



**Sistema para clasificación y reconocimiento de nódulos en imágenes mamográficas
utilizando procesamiento digital de imágenes y redes neuronales.**

Bermeo Calle, Ivonne Alexandra y Chicaiza Sangurima, Jairo Alejandro

Departamento De Eléctrica y Electrónica

Carrera de Ingeniería en Electrónica e Instrumentación

Trabajo de titulación, previo a la obtención del título de Ingeniero en Electrónica e
Instrumentación

Ing. Galarza Zambrano, Eddie Egberto

31 de agosto del 2020



**DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA
CARRERA DE INGENIERÍA EN ELECTRÓNICA E INSTRUMENTACIÓN**

CERTIFICACIÓN

Certifico que el trabajo de titulación, ***“Sistema para clasificación y reconocimiento de nódulos en imágenes mamográficas utilizando procesamiento digital de imágenes y redes neuronales”*** fue realizado por los señores **Bermeo Calle Ivonne Alexandra y Chicaiza Sangurima Jairo Alejandro**, el mismo que ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos técnicos, metodológicos y legales establecidos por la Universidad de las Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Latacunga, 31 de agosto del 2020

Ing. Galarza Zambrano, Eddie Egberto

C. C: 1303128514






REPORTE DE VERIFICACIÓN



Document Information

| | |
|--------------------------|---|
| Analyzed document | Proyecto de Investigación Final Agosto Bermeo_Chicaiza.docx (D78422580) |
| Submitted | 8/31/2020 10:10:00 PM |
| Submitted by | Eddie Galarza |
| Submitter email | eegalarza@espe.edu.ec |
| Similarity | 2% |
| Analysis address | eegalarza.espe@analysis.arkund.com |

Sources included in the report

| | | | |
|-----------|--|---|---|
| SA | tesis_final_JC_AE.pdf Document tesis_final_JC_AE.pdf (D63490085) |  | 2 |
| SA | Universidad de las Fuerzas Armadas ESPE / Tesis Carrasco.docx Document Tesis Carrasco.docx (D26102364) Submitted by: andres_867@hotmail.com Receiver: evcarrera.espe@analysis.arkund.com |  | 4 |
| W | URL: https://repositorio.upeu.edu.pe/bitstream/handle/UPEU/1961/Nilda_Tesis_Licenciatur ... Fetched: 6/13/2020 3:54:06 PM |  | 2 |
| W | URL: https://core.ac.uk/download/pdf/11056791.pdf Fetched: 6/12/2020 8:16:27 PM |  | 6 |
| SA | QUESQUEN_LIZA_RAUL.docx Document QUESQUEN_LIZA_RAUL.docx (D54096250) |  | 2 |


Ing. Eddie Galarza Zambrano
Director del Proyecto



DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA
CARRERA DE INGENIERÍA EN ELECTRÓNICA E INSTRUMENTACIÓN

AUTORÍA DE RESPONSABILIDAD

Nosotros, **Bermeo Calle, Ivonne Alexandra** con cédula de identidad N°1714219787 y **Chicaiza Sangurima, Jairo Alejandro** con cédula de identidad N°1726641572, declaramos que el contenido, ideas y criterios del trabajo de titulación: **Sistema para clasificación y reconocimiento de nódulos en imágenes mamográficas utilizando procesamiento digital de imágenes y redes neuronales** es de nuestra autoría y responsabilidad, cumpliendo con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Consecuentemente declaro que este trabajo veraz.

Latacunga, 31 de agosto del 2020

Bermeo Calle, Ivonne Alexandra

C.C.: 1714219787

Chicaiza Sangurima, Jairo Alejandro

C.C.: 1726641572



**DEPARTAMENTO DE ELÉCTRICA Y ELECTRÓNICA
CARRERA DE INGENIERÍA DE ELECTRÓNICA E INSTRUMENTACIÓN**

AUTORIZACIÓN

Nosotros, **Bermeo Calle, Ivonne Alexandra** con cédula de identidad N°1714219787 y **Chicaiza Sangurima, Jairo Alejandro** con cédula de identidad N°1726641572, autorizamos a la Universidad de las Fuerzas Armadas ESPE publicar el presente trabajo de titulación: **Sistema para clasificación y reconocimiento de nódulos en imágenes mamográficas utilizando procesamiento digital de imágenes y redes neuronales** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de nuestra responsabilidad.

Latacunga, 31 de agosto del 2020

Bermeo Calle, Ivonne Alexandra
C.C.: 1714219787

Chicaiza Sangurima, Jairo Alejandro
C.C.: 1726641572

DEDICATORIA

Quiero dedicar este trabajo a mi Dios, que me ha dado la sabiduría y la salud para culminar con este trabajo.

A mi padre y a mi madre, que han sido un apoyo incondicional en toda mi carrera, gracias por toda su paciencia, sus consejos, por inculcarme a cumplir con mis metas y darme las fuerzas para no rendirme nunca.

A mis hermanos, que siempre han estado conmigo brindándome su apoyo y ayuda para resolver problemas.

A mi novio, por todo el cariño sincero entregado, por acompañarme en esos días de alegrías y tristezas, por motivarme a seguir adelante.

Ivonne Alexandra Bermeo Calle

DEDICATORIA

El presente trabajo lleno de esfuerzo, entrega y cariño se lo dedico primero a Dios y a la Virgen María Auxiliadora por su infinito amor y bendiciones concedidas.

A mi padre Marco Chicaiza por su apoyo en estos años, a mi abuela Rosa Jaime por compartir su cariño y oraciones. Sobre todo, a mi madre Clementina Sangurima, mi inspiración, motivación y el pilar fundamental de mi vida, por hacer de mi la persona que soy, por todo su sacrificio, trabajo, amor, cuidados y consejos.

A mi familia, novia, amigos en quienes encontré siempre una mano solidaria y han sido partícipes de este logro alcanzado.

Jairo Alejandro Chicaiza Sangurima

AGRADECIMIENTO

Agradezco primero a Dios por permitirme cumplir con esta meta profesional, por ayudarme a crecer como persona física y moralmente.

Agradezco infinitamente a mis padres Hernán y Norma, que han sido un ejemplo de perseverancia, de trabajo y de lucha, gracias por creer y confiar en mi en todo momento y depositar todo su amor y apoyo en el transcurso de mi vida universitaria.

Gracias a mis hermanos, por ser personas de bien y responsables, por ayudarme en cada paso que doy.

Gracias a mi novio, por todo el cariño sincero que me brinda, por ser mi apoyo y por acompañarme en esta etapa de mi vida.

Agradezco a mi director de tesis Ing. Eddy Galarza por confiar en nosotros, por guiarnos y brindarnos sus conocimientos para cumplir con esta meta.

Un Agradecimiento especial a la Dra. Norma Orbea y Dr. Edgar Espinel por brindarnos su ayuda, para la realización de este proyecto.

Agradecer también a mi compañero de Tesis, por su paciencia para explicarme y ayudarme en todo lo que ha podido, y por ser una gran persona.

Agradezco a todas las personas, que de una u otra forma, han estado presentes en mi vida, y han deseado el bien para mi familia y para mí, gracias por sus palabras y ocurrencias.

Ivonne Alexandra Bermeo Calle

AGRADECIMIENTO

Gracias a Dios y a María Auxiliadora por ser mis guías y protectores en el transcurso de mi vida, permitiéndome culminar esta etapa tan importante.

Agradezco a mi padre Marco, principalmente a mi madre Clementina por su amor infinito y apoyo incondicional en todo momento. Es mi mayor orgullo, mi ejemplo a seguir, gracias por la oportunidad, esfuerzo y trabajo dedicado. Me faltará vida para agradecer todo lo que has hecho por mí. Para mi abuela Rosa por estar pendiente, con su amor y oraciones me motivó a seguir adelante.

Gracias a la Universidad de las Fuerzas Armadas ESPE, por acogerme y formarme como profesional. Al ingeniero Eddie Galarza, mentor de este proyecto, su confianza, guía y conocimiento fueron esenciales en el transcurso de este trabajo.

Agradezco a mi novia Grace por entregar en todo este trayecto su apoyo, amor, tiempo y palabras de aliento las cuales me han ayudado a superar varios obstáculos.

De manera especial agradezco al Centro Médico Láser, a la Dra. Norma Orbea y Dr. Edgar Espinel por abrirnos las puertas y brindarnos asesoría necesaria para realizar este trabajo.

A mi compañera de tesis agradezco por compartir su amistad, conocimiento, esfuerzo y dedicación en este proyecto.

De igual manera agradezco a todas las personas que de forma directa e indirecta estuvieron apoyándome en todo momento.

Jairo Alejandro Chicaiza Sangurima

ÍNDICE DE CONTENIDO

| | |
|--|-----------|
| CARÁTULA | 1 |
| CERTIFICACIÓN | 2 |
| REPORTE DE VERIFICACIÓN..... | 3 |
| AUTORÍA DE RESPONSABILIDAD | 4 |
| AUTORIZACIÓN | 5 |
| DEDICATORIA | 6 |
| DEDICATORIA | 7 |
| AGRADECIMIENTO | 8 |
| AGRADECIMIENTO | 9 |
| ÍNDICE DE CONTENIDO | 10 |
| ÍNDICE DE TABLAS..... | 14 |
| ÍNDICE DE FIGURAS | 15 |
| RESUMEN | 22 |
| ABSTRACT..... | 23 |
| 1 Introducción | 24 |
| 1.1 Planteamiento del problema..... | 24 |
| 1.2 Antecedentes | 24 |
| 1.3 Justificación e Importancia | 25 |
| 1.4 Objetivos | 26 |
| 1.4.1 Objetivo General | 26 |

| | | | |
|-----|-------|---|----|
| | 1.4.2 | Objetivos Específicos..... | 26 |
| 1.5 | | Variables de la investigación | 27 |
| | 1.5.1 | Variable Independiente | 27 |
| | 1.5.2 | Variable Dependiente..... | 27 |
| 1.6 | | Hipótesis | 27 |
| 2 | | Marco Teórico..... | 28 |
| | 2.1 | Anatomía de la Mama | 28 |
| | 2.1.1 | Cuadrantes de las Mamas | 29 |
| | 2.2 | Mamografía..... | 31 |
| | 2.2.1 | Mamografía digital | 32 |
| | 2.3 | Lesiones Mamográficas..... | 32 |
| | 2.3.1 | Asimetría global y focal | 32 |
| | 2.3.2 | Calcificaciones | 33 |
| | 2.3.3 | Masas | 34 |
| | 2.4 | Clasificación BI-RADS para Nódulos..... | 35 |
| | 2.4.1 | Forma | 35 |
| | 2.4.2 | Margen..... | 36 |
| | 2.4.3 | Densidad..... | 37 |
| | 2.4.4 | Composición del Tejido..... | 38 |
| | 2.5 | Categorías de Evaluación | 38 |
| | 2.6 | Procesamiento Digital de Imágenes. | 39 |

| | | |
|-----|-------|--|
| | | 12 |
| | 2.6.1 | Adquisición39 |
| | 2.6.2 | Realce y mejora40 |
| | 2.6.3 | Segmentación42 |
| | 2.6.4 | Basado en Regiones.....45 |
| | 2.6.5 | Algoritmos de Clustering46 |
| | 2.6.6 | Extracción de características47 |
| | 2.6.7 | Operaciones Morfológicas.....48 |
| 2.7 | | Redes Neuronales.....48 |
| | 2.7.1 | Neurona Biológica.....48 |
| | 2.7.2 | Neurona Artificial.....50 |
| | 2.7.3 | Red Neuronal Artificial52 |
| | 2.7.4 | Redes Neuronales Convolucionales54 |
| 2.8 | | Python.....55 |
| | 2.8.1 | Anaconda Navigator.....56 |
| | 2.8.2 | TensorFlow56 |
| | 2.8.3 | Keras57 |
| 3 | | Diseño e Implementación.....58 |
| | 3.1 | Especificaciones del software y equipo58 |
| | 3.2 | Anaconda Navigator.....58 |
| | 3.3 | Jupyter Notebook60 |
| | 3.4 | Instalación de librerías de Python.....63 |

| | | |
|------|--|------------|
| | | 13 |
| 3.5 | Estructura del sistema | 66 |
| 3.6 | Diagrama de bloques | 67 |
| 3.7 | Mamografías | 68 |
| 3.8 | Procesamiento Digital de Mamografías | 69 |
| 3.9 | Extracción de Características | 92 |
| 3.10 | Red Neuronal Convolutiva (CNN) | 95 |
| 3.11 | Interfaz gráfica..... | 103 |
| 4 | Análisis y Resultados..... | 108 |
| 4.1 | Análisis etapa de procesamiento digital de mamografías | 108 |
| 4.2 | Análisis de la red neuronal convolutiva..... | 120 |
| 4.3 | Análisis del sistema para clasificación y reconocimiento de nódulos mamarios..... | 126 |
| 4.4 | Comprobación de la hipótesis | 134 |
| 4.5 | Alcances..... | 134 |
| 4.6 | Limitaciones | 135 |
| 5 | Conclusiones y Recomendaciones | 136 |
| 5.1 | Conclusiones..... | 136 |
| 5.2 | Recomendaciones..... | 137 |
| | REFERENCIAS BIBLIOGRÁFICAS | 139 |
| | ANEXOS..... | 144 |

ÍNDICE DE TABLAS

| | |
|--|-----|
| Tabla 1 <i>Características de los Nódulos Mamarios</i> | 93 |
| Tabla 2 <i>Esquema General Matriz de Confusión</i> | 121 |
| Tabla 3 <i>Resultados del Entrenamiento de Clasificación BIRADS</i> | 124 |
| Tabla 4 <i>Resultados de Validación de la Clasificación BIRADS</i> | 126 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 1 <i>Glándulas Mamarias en el Interior de las Mamas</i> | 28 |
| Figura 2 <i>Orientación de la Mama Según Planos</i> | 29 |
| Figura 3 <i>Cuadrantes Mamarios</i> | 30 |
| Figura 4 <i>Orientación Clínico-Topográfica</i> | 30 |
| Figura 5 <i>Proyección Media-Lateral-Oblicua (MLO)</i> | 31 |
| Figura 6 <i>Detección de Asimetría Bilateral</i> | 33 |
| Figura 7 <i>Calcificaciones: a) Macrocalcificaciones, b) Microcalcificaciones</i> | 34 |
| Figura 8 <i>Masa o Nódulo con Forma Ovalada y Margen Circunscrito</i> | 35 |
| Figura 9 <i>Formas de Una Masa</i> | 36 |
| Figura 10 <i>Márgenes de Una Masa</i> | 36 |
| Figura 11 <i>Formas y Márgenes: (a) Margen Circunscrito, (b) Margen Indefinido, (c)</i> <i>Margen Microlobulado y (d) Margen Espiculado</i> | 37 |
| Figura 12 <i>Categorías BI-RADS para Masas</i> | 38 |
| Figura 13 <i>Mamografía en Escala de Grises</i> | 40 |
| Figura 14 <i>Funcionamiento Filtrado de Mediana</i> | 41 |
| Figura 15 <i>Histograma de Densidad de Una Mamografía</i> | 43 |
| Figura 16 <i>Binarización de Una Mamografía</i> | 44 |
| Figura 17 <i>Eliminación del Músculo Pectoral por el Método Basado en Regiones</i> | 45 |
| Figura 18 <i>Resultado de la imagen aplicando el Filtro Iris</i> | 47 |
| Figura 19 <i>Estructura básica de una Neurona biológica</i> | 49 |
| Figura 20 <i>Modelo de Neurona Artificial</i> | 50 |
| Figura 21 <i>Tipos de Neuronas Artificiales</i> | 52 |
| Figura 22 <i>Modelo de Una Red Neuronal Convolutiva Típica</i> | 55 |
| Figura 23 <i>Logotipo del Software Python</i> | 55 |
| Figura 24 <i>Aplicación Anaconda Navigator</i> | 59 |

| | |
|--|----|
| Figura 25 <i>Aplicación Anaconda – Enviroments</i> | 59 |
| Figura 26 <i>Ventana para Crear un Nuevo Entorno</i> | 60 |
| Figura 27 <i>Entorno Creado (Tesis)</i> | 60 |
| Figura 28 <i>Selección del Entorno</i> | 61 |
| Figura 29 <i>Instalación de Jupyter Notebook</i> | 61 |
| Figura 30 <i>Ejecución de Jupyter Notebook</i> | 62 |
| Figura 31 <i>Elección del Navegador Web</i> | 62 |
| Figura 32 <i>Ventana de Jupyter Notebook</i> | 63 |
| Figura 33 <i>Área de Edición y Ejecución de Código de Jupyter Notebook</i> | 63 |
| Figura 34 <i>Ventana Anaconda Prompt</i> | 64 |
| Figura 35 <i>Activación del Entorno para Instalación de Librerías</i> | 64 |
| Figura 36 <i>Instalación de la Librería Numpy</i> | 64 |
| Figura 37 <i>Librería Numpy Instalada</i> | 65 |
| Figura 38 <i>Instalación de la Librería Pydicom</i> | 66 |
| Figura 39 <i>Estructura del Sistema</i> | 67 |
| Figura 40 <i>Diagrama de Bloques del Sistema</i> | 68 |
| Figura 41 <i>Diagrama de Flujo del Procesamiento Digital de Mamografías</i> | 70 |
| Figura 42 <i>Imágenes Mamográficas (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC</i> | 73 |
| Figura 43 <i>Histograma de la Mamografía: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC</i> | 74 |
| Figura 44 <i>Histograma, Media Aritmética de la Imagen y Promedio entre Píxel Mayor y Menor: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC</i> | |

| | | |
|------------------|--|-----------|
| | <i>(c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | <i>75</i> |
| Figura 45 | <i>Presentación Estandarizada de las Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | <i>76</i> |
| Figura 46 | <i>Histograma de Imágenes Mamográficas Re-dimensionadas: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | <i>78</i> |
| Figura 47 | <i>Ruido Presente en las Imágenes: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | <i>79</i> |
| Figura 48 | <i>Imágenes Mamográficas Después de Aplicar el Filtro de Mediana: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | <i>80</i> |
| Figura 49 | <i>Histograma de Imágenes Mamográficas Después de Aplicar el Filtro de Mediana: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | <i>81</i> |
| Figura 50 | <i>Histograma de Imágenes Mamográficas con su Respectivo Valor de Moda: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | <i>82</i> |
| Figura 51 | <i>Mamografías Binarizadas: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | <i>83</i> |
| Figura 52 | <i>Mamografías Sin Etiquetas: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista</i> | <i>84</i> |

| | |
|--|----|
| Figura 53 <i>Región de Interés de la Mamografía: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | 85 |
| Figura 54 <i>Resumen de los Metadatos de la Mamografía DICOM.....</i> | 86 |
| Figura 55 <i>Estiramiento de Contraste Aplicado a Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | 87 |
| Figura 56 <i>K-means con n=3 Aplicado a Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista C.....</i> | 88 |
| Figura 57 <i>Extracción del Conjunto de Píxeles con Valor más Alto en Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | 89 |
| Figura 58 <i>Exclusión del Músculo de las Mamografías de Vista MLO: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | 91 |
| Figura 59 <i>Nódulos Extraídos en Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.....</i> | 92 |
| Figura 60 <i>Nódulo tipo Birads 3 de la Base de datos CBIS-DDSM, Formato: (a) DICOM (b) JPG</i> | 94 |
| Figura 61 <i>Histograma de Nódulo Tipo Birads 3 de la Base de Datos CBIS-DDSM, Formato: (a) DICOM (b) JPG.....</i> | 94 |
| Figura 62 <i>Nódulo tipo Birads 3 de la Base de Datos CBIS-DDSM: (a) Formato JPG (b) Nódulo con Proceso de Extracción de características.....</i> | 95 |
| Figura 63 <i>Arquitectura de la Red Neuronal Convolutiva Implementada</i> | 97 |

| | |
|---|-----|
| Figura 64 <i>Estructura y Parámetros de la Red Neuronal Convolutacional.....</i> | 100 |
| Figura 65 <i>BIRADS 2-3 (a) Imagen Formato jpg, (b) Extracción de Características, BIRADS 4 (c) Imagen Formato jpg, (d) Extracción de Características, BIRADS 5 (e) Imagen Formato jpg, (f) Extracción de Características.....</i> | 101 |
| Figura 66 <i>Interfaz gráfica: (a) Pestaña de Inicio (b) Pestaña de Aplicación</i> | 106 |
| Figura 67 <i>Distribución de los Cuadrantes en la Mama Izquierda y Derecha</i> | 109 |
| Figura 68 <i>Mamografía con Proyección: (a) Craneocaudal en la Cual se Distinguen los Cuadrantes Externos e Internos de la Mama, (b) Medio Lateral Oblicua en la Cual se Distinguen los Cuadrantes Superiores e Inferiores de la Mama.....</i> | 109 |
| Figura 69 <i>Mamografía del Paciente 1 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara del Nódulo Extraído Ubicado en el Cuadrante Superior.....</i> | 110 |
| Figura 70 <i>Mamografía del Paciente 1 Con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara del Nódulo Extraído Ubicado en el Cuadrante Externo.....</i> | 110 |
| Figura 71 <i>Mamografía del Paciente 2 con proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Imagen Sin Nódulos (BIRADS 1).....</i> | 111 |
| Figura 72 <i>Mamografía del Paciente 2 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Tejido Extraído.....</i> | 112 |
| Figura 73 <i>Mamografía del Paciente 3 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulos Extraídos Ubicados en el Cuadrante Superior e Inferior de la Mama.....</i> | 113 |

- Figura 74** *Mamografía del Paciente 3 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído ubicado en el Cuadrante Superior de la Mama.....113*
- Figura 75** *Mamografía del Paciente 4 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído Ubicado en el Cuadrante Superior y Máscara de Tejido Ubicada en el Cuadrante Inferior de la Mama.....114*
- Figura 76** *Mamografía del Paciente 4 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído Ubicada en el Cuadrante Exterior e Inferior de la Mama.....115*
- Figura 77** *Mamografía del Paciente 5 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído Ubicado en el Cuadrante Superior de la Mama.....116*
- Figura 78** *Mamografía del Paciente 5 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulos Extraídos Ubicados en el Cuadrante Externo de la Mama.....116*
- Figura 79** *Mamografía del Paciente 6 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído Localizado en el Cuadrante Superior y Máscara de Lesión Ubicada en el Músculo de la Mama117*
- Figura 80** *Mamografía del Paciente 6 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído.....118*
- Figura 81** *Mamografía del Paciente 7 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Mama Densa Categorizada Como BIRADS 0.....119*

| | |
|--|-----|
| Figura 82 <i>Mamografía del Paciente 7 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído ubicado en el cuadrante interno de la mama.....</i> | 119 |
| Figura 83 <i>Evolución de Accuracy (Precisión) Durante las 250 Épocas de Entrenamiento.....</i> | 122 |
| Figura 84 <i>Evolución de Loss (Pérdida) Durante las 250 Épocas de Entrenamiento....</i> | 123 |
| Figura 85 <i>Matriz de Confusión con Imágenes Utilizadas en el Entrenamiento.....</i> | 123 |
| Figura 86 <i>Matriz de Confusión con Imágenes Nuevas.....</i> | 125 |
| Figura 87 <i>Diagnóstico Proporcionado por el Sistema al Paciente 1: Imagen Nodular de Categoría BIRADS 4, Proyección MLO.....</i> | 127 |
| Figura 88 <i>Diagnóstico Proporcionado por el Sistema al Paciente 1: Imagen Nodular de Categoría BIRADS 4, Proyección CC.....</i> | 127 |
| Figura 89 <i>Diagnóstico Proporcionado por el Sistema al Paciente 2: Imagen sin Nódulos, BIRADS 1, Proyección MLO.....</i> | 128 |
| Figura 90 <i>Diagnóstico Proporcionado por el Sistema al Paciente 2: Imágenes Nodulares de Categorías BIRADS 3 y 5, Proyección CC.....</i> | 129 |
| Figura 91 <i>Diagnóstico Proporcionado por el Sistema al Paciente 3: Imágenes Nodulares de Categorías BIRADS 4 y 5, Proyección MLO.....</i> | 130 |
| Figura 92 <i>Diagnóstico Proporcionado por el Sistema al Paciente 3: Imágenes Nodulares de Categorías BIRADS 3 y 4, Proyección CC.....</i> | 130 |
| Figura 93 <i>Diagnóstico Proporcionado por el Sistema al Paciente 4: Imagen Nodular de Categoría BIRADS 4, Proyección MLO.....</i> | 131 |
| Figura 94 <i>Diagnóstico Proporcionado por el Sistema al Paciente 4: Imágenes de Categorías BIRADS 4, Proyección CC.....</i> | 132 |
| Figura 95 <i>Diagnóstico Proporcionado por el Sistema al Paciente 5: Imágenes Nodulares de Categoría BIRADS 5, Proyección CC.....</i> | 133 |

RESUMEN

En el presente trabajo de titulación se desarrolló un sistema de reconocimiento y clasificación de nódulos mamarios, mismo que tiene como primer paso crear una base de datos con mamografías que contenga todos los tipos de lesiones nodulares, luego se aplican técnicas de procesamiento digital de imágenes como: binarización, segmentación, filtros, estiramiento de contraste, agrupación de datos, operaciones morfológicas como: cierre y apertura, entre otras con el fin de extraer el nódulo presente en la imagen mamográfica procesada. Se realiza un clasificador basado en redes neuronales convolucionales con entrenamiento supervisado, el cual necesita que las imágenes estén etiquetadas de acuerdo al número de categorías a clasificar, el objetivo de la red es distinguir el tipo de nódulo según la categoría de evaluación BI-RADS (Sistema de informes y datos de imágenes mamarias). Todo el sistema está desarrollado en el lenguaje de programación Python, es decir software libre, el cual es un lenguaje de alto nivel y no requiere de pago de licencias para su uso. La edición se realizó en la Suite Anaconda con la aplicación Jupyter Notebook que nos brinda facilidades de ejecución de código, y librerías para el desarrollo de la etapa de procesamiento, de la red neuronal convolucional y para la interfaz gráfica. El sistema tiene una interfaz gráfica, que permite ingresar al usuario una imagen mamográfica, ésta a su vez es procesada y analizada para clasificarla según la categoría de evaluación (BI-RADS).

Palabras clave:

- **NÓDULOS MAMARIOS**
- **PROCESAMIENTO DIGITAL DE IMÁGENES**
- **REDES NEURONALES**
- **BI-RADS**

ABSTRACT

The present work presents a system for recognition and classification of breast nodules. The system firstly create a database with mammograms containing all types of nodular lesions, then digital imaging techniques such as: binarization, segmentation, filters, contrast stretching, data grouping, morphological operations as: closing and opening, among others, are used to extract the nodule present in the processed image. A classifier based on convolutional neural networks with supervised training is used, requires that the images be labeled according to the number of categories to be classified, the objective of the network is to distinguish the type of nodule according to the BI-RADS (Breast Image Data and Reporting System) evaluation category. The entire system is developed in Python language, a free software, which is a high-level language that do not need the use of licenses. The edition was carried out in the Anaconda Suite with the Jupyter Notebook application which provides us with code execution facilities, and libraries for the development of the processing, the convolutional neural network and the graphical interface. The system has a graphical interface, which allows the user to enter a mammographic image, which is at the same time processed and analyzed to classify it according to the evaluation category (BI-RADS).

Keywords:

- **MAMMARY NODULES**
- **DIGITAL IMAGE PROCESSING**
- **NEURONAL NETWORKS**
- **BI-RADS**

1 Introducción

1.1 Planteamiento del problema

Según los datos de la Organización Mundial de la Salud (OMS, 2018), se reportan 1.3 millones de casos y 458.000 muertes por cáncer de mama cada año, siendo ésta una de las enfermedades más comunes en mujeres. En el Ecuador, el Instituto Nacional de Estadísticas y Censos (INEC, 2017), indica que entre los años 2014 y 2017, 2.760 personas murieron por cáncer de mama. Por esta razón, es importante detectar a tiempo lesiones mamarias con la finalidad de que haya una reducción en la tasa de mortalidad por esta causa.

La mamografía es un examen radiológico que utiliza rayos x para obtener en forma de imagen el tejido mamario, es el más confiable y común para detectar a tiempo lesiones en la mama (Clavero, Molina, & Piccinini, 2016), cuya interpretación y diagnóstico depende en ocasiones de la experiencia del radiólogo. Una de las principales lesiones mamarias, son los nódulos, mismos que son objeto de análisis constante para determinar su benignidad y malignidad. Esto se lo realiza a través de un estudio imagenológico mamario, en el cual se evalúan rasgos específicos como: forma, tamaño, márgenes y calcificaciones, que permitan su adecuada categorización y manejo.

Actualmente nuestro país no cuenta con sistemas inteligentes que apoyen a los especialistas en la detección y diagnóstico de nódulos mamarios en imágenes mamográficas. Normalmente se aplican técnicas convencionales que provocan en ocasiones falsos positivos, demora en el diagnóstico y la aplicación de un tratamiento médico inadecuado al paciente.

1.2 Antecedentes

El procesamiento digital de imágenes tiene por objetivo extraer características de una imagen. En sus inicios por falta de recursos computacionales las aplicaciones

estaban destinadas a campos de investigación tales como matemática, computación, astronomía, y se consideraba costoso, sofisticado y complejo. Actualmente con el avance tecnológico el campo de las aplicaciones se ha expandido, llegando a ser indispensable en: tareas de control de calidad, sistemas de locomoción de robots, ciencias ambientales, seguridad, medicina, etc. (Cuevas, Zaldivar, & Pérez, 2010)

La medicina es la ciencia que da mayor utilidad a los diferentes métodos de procesamiento digital de imágenes, pues se sirven de estos para, mejorar la detección de anomalías en imágenes radiológicas, realizar reconstrucción, restauración, segmentación de imágenes biomédicas, análisis de texturas sólidas en imágenes 2D Y 3D, entre otras tareas. Con el fin de mejorar la interpretación de imágenes, se han implementado sistemas de diagnóstico asistido por computadora los cuales realizan extracción y selección de características de una zona de interés predefinida, estos sistemas tienen la particularidad de poseer algoritmos inteligentes mismos que se aplican para resolver tareas complejas y no lineales. (Stoitsis, y otros, 2006)

Las redes neuronales tienen el objetivo de procesar información de forma similar al de un sistema nervioso biológico, es decir, realizar múltiples operaciones de forma simultánea. La principal ventaja de una red neuronal es la capacidad natural de adquirir conocimiento a través de experiencia o entrenamiento, por tal motivo pueden ser ocupadas de forma individual o combinada con el fin de resolver tareas que necesiten aprendizaje en tiempo de ejecución y con tolerancia a fallos. (Izaurieta & Saavedra)

1.3 Justificación e Importancia

En el Ecuador existen aproximadamente 2.787 nuevos casos de cáncer de mama en mujeres por año, con una tasa estandarizada de 31.8 x 100.000, es decir que, por cada 100.000 mujeres, 33 casos son diagnosticados con cáncer de mama (GCO, 2018), esto lo convertiría en el tipo de cáncer de mayor incidencia en mujeres.

En la actualidad, la mamografía es el examen radiológico que más se practica a mujeres para detectar anomalías en el tejido de la mama. Los médicos pueden encontrar nódulos, que muchas de las veces no son palpables por su tamaño, o calcificaciones mamarias con este examen. Sin embargo, la mamografía no es 100% exacta y puede ocasionar un diagnóstico falso positivo de parte del especialista.

De esta manera se busca nuevas herramientas y tecnologías, que se centren en investigaciones de manejo de algoritmos y redes neuronales, para realizar aplicaciones en imágenes biomédicas con el fin de ayudar y mejorar el diagnóstico de diferentes padecimientos. Un sistema basado en Inteligencia Artificial (IA) trata de reducir el número de regiones bajo sospecha y ofrecer información sobre la posible presencia de una anomalía gracias a la utilización de algoritmos predictivos y redes neuronales.

Por ello, el presente trabajo propone el uso de mamografías para encontrar áreas con características anómalas mediante procesamiento de imágenes y redes neuronales que permitan reconocer ciertas patologías mamarias en su estado inicial y de esta manera permita al especialista auxiliar y facilitar en el diagnóstico de las enfermedades.

1.4 Objetivos

1.4.1 Objetivo General

Desarrollar un sistema para la clasificación y reconocimiento de nódulos en imágenes mamográficas utilizando procesamiento digital de imágenes y redes neuronales.

1.4.2 Objetivos Específicos

- Identificar los tipos habituales de nódulos que se pueden encontrar en imágenes mamográficas.
- Crear una base de datos con imágenes mamográficas proporcionadas por el centro médico Láser, que presenten nódulos de diferentes tipos.

- Realizar el procesamiento de imagen para extraer las características esenciales de los nódulos en mamografías.
- Desarrollar un método de aprendizaje basado en redes neuronales artificiales para la clasificación de nódulos en imágenes mamográficas.
- Evaluar los resultados de detección de nódulos en mamografías mediante pruebas que validen el sistema implementado.

1.5 Variables de la investigación

1.5.1 *Variable Independiente*

Procesamiento digital de imágenes y redes neuronales

1.5.2 *Variable Dependiente*

Clasificación y reconocimiento de nódulos en imágenes mamográficas.

1.6 Hipótesis

El desarrollo del sistema basado en procesamiento digital de imágenes y redes neuronales permitirá la clasificación y reconocimiento de nódulos en imágenes mamográficas.

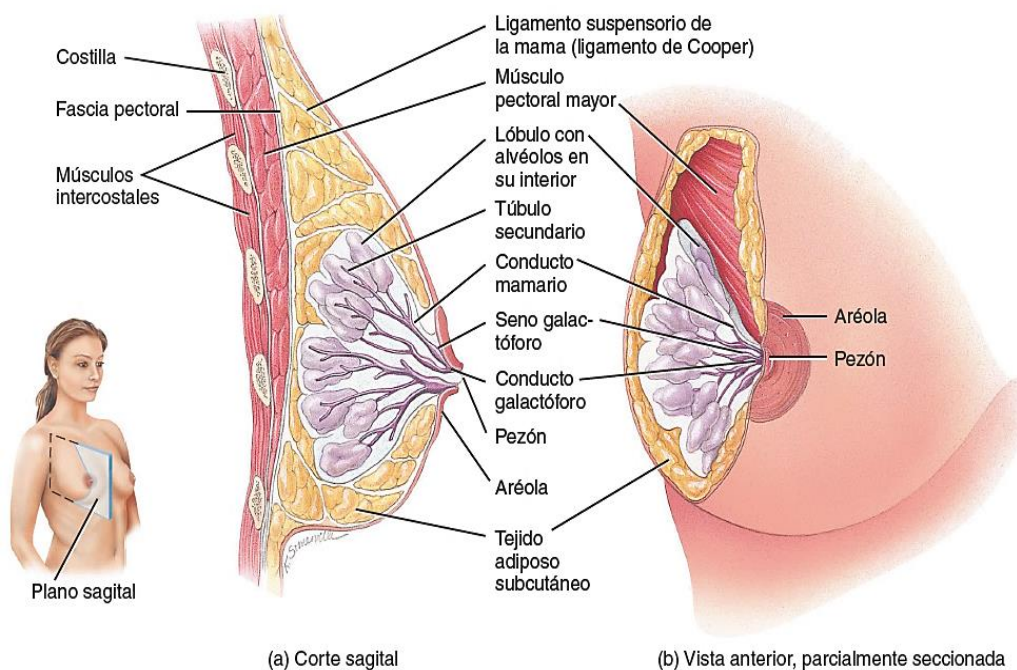
2 Marco Teórico

2.1 Anatomía de la Mama

La mama consiste en la glándula mamaria y un estroma fibroso como se observa en la Figura 1. Cada mama está constituida entre quince a veinte lóbulos, distanciados por tejido denso o graso. Los espacios pequeños que se encuentran entre los lóbulos son conocidos como lobulillos, conformado por: ramificaciones de células productoras de leche y alvéolos rodeados de tejido conectivo. Así mismo, se observa que la glándula mamaria no posee músculo, pero si se encuentra el músculo pectoral debajo de ella. (Tortora & Derrickson, 2013)

Figura 1

Glándulas Mamarias en el Interior de las Mamas



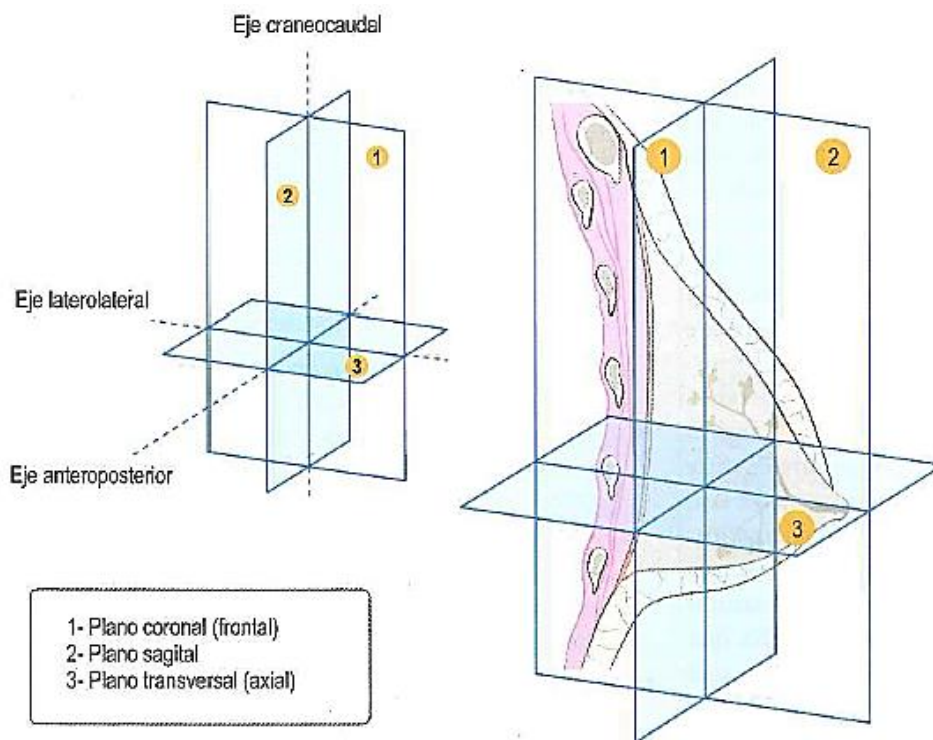
Nota. Tomado de *Principios de Anatomía y Fisiología*, por Tortora & Derrickson, 2013, Panamericana.

2.1.1 Cuadrantes de las Mamas

Para el análisis de la mama se consideran los planos transversal (o axial), sagital y coronal (o frontal) como se observa en la Figura 2.

Figura 2

Orientación de la Mama Según Planos.



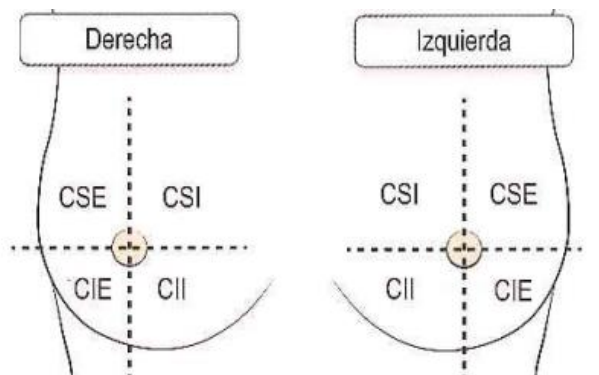
Nota. Tomado de *Manual de Radiología Mamaria*, por López Ruiz & Pina Insausti, 2016, Medica Panamericana.

Para el estudio y localización de lesiones mamarias, como se observa en la Figura 3, la mama se divide en cuatro cuadrantes, denominados:

- Superoexterno (CSE)
- Superointerno (CSI)
- Inferoexterno (CIE)
- Inferointerno (CII) (López Ruiz & Pina Insausti, 2016)

Figura 3

Cuadrantes Mamarios.

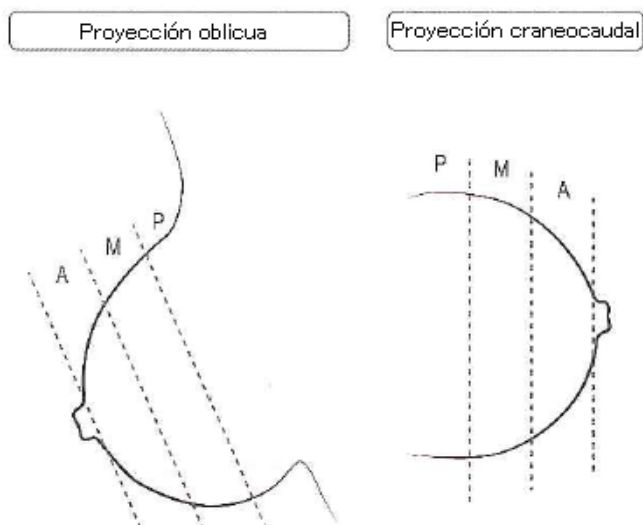


Nota. Tomado de *Manual de Radiología Mamaria*, por López Ruiz & Pina Insausti, 2016, Medica Panamericana.

Para la interpretación y descripción de imágenes radiológicas, específicamente mamografías, la mama se divide en 3 zonas según la profundidad, como se indica en la Figura 4, siendo A: anterior, M: media y P: posterior. (López Ruiz & Pina Insausti, 2016)

Figura 4

Orientación Clínico-Topográfica



Nota. Tomado de *Manual de Radiología Mamaria*, por López Ruiz & Pina Insausti, 2016, Medica Panamericana.

2.2 Mamografía

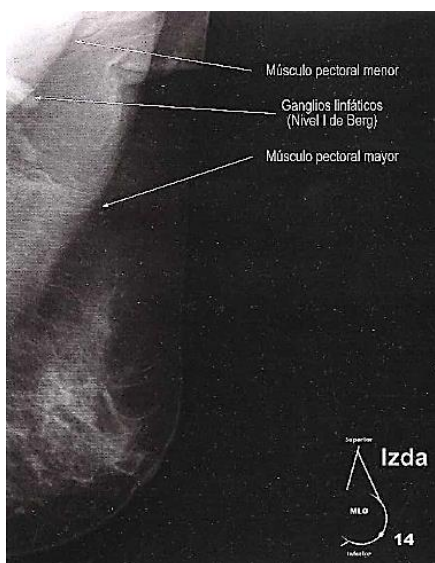
La mamografía es una imagen radiológica plana que permite visualizar el tejido y la estructura de la mama a través de una incidencia de rayos X. Es el análisis más utilizado para detectar lesiones mamarias, especialmente para el diagnóstico de cáncer en mujeres. (Reyes López, 2015)

Según (Reyes López, 2015), Existen dos proyecciones con las que se realiza una mamografía:

- Cráneo-Caudal: Es una proyección vertical, vista superior a inferior de la mama, en la que se puede observar con mayor precisión el tejido medial. La proyección
- MLO (Media-Lateral-Oblicua): Se realiza a 45 grados con respecto del eje de simetría del cuerpo de la paciente, una ventaja es que se observa el músculo pectoral y los ganglios linfáticos, como se indica en la Figura 5. (pág.9)

Figura 5

Proyección Media-Lateral-Oblicua (MLO).



Nota. Tomado de *Manual de Radiología Mamaria*, por López Ruiz & Pina Insausti, 2016, Medica Panamericana.

2.2.1 Mamografía digital

La mamografía digital de campo completo (MDCC) consiste en un sistema de mamografía en el que la película de Rayos X es sustituida por dispositivos electrónicos que transforman los Rayos X en señales eléctricas, los cuales se utilizan para obtener la imagen de la mama con una menor dosis de radiación, este tipo de imagen permite una mejor visualización y un almacenaje directo en una computadora. (Narváez, 2010)

Actualmente el uso constante de imágenes digitales en diferentes exámenes médicos como: radiología, cardiología, etc, ha hecho necesario crear un formato único entre éstas, así se crea el estándar de imágenes médicas es el DICOM (Digital Imaging and Communications in Medicine, traducido como: Imágenes digitales y comunicaciones en medicina), que permite visualizar, imprimir, almacenar y facilitar la operabilidad, transmisión y conectividad entre múltiples equipos médicos. Además de tener la particularidad de guardar información del paciente en la imagen. (Narváez, 2010)

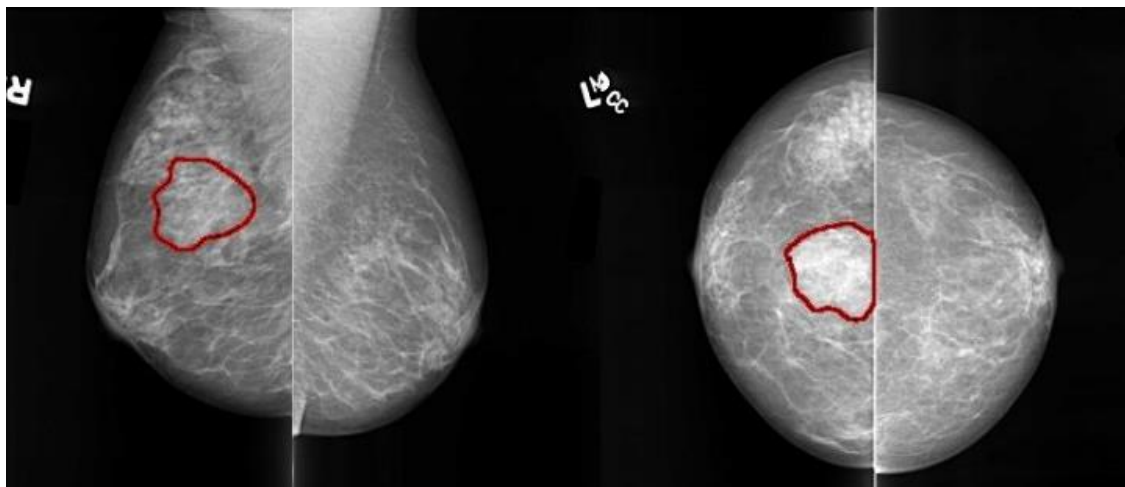
2.3 Lesiones Mamográficas

2.3.1 Asimetría global y focal

El tejido se distribuye de una forma diferente en cada mama como se puede observar en la Figura 6, la mama izquierda presenta un cambio en la estructura del tejido respecto a la mama derecha. La asimetría global es cuando en más de un cuadrante de la mama se aprecia un tejido glandular que no es similar en localización en comparación de la otra mama. En la asimetría focal se aprecia un tejido glandular en menos de un cuadrante, que puede significar el inicio a la formación de una masa, pero no contiene ninguna propiedad asociada a ésta. (Narváez, 2010)

Figura 6

Detección de Asimetría Bilateral



Nota. Tomado de *Sistema de Anotación para Apoyo en el Seguimiento y Diagnóstico de Cáncer de Seno* (p.5), por F. Narváez, 2010, Universidad Nacional de Colombia.

2.3.2 Calcificaciones

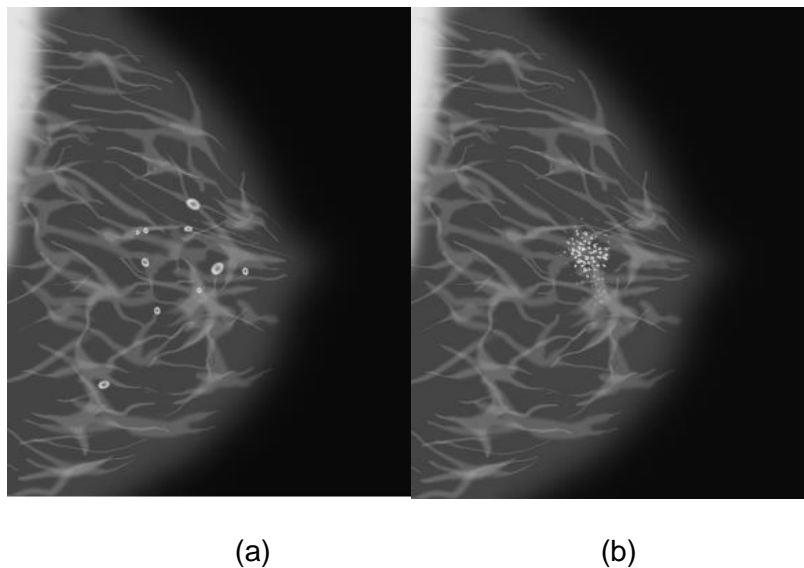
Las calcificaciones son pequeños cúmulos de calcio, que se observan en el tejido mamario, una de sus características más importantes es la de presentar un mayor contraste en la mamografía. (Narváez, 2010)

Según (Komen, 2016) se pueden encontrar 2 tipos de calcificaciones:

- **Macrocalcificaciones.** Son depósitos de calcio más grandes como se observa en la Figura 7 (a), que en su mayoría se debe al envejecimiento de las arterias mamarias, lesiones viejas o inflamación.
- **Microcalcificaciones.** Son diminutas manchas de calcio en el seno como se observa en la Figura 7 (b). la presencia de éstas en la imagen mamográfica puede indicar que el paciente tiene cáncer, pero no siempre es así. (pág. 1-2)

Figura 7

Calcificaciones: a) Macrocalcificaciones, b) Microcalcificaciones



Nota. Tomado de *Las calcificaciones en el seno*, por S. Komen (2016), Cancer Foundation.

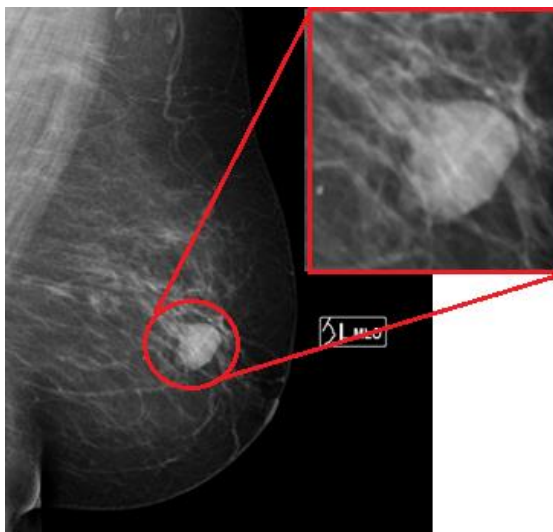
2.3.3 Masas

Una masa, conocida también como nódulo, es una protuberancia o un bulto que tiene forma y márgenes, como se distingue en la Figura 8, que resaltan en el tejido mamario. (Narváez, 2010)

(Narváez, 2010) explica que un nódulo se define como: “una lesión que ocupa espacio observado en dos proyecciones diferentes; si no se puede observar en ambas proyecciones se considera una asimetría”. Este concepto fue establecido por el Colegio Americano de Radiología (American College of Radiology) en el estándar BI-RADS de nódulos o masas, de sus siglas en inglés: Breast Imaging Reporting and Data Systems. (pág. 6)

Figura 8

Masa o Nódulo con Forma Ovalada y Margen Circunscrito.



Nota. Tomado de Clínica Láser, Latacunga

2.4 Clasificación BI-RADS para Nódulos

En 1992 se propone el informe mamográfico conocido como BI-RADS (Breast Imaging Reporting and Data System) que traducido al idioma español significa: Sistema de informes y datos de imágenes mamarias. Propuesto por el Colegio Americano de Radiólogos (ACR) en trabajo mutuo con el Colegio Americano de Cirujanos y el Colegio Americano de Patólogos. Lo describen como un sistema de nomenclatura utilizado para interpretar los informes radiológicos para la detección temprana de cáncer de mama a nivel internacional. (López Ruiz & Pina Insausti, 2016)

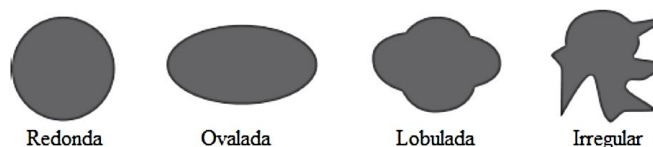
Las características que se definen para los nódulos en el estándar BI-RADS son las siguientes:

2.4.1 Forma

Los nódulos se clasifican por su configuración o apariencia en: Redonda, Ovalada, Lobulada e Irregular. La Figura 9 muestra las posibles formas de un nódulo. (Narváez, 2010)

Figura 9

Formas de Una Masa



Nota. Tomado de *Segmentación de masas en imágenes de mamografía mediante agrupamiento restringido* (p.13), por S. Hernández, 2018, Universidad Tecnológica de la Mixteca.

2.4.2 Margen

En 1992 se propone el informe mamográfico conocido como BI-RADS (Breast Imaging Reporting and Data System) que traducido al idioma español significa: Sistema de informes y datos de imágenes mamarias. Propuesto por el Colegio Americano de Radiólogos (ACR) en trabajo mutuo con el Colegio Americano de Cirujanos y el Colegio Americano de Patólogos. Lo describen como un sistema de nomenclatura utilizado para interpretar los informes radiológicos para la detección temprana de cáncer de mama a nivel internacional. (López Ruiz & Pina Insausti, 2016)

Las características que se definen para los nódulos en el estándar BI-RADS son las siguientes:

Figura 10

Márgenes de Una Masa



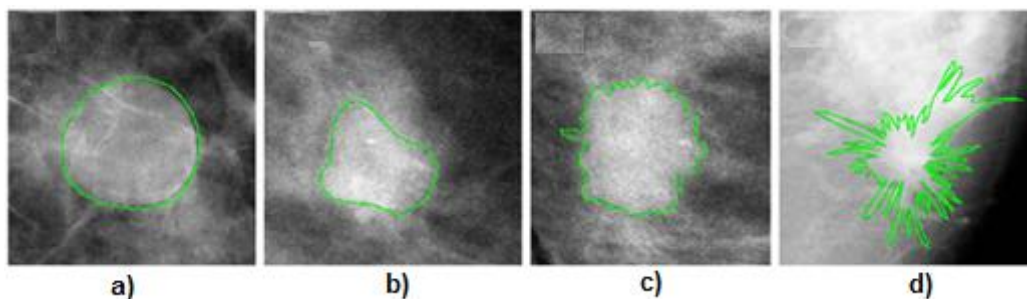
Nota. Tomado de *Segmentación de masas en imágenes de mamografía mediante agrupamiento restringido* (p.13), por S. Hernández, 2018, Universidad Tecnológica de la Mixteca.

El BI-RADS estima el nivel de severidad (benigna o maligna), según las características morfológicas. Generalmente las masas benignas tienen márgenes circunscritos, semejante a formas circulares o elípticos. Las masas malignas usualmente tienen márgenes oscurecidos, con forma irregular, algunas veces describen patrones radiantes o linealmente espiculados. (Narváez, 2010)

La Figura 11 presenta diferentes tipos de masas en la cual se observan las características correspondientes a formas y márgenes. (Narváez, 2010)

Figura 11

Formas y Márgenes: (a) Margen Circunscrito, (b) Margen Indefinido, (c) Margen Microlobulado y (d) Margen Espiculado.



Nota. Tomado de *Sistema de Anotación para Apoyo en el Seguimiento y Diagnóstico de Cáncer de Seno* (p.9), por F. Narváez, 2010, Universidad Nacional de Colombia.

2.4.3 Densidad

La densidad de un nódulo o masa es la disminución de intensidad de los rayos X por la lesión, considerando la atenuación esperada para un volumen de tejido glandular similar. Esta juega un papel importante ya que al tener una mayor densidad mamográfica existe un mayor riesgo de la formación de cáncer de mama. (Shinde, 2003). Según (Rodríguez López, 2012) la densidad de una masa está también relacionada con el tejido graso de la mama esta puede ser: alta, isodensa, baja y radiolúcida (pág. 41)

2.4.4 Composición del Tejido

Entrega una visión general del tejido en la mama. De esta forma ayuda a encontrar alguna anomalía en el tejido. En este caso influye la edad, el ciclo menopáusico y alteraciones hormonales del paciente. Generalmente incluye la presencia de tejido graso mezclado con las glándulas. (Narváez, 2010)













2.5 Categorías de Evaluación

Las categorías de evaluación definen una interpretación estándar de los hallazgos en las mamografías. Cada una de estas categorías se utiliza para el seguimiento, control y evaluación del caso de estudio.

La Figura 12 muestra una descripción con los hallazgos gráficos más representativos de cada categoría BI-RADS.

Figura 12

Categorías BI-RADS para Masas.

| Categorías BI-RADS 2004. Masas <small>Breast Imaging Reporting And Data System (BI-RADS), American College of Radiology (ACR)</small> | | | | Román Rostagno y col. |
|--|--------------------------|--|--|--------------------------|
| BI-RADS | DESCRIPCIÓN | VPP <small>Very Probable Positive</small> | HALLAZGOS | SUGERENCIA |
| BR1 | mama normal | | | control habitual |
| BR2 | patología benigna | |   | control habitual |
| BR3 | sugestiva de benignidad | < 2% |    | control 6 meses |
| BR4A | baja a moderada sospecha | ~5% 2 al 10% |   | punción cito/histológica |
| BR4B | moderada sospecha | ~25% 11 al 40% |   | estudio histológico |
| BR4C | moderada a alta sospecha | ~70% 41 al 94% |   | estudio histológico |
| BR5 | alta sospecha (clásica) | ~95% |  | estudio histológico |
| BR6 | malignidad confirmada | ~100% | anatomía patológica positiva | no hay sugerencias |
| BR0 | estudio insuficiente | | opacidades, asimetrías formaciones palpables | estudios complementarios |

Nota. Tomado de Diagnóstico por Imágenes – Imágenes Mamarias (p.6), por S. Ciardullo.

2.6 Procesamiento Digital de Imágenes.

El procesamiento digital de imágenes es una secuencia de operaciones realizadas con la ayuda de un computador, que se tienen como tarea principal modificar las propiedades o características de una imagen digital; mejorando así su calidad, cambiando sus dimensiones, entre otros, obteniendo de esta manera una nueva imagen. (Carrasco LLerena, 2017)

Las técnicas que se ocupan en el procesamiento de imagen se utilizan para optimizar el aspecto de las imágenes, la adapta para ser receptada por parte de una máquina. (Gómez & Guerrero, 2016).

El proceso digital de imágenes incluye las siguientes etapas:

- **Adquisición o captura:** Importar imágenes por diferentes herramientas.
- **Realce y mejora:** Es la etapa en donde mediante técnicas se resalta el aspecto visual de regiones importantes de las imágenes.
- **Segmentación:** Se encarga de delimitar las imágenes en regiones o áreas significativas.
- **Extracción de características:** Se encarga del reconocimiento y localización de objetos geométricos simples y complejos. Estas pueden ser líneas, puntos, curvas y cuádricas. (Gómez & Guerrero, 2016)

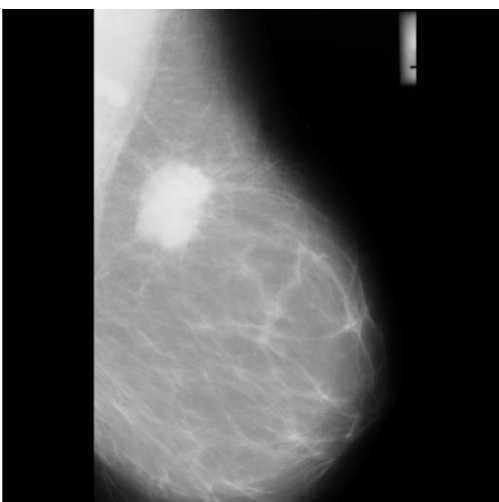
2.6.1 *Adquisición*

- **Imagen en Escala de Grises.** Cada celda de la matriz de una imagen en escala de grises está representada por 1 píxel, cuyo valor dependerá del número de niveles de grises que posea la imagen. El número de niveles posibles al aumentar el número de bits será 2^n , siendo n el número de bits, es decir obtendremos 256 niveles de grises con ocho bits. Por lo general las imágenes en escalas de grises se crean con ocho bits, con los valores de (0)

negro, (255) blanco, los 254 tonos grises restantes son suficientes para que el ojo humano no detecte transiciones agresivas, como se observa en la Figura 13. Una buena imagen depende del número de píxeles y de niveles de grises que posea. (Costa Campos & Fernández Bozal, 2005)

Figura 13

Mamografía en Escala de Grises



Nota. Tomado de *MIAS Mammography*, por Kaggle Dataset.

2.6.2 Realce y mejora

Teniendo una imagen digital, se aplica esta técnica para mejorar el aspecto visual y la calidad de la imagen, y de esta manera facilitar su procesamiento para una aplicación específica. (Gómez & Guerrero, 2016)

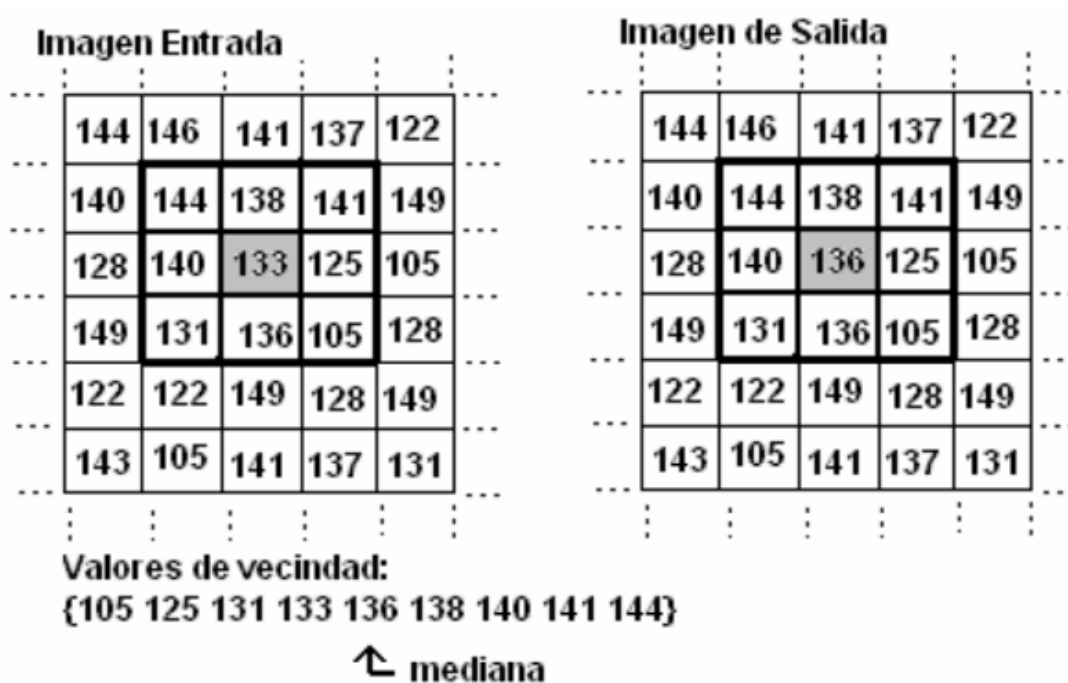
Durante la etapa de adquisición y digitalización de la imagen, surge el ruido, que son datos no deseados que dificultan el procesamiento. Por esto, es necesario implementar un método de reducción de ruido, y disminuir los efectos adversos que produce. (Álvarez G., Guevara G., & Holguín L., 2006)

- **Filtrado de mediana.** Este filtro se usa para reducir el ruido encontrado en una imagen. Su funcionamiento consiste en recorrer cada píxel de la imagen y cambiar el valor del píxel analizado por el valor de la mediana de los

píxeles vecinos. La característica principal de este filtro es hacer que los puntos de una imagen de muy alta o muy baja intensidad sea reemplazado por puntos de intensidad parecida a sus vecinos, la forma de aplicación de este filtro se ilustra en la Figura 14. (Álvarez G., Guevara G., & Holguín L., 2006)

Figura 14

Funcionamiento Filtrado de Mediana.



Nota. Tomado de *Preprocesamiento de Imágenes Aplicadas a Mamografías Digitales* (p. 2), por Álvarez et al., 2006, Universidad Tecnológica de Pereira.

- **Realce de contraste.** Generalmente las imágenes médicas tienen un pobre contraste y las mamografías no son la excepción. Por esto, es necesario aplicar un método que permita aumentar el contraste de estas imágenes para un procesamiento posterior o simplemente realizar un análisis. (Álvarez G., Guevara G., & Holguín L., 2006)

- **Técnicas de realce por ampliación del contraste.** Se aplica para obtener un aumento o dilatación del contraste que el usuario requiera para la imagen. Como el nombre lo indica, realza el contraste y permite resaltar la estructura del tejido mamario, mediante el ajuste de histograma el cual incrementa la separación del nivel de grises del fondo y los objetos. (Álvarez G., Guevara G., & Holguín L., 2006)

La Ecuación 1, explica (Álvarez G., Guevara G., & Holguín L., 2006) describe un modelo de transformación de la técnica de realce de contraste:

$$f(x) = \begin{cases} f_1(x), & 0 \leq x \leq S_1 \\ f_2(x), & S_1 \leq x \leq S_2 \\ \vdots & \\ f_n(x), & S_{n-1} \leq x \leq L \end{cases} \quad (1)$$

Donde:

- S_i : son los valores que separan los diferentes modos del histograma
- L : el valor de nivel de gris máximo (pág. 4)

2.6.3 Segmentación

La segmentación consiste en simplificar o subdividir una imagen en varios píxeles, con el fin de aislar un objeto o región de interés y poder obtener sus propiedades para analizarla con mayor facilidad. Las propiedades que sean extraídas de la imagen definirán la efectividad de la técnica que se elija utilizar. (Carrero, 2007).

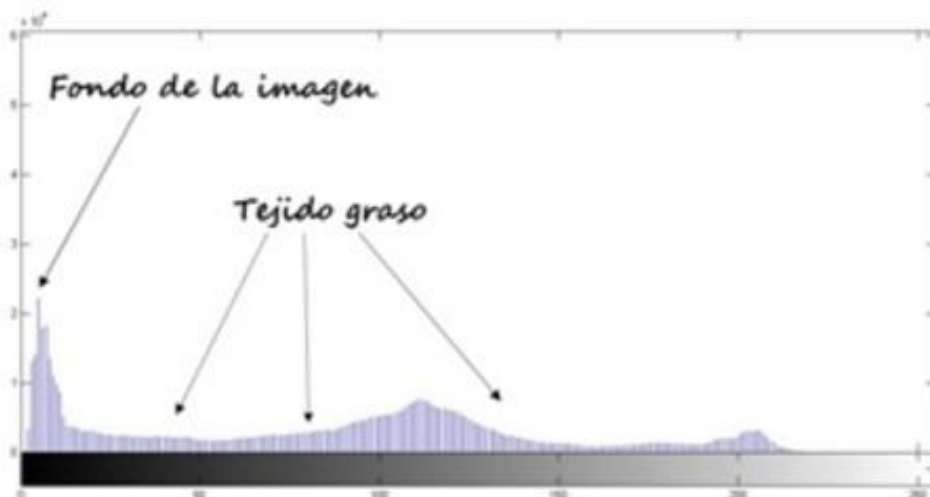
Los métodos de segmentación son:

- a) Detección de Bordes.** Cuando la imagen está compuesta por áreas de diversas características, la detección de bordes se sitúa en la frontera o límite de dos áreas diferentes para poder determinar sus regiones e información. Los bordes pueden distinguirse por el ancho, el ángulo de su pendiente de variación, y las coordenadas de su punto medio. (Gómez & Guerrero, 2016)

- b) Umbralización.** Es una técnica que convierte una imagen, que presenta diferentes tonos de grises, a una imagen expresamente de dos colores: blanco y negro. De esta manera, permite extraer los objetos y por lo tanto los separa del fondo. (Gómez & Guerrero, 2016)
- c) Histograma.** Es una representación gráfica con valores discretos que muestra el número de píxeles de los diferentes tonos de gris de la imagen. Su proceso de segmentación consiste en elegir un valor de umbral o umbrales adecuados para poder agrupar los píxeles de la imagen de acuerdo a sus características y a sus niveles de gris. (Gómez & Guerrero, 2016). Por ejemplo, en la Figura 15, se encuentra el histograma de la densidad de una mamografía, donde se observa que el fondo de la imagen tiende a tonos oscuros con una gran cantidad de píxeles, en cambio, el tejido graso tiende a tonos claros y cantidades pequeñas de píxeles. (Menéndez Alonso, 2016)

Figura 15

Histograma de Densidad de Una Mamografía



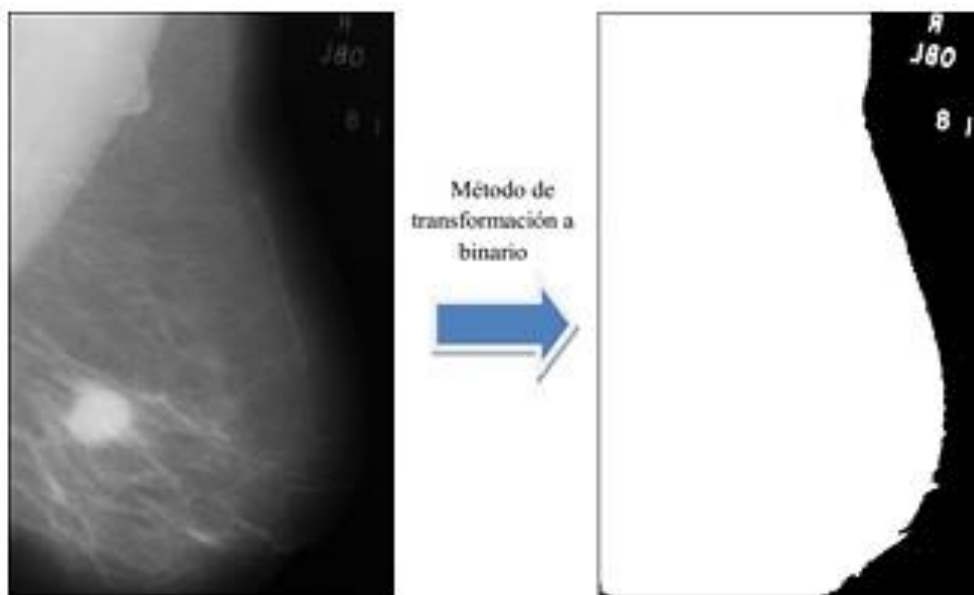
Nota. Tomado de *Detección de masas en mamografías asistida por computadora*, por Menéndez et al., 2016, Revista Cubana de Informática Médica.

d) **Binarización.** Permite que el objeto de una imagen sobresalga del fondo, Por ejemplo, una imagen con fondo de tonos bajos, los píxeles de esta región serán definidos en “0”, mientras que las regiones de tonos claros serán definidos con “1”. Si tenemos casos opuestos, el procedimiento también es inverso. (Gómez & Guerrero, 2016)

Las imágenes mamográficas presentan ciertas particularidades en su contraste, debido a las densidades de la mama. En objetos como el músculo, nódulos y tejidos se tiene una alta densidad. Mientras que, en el tejido graso se tiene una baja densidad. (Menéndez Alonso, 2016). Como se observa en la Figura 16, se realiza un proceso de binarización en una mamografía, donde la región de interés está delimitada sin afectar la silueta de la imagen.

Figura 16

Binarización de Una Mamografía



Nota. Tomado de *Detección de masas en mamografías asistida por computadora*, por Menéndez et al.,2016, Revista Cubana de Informática Médica.

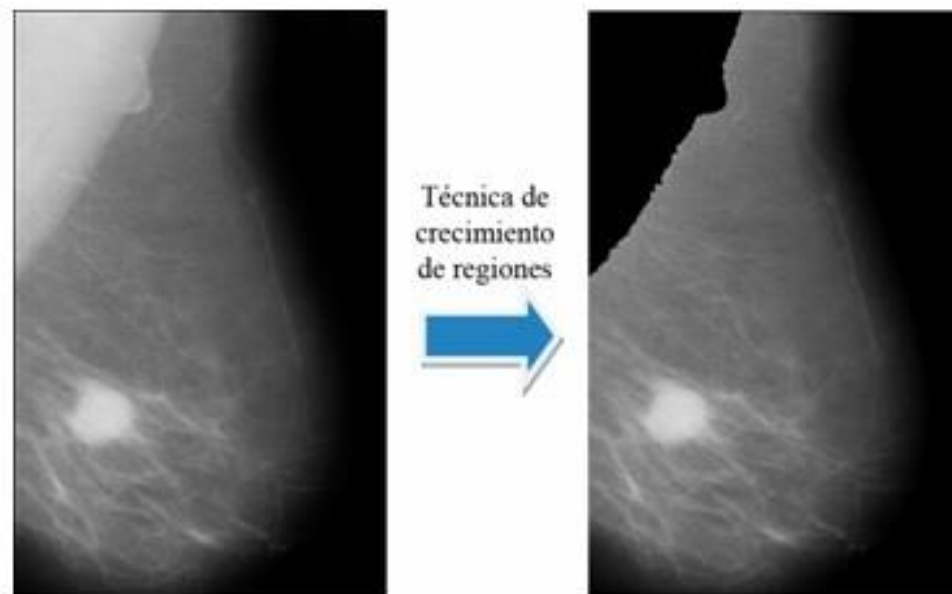
2.6.4 Basado en Regiones

Esta técnica se enfoca en las áreas de interés de la imagen. La imagen se divide en distintas áreas, que poseen propiedades particulares. Se toman en cuenta características como intensidad, textura, etc.; para distinguir un área de la otra. (Gómez & Guerrero, 2016)

En procesos de reconocimiento de objetos, la aplicación de la técnica basada en regiones es muy empleada. Así encontramos ejemplos aplicados en imágenes mamográficas donde se emplean métodos correctores y limitadores para que el músculo pectoral, que se observa en la mamografía de la Figura 17, sea extraído y separado del área de la mama. (Menéndez Alonso, 2016)

Figura 17

Eliminación del Músculo Pectoral por el Método Basado en Regiones



Nota. Tomado de *Detección de masas en mamografías asistida por computadora*, por Menéndez et al., 2016, Revista Cubana de Informática Médica.

2.6.5 Algoritmos de Clustering

Esta técnica conocida también como algoritmo de agrupación tiene por objetivo realizar un proceso de organización sucesiva de la información hasta convertirla en conjuntos comprensibles y manejables a los que se denomina clústeres. (Bóveda Álvarez, 2014)

Cada grupo que se forma tiene propiedades similares, y difieren de cada grupo, estas agrupaciones pueden ser formados aplicando métodos como distancia o similitud. El primer método, se basa en funciones de distancia como, por ejemplo: la función euclídea, y el segundo emplea algoritmos de funciones de verosimilitud. (López J. &, 2014)

En este proceso primero se recibe la imagen a segmentar, el vector de características y parámetros específicos; y para luego agrupar automáticamente los píxeles de la imagen de entrada en clústeres a partir de la información disponible. Así los clústeres formados a partir de la segmentación comparten propiedades similares. (Bóveda Álvarez, 2014)

El algoritmo más conocido de Clustering es el siguiente:

- **K-Means.** El algoritmo K-Means, se basa en los conceptos de clustering dichos anteriormente, en donde se realiza un proceso de agrupamiento de datos basándose en sus propiedades, su principal característica es que emplea la media estadística para el cálculo de los nuevos clústeres. El centroide es calculado por el clúster, cuyo valor es la media de todos los elementos asignados al clúster en cuestión.

Para iniciar el algoritmo, primero se define el número de grupos k , el cual establece los centroides del conjunto de datos, Cada punto de los datos es enviado al centroide más próximo y finalmente se actualiza el nuevo centroide de cada grupo de

datos, con el cálculo resultante de la media estadística de los datos que conforman dicha agrupación. (López J. &, 2014)

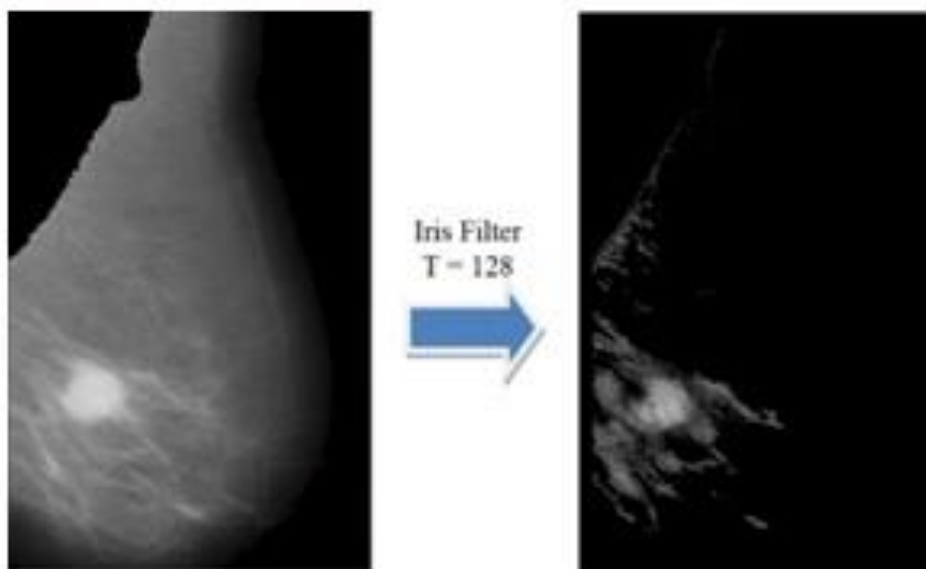
2.6.6 Extracción de características

Esta técnica consiste en extraer las características de interés de la imagen tales como tamaño, colores, área, forma, perímetro; entre otras. También permite reconocer objetos geográficos y el tamaño de la superficie. (Gómez & Guerrero, 2016)

Por ejemplo, la apariencia de una masa en una mamografía consiste en una región de forma circular con un contraste alto. Teniendo en cuenta estas características, se obtiene un mapa de gradientes de la mamografía, la masa se mostrará como una región circular con muchos gradientes señalando hacia el centro. Una manera de evaluar el grado de convergencia del vector gradiente en la vecindad de un píxel de interés es mediante el Filtro Iris como se indica en la Figura 18. (Menéndez Alonso, 2016)

Figura 18

Resultado de la imagen aplicando el Filtro Iris.



Nota. Tomado de *Detección de masas en mamografías asistida por computadora*, por Menéndez et al., 2016, Revista Cubana de Informática Médica.

2.6.7 Operaciones Morfológicas

La morfología se basa en la geometría y forma de las áreas de interés mediante un procesamiento no lineal, permitiendo preservar las características esenciales eliminando aspectos que puedan ser irrelevantes. La morfología se fundamenta en la teoría matemática de los conjuntos que se encuentran como espacios matriciales binarios. (Carrasco LLerena, 2017)

Las operaciones morfológicas matemáticas, son muy utilizadas en:

- La generación de límites o bordes en zonas de interés.
- Adelgazar, ampliar o reducir secciones
- Como filtro morfológico en la supresión de ruido
- Simplificación de formas
- Extracción los esqueletos
- Marcado de objetos,

La imagen resultante de aplicar las operaciones morfológicas es una imagen con apariencia similar a la imagen original, pero con las características de interés realzadas, es decir, que sea más evidente para el usuario. (Álvarez G., Guevara G., & Holguín L., 2006)

2.7 Redes Neuronales

2.7.1 Neurona Biológica

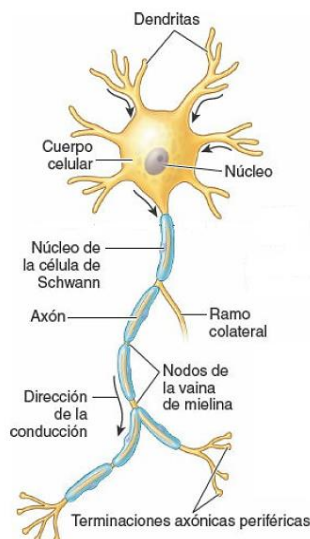
Las neuronas son las unidades estructurales y funcionales del sistema nervioso y están especializadas en el procesamiento de información. Una neurona está formada por un cuerpo celular con espiculaciones denominadas dendritas y un axón, como se indica en la Figura 19, que llevan los impulsos hacia y desde el cuerpo celular, respectivamente. (Agur, Dalley, & Moore, 2019)

Teniendo como ejemplo un sistema, las dendritas serían las entradas, por donde una neurona recibe información de otras y las procesa, forma ramificaciones muy finas desde el cuerpo celular de la neurona. Si la información es tal que genere un potencial eléctrico, se dice que la neurona se dispara, generando una señal de salida que se transporta a otras neuronas a través del axón que se define como la salida del sistema. (Quintanilla Domínguez, 2015)

Las neuronas se comunican entre sí por conexiones que se conocen como sinapsis. La sinapsis está rellena de fluidos neurotransmisores que excitan o inhiben la propagación de la información. Al ser un proceso químico, este crea señales eléctricas que llegan al núcleo celular de la neurona. (Quintanilla Domínguez, 2015).

Figura 19

Estructura básica de una Neurona biológica.



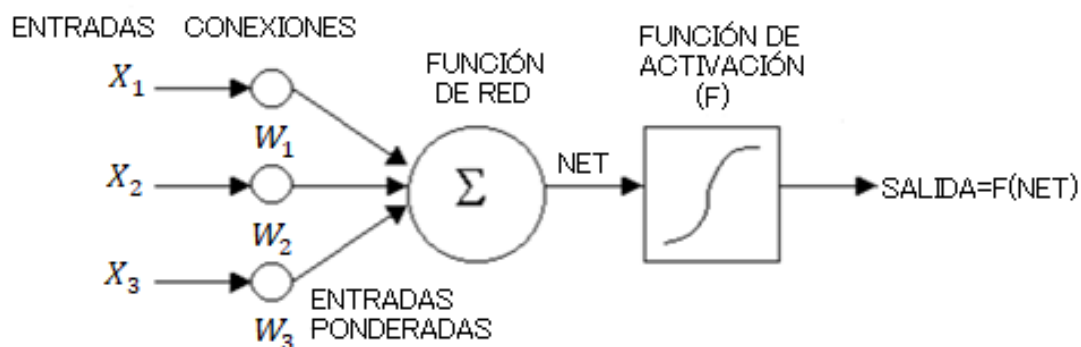
Nota. Tomado de Moore *Fundamentos de anatomía con orientación clínica*, por Agur et al., 2019, Wolters Kluwer.

2.7.2 Neurona Artificial

La neurona artificial es un modelo matemático diseñado para que funcione similar a como trabajaría una neurona biológica. El modelo de una neurona artificial, como se indica en la Figura 20, consiste en un conjunto de entradas a la neurona, cada una de las cuales representa una salida de otra neurona. Las entradas son multiplicadas independientemente por su peso correspondiente en forma análoga al grado de conexión de la sinapsis. Todas las entradas ponderadas se suman y se determina el nivel de activación de la neurona. (Basogain Olabe, 1998)

Figura 20

Modelo de Neurona Artificial.



Nota. Tomado de *Redes Neuronales Artificiales*, por InChaos, <http://rna.50webs.com/>

Una representación vectorial del funcionamiento básico de una neurona artificial se indica en la siguiente expresión de la Ecuación 2.

$$NET = X * W \quad (2)$$

Donde:

- NET es la salida,
- X es el vector de entrada
- W es el vector de pesos.

Comúnmente la señal de salida *NET* va acompañada por un factor conocido como función de activación *F*, cuyo resultado produce la señal de *Salida* de la neurona. Para la función *F* se puede emplear: una función lineal o no lineal, una función umbral, que permiten simular con mejor precisión las propiedades de transferencia no lineales de las neuronas biológicas. (Basogain Olabe, 1998)

También puede ser expresada en un modelo matemático como se indica en la Ecuación 3:

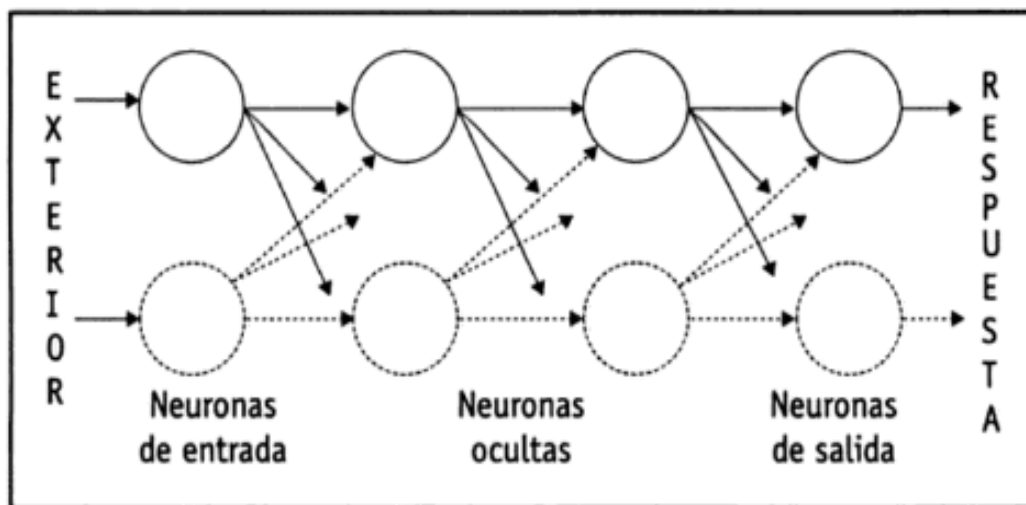
$$Salida = F(\sum_{i=0}^n W_i \cdot X_i) \quad (3)$$

2.7.2.1 Tipos de Neuronas Artificiales. Se encuentran 3 tipos que son:

- Las neuronas de entrada captan información del exterior, la cual es proporcionada por sensores o por otras etapas del sistema.
- Las neuronas de salida envían su respuesta hacia el exterior del sistema, después de que esta fue procesada por los demás elementos de la red (salidas de la red).
- Las neuronas ocultas reciben y entregan información, sin mantener contacto con el exterior. En ellas realizan el procesamiento básico de la información, la representación interna de ésta se observa en la Figura 21. (López & Fernández, 2008)

Figura 21

Tipos de Neuronas Artificiales.



Nota. Tomado de *Las redes neuronales artificiales*, por López & Fernández, 2008, Netbiblo.

2.7.3 Red Neuronal Artificial

Una red neuronal artificial (RNA) es un procesamiento de datos que se asemeja al sistema nervioso de los seres humanos. La arquitectura de una red neuronal es formada interconectando múltiples procesadores elementales, este sistema posee un algoritmo para ajustar sus pesos para alcanzar los requerimientos de un problema en específico basado en muestras representativas. (Salas, 2004)

En consecuencia, obtenemos que una RNA es un sistema de computación distribuida representado por:

- Un conjunto de unidades elementales, las cuales posee escasas capacidades de procesamiento o una estructura interconectada aplicando enlaces ponderados.
- Parámetros libres se ajustan con el fin de satisfacer los requerimientos y obtener un correcto funcionamiento.

- Paralelismo.

La característica más importante de las redes neuronales artificiales es que, a partir de un entrenamiento adecuado del procesamiento de datos, pueden aprender, es decir, es capaz de encontrar un modelo que ajuste los datos. El proceso de entrenamiento de la red puede ser de aprendizaje supervisado o no supervisado.

(Salas, 2004)

- a) Aprendizaje Supervisado.** El aprendizaje supervisado consiste en el entrenamiento de la red neuronal a través de un conjunto de datos compuesto por pares de entrada y salida deseados. El objetivo es encontrar un modelo entre los valores de entrada y salida del proceso. De este modo, el algoritmo ajusta los valores de pesos de la red neuronal y hace que la salida generada por ésta, obtenga el resultado más próximo posible a la verdadera salida dada una entrada. Este tipo de aprendizaje suele ser empleado para procesos de clasificación y problemas de regresión (Simeone, 2018)

En el aprendizaje supervisado se consideran los siguientes pasos: Primero se establece el tipo de red neuronal artificial para el entrenamiento. Segundo, se recopila un conjunto de datos de entrenamiento que describa correctamente un problema dado. Tercero, se clasifica el conjunto de datos de entrenamiento recopilados en una forma comprensible para la red neuronal artificial. Cuarto, se realiza el aprendizaje y después se prueba el rendimiento de la red neuronal artificial aprendida con el conjunto de datos de prueba, esto se conoce como proceso de validación. El conjunto de datos de prueba está formado por datos que no se han insertado en el entrenamiento de la red neuronal artificial. (Krenker, Bešter, & Kos, 2011)

- b) Aprendizaje No Supervisado.** El aprendizaje no supervisado presenta sólo un conjunto de patrones a la RNA, el conjunto de entrenamiento está formado de entradas sin ninguna salida deseada definida. Este aprendizaje describe la

estructura de los datos, y busca algún tipo de organización o agrupación de los datos por ejemplo por similitud. (Simeone, 2018)

2.7.4 Redes Neuronales Convolucionales

Las redes neuronales convolucionales son redes diseñadas con múltiples capas que permiten trabajar específicamente con imágenes bidimensionales, que la hacen adecuada para clasificar patrones. (Ortiz Suárez, 2016)

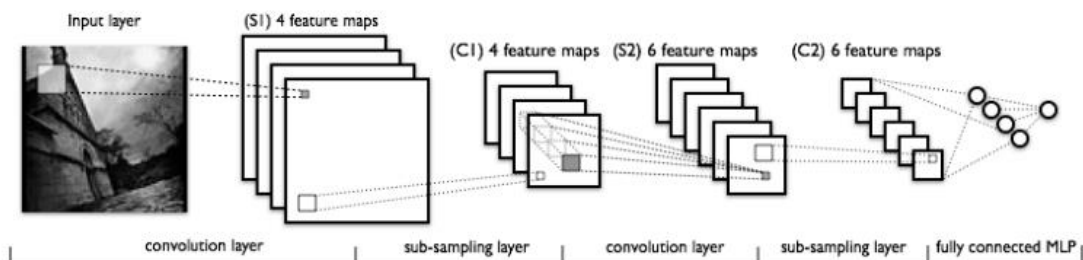
Se observa un modelo o estructura de una red convolucional en la Figura 22, donde se indican las capas con las que está formada.

(Ortiz Suárez, 2016), clasifica y describe el funcionamiento de cada capa que forma una red neuronal convolucional como:

- **Capa de Entrada:** Una imagen de dimensiones ancho, largo y profundidad, por ejemplo, imagen de color RGB (rojo, verde y azul) o una imagen a blanco y negro.
- **Capa convolucional:** Procesa la salida de las neuronas que están conectadas a píxeles más cercanos de la entrada, realiza el cálculo de producto escalar entre los pesos de cada neurona y la región a la cual están conectadas en el volumen de entrada, extraen características mediante la aplicación de filtros y reduce el número de parámetros a analizar. Luego se aplica una función de activación, siendo la más recomendada ReLU (Rectified Linear Units), que permite pasar los valores de convolución a la siguiente capa.
- **Capa sub-sampling o muestreo:** conocida también como pooling, se encarga de reducir los parámetros, por lo que el mapa de características reduce en dimensiones ancho y largo, pero mantiene su profundidad.
- **Capa fully conectada o de clasificación:** como su nombre indica, es una capa clasificadora de capas completamente conectadas, donde el número de clases a predecir será igual a la cantidad de neuronas. (pág.14)

Figura 22

Modelo de Una Red Neuronal Convolutiva Típica.



Nota. Tomado de *Evaluación de métodos con redes neuronales convolucionales para la clasificación de la enfermedad de alzheimer en imágenes de resonancia magnética cerebral* (p.15), por J. Ortiz, 2016, Universidad Industrial De Santander.

2.8 Python

Python, Figura 23, es un lenguaje de programación orientado a objetos, interpretado e interactivo. Conocido por ser software libre y además open-source.

(Holden , 2018)

Figura 23

Logotipo del Software Python



Nota. Tomado de *Python*, por S. Holden, 2018, Python Corporation.

Python tiene una base de librerías y funciones incorporadas en el propio lenguaje, facilitando la realización de programas sin necesidad de tener que empezar desde cero. Lo cual ha llevado a este lenguaje a ampliar su entorno, dándole robustez a sus funciones y lo convierte en un lenguaje de más fácil de aprendizaje. (Berrones Reyes, 2019)

2.8.1 Anaconda Navigator

Para gestionar y compilar los paquetes de Python se utiliza la Suite:

La Suite Anaconda Navigator es una interfaz gráfica de usuario de escritorio incluida en la distribución de Anaconda® que le permite:

- Gestionar
- Compilar
- Iniciar aplicaciones
- Administrar paquetes (Anaconda, 2020)

La edición y ejecución de código se realiza a través de la siguiente aplicación que posee Anaconda Navigator:

2.8.1.1 Jupyter Notebook. Es una aplicación de Anaconda, que combina su código, texto descriptivo, salida, imágenes e interfaces interactivas en un solo archivo de notebook que se edita, visualiza y utiliza en un navegador web. (Anaconda, 2020)

2.8.2 TensorFlow

TensorFlow es un sistema de aprendizaje automático de Google Brain, que trabaja en entornos a gran escala. Es una biblioteca de código abierto para cálculo numérico y que usan una gráfica de nodos enlazados entre sí, que representan las posibilidades, para describir su programación. (Zamorano Ruiz, 2019)

TensorFlow permite la compilación y entrenamiento de modelos por medio de la API de alto nivel de Keras, y hace sencillo el proceso de realización de aprendizaje automático para el usuario. (TensorFlow, 2016)

TensorFlow 2 tiene como características principales simplicidad y facilidad de uso, posee actualizaciones como ejecución inmediata, API's de nivel superior y compila modelos indistintamente de la plataforma. (TensorFlow, 2016)

2.8.3 Keras

Para empezar a realizar una red neuronal en TensorFlow, utilizando Python, se necesita la librería de Keras, que ofrece un conjunto de funciones establecidos para reducir la cantidad de acciones al usuario para la creación de modelos. Una de las ventajas es que proporciona guías y documentación en general, para adentrarse en el mundo del aprendizaje automático. (Keras, 2019)

Según (Berrones Reyes, 2019) el motivo del porqué elige Keras sobre otras librerías es porque es más práctica o sencilla de aplicar, ya que solo se necesita cambiar los parámetros de sus funciones. Así mismo, explica que es necesario instalar todas las librerías en la computadora para poder crear el modelo, recomienda tener una noción de la estructura de la red neuronal que se va a utilizar y buscar las funciones en la página de Keras. (pág. 68)

3 Diseño e Implementación

En el presente capítulo, se describe de manera detallada el diseño e implementación del sistema de detección de lesiones mamarias, la base de datos de mamografías utilizada y la forma de empleo del software Anaconda con el editor de texto de Python, Jupyter Notebook, el cual se utiliza para el reconocimiento, clasificación del nódulo mamario y desarrollo de la interfaz.

3.1 Especificaciones del software y equipo

Para el desarrollo del proyecto se tienen las siguientes especificaciones de software y equipo:

- Python versión 3.7
- Aplicación Anaconda 3
- Sistema operativo Windows 10 de 64 bits
- Equipo con procesador Intel Core i7
- Tarjeta gráfica Nvidia 1080
- Memoria RAM de 16 Gb

3.2 Anaconda Navigator

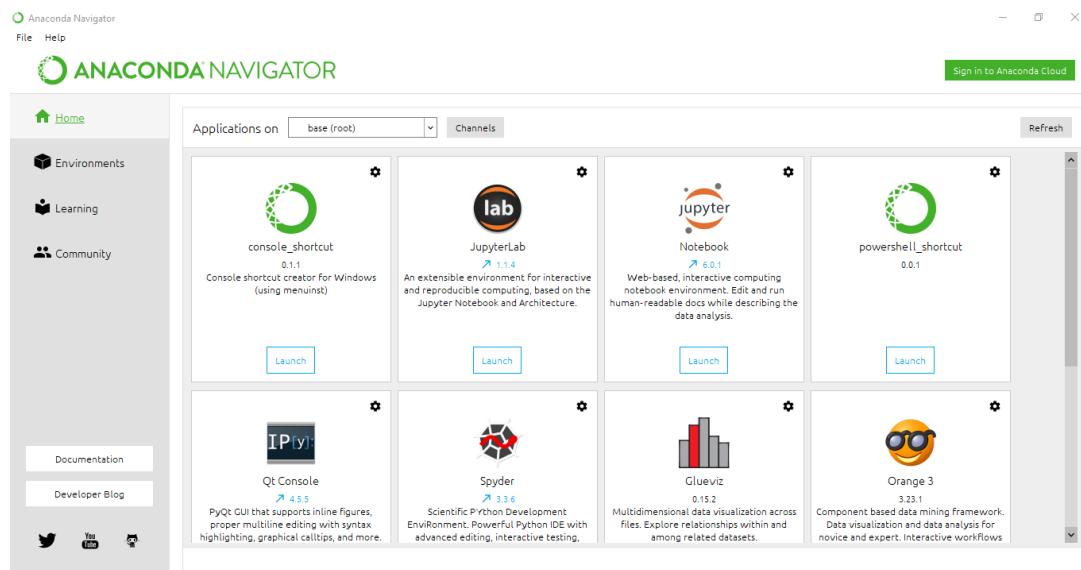
Anaconda Navigator es una interfaz que contiene aplicaciones para desarrollar código R y Python, a través de esta aplicación se puede gestionar entornos y paquetes.

A continuación, se describe el proceso para crear un entorno en el cual se instalarán los paquetes que necesite el editor de lenguaje Python, Jupyter Notebook.

En la aplicación Anaconda Navigator, Figura 24, se selecciona *Environments*.

Figura 24

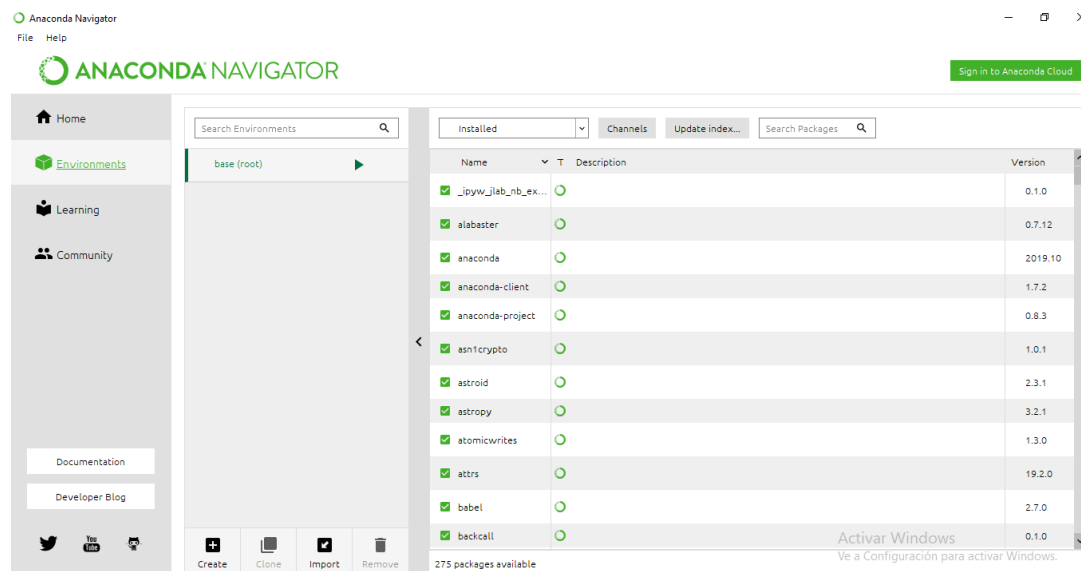
Aplicación Anaconda Navigator



En Enviroments, Figura 25, se selecciona el ícono Create que se encuentra en el inferior de la ventana.

Figura 25

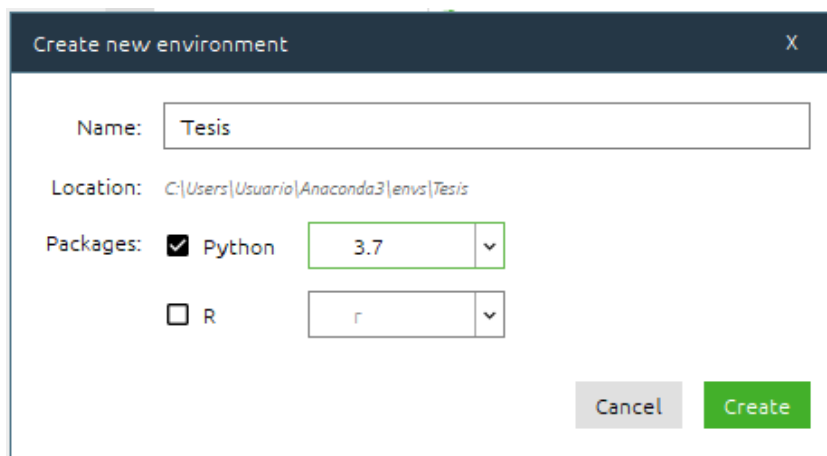
Aplicación Anaconda – Enviroments



Ingresamos el nombre que deseamos dar al entorno, Figura 26, en este caso se llamará *Tesis*, luego elegimos la versión de Python, que será 3.7 y se selecciona *Create*.

Figura 26

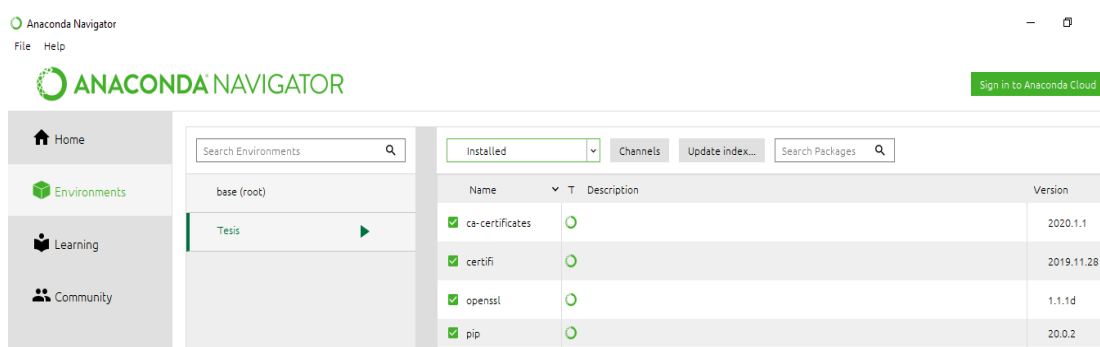
Ventana para Crear un Nuevo Entorno.



En la Figura 27, se observa el nuevo entorno creado llamado *Tesis* y a su derecha se indica las versiones de paquetes que contiene el mismo.

Figura 27

Entorno Creado (Tesis).

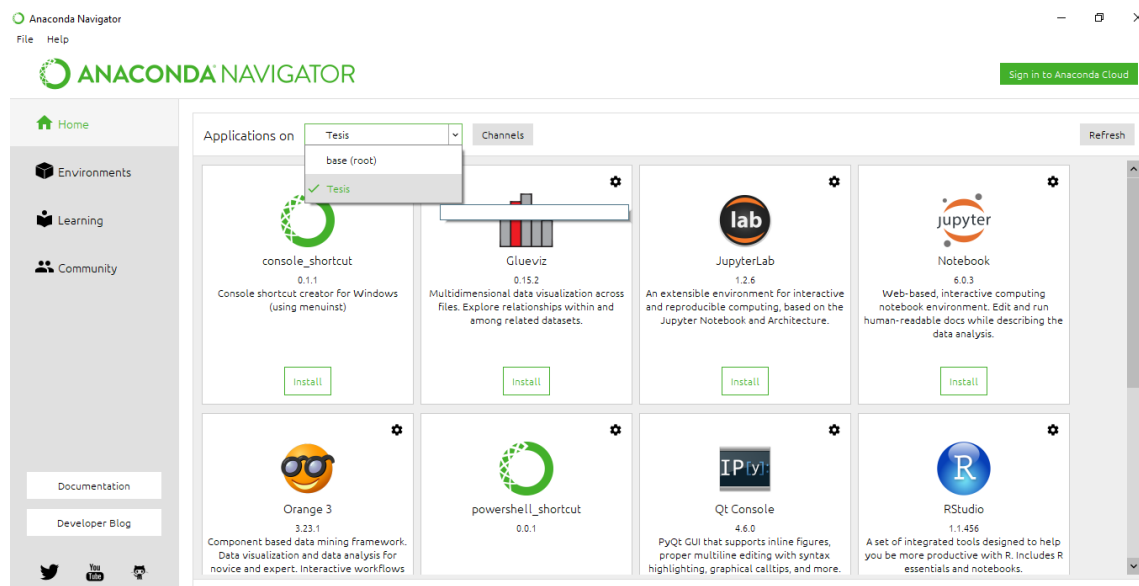


3.3 Jupyter Notebook

Después de haber creado el entorno en Anaconda Navigator, retornamos a la pestaña Home y seleccionamos el entorno *Tesis* en el cual se va a realizar la aplicación, como indica la Figura 28.

Figura 28

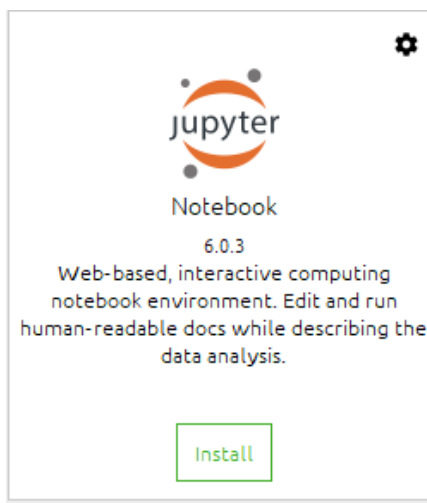
Selección del Entorno



Si es la primera vez que ejecutamos la aplicación de Jupyter Notebook, seleccionamos *Install* para su instalación, como se observa en la Figura 29.

Figura 29

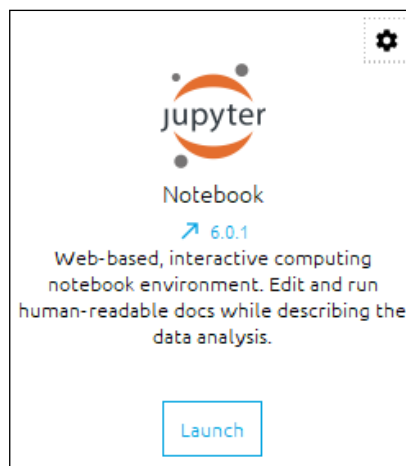
Instalación de Jupyter Notebook



Después ejecutamos la aplicación, seleccionando *Launch*, Figura 30. Todo esto en la misma Suite de Anaconda.

Figura 30

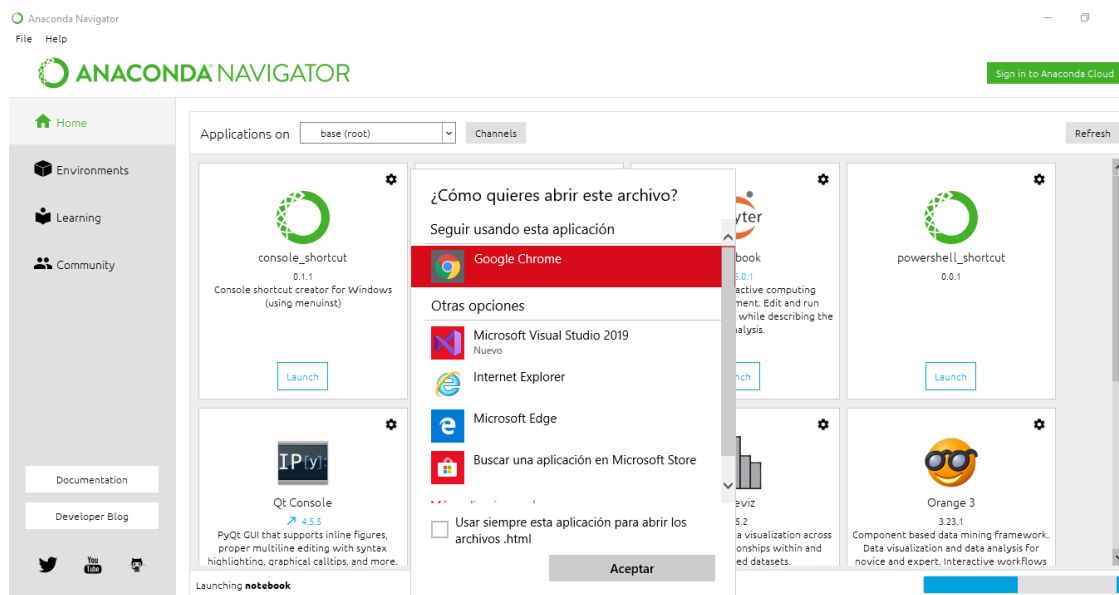
Ejecución de Jupyter Notebook



La aplicación requiere de un navegador para ejecutarse, en este caso utilizamos Google Chrome, Figura 31.

Figura 31

Elección del Navegador Web.



En el navegador se observa una ventana como de la Figura 32, en la parte superior derecha, seleccionamos *New*, y luego *Python 3*, se creará un archivo con extensión *.ipynb* el cuál contendrá el código desarrollado posteriormente.

Figura 32

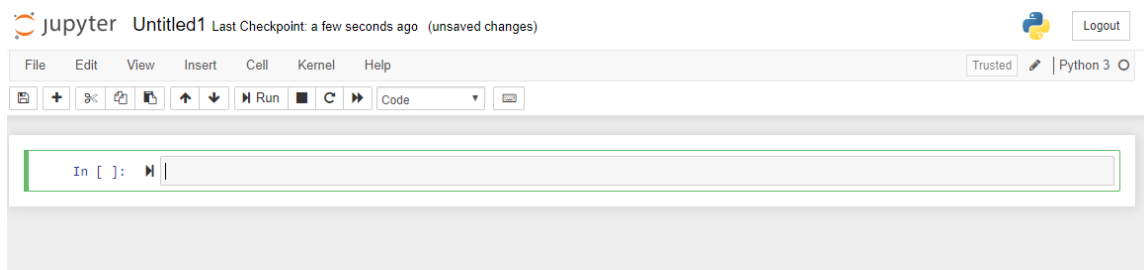
Ventana de Jupyter Notebook.



Finalmente, en la Figura 33, observamos el área en la cual se editará y ejecutará el código del proyecto.

Figura 33

Área de Edición y Ejecución de Código de Jupyter Notebook.



3.4 Instalación de librerías de Python

Las librerías son instaladas a través de Anaconda Prompt, Figura 34, se lo puede encontrar mediante el buscador de Windows.

Figura 34

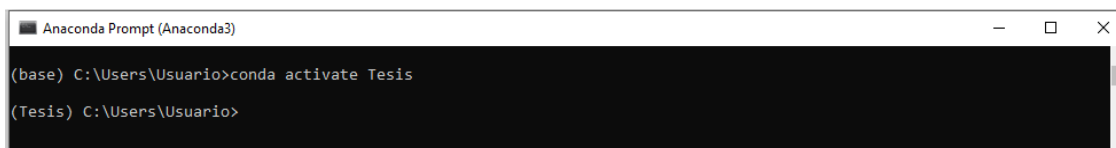
Ventana Anaconda Prompt.



Se digita el comando *conda activate Tesis* (nombre del entorno), Figura 35, de esta forma se garantiza que las librerías a instalar solo se instalarán en el entorno mencionado.

Figura 35

Activación del Entorno para Instalación de Librerías.



Para la instalación de librerías se usa el comando *conda install (nombre de librería)* o *conda install (caracteres + nombre de librería)*, para conocer la forma correcta de instalar la librería requerida, es necesario visitar la página oficial de la librería y verificar sus requisitos ya que algunas librerías dependen de otras. Por ejemplo, en la Figura 36 se instala la librería *numpy*, el comando a digitar es: *conda install numpy*.

Figura 36

Instalación de la Librería Numpy.



La Figura 37 muestra la instalación de la librería *numpy* ya terminada.

Figura 37

Librería Numpy Instalada.

```

Anaconda Prompt (Anaconda3) - conda install numpy
icc_rt                pkgs/main/win-64::icc_rt-2019.0.0-h0cc432a_1
intel-openmp          pkgs/main/win-64::intel-openmp-2020.0-166
mkl                   pkgs/main/win-64::mkl-2020.0-166
mkl-service           pkgs/main/win-64::mkl-service-2.3.0-py37hb782905_0
mkl_fft               pkgs/main/win-64::mkl_fft-1.0.15-py37h14836fe_0
mkl_random            pkgs/main/win-64::mkl_random-1.1.0-py37h675688f_0
numpy                 pkgs/main/win-64::numpy-1.18.1-py37h93ca92e_0
numpy-base           pkgs/main/win-64::numpy-base-1.18.1-py37hc3f5095_1
six                   pkgs/main/win-64::six-1.14.0-py37_0

Proceed ([y]/n)? y

Downloading and Extracting Packages
numpy-base-1.18.1      | 3.8 MB | ##### | 100%
intel-openmp-2020.0    | 1.5 MB | ##### | 100%
mkl_fft-1.0.15         | 118 KB | ##### | 100%
mkl-2020.0             | 98.9 MB | ##### | 100%
numpy-1.18.1           | 6 KB | ##### | 100%
six-1.14.0             | 27 KB | ##### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done

```

De la misma manera se procede con la instalación de las librerías matplotlib, scikit-image.

- `conda install matplotlib.`
- `conda install scikit-image.`
- `conda install threadpoolctl`

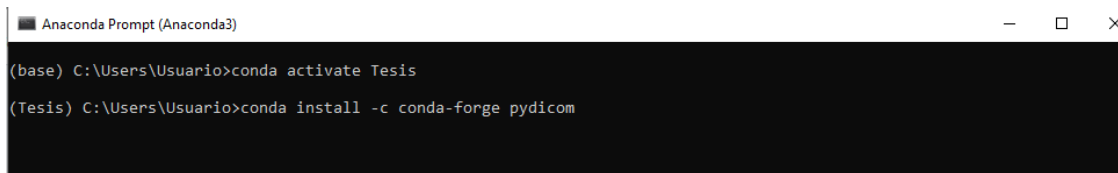
Para librerías como pydicom, opencv, scikit-learn, y keras, se repite el proceso anterior, y se digitan los siguientes comandos para la instalación:

- `conda install -c conda-forge pydicom`
- `conda install -c conda-forge opencv`
- `conda install -c intel scikit-learn`
- `conda install -c conda-forge keras`

Se utiliza como ejemplo la librería pydicom y se ingresa el código `conda install -c conda-forge pydicom` como se observa en la Figura 38.

Figura 38

Instalación de la Librería Pydicom



```

Anaconda Prompt (Anaconda3)
(base) C:\Users\Usuario>conda activate Tesis
(Tesis) C:\Users\Usuario>conda install -c conda-forge pydicom

```

La instalación de la librería tensorflow presenta dos opciones, se recomienda utilizar la segunda opción con el fin de utilizar los recursos de la gpu instalada en el computador, para esto se debe verificar que se cuente con una tarjeta gráfica. Su beneficio principal será la disminución de tiempo al momento de entrenar la red neuronal.

- `conda install -n [my-env-name] tensorflow`
- `conda install -n [my-env-name] tensorflow-gpu`

3.5 Estructura del sistema

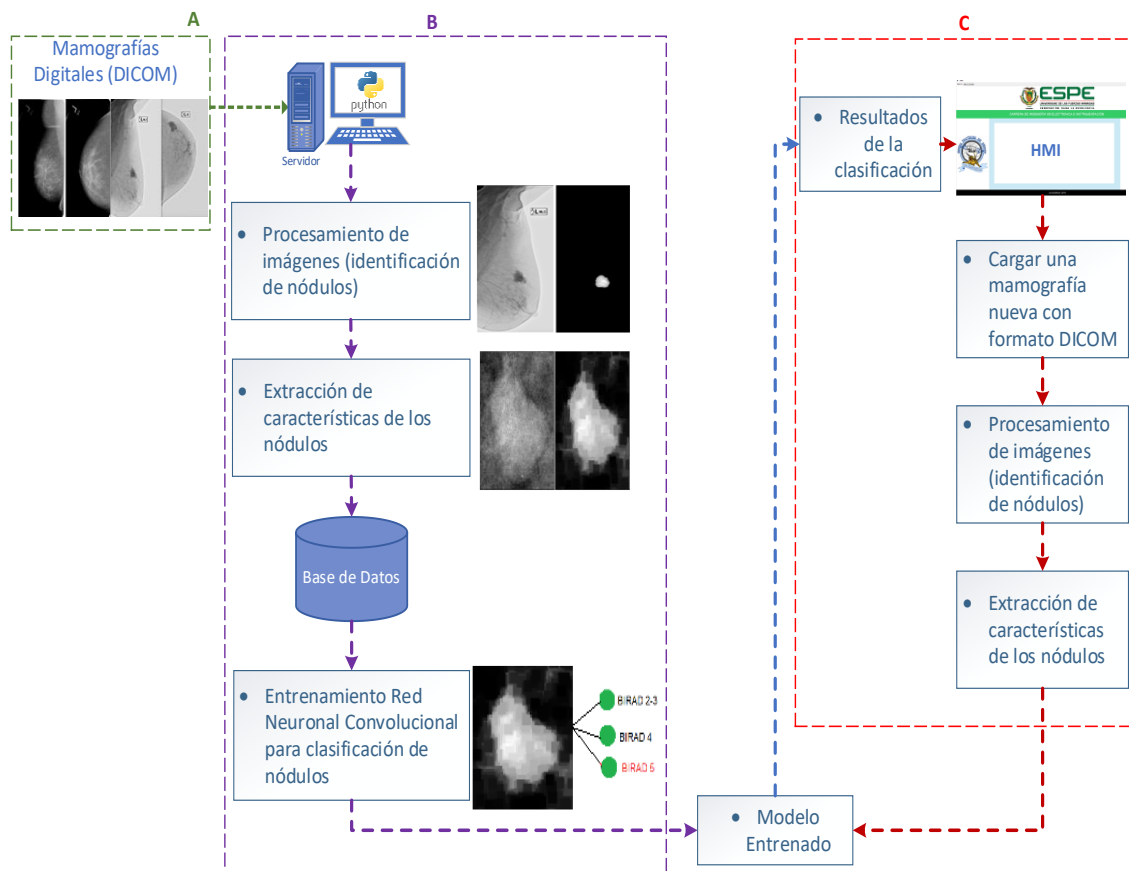
En la Figura 39, el bloque A representa la recolección de imágenes mamográficas digitales de diferentes personas con la finalidad de obtener una variedad de nódulos mamarios, las imágenes serán almacenadas en una carpeta del servidor para su posterior uso.

Para el bloque B se utiliza el editor de código Jupyter Notebook, a través de esta herramienta se realiza el procesamiento digital de las imágenes mamográficas con el fin de detectar los nódulos presentes y la extracción de las características relevantes de los mismos formando una base de datos con los diferentes tipos de nódulos. Esta base de datos debe estar balanceada, es decir, el número de imágenes es similar para cada tipo de masa. Se elabora una red neuronal convolucional (CNN), misma que es entrenada con la base de datos balanceada. Al finalizar este proceso obtendremos un modelo entrenado de la CNN con el cual se podrá reconocer y clasificar las lesiones mamarias encontradas en mamografías.

En el bloque C se desarrolla una interfaz gráfica que permite cargar una nueva imagen mamográfica en la cual se detectará y clasificará nódulos mamarios en caso de haberlos, empleando el modelo entrenado de la red CNN.

Figura 39

Estructura del Sistema

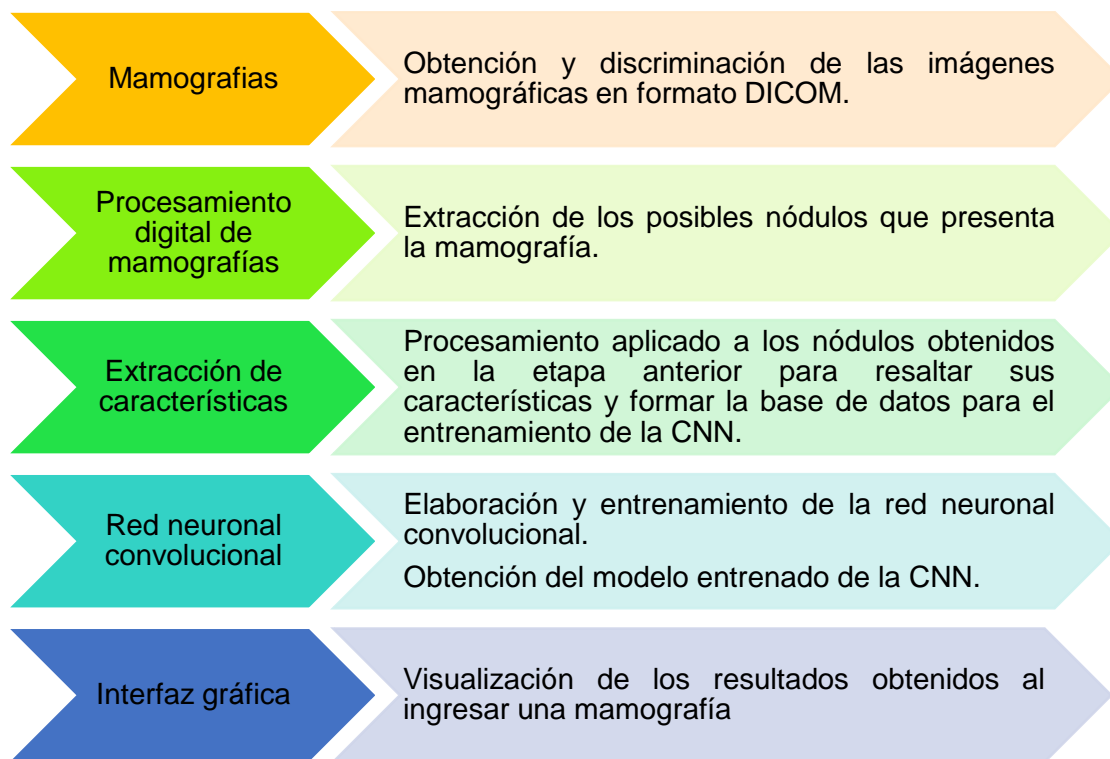


3.6 Diagrama de bloques

En la Figura 40 se observa las etapas que cumple el sistema desarrollado.

Figura 40

Diagrama de Bloques del Sistema



3.7 Mamografías

Para el desarrollo del sistema es necesario una base de datos que cuente con imágenes mamográficas en formato DICOM, en donde se encuentren todos los casos de la categoría BIRADS. Además, debe contar con estudios validados por médicos especialistas, para estar seguros que la información utilizada sea la correcta y evitar errores en el sistema. La recolección de imágenes mamográficas se realizó en:

- **CBIS-DDSM**

La página web CANCER IMAGING ARCHIVE (TCIA), contiene gran colección de imágenes médicas de cáncer de diferentes tipos. Aquí se encuentra la base de datos mamográficos CBIS-DDSM, misma que tiene casos de pacientes con nódulos mamarios y calcificaciones por separado. Al extraer la información de esta base de datos obtenemos de cada paciente lo siguiente: imagen mamográfica completa, máscara de

los nódulos encontrados y la región de interés ROI de cada nódulo, todos estos archivos en formato DICOM. (Sawyer Lee, Giménez, Hoogi, & Rubin, 2016)

- **Centro Médico Láser Diagnóstico**

Es un centro radiológico de la ciudad de Latacunga que cuenta con médicos especialistas en imagenología, realizan estudios de rayos x, mamografías, entre otros.

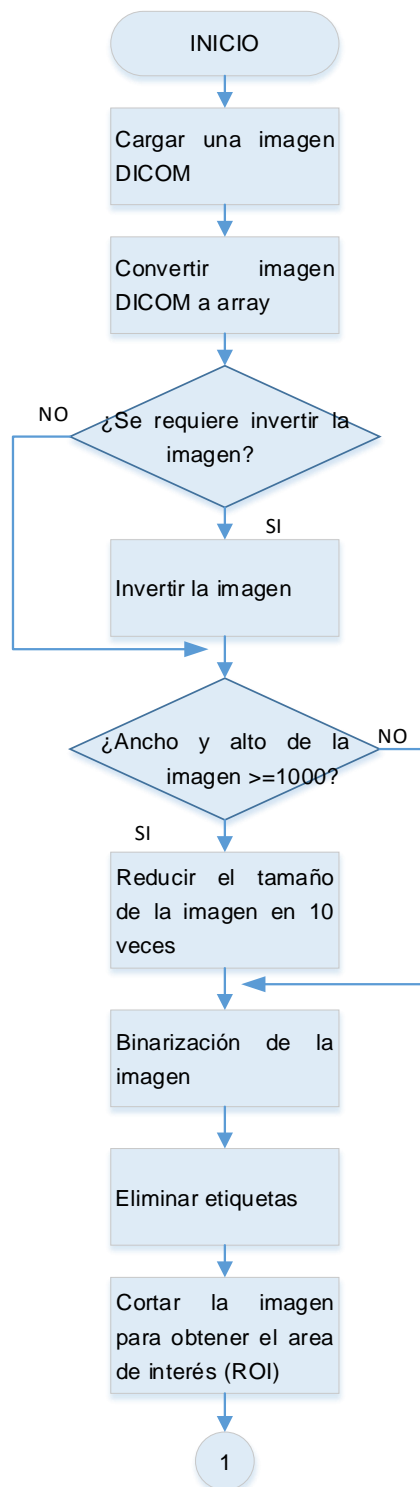
Las imágenes proporcionadas por el centro médico Láser, incluyen 48 casos de pacientes con sus respectivos informes y el diagnóstico proporcionado por el médico especialista. Cada paciente presenta cuatro imágenes en formato DICOM de las vistas cráneo caudales derecho (CCR) e izquierdo (CCL) y medio laterales oblicuos de igual forma derecho (MLOR) e izquierdo (MLOL).

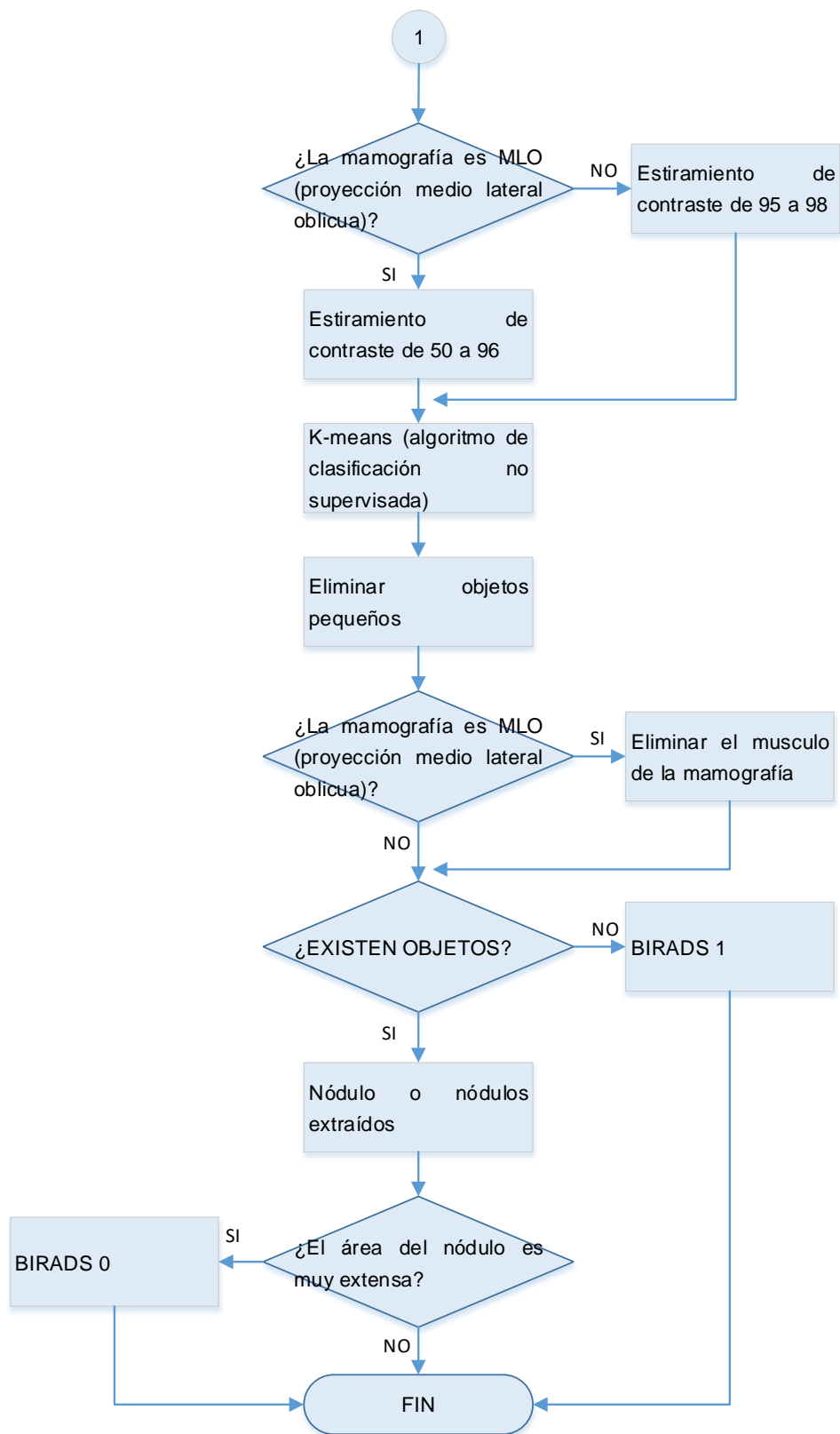
3.8 Procesamiento Digital de Mamografías

Para realizar la extracción del nódulo o los nódulos mamarios que presente una mamografía, se utilizó el procedimiento representado por el diagrama de flujo de la Figura 41.

Figura 41

Diagrama de Flujo del Procesamiento Digital de Mamografías





A continuación, se explica el código desarrollado en Jupyter Notebook para la implementación de esta etapa del sistema.

Una vez creado el entorno en Anaconda e instalar las librerías necesarias, el primer paso es importar estas librerías en Jupyter Notebook.

```
import matplotlib.pyplot as plt
import pydicom
import numpy as np
import cv2
import os, glob

import skimage.morphology as morp
from skimage import exposure
from skimage.filters.rank import median
from skimage.transform import resize
from skimage.measure import label, regionprops
from skimage.morphology import remove_small_objects
from skimage.morphology import disk
from skimage.morphology import erosion, dilation, opening, closing
from skimage.filters import rank
from scipy import stats
from sklearn.cluster import KMeans
```

Para la lectura de una mamografía se utilizó el comando

`pydicom.dcmread('nombre del archivo')`, con el fin de evitar colocar la dirección de las imágenes, estas se encuentran en la carpeta con el archivo de Jupyter Notebook. Es necesario convertir el archivo DICOM en una matriz, de esta forma se puede utilizar la librería `numpy` y realizar operaciones matriciales a la imagen, `DICOM.pixel_array`.

Finalmente visualizamos la matriz con el comando `plt.imshow(matriz)`, la Figura 42 muestra 4 imágenes con el fin de contrastar las imágenes obtenidas de las dos bases de datos.

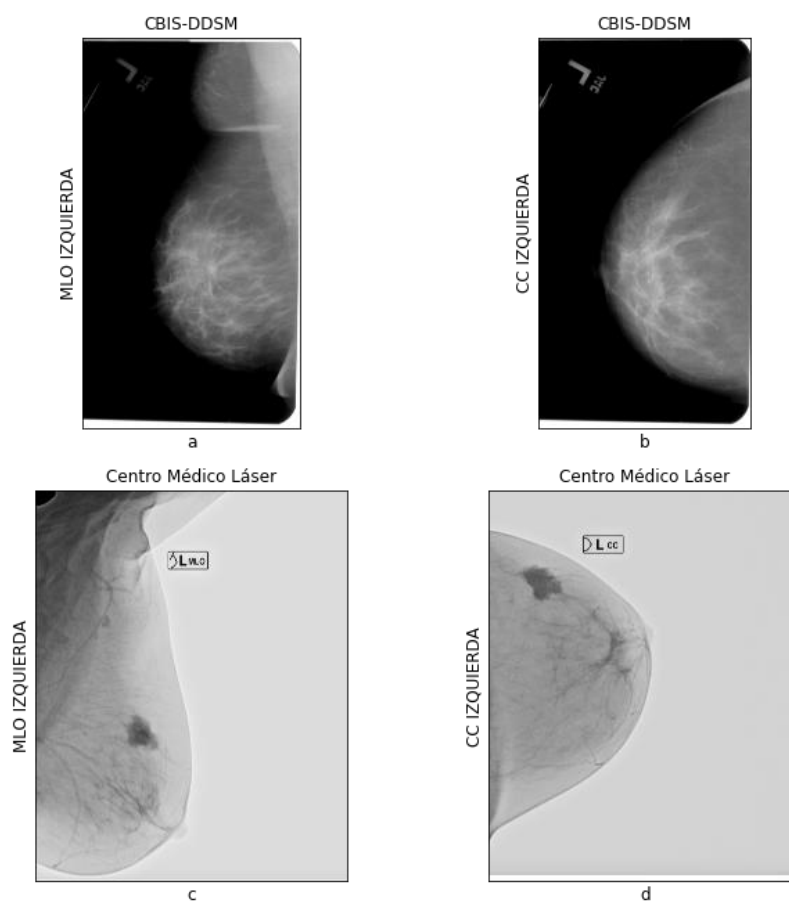

```

ds1 = pydicom.dcmread("P51CD3.dcm")
ds2=ds1.pixel_array
plt.imshow(ds2, cmap=plt.cm.gray)

```

Figura 42

Imágenes Mamográficas (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC

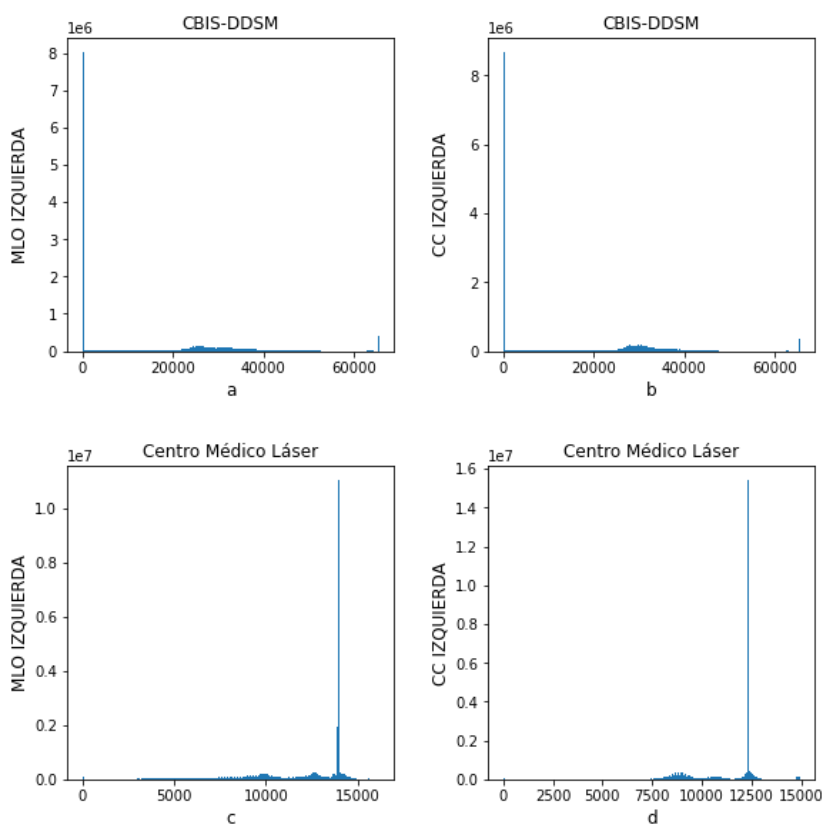


Todas las imágenes recolectadas son de tipo uint 16 (enteros almacenados en 16 bits) poseen 2^{16} tonalidades, es decir los valores de sus elementos están en el rango de 0 a 65535. Para verificar esto ejecutamos el comando `ds2.dtype`, obteniendo como respuesta: `dtype('uint16')`.

Observando las imágenes presentadas en la Figura 42, se deduce que las mamografías tienen diferente presentación dependiendo del lugar en el cual se hizo la recolección. Para corregir este error es necesario analizar el histograma de las imágenes presentado en la Figura 43, el eje x presenta el valor de los píxeles, mientras que el eje y muestra la cantidad de píxeles. Es importante recalcar que los valores de los píxeles de la base de datos CBIS-DDSM ocupan todo el rango de trabajo disponible (0-65535), mientras que los píxeles de la base de datos del centro médico Láser ocupan aproximadamente una cuarta parte de este rango.

Figura 43

Histograma de la Mamografía: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC.



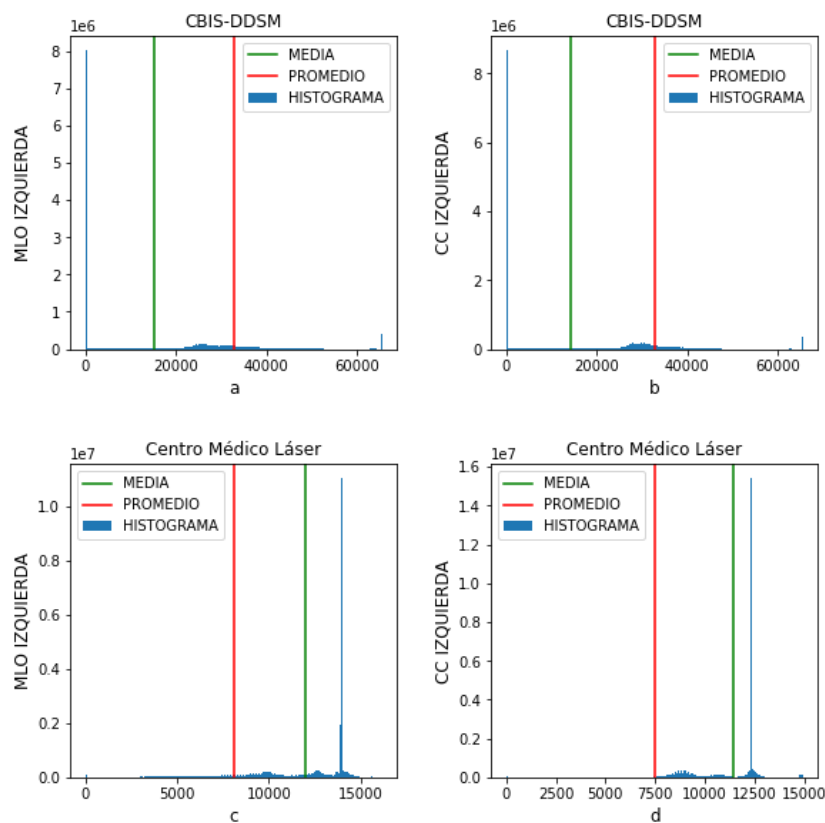
En función de los histogramas presentados, se concluye que la media aritmética de las imágenes obtenidas de la base de datos CBIS-DDSM estará próxima al cero,

mientras que la media aritmética de las mamografías obtenidas en el Centro Médico Láser estará próxima al valor del píxel más alto, esta característica se da por el número de píxeles que presenta el fondo de la mamografía. En el primer caso tenemos un fondo negro en el cual predominan píxeles con valores de cero o cercanos, mientras que en el segundo caso el fondo tiene valores de píxeles altos.

En conclusión, la media aritmética estará cerca de los extremos, por lo cual, se requiere comparar esta característica con el punto medio del histograma, este valor obtenemos del promedio entre el píxel con valor más alto y el píxel con menor valor, en la Figura 44 se muestra la relación de los valores mencionados.

Figura 44

Histograma, Media Aritmética de la Imagen y Promedio entre Píxel Mayor y Menor: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



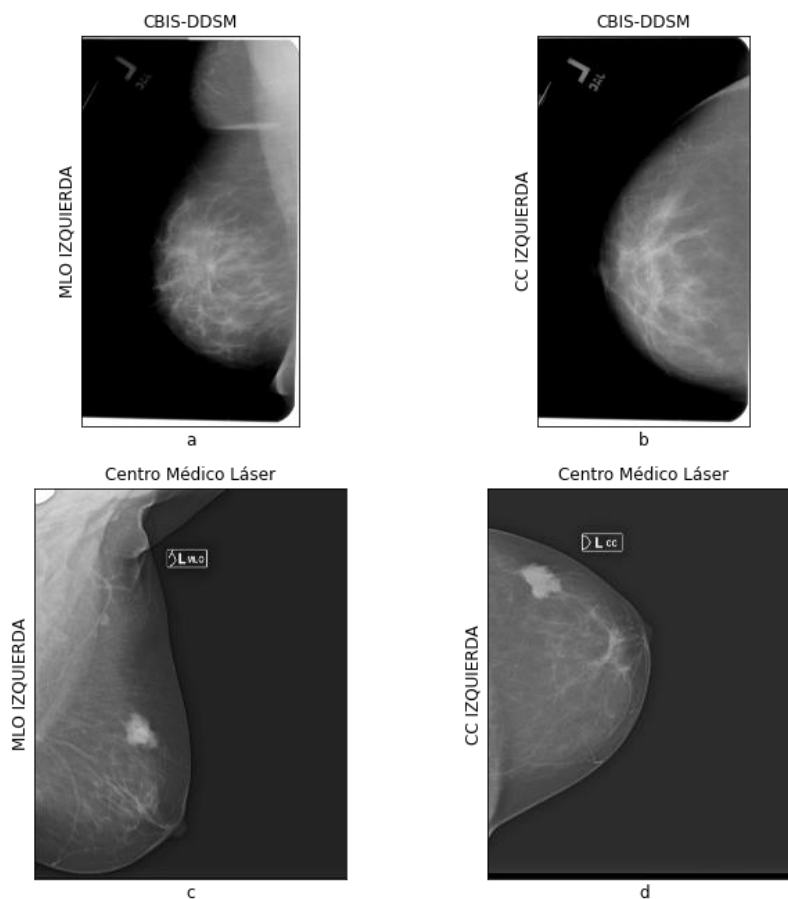
El código implementado en Jupyter Notebook para estandarizar la presentación de las mamografías es el siguiente:

```
if ((np.max(ds2)+np.min(ds2))/2)<(np.mean(ds2)):
    pixel_range = np.amin(ds2) + np.amax(ds2)
    invertir = pixel_range - ds2
else:
    invertir=ds2
```

Los resultados obtenidos se muestran en la Figura 45.

Figura 45

Presentación Estandarizada de las Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



Al ser matrices de tipo uint16, poseen valores que van de 0 a 65535, además sus dimensiones superan los 2500 x 2500, por ende, para facilitar las operaciones

matriciales posteriores, es necesario reducir el tamaño de la matriz en diez veces y cambiar su tipo a uint8 para obtener valores de 0 a 255.

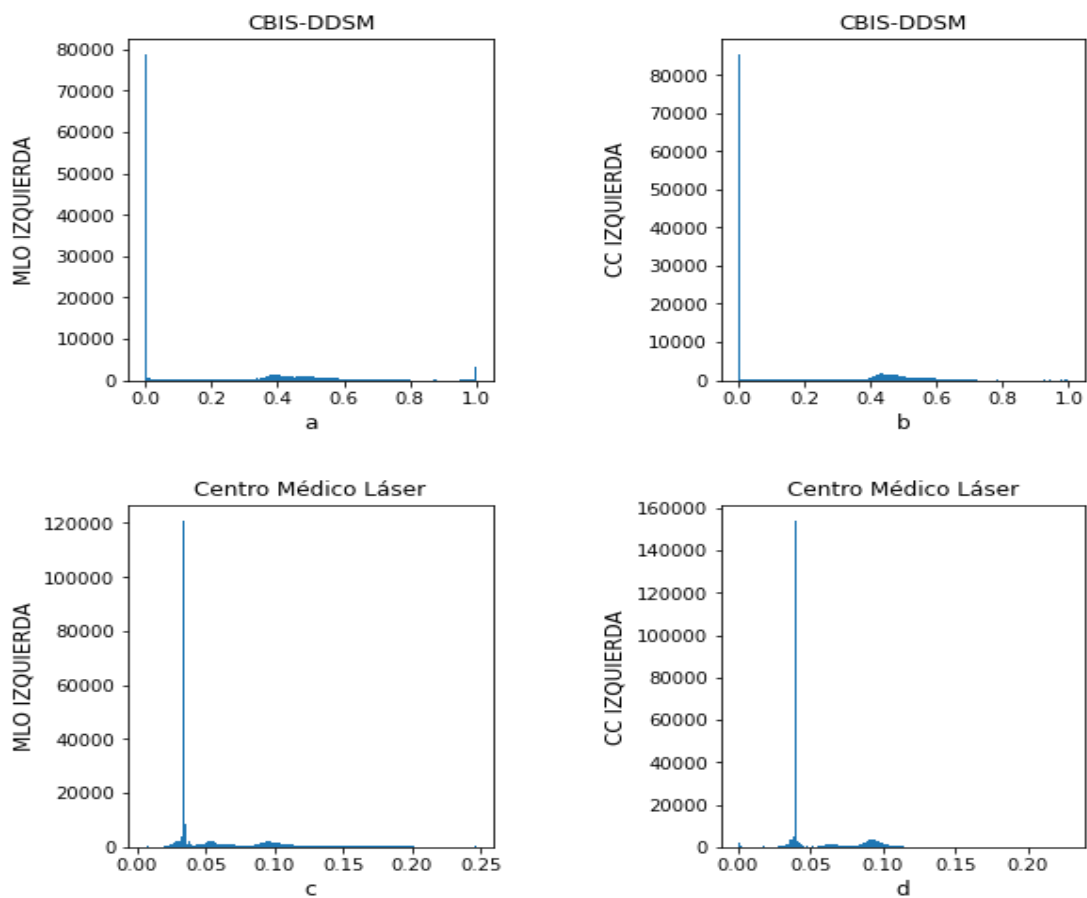
Las dimensiones de la matriz se obtienen con el comando *matriz.shape*, el cual retorna un array de dos posiciones con el valor del alto y ancho, posteriormente se realiza una sentencia condicional y se reducirá el tamaño de la matriz cuando supere en ancho y alto a 999 utilizando el comando *resize(matriz, (nuevo_tamaño_alto, nuevo_tamaño_ancho))*

```
tamaño=ds1.pixel_array.shape
if tamaño[0]>=1000 and tamaño[1]>=1000:
    alto=int(tamaño[0]/10)
    ancho=int(tamaño[1]/10)
else:
    alto=tamaño[0]
    ancho=tamaño[1]
ds3 = resize(invertir, (alto, ancho), anti_aliasing=True)
```

A través del histograma de las imágenes, verificamos que existe una gran disminución de píxeles y un cambio en el valor de estos, ahora están en el rango de 0 y 1, Figura 46.

Figura 46

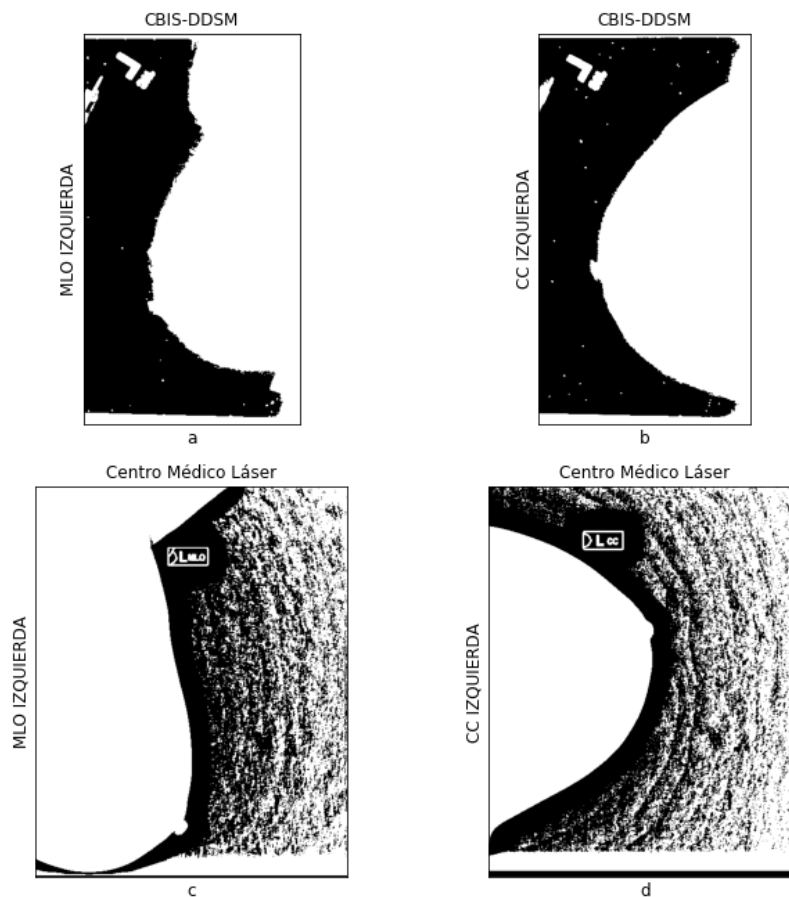
Histograma de Imágenes Mamográficas Re-dimensionadas: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



Al realizar el proceso de binarización de la imagen, se evidencia la presencia de ruido, Figura 47, el cual debe ser corregido para evitar errores en las demás etapas.

Figura 47

Ruido Presente en las Imágenes: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC
(c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



Para suprimir el ruido presente, recurrimos al filtro de mediana cuya función principal es eliminar el ruido presente en la imagen, además de mejorar bordes, nitidez y contraste.

$$f_{med} = \text{median}(ds_3)$$

En la Figura 48 se grafica las imágenes y sus respectivos histogramas se presentan en la Figura 49, en ambos casos después de utilizar el filtro de mediana.

Figura 48

Imágenes Mamográficas Después de Aplicar el Filtro de Mediana: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC

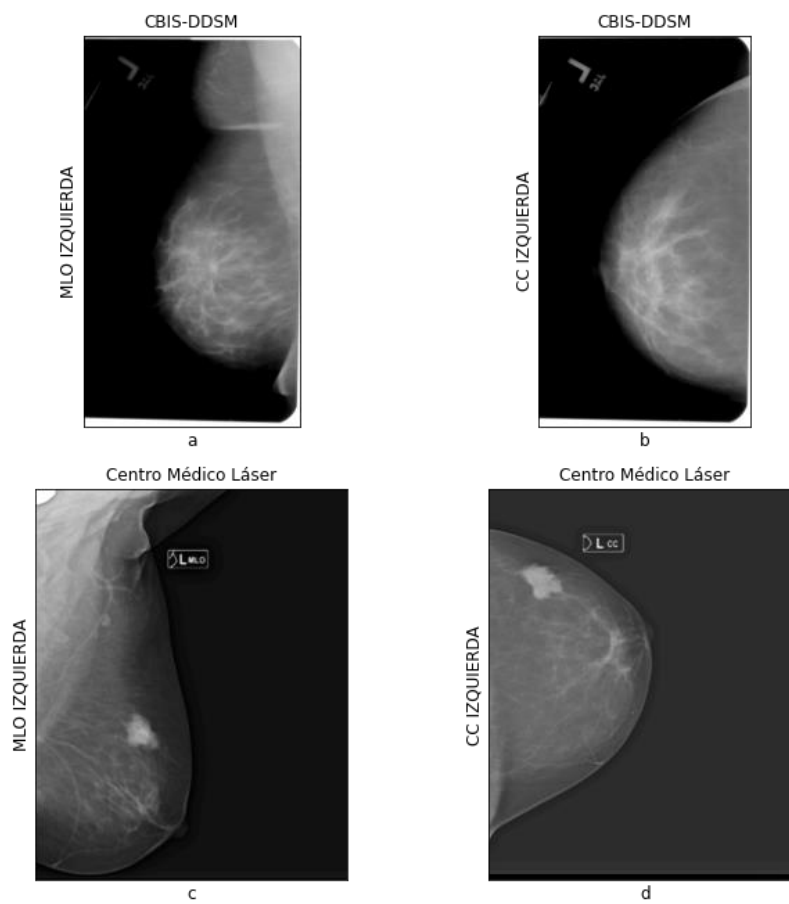
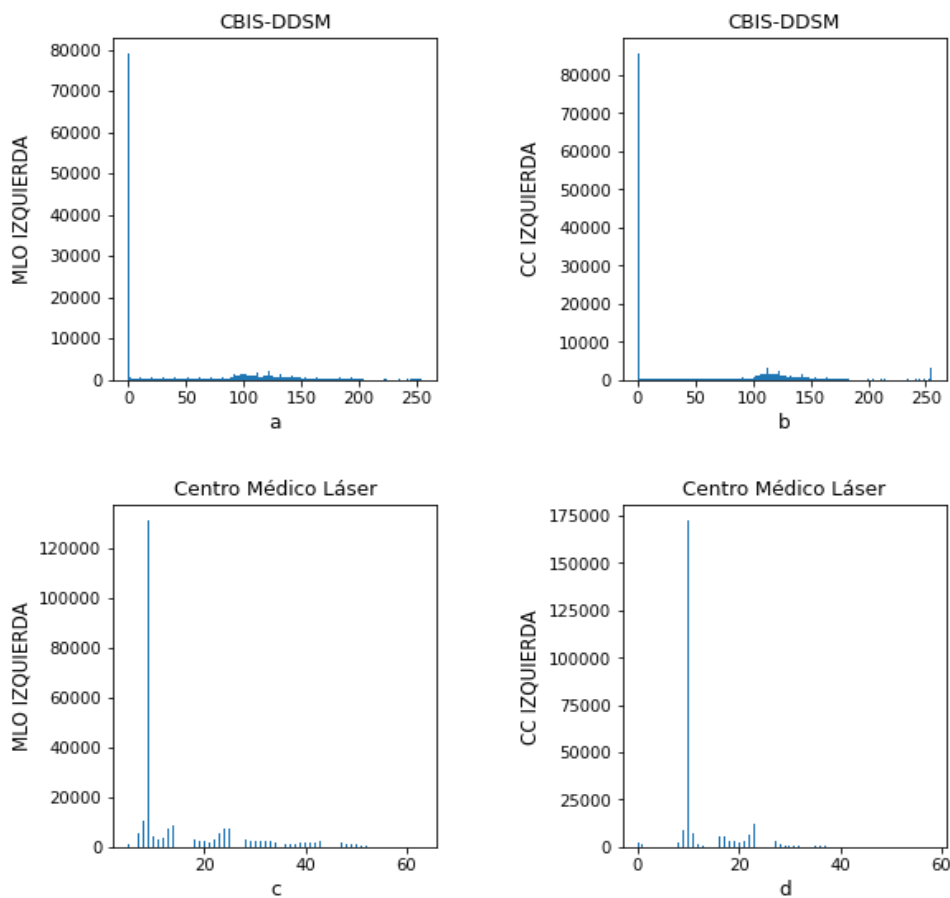


Figura 49

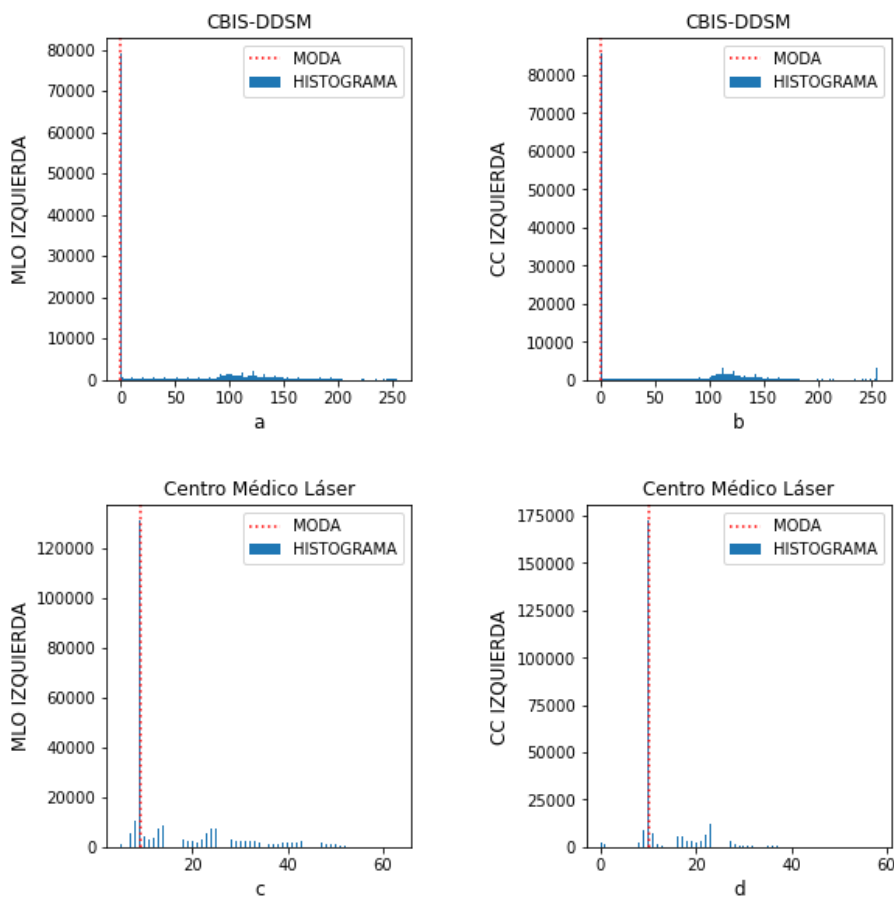
Histograma de Imágenes Mamográficas Después de Aplicar el Filtro de Mediana: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



Con la imagen sin ruido procedemos a realizar el proceso de binarización, para esto se calcula el valor de la moda, es decir buscamos el valor de píxel que más veces se repite en la imagen. Los valores que se encuentren por debajo de la moda serán 0 y los que se encuentren sobre este serán 1. En la Figura 50 se muestra el histograma de las imágenes y el valor de la moda.

Figura 50

Histograma de Imágenes Mamográficas con su Respectivo Valor de Moda: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



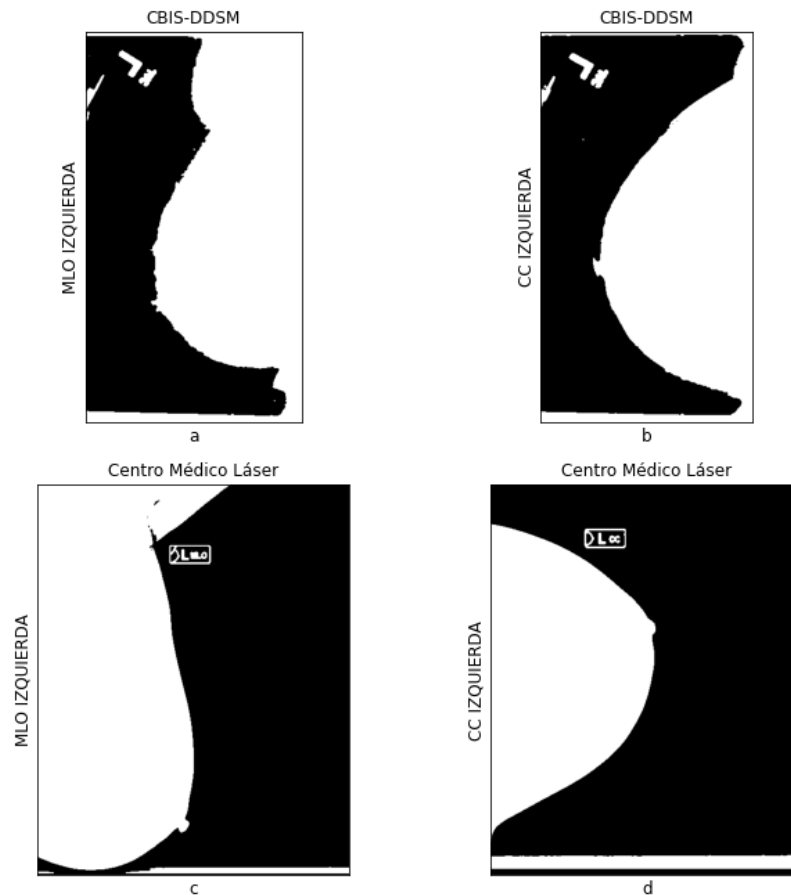
Las mamografías binarizadas se presentan en la Figura 51 y el código utilizado para calcular la moda y realizar la binarización es:

```
a=fmed.ravel()
```

```
ret,thresh = cv2.threshold(fmed,stats.mode(a)[0],1,0)
```

Figura 51

Mamografías Binarizadas: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC

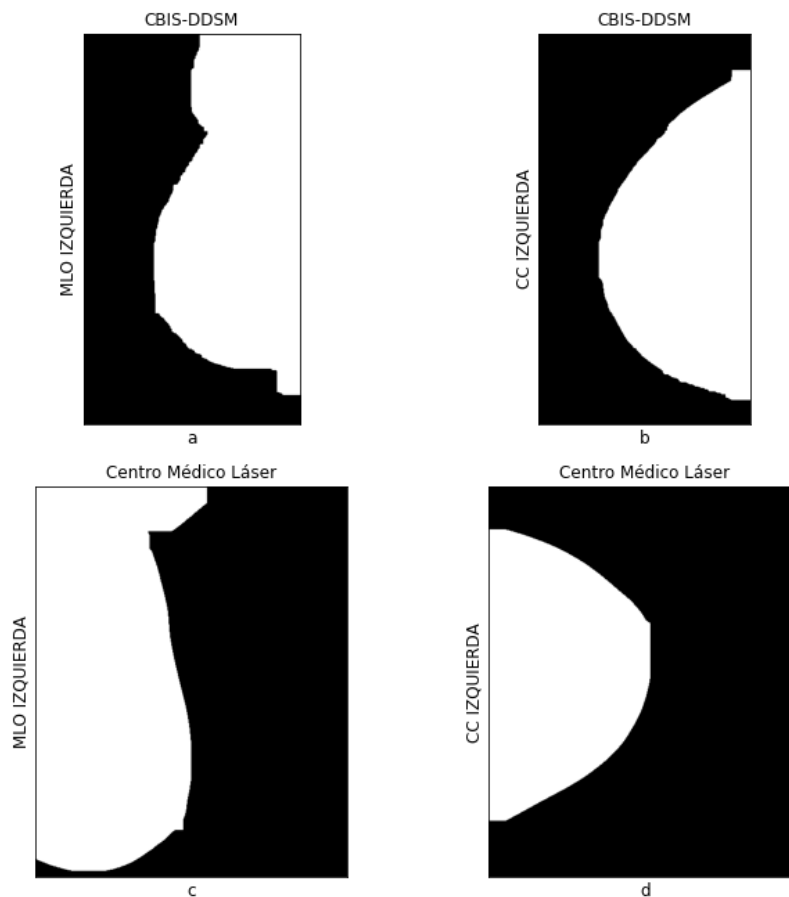


Se debe eliminar la información innecesaria, en este caso las etiquetas y objetos adicionales. Las operaciones morfológicas son las herramientas que aplicamos para este proceso. El cierre se utiliza para rellenar discontinuidades o agujeros presentes en el objeto, mientras que la apertura se aplica para eliminar objetos que sobresalgan o no pertenezcan al objeto principal. En la Figura 52 se muestran los resultados, y el código implementado es el siguiente:

```
cierre = morp.binary_closing(thresh,np.ones([3,3]))
apertura = morp.binary_opening(cierre,np.ones([51,51]))
```

Figura 52

Mamografías Sin Etiquetas: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



Se realiza un corte a la imagen con el fin de obtener el área de interés y volvemos los píxeles a su valor original, el resultado se observa en la Figura 53, el código para este propósito es:

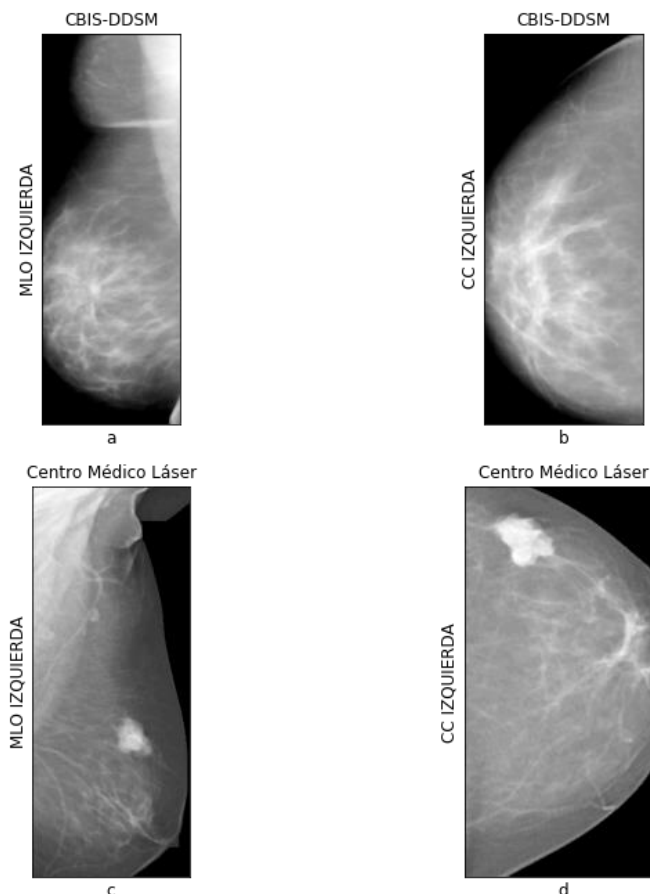
```

sinetiqueta = (fmed * apertura)
pts = np.argwhere(apertura>0)
y1,x1 = pts.min(axis=0)
y2,x2 = pts.max(axis=0)
cortada = sinetiqueta[y1+20:y2-20, x1+20:x2-20]

```

Figura 53

Región de Interés de la Mamografía: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



En este punto, se aplica la técnica de estiramiento de contraste que es un método alternativo a la ecualización del histograma. Es necesario conocer el tipo de vista de la mamografía ya que dependiendo de esto se otorga un rango para realizar la técnica mencionada. Al trabajar con formato DICOM se puede acceder a los metadatos, información adicional que contiene la mamografía en la cual se describe el nombre del paciente, tipo de examen, tipo de vista, etc. Para acceder a estos datos, se digita el comando:

```
print(ds1)
```

Un resumen de los datos presentes en la imagen se presenta en la Figura 54:

Figura 54*Resumen de los Metadatos de la Mamografía DICOM*

```

(0008, 0005) Specific Character Set          CS: 'ISO_IR 100'
(0008, 0016) SOP Class UID                  UI: Secondary
Capture Image Storage
(0008, 0020) Study Date                     DA: '20160720'
(0008, 0023) Content Date                   DA: '20160414'
(0008, 103e) Series Description             LO: 'full mammogram
images'
(0010, 0010) Patient's Name                 PN: 'Mass-
Training_P_00004_LEFT_MLO'
(0020, 0020) Patient Orientation            CS: 'MLO'
(0020, 0060) Laterality                     CS: 'L'

```

El código para aplicar el estiramiento de contraste en función de la vista de la mamografía se presenta a continuación y sus resultados se visualizan en la Figura 55:

```

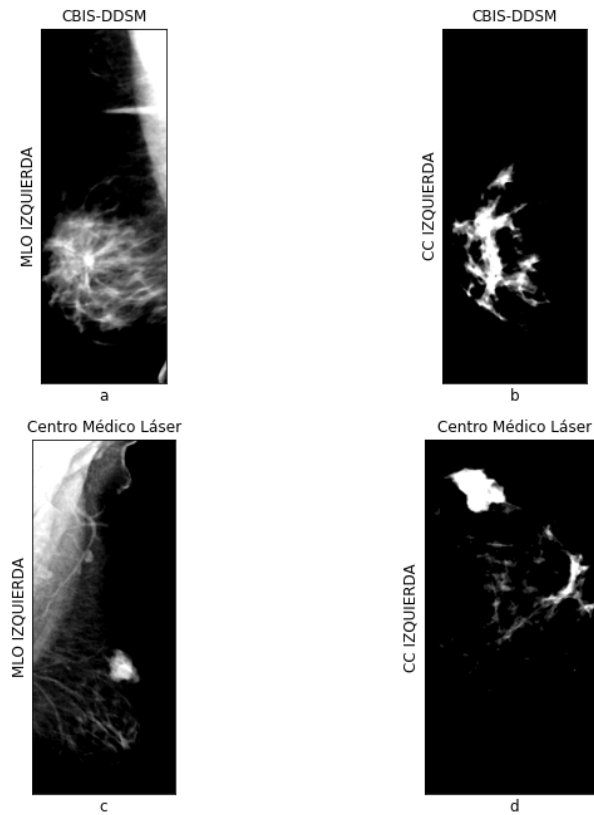
if ds1.PatientOrientation[1][0]!='F' and
ds1.PatientOrientation[0]!='M':
    #CC
    print('CC')
    p22, p988 = np.percentile(cortada, (95, 98))
    I2 = (exposure.rescale_intensity(cortada, in_range=(p22,
p988)))
else:
    #MLO
    print('MLO')
    p22, p988 = np.percentile(cortada, (50, 96))

I2=(exposure.rescale_intensity(cortada,in_range=(np.median(estadi
stica), p988)))

```

Figura 55

Estiramiento de Contraste Aplicado a Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



Ahora se utiliza el algoritmo de clasificación k-means, con el cual agruparemos los píxeles en tres grupos, de esta forma se eliminará los píxeles que no se asemejen al nódulo, el código aplicado se muestra a continuación y sus resultados en la Figura 56.

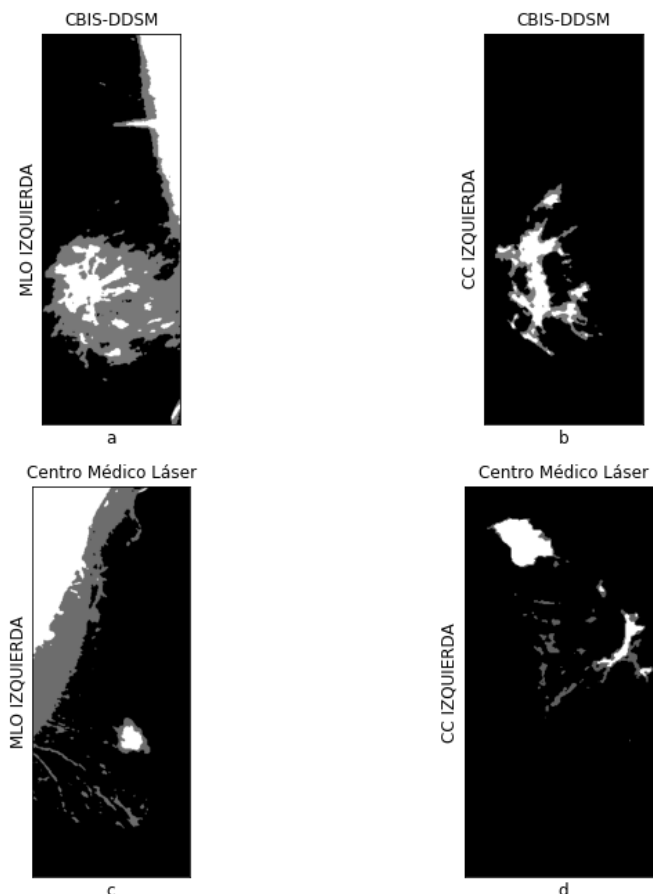
```

Z = I2.reshape((-1, 1))
k_means = KMeans(n_clusters= 3)
k_means.fit(Z)
centroides = k_means.cluster_centers_
etiquetas = k_means.labels_
I2_compressed = np.choose(etiquetas, centroides)
I2_compressed.shape = I2.shape

```

Figura 56

K-means con n=3 Aplicado a Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



Después de haber aplicado el algoritmo k-means se debe segmentar la imagen, separaremos el grupo que tenga los pixeles con mayor valor ya que los nódulos se encuentran en este grupo, y eliminamos objetos pequeños que pueden ser parte de otras lesiones mamarias como calcificaciones. El código implementado para este propósito se muestra a continuación y su resultado en la Figura 57.

```

musculo_mas=(np.where(I2_compressed < np.max(centroides), 0, 1))
musculo_mask = morp.binary_opening(musculo_mas,np.ones([3,3]))
musculo_mask_label= label(musculo_mask)
musculo_mask_delete = remove_small_objects(musculo_mask_label,
160)

```



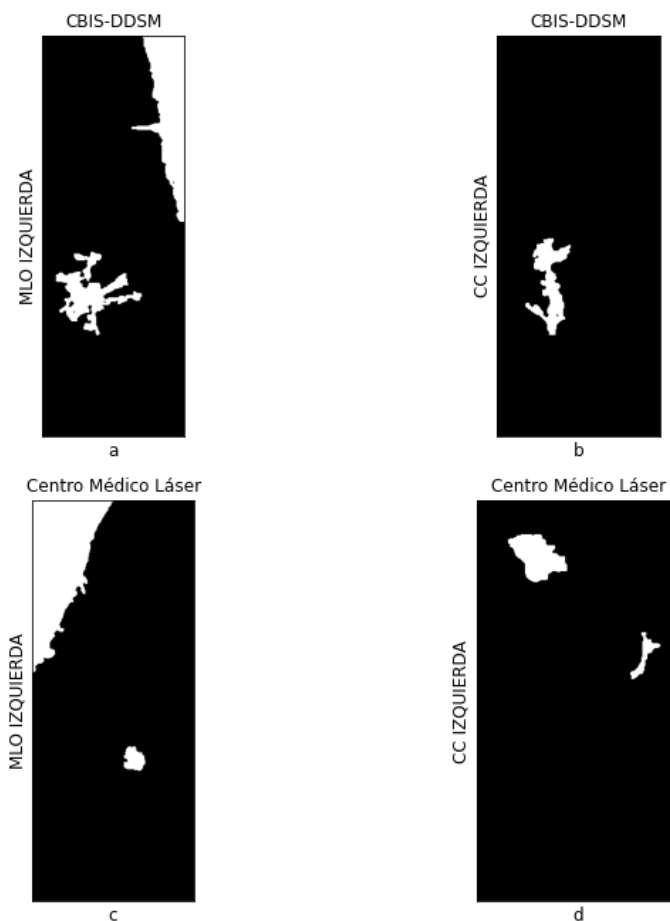
```

musculo_mask_mn=(np.where(musculo_mask_delete
                             ==0,
                             musculo_mask_delete, 1))

```

Figura 57

Extracción del Conjunto de Píxeles con Valor más Alto en Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



Revisando la Figura 57, observamos que en las mamografías de vista MLO, es necesario eliminar el músculo de la mama y de esta forma obtener el nódulo o nódulos presentes. Para esto se debe extraer las propiedades de todos los objetos presentes en la imagen y eliminar el objeto que más cerca esté del vértice superior a través del comando de *regionprops*, se utiliza el siguiente código y sus resultados se grafican en la Figura 58.

```

if ds1.PatientOrientation[1][0]!='F' and
ds1.PatientOrientation[0]!='M':
    musculo_s =musculo_mask_mn
plt.imshow((musculo_s), cmap=plt.cm.gray)
else:
    label_img = label(musculo_mask_mn)
    regions = regionprops(label_img)

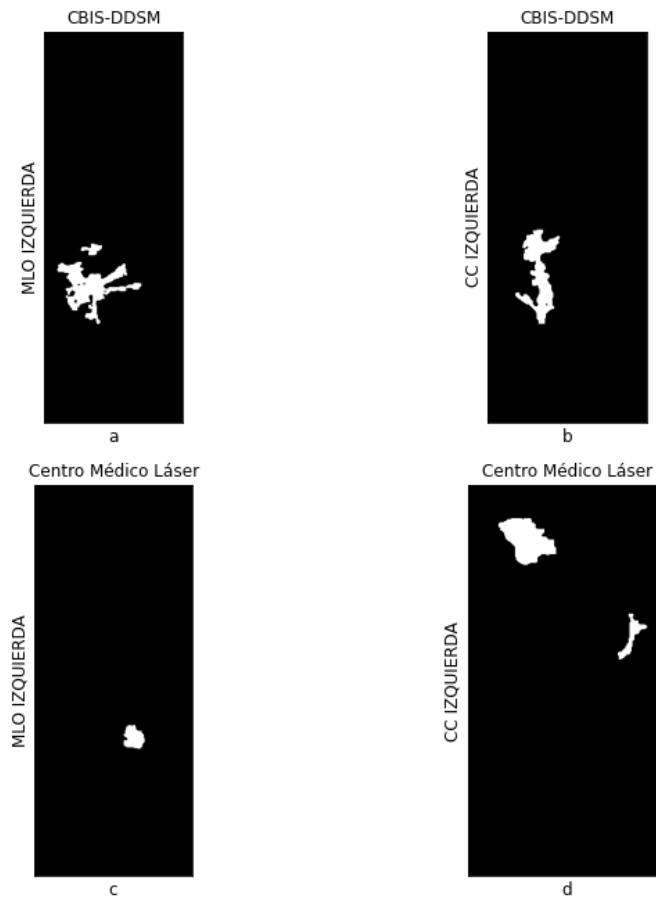
fig, ax = plt.subplots()
ax.imshow(musculo_mask_mn, cmap=plt.cm.gray)
good_labels=[]
lado=cortada.sum(axis=0)
if (lado[0])>(lado[len(lado)-1]):
    pr=(0,0)
    print('izquierda')
else:
    pr=(0,len(lado)-1)
    print('derecha')
for props in regions:
    y0, x0 = props.centroid
    dst = distance.euclidean(pr,props.centroid)
    if dst>216:
        good_labels.append(props.label)
mask
=
np.ndarray([musculo_mask_mn.shape[0],musculo_mask_mn.shape[1]],dt
ype=np.int8)
mask[:] = 0

for N in good_labels:
    mask = mask + np.where(label_img==N,1,0)
nodulo=mask
plt.imshow((nodulo), cmap=plt.cm.gray)

```

Figura 58

Exclusión del Músculo de las Mamografías de Vista MLO: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC

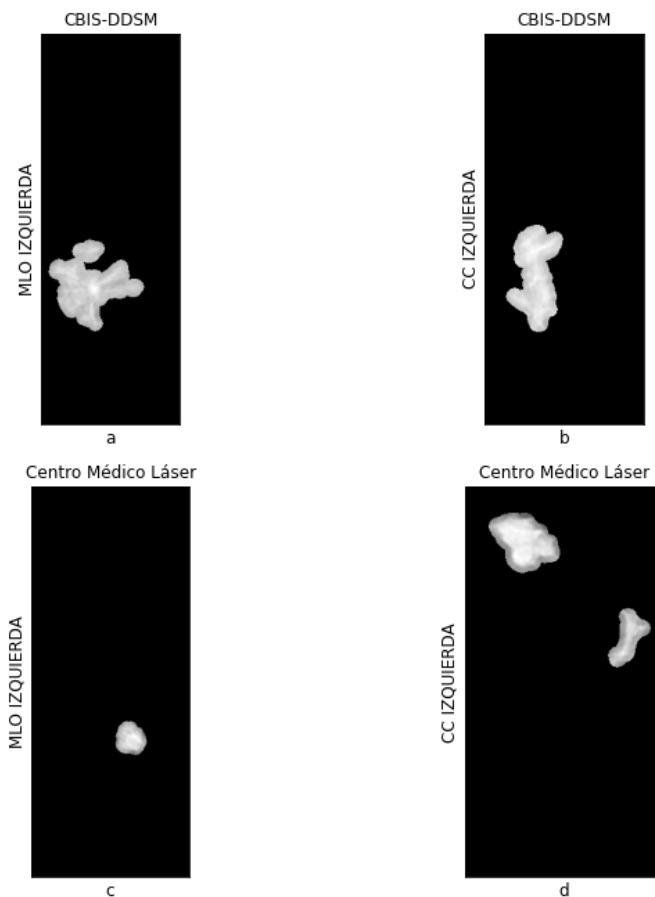


Para finalizar este proceso, dilatamos la máscara del nódulo y regresamos los píxeles a sus valores originales con el siguiente código. Los resultados se grafican en la Figura 59.

```
selem=disk(7)
dilation_nodulo = dilation(nodulo,selem)
nod_orig= dilation_nodulo*cortada
plt.imshow(dilation_nodulo, cmap=plt.cm.gray)
```

Figura 59

Nódulos Extraídos en Mamografías: (a) CBIS-DDSM Vista MLO (b) CBIS-DDSM Vista CC (c) Centro Médico Láser Vista MLO (d) Centro Médico Láser Vista CC



3.9 Extracción de Características

Esta etapa tiene por objetivo disminuir la cantidad de características presentes en los nódulos y resaltar aquellas que generalicen su categoría, siendo importante al momento de entrenar la red neuronal para evitar el sobreajuste del modelo, maximizar el alcance de la CNN y obtener una buena predicción con datos nuevos.

Las características representativas de las masas mamarias son: forma, márgenes, densidad, detalladas en la siguiente Tabla 1: (Obenauer, Hermann, & Grabbe, 2005).

Tabla 1*Características de los Nódulos Mamarios*

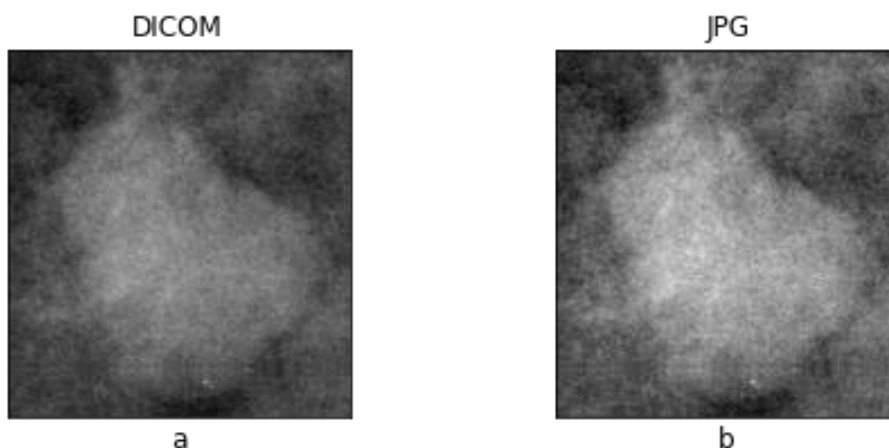
| Feature | Categoría ≤3 | Categoría 4 | Categoría 5 |
|----------------|-------------------------|------------------------|------------------------|
| Shape | | | |
| Round | x | | |
| Oval | x | | |
| Lobular | | x | |
| Irregular | | x | x |
| Margin | | | |
| Circumscribed | x | | |
| Microlobulated | | x | |
| Obscured | | x | |
| Indistinct | | x | x |
| Spiculated | | | x |
| Density | | | |
| High | | | x |
| Equal | x | x | |
| Low | x | | |
| Fat-containing | x | | |

Nota. Tomado de *Applications and literature review of the BI-RADS classification*, por Obenauer et al., 2005, PubMed. (Obenauer, 2005)

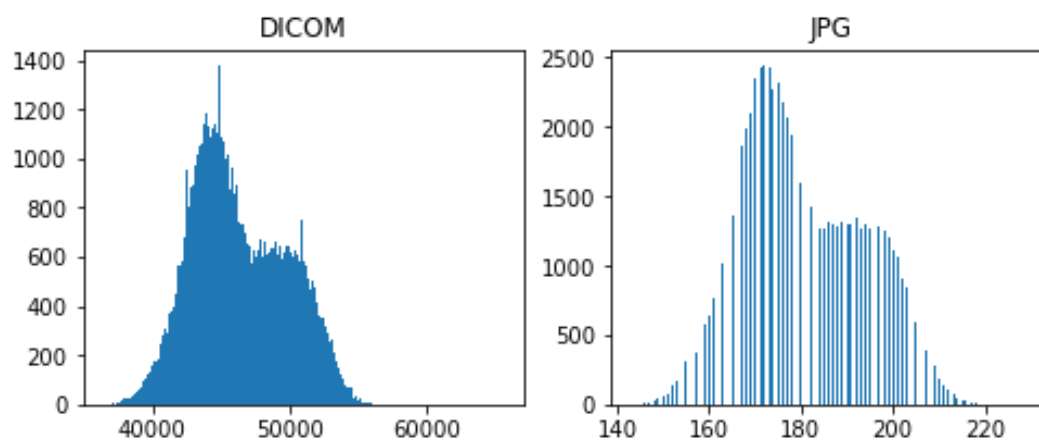
Al tener imágenes en formato DICOM, fue necesario convertirlos a formato JPG, con el fin de disminuir la cantidad de píxeles y en consecuencia de información, en la Figura 60 se presenta la imagen del nódulo en los dos formatos, mientras que en la Figura 61 se grafican sus histogramas.

Figura 60

Nódulo tipo Birads 3 de la Base de datos CBIS-DDSM, Formato: (a) DICOM (b) JPG

**Figura 61**

Histograma de Nódulo Tipo Birads 3 de la Base de Datos CBIS-DDSM, Formato: (a) DICOM (b) JPG



La información innecesaria que contienen los nódulos debe ser eliminada, en nuestro caso el tejido mamario presente. Se resalta características como la forma o márgenes de la imagen que son patrones generales que presentan los nódulos dependiendo su categoría, de esta forma se distingue la masa. Para esto se aplica herramientas como: ecualización del histograma, operación morfológica de apertura y erosión, en la Figura 62 se muestra el resultado.

```

img = cv.imread('./prueba/imagen_3_1.jpg',0)
img1 = cv.equalizeHist(img)
kernel = np.ones((21,21),np.uint8)
op = cv.morphologyEx(img1, cv.MORPH_OPEN, kernel)
img2 = cv.erode(op,kernel,iterations = 1)
plt.imshow(img2,cmap='gray')

```

Figura 62

Nódulo tipo Birads 3 de la Base de Datos CBIS-DDSM: (a) Formato JPG (b) Nódulo con Proceso de Extracción de características.



3.10 Red Neuronal Convolutacional (CNN)

Después de haber realizado la detección y extracción de características de los nódulos mamarios presentes en una mamografía se realiza la clasificación de las masas en función del estándar BIRADS, para lo cual se implementa una red neuronal convolutacional que consta de las siguientes capas:

Las capas convolucionales poseen filtros y una matriz (kernel [mxm]) cuya función es recorrer la imagen de entrada píxel por píxel, horizontal y verticalmente realizando la operación de convolución, de esta manera se extraen atributos que formarán una nueva matriz. El total de nuevas matrices será igual al número de filtros.

Capa de reducción o max pooling realiza una función similar a la capa convolutacional, posee una matriz (kernel [mxm]) que recorre la imagen de forma

horizontal y vertical con pasos de m , eligiendo el valor más alto entre los píxeles seleccionados. Al final disminuirá las dimensiones de la imagen de entrada.

Capas fully connected o totalmente conectadas se utiliza para realizar la clasificación en base a los resultados entregados por las capas convolucionales o de reducción. Al ser estos resultados matrices se debe utilizar una capa Flatten o capa aplanada la cual transforma una matriz en un vector permitiendo conectar las capas de reducción o convolución con las fully connected. La última capa fully connected utiliza la activación softmax para presentar los resultados en porcentajes. (Google Developers, 2020)

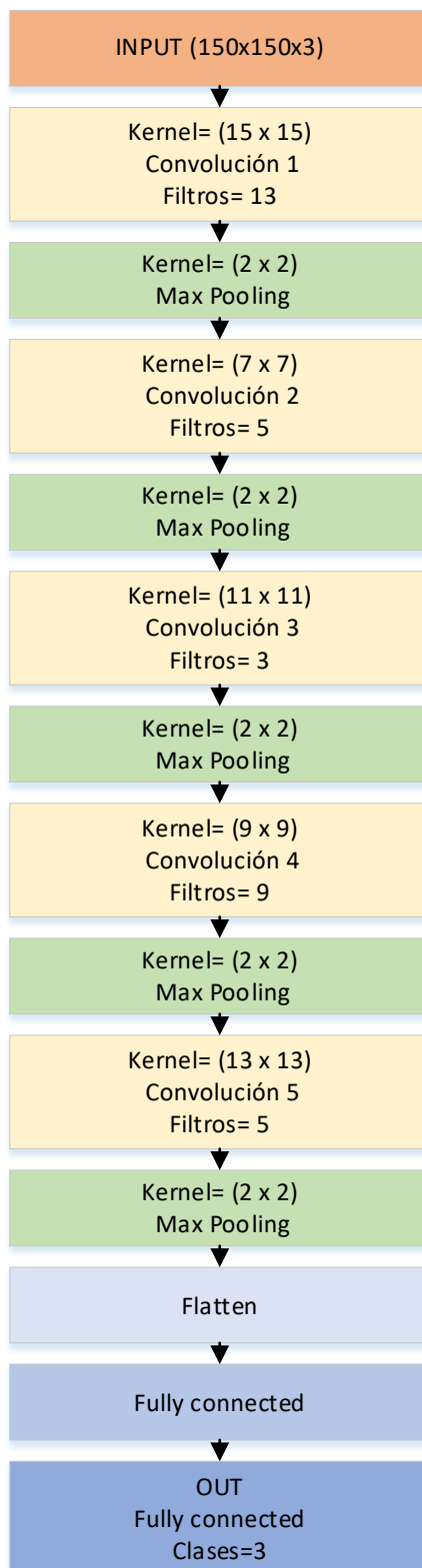
Para crear una red neuronal convolucional en Jupyter Notebook se utiliza las siguientes librerías:

```
import sys
import os
from tensorflow.python.keras.preprocessing.image import
ImageDataGenerator
from tensorflow.python.keras import optimizers
from tensorflow.python.keras.models import Sequential
from tensorflow.python.keras.layers import Dropout, Flatten, Dense,
Activation
from tensorflow.python.keras.layers import Convolution2D,
MaxPooling2D
from tensorflow.python.keras import backend as K
import matplotlib.pyplot as plt
from tensorflow.python.keras.optimizers import Adam
import tensorflow as tf
```

En nuestro caso se implementa la red neuronal convolucional con la arquitectura descrita en la Figura 63.

Figura 63

Arquitectura de la Red Neuronal Convolutiva Implementada



El código desarrollado en Jupyter Notebook para armar la CNN es:

```
K.clear_session()
```

Cantidad de filtros por capa de convolución.

```
filtrosConv1 = 13
```

```
filtrosConv2 = 5
```

```
filtrosConv3 = 3
```

```
filtrosConv4 = 9
```

```
filtrosConv5 = 5
```

Tamaño de Kernel, el cual realizará la operación de convolución con la imagen.

```
tamano_filtro1 = (15,15)
```

```
tamano_filtro2 = (7, 7)
```

```
tamano_filtro3 = (11, 11)
```

```
tamano_filtro4 = (9, 9)
```

```
tamano_filtro5 = (13, 13)
```

Tamaño de Kernel, el cual realizará la operación de max pooling con la imagen.

```
tamano_pool = (2, 2)
```

Número de categorías que obtendremos a la salida

```
clases = 3
```

El optimizador elegido es Adam, cuya función es actualizar los pesos de la CNN

con el fin de mejorar los parámetros de rendimiento de la red.

```
lr = 0.0001
```

```
optimizer = tf.keras.optimizers.Adam(learning_rate=lr)
```

La CNN será secuencial, es decir las capas irán una tras otra.

```
cnn = Sequential()
```

```
cnn.add(Convolution2D(filtrosConv1, tamano_filtro1, padding="same", input_shape=(longitud, altura, 3), activation='relu'))
```

```
cnn.add(MaxPooling2D(pool_size=tamano_pool))
```

```

cnn.add(Convolution2D(filtrosConv2, tamaño_filtro2, padding
="same"))
cnn.add(MaxPooling2D(pool_size=tamaño_pool))

cnn.add(Convolution2D(filtrosConv3, tamaño_filtro3, padding
="same"))
cnn.add(MaxPooling2D(pool_size=tamaño_pool))

cnn.add(Convolution2D(filtrosConv4, tamaño_filtro4, padding
="same"))
cnn.add(MaxPooling2D(pool_size=tamaño_pool))
cnn.add(Convolution2D(filtrosConv5, tamaño_filtro5, padding
="same"))
cnn.add(MaxPooling2D(pool_size=tamaño_pool))

cnn.add(Flatten())
cnn.add(Dense(256, activation='relu'))
cnn.add(Dropout(0.5))
cnn.add(Dense(clases, activation='softmax'))

```

Al momento de compilar la CNN se define parámetros como: el tipo de pérdida, optimizador y las métricas que deseamos obtener.

```

cnn.compile(optimizer=optimizer, loss='categorical_crossentropy', m
etrics=['accuracy'])

```

Con la función `summary` se visualiza las características de la CNN como se

indica en la Figura 64:

```

cnn.summary()

```

Figura 64

Estructura y Parámetros de la Red Neuronal Convolutiva

```

Model: "sequential"

```

| Layer (type) | Output Shape | Param # |
|--------------------------------|----------------------|---------|
| conv2d (Conv2D) | (None, 150, 150, 13) | 8788 |
| max_pooling2d (MaxPooling2D) | (None, 75, 75, 13) | 0 |
| conv2d_1 (Conv2D) | (None, 75, 75, 5) | 3190 |
| max_pooling2d_1 (MaxPooling2D) | (None, 37, 37, 5) | 0 |
| conv2d_2 (Conv2D) | (None, 37, 37, 3) | 1818 |
| max_pooling2d_2 (MaxPooling2D) | (None, 18, 18, 3) | 0 |
| conv2d_3 (Conv2D) | (None, 18, 18, 9) | 2196 |
| max_pooling2d_3 (MaxPooling2D) | (None, 9, 9, 9) | 0 |
| conv2d_4 (Conv2D) | (None, 9, 9, 5) | 7610 |
| max_pooling2d_4 (MaxPooling2D) | (None, 4, 4, 5) | 0 |
| flatten (Flatten) | (None, 80) | 0 |
| dense (Dense) | (None, 256) | 20736 |
| dropout (Dropout) | (None, 256) | 0 |
| dense_1 (Dense) | (None, 3) | 771 |

```

Total params: 45,109
Trainable params: 45,109
Non-trainable params: 0

```

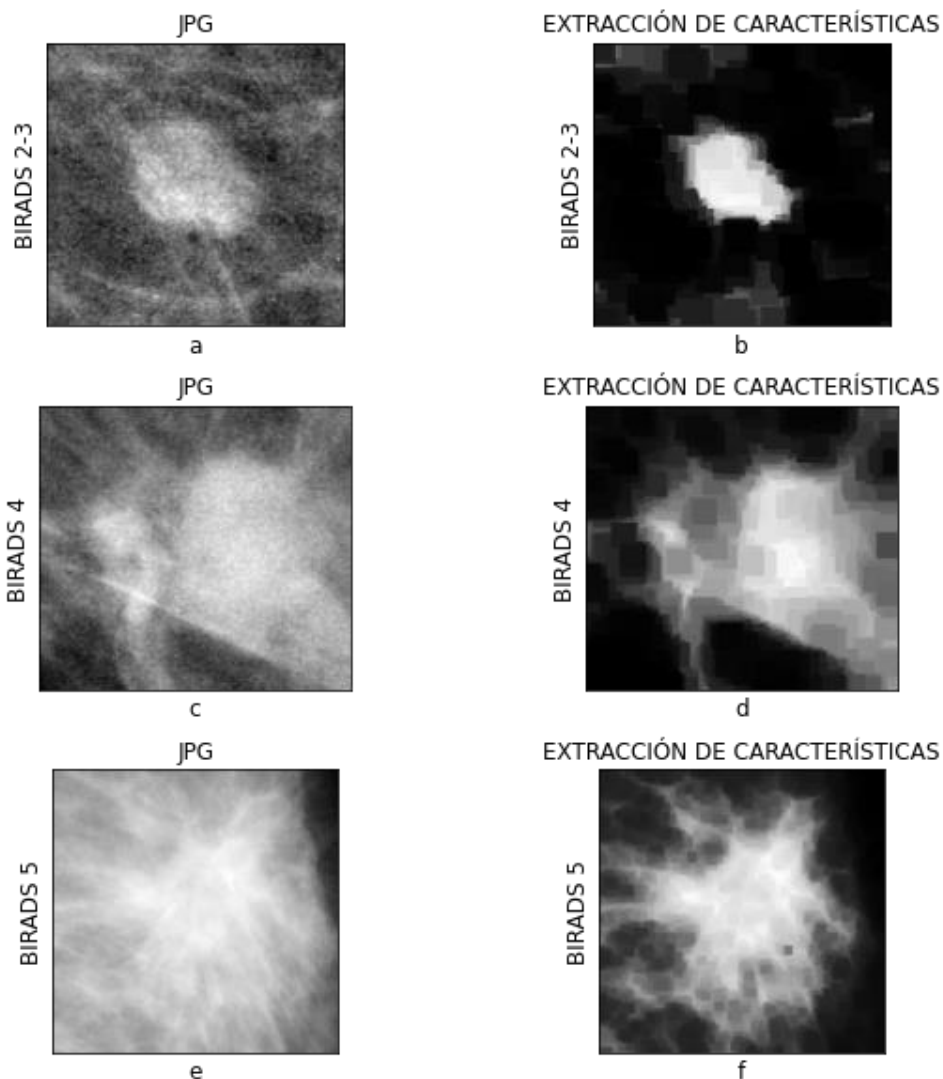
Para entrenar la CNN se utiliza la base de datos CBIS-DDSM ya que posee imágenes de todos los tipos de BIRADS, además tienen imágenes de los nódulos extraídos en formato DICOM.

Se agruparon las imágenes de los nódulos que tengan diagnóstico BIRADS 2 o 3 como sugiere la Tabla 1. Para evitar sobreajuste en el modelo se realiza lo siguiente:

- Seleccionar imágenes que estén sin ruido, sin objetos extraños, y que los nódulos se distingan a simple vista.
- Asegurar que los tipos de BIRADS contengan similar o igual número de imágenes
- Aplicar la etapa de extracción de características a las imágenes seleccionadas y guardar los resultados en formato PNG. En la Figura 65 se visualizan un ejemplo de cada BIRADS.

Figura 65

BIRADS 2-3 (a) Imagen Formato jpg, (b) Extracción de Características, BIRADS 4 (c) Imagen Formato jpg, (d) Extracción de Características, BIRADS 5 (e) Imagen Formato jpg, (f) Extracción de Características



Después de haber realizado la extracción de características a todas las imágenes con las cuales entrenaremos la CNN, dividimos estas imágenes en dos grupos, uno para entrenamiento (aproximadamente 80% de imágenes) y otro de pruebas (aproximadamente 20% de imágenes). Las imágenes se alimentan a la red en lotes de 32 a través de dos generadores, uno de entrenamiento que tendrá la función de

reescalar los valores de 0 a 1, estandarizar el tamaño de las imágenes en 150 x 150, aumentar las imágenes realizando operaciones de espejo, rotación y zoom. El segundo será para validación el cual reescala los valores de las imágenes de 0 a 1, estandariza el tamaño de las imágenes en 150 x 150.

Dirección de las carpetas que contienen las imágenes

```
data_entrenamiento = './nodulo/entrenamiento'
```

```
data_validacion = './nodulo/validacion'
```

```
longitud, altura = 150, 150
```

```
batch_size = 32
```

Generador de imágenes

```
entrenamiento_datagen = ImageDataGenerator( rescale=1. / 255,
                                             rotation_range=10,
                                             shear_range=0.1,
                                             zoom_range=0.1, #0.1
                                             horizontal_flip=True)
```

```
test_datagen = ImageDataGenerator(rescale=1. / 255)
```

```
entrenamiento_generador =
entrenamiento_datagen.flow_from_directory(
    data_entrenamiento,
    target_size=(altura, longitud),
    batch_size=batch_size,
    class_mode='categorical')
```

```
validacion_generador = test_datagen.flow_from_directory(
    data_validacion,
    target_size=(altura, longitud),
    batch_size=batch_size,
    class_mode='categorical')
```

Una vez armada la red neuronal convolucional y cargadas las imágenes que se utilizan para entrenar y validar el modelo, se realiza el entrenamiento de la CNN. Los resultados se presentan y analizan en el siguiente capítulo.

```
Parámetros para el entrenamiento de la red neuronal, numero de épocas, pasos,
epocas=250
pasos = 2000
validation_steps = 200
```

```
history=cnn.fit_generator(entrenamiento_generador, steps_per_epoch
=pasos, epochs=epocas, validation_data=validacion_generador, validat
ion_steps=validation_steps)
```

3.11 Interfaz gráfica

Para mejorar la interacción entre el usuario y la máquina se realiza una interfaz gráfica, Figura 65, la cual es desarrolla en Jupyter Notebook utilizando la librería Tkinter.

Para esto se debe importar esta librería.

```
import tkinter as tk
from tkinter import *
from tkinter import ttk
from tkinter import filedialog
```

Se crea una ventana, que contendrá los demás elementos de la interfaz

```
window = tk.Tk()
window.title("HMI")
```

Añadimos las pestañas de inicio y aplicación

```
tab_control = ttk.Notebook(window)
tab1 = ttk.Frame(tab_control)
tab_control.add(tab1, text='INICIO')
tab2 = ttk.Frame(tab_control)
```

```
tab_control.add(tab2, text='APLICACIÓN')
tab_control.pack(expand=1, fill='both')
```

Añadimos los elementos de la pestaña de inicio

```
icon = tk.PhotoImage(file = "espe.png")
tk.Label(tab1, image = icon, bg = "white", bd=5).pack(fill = "x")
tk.Label(tab1, text = "LATACUNGA -2020", font=("Arial Bold", 10),
fg = "white", bg = "black", bd=10).pack(fill = "x",side = "bottom")
tk.Label(tab1, text = "CARRERA DE INGENIERÍA EN ELECTRÓNICA E
INSTRUMENTACIÓN.", fg = "white", height = 2 ,width = 0, bg =
"#33cc66",font=("Arial Bold", 15)).pack(fill = "x")
icon2 = tk.PhotoImage(file = "electronica.png")
tk.Label(tab1, image = icon2, bg = "white").pack(side = "left",
fill = "y")
tk.Label(tab1, text = "SISTEMA PARA CLASIFICACIÓN Y RECONOCIMIENTO
\n DE NÓDULOS EN IMÁGENES MAMOGRÁFICAS \n UTILIZANDO PROCESAMIENTO
DIGITAL DE IMÁGENES Y REDES NEURONALES.", fg = "black", width = 80,
bg = "#cae9f5",font=("Arial Bold", 15)).pack(side = "left",fill =
"y")
tk.Label(tab1, text = "AUTORES: \n\n\n Ivonne Bermeo \n Jairo
Chicaiza", fg = "black", bg = "white", width = 80, font=("Arial
Bold", 15)).pack(side = "left",fill = "y")
```

Añadimos los elementos de la pestaña aplicación como etiquetas, botones.

```
lbl=Label(tab2, text = "SELECCIONAR ARCHIVO", font=("Arial Bold",
10))
lbl.grid(column=0, row=0)
```

```
def clicked():
```

```
# selecciona y abre una mamografía en formato DICOM
```

```
def procesar():
```



```
# Procesamiento digital de mamografías
# Extracción de características
def identificar():
# Clasificación de nódulos mamarios (modelo entrenado CNN)
def salir():
#cierra la aplicación

# Botones
btn = Button(tab2, text="Abrir", command=clicked, font = ('Agency
FB:',12),height = 2 , width = 15)
btn.place(x = 25, y =100)
btn2 = Button(tab2, text="Procesar", command=procesar, font =
('Agency FB:',12),height = 2 , width = 15)
btn3 = Button(tab2, text="Identificar", command=identificar, font
= ('Agency FB:',12),height = 2 , width = 15)
btn4 = Button(tab2, text="Salir", command=salir, font = ('Agency
FB:',12),height = 2 , width = 15)
btn4.place(x = 25, y =250)

window.mainloop()
```

Figura 66

Interfaz gráfica: (a) Pestaña de Inicio (b) Pestaña de Aplicación



(a)



(b)

La interfaz gráfica presenta 2 pestañas: Inicio y Aplicación. La pestaña de Inicio contiene una portada de la Universidad, la carrera y autores. La pestaña Aplicación contiene 4 botones:

- Abrir: Permite al usuario abrir una imagen mamográfica en formato DICOM, la cual se visualiza en el recuadro Mamografía Original.
- Procesar: Realiza el procesamiento digital de imágenes y extracción de características de la imagen.
- Identificar: Encierra en un cuadro el nódulo o nódulos detectados, maximiza la imagen del nódulo que se visualiza en el recuadro Obtención de nódulos, e indica el resultado posible de clasificación BIRADS a través del modelo entrenado de la CNN.
- Salir: Permite cerrar o salir de la aplicación.

4 Análisis y Resultados

En este capítulo se presenta pruebas y resultados obtenidos en las etapas de procesamiento digital de mamografías y entrenamiento de la red neuronal convolucional. Además, se evalúa el sistema de detección e identificación de nódulos mamarios a través de los diagnósticos emitidos por especialistas.

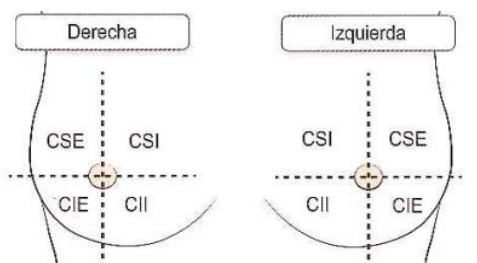
4.1 Análisis etapa de procesamiento digital de mamografías

Al aplicar la etapa de procesamiento digital a una mamografía se extraen los nódulos presentes en la misma, en caso de no presentar nódulos se emite el mensaje "BIRADS 1" y si el área extraída es demasiada grande se emite el mensaje "BIRADS 0" lo que indica que se trata de una mama densa. A continuación, se evalúa esta etapa del sistema contrastando los resultados obtenidos con el diagnóstico médico del paciente otorgado por el centro médico Laser, el cual describe la cantidad de masas presentes en la mama y su ubicación.

Para realizar una correcta interpretación del diagnóstico médico se presenta en la Figura 67 la distribución de los cuatro cuadrantes de la mama: cuadrante superior externo (CSE), cuadrante superior interno (CSI), cuadrante inferior externo (CIE), cuadrante inferior interno (CII). Mientras que en la Figura 68 se ubica los cuadrantes superiores, inferiores, externos e internos dependiendo la proyección en la que se realiza la mamografía.

Figura 67

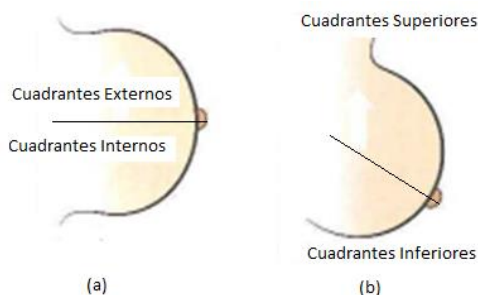
Distribución de los Cuadrantes en la Mama Izquierda y Derecha



Nota. Tomado de *Manual de Radiología Mamaria*, por López Ruiz & Pina Insausti, 2016, Medica Panamericana.

Figura 68

Mamografía con Proyección: (a) Craneocaudal en la Cual se Distinguen los Cuadrantes Externos e Internos de la Mama, (b) Medio Lateral Oblicua en la Cual se Distinguen los Cuadrantes Superiores e Inferiores de la Mama.



Nota. Tomado de *Manual de Radiología Mamaria*, por López Ruiz & Pina Insausti, 2016, Medica Panamericana.

Para iniciar el proceso de evaluación se presenta el diagnóstico del paciente 1 emitido por centro médico Laser: Imagen nodular densa, irregular, con márgenes lobulados de 23 x 33 mm localizada en el cuadrante supero-externo de mama izquierda. En contraste, el sistema detecta una imagen nodular ubicada en el cuadrante superior de la mamografía con proyección MLO de la Figura 69 y completa su ubicación con la mamografía de proyección CC de la Figura 70 en la cual localiza una masa en el

cuadrante externo. De esta forma se concluye que el sistema concuerda con el diagnóstico médico, reconociendo la masa en el cuadrante superior externo de la mama izquierda.

Figura 69

Mamografía del Paciente 1 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara del Nódulo Extraído Ubicado en el Cuadrante Superior

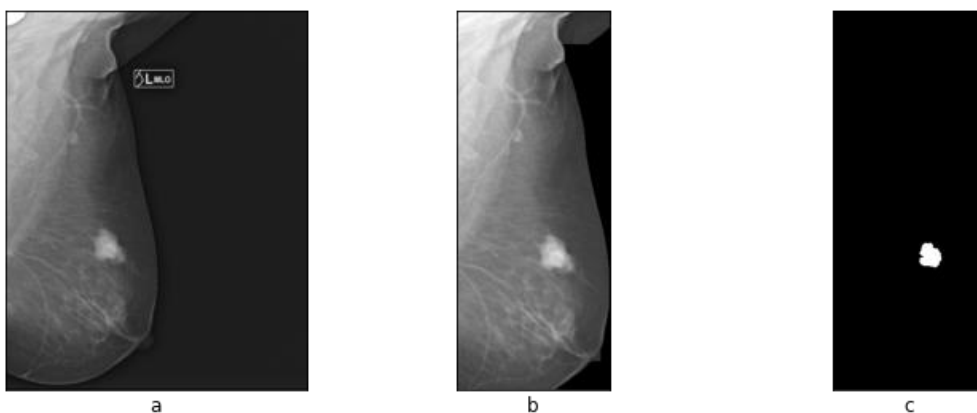
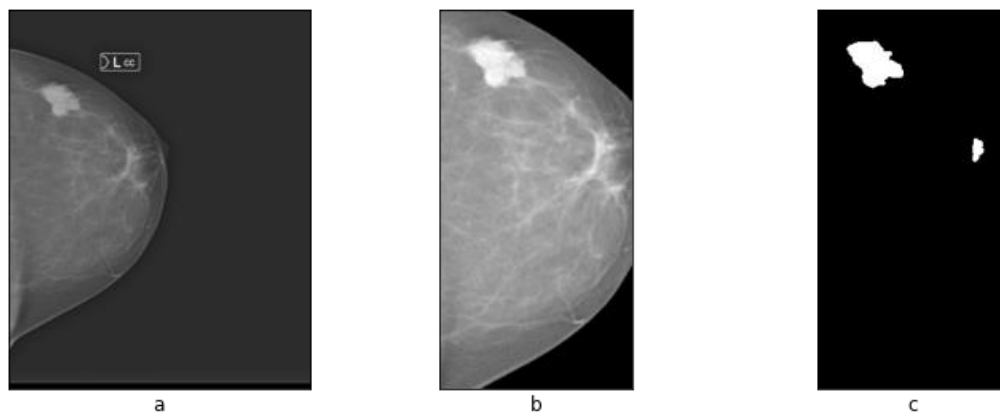


Figura 70

Mamografía del Paciente 1 Con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara del Nódulo Extraído Ubicado en el Cuadrante Externo



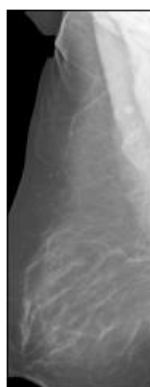
El diagnóstico médico del paciente 2 carece de hallazgos nodulares, por ende, lo categorizan como BIRADS 1. En contraste, la respuesta del sistema al procesar la mamografía con proyección MLO de la Figura 71 es un BIRADS 1 (no encontró nódulos), mientras que en la proyección CC de la Figura 72 localiza masas en el cuadrante interno y externo de la mama. De esta forma se concluye que el sistema concuerda con el diagnóstico médico al procesar la mamografía con proyección MLO y erra al procesar la mamografía con proyección CC ya que detecta tejido.

Figura 71

Mamografía del Paciente 2 con proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Imagen Sin Nódulos (BIRADS 1)



a



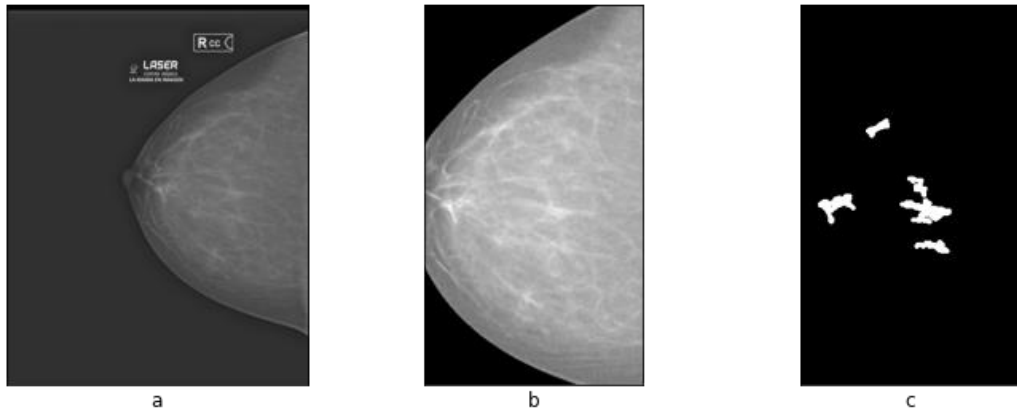
b



c

Figura 72

Mamografía del Paciente 2 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Tejido Extraído



El paciente 3 presenta el siguiente diagnóstico médico: presencia de 2 imágenes densas de márgenes indefinidos localizadas en el cuadrante inferior externo y una imagen de similares características localizada en el cuadrante superior interno de la mama derecha. Al procesar la mamografía con proyección MLO de la Figura 73 se detectan 2 masas ubicadas en el cuadrante inferior y una en el cuadrante superior, mientras que al procesar la mamografía con proyección CC de la Figura 74 solo se detecta una masa ubicada en el cuadrante externo. De esta forma se concluye que el sistema acierta al procesar la mamografía con proyección MLO y erra al procesar la mamografía con proyección CC ya que solo obtiene un hallazgo.

Figura 73

Mamografía del Paciente 3 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulos Extraídos Ubicados en el Cuadrante Superior e Inferior de la Mama

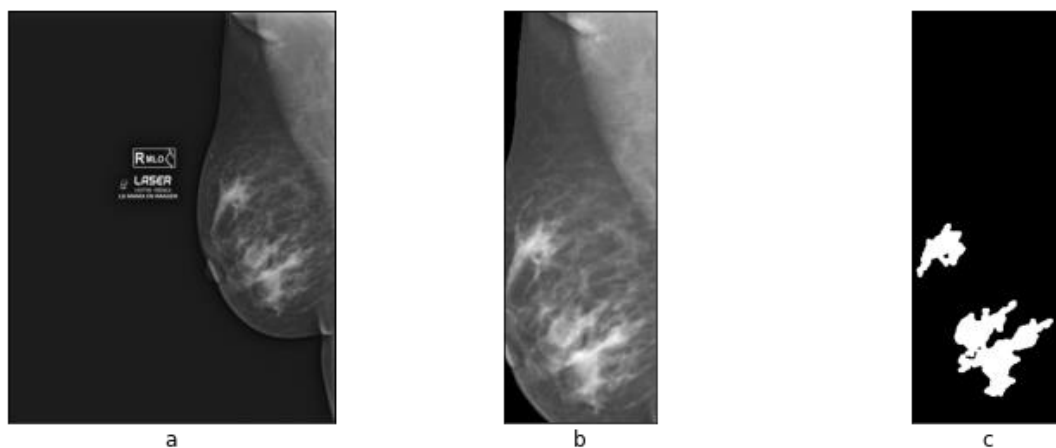
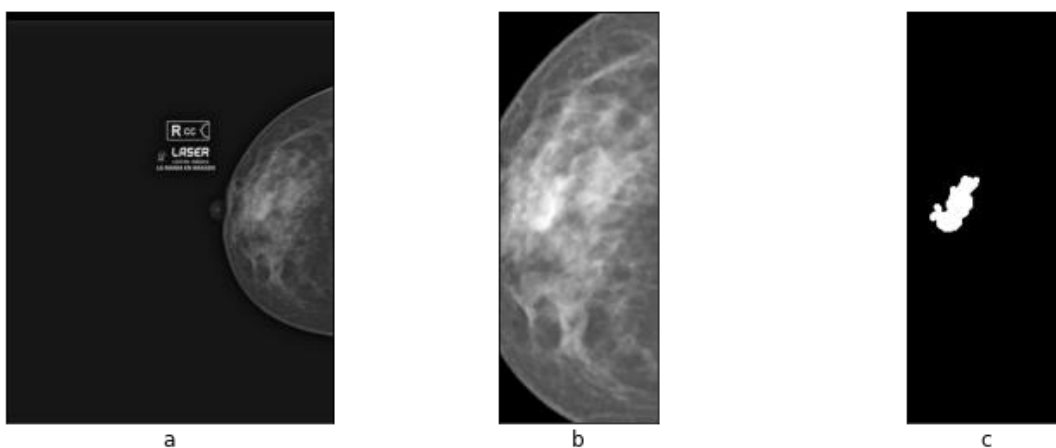


Figura 74

Mamografía del Paciente 3 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído ubicado en el Cuadrante Superior de la Mama



El diagnóstico médico del paciente 4 es: Presencia de dos imágenes densas localizadas en el cuadrante superior externo y superior interno de la mama izquierda

visualizado en la proyección cráneo caudal, que miden 15 y 10 mm respectivamente. Por su parte el sistema al procesar la mamografía con proyección MLO de la Figura 75 detecta una masa ubicada el cuadrante superior de la mama y otra ubicada en el cuadrante inferior, por otra parte, al procesar la mamografía con proyección CC de la Figura 76 el sistema detecta una masa ubicada en el cuadrante externo de la mama y otra en el cuadrante interno. En conclusión, el sistema erra al momento de procesar la mamografía con proyección MLO ya que detecta tejido en el cuadrante inferior de la mama y acierta al procesar la mamografía con proyección CC ya que detecta las masas en el cuadrante externo e interno. En este caso el médico especialista emite su diagnóstico apoyándose en la proyección CC.

Figura 75

Mamografía del Paciente 4 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído Ubicado en el Cuadrante Superior y Máscara de Tejido Ubicada en el Cuadrante Inferior de la Mama.

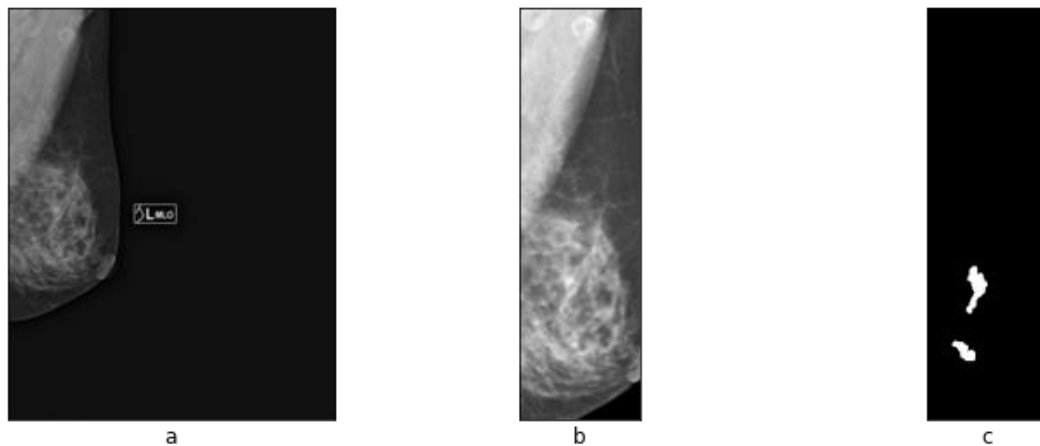
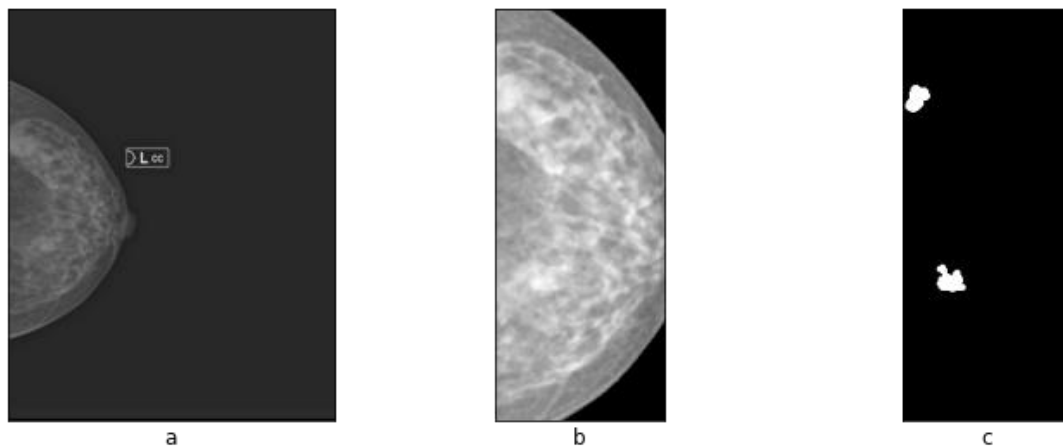


Figura 76

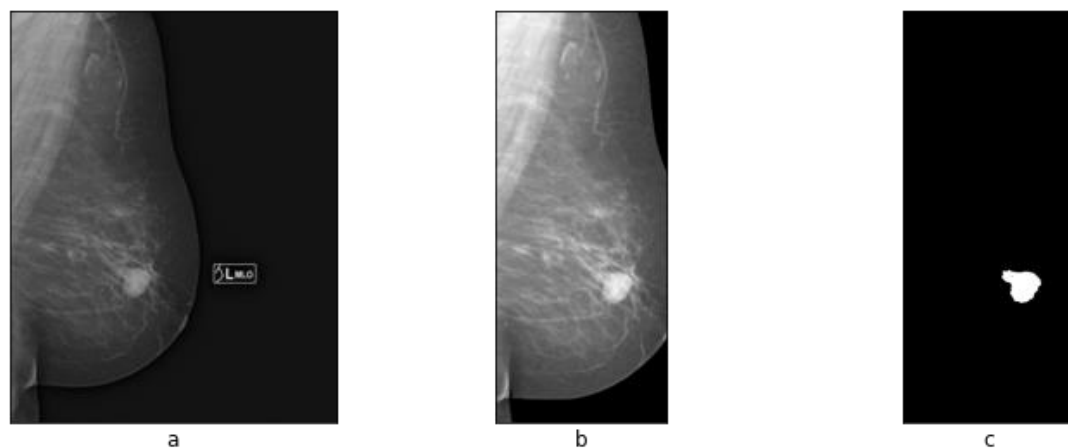
Mamografía del Paciente 4 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído Ubicada en el Cuadrante Exterior e Inferior de la Mama.



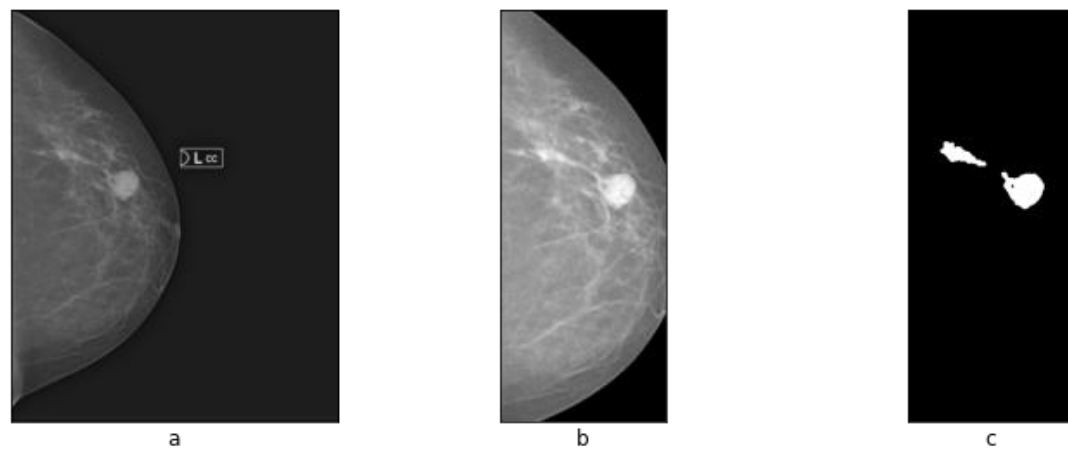
El diagnóstico médico del paciente 5 es: Imagen nodular densa con márgenes circunscritos, de 20 mm, otra imagen de similares características con microcalcificaciones pleomórficas de 11 mm localizadas en el cuadrante superior-externo de la mama izquierda. Por su parte el sistema procesa la mamografía con proyección MLO de la Figura 77 y encuentra una masa ubicada en el cuadrante superior, mientras que al procesar la mamografía con proyección CC de la Figura 78 se encuentra dos nódulos ubicados en el cuadrante externo de la mama. De esta forma se concluye que el sistema erra al procesar la mamografía con proyección MLO ya que detecta un solo nódulo y acierta al procesar la mamografía con proyección CC al detectar dos masas.

Figura 77

Mamografía del Paciente 5 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído Ubicado en el Cuadrante Superior de la Mama

**Figura 78**

Mamografía del Paciente 5 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulos Extraídos Ubicados en el Cuadrante Externo de la Mama.



El paciente 6 presenta el siguiente diagnóstico médico: Imagen hiperdensa, espiculada de 45 x 40 mm más calcificaciones amorfas, localizada en el cuadrante superior externo de la mama izquierda. Por otra parte, el sistema procesa la mamografía con proyección MLO de la Figura 79 e identifica dos masas en el cuadrante superior de la mama, mientras que al procesar la mamografía con proyección CC de la Figura 80 se identifica una masa ubicada en el cuadrante externo de la mama. De esta forma se concluye que el sistema erra al procesar la mamografía con proyección MLO ya que identifica una lesión ubicada en el músculo de la mama, y acierta al procesar la mamografía con proyección CC ya que localiza una masa en el cuadrante externo de la mama.

Figura 79

Mamografía del Paciente 6 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído Localizado en el Cuadrante Superior y Máscara de Lesión Ubicada en el Músculo de la Mama.

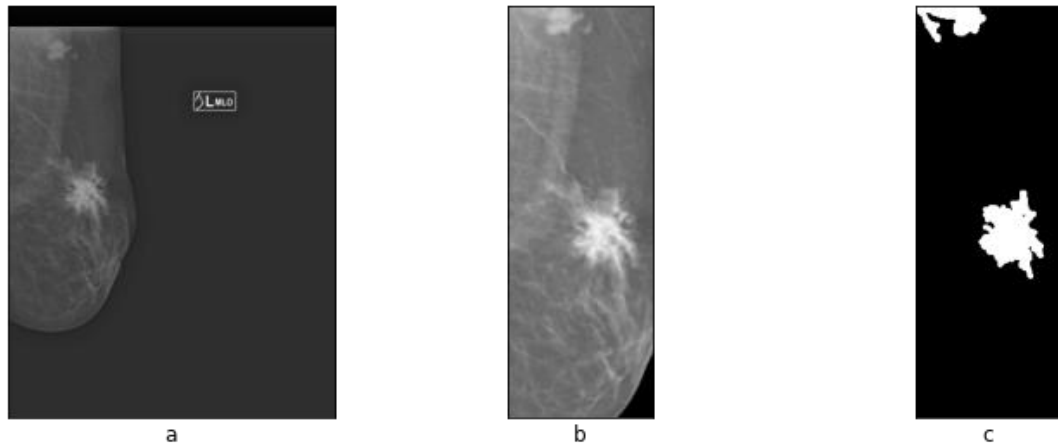
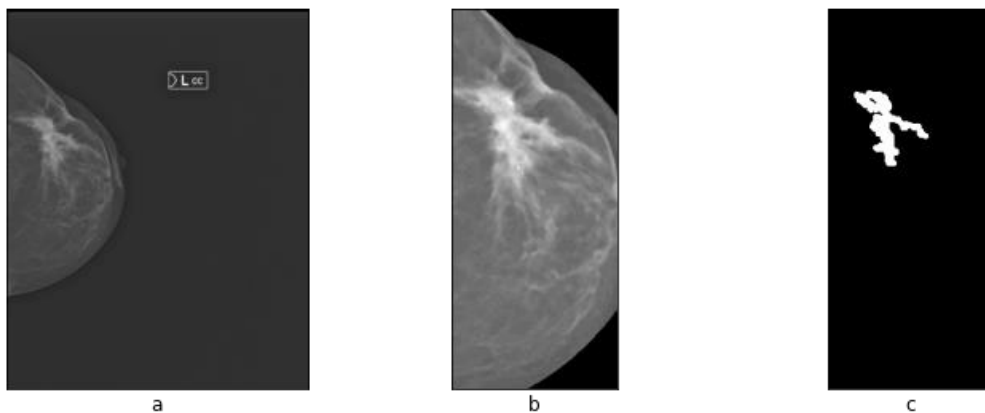


Figura 80

Mamografía del Paciente 6 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído



El paciente 7 presenta el siguiente diagnóstico médico: Imagen nodular densa, irregular con márgenes oscurecido localizada en el cuadrante infero-interno de la mama izquierda. Estudio incompleto: Por la densidad mamaria se sugiere complementar con ecografía. El sistema procesa la mamografía con proyección MLO de la Figura 81 y determina que es una mama densa, mientras que al procesar la mamografía con proyección CC de la Figura 82 se reconoce una masa ubicada en el cuadrante interno de la mama. En conclusión, el sistema acierta al categorizarla como BIRADS 0, y acierta al localizar una masa localizada en el cuadrante interno de la mama.

Figura 81

Mamografía del Paciente 7 con Proyección MLO (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Mama Densa Categorizada Como BIRADS 0

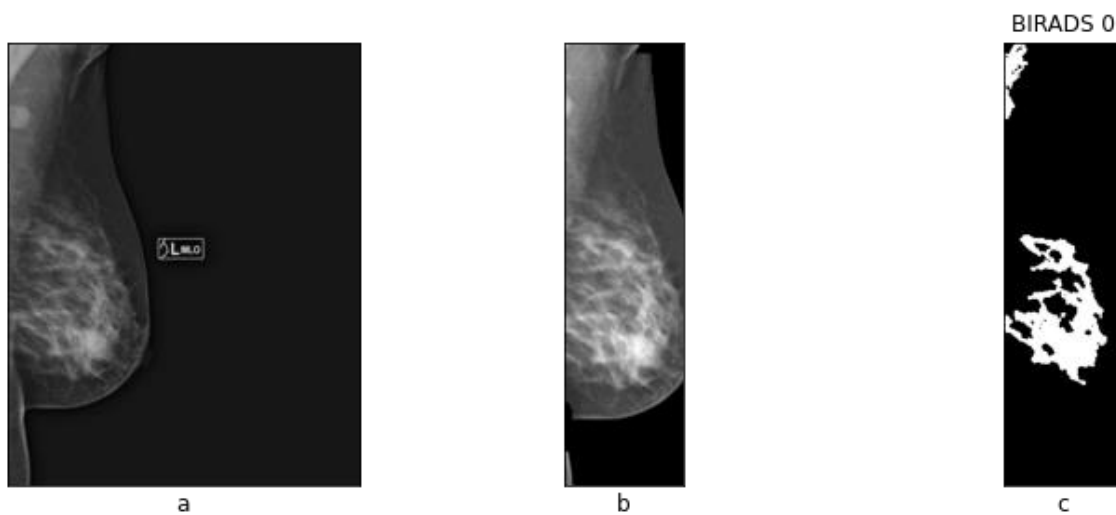
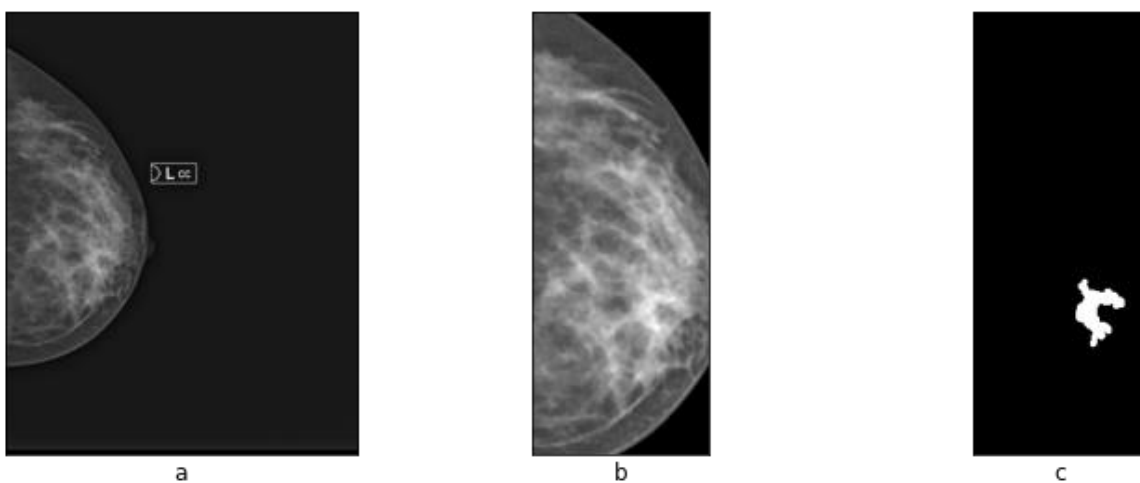


Figura 82

Mamografía del Paciente 7 con Proyección CC (a) Mamografía de Entrada, (b) Mamografía Recortada (c) Máscara de Nódulo Extraído ubicado en el cuadrante interno de la mama.



Los resultados globales arrojados por la etapa de procesamiento digital de mamografías, adjudican un buen porcentaje de acierto al momento de reconocer y

extraer nódulos mamarios, dependiendo la proyección en la que se realiza la mamografía se pueden solapar o descubrir nuevas anormalidades mismas que deben ser ratificadas o descartadas revisando las demás proyecciones del examen mamográfico.

La respuesta de esta etapa puede verse afectada en casos de mamas densas, cuando el tejido glandular mamario es mayor al adiposo, el porcentaje de extracción y reconocimiento baja ya que el nódulo se confunde con el tejido glandular.

4.2 Análisis de la red neuronal convolucional

En esta sección se presenta la evaluación del entrenamiento de la red neuronal convolucional con la arquitectura descrita en el capítulo anterior, adicional a esto se presenta una matriz de confusión utilizando las imágenes ocupadas en el entrenamiento y otra con imágenes nuevas, con las cuales se analiza parámetros de desempeño de la CNN como: exactitud, sensibilidad, precisión y f1 score.

La matriz de confusión es una herramienta visual que resume el rendimiento de la CNN, muestra los errores y aciertos que ha tenido nuestro modelo al momento de realizar predicciones. La mayor parte de métricas de rendimiento se basan en ella. (Artola Moreno, 2019). Para calcular los parámetros de la CNN a partir de la matriz de confusión se debe aclarar los conceptos de true positives (TP: casos correctamente clasificados como positivos), true negative (TN: casos correctamente clasificados como negativos), false positives (FP: casos erróneamente clasificados como positivos) y false negative (FN: casos erróneamente clasificados como negativos), cuya relación se visualiza en la Tabla 2.

Tabla 2

Esquema General Matriz de Confusión

| | | Clase predicha | |
|--------------|-----------|----------------|-----------|
| | | Positivos | Negativos |
| Clase actual | Positivos | TP | FP |
| | Negativos | FN | TN |

Las principales métricas que cuantifican el rendimiento de la CNN son:

- **Accuracy:** Cantidad de casos predichos correctamente por la CNN sobre todo tipo de predicción realizada, se calcula con la Ecuación 4.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

- **Precisión:** Relación entre los casos clasificados en una categoría y los que realmente pertenecen a esa categoría, su valor se calcula con la Ecuación 5.

$$Precisión = \frac{TP}{TP+FP} \quad (5)$$

- **Recall (sensibilidad):** Tasa de verdaderos positivos, relación entre los casos clasificados en una categoría y los casos que efectivamente son de esa categoría, su valor se calcula con la Ecuación 6.

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

- **F1 score:** Es la media ponderada de precisión y sensibilidad. Resume la precisión y la sensibilidad, se calcula con la Ecuación 7.

$$F1 \text{ score: } = \frac{2*(Recall*Precisión)}{Recall+Precision} \quad (7)$$

Después de realizar el entrenamiento de la CNN con 224 imágenes, aproximadamente 75 por clase, se obtuvo un accuracy (exactitud) aproximado de 0.91 y un loss (pérdida) mínima de 0.23. La evolución de los parámetros de accuracy y loss

durante el entrenamiento realizado se presenta en la Figura 83 y Figura 84 respectivamente.

El porcentaje de accuracy puede mejorar al incrementar la cantidad de imágenes que se utilizan para el entrenamiento, o incluso implementando una red neuronal con una estructura más compleja, aunque se debe tener en cuenta que un 100% de accuracy no es recomendable ya que la red neuronal convolucional aprenderá incluso del ruido presente en las imágenes, obteniendo como resultado una generalización pobre, es decir no se obtendrán buenos resultados de predicción cuando el modelo se enfrente a imágenes nuevas. (Jiménez, 2012)

Figura 83

Evolución de Accuracy (Precisión) Durante las 250 Épocas de Entrenamiento

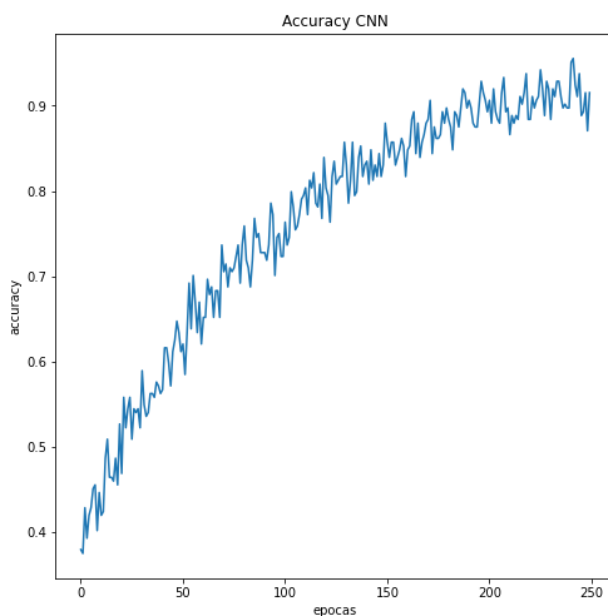
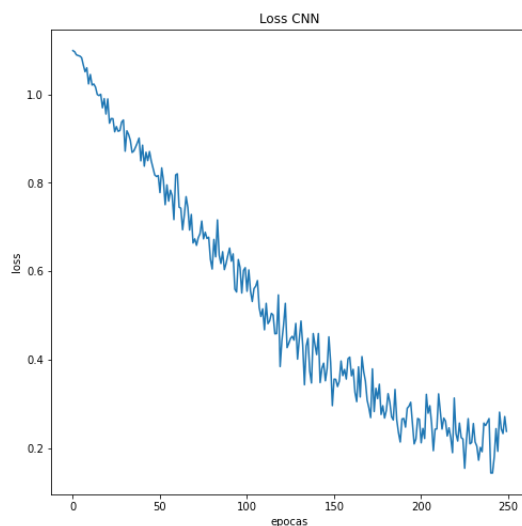


Figura 84

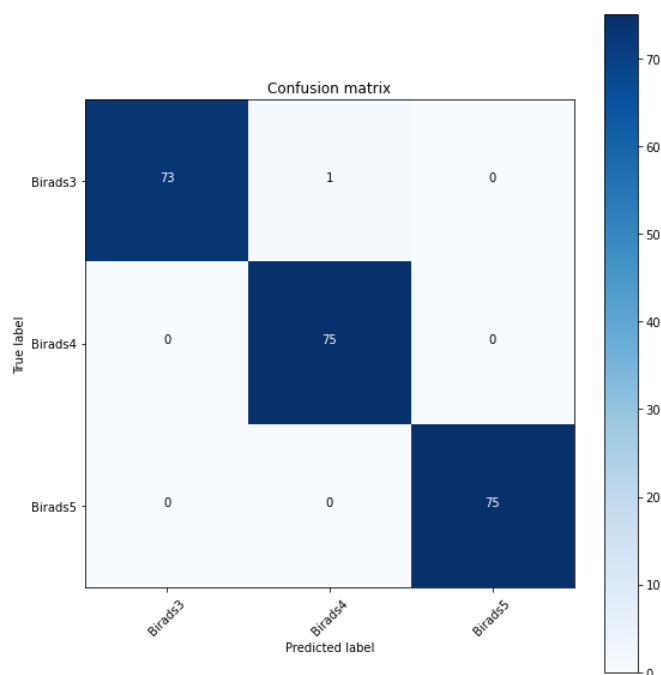
Evolución de Loss (Pérdida) Durante las 250 Épocas de Entrenamiento



Se carga el modelo entrenado y se procede a realizar predicciones utilizando las imágenes de entrenamiento, en la Figura 85 se muestra la matriz de confusión generada por el proceso de comparación de la etiqueta original y la predicha.

Figura 85

Matriz de Confusión con Imágenes Utilizadas en el Entrenamiento



Por medio de la librería sklearn calculamos los parámetros de la CNN los cuales se describen en la Tabla 3. El accuracy del sistema es de 0.996, mientras que la clase que menor precisión presenta es la BIRADS 4 al poseer un falso positivo, la categoría BIRADS 3 tiene un falso negativo el cual afecta al Recall es esta clase, la categoría BIRADS 5 no presenta falsos negativos y tampoco falsos positivos, por último, de las 224 imágenes el sistema clasificó mal solo 1 imagen.

Tabla 3

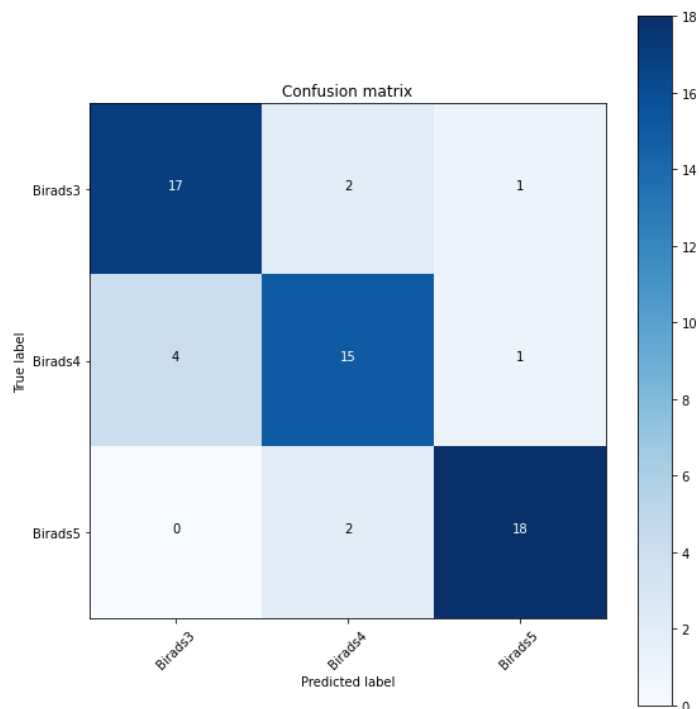
Resultados del Entrenamiento de Clasificación BIRADS

| | Precisión | Recall | F1-score | Support |
|---------------------|------------------|---------------|-----------------|----------------|
| BIRAD 2-3 | 1.000 | 0.986 | 0.993 | 74 |
| BIRAD 4 | 0.987 | 1.000 | 0.993 | 75 |
| BIRAD 5 | 1.000 | 1.000 | 1.000 | 75 |
| accuracy | | | 0.996 | 224 |
| Macro avg | 0.996 | 0.995 | 0.996 | 224 |
| Weighted avg | 0.996 | 0.996 | 0.996 | 224 |

Estos parámetros presentados no aseguran que el modelo se desempeñe de igual forma con nuevas imágenes, por lo que se realiza una matriz de confusión utilizando 20 imágenes nuevas por categoría, se obtendrán parámetros los cuales muestran el desempeño real del modelo entrenado. En la Figura 86 se grafica la matriz de confusión utilizando 60 imágenes nuevas.

Figura 86

Matriz de Confusión con Imágenes Nuevas.



En la matriz de confusión se observa que la categoría BIRADS 3 presenta 4 falsos positivos, la categoría BIRADS 4 tiene 4 falsos positivos y la categoría BIRADS 5 posee 2 falsos positivos, los cuales afectan a la precisión de la clase. El Recall por clase se ve afectado por los falsos negativos, la clase BIRADS 3 tiene 3, la clase BIRADS 4 presenta 5 y la clase BIRADS 5 posee 2. De esta forma deducimos que la clase que mayor precisión, recall, f1 score presenta es la BIRADS 5.

Con los nuevos datos obtenidos, que se indica en la Tabla 4, se concluye que el modelo responde de manera correcta cuando se enfrenta a nuevas imágenes, a pesar que el accuracy del sistema redujo a 0.833 (83.3%).

Tabla 4*Resultados de Validación de la Clasificación BIRADS*

| | Precisión | Recall | F1-score | Support |
|---------------------|------------------|---------------|-----------------|----------------|
| BIRAD 2-3 | 0.810 | 0.850 | 0.829 | 20 |
| BIRAD 4 | 0.789 | 0.750 | 0.769 | 20 |
| BIRAD 5 | 0.900 | 0.900 | 0.900 | 20 |
| Accuracy | | | 0.833 | 60 |
| Macro avg | 0.833 | 0.833 | 0.833 | 60 |
| Weighted avg | 0.833 | 0.833 | 0.833 | 60 |

De esta forma se afirma que el modelo entrenado tiene un buen porcentaje de acierto, y puede ser utilizado para clasificar nódulos mamarios en las categorías: BIRADS 2-3, BIRADS 4, BIRADS 5.

4.3 Análisis del sistema para clasificación y reconocimiento de nódulos mamarios

Una vez realizado el análisis a las etapas que forman el sistema para clasificación y reconocimiento de nódulos mamarios, se procede a evaluar el sistema completo con el fin de verificar el funcionamiento del mismo. Para esto se utilizan algunos casos proporcionados por el centro médico Láser cuyo diagnóstico se contrasta con el proporcionado por el sistema.

El diagnóstico del paciente 1 proporcionado por el centro médico Láser es: BIRADS 4B (hallazgo sospechoso. Por su parte el sistema determina que la masa encontrada en la mamografía de proyección MLO de la Figura 87 es un BIRADS 4, y para el nódulo encontrada en la mamografía de proyección CC de la Figura 88, el

sistema le asigna la categoría BIRADS 4. Con esto verificamos que el sistema concuerda con el diagnóstico médico del paciente.

Figura 87

Diagnóstico Proporcionado por el Sistema al Paciente 1: Imagen Nodular de Categoría BIRADS 4, Proyección MLO

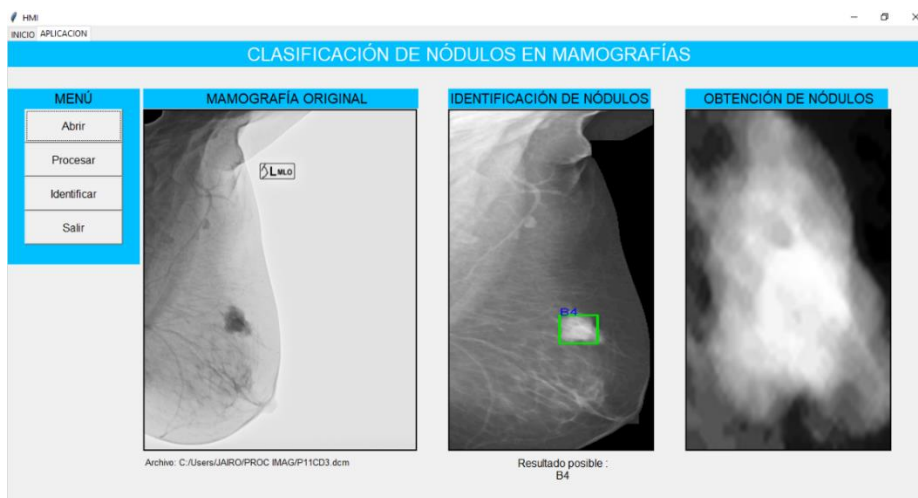
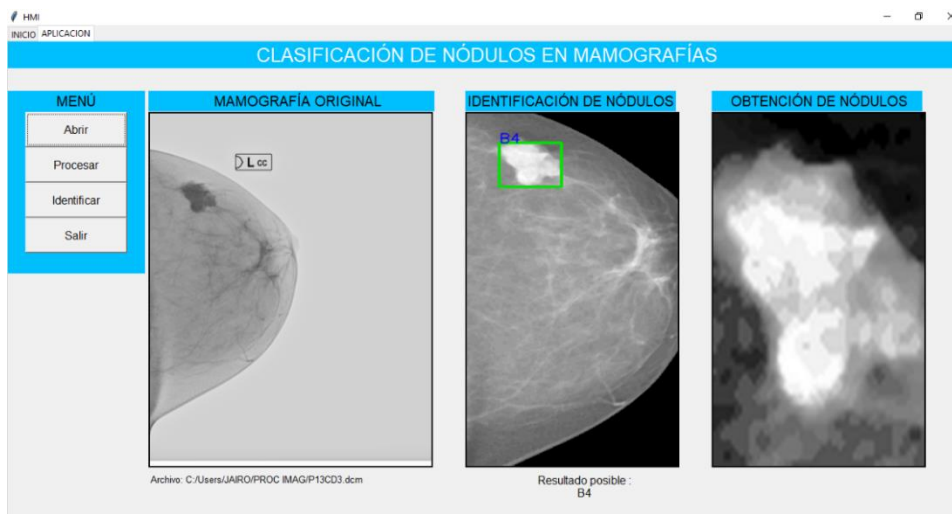


Figura 88

Diagnóstico Proporcionado por el Sistema al Paciente 1: Imagen Nodular de Categoría BIRADS 4, Proyección CC



Para el paciente 2 el diagnóstico médico proporcionado por el centro médico Láser es el siguiente: BIRADS 1 (Se observan sectores dispersos de densidad

fibroglandular, ACR tipo B). El sistema arroja dos resultados dependiendo la proyección en la que se tomó la mamografía. En la Figura 88 con proyección MLO el sistema acierta y no encuentra nódulos, mientras que en la Figura 89 con proyección CC el sistema detecta tejido y lo confunde con nódulos clasificándolos en diferentes categorías BIRDAS.

Figura 89

Diagnóstico Proporcionado por el Sistema al Paciente 2: Imagen sin Nódulos, BIRADS 1, Proyección MLO

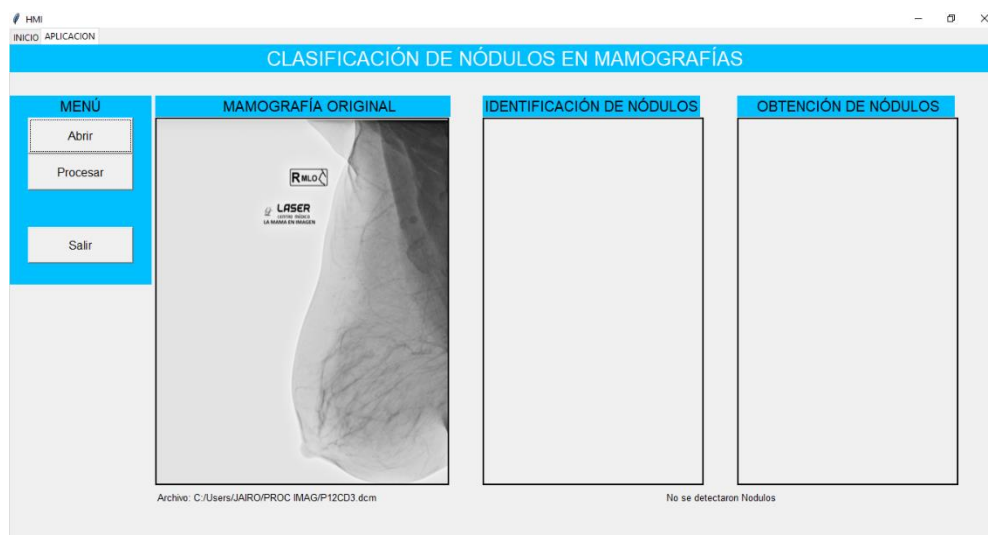
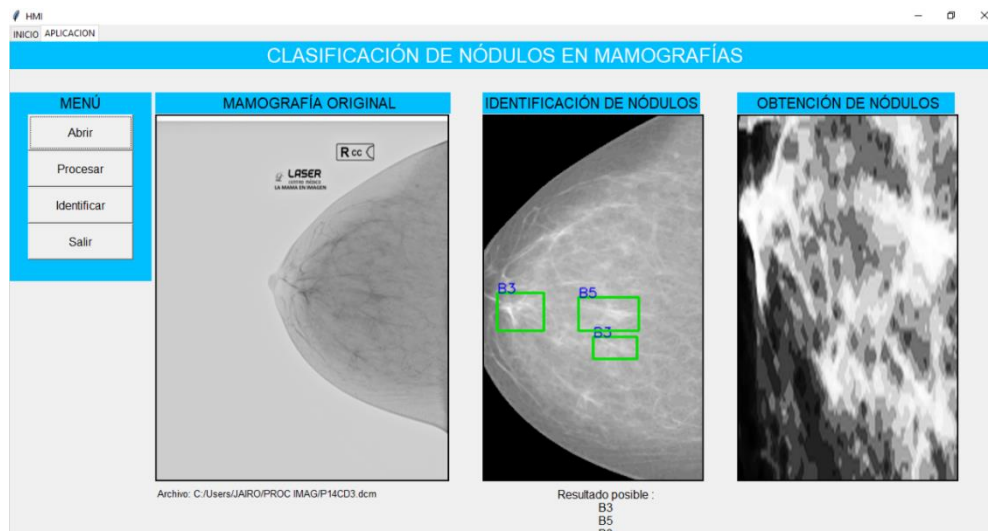


Figura 90

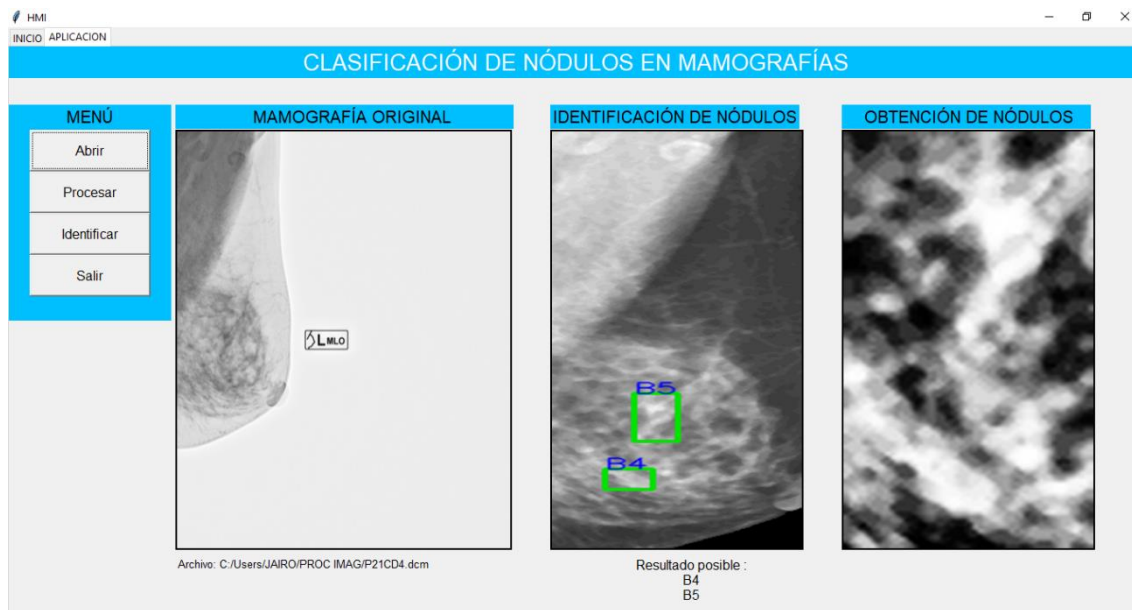
Diagnóstico Proporcionado por el Sistema al Paciente 2: Imágenes Nodulares de Categorías BIRADS 3 y 5, Proyección CC



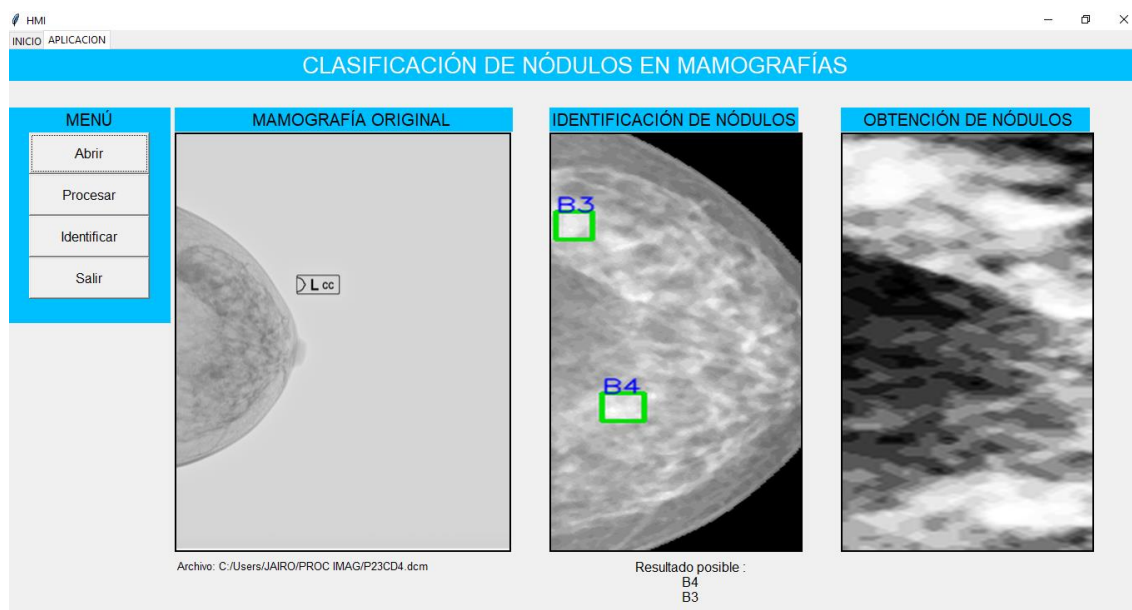
Para el paciente 3 el diagnóstico médico es BIRADS 0 (Presencia de dos imágenes densas localizadas en el cuadrante superior externo y superior interno de la mama izquierda visualizado en la proyección cráneo caudal, que miden 15 y 10 mm respectivamente). De acuerdo al diagnóstico del centro médico Laser este caso no puede ser diagnosticado únicamente con la mamografía por lo que sugieren realizar estudios complementarios. El sistema arroja dos resultados dependiendo la proyección en la que se tomó la mamografía. En la Figura 91 con proyección MLO el sistema encuentra dos masas con categorías BIRADS 5 y 4, mientras que en la Figura 92 con proyección CC el sistema encuentra 2 masas con categorías BIRADS 3 y 4. En conclusión el sistema encuentra dos nódulos y los clasifica de acuerdo a las características que presentan.

Figura 91

Diagnóstico Proporcionado por el Sistema al Paciente 3: Imágenes Nodulares de Categorías BIRADS 4 y 5, Proyección MLO

**Figura 92**

Diagnóstico Proporcionado por el Sistema al Paciente 3: Imágenes Nodulares de Categorías BIRADS 3 y 4, Proyección CC



El diagnóstico médico para el paciente 4 es el siguiente: BIRADS 4A (Imagen nodular densa con márgenes circunscritos, de 20 mm, otra imagen de similares características con microcalcificaciones pleomórficas de 11 mm localizadas en el cuadrante superior-externo de la mama izquierda). El diagnóstico del centro médico Laser generaliza los hallazgos, por ende, se tiene un solo diagnóstico. En la Figura 93 con proyección MLO el sistema encuentra una masa y la categoriza como BIRADS 4, mientras que en la Figura 94 con proyección CC el sistema detecta 2 masas de categoría BIRDAS 4, concordando con el número de nódulos descritos en el diagnóstico.

Figura 93

Diagnóstico Proporcionado por el Sistema al Paciente 4: Imagen Nodular de Categoría BIRADS 4, Proyección MLO

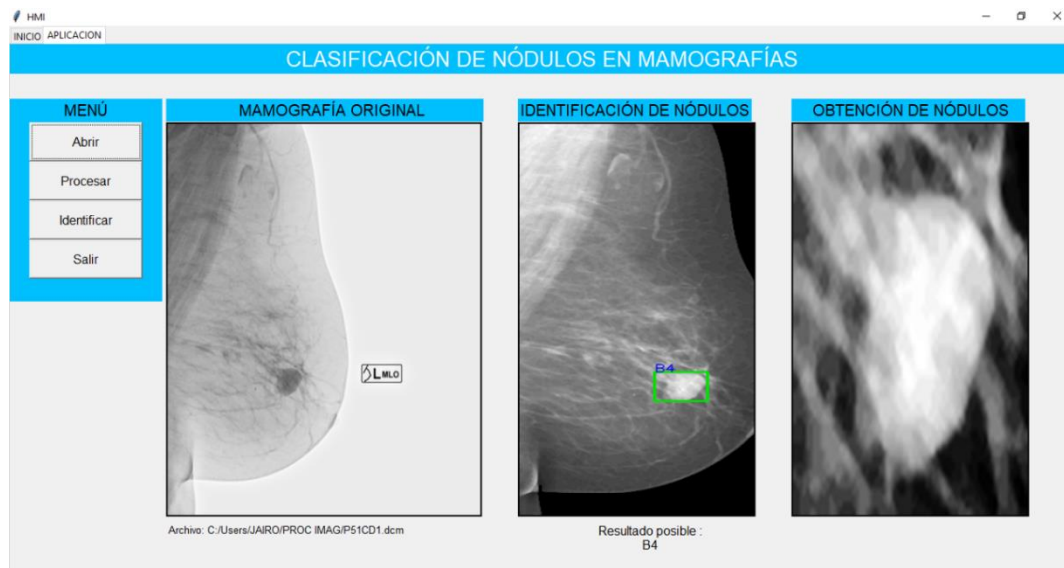
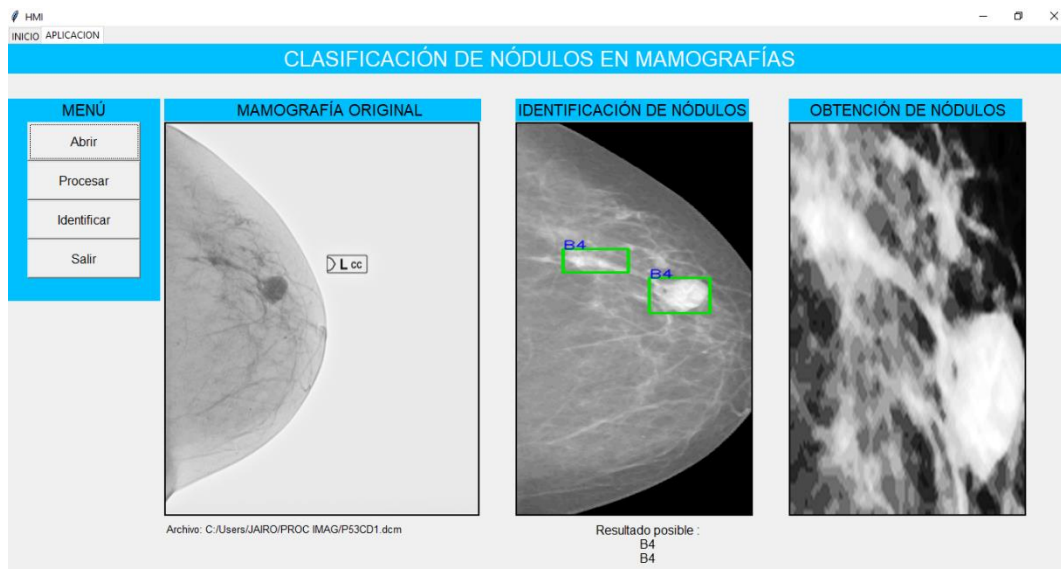


Figura 94

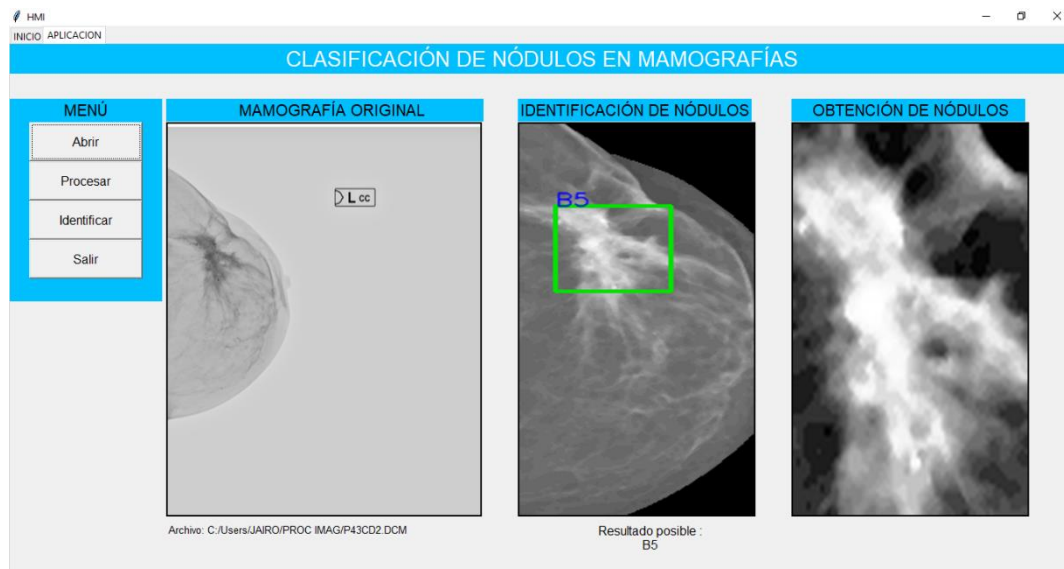
Diagnóstico Proporcionado por el Sistema al Paciente 4: Imágenes de Categorías BIRADS 4, Proyección CC



El diagnóstico médico para el paciente 5 es: BIRADS 5 (Imagen hiperdensa, espiculada de 45 x 40 mm más calcificaciones amorfas, localizada en el cuadrante superior externo de la mama izquierda.). El sistema analiza la mamografía con proyección CC, Figura 95, identifica un nódulo y lo clasifica como BIRADS 5. Por ende, concuerdan los resultados.

Figura 95

Diagnóstico Proporcionado por el Sistema al Paciente 5: Imágenes Nodulares de Categoría BIRADS 5, Proyección CC



Para finalizar el sistema tiene un buen porcentaje de reconocimiento y clasificación de nódulos mamarios, pero este porcentaje depende de la cantidad de tejido que presente la mama, es decir al tratar casos con porciones altas de tejido glandular y tejido conectivo fibroso el porcentaje disminuirá, ya que tiende a reconocer este tejido como nódulos y en otros casos si el tejido está bordeando al nódulo el sistema no reconoce la masa y lo categoriza como BIRADS 0.

El sistema clasifica los nódulos en función de sus características, las cuales varían según la proyección en la que fue realizada la mamografía, este efecto se puede notar en la base de datos CBIS-DDSM en la cual se encuentra un mismo nódulo con diferente diagnóstico. Por esto es importante para los médicos especialistas realizar mamografías con diferentes proyecciones y analizarlas de manera individual para descartar o afirmar los hallazgos, e inclusive en algunas ocasiones es necesario complementar con otros exámenes para emitir un diagnóstico.

4.4 Comprobación de la hipótesis

Para comprobar la hipótesis propuesta se realiza un análisis dividido en dos partes.

La primera parte consiste en evaluar el proceso de reconocimiento de nódulos mamarios el cual se realiza en la sección 4.1, para esto se utilizó el diagnóstico médico de algunos casos tratados por el centro médico Laser. Al contrastar el diagnóstico con los resultados de la etapa se visualiza un buen porcentaje de reconocimiento de masas por parte del sistema, aunque estos resultados pueden variar según la proyección de la mamografía analizada.

La segunda parte es evaluar el modelo entrenado de la red neuronal convolucional, este análisis se lo realizó en el apartado 4.2 en el cual se hizo predicciones con 60 imágenes nuevas divididas en 3 casos, obteniendo un porcentaje de accuracy del sistema de 0.833 (83.3%), así se concluye que el modelo entrenado tiene un alto grado de exactitud al clasificar imágenes nuevas.

4.5 Alcances

- El sistema desarrollado realiza el reconocimiento y clasificación de nódulos mamarios, convirtiéndose en un asistente que ayuda al especialista en imagenología al momento de emitir su diagnóstico.
- La interfaz gráfica desarrollada se la puede utilizar de manera fluida, posee etiquetas que proporcionan información al usuario para evitar errores en su manipulación, presenta solo la información requerida, por esto se la describe como simple, clara y concisa. Su función es mostrar los resultados de la clasificación de los nódulos detectados en una mamografía de formato DICOM.

4.6 Limitaciones

- El sistema depende mucho de la densidad de las mamas, es decir existen casos en los cuales no se obtendrá una respuesta concisa y será categorizada como BIRADS 0.
- Las características de los nódulos mamarios pueden variar dependiendo de la proyección en la que se realiza la mamografía, por ende, el sistema emitirá un diagnóstico acorde a la imagen tratada. Para evitar este efecto el médico imagenólogo revisa las mamografías con diferentes proyecciones y así afirma o descarta ciertos casos.
- Para el entrenamiento de la red neuronal convolucional se requiere de un equipo computacional con un procesador avanzado y tarjeta gráfica de muy buena calidad estos pueden llegar a ser muy costosos.

5 Conclusiones y Recomendaciones

5.1 Conclusiones

- El sistema de clasificación y reconocimiento de nódulos en imágenes mamográficas utilizando procesamiento digital de imágenes y redes neuronales convolucionales, permite extraer y categorizar masas de acuerdo al método de estandarización BIRADS, para esto la mamografía debe estar en formato DICOM y poseer metadatos que permita conocer la proyección en la que fue realizada.
- El sistema se desarrolló totalmente en software libre, para esto se utilizó la plataforma Anaconda Navigator la cual administra aplicaciones, paquetes, entornos, dedicados al desarrollo de lenguaje Python. Como editor de código se utiliza Jupyter Notebook y sus diferentes librerías para realizar: análisis de datos, procesamiento digital de imágenes, redes neuronales, interfaces gráficas, etc.
- La etapa de procesamiento digital de mamografías realiza el reconocimiento de nódulos mamarios, y extrae el área en la cual se encuentran las masas localizadas. La eficiencia de esta etapa está limitada por la cantidad de tejido y otras lesiones mamarias presentes en la mamografía. En esta etapa se detecta si una mama es densa o carece de nódulos, emitiendo el diagnóstico de BIRADS 0 Y BIRADS 1 respectivamente.
- A los nódulos mamarios se les otorga una categoría BIRADS en base a las características que presentan, siendo las más importantes: forma, márgenes, densidad. Acorde a la información y a la cantidad de imágenes recopiladas, se agrupó las categorías 2 y 3 del BIRADS obteniendo una base de datos con 3 categorías, BIRADS 2-3, BIRADS 4,

BIRADS 5, la cual se utiliza para el entrenamiento de la red neuronal convolucional.

- La red neuronal convolucional implementada consta de 5 capas de convolución, 5 capas de max pooling, dos totalmente conectadas, para la cual se ocupa el método de aprendizaje supervisado. Los resultados de esta red se evaluaron a través de su matriz de confusión, en la cual se obtiene un porcentaje de exactitud de 83% suficientemente bueno para realizar la clasificación de nódulos mamarios.
- La interfaz gráfica del sistema se desarrolló por medio de la librería Tkinter, muestra los resultados de identificación y categorización de nódulos mamarios, que se encuentren en una imagen mamográfica de formato DICOM. La interfaz gráfica es sencilla, proporciona y presenta solo información necesaria.

5.2 Recomendaciones

- Se requiere de una base de datos la cual contenga un gran número de imágenes y todos los casos de las diferentes categorías BIRADS. Además, debe ser confiable y estar avalada por médicos especialistas con el fin de evitar errores al momento de entrenar la red neuronal convolucional.
- Es necesario el uso de un equipo con una buena tarjeta gráfica, ya que la GPU aumenta la velocidad con la que se realiza el proceso de aprendizaje de la red neuronal.
- Para lograr un mejor entrenamiento de la red neuronal se recomienda descartar imágenes que presenten ruido y objetos ajenos al estudio, además se debe balancear el número de imágenes en cada categoría, de esta forma se evitará el sobreajuste del modelo.

- En caso de tener una base de datos con pocas imágenes, se debe aplicar operaciones de rotación, zoom, espejo a las imágenes de entrenamiento, con el fin de aumentar imágenes y mejorar el porcentaje de exactitud del modelo.

REFERENCIAS BIBLIOGRÁFICAS

- Agur, A., Dalley, A., & Moore, K. (2019). *Moore Fundamentos de anatomía con orientación clínica*. Wolters Kluwer.
- Álvarez G., D. A., Guevara G., M. L., & Holguín L., G. A. (2006). *Preprocesamiento de Imágenes Aplicadas a Mamografías Digitales*, 2-4. Universidad Tecnológica de Pereira.
- Anaconda. (2020). *Anaconda*. Recuperado el 10 de febrero de 2020, de Anaconda Documentation: <https://docs.anaconda.com/anaconda/navigator/>
- Artola Moreno, Á. (2019). *Clasificación de imágenes usando redes neuronales*. Universidad de Sevilla.
- Basogain Olabe, X. (1998). *Redes Neuronales artificiales y sus aplicaciones*. Escuela Superior de Ingeniería de Bilbao, EHU.
- Berrones Reyes, M. C. (2019). *Clasificación de mamografías mediante redes neuronales convolucionales*, 68. Universidad Autónoma De Nuevo León.
- Bóveda Álvarez, M. C. (2014). *Optimización paramétrica de técnicas avanzadas de segmentación de imágenes biomédicas: una contextualización a la patología*. Universidad de Coruña.
- Carrasco LLerena, C. A. (2017). *Reconocimiento de cáncer de mama utilizando técnicas de procesamiento digital de imágenes mamográficas y aprendizaje de máquina*. Carrera de Ingeniería en Electrónica y Telecomunicaciones. Universidad de las Fuerzas Armadas ESPE. Recuperado el 17 de noviembre de 2019
- Carrero, G. (2007). *Caracterización de imágenes nodulares en mamografía por medio de dimensión fractal*, 129. Universidad Central de Venezuela. Recuperado el 18 de septiembre de 2019, de https://www.researchgate.net/profile/Miguel_Martin-Landrove/publication/258516676_Caracterizacion_de_Imagenes_Nodulares_en_

Mamografia_por_Medio_de_Dimension_Fractal/links/004635288087d28b4d000000/Caracterizacion-de-Imágenes-Nodulares-en-Mamografia-por-Med

Ciardullo, S. F. (s.f.). *Diagnóstico por Imágenes – Imágenes Mamarias*. Recuperado el 6 de octubre de 2019, de https://diagnosticoporimagenes.weebly.com/uploads/9/4/6/5/94652816/apunte_imagenesmama.pdf

Clavero, A., Molina, E., & Piccinini, P. (2016). *Calidad en estudios de Mamografía*. Ingeniería Biomédica, Universidad Nacional de Córdoba.

Costa Campos, A., & Fernández Bozal, J. (2005). *La imagen digital*. Recuperado el 8 de noviembre de 2019, de http://www.revistadeortodoncia.es/files/2005_35_3_255-266.pdf

GCO. (2018). *Global Cancer Observatory*. Recuperado el 3 de octubre de 2019, de GLOBOCAN: <https://gco.iarc.fr/today/>

Gómez, D., & Guerrero, A. (2016). *Estudio y análisis de técnicas para procesamiento digital de imágenes*. Colombia: Universidad tecnológica de Pereira.

Google Developers. (20 de julio de 2020). *Google Developers*. Recuperado el 15 de febrero de 2020, de <https://developers.google.com/machine-learning/practica/image-classification/convolutional-neural-networks?hl=es-419>

Hernández, S. (2018). *Segmentación de masas en imágenes de mamografía mediante agrupamiento restringido.*, 13. Universidad Tecnológica de la Mixteca.

Holden , S. (2018). *Python*. Recuperado el 13 de noviembre de 2019, de <https://wiki.python.org/moin/FrontPage>

InChaos. (s.f.). *Redes Neuronales Artificiales*. Recuperado el 24 de noviembre de 2019, de <http://rna.50webs.com/>

INEC. (Octubre de 2017). Recuperado el 15 de septiembre de 2019, de <https://www.ecuadorencifras.gob.ec/el-cancer-de-mama-en->

- Obenauer. (2005). *Applications and literature review of the BI-RADS classification [Tabla]*. PubMed.
- Obenauer, S., Hermann, K. P., & Grabbe, E. (15 de 05 de 2005). Applications and literature review of the BI-RADS classification. *PubMed*. Recuperado el 10 de diciembre de 2019, de <https://pubmed.ncbi.nlm.nih.gov/15856253/>
- OMS. (Septiembre de 2018). *Cáncer de mama*. Organización Mundial de la Salud. Recuperado el 3 de agosto de 2019
- Ortiz Suárez, J. M. (2016). *Evaluación de métodos con redes neuronales convolucionales para la clasificación de la enfermedad de alzheimer en imágenes de resonancia magnética cerebral*, 14-15. Universidad Industrial De Santander.
- Quintanilla Domínguez, J. (2015). *Aplicación de técnicas de inteligencia artificial para contribuir en la detección de microcalcificaciones en mamografía digitalizada*. Universidad Politécnica de Madrid.
- Reyes López, M. (2015). *Implementación y evaluación de dos metodologías para la detección de masas espiculadas en imágenes de mamografías*, 9-10. Universidad Tecnológica de la Mixteca.
- Rodríguez López, V. (2012). *Análisis de imágenes de mamografía para la detección de cáncer de mama*, 41. Universidad Tecnológica de la Mixteca.
- Salas, R. (2004). *Redes neuronales artificiales*. Universidad de Valparaíso, Departamento de Computación.
- Sawyer Lee, R., Giménez, F., Hoogi, A., & Rubin, D. (2016). The Cancer Imaging Archive. *Subconjunto de imágenes mamarias curadas de DDSM [Dataset]*. TCIA.
- Shinde, M. (2003). *Computer aided diagnosis in digital mammography: Classification of mass and normal tissue*. University of South Florida.
- Simeone, O. (2018). *A Very Brief Introduction to Machine Learning With Applications to Communication Systems*. IEEE. doi:10.1109/TCCN.2018.2881442

TensorFlow. (2016). *TensorFlow*. Recuperado el 20 de agosto de 2019, de <https://www.tensorflow.org/about>

Tortora, G. J., & Derrickson, B. (2013). *Principios de Anatomía y Fisiología*. Panamericana.

Zamorano Ruiz, J. (2019). *Comparación y análisis de métodos de clasificación con las bibliotecas scikit-learn y TensorFlow en Python*. Universidad de Málaga.

ANEXOS