para datos de la vida real (Gawlikowski et al., 2023).

### Red VGG16

Es una red neuronal convolutiva (CNN) profunda utilizada especialmente en clasificación de imágenes y extracción de características. Fue introducida por el grupo de investigación de Visual Geometry Group (VGG) de la Universidad de Oxford en su trabajo titulado "Very Deep Convolutional Networks for Large-Scale Image Recognition" en 2014.



Sus características principales son:

1. Arquitectura:

- 16 capas entrenables: De ahí data su nombre, VGG16 tiene 13 capas convolucionales y 3 capas completamente conectadas.
- Las capas convolucionales usan filtros pequeños de tamaño 3x3, lo que permite capturar características más detalladas y locales.
- Las capas de agrupamiento o pooling 2x2 se usan para reducir la dimensionalidad espacial mientras mantienen las características más importantes.

2. Estructura:

La arquitectura sigue un patrón simple y uniforme: múltiples capas convolucionales seguidas de una capa de max pooling para reducir las dimensiones. Al final, las características extraídas pasan por 3 capas completamente conectadas y una capa softmax para la clasificación. Por lo general, se utiliza como entrada una imagen de tamaño fijo, típicamente 224x224.

3. Tamaño del modelo:

La red tiene aproximadamente 138 millones de parámetros, lo que la hace relativamente grande y costosa en términos de almacenamiento y tiempo de cómputo.

4. Sus ventajas son:

- La simplicidad en el diseño hace que sea fácil de entender y adaptar.
- Ofrece un rendimiento robusto en tareas de clasificación de imágenes, como las del conjunto de datos ImageNet.
- Es muy útil para transfer learning, ya que los pesos preentrenados pueden utilizarse en otros problemas de visión por computadora.

5. Las desventajas:

- Su gran número de parámetros puede causar problemas de sobreajuste en conjuntos de datos pequeños.
- Es computacionalmente costosa en comparación con arquitecturas más modernas como ResNet o EfficientNet.

### Autoencoders

Son un tipo de red neuronal que consta de una fase de codificación, en la que los vectores de características se asignan a un espacio dimensional inferior o superior, por lo cual el vector de características original se puede reconstruir en una fase de decodificación posterior con un error de reconstrucción mínimo. El codificador automático consta de una capa de entrada de  $n$  dimensiones, una capa oculta de  $h$  dimensiones y una capa de salida de  $n$  dimensiones (Sagha et al., 2017). A continuación, en la ecuación se muestra la función de autoencoder:

$$L(W) = \sum_i D(x_i, \hat{x}_i) = h_i = g(Wx_i + b), \hat{x}_i = f(W'h_i + b') \quad (31)$$

En la figura 10 se puede observar la representación de una red autoencoder.

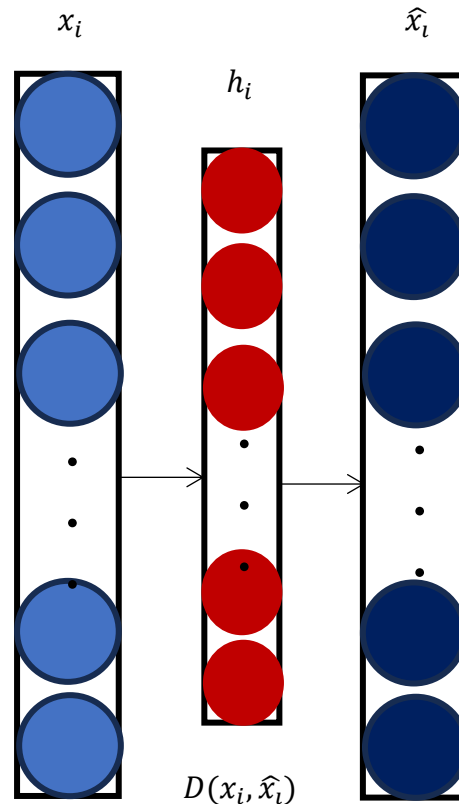


Figura 2.12 Autoencoder

## VII. HIPÓTESIS

La detección de anomalías asociadas al cáncer de mama se detectará de manera automática al utilizar imágenes termográficas infrarrojas, para obtener características de textura (homogeneidad, contraste, cluster shade y cluster prominence) como datos de entrada de un clasificador basado en redes neuronales que se implementará en una metodología de inteligencia artificial.

## VIII. OBJETIVOS

### a) Objetivo general

Diseñar un sistema de clasificación automática para detección de anomalías asociadas al cáncer de mama en imágenes termográficas utilizando técnicas de inteligencia artificial.



## b) Objetivos específicos

- Desarrollar una metodología para segmentar la región de las mamas de en imágenes termográficas infrarrojas.
- Implementar algoritmos para obtener características de textura (homogeneidad, contraste, cluster shade y cluster prominence) de las termografías que servirán como datos de entrada del clasificador.
- Programar diferentes clasificadores basados: en redes neuronales convolucionales, redes neuronales profundas y autoencoders.
- Seleccionar el mejor clasificador mediante un análisis de desempeño basado en la especificidad y sensibilidad obtenidas para cada clasificador.

## IX. METODOLOGÍA

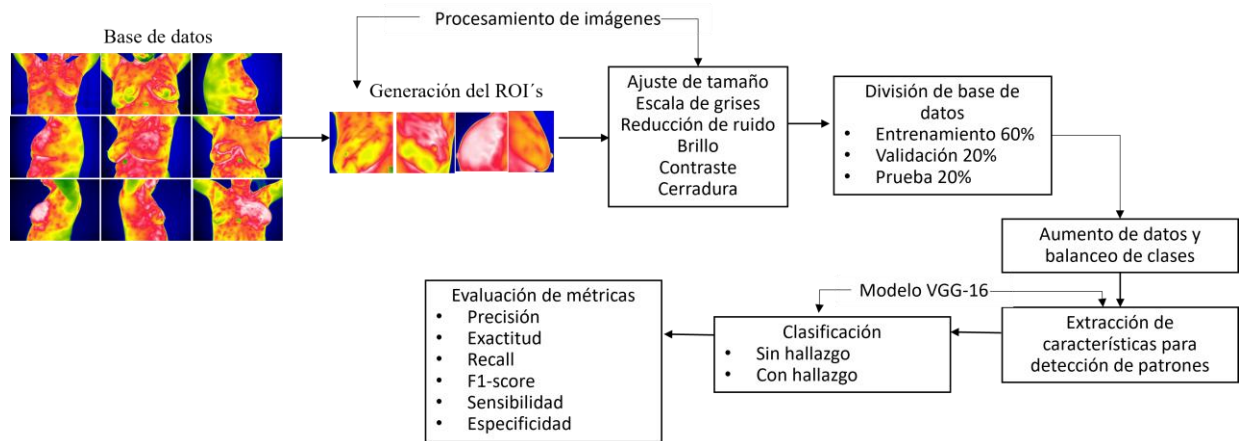


Figura 3.1 Metodología propuesta

A continuación, en la figura 3.1 se presenta la metodología aplicada en el trabajo.

### Base de datos

El conjunto de datos DMR-IR del Hospital Universitario A Antonio Pedro Silva (2014), que contiene un total de 6749 imágenes de 149 pacientes en tamaño de 640x480 píxeles, con datos de pacientes sanos y de aquellos con cáncer de mama, incluye 37 secuencias de pacientes con cáncer de mama confirmado histopatológicamente y 19 secuencias de individuos sanos sin hallazgos benignos. El conjunto de datos contiene imágenes segmentadas que se centran exclusivamente en las temperaturas de las mamas y omiten las temperaturas de otras áreas del cuerpo. Además, los registros médicos de cada paciente contienen detalles como edad, etnia, antecedentes personales, antecedentes familiares, historial médico y recomendaciones de protocolo.

### Generación del ROI

Para esta etapa se dejó únicamente el área de interés al emplear detección de curvas para delinear el contorno de la mama y eliminar todo el fondo. Esta limpieza y segmentación mejoraron la calidad de las imágenes para las etapas posteriores.

### Procesamiento de imágenes

En la primera etapa, se limpiaron las imágenes utilizando filtros de eliminación de ruido gaussiano y técnicas avanzadas de procesamiento de imágenes para segmentar la región de las mamas.

### Aumento de datos

Para incrementar la precisión del modelo y equilibrar la cantidad de imágenes en la clase de pacientes con hallazgos, se llevó a cabo un proceso de aumento de datos en el conjunto de entrenamiento. Esto fue necesario debido a la gran diferencia entre el número de pacientes sanos y aquellos con la enfermedad. Para ello, se aplicaron las siguientes técnicas: rotación, modificación de alto y ancho de la imagen, recorte, zoom, relleno y rotación horizontal.

### Extracción de características

En la segunda etapa, se extrajeron características de textura para las como variables de entrada para los algoritmos de clasificación y permitiendo una representación detallada y diferenciada de las imágenes.

### Clasificación

Para la clasificación, se exploraron y programaron diversos algoritmos avanzados con el objetivo principal de categorizar los datos según sus características intrínsecas. Se utilizaron redes neuronales convolucionales (CNN), efectivas en la clasificación de imágenes al emplear capas convolucionales que extraen las características más relevantes y capas de pooling para reducir la dimensionalidad.

Se compararon principalmente tres redes neuronales convolucionales, EfficientNet, ResNet50 y VGG16 debido a que cada una cuenta con características distintas como es la eficiencia computacional en la primera, la siguiente propuesta debido a su capacidad para evitar problemas del desvanecimiento de gradiente y la última red debido a su simplicidad en la estructura de la red y su alta efectividad en clasificación de imágenes.

### Métricas

El análisis del desempeño de los clasificadores se llevó a cabo utilizando métricas clave como la especificidad, que mide la capacidad del clasificador para identificar correctamente los casos negativos; la sensibilidad, para evaluar la identificación de casos positivos; la exactitud (accuracy), para medir el rendimiento global del clasificador; y, por último, el área bajo la curva ROC (AUC), que evalúa la capacidad para distinguir entre clases positivas y negativas. Al obtener y analizar estas métricas, se permitió seleccionar el clasificador más adecuado basándose en su desempeño cuantitativo.

Se implementó una validación cruzada para garantizar la fiabilidad y generalización de los resultados. Este enfoque proporcionó una evaluación integral de cada modelo, facilitando una comparación objetiva entre los diferentes algoritmos de clasificación empleados. Los resultados obtenidos informaron la elección del modelo óptimo para la tarea específica de clasificación en imágenes mamográficas.

Los valores métricos tomados en cuenta están representados en la tabla 3, la cual brinda su definición y su fórmula.

*Tabla 3. Evaluación de métricas*

Métrica	Definición	Formula
---------	------------	---------

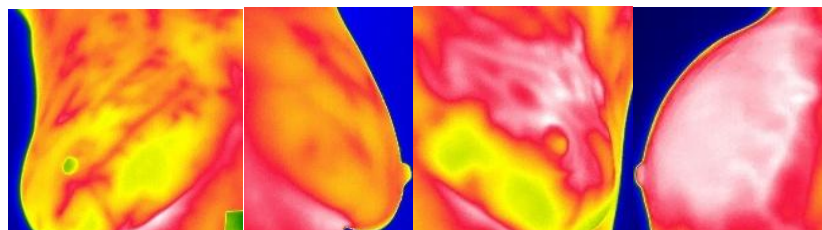
Verdadero positivo (VP)	Número de imágenes clasificadas como positiva y efectivamente es positiva	
Verdadero negativo (VN)	Número de imágenes clasificadas como negativas y efectivamente es negativa	
Falso positivo (FP)	Número de imágenes clasificadas como positivas y es negativa	
Falso negativo (FN)	Número de imágenes clasificadas como negativas y es positiva	
Exactitud	Proporción de clasificaciones correctas	$(VP + VN)/(VN + VP + FN + FP)$
Sensibilidad (Recall)	Proporción de la clase positiva que fue clasificada correctamente	$(VP)/(VP + FN)$
Especificidad	Proporción de la clase negativa que fue clasificada correctamente	$(VN)/(VN + FP)$
Precisión	Que tan bueno es es modelo prediciendo los casos positivos	$(VP)/(VP + FP)$
F1- Score	Evalua la exactitud del modelo entre ambas clases	$\frac{VP}{VP + \frac{1}{2}(FP + FN)}$

## X. RESULTADOS

Se llevo a cabo la experimentación en las fases siguientes, el procesamiento de imágenes, la división de la base de datos, aumento de datos y balanceo de clases, seguido de la extracción de patrones de detección para poder aplicar el modelo CNN de clasificación de imágenes para evaluación de sus métricas. A continuación, se explicará los detalles de cada una de las fases.

### Generación de ROI

A continuación, se presentan los resultados obtenidos de hacer el recorte del área de interés.



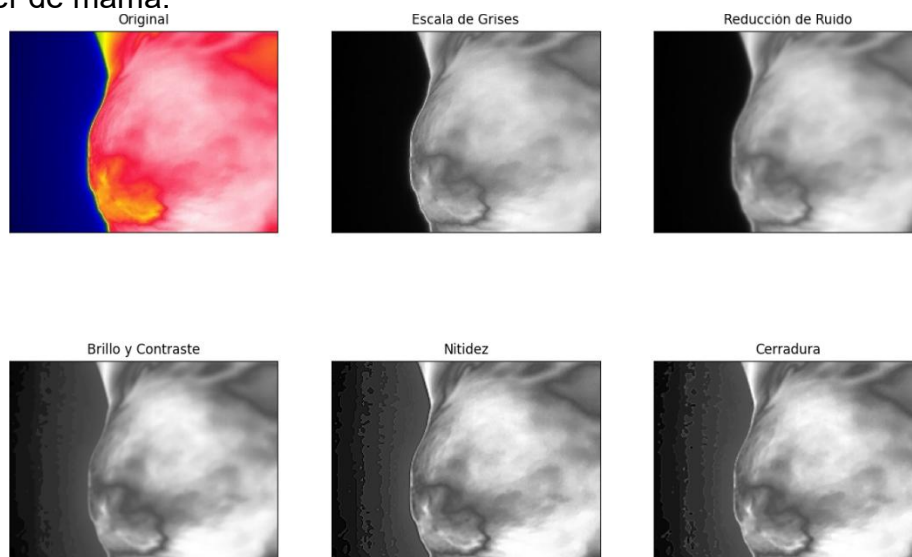
a) b) c) d)

*Figura 3.2 Generación de ROI a) vista frontal mama derecha b) mama derecha 90°  
c) Vista frontal mama izquierda c) mama izquierda 90°.*

Como se puede observar, la región de las mamas se selecciona correctamente y se separa cada una de las mamas en imágenes independientes para su estudio individual, lo que facilita la segmentación para su extracción de características y clasificación.

### Procesamiento de imágenes

A continuación, se presenta los pasos que se siguieron para detectar los cambios de temperatura en las imágenes de los pacientes que presentan un tumor de cáncer de mama.



*Figura 3.3 Procesamiento de imágenes*

Primeramente, un ajuste de tamaño de imagen de 350x230 para poder realizar la escala de grises de las imágenes, una reducción de ruido haciendo uso de un filtro Gaussiano, posteriormente un aumento de brillo y contraste, seguido de un aumento de nitidez y por último la aplicación de una operación morfológica de cerradura, con la cual la imagen se puede observar de manera más clara el cambio de temperatura en las áreas de la mama.

### División de la base de datos

En esta sección se llevó a cabo la división de la base de datos en entrenamiento un 60%, validación 20% y prueba de 20%. Quedando de la siguiente manera como se muestra en la tabla 4.

Tabla 4. División de la base de datos

	Entrenamiento	Validación	Prueba
<b>Sin hallazgo</b>	4174	1379	1373
<b>Con hallazgo</b>	945	327	334
<b>Total</b>	5119	1706	1707





Como se puede observar existe un gran desbalance entre las clases, existiendo un porcentaje mayor en imágenes sin hallazgo, lo que puede provocar en el clasificador una alta detección en imágenes de esta clase y mucho más baja en la clase de imágenes con hallazgo, para lo cual se necesita la aplicación de un aumento de clases en la sección de entrenamiento.

### Aumento de datos

El aumento de datos realizado permitió el balance de las clases y que la base de datos tuviera más imágenes para mejorar las métricas del modelo, quedando de la siguiente manera como en la tabla 5.

*Tabla 5. División de la base de datos después del balanceo de clases*

	Entrenamiento	Validación	Prueba
<b>Sin hallazgo</b>	4174	1379	1373
<b>Con hallazgo</b>	3780	327	334
<b>Total</b>	7954	1706	1707

Tras realizar este balanceo, se observó un incremento del 25% en las imágenes con hallazgo, lo cual contribuye a mejorar el desempeño del clasificador.

### Extracción de características y clasificación

Al aplicar los tres distintos modelos de CNN, se presentan los mejores resultados en la tabla 5 obtenidos en cada uno de ellos, observando un mejor desempeño en las métricas en el modelo VGG16.

*Tabla 6. Comparativa de modelos*

Reporte de modelos				
	Exactitud	Precisión	Pérdida	Recall
ResNet50	0.92	0.84	0.16	0.83
EfficientNet	0.84	0.76	0.23	0.75
VGG16	0.95	0.89	0.13	0.87

Al elegir este modelo, se presentan detalles de la configuración computacional utilizada en el modelo final. A continuación, presentados en la tabla 7.

*Tabla 7. Configuración computacional*

Parámetros	Valores
Tamaño de la imagen	350x230
Tasa de aprendizaje	$1 \times 10^{-4}$
Tamaño de lote	64
Número de épocas	40
Optimizador	Adam



A continuación, se presenta en la figura 14 la arquitectura de la red neuronal convolucional VGG16, en la cual se muestran el número de capas de convolucionales, maxPooling, de aplanamiento, capas densas y su salida, así mismo donde ocurre la extracción de características y el clasificador.

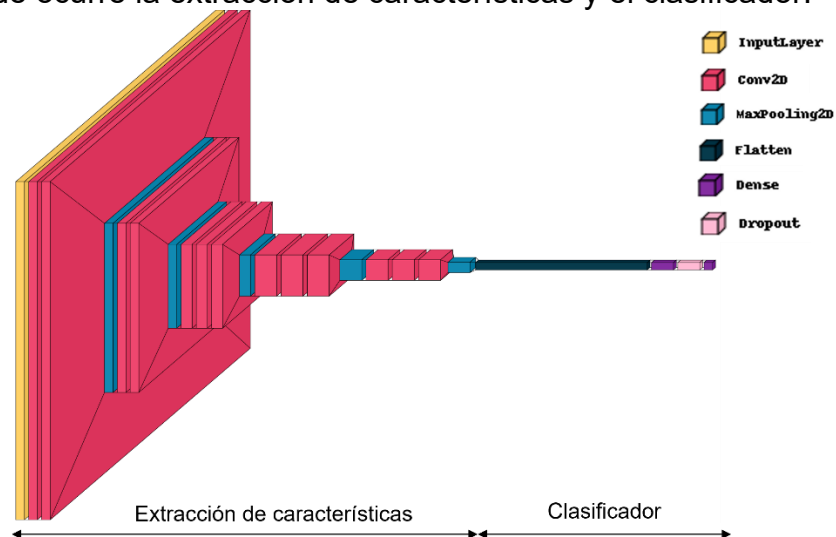
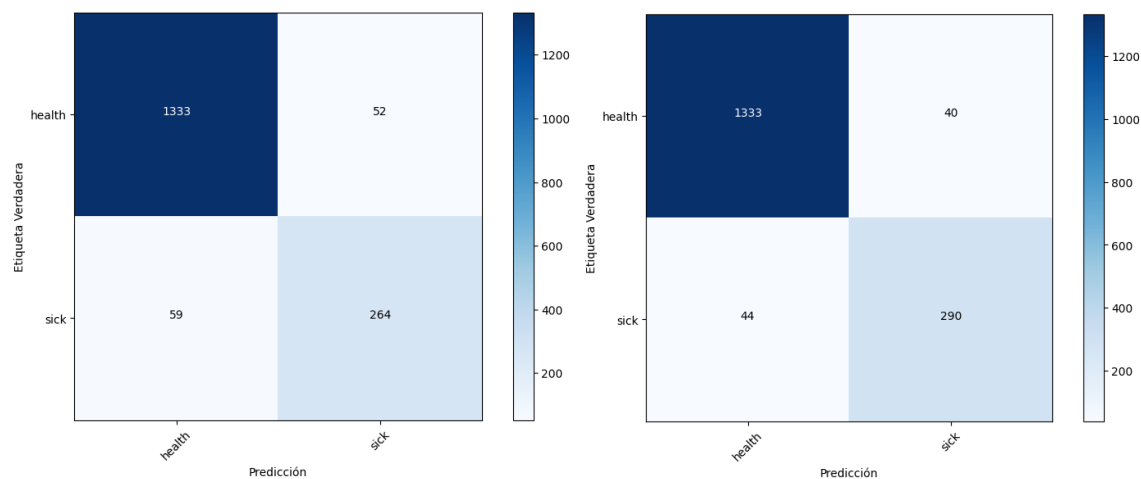


Figura 3.4 Arquitectura del modelo

La primera parte como se puede observar es donde ocurre la extracción de características de las imágenes y lo continuo donde aplicó para la clasificación, una capa de aplanamiento (Flatten), seguida de una capa densa de 512 neuronas con activación ReLU. Para evitar el sobreajuste, se incluye una capa de Dropout con una tasa del 50%. Se adaptó la red VGG16 para realizar la clasificación, donde se le eliminan las capas de clasificación originales, y se le añaden nuevas capas: una capa de aplanamiento (Flatten), seguida de una capa densa de 512 neuronas con activación ReLU. Para evitar el sobreajuste, se incluye una capa de Dropout con una tasa del 50%.

Al realizar la clasificación durante distintos números de épocas, se obtuvieron los siguientes valores de verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos, representados a continuación en la figura 15.



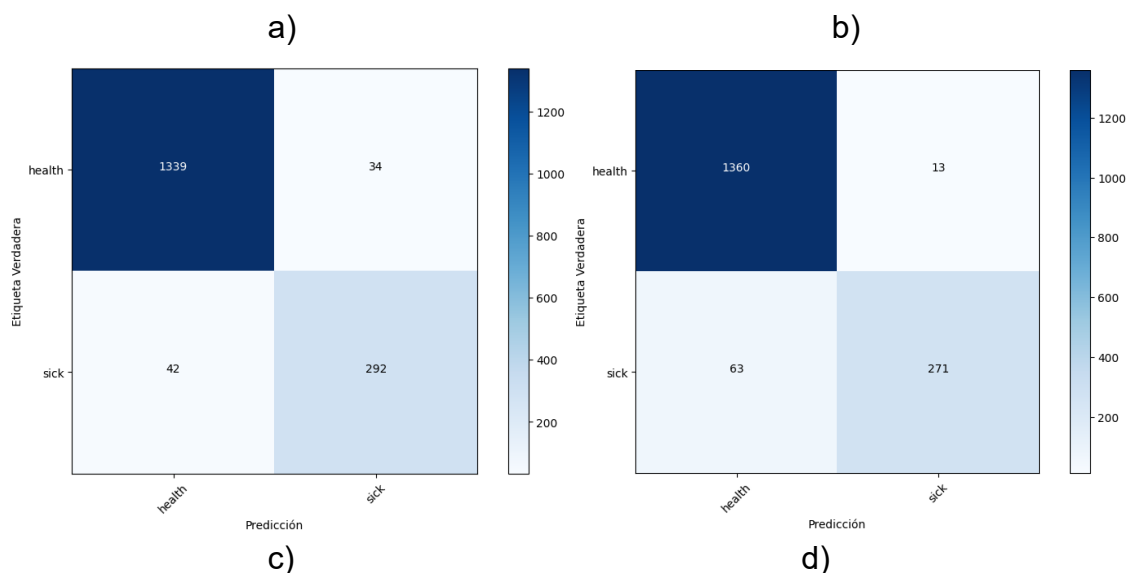


Figura 3.5 Matriz de confusión a) En siete épocas  
b) treinta épocas c) cuarenta épocas d) cincuenta épocas

Como se puede observar, durante las primeras siete épocas existe un valor alto entre los falsos positivos y falsos negativos, indicando que el modelo se equivoca mayormente en clasificar a pacientes como sin hallazgo cuando su etiqueta verdadera es con hallazgo e igualmente de forma contraria al etiquetar como pacientes con hallazgo cuando su etiqueta verdadera es sin hallazgo. Al realizar el modelo en treinta épocas, esta muestra una mejora en la reducción errónea de clasificación de falsos positivos y falsos negativos y un aumento en clasificación de verdaderos negativos. Continuando con el modelo a cuarenta épocas, esta muestra el mejor desempeño en cuanto al número de mayor clasificación como verdaderos positivos y verdaderos negativos, al igual que al menor número en falsos negativos y una reducción en falsos positivos. Finalmente, al observar el modelo en cincuenta épocas, este demuestra un aumento en verdaderos positivos mayor a cualquiera de los anteriores y una disminución significativa en falsos positivos, sin embargo es el que peor se desempeña en clasificación de falsos negativos y reduciendo significativamente el número de verdaderos negativos comparándolo con los anteriores.

Por lo cual se eligió el modelo en cuarenta épocas, debido a sus resultados de desempeño en clasificación demostrando que las tanto las imágenes de pacientes sin hallazgo como pacientes con hallazgo son bien clasificados y el número de falsos positivos y falsos negativos representan una minoría.

### Métricas

La evaluación de la precisión y pérdida de los modelos durante las siguientes etapas fueron evaluadas de igual manera, las cuales se representan en las siguientes gráficas.

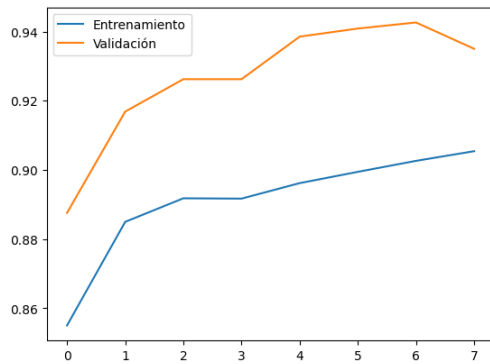


Figura 3.6 Desempeño en 7 épocas

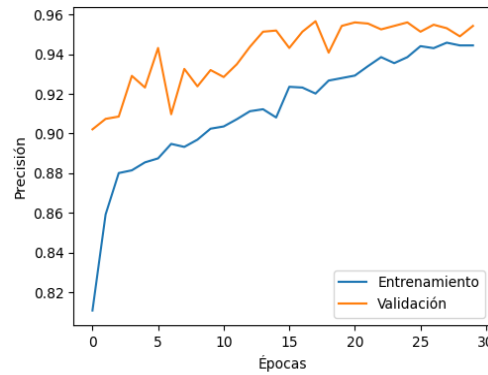


Figura 3.7 Desempeño en 30 épocas

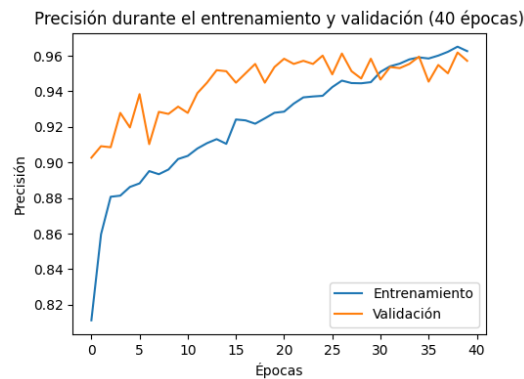


Figura 3.8 Desempeño en 40 épocas

El sistema propuesto, basado en aprendizaje profundo, fue entrenado con 7,954 imágenes de 149 pacientes, logrando una precisión del 94%, una sensibilidad del 95.76%, y una especificidad del 83.54%, demostrando su eficacia en la detección de tumores malignos en las mamas. A continuación, se presentan las métricas obtenidas.

Al obtener las mejores métricas en la época cuarenta como se observó anteriormente, se presenta el desglose de su desempeño durante cada época en cuanto a accuracy y pérdida, su reporte de clasificación y la sensibilidad y especificidad del modelo, en donde se demuestra su eficacia en la detección de hallazgos en cuanto a las imágenes evaluadas.

Tabla 8. Exactitud del modelo a 30 épocas

Reporte de métricas				
	Precisión	Recall	F1-score	Soporte
Sin hallazgo	0.97	0.97	0.97	1373
Con hallazgo	0.88	0.87	0.87	334
Exactitud	0.95			1707
Media de métricas	0.92	0.92	0.92	1707



Promedio ponderado	0.95	0.94	0.95	1707
--------------------	------	------	------	------

*Tabla 9. Reporte de modelo a 30 épocas*

Reporte de métricas de prueba	
Sensibilidad	0.95
Especificidad	0.83

*Tabla 8. Exactitud del modelo a 30 épocas*

Reporte de métricas				
	Precisión	Recall	F1-score	Soporte
Sin hallazgo	0.97	0.98	0.97	1373
Con hallazgo	0.90	0.87	0.88	334
Exactitud 0.96 1707				
Media de métricas	0.93	0.92	0.93	1707
Promedio ponderado	0.96	0.96	0.96	1707

*Tabla 9. Reporte de modelo a 30 épocas*

Reporte de métricas de prueba	
Sensibilidad	0.96
Especificidad	0.87

*Tabla 10. Evaluación del conjunto de prueba a 30 épocas*

Evaluación del conjunto de prueba
-----------------------------------



Pérdida	Exactitud	Precisión	Recall
0.14	0.95	0.87	0.86

*Tabla 11. Reporte de métricas de prueba a 30 épocas*

Reporte de métricas de prueba	
Sensibilidad	0.95
Especificidad	0.83

*Tabla 12. Evaluación del conjunto de prueba a 40 épocas*

Evaluación del conjunto de prueba			
Pérdida	Exactitud	Precisión	Recall
0.13	0.95	0.89	0.87

*Tabla 13. Reporte de métricas de prueba a 40 épocas*

Reporte de métricas de prueba	
Sensibilidad	0.96
Especificidad	0.87

Reporte de métricas de prueba	
Sensibilidad	0.95
Especificidad	0.83
AUC	0.95

Figura 15. Reporte de métricas

En la cual se puede observar una alta sensibilidad en los casos de personas con algún hallazgo de cáncer de mama de un 96% y un 87% en casos de personas sanas sin hallazgo en la enfermedad, lo que representa buenos resultados en la clasificación de imágenes.

## XI. DISCUSIONES

La implementación de una metodología automática representa un avance significativo en términos de precisión y eficiencia para el análisis de imágenes médicas. Se puede destacar cómo este enfoque facilita la detección y diagnóstico, eliminando la necesidad de segmentaciones manuales que son más propensas a errores y más lentas.

Se compararon varias arquitecturas robustas de aprendizaje profundo (ResNet50, EfficientNet y VGG16), lo cual demuestra un enfoque meticuloso para seleccionar el modelo más adecuado según el contexto.

Cada red tiene características específicas que las hacen útiles para distintas aplicaciones; por ejemplo, ResNet50 es conocida por su capacidad para prevenir el problema de gradientes desvanecientes, mientras que EfficientNet está optimizada para eficiencia computacional y VGG16 ofrece una estructura más simple pero efectiva para clasificación de imágenes.

La elección del modelo VGG16 basado en las métricas obtenidas (exactitud promedio del 95%, especificidad del 87% y sensibilidad del 96%) resalta su capacidad para manejar imágenes médicas, donde la sensibilidad y especificidad son cruciales. La sensibilidad del 96% indica que el modelo tiene un excelente desempeño al identificar casos positivos, mientras que la especificidad del 87% sugiere una buena capacidad para evitar falsos positivos, aunque se podría mejorar este aspecto.

Se pueden explorar estrategias para mejorar la especificidad del modelo, como el ajuste fino de hiperparámetros o el uso de técnicas de aumento de datos específicas para imágenes médicas. También es posible integrar este modelo en sistemas más amplios para asistir en el diagnóstico, con una validación clínica adicional.



## XII. CONCLUSIONES

Se desarrollo una metodología de segmentación automática para la región de las mamas., también se implementaron distintos clasificadores de redes neuronales convolucionales de aprendizaje profundo obteniendo buenos resultados en una red VGG-16. Los resultados obtenidos del clasificador reportan una precisión promedio del 96%.

Se desarrollo una metodología de segmentación automática para la región de las mamas para lograr la clasificación en el modelo.

Se implementaron distintos clasificadores de redes neuronales convolucionales de aprendizaje profundo como lo fue ResNet50, EfficientNet y VGG16 y al evaluar las diferentes métricas obtenidas, se seleccionó la última red implementada.

Los resultados obtenidos del modelo VGG16 reportan una exactitud promedio del 96%, especificidad del 87% y sensibilidad del 96%, por lo que se seleccionó debido a su alta clasificación en imágenes médicas y que en todos los parámetros obtuvo el mejor desempeño

## XIII. REFERENCIAS BIBLIOGRÁFICAS

- Abdel-Nasser, M., Moreno, A., & Puig, D. (2019). Breast cancer detection in thermal infrared images using representation learning and texture analysis methods. *Electronics (Switzerland)*, 8(1). <https://doi.org/10.3390/electronics8010100>
- Abdelrahman, L., Al Ghamdi, M., Collado-Mesa, F., & Abdel-Mottaleb, M. (2021). Convolutional neural networks for breast cancer detection in mammography: A survey. In *Computers in Biology and Medicine* (Vol. 131). <https://doi.org/10.1016/j.combiomed.2021.104248>
- Alanazi, S. A., Kamruzzaman, M. M., Islam Sarker, M. N., Alruwaili, M., Alhwaiti, Y., Alshammari, N., & Siddiqi, M. H. (2021). Boosting Breast Cancer Detection Using Convolutional Neural Network. *Journal of Healthcare Engineering*, 2021. <https://doi.org/10.1155/2021/5528622>
- Ali, M. A. S., Sayed, G. I., Gaber, T., Hassanien, A. E., Snasel, V., & Silva, L. F. (2015). Detection of breast abnormalities of thermograms based on a new segmentation method. *Proceedings of the 2015 Federated Conference on Computer Science and Information Systems, FedCSIS 2015*. <https://doi.org/10.15439/2015F318>
- Arisoy, E., Sainath, T. N., Kingsbury, B., & Ramabhadran, B. (2012). Deep neural network language models. WLM 2012 - Will We Ever Really Replace the N-Gram Model? On the Future of Language Modeling for HLT at the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2012.
- Ayana, G., Dese, K., & Choe, S. W. (2021). Transfer learning in breast cancer diagnoses via ultrasound imaging. In *Cancers* (Vol. 13, Issue 4). <https://doi.org/10.3390/cancers13040738>
- Bartels, P. H., Thompson, D., Bibbo, M., & Weber, J. E. (1992). Bayesian belief networks in quantitative histopathology. *Analytical and Quantitative Cytology and Histology*, 14(6).
- Beura, S., Majhi, B., & Dash, R. (2015). Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for





- detection of breast cancer. *Neurocomputing*, 154.  
<https://doi.org/10.1016/j.neucom.2014.12.032>
- Cai, X., Li, X., Razmjoo, N., & Ghadimi, N. (2021). Breast Cancer Diagnosis by Convolutional Neural Network and Advanced Thermal Exchange Optimization Algorithm. *Computational and Mathematical Methods in Medicine*, 2021.  
<https://doi.org/10.1155/2021/5595180>
- Cary, J., Kalisher, L., Sadowsky, N., & Mikic, B. (1975). Thermal evaluation of breast disease using local cooling. *Radiology*, 115(1). <https://doi.org/10.1148/115.1.73>
- De Weerd, L., Mercer, J. B., & Weum, S. (2011). Dynamic Infrared Thermography. In *Clinics in Plastic Surgery* (Vol. 38, Issue 2).  
<https://doi.org/10.1016/j.cps.2011.03.013>
- Du, R., Chen, Y., Li, T., Shi, L., Fei, Z., & Li, Y. (2022). Discrimination of Breast Cancer Based on Ultrasound Images and Convolutional Neural Network. *Journal of Oncology*, 2022. <https://doi.org/10.1155/2022/7733583>
- Ekici, S., & Jawzal, H. (2020a). Breast cancer diagnosis using thermography and convolutional neural networks. *Medical Hypotheses*, 137.  
<https://doi.org/10.1016/j.mehy.2019.109542>
- Ekici, S., & Jawzal, H. (2020b). Breast cancer diagnosis using thermography and convolutional neural networks. *Medical Hypotheses*, 137.  
<https://doi.org/10.1016/j.mehy.2019.109542>
- Etehadavakol, M., Chandran, V., Ng, E. Y. K., & Kafieh, R. (2013). Breast cancer detection from thermal images using bispectral invariant features. *International Journal of Thermal Sciences*, 69. <https://doi.org/10.1016/j.ijthermalsci.2013.03.001>
- Francis, S. V., Sasikala, M., Bhavani Bharathi, G., & Jaipurkar, S. D. (2014). Breast cancer detection in rotational thermography images using texture features. *Infrared Physics and Technology*, 67. <https://doi.org/10.1016/j.infrared.2014.08.019>
- Fu, Q., & Dong, H. (2022). Spiking Neural Network Based on Multi-Scale Saliency Fusion for Breast Cancer Detection. *Entropy*, 24(11).  
<https://doi.org/10.3390/e24111543>
- Gawlikowski, J., Tassi, C. R. N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., & Zhu, X. X. (2023). A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, 56.  
<https://doi.org/10.1007/s10462-023-10562-9>
- Gerasimova, E., Audit, B., Roux, S. G., Khalil, A., Gileva, O., Argoul, F., Naimark, O., & Arneodo, A. (2014). Wavelet-based multifractal analysis of dynamic infrared thermograms to assist in early breast cancer diagnosis. *Frontiers in Physiology*, 5 MAY. <https://doi.org/10.3389/fphys.2014.00176>
- Gonzalez-Hernandez, J. L., Recinella, A. N., Kandlikar, S. G., Dabydeen, D., Medeiros, L., & Phatak, P. (2019). Technology, application and potential of dynamic breast thermography for the detection of breast cancer. In *International Journal of Heat and Mass Transfer* (Vol. 131).  
<https://doi.org/10.1016/j.ijheatmasstransfer.2018.11.089>
- Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Yener, B. (2009). Histopathological Image Analysis: A Review. *IEEE Reviews in Biomedical Engineering*, 2, 147–171. <https://doi.org/10.1109/RBME.2009.2034865>
- Hamilton, P. W., Anderson, N., Bartels, P. H., & Thompson, D. (1994). Expert system support using Bayesian belief networks in the diagnosis of fine needle aspiration



- biopsy specimens of the breast. *Journal of Clinical Pathology*, 47(4).  
<https://doi.org/10.1136/jcp.47.4.329>
- Hinton, G. E. (2007). Learning multiple layers of representation. In *Trends in Cognitive Sciences* (Vol. 11, Issue 10). <https://doi.org/10.1016/j.tics.2007.09.004>
- INEGI. (2022). “Estadísticas a Propósito Del... Día Mundial Contra El Cáncer (4 De Febrero)” Datos Nacionales. *Comunicado de Prensa Num. 105/21, 2021*.
- INEGI. (2023). ESTADÍSTICAS A PROPÓSITO DEL... DÍA MUNDIAL DE LA LUCHA CONTRA EL CÁNCER DE MAMA. *Revista Nacional Del INEGI*.
- Jia, Y., Wu, R., Lu, X., Duan, Y., Zhu, Y., Ma, Y., & Nie, F. (2023). Deep Learning with Transformer or Convolutional Neural Network in the Assessment of Tumor-Infiltrating Lymphocytes (TILs) in Breast Cancer Based on US Images: A Dual-Center Retrospective Study. *Cancers*, 15(3).  
<https://doi.org/10.3390/cancers15030838>
- Jochelson, M. (2012). Advanced Imaging Techniques for the Detection of Breast Cancer. *American Society of Clinical Oncology Educational Book*, 32.  
[https://doi.org/10.14694/edbook\\_am.2012.32.223](https://doi.org/10.14694/edbook_am.2012.32.223)
- Kakileti, S. T., Manjunath, G., Madhu, H., & Ramprakash, H. V. (2017). Advances in Breast Thermography. In *New Perspectives in Breast Imaging*.  
<https://doi.org/10.5772/intechopen.69198>
- Krishna, S., & George, B. (2021). An affordable solution for the recognition of abnormality in breast thermogram. *Multimedia Tools and Applications*, 80(18).  
<https://doi.org/10.1007/s11042-021-11082-w>
- Le, E. P. V., Wang, Y., Huang, Y., Hickman, S., & Gilbert, F. J. (2019). Artificial intelligence in breast imaging. In *Clinical Radiology* (Vol. 74, Issue 5).  
<https://doi.org/10.1016/j.crad.2019.02.006>
- Mammoottil, M. J., Kulangara, L. J., Cherian, A. S., Mohandas, P., Hasikin, K., & Mahmud, M. (2022). Detection of Breast Cancer from Five-View Thermal Images Using Convolutional Neural Networks. *Journal of Healthcare Engineering*, 2022.  
<https://doi.org/10.1155/2022/4295221>
- Penm, J., Chaar, B., Moles, R., & Penm, J. (2013). Predicting ASX Health Care Stock Index Movements After the Recent Financial Crisis Using Patterned Neural Networks. In *Rethinking Valuation and Pricing Models*.  
<https://doi.org/10.1016/B978-0-12-415875-7.00037-3>
- Resmini, R., Faria da Silva, L., Medeiros, P. R. T., Araujo, A. S., Muchaluat-Saade, D. C., & Conci, A. (2021). A hybrid methodology for breast screening and cancer diagnosis using thermography. *Computers in Biology and Medicine*, 135.  
<https://doi.org/10.1016/j.combiomed.2021.104553>
- Sagha, H., Cummins, N., & Schuller, B. (2017). Stacked denoising autoencoders for sentiment analysis: a review. In *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (Vol. 7, Issue 5). <https://doi.org/10.1002/widm.1212>
- Sánchez-Cauce, R., Pérez-Martín, J., & Luque, M. (2021). Multi-input convolutional neural network for breast cancer detection using thermal images and clinical data. *Computer Methods and Programs in Biomedicine*, 204.  
<https://doi.org/10.1016/j.cmpb.2021.106045>
- Sathish, D., Kamath, S., Prasad, K., Kadavigere, R., & Martis, R. J. (2017). Asymmetry analysis of breast thermograms using automated segmentation and texture



- features. *Signal, Image and Video Processing*, 11(4).  
<https://doi.org/10.1007/s11760-016-1018-y>
- Shada, A. L., Dengel, L. T., Petroni, G. R., Smolkin, M. E., Acton, S., & Slingluff, C. L. (2013). Infrared thermography of cutaneous melanoma metastases. *Journal of Surgical Research*, 182(1). <https://doi.org/10.1016/j.jss.2012.09.022>
- Silva, L. F., Santos, A. A. S. M. D., Bravo, R. S., Silva, A. C., Muchaluat-Saade, D. C., & Conci, A. (2016). Hybrid analysis for indicating patients with breast cancer using temperature time series. *Computer Methods and Programs in Biomedicine*, 130. <https://doi.org/10.1016/j.cmpb.2016.03.002>
- Singh, P., Gupta, S., & Gupta, V. (2023). Multi-objective hyperparameter optimization on gradient-boosting for breast cancer detection. *International Journal of System Assurance Engineering and Management*. <https://doi.org/10.1007/s13198-023-01955-8>
- WHO. (2023). WHO launches new roadmap on breast cancer. In *The Global Breast Cancer Initiative (GBCI)*.
- Yadav, S. S., & Jadhav, S. M. (2022). Thermal infrared imaging based breast cancer diagnosis using machine learning techniques. *Multimedia Tools and Applications*, 81(10). <https://doi.org/10.1007/s11042-020-09600-3>
- Yang, X., Tridandapani, S., Beitler, J. J., Yu, D. S., Yoshida, E. J., Curran, W. J., & Liu, T. (2012). Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: An in vivo study of late toxicity. *Medical Physics*, 39(9). <https://doi.org/10.1118/1.4747526>
- Zhang, G., Zhao, K., Hong, Y., Qiu, X., Zhang, K., & Wei, B. (2021). SHA-MTL: soft and hard attention multi-task learning for automated breast cancer ultrasound image segmentation and classification. *International Journal of Computer Assisted Radiology and Surgery*, 16(10). <https://doi.org/10.1007/s11548-021-02445-7>