



**FACULTAD DE INGENIERÍA, ARQUITECTURA Y
URBANISMO**

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

TESIS

**Reconocimiento de microcalcificaciones en mamografías
haciendo uso de una máquina de aprendizaje**

**PARA OPTAR EL TÍTULO PROFESIONAL
DE INGENIERO DE SISTEMAS**

Autor

**Bach. Wildor Alexander Flores Vasquez
<https://orcid.org/0009-0002-5542-4751>**

Asesor

**Dr. Jorge Luis Gutiérrez Gutiérrez
<https://orcid.org/0000-0002-4989-1196>**

**Línea de Investigación
Infraestructura, Tecnología y Medioambiente**

Pimentel – Perú

**RECONOCIMIENTO DE MICROCALCIFICACIONES EN MAMOGRAFÍAS
HACIENDO USO DE UNA MÁQUINA DE APRENDIZAJE**

Aprobación del jurado

DR. JORGE LUIS GUTIÉRREZ GUTIÉRREZ

Presidente del Jurado de Tesis

MG. HEBER IVAN MEJIA CABRERA

Secretario del Jurado de Tesis

DR. VICTOR ALEXCI TUESTA MONTEZA

Vocal del Jurado de Tesis


DECLARACIÓN JURADA DE ORIGINALIDAD

Quien suscribe la DECLARACIÓN JURADA, soy egresado (s) **Wildor Alexander Flores Vasquez**, del Programa de Estudios de **Ingeniería de Sistemas** de la Universidad Señor de Sipán S.A.C, declaro bajo juramento que soy autor del trabajo titulado:

**RECONOCIMIENTO DE MICROCALCIFICACIONES EN MAMOGRAFÍAS HACIENDO
USO DE UNA MÁQUINA DE APRENDIZAJE**

El texto de mi trabajo de investigación responde y respeta lo indicado en el Código de Ética del Comité Institucional de Ética en Investigación de la Universidad Señor de Sipán, conforme a los principios y lineamientos detallados en dicho documento, en relación con las citas y referencias bibliográficas, respetando el derecho de propiedad intelectual, por lo cual informo que la investigación cumple con ser inédito, original y auténtico.

En virtud de lo antes mencionado, firman:

WILDOR ALEXANDER FLORES VASQUEZ	DNI: 75342738	
------------------------------------	---------------	---

Pimentel, 18 de julio de 2024

Dedicatoria

A mis padres Flor del Carmen y José Antero, hermanos Dayro y Andy y a mis amigos porque me otorgan la fuerza y aliento para conseguir nuevos objetivos.

Agradecimientos

Agradezco al Dr. Jorge Luis Gutiérrez Gutiérrez por facilitarme la ayuda de especialista además por el seguimiento y la supervisión de este trabajo de investigación, también agradezco al Dr. Víctor Tuesta y a la Mg. Rosa América Cobeñas por la orientación y sugerencias recibidas durante las sesiones de Habilidades Investigativas.

índice

Dedicatoria	IV
Agradecimientos	V
Índice de Tablas, figuras y fórmulas:.....	VII
Abstract	XI
I. INTRODUCCIÓN	13
1.1. Realidad problemática.....	13
1.2. Formulación del problema	16
1.3. Hipótesis.....	17
1.4. Objetivos.	17
1.5. Teorías relacionadas al tema	18
II. MATERIAL Y MÉTODO	53
2.1. Tipo y Diseño de Investigación:	53
2.2. Variables, Operacionalización.....	53
2.3. Población y muestra de estudio	56
2.4. Técnicas e instrumentos de recolección de datos	57
2.5. Procedimiento de análisis de datos:	58
2.6. Criterios éticos.....	60
2.7. Criterios de Rigor Científico.....	60
III. RESULTADOS Y DISCUSIÓN	61
3.1. Resultados	61
3.2. Discusión	96
IV. CONCLUSIONES Y RECOMENDACIONES	105
4.1. Conclusiones:	105
4.2. Recomendaciones:	107
REFERENCIAS.....	108
ANEXOS:	110

Índice de Tablas, figuras y fórmulas:

Tabla 1. BI-RADS Mammography Lexicon Overview.....	20
Tabla 2. Diferencia entre microcalcificaciones.....	44
Tabla 3. Operacionalización de variables.	54
Tabla 4. Método propuesto dado los resultados 3	72
Tabla 5. Matriz de confusión: aplicando segmentación piramidal con técnicas de preprocesamiento.....	73
Tabla 6. Matriz de confusión: aplicando Segmentación Thresholding con técnicas de Preprocesamiento (Método Propuesto)	74
Tabla 7. Resultados de la extracción de características	83
Tabla 8. Capa de entrada – época 1	90
Tabla 9. Capa de entrada – época 2	90
Tabla 10. Capa de entrada – época 3	91
Tabla 11. Capa oculta – época 1	92
Tabla 12. Capa oculta – época 2.....	92
Tabla 13. Capa oculta – época 3.....	93
Tabla 14. Capa de salida – época 1	94
Tabla 15. Capa de salida – época 2	94
Tabla 16. Capa de salida – época 3	95
Tabla 17. Sensibilidad	96
Tabla 18. Especificidad	97
Tabla 19. Valor predictivo positivo.....	98
Tabla 20. Valor predictivo negativo.....	99
Tabla 21. Razón de verosimilitudes positivas	100
Tabla 22. Razón de verosimilitudes negativas.....	101
Tabla 23. Precisión.....	102
Tabla 24. Ratio o razón de falsos descubrimientos.....	103

Tabla 25. Ratio o razón de falsos positivos	104
---	-----

Índice de Figuras:

Figura 1. Procedimiento de evaluación y diagnóstico mamográfico.....	14
Figura 2. Casos registrados de cáncer de localización topográfica y sexo en la región Lambayeque.....	14
Figura 3. Masa micro-lobulada.	24
Figura 4. Masa indistinta.....	25
Figura 5. Masa espiculada.....	26
Figura 6. Masa de alta densidad.....	27
Figura 7. Masa de igual densidad.....	28
Figura 8. Masa de baja densidad.....	29
Figura 9. Masa contenedora de grasa.	30
Figura 10. Microcalcificación típicamente benigna de piel (a).	31
Figura 11. Microcalcificación típicamente benigna de piel (b).	32
Figura 12. Microcalcificación típicamente benigna vascular (a).	33
Figura 13. Microcalcificación típicamente benigna vascular (b).	33
Figura 14. Microcalcificación típicamente Grueso o “como palomita de maíz” (popcorn-like) (a).	34
Figura 15. Microcalcificación típicamente Grueso o “como palomita de maíz” (popcorn-like) (b).	34
Figura 16. Microcalcificación típicamente benigna larga en forma de varilla (a).....	35
Figura 17. Microcalcificación típicamente benigna larga en forma de varilla (b).....	36
Figura 18. Microcalcificación típicamente benigna redondas.	37
Figura 19. Microcalcificación típicamente benigna bordeada.	38
Figura 20. Microcalcificación típicamente benigna distrófico.....	39
Figura 21. Microcalcificación típicamente benigna leche de calcio.	39

Figura 22. Microcalcificación típicamente benigna saturada.	40
Figura 23. Microcalcificación típicamente maligna amorfo.	41
Figura 24. Microcalcificación típicamente maligna grueso heterogéneo.	42
Figura 25. Microcalcificación típicamente maligna pleomorfo fino.....	43
Figura 26. Microcalcificación típicamente maligna bifurcación lateral fina.....	44
Figura 27. Mamografía digital.	46
Figura 28. Histograma para los niveles de intensidad de una imagen, con L= 64....	47
Figura 29. Diagrama de flujo k-means.	50
Figura 30. Proceso de backpropagation.	52
Figura 31. Métodos propuestos.	63
Figura 32. Gráfico Total de error de la red en Neuroph Studio.....	64
Figura 33. Total MSE en Neuroph Studio.	64
Figura 33. Total de error de la red en Neuroph Studio – resultados 2.....	65
Figura 35. Total MSE en Neuroph Studio - resultados 2.....	66
Figura 36. Total de error de la red en Neuroph Studio – resultados 3.....	66
Figura 37. Total MSE en Neuroph Studio - resultados 3.....	67
Figura 38. Método 2 - Gráfico Total de error de la red en Neuroph Studio.....	68
Figura 39. Método 2 - Total MSE en Neuroph Studio.	69
Figura 33. Método 2 - Total de error de la red en Neuroph Studio – resultados 2....	69
Figura 41. Método 2 - Total MSE en Neuroph Studio - resultados 2.....	70
Figura 42. Método 2 - Total de error de la red en Neuroph Studio – resultados 3....	70
Figura 43. Método 2 - Total MSE en Neuroph Studio - resultados 3.	71
Figura 44. Preprocesamiento del método propuesto.	78
Figura 45. Segmentación piramidal.	79
Figura 46. Resultado de la segmentación.....	79
Figura 47. Diagrama de flujo de trabajo de la extracción de la ROI.	80
Figura 48. Diseño de la Red Neuronal bajo el Framework Neuroph.	85

Figura 49. Diseño de la Red Neuronal bajo el Framework Neuroph – Capa de Entrada	87
Figura 50. Diseño de la Red Neuronal bajo el Framework Neuroph – Capa de Salida.	87
Figura 51. Formato txt de las características extraídas.....	88
Figura 52. Código Java Aplicando Neuroph – Almacenando el DataSet.....	88
Figura 53. Código Java Aplicando Neuroph – Especificando Parámetros para el entrenamiento de la RNA.	89
Figura 54. Código Java Aplicando Neuroph – Arrojando Resultados del Entrenamiento.....	89
Figura 55. Porcentaje de sensibilidad.....	97
Figura 56. Porcentaje de especificidad.....	98
Figura 57. Valor predictivo positivo.....	99
Figura 58. Valor predictivo negativo.	100
Figura 59. Razón de verosimilitudes positivas.....	101
Figura 60. Razón de verosimilitudes negativas.....	102
Figura 61. Precisión.....	103
Figura 62. Ratio o razón de falsos descubrimientos.	104
Figura 63. Ratio o razón de falsos positivos.	105

Abstract

Breast cancer is one of the most prevalent neoplasms in Peru, ranking third in incidence after gastric cancer, with an annual average of 11,340 cases reported by the Ministry of Health (MINSA). The National Institute of Neoplastic Diseases (INEN), specializing in oncology, currently relies on the knowledge and expertise of oncologists, as detailed in its Manual of Procedural Processes for Medical Assistance. Types of breast cancer are classified into various patterns, one of which includes microcalcifications, a primary indicator of the disease.

This research study aims to demonstrate that it is possible to support the diagnosis of breast cancer by identifying microcalcifications in mammographic images using artificial vision techniques. For this purpose, 36 images were selected from a population of 322, to which preprocessing algorithms were applied. Subsequently, the necessary features were automatically extracted to classify them using an artificial neural network. The analysis considered the results of preprocessing, processing, and classification, employing indicators such as sensitivity percentage, specificity percentage, positive predictive value, and negative predictive value. These results were compared with similar research, achieving an accuracy of 80.55%, highlighting the significance of this method and its potential for future research or continuous application.

Keywords: Breast cancer, Microcalcifications, Artificial vision, Artificial neural network, Assisted diagnosis.

Resumen

El cáncer de mama es una de las neoplasias más prevalentes en Perú, ocupando el tercer lugar en incidencia después del cáncer gástrico, con un promedio anual de 11,340 casos reportados por el Ministerio de Salud (MINSA). El Instituto Nacional de Enfermedades Neoplásicas (INEN), especializado en oncología, actualmente depende del conocimiento y la experiencia de oncólogos, como se detalla en su Manual de Procesos Procedimentales para la Asistencia Médica. Los tipos de cáncer de mama se clasifican en varios patrones, uno de los cuales incluye las microcalcificaciones, un indicador primario de la enfermedad.

Este estudio de investigación tiene como objetivo demostrar que es posible apoyar el diagnóstico del cáncer de mama identificando microcalcificaciones en imágenes mamográficas utilizando técnicas de visión artificial. Para este fin, se seleccionaron 36 imágenes de una población de 322, a las cuales se aplicaron algoritmos de preprocesamiento. Posteriormente, se extrajeron automáticamente las características necesarias para clasificarlas utilizando una red neuronal artificial. El análisis consideró los resultados de preprocesamiento, procesamiento y clasificación, empleando indicadores como el porcentaje de sensibilidad, porcentaje de especificidad, valor predictivo positivo y valor predictivo negativo. Estos resultados se compararon con investigaciones similares, logrando una precisión del 80.55%, destacando la importancia de este método y su potencial para investigaciones futuras o aplicación continua.

Palabras clave: Cáncer de mama, Microcalcificaciones, Visión artificial, Red neuronal artificial, Diagnóstico asistido.

I. INTRODUCCIÓN

1.1. Realidad problemática

En la actualidad, el cáncer de mama se ha identificado como una de las principales causas de mortalidad en mujeres, representando un significativo problema de salud pública. Por ello, es fundamental implementar planes efectivos de detección temprana para prevenir y controlar esta patología. La mamografía es una de las herramientas más tradicionales y efectivas en la detección del cáncer de mama. En mujeres posmenopáusicas, las microcalcificaciones mamarias, pequeños depósitos de calcio en el tejido mamario, son hallazgos recurrentes y cruciales en las imágenes mamográficas [1].

La evaluación temprana de esta enfermedad permite que las mujeres afectadas reciban tratamiento antes de que el cáncer progrese, previniendo así posibles complicaciones [2]. Los especialistas en oncología consideran que la cantidad mínima significativa de microcalcificaciones para destacar en un diagnóstico debe ser de cuatro a seis, ya que un número menor rara vez conlleva a la detección de cáncer de mama [3].

Actualmente el Instituto Nacional de Enfermedades Neoplásicas (INEN) usa la experiencia del médico especialista para el procesamiento de imágenes médicas, en la figura 1 se aprecia el Procedimiento, evaluación y diagnóstico mamográfico.

PROCEDIMIENTO: EVALUACIÓN Y DIAGNÓSTICO MAMOGRAFICO		PROCESO: EVALUACIÓN DE IMAGENES EN RADIOLOGIA		SUBPROCESO: RADIOLOGIA			
Nº Ord	Tipo Act	Descripción de Actividad	Producto o Resultado	Responsable	Unidad Orgánica	Indicador	Procedimiento relacionado
1	I	Recabar H.C. de Pacientes citados y entrega a Secretaria	H.C. del Paciente	Técnico Administrativo	Dpto. de Radiodiagnóstico - Archivo		Procedimiento Programación y Evaluación
2	D	Recepción del Paciente, verificación de requisitos, registro de turno de atención y entrega de listados a los responsables de las tomas de exámenes mamográficos a) Si el paciente es derivado de Emergencia, hacerlo pasar de inmediato, recabar historia clínica, registrar su atención y comunicar al E/la Director/a Ejecutivo/a del Departamento o Médico de Guardia b) Si el Paciente no ha sido derivado de Emergencia y/o Hosp. verificar la cita programada y registrar su asistencia y turno de atención, quedando en espera	Listado de Atención de mamografías e H.C. de Paciente de Emergencia	Secretaria.	Dpto. de Radiodiagnóstico		Procedimiento Programación y Evaluación
3	O	Llamado, preparación, información y orientación del Paciente sobre el examen.	Paciente preparado	Tecnólogo Médico en Radiología y/o Técnica en Radiología	Dpto. de Radiodiagnóstico		
4	O	Preparación del Equipo de Rayos X y el Material Placas y Cassetes para toma de examen mamográfico. C17	Equipo y placas listas	Tecnólogo Médico en Radiología y/o Técnica en Radiología	Dpto. de Radiodiagnóstico		
5	O	Preparación y ubicación del Paciente para la toma del examen mamográfico en Equipo Radiográfico.	Paciente en posición para toma de examen	Tecnólogo Médico en Radiología y/o Técnica en Radiología	Dpto. de Radiodiagnóstico		
6	O	Toma de placa y/o imagen digitalizada del examen mamográfico en la posición y topografía especificada, según las indicaciones y especificaciones requeridas.	Placa tomada y/o imagen digitalizada	Tecnólogo Médico en Radiología y/o Técnica en Radiología	Dpto. de Radiodiagnóstico		
7	O	Procesamiento de revelado automático de placas mamográficas y/o visualización de imagen digitalizada. En caso de imagen difusa, repetir toma, registrar la placa desechada y retomar a la Actividad N° 6. En caso que la imagen no presente defectos continuar	Placa procesada y/o imagen digitalizada verificada, codificada y archivada	Tecnólogo Médico en Radiología y/o Técnica en Radiología	Dpto. de Radiodiagnóstico		
8	O	Comunicar al Paciente que se vista e invitarlo a que se retire, con la orientación sobre el procedimiento.	Paciente orientado e informado	Tecnólogo Médico en Radiología y/o Técnica en Radiología	Dpto. de Radiodiagnóstico		
9	O	Registro y remisión al Archivo de Radiodiagnóstico de placas y/o impresión de imagen mamográfica tomada y reporte de placas utilizadas y desechadas.	Registro de entrega de Placa	Tecnólogo Médico en Radiología y/o Técnica en Radiología	Dpto. de Radiodiagnóstico		
10	O	Registro de recepción de Placa mamográfica y/o impresión de imagen mamográfica del paciente	Registro de Archivo	Tecnólogo Médico en Radiología y/o Técnica en Radiología	Dpto. de Radiodiagnóstico		

Manual de Procedimientos de Procesos Asistenciales del INEN

Página 57 de 461

Versión 01B

Dr. JORGE HUACHTAY SINTOS
C.O.P. 15072
Inst. de Enfermedades Neoplásicas
"Dr. EDUARDO CASERES GRAJALCA"



Figura 1. Procedimiento de evaluación y diagnóstico mamográfico.

Fuente: Manual de Procedimientos de Procesos Asistenciales del INEN.

En la Figura 2, se evidencia que esta enfermedad puede afectar tanto a hombres como a mujeres, aunque su incidencia es significativamente mayor en mujeres. Estudios indican que una de cada diez mujeres desarrollará esta patología a lo largo de su vida.

LOCALIZACIÓN TOPOGRÁFICA	MASCULINO	FEMENINO	TOTAL	%
Cérvix	-	1565	1565	16.2
Mama	3	1,016	1019	10.5
Estómago	493	519	1012	10.5
Próstata	582	-	582	6.0
Piel	247	246	493	5.1
Pulmón	242	220	462	4.8
Hematológico	228	185	413	4.3
Hígado y vías biliares intrahepáticas	158	216	374	3.9
Otros	1621	2132	3753	38.8
TOTAL	3574	6099	9673	100.0

Figura 2. Casos registrados de cáncer de localización topográfica y sexo en la región Lambayeque.

Fuente: Ministerio de Salud – Dirección General de Epidemiología.

El cáncer de mama en Perú ha representado un problema de gran relevancia en la salud pública durante los últimos años. Muchas personas carecen de los

recursos económicos necesarios para acceder a atención médica temprana, lo que a menudo resulta en diagnósticos tardíos que dificultan el tratamiento efectivo de la enfermedad.

Los primeros esfuerzos por registrar enfermedades mamarias en una placa radiológica están próximos a cumplir un siglo. Fue un cirujano quien inicialmente trabajó con piezas de mastectomía, aunque su investigación no se aplicó en la práctica clínica. Durante más de la mitad de ese tiempo, la radiografía mamaria quedó relegada a trabajos artesanales que solo buscaban demostrar cómo se manifestaban las lesiones mamarias en la escala de grises, ya confirmadas por el tacto y la vista. Una vez logrado esto, el desafío se trasladó a detectar las lesiones antes que el clínico, utilizando no solo rayos X, sino también ondas sonoras, cambios de temperatura, electrones, entre otros métodos. Aún más ambicioso fue el objetivo de predecir la respuesta a diferentes [4].

Con referencia a lo previamente mencionado, la relevancia del problema es equiparable a la necesidad de soluciones para la prevención o diagnóstico. Recientemente se han desarrollado investigaciones para apoyar en el diagnóstico asistido por computadora para la identificación del cáncer de mama mediante imágenes [5] [3] [6] [7] , utilizando datos genéticos [8], y utilizando diferentes técnicas, como redes neuronales convolucionales [9] [10] .

La investigación titulada "Sistema de auxílio na avaliação de calcificações mamárias por processamento digital de imagens e inteligência artificial" (sistema de ayuda en la evaluación de calcificaciones mamarias mediante procesamiento digital de imágenes e inteligencia artificial), emplea técnicas de procesamiento de imágenes para entrenar una red neuronal en el software

MatLab, logrando resultados exitosos en la clasificación de las calcificaciones [11].

Asimismo, la investigación titulada "Consistent performance measurement of a system to detect masses in mammograms based on blind feature extraction", se centra en la extracción de características ciegas a través del procesamiento digital de imágenes, seguidas por su clasificación mediante redes neuronales y SVM, obteniendo resultados positivos. Estas investigaciones demuestran el interés en mejorar los resultados y apoyar el diagnóstico utilizando herramientas de procesamiento digital de imágenes e inteligencia artificial [12]. En la investigación [13], se estudia un algoritmo de aprendizaje conjunto basado en una máquina de vectores de soporte (SVM) para el diagnóstico del cáncer de mama. El estudio [14] utilizó dos tipos de redes neuronales convolucionales (CNN) para mejorar el diagnóstico del cáncer de mama utilizando imágenes de mamografía: Una CNN profunda con imágenes de Contraste Mejorado de Mamografía Digital (CEDM) de Mayo Clinic mostró una precisión de diagnóstico del 85% y un área bajo la curva (AUC) de 0.84 con imágenes de LE, y del 89% de precisión y AUC de 0.91 con imágenes tanto de LE como recombinadas. Una CNN superficial entrenada con los mismos casos de CEDM aprendió a mapear de manera no lineal imágenes de LE a imágenes recombinadas, mejorando la precisión diagnóstica a 90% (AUC = 0.92) utilizando una base de datos pública de Mamografía Digital de Campo Completo (FFDM) llamada INbreast.

El presente proyecto tiene como objetivo reconocer microcalcificaciones en mamografías mediante el procesamiento de imágenes del tejido mamario.

1.2. Formulación del problema

¿Cómo puede una máquina de aprendizaje identificar automáticamente las microcalcificaciones en mamografías digitales que son indicativas de la presencia de tumores?

1.3. Hipótesis

La aplicación de clasificadores basados en máquinas de aprendizaje permitirá una identificación automática de las microcalcificaciones en mamografías digitales, mejorando la detección temprana de tumores mamarios.

1.4. Objetivos.

Objetivo general

Reconocimiento de microcalcificaciones en mamografías haciendo uso de una máquina de aprendizaje en análisis de imágenes.

Objetivos específicos

- a) Clasificar las imágenes de la base de datos utilizadas.
- b) Aplicar técnicas de pre-procesamiento en las imágenes mamográficas.
- c) Aplicar técnicas de segmentación
- d) Aplicar algoritmos para la identificación de las características a detectar en las imágenes mamográficas.
- e) Clasificar y evaluar los resultados.

1.5. Teorías relacionadas al tema

1.5.1. Cáncer de Mama

El cáncer de mama abarca un grupo de enfermedades biológicamente y molecularmente diversas que se originan en el tejido mamario. Aunque los factores de riesgo asociados a este tipo de cáncer pueden diferir de los de otros tipos de cáncer, la predisposición genética, especialmente las mutaciones en los genes BRCA1 y BRCA2, constituye un factor causal significativo para esta neoplasia. Los cánceres de mama pueden iniciarse en distintas áreas de la mama, como los conductos, los lobulillos o el tejido intersticial. Dentro del amplio grupo de carcinomas mamarios, existen varios subtipos que se distinguen por su grado de invasividad y la ubicación del tumor primario. Es crucial identificar los diferentes subtipos debido a sus variaciones en pronóstico y en las estrategias terapéuticas recomendadas [15].

Cáncer de mama tipo Luminal

Los tumores Luminal A constituyen entre el 50% y el 60% de los casos, caracterizándose por altos niveles de expresión de genes regulados por el receptor estrogénico (RE) y bajos niveles de expresión de genes asociados con la proliferación celular. Estos tumores suelen mostrar una expresión génica típica del epitelio luminal que recubre los ductos mamarios. De acuerdo con su perfil molecular, la mayoría de los casos de carcinoma lobulillar in situ se clasifican como Luminal A, y una proporción significativa de carcinomas lobulillares infiltrantes también presentan este perfil característico. Los tumores Luminal A generalmente tienen un bajo grado histológico y un pronóstico favorable. Desde una perspectiva inmunohistoquímica, el cáncer de mama Luminal A se distingue por la expresión

del RE, del receptor de progesterona (RPg), de Bcl-2 y de las citoqueratinas CK8/18, así como por la ausencia de sobreexpresión de HER2. Además, el factor de transcripción GATA3 se expresa en el tipo Luminal, con niveles particularmente altos en el subtipo Luminal A. En contraste, los tumores Luminal B, que constituyen entre el 10% y el 20% de los casos, pueden exhibir un fenotipo más agresivo, caracterizado por un mayor grado histológico y un índice proliferativo elevado, lo que se traduce en un pronóstico significativamente menos favorable en comparación con los tumores Luminal A [16].

Cáncer de mama tipo HER2

Este tipo de variante presenta el genotipo de RE negativos y se asocia con la sobreexpresión de genes relacionados con un alto grado de agresividad biológica, lo que sufre un peor pronóstico. Este viene a presentar una resistencia relativa a los determinados agentes quimioterápicos, como el esquema clásico CMF (Ciclofosfamida, Metotrexate y Fluoruracilo) y al tamoxifeno. La historia natural ha venido sufriendo una variación en el mercado del anticuerpo monoclonal anti-HER-2, denominado trastuzumab [17].

Cáncer de mama Basal-like

El fenotipo triple negativo (FTN) es denominado así debido a la falta de expresión de los receptores de estrógeno (RE), los receptores de progesterona (RP) y los receptores HER-2. En aproximadamente el 50% de los casos, los carcinomas mamarios de tipo "Basal-like" presentan una sobreexpresión del receptor del factor de crecimiento epidérmico (EGFR). Estas características histopatológicas suelen estar asociadas con un mal pronóstico, incluyendo un alto grado nuclear, pleomorfismo y un índice mitótico elevado [17].

Cáncer de mama similar a la mama normal / inclasificados

Estos tumores, que constituyen entre el 6% y el 10% de todos los casos de cáncer de mama, están insuficientemente caracterizados. En el dendrograma, estos tumores se agrupan de manera jerárquica con fibroadenomas y muestras de tejido mamario normal, a veces cercanos a carcinomas tipo basal y otras veces ubicados en la rama de tumores negativos para el receptor estrogénico (RE). El concepto clínico de estos tumores, que presentan similitudes con el tejido mamario normal, aún no está completamente definido. Existe la sugerencia de que este subgrupo podría ser el resultado de una notable contaminación con tejido normal. Estudios que han aislado células neoplásicas mediante microdissección no han identificado ningún carcinoma de mama "normal-like", lo que sugiere que este subtipo molecular podría surgir de la mezcla con células normales del parénquima mamario.

1.5.2. Patrones tumorales de mama

Según American College of Radiology(ACR), otorga un compendio establecido y estudiado, se muestra en la siguiente tabla:

Tabla 1. BI-RADS Mammography Lexicon Overview.

Tejido Mamario	Condiciones
Composición de la mama	A. La mama son casi en su totalidad graso. B. Hay áreas dispersas de densidad fibroglandular. C. La mama son de forma heterogénea denso, que puede oscurecer pequeñas masas. D. La mama son extremadamente densa, lo que disminuye la sensibilidad de la mamografía

Hallazgos	Condiciones	
A. Masas	A.1. Forma	A.1.A. Ovalo A.1.B. Redondo A.1.C. Irregular
	A.2. Margen	A.2.A. Circunscrito A.2.B. Oscurecido A.2.C. Micro-lobulado A.2.D. Indistinto A.2.E. Espículada
	A.3. Densidad	A.3.A. Alta densidad A.3.B. Igual densidad A.3.C. Baja densidad A.3.D. Contenedor de grasa (fat-containing)
B. Calcificaciones	B.1. Típicamente benigno	B.1.A. Piel B.1.B. Vascular B.1.C. Grueso o “como palomita de maíz” (popcorn-like) B.1.D. Largo en forma de varilla B.1.E. Redondo B.1.F. Bordeado B.1.G. Distrófico B.1.H. Leche de calcio B.1.I. Suturada

	B.2. Morfología sospechosa	B.2.A. Amorfo B.2.B. Heterogénea gruesa B.2.C. Pleomórfico fina B.2.D. Lineal fino o lineal fino de ramificación
	B.3. Distribución	B.3.A. Difuso B.3.B. Regional B.3.C. Agrupado B.3.D. Lineal B.3.E. Segmentado
C. Distorsiones de la arquitectura		
D. Asimetrías	D.1. Asimetría D.2. Asimetría global D.3. Asimetría focal D.4. Asimetría en desarrollo	
E. Ganglio linfático intramamario		
F. Lesiones en la piel		
G. Conducto dilatado solitario		
H. Características asociadas	H.1. Piel retraída H.2. Pezón retraído H.3. Piel gruesa H.4. Engrosamiento trabecular H.5. Adenopatía axilar H.6. Calcificaciones	

I. Localización de la lesión	I.1. Lateralidad I.2. Cuadrante y sentido de reloj I.3. Profundidad I.4. Distancia desde el pezón
------------------------------	--

Fuente: ACR BI-RADS Mammography [7].

a. Masas

Una masa es de 3 dimensiones y ocupa espacio. Se ve en dos diferentes proyecciones mamográficas. Tiene total o parcialmente fronteras convexas hacia el exterior y (radiodensos) aparece crecidamente densa en el centro que en la periferia. Si una masa potencial se ve sólo en una sola proyección, debe ser llamado una asimetría hasta su 3-dimensionalidades es confirmado [18].

Formas

Ovalo: Una masa ovalada es elíptica o en forma de huevo (puede incluir dos o tres ondulaciones).

Redondo: Una masa redonda es esférica, en forma de bola, circular, o de forma globular.

Irregular: La forma de la masa no es ni redondo ni oval. Para la mamografía, el uso de este descriptor por lo general implica un hallazgo sospechoso.

Margen

El margen es el borde o borde de la lesión. Los descriptores de margen, al igual que los descriptores de forma, son importantes predictores de si una masa es benigna o maligna [18].

Circunscrito: El margen está claramente delineado, con una transición abrupta

entre la lesión y el tejido circundante. En mamografía, si parte del margen está oscurecida, al menos el 75% del margen debe estar bien definido para que una masa sea calificada como circunscrita. Una masa cuya cualquier porción del margen sea indistinta, microlobulada o espiculada debe clasificarse según esta última característica (el componente más sospechoso).

Oscurecido: Un margen oscurecido es aquel que está oculto por tejido fibroglandular superpuesto o adyacente. Este término se utiliza principalmente cuando algunos márgenes de la masa están circunscritos, pero el resto (> 25%) está oculto.

Micro-lobulado: El margen se caracteriza por ondulaciones de ciclo cortos. Para la mamografía, el uso de este descriptor generalmente implica un hallazgo sospechoso.

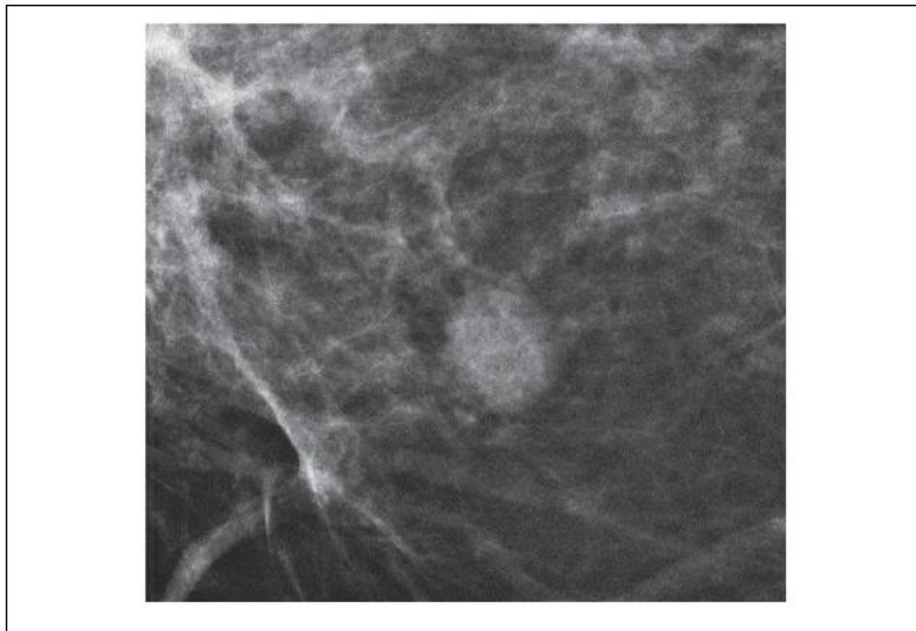


Figura 3. Masa micro-lobulada.

Fuente: ACR BI-RADS Mammography [7].

Indistinto: La demarcación no es claramente visible en ninguna porción o en

todo el margen del tejido circundante. En mamografía, este descriptor no debe utilizarse cuando el radiólogo sospecha que la indistinción se debe al tejido mamario adyacente inmediato. Generalmente, el uso de este descriptor implica un hallazgo sospechoso.

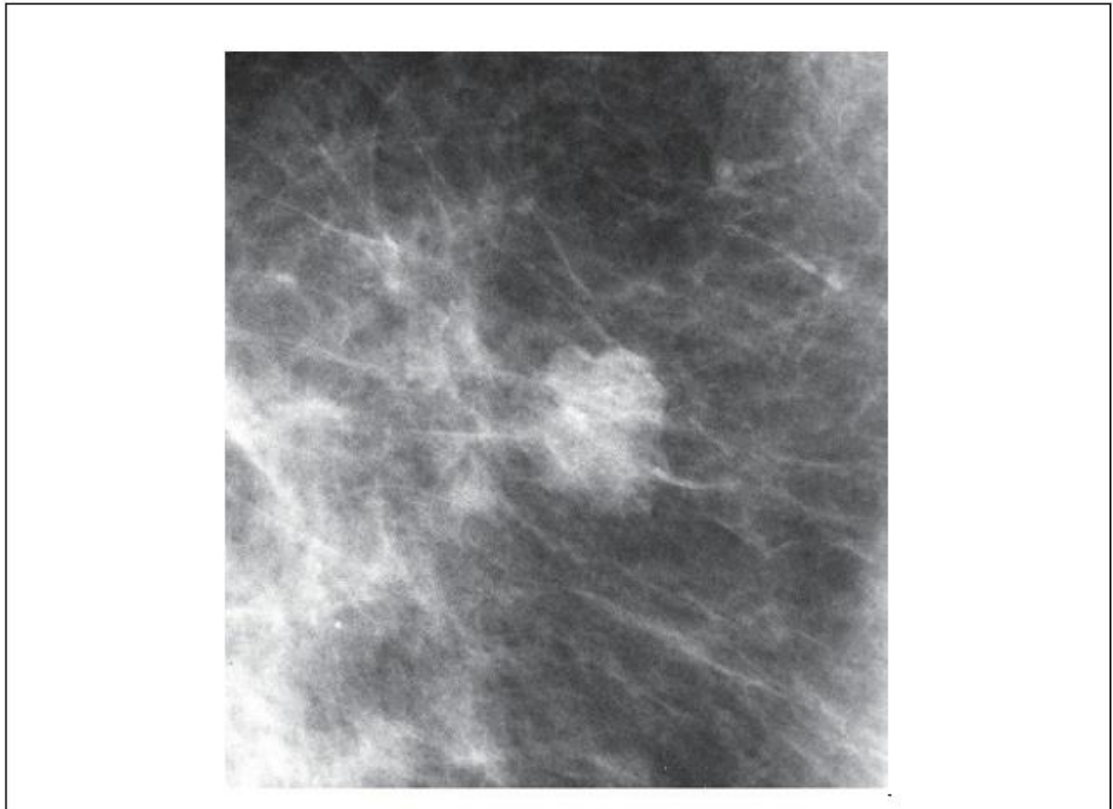


Figura 4. Masa indistinta.

Fuente: ACR BI-RADS Mammography [7].

Espiculada: El margen se caracteriza por líneas que irradian desde la masa. Este descriptor generalmente sugiere un hallazgo sospechoso.



Figura 5. Masa espiculada.

Fuente: ACR BI-RADS Mammography [7].

Densidad

Este descriptor se utiliza para definir la atenuación de rayos X de la masa en comparación con la atenuación esperada de un volumen igual de tejido mamario fibroglandular normal. La mayoría de los cánceres de mama aparecen como masas con densidad igual o superior a la de un volumen igual de tejido fibroglandular normal. La probabilidad de malignidad de una masa de alta densidad (70%) es significativamente mayor que la de masas de densidad igual o baja (22%). Es raro, aunque no imposible, que el cáncer de mama sea de menor densidad. Sin embargo, la densidad mamaria es una evaluación subjetiva y menos confiable en comparación con otras características mamográficas de las masas, como la forma y el margen. Los cánceres de mama nunca son radiolúcidos (contenedores de grasa), aunque pueden atrapar grasa [7].

Alta densidad: La atenuación de los rayos X de la masa es mayor que la

atenuación esperada de un volumen de igual tejido mamario fibroglandular.

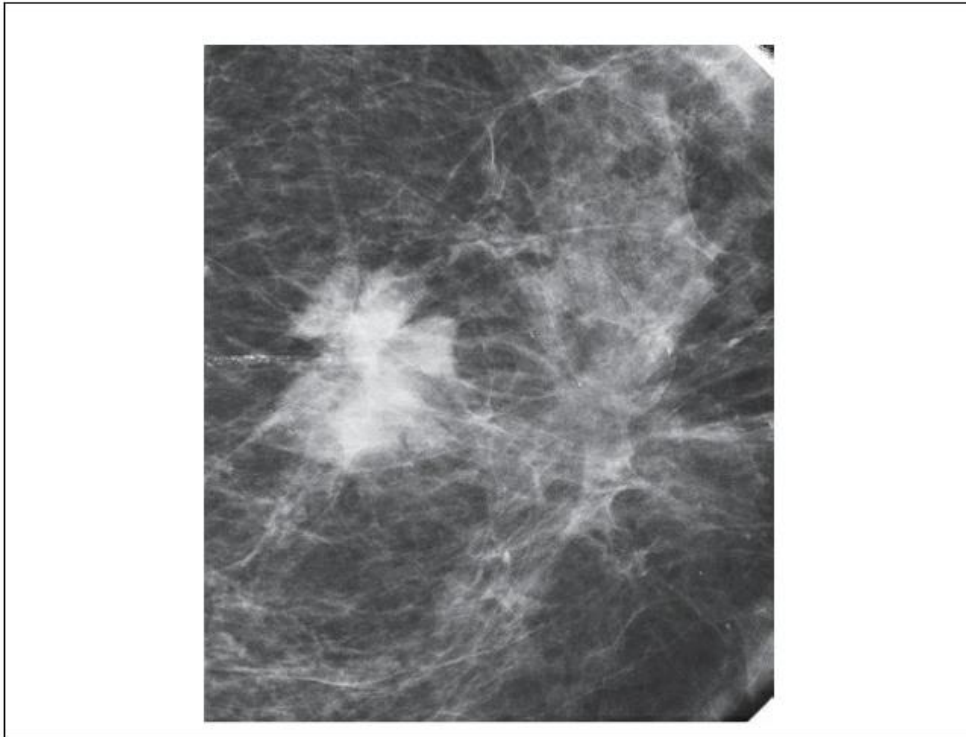


Figura 6. Masa de alta densidad.

Fuente: ACR BI-RADS Mammography [7].

Igual densidad: La atenuación de los rayos X de la masa es la misma que la atenuación esperada de un volumen igual de tejido mamario fibroglandular.

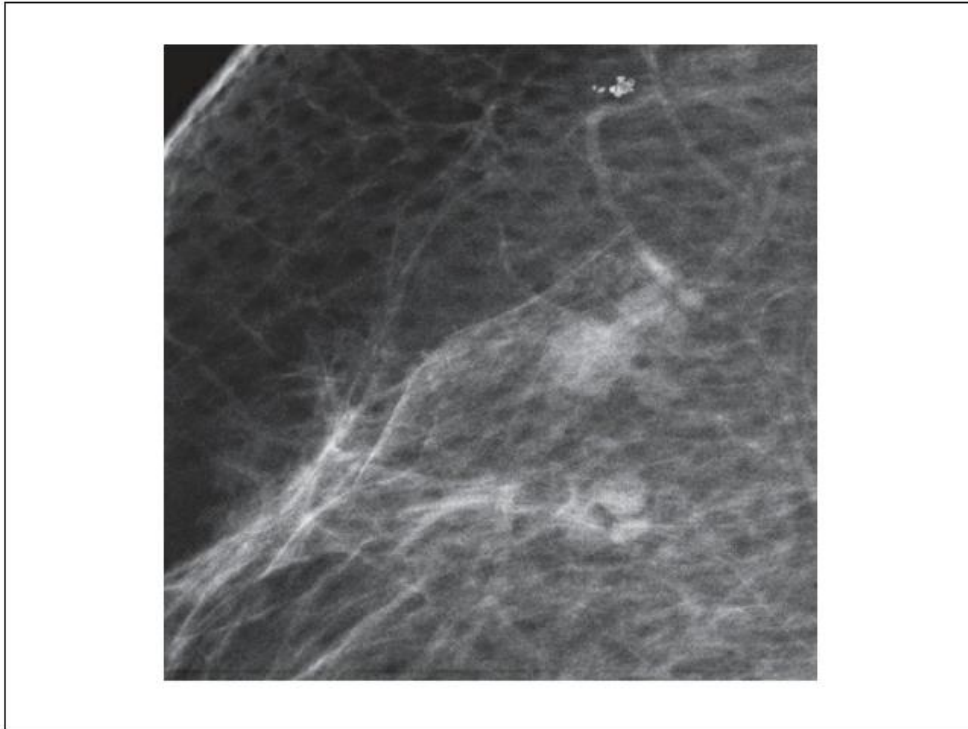


Figura 7. Masa de igual densidad.

Fuente: ACR BI-RADS Mammography [7].

Baja densidad: La atenuación de los rayos X de la masa es menor que la atenuación que se espera de un volumen igual de tejido mamario fibroglandular. Una masa de baja densidad puede ser un grupo de microquistes.

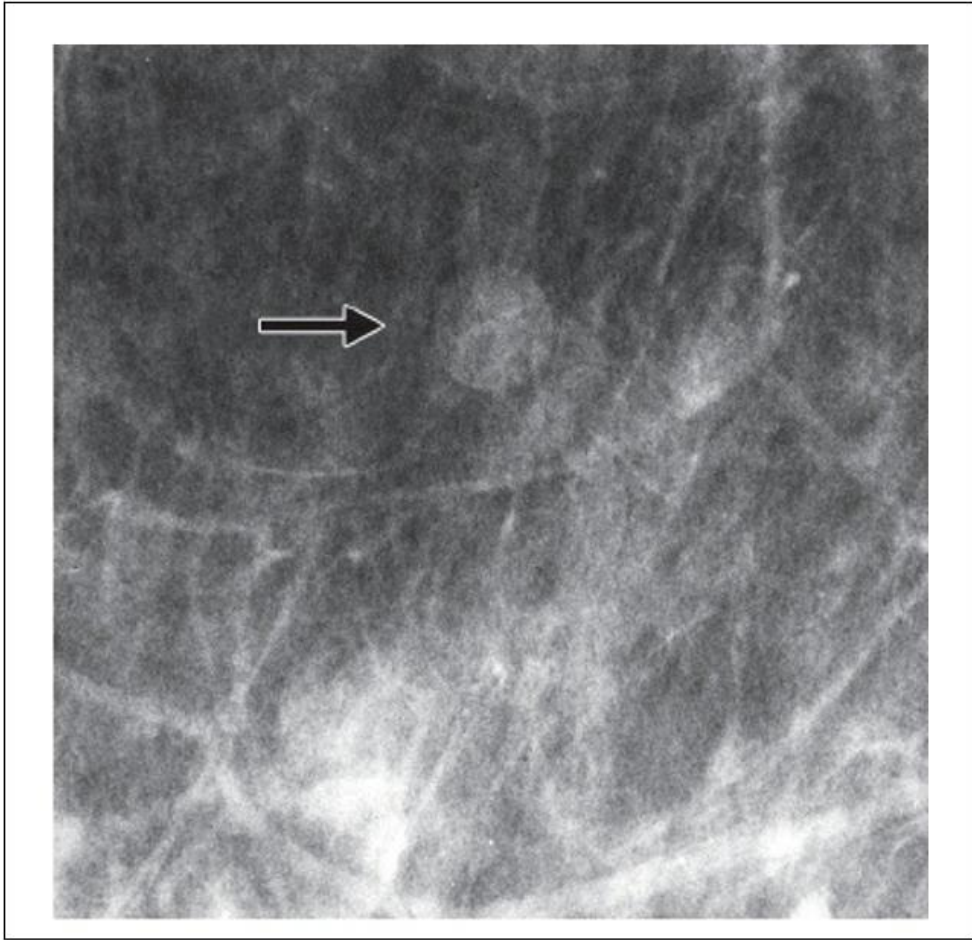


Figura 8. Masa de baja densidad.

Fuente: ACR BI-RADS Mammography [7].

Contenedor de grasa (fat-containing): Este término abarca todas las masas que contienen grasa, como los quistes oleosos, lipomas o galactoceles, así como las masas de densidad mixta, como los hamartomas. Una masa que contiene grasa casi siempre se considera una masa benigna.

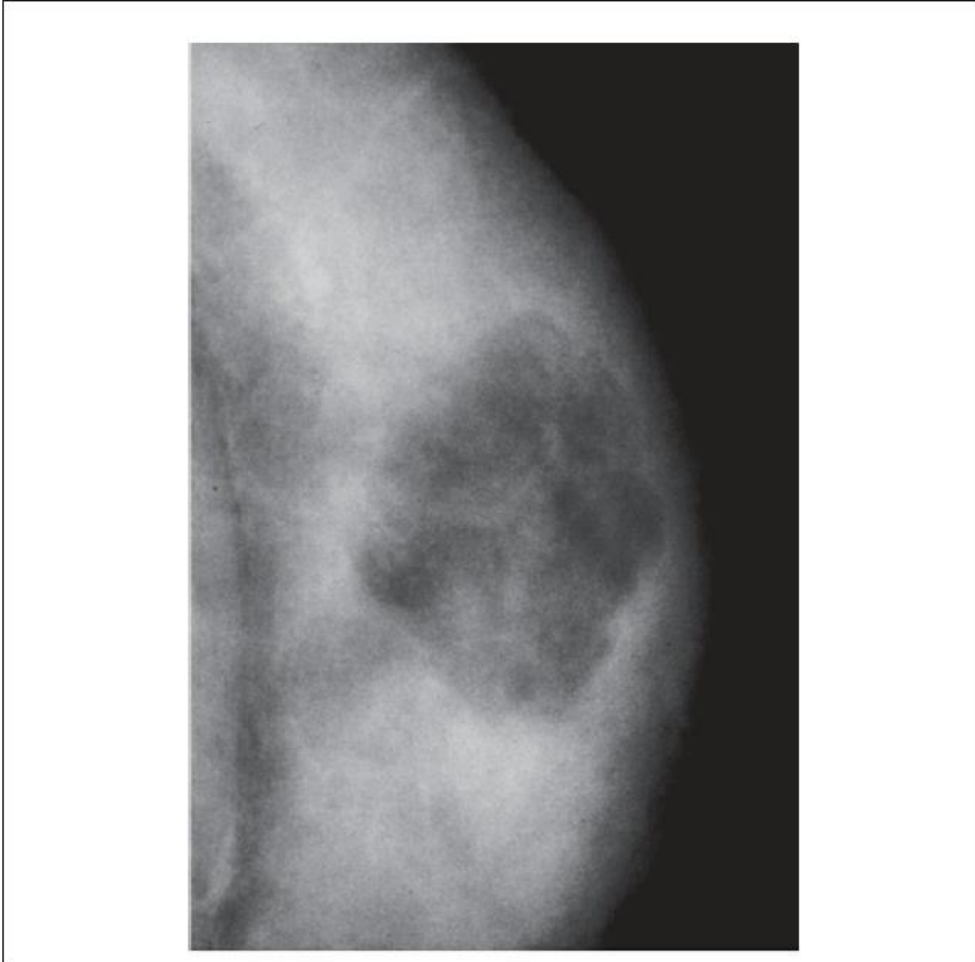


Figura 9. Masa contenedora de grasa.

Fuente: ACR BI-RADS Mammography [7].

b. Microcalcificaciones

Las microcalcificaciones evaluadas como benignas en la mamografía son típicamente más grandes, gruesas, redondeadas, con bordes suaves y más fáciles de visualizar que las calcificaciones malignas. Las microcalcificaciones asociadas con malignidad (y muchas calcificaciones benignas también) son generalmente muy pequeñas y, a menudo, requieren el uso de magnificación para ser observadas con claridad. Cuando no se puede determinar una etiología típicamente benigna, una descripción de las calcificaciones debe incluir su morfología y distribución [7].

Típicamente benigno

Piel: Estas microcalcificaciones suelen ser centradas y patognomónicas en su apariencia. Las microcalcificaciones en la piel son más comunes a lo largo del pliegue inframamario, paraesternal, que cubre la axila y alrededor de la areola. Las partículas individuales suelen estar agrupadas estrechamente, con grupos de menos de 5 mm en su mayor dimensión. Las formas atípicas pueden confirmarse como depósitos en la piel mediante proyecciones mamográficas tangenciales adicionales. Es importante notar que, si las microcalcificaciones de apariencia sospechosa están adyacentes a una superficie de la piel en una vista mamográfica, podrían ser dérmicas (y por lo tanto benignas). Por ello, se debe realizar una vista tangencial mamográfica con o sin magnificación antes de cualquier intervención [7].

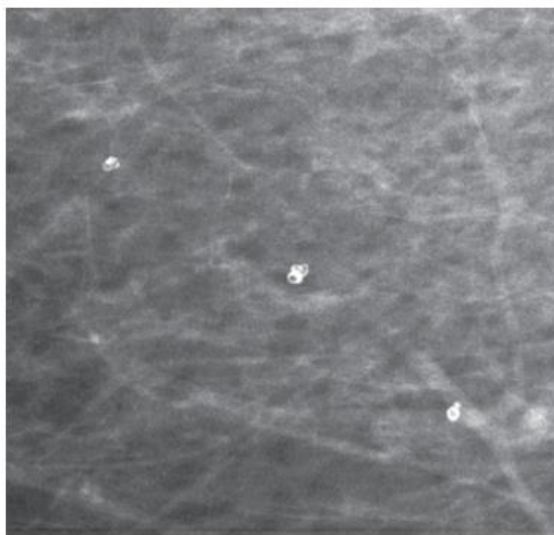


Figura 10. Microcalcificación típicamente benigna de piel (a).

Fuente: ACR BI-RADS Mammography [7].

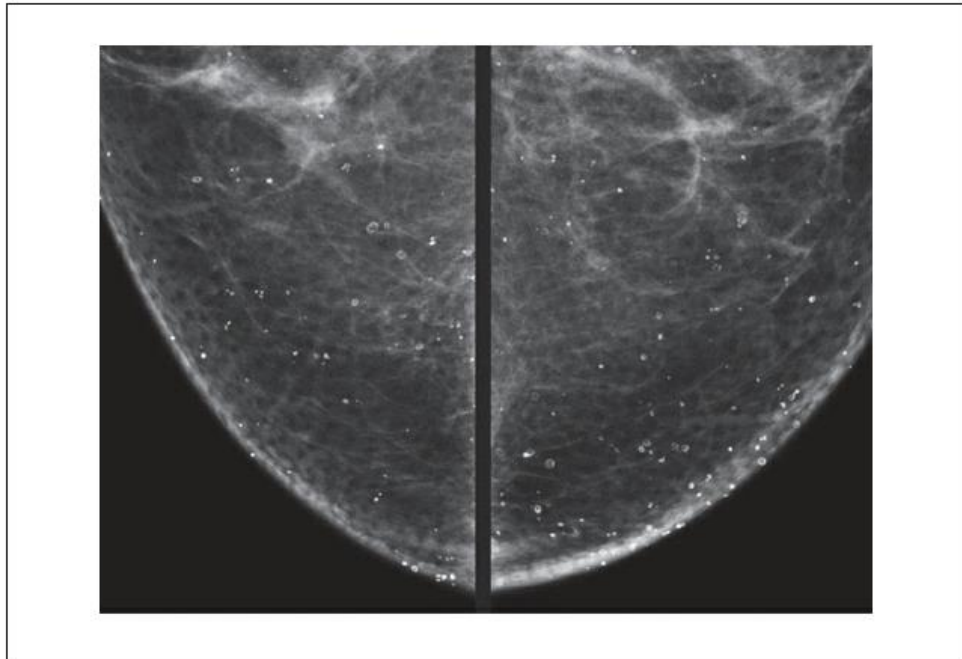


Figura 11. Microcalcificación típicamente benigna de piel (b).

Fuente: ACR BI-RADS Mammography [7].

Vascular: Son calcificaciones lineales tubulares o paralelas claramente asociadas con los vasos sanguíneos. Si bien las calcificaciones vasculares mayores son fáciles de identificar, unas pocas partículas calcificadas discontinuas pueden requerir vistas adicionales de ampliación y compresión para caracterizar mejor su naturaleza.

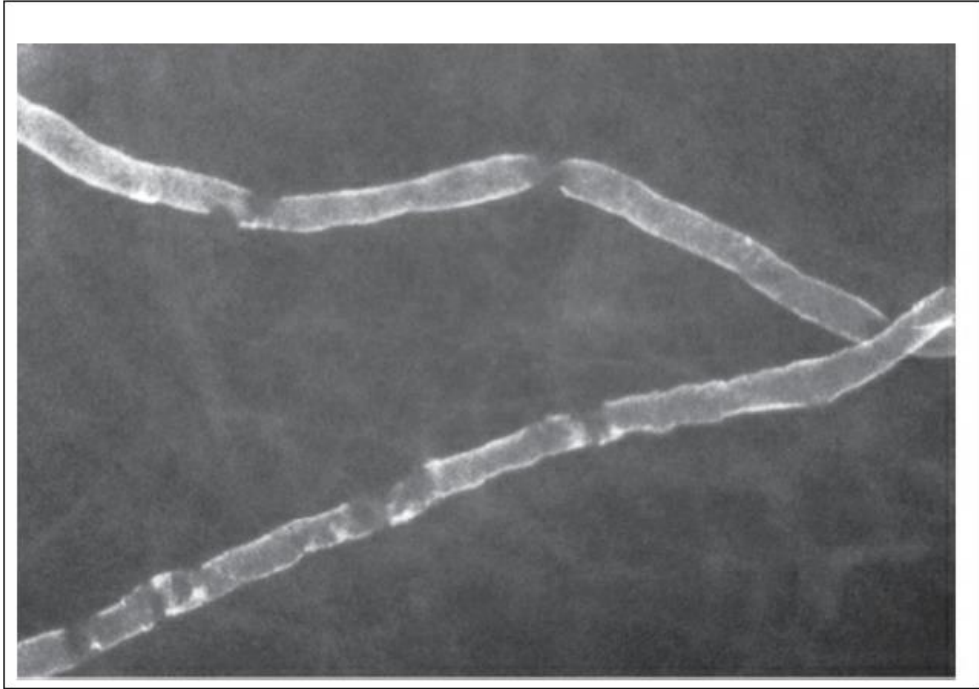


Figura 12. Microcalcificación típicamente benigna vascular (a).

Fuente: ACR BI-RADS Mammography [7].



Figura 13. Microcalcificación típicamente benigna vascular (b).

Fuente: ACR BI-RADS Mammography [7].

Grueso o “como palomita de maíz” (popcorn-like): Estas calcificaciones son clásicamente grandes (mayores de 2-3 mm de diámetro) y son producidas por un fibroadenoma involutivo.

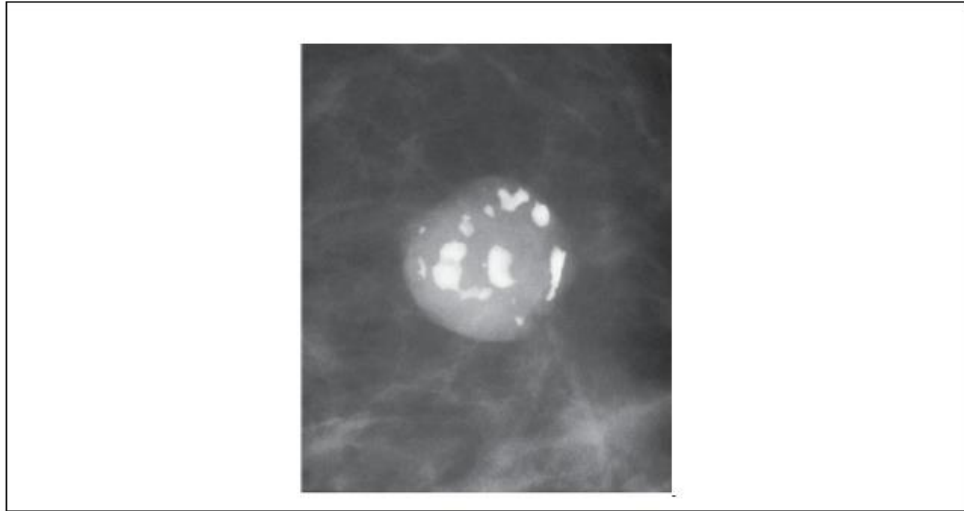


Figura 14. Microcalcificación típicamente Grueso o “como palomita de maíz” (popcorn-like) (a).

Fuente: ACR BI-RADS Mammography [7].

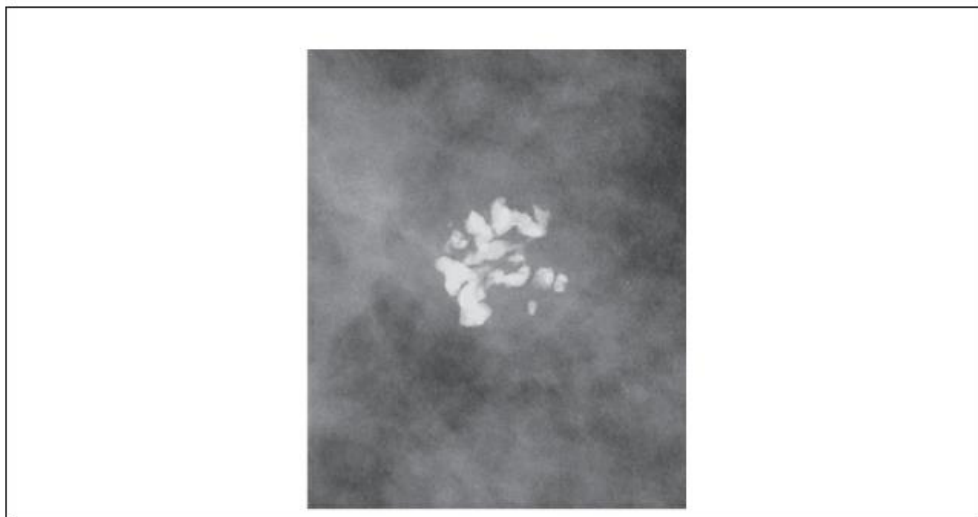


Figura 15. Microcalcificación típicamente Grueso o “como palomita de maíz” (popcorn-like) (b).

Fuente: ACR BI-RADS Mammography [7].

Largo en forma de varilla: Estas microcalcificaciones benignas, asociadas con

ectasia ductal, pueden formar varillas lineales lisas, sólidas o discontinuas, generalmente de 0.5 mm o más de diámetro. Un pequeño porcentaje de estas microcalcificaciones pueden tener un centro de calcio en la pared del conducto (peri ductal), pero la mayoría son intraductales, formándose dentro del lumen del conducto. Todas las calcificaciones similares a barras siguen una distribución ductal, que irradian hacia el pezón y ocasionalmente se ramifican. Generalmente son bilaterales, aunque pueden observarse solo en un seno, especialmente cuando hay pocas partículas visibles. Estas calcificaciones son comunes en mujeres mayores de 60 años.

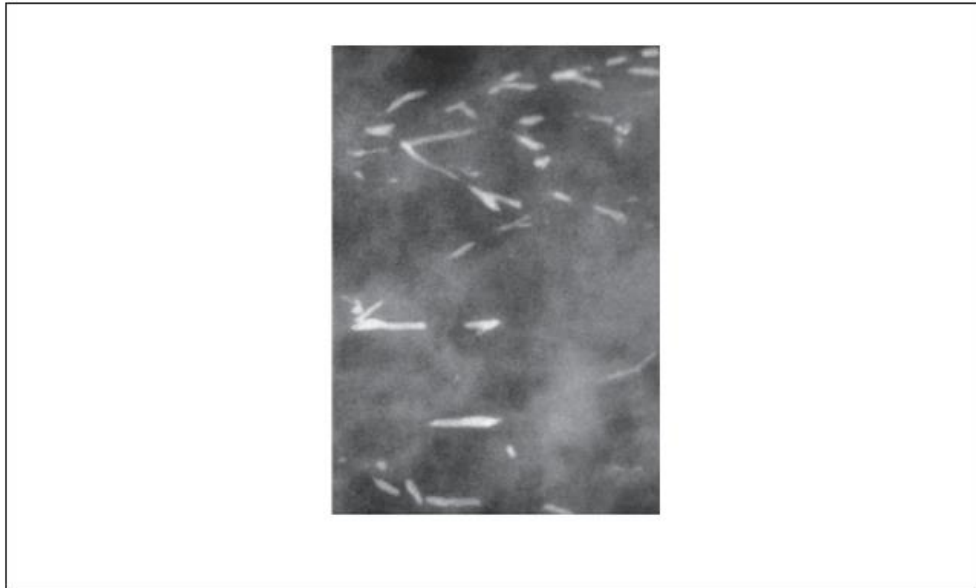


Figura 16. Microcalcificación típicamente benigna larga en forma de varilla (a).

Fuente: ACR BI-RADS Mammography [7].

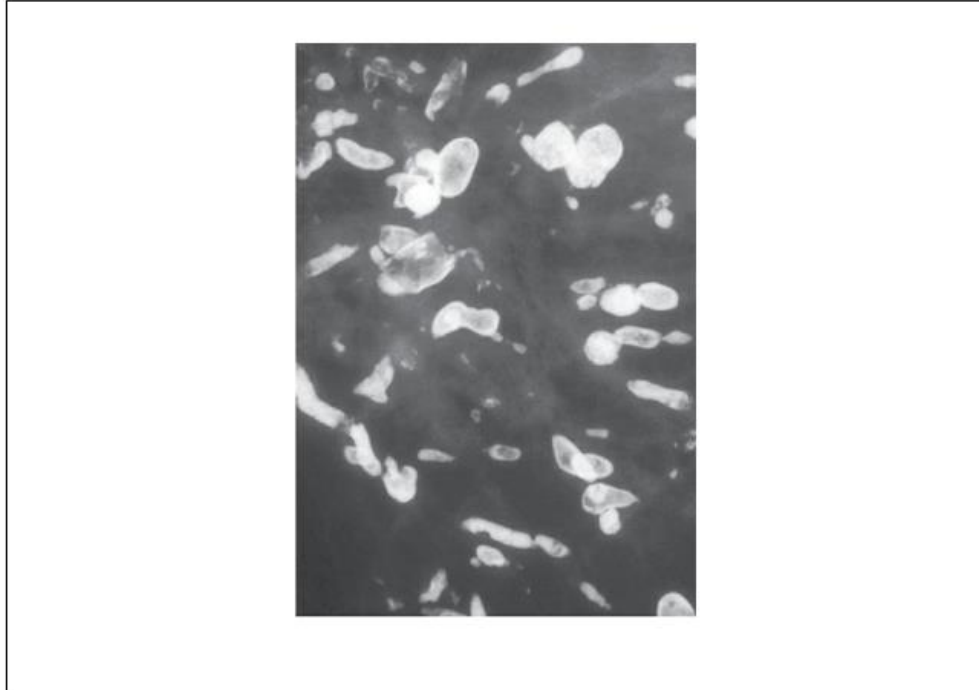


Figura 17. Microcalcificación típicamente benigna larga en forma de varilla (b).

Fuente: ACR BI-RADS Mammography [7].

Redondo: Estas microcalcificaciones pueden variar en tamaño y opacidad. Se consideran benignas cuando son difusas y pequeñas (menos de 1 mm) y se forman frecuentemente en los acinos de los lóbulos. Si son menores de 0.5 mm, se utiliza el término "punteada". Un grupo aislado de calcificaciones punteadas puede justificar una evaluación como probablemente benignas y seguimiento mamográfico, o una biopsia guiada por imagen si el grupo es nuevo, está aumentando, es lineal o segmentario en distribución, o está adyacente a un cáncer conocido.



Figura 18. Microcalcificación típicamente benigna redondas.

Fuente: ACR BI-RADS Mammography [7].

Bordeado: Estas son microcalcificaciones benignas que aparecen como depósitos de calcio en el borde de una esfera. Los depósitos son generalmente menores de 1 mm de grosor. Estas microcalcificaciones no agrupadas pueden oscilar entre menos de 1 mm hasta más de un centímetro. Las microcalcificaciones son redondas u ovaladas, con superficies lisas y centros lucentes. La necrosis grasa y las calcificaciones en las paredes de los quistes son las formas más comunes de calcificaciones en llanta, aunque pueden observarse formas más extensas (y ocasionalmente más gruesas) en quistes oleosos o simples.

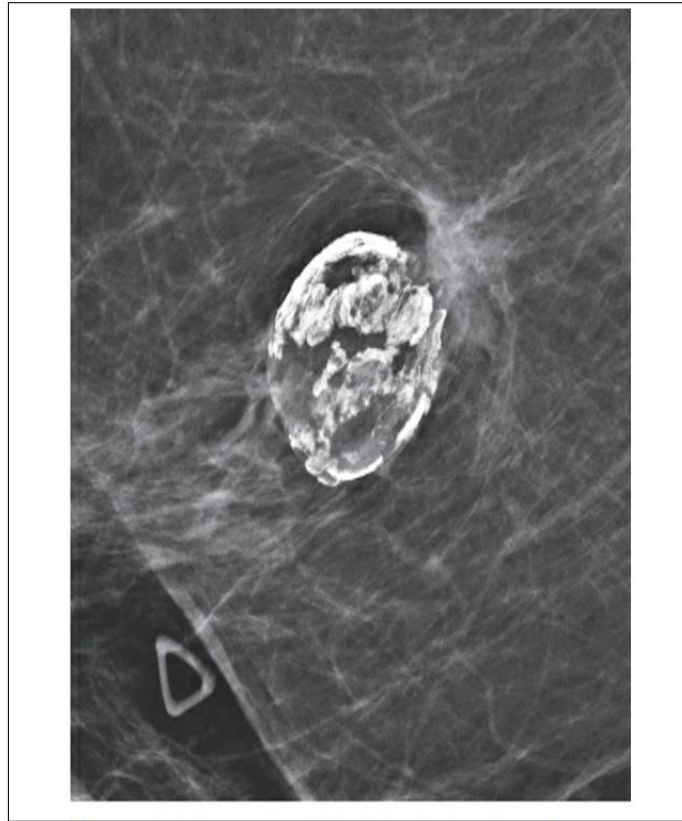


Figura 19. Microcalcificación típicamente benigna bordeada.

Fuente: ACR BI-RADS Mammography [7].

Distrófico: Estas calcificaciones suelen formarse en el seno irradiado o después de un trauma o cirugía. Son de forma irregular y generalmente mayores de 1 mm.

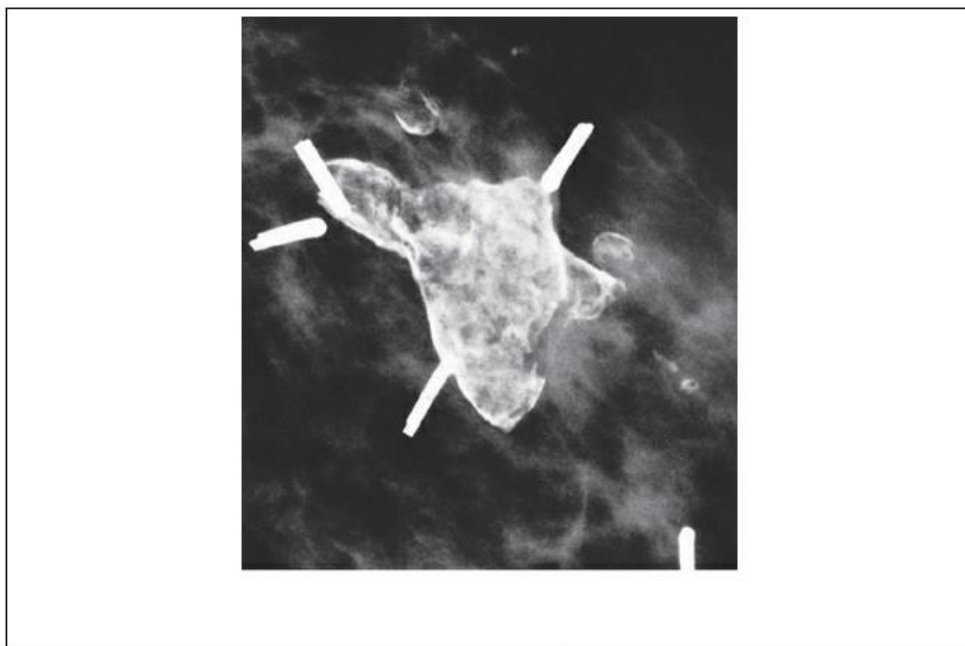


Figura 20. Microcalcificación típicamente benigna distrófico.

Fuente: ACR BI-RADS Mammography [7].

Leche de calcio: Estas microcalcificaciones se presentan como depósitos sedimentados en macro o microquistes. En la imagen craneocaudal suelen ser menos evidentes y aparecen como depósitos redondos y difusos, mientras que en vistas laterales a 90° (ML/LM) se definen más claramente y a menudo son semilunares, curvilíneos o lineales, delineando la parte dependiente del quiste. La característica más importante es el cambio aparente en la forma de las partículas calcificadas en diferentes proyecciones mamográficas (CC vs. MLO y especialmente LM/ML). A veces, las microcalcificaciones de leche de calcio se observan adyacentes a otros tipos de calcificaciones que pueden estar asociadas con malignidad, por lo que es crucial buscar formas más sospechosas, especialmente aquellas que no cambian de forma en la proyección lateral a 90° comparada con la proyección CC.

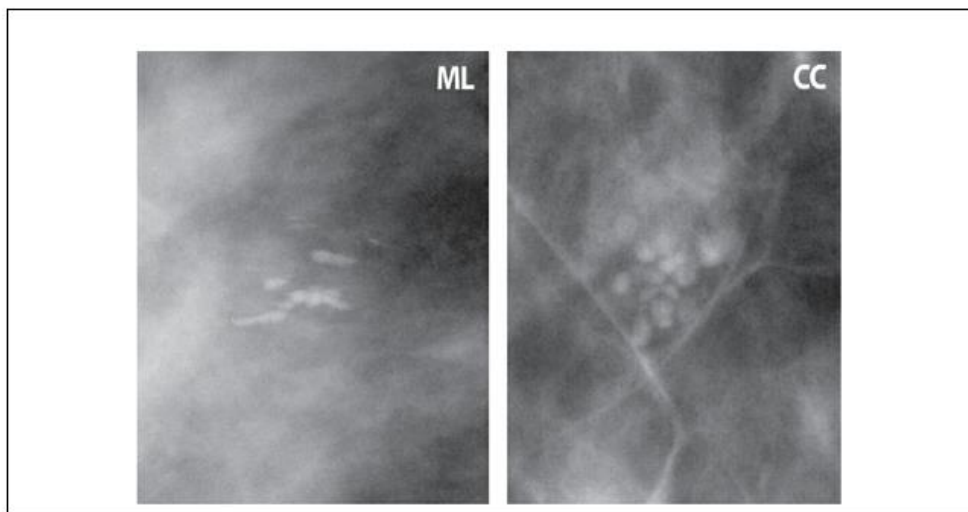


Figura 21. Microcalcificación típicamente benigna leche de calcio.

Fuente: ACR BI-RADS Mammography [7].

Suturada: Estas calcificaciones representan depósitos de calcio sobre material de sutura. Típicamente son lineales o tubulares en apariencia y, cuando están presentes, los nudos de sutura suelen ser visibles.

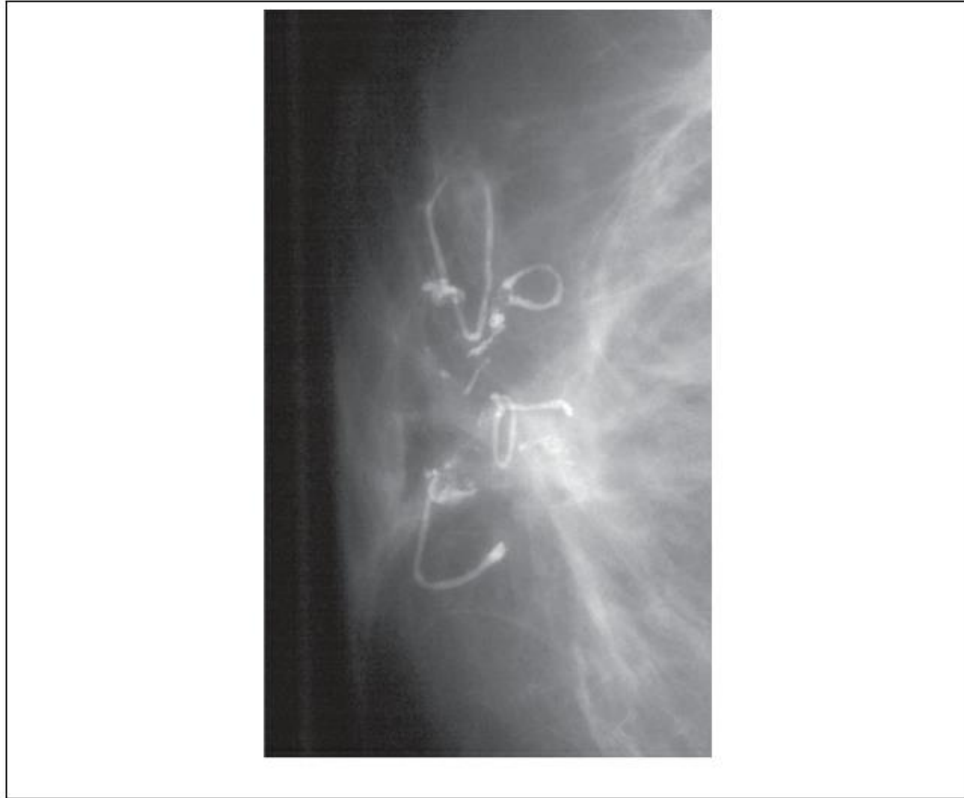


Figura 22. Microcalcificación típicamente benigna saturada.

Fuente: ACR BI-RADS Mammography [7].

Típicamente maligna

Amorfo: Estas microcalcificaciones son tan pequeñas y difusas que no se puede determinar una forma específica. Cuando las microcalcificaciones amorfas están agrupadas, en una disposición lineal o segmentaria, son sospechosas y generalmente justifican una biopsia. Si son bilaterales y difusas, suelen considerarse benignas, aunque las vistas aumentadas pueden ser útiles para una evaluación más precisa. El valor predictivo positivo (VPP) de las microcalcificaciones amorfas es aproximadamente del 20%. Por lo tanto, estas microcalcificaciones deben clasificarse adecuadamente en la categoría de evaluación BI-RADS 4B (VPP entre $>10\%$ y $\leq 50\%$).



Figura 23. Microcalcificación típicamente maligna amorfo.

Fuente: ACR BI-RADS Mammography [7].

Grueso heterogéneo: Estas microcalcificaciones son irregulares y conspicuas, generalmente oscilan entre 0.5 mm y 1 mm, y tienden a fusionarse. Son más pequeñas que las microcalcificaciones distróficas y pueden estar asociadas con neoplasias malignas, aunque con mayor frecuencia se encuentran en fibroadenomas, áreas de fibrosis o trauma con evolución distrófica. Numerosos grupos bilaterales de microcalcificaciones heterogéneas gruesas suelen considerarse benignos, aunque las vistas aumentadas pueden ser útiles. Un solo grupo de microcalcificaciones heterogéneas gruesas tiene un VPP de poco menos del 15%, por lo que este hallazgo debe clasificarse en la categoría de evaluación BI-RADS 4B (VPP entre $>10\%$ y $\leq 50\%$).

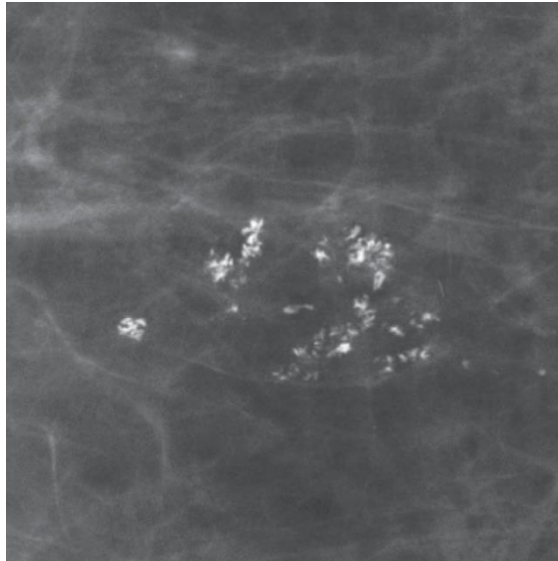


Figura 24. Microcalcificación típicamente maligna grueso heterogéneo.

Fuente: ACR BI-RADS Mammography [7].

Pleomorfo fino: Estas microcalcificaciones son más visibles que las amorfas y tienen formas discretas e irregulares. Se distinguen de las formas lineales finas y ramificadas por la ausencia de partículas lineales finas. Las microcalcificaciones pleomórficas varían en tamaño y forma, generalmente miden menos de 0.5 mm de diámetro y tienen un VPP para malignidad del 29%. Deben clasificarse en la categoría de evaluación BI-RADS 4B (VPP entre >10% y ≤50%).

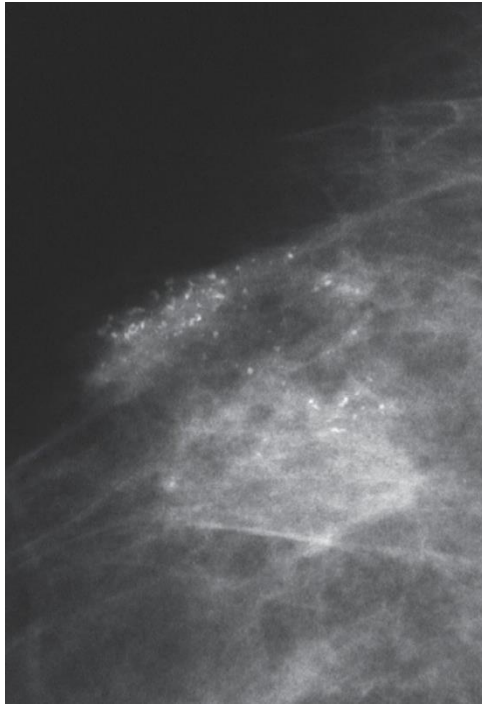


Figura 25. Microcalcificación típicamente maligna pleomorfo fino.

Fuente: ACR BI-RADS Mammography [7].

Bifurcación lineal fina: Estas microcalcificaciones son delgadas, lineales, irregulares y a veces discontinuas, con un calibre inferior a 0.5 mm. Ocasionalmente, pueden observarse formas ramificadas. Su apariencia sugiere el llenado de la luz de un conducto afectado por cáncer de mama. Entre las microcalcificaciones sospechosas, las de ramificación lineal fina tienen el VPP más alto (70%), por lo que deben clasificarse en la categoría de evaluación BI-RADS 4C (VPP entre >50% y <95%) independientemente de su distribución.

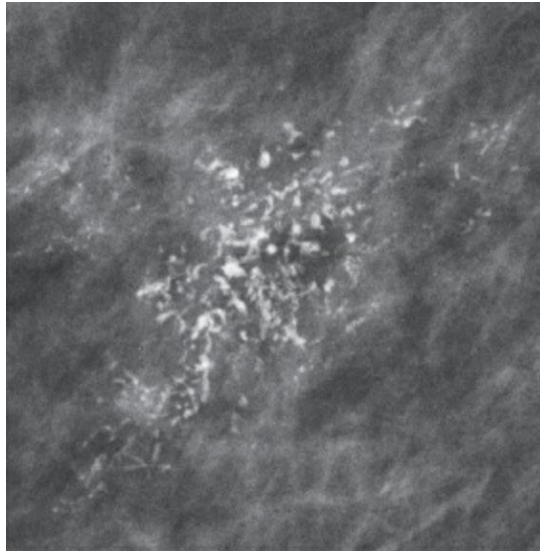
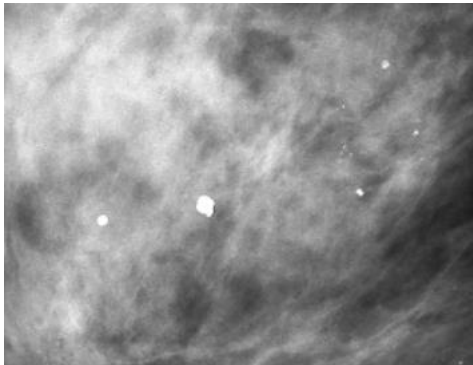


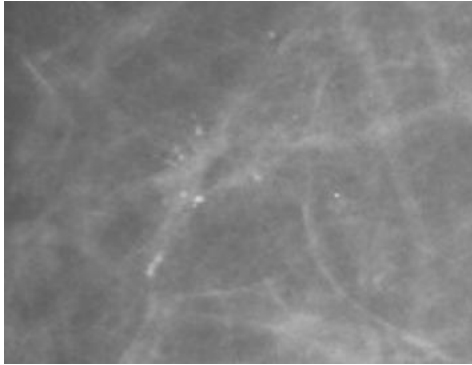
Figura 26. Microcalcificación típicamente maligna bifurcación lateral fina.

Fuente: ACR BI-RADS Mammography [7].

c. Principales Diferencias - Microcalcificaciones

Tabla 2. Diferencia entre microcalcificaciones.

 <p>Microcalcificación Benigna</p>	<ul style="list-style-type: none"> - Las microcalcificaciones benignas se caracterizan por ser más grandes, más gruesos, redondos con bordes suaves y más fáciles de ver. - No requiere magnificación - Están conformadas principalmente por oxalato de calcio.
--	--

 <p>Microcalcificación Maligna</p>	<ul style="list-style-type: none"> - Las microcalcificaciones malignas suelen ser más pequeñas. - Deben ser estudiadas con placas magnificadas para su caracterización. - Están conformadas predominantemente por fosfato de calcio.
--	---

1.5.3. Rayos x

Esta categoría de radiación electromagnética es similar a la luz visible, pero con una longitud de onda considerablemente menor. Sus propiedades físicas más destacadas incluyen su capacidad para atravesar el cuerpo humano, producir fluorescencia e impresionar emulsiones fotográficas. Además, esta radiación puede inducir diversos cambios en los tejidos biológicos y provocar la ionización de materiales gaseosos. Se utiliza ampliamente en la producción de imágenes médicas, desempeñando un papel crucial en la creación de radiografías convencionales, vídeo-angiografías o fluoroscopias, y tomografías [19].

Mamografía Digital

La mamografía digital se basa en imágenes de rayos X y utiliza el estándar DICOM (Digital Imaging and Communication in Medicine) para el cribado del cáncer de

mama. Este proceso implica la toma de cuatro radiografías: dos vistas cefalocaudales (CC) (desde arriba hacia abajo) y dos vistas medio laterales oblicuas (MLO) (desde el lateral derecho o izquierdo con una inclinación de 45 grados).

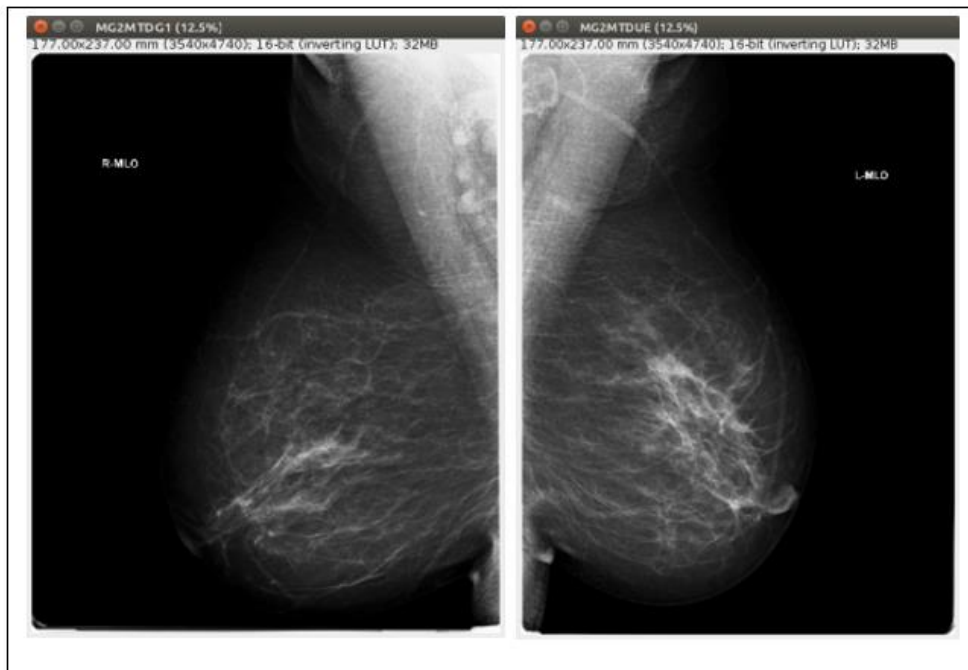


Figura 27. Mamografía digital.

Fuente: Oncosalud - Perú

1.5.4. Procesamiento de imágenes médicas

Estas técnicas operan sobre la representación de una imagen con el objetivo de facilitar su análisis por parte de un especialista médico o un sistema de visión artificial.

Adquisición y almacenamiento de imágenes digitales

Este proceso implica el uso de un sensor adecuado para detectar la información visual o alguna emisión, convirtiéndola en una señal eléctrica. Esta señal se transforma en un arreglo de cantidades binarias, que se almacenan o procesan utilizando una computadora. La imagen digital se representa como un arreglo 2D,

denotado por $f(x, y)$, donde cada punto, llamado píxel, tiene coordenadas espaciales definidas por x e y . El tamaño de la imagen es $N \times M$ píxeles, donde N es el ancho y M es el largo de la imagen. Cada píxel representa un valor de intensidad de la información visual o de la emisión obtenida. Este valor binario necesita un determinado número de bits para representar la información, comúnmente 8 bits (1 byte), 16 bits (2 bytes) o 32 bits (4 bytes) [19].

Histograma de una imagen

El histograma es una herramienta útil tanto cualitativa como cuantitativamente. Representa gráficamente la distribución de las intensidades de los píxeles en una imagen en niveles de gris. Denotamos $h(i)$ como el número de píxeles en la región de interés con un valor de intensidad i , donde $i = 0, 1, 2, \dots, L-1$ es el número posible de niveles de gris para la imagen. Los valores $h(i)$ corresponden a los valores del histograma. Este gráfico es bidimensional y representa $h(i)$ en función de i . Así, el histograma proporciona información crucial sobre el brillo, el contraste y el rango dinámico de una imagen [19].

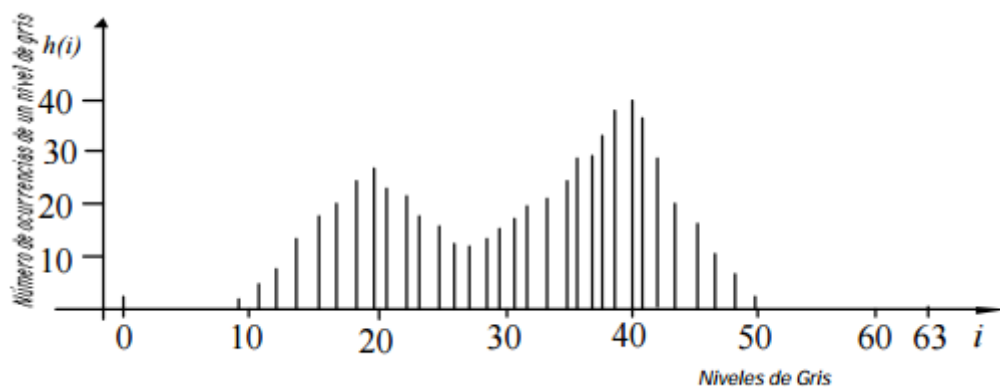


Figura 28. Histograma para los niveles de intensidad de una imagen, con $L=64$.

Fuente: [19].

Realce de imágenes por modificación del contraste

Una de las imperfecciones más frecuentes en las imágenes digitales es el bajo contraste, que ocurre cuando el rango de intensidades es limitado en comparación con el rango completo de niveles de gris disponibles (por ejemplo, de 0 a 255 niveles). El contraste de una imagen puede mejorarse modificando la intensidad de cada píxel en la imagen de entrada, denotada por i , mediante una transformación específica. Esta transformación, $g=T(i)$, relaciona la intensidad de entrada i con la intensidad de salida g y se representa habitualmente mediante un gráfico o una tabla [19].

1.5.5. Algoritmos de procesamiento de imágenes

Region Growing

Este algoritmo se utiliza para agrupar píxeles en regiones más extensas. Comienza con un conjunto de puntos semilla y expande las regiones a partir de estos puntos al incluir píxeles vecinos que comparten atributos similares, como intensidad, textura, niveles de gris o color. El proceso es iterativo: cada píxel semilla se expande progresivamente hasta que todos los píxeles han sido procesados, formando así regiones distintas cuyos límites están definidos por polígonos cerrados [20].

Canny

El algoritmo Canny es un método avanzado para la detección de bordes, que se realiza en tres etapas principales: filtrado, toma de decisiones inicial y aplicación de histéresis [21]. En la primera fase, se aplica un filtrado mediante convolución utilizando la primera derivada de una función gaussiana

normalizada discreta en dos direcciones: horizontal y vertical. La función gaussiana se caracteriza por dos parámetros fundamentales: el valor medio m y la desviación estándar σ . En este contexto, el valor medio es cero, y la ecuación del filtro gaussiano es:

$$g(x) = k \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (1)$$

El valor de K se elige para que el máximo de $g(x)$ sea uno en su versión discreta. La fase de filtrado utiliza esta derivada para resaltar los bordes en la imagen.

$$g'(x) = \frac{d}{dx} g(x) = -\frac{k}{\sigma^2} x \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (2)$$

K-Means

El algoritmo de agrupamiento K-means es un método sencillo para la segmentación de datos, que implica determinar el número de grupos K y los centroides de dichos grupos [22]. Hay dos enfoques para seleccionar los centroides iniciales: el primero consiste en elegir aleatoriamente K objetos como centroides iniciales, y el segundo en utilizar los primeros K objetos en secuencia. Posteriormente, el proceso sigue estos pasos:

- a) Establecer los centroides iniciales según el número de clusters esperado.
- b) Calcular la distancia de cada objeto con respecto a los centroides.
- c) Agrupar los objetos en función de la distancia mínima a los centroides.

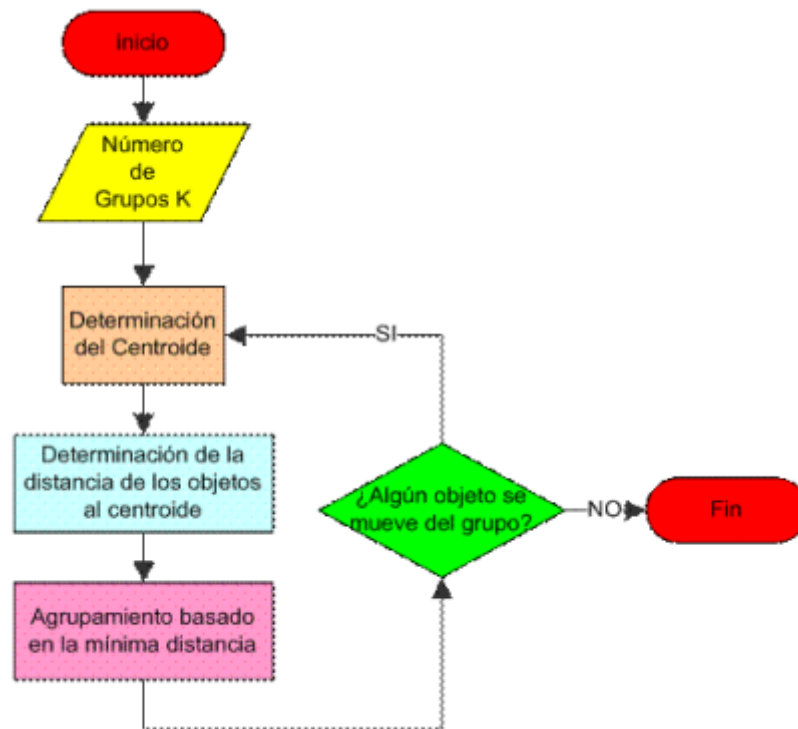


Figura 29. Diagrama de flujo k-means.

Fuente: [22].

1.5.6. Redes neuronales artificiales

a. Redes no lineales

Las redes neuronales lineales son conocidas por su carácter predecible y la aplicación de descripciones matemáticas simples, lo cual facilita la obtención de soluciones. Sin embargo, las redes neuronales no lineales ofrecen una mayor capacidad de resolución de una amplia gama de problemas, ya que pueden manejar casos en los que la relación entre las señales de entrada y salida no es lineal. En contraste, las redes neuronales lineales sólo pueden abordar problemas con correlaciones lineales entre las señales de entrada y salida.

Backpropagation

El algoritmo de backpropagation (propagación hacia atrás) es un método comúnmente empleado para estimar los errores en las neuronas de las capas ocultas dentro de una red neuronal. Este enfoque es tan prevalente que la mayoría de los programas para el entrenamiento de redes neuronales lo utilizan como método estándar [23]. Existen otros métodos de entrenamiento, como el algoritmo de propagación rápida, así como sistemas basados en métodos matemáticos avanzados, como el gradiente conjugado y los métodos de Levenberg-Marquardt.

Aunque backpropagation es reconocido por su eficiencia, esta ventaja es evidente principalmente cuando se resuelve un problema específico utilizando una red neuronal. Una característica notable de backpropagation es que opera sin depender de supuestos teóricos subyacentes, a diferencia de algunas técnicas complejas que pueden ser ocasionalmente ineficaces. Por tanto, backpropagation ofrece una solución fiable y efectiva para el entrenamiento de redes neuronales.

La figura a continuación ilustra el proceso de backpropagation en una red neuronal de tres capas con dos entradas y una salida.

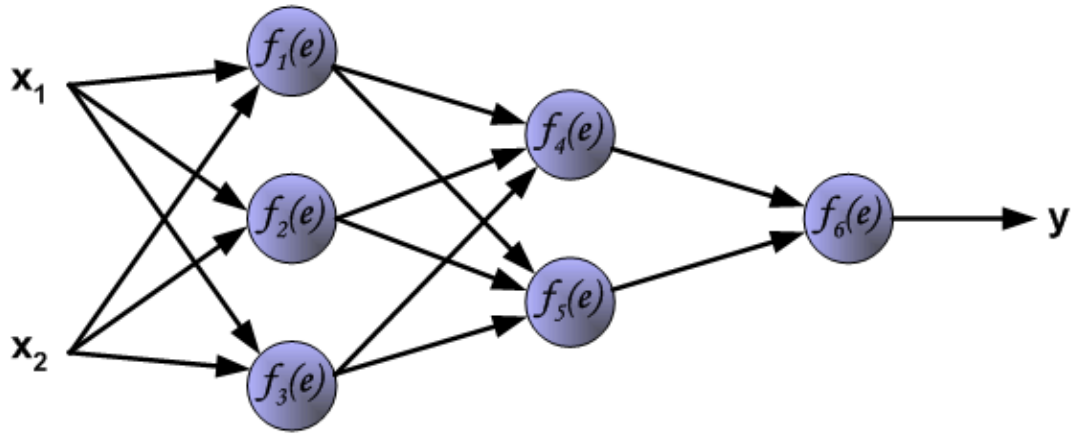


Figura 30. Proceso de backpropagation.

II. MATERIAL Y MÉTODO

2.1. Tipo y Diseño de Investigación:

Tipo de Investigación: La investigación sobre el reconocimiento de microcalcificaciones en mamografías utilizando técnicas de aprendizaje automático se clasifica principalmente como investigación aplicada. Este tipo de investigación busca desarrollar y aplicar tecnologías avanzadas para resolver problemas específicos del mundo real. En este caso, el objetivo es mejorar la detección y el diagnóstico de microcalcificaciones en imágenes mamográficas, un aspecto crucial en la detección temprana del cáncer de mama.

Diseño de investigación: La presente investigación se clasifica como cuasi-experimental debido a su enfoque en la evaluación de un modelo en un entorno no controlado y sin asignación aleatoria de tratamientos. La investigación se centra en el análisis y la comparación del rendimiento del modelo frente a datos reales y previamente etiquetados, en lugar de seguir un diseño experimental rígido con grupos de intervención y control aleatorios. Este enfoque permite examinar la efectividad del modelo en un contexto práctico y realista, aunque no cumple con los criterios estrictos de un experimento controlado aleatorio.

2.2. Variables, Operacionalización

Variable Independiente:

Uso de una máquina de aprendizaje.

Variable Dependiente:

Reconocimiento de microcalcificaciones en mamografías.

Tabla 3. Operacionalización de variables.

Variable de estudio	Definición conceptual	Definición operacional	Dimensiones	Indicadores	Ítems	Instrumento	Valores finales	Tipo de variable	Escala de medición
Uso de una máquina de aprendizaje	Uso de máquinas de aprendizaje para entrenamiento de modelos para procesar y analizar mamografías.	Implementación específica de un algoritmo de aprendizaje automático.	Eficiencia	Cantidad de épocas de entrenamiento	$E = 1415 \text{ ep}$	Ficha de observación	0-2000	Variable independiente	Numérica
				Porcentaje de Error	$Pe \leq 1\%$	Ficha de observación	0-100.00		Porcentaje

Reconocimiento de microcalcificaciones en mamografías	proceso de identificar y clasificar microcalcificaciones presentes en imágenes de mamografías	Evalúa el desempeño del modelo de aprendizaje automático en la detección.	Desempeño	Sensibilidad	$P_s = \frac{M_{vp}}{M_{vp} + M_{fn}}$	Ficha Observación	0-100	Variable dependiente	Porcentaje
				Especificidad	$P_e = \frac{M_{vn}}{M_{vn} + M_{fp}}$	Ficha Observación	0-100		Porcentaje
				Valor Predictivo Positivo	$V_{pp} = \frac{M_{vp}}{M_{vp} + M_{fp}}$	Ficha Observación	0-100		Porcentaje
				Valor Predictivo Negativo	$V_{pn} = \frac{M_{vn}}{M_{fn} + M_{vn}}$	Ficha Observación	0-100		Porcentaje

Donde:

Sensibilidad (Ps): Se refiere a la capacidad de nuestro estimador para identificar correctamente como positivos los casos que realmente están enfermos; es la proporción de personas enfermas que han sido correctamente identificadas.

Especificidad (Pe): Es la capacidad de nuestro estimador para dar como casos negativos los casos realmente sanos; en una proporción de sanos correctamente identificados.

Valor Predictivo Positivo (Vpp): Es la probabilidad de tener la enfermedad si el resultado de la prueba diagnóstica es positivo.

Valor Predictivo Negativo (Vpn): Probabilidad de no presentar la enfermedad.

Muestra verdadero positivo (Mvp): Muestra con resultado correcto.

Muestra verdadero negativo (Mvn): Muestra con resultado correcto.

Muestra falso positivo (Mfp): Muestra con resultado incorrecto.

Muestra falso negativo (Mfn): Muestra con resultado incorrecto.

2.3. Población y muestra de estudio

Población:

En esta investigación se ha utilizado una base de datos compuesta por 322 imágenes de diagnóstico médico proporcionada por Databases you can use for benchmarking - E.E.U.U.

Muestra:

La muestra para la investigación es probabilística. Se utiliza la fórmula 3 para calcular el tamaño de muestra para una población finita.

$$n = \frac{Z^2 \times p \times q \times N}{E^2 \times (N - 1) + Z^2 \times p \times q} \quad (3)$$

Donde:

N = Población

Z = Nivel de confianza (1.025)

p = Proporción de la característica de interés (0.5)

q = Proporción que no tiene la característica de interés (0.5)

E = Error máximo (0.08)

Imágenes
$n = \frac{1.025^2 \times 0.5 \times 0.5 \times 322}{0.08^2 \times (322 - 1) + 1.025^2 \times 0.5 \times 0.5}$
$Total = 36.590 \cong 36$ <p>Imágenes</p>

Aplicando la fórmula 3 se obtiene como muestra a evaluar tan solo de 36 imágenes, dichas imágenes mamográficas servirán para poner a prueba el método desarrollado en esta investigación.

2.4. Técnicas e instrumentos de recolección de datos

Análisis Documental

La recolección de información se llevó a cabo mediante una revisión exhaustiva de libros, investigaciones previas, documentos y otros materiales relevantes. Este proceso implica obtener datos de diversas fuentes documentales, como libros, artículos académicos y otros informes. Con base en esta información, se determinará el modelo y el algoritmo a seguir.

Observación

Técnica de investigación que consiste en la observación y registro sistemático de comportamientos y eventos relevantes dentro de un contexto específico. El objetivo es obtener datos detallados y precisos sobre el comportamiento

humano en situaciones reales.

Procedimiento para la recolección de datos

Análisis Documental: El proceso de análisis documental se desarrolla de la siguiente manera:

Adquisición de imágenes mamográficas que incluyan microcalcificaciones, obtenidas del dataset disponible en este enlace <http://peipa.essex.ac.uk/pix/mias/>.

Investigación de las técnicas de reconocimiento de imágenes más utilizadas para la detección de microcalcificaciones, abarcando tanto métodos desarrollados como propuestas de investigaciones previas.

Exploración de métodos de visión artificial pertinentes para alcanzar los objetivos del proyecto.

2.5. Procedimiento de análisis de datos:

Para la evaluación del método, se calculará la siguiente fórmula estadística.

Razón de verosimilitudes positiva

$$Rvp = \frac{\text{Sensibilidad}}{1 - \text{Especificidad}} \quad (4)$$

Razón de verosimilitudes negativa

$$Rvp = \frac{1 - \text{Sensibilidad}}{\text{Especificidad}} \quad (5)$$

Ratio o razón de falsos descubrimientos

$$Rfd = \frac{Mfp}{Mfp + Mvp} \quad (6)$$

Ratio o razón de falsos positivos

$$Rfp = \frac{Mfp}{Mfp + Mvn} \quad (7)$$

Precisión

Es la relación de muestras precisas diagnosticadas de forma correcta del total de muestras evaluadas.

$$P = \frac{Mvp + M}{p + n}$$

(8)

2.6. Criterios éticos

Confidencialidad: Toda muestra obtenida y procesada se mantendrá en estricto secreto durante el proceso de investigación y desarrollo.

Derechos de Autor: Todos los materiales intelectuales utilizados en esta investigación están adecuadamente citados y referenciados.

Responsabilidad Social: Las muestras obtenidas y procesadas cumplen con los estándares internacionales sobre radiación, asegurando que ningún participante haya sufrido daños en su integridad física.

2.7. Criterios de Rigor Científico

Fiabilidad: Los resultados obtenidos reflejan datos verídicos y precisos.

Validez: La interpretación de los resultados será validada por expertos en la materia.

Consistencia: Esta investigación proporcionará material coherente y será certificado por la comunidad científica.

III. RESULTADOS Y DISCUSIÓN

3.1. Resultados

3.1.1. Pruebas realizadas

En este capítulo se presenta la cantidad de pruebas realizadas a los métodos propuestos, los cuales se basan en técnicas de Inteligencia Artificial para el reconocimiento de microcalcificaciones en mamografías. Se emplearon los lenguajes de programación C++ y Java, utilizando los entornos de desarrollo QT Creator y Netbeans, respectivamente, junto con las bibliotecas de visión artificial OpenCV y Neuroph para redes neuronales.

Se utilizaron un total de 286 imágenes con microcalcificaciones benignas y malignas para el entrenamiento, y 36 imágenes adicionales para las pruebas. El número de imágenes clasificadas fue:

Imágenes	Total
Microcalcificaciones benignas	176
Microcalcificaciones malignas	110
Total	286

El número de imágenes para pruebas fue:

Imágenes	Total
Microcalcificaciones benignas	24
Microcalcificaciones malignas	12
Total	36

a. Aplicación desarrollada

Se ha creado una aplicación en el lenguaje de programación C++ utilizando el IDE Qt Creator 3.2.1,

y en Java empleando el IDE Netbeans 8.0, junto con las bibliotecas OpenCV 2.4.9 y Neuroph.

b. Equipo para la puesta a prueba de los métodos

Las especificaciones principales de la computadora utilizada para las pruebas son las siguientes:

- a) Sistema Operativo: Windows 10 64 Bits.
- b) Memoria RAM: 8 GB.
- c) Microprocesador: Intel Core i7-4700MQ 2.40 GHz.
- d) GPU (Unidad de Procesamiento Gráfico): NVIDIA GeForce 840M.

En este capítulo se explicará el método propuesto para el reconocimiento de microcalcificaciones en mamografías, así como la interpretación de los resultados. Se utilizó una red neuronal perceptrón multicapa como clasificador. En total, se emplearon 36 imágenes mamográficas, incluyendo aquellas con microcalcificaciones benignas y malignas.

c. Indicadores

Los métodos propuestos fueron probados para poder medir y comparar los indicadores:

Porcentaje de sensibilidad

Porcentaje de especificidad

Valor predictivo positivo

Valor predictivo negativo

Razón de verosimilitudes positivas

Razón de verosimilitudes negativas

Precisión o exactitud - Accuracy

Ratio o razón de falsos descubrimientos

Ratio o razón de falsos positivos

d. Métodos propuestos

En la siguiente figura se muestran los métodos propuestos.

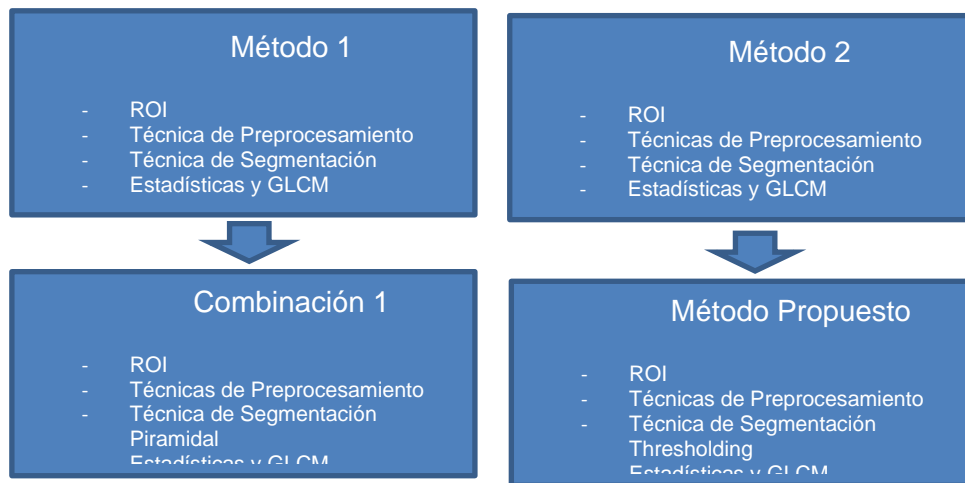


Figura 31. Métodos propuestos.

Primer Método

Combinación 1

Se han utilizado las siguientes técnicas:

Técnicas de Preprocesamiento: Dilatación y Escalado de Grises.

Técnica de Segmentación: Piramidal.

Extracción de Características: Para el reconocimiento de microcalcificaciones benignas o malignas se obtuvieron los siguientes parámetros: estadísticos (media, desviación estándar, asimetría, curtosis) y mediante GLCM (contraste, correlación, máxima probabilidad, energía, entropía, homogeneidad, sombreado de conglomerados, tendencia de conglomerados, información de correlación 1, información de correlación 2).

Resultados 1:

Para el reconocimiento de imágenes se empleó una Red Neuronal Perceptrón Multicapa con 14 neuronas en la capa de entrada, 10 neuronas en la capa oculta y una neurona en la capa de salida. Se realizaron 45000 iteraciones, con un error máximo de 0.01, una tasa de aprendizaje de 0.2 y un momentum de 0.7 durante el entrenamiento utilizando la función sigmoideal.

El gráfico del error total de la red presenta el siguiente aspecto:

Total Network Error Graph

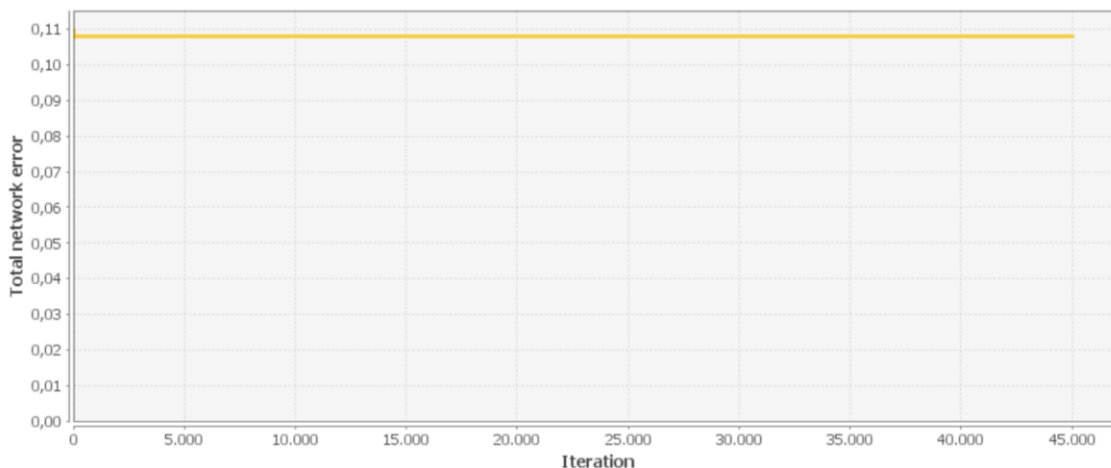


Figura 32. Gráfico Total de error de la red en Neuroph Studio.

El total de error de la red es superior al máximo error indicado (0.01) el total de error de la red nos da un resultado de 0.108038427327329.

Error cuadrático Medio (MSE): El error cuadrático medio es un criterio de evaluación más usado para problemas de regresión, su uso es mayormente para el aprendizaje automático supervisado:

$$error\ cuadrático = (real - estimado)^2 \quad (9)$$

Para que el error cuadrático medio sea perfecto tiene que ser 0 o aproximarse a ello. En la siguiente imagen podemos ver el resultado MSE de 0.23860983492404692.

```

Input: 208,7906; 32,6246; -0,4233; -1,0115; 0,1026; 511,9646; 0,7436; 0,0031; 5,8422; -116,548,344; 3,481,9828; 0,0101; -0,0131; 0,9981; Output: 0,543; Desired output: 1; Error: -0,
Input: 205,8487; 46,6406; -1,0254; -0,0574; 0,1068; 801,5143; 0,7934; 0,0034; 5,8112; -709,814,0078; 6,958,1101; 0,0143; -0,013; 0,998; Output: 0,543; Desired output: 1; Error: -0,4
Input: 193,66; 55,6753; -0,6811; -0,976; 0,0797; 806,2228; 0,8706; 0,0015; 6,7076; -1,011,570,6193; 11,652,0417; 0,0082; -0,0134; 0,9974; Output: 0,543; Desired output: 1; Error: -0
Input: 193,5767; 55,9383; -0,6865; -0,9728; 0,0735; 820,6087; 0,8691; 0,0014; 6,7435; -1,024,633,176; 11,713,5728; 0,0072; -0,0137; 0,9977; Output: 0,543; Desired output: 1; Error:
Input: 211,88; 44,1707; -1,4858; 2,289; 0,0769; 684,1086; 0,8068; 0,0021; 6,5575; -835,829,3906; 6,399,0748; 0,019; -0,0106; 0,9947; Output: 0,543; Desired output: 0; Error: 0,543;
Input: 230,5357; 16,5691; -0,6942; 0,5737; 0,0757; 275,7702; 0,4587; 0,0011; 6,9788; -13,584,79; 743,144; 0,0048; -0,0044; 0,9319; Output: 0,543; Desired output: 0; Error: 0,543;
Input: 205,9444; 30,9389; -0,7749; 0,0371; 0,0664; 324,3581; 0,8212; 0,0006; 7,5967; -160,591,966; 3,303,3187; 0,0018; -0,0083; 0,9836; Output: 0,543; Desired output: 0; Error: 0,5
Input: 208,4092; 23,4604; -0,3839; -0,5368; 0,0782; 215,3477; 0,7997; 0,0008; 7,2746; -32,842,3941; 1,935,2657; 0,0022; -0,0078; 0,9806; Output: 0,543; Desired output: 0; Error: 0,
Input: 193,376; 19,9404; -0,0809; -0,0602; 0,0943; 188,7846; 0,7618; 0,0009; 7,1018; -5,294,4535; 1,396,1752; 0,0028; -0,008; 0,9776; Output: 0,543; Desired output: 0; Error: 0,543;
Input: 211,1771; 29,6658; -1,6881; 3,9711; 0,0704; 379,775; 0,7562; 0,0006; 7,6407; -260,300,3656; 2,735,3342; 0,0024; -0,0064; 0,9669; Output: 0,543; Desired output: 0; Error: 0,5
Input: 185,1905; 21,9586; 0,0064; -0,9879; 0,0643; 280,8281; 0,7044; 0,0021; 6,1985; 13,209,2539; 1,619,1611; 0,0053; -0,0127; 0,9952; Output: 0,4651; Desired output: 0; Error: 0,4
Input: 194,5379; 25,0635; -1,9138; 5,762; 0,1038; 237,8144; 0,7792; 0,0006; 7,7333; -134,282,7897; 1,916,4589; 0,0018; -0,0051; 0,9301; Output: 0,543; Desired output: 0; Error: 0,5
Input: 236,1119; 16,3357; -2,8334; 12,7781; 0,1765; 90,3951; 0,8145; 0,0016; 6,8724; -72,719,4435; 884,1474; 0,0056; -0,003; 0,8691; Output: 0,543; Desired output: 0; Error: 0,543;
Input: 201,6618; 39,2496; -1,1693; 1,7088; 0,1102; 297,4607; 0,8964; 0,0007; 7,5641; -489,083,2868; 5,443,2386; 0,0051; -0,0098; 0,9907; Output: 0,543; Desired output: 0; Error: 0,
Input: 181,6683; 38,4009; -0,6301; -0,5839; 0,1112; 169,3416; 0,9408; 0,0009; 7,1872; -267,231,0388; 5,556,069; 0,0032; -0,0131; 0,9959; Output: 0,543; Desired output: 0; Error: 0,
Input: 222,3977; 39,5223; -1,702; 1,5023; 0,2181; 121,1461; 0,9597; 0,0018; 6,9077; -805,761,1132; 5,883,9699; 0,0061; -0,0064; 0,9715; Output: 0,543; Desired output: 0; Error: 0,5
Input: 149,2978; 13,3166; 0,2393; -0,5501; 0,2616; 17,9911; 0,9479; 0,0015; 6,7136; 4,592,9613; 672,4594; 0,0044; -0,0079; 0,952; Output: 0,4651; Desired output: 0; Error: 0,4651;
Input: 113,9505; 21,534; 0,4886; 0,3482; 0,1277; 106,3603; 0,8854; 0,0006; 7,6306; 32,510,1432; 1,749,8471; 0,0031; -0,0109; 0,9577; Output: 0,4651; Desired output: 0; Error: 0,46
Input: 75,524; 10,8023; 0,8642; 0,1283; 0,3283; 12,2761; 0,9452; 0,0034; 6,0542; 7,898,0589; 435,6359; 0,0102; -0,0162; 0,9555; Output: 0,4651; Desired output: 0; Error: 0,4651;
Input: 220,0444; 6,8861; -1,0248; 1,1667; 0,5555; 2,7358; 0,9708; 0,0114; 5,0212; -2,585,7277; 184,9699; 0,041; -0,0065; 0,9711; Output: 0,543; Desired output: 0; Error: 0,543;
Input: 152,2897; 9,3657; 0,317; -0,5354; 0,31; 11,2292; 0,9358; 0,0028; 6,1548; 2,144,589; 338,7344; 0,0076; -0,007; 0,9379; Output: 0,4651; Desired output: 0; Error: 0,4651;
Input: 154,6726; 8,2907; 0,1266; -0,3534; 0,3294; 9,6398; 0,9294; 0,0033; 6,0047; 580,1925; 263,5189; 0,0075; -0,0066; 0,9331; Output: 0,4651; Desired output: 0; Error: 0,4651;
Input: 148,5788; 24,4708; -2,4184; 6,0868; 0,2842; 27,9629; 0,9741; 0,0027; 6,5223; -245,876,6616; 2,135,3931; 0,0099; -0,0101; 0,9755; Output: 0,543; Desired output: 0; Error: 0,5
Input: 175,4577; 9,4608; 0,027; -0,5815; 0,3119; 10,9676; 0,9388; 0,0027; 6,17; 263,7394; 347,6167; 0,0067; -0,0062; 0,9415; Output: 0,4651; Desired output: 0; Error: 0,4651;
Input: 146,3604; 10,2862; -0,2086; -0,1305; 0,2486; 19,1855; 0,9082; 0,0019; 6,5165; -1,597,2708; 398,9516; 0,0056; -0,0063; 0,9185; Output: 0,543; Desired output: 0; Error: 0,543;
Total Mean Square Error: 0,23860983492404692
    
```

Figura 33. Total MSE en Neuroph Studio.

Resultados 2:

Para el reconocimiento de imágenes se usó la Red Neuronal Perceptrón Multicapa con 14 neuronas en la capa de entrada, 50 neuronas en la capa oculta y una neurona en la capa de salida tomando 9000 iteraciones, un error máximo de 0.01, una tasa de aprendizaje de 0.3 y un momentum de 0.6 esto de prueba para el entrenamiento usando la función sigmoideal.

El gráfico total de error de la red muestra el siguiente aspecto:

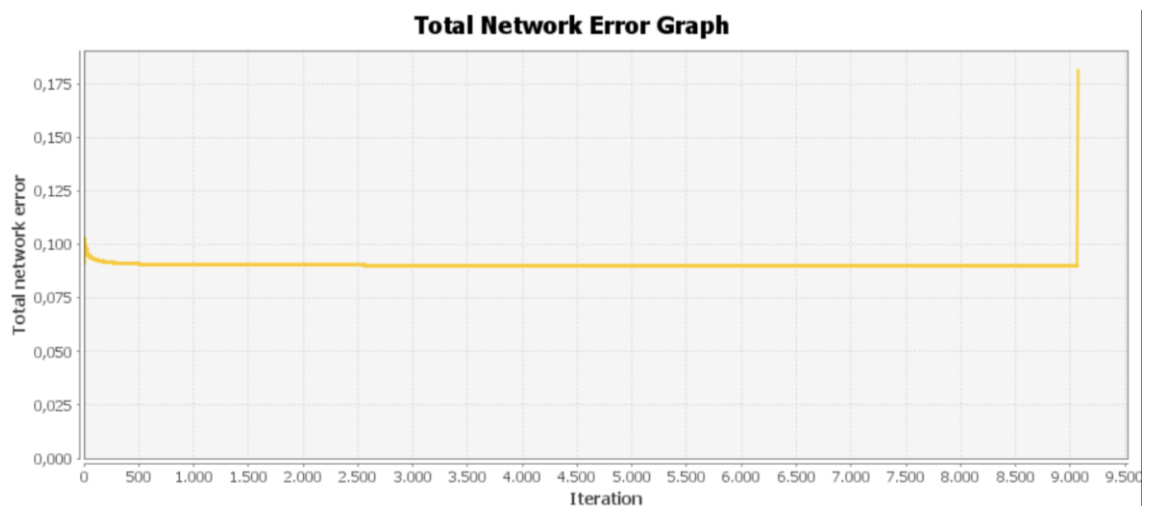


Figura 34. Total de error de la red en Neuroph Studio – resultados 2.

El total de error de la red es superior al máximo error indicado (0.01) el total de error de la red nos da un resultado de 0.17893473543953.

Error cuadrático Medio (MSE): El error medio cuadrático tiene un resultado de 0.3213659900497219 teniendo un valor más alto a la anterior, así que volveremos a probar con diferentes parámetros ajustando las neuronas ocultas, tasa de aprendizaje, momentum, etc.

```

Input: 208,7906; 32,6246; -0,4233; -1,0115; 0,1026; 511,9646; 0,7436; 0,0031; 5,8422; -116,548,344; 3,481,9828; 0,0101; -0,0131; 0,9981; Output: 0,2372; Desired output: 1; Error: -0,2372
Input: 205,8487; 46,6406; -1,0254; -0,0574; 0,1068; 801,5143; 0,7934; 0,0034; 5,8112; -709,814,0078; 6,958,1101; 0,0143; -0,013; 0,998; Output: 0,2372; Desired output: 1; Error: -0,2372
Input: 193,66; 55,6753; -0,6811; -0,976; 0,0797; 806,2228; 0,8706; 0,0015; 6,7076; -1,011,570,6193; 11,652,0417; 0,0082; -0,0134; 0,9974; Output: 0,2372; Desired output: 1; Error: -0,2372
Input: 193,5767; 55,9383; -0,6865; -0,9728; 0,0735; 820,6087; 0,8691; 0,0014; 6,7435; -1,024,633,176; 11,713,5728; 0,0072; -0,0137; 0,9977; Output: 0,2372; Desired output: 1; Error: -0,2372
Input: 211,88; 44,1707; -1,4858; 2,289; 0,0769; 684,1086; 0,8068; 0,0021; 6,5575; -835,829,3906; 6,399,0748; 0,019; -0,0106; 0,9947; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 230,5357; 16,5691; -0,6942; 0,5737; 0,0757; 275,7702; 0,4587; 0,0011; 6,9788; -13,584,79; 743,144; 0,0048; -0,0044; 0,9319; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 205,9444; 30,9389; -0,7749; 0,0371; 0,0664; 324,3581; 0,8212; 0,0006; 7,5967; -160,591,966; 3,303,3187; 0,0018; -0,0083; 0,9836; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 208,4092; 23,4604; -0,3839; -0,5368; 0,0782; 215,3477; 0,7997; 0,0008; 7,2746; -32,842,3941; 1,935,2657; 0,0022; -0,0078; 0,9806; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 193,376; 19,9404; -0,0809; -0,0602; 0,0943; 188,7846; 0,7618; 0,0009; 7,1018; -5,294,4535; 1,396,1752; 0,0028; -0,008; 0,9776; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 211,1771; 29,6658; -1,6881; 3,9711; 0,0704; 379,775; 0,7562; 0,0006; 7,6407; -260,300,3656; 2,735,3342; 0,0024; -0,0064; 0,9669; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 185,1905; 21,9586; 0,0064; -0,9879; 0,0643; 280,8281; 0,7044; 0,0021; 6,1985; 13,209,2539; 1,619,1611; 0,0053; -0,0127; 0,9952; Output: 0,4501; Desired output: 0; Error: 0,4501
Input: 194,5379; 25,0635; -1,9138; 5,762; 0,1038; 237,8144; 0,7792; 0,0006; 7,7333; -134,282,7897; 1,916,4589; 0,0018; -0,0051; 0,9301; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 236,1119; 16,3357; -2,8334; 12,7781; 0,1765; 90,3951; 0,8145; 0,0016; 6,8724; -72,719,4435; 884,1474; 0,0056; -0,003; 0,8691; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 201,6618; 39,2496; -1,1693; 1,7088; 0,1102; 297,4607; 0,8964; 0,0007; 7,5641; -489,083,2868; 5,443,2386; 0,0051; -0,0098; 0,9907; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 181,6683; 38,4009; -0,6301; -0,5839; 0,1112; 169,3416; 0,9408; 0,0009; 7,1872; -267,231,0388; 5,556,069; 0,0032; -0,0131; 0,9959; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 222,3977; 39,5223; -1,702; 1,5023; 0,2181; 121,1461; 0,9597; 0,0018; 6,9077; -805,761,1132; 5,883,9699; 0,0061; -0,0064; 0,9715; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 149,2978; 13,3166; 0,2393; -0,5501; 0,2616; 17,9911; 0,9479; 0,0015; 6,7136; 4,592,9613; 672,4594; 0,0044; -0,0079; 0,952; Output: 0,4501; Desired output: 0; Error: 0,4501
Input: 113,9505; 21,534; 0,4886; 0,3482; 0,1277; 106,3603; 0,8854; 0,0006; 7,6306; 32,510,1432; 1,749,8471; 0,0031; -0,0109; 0,9577; Output: 0,4501; Desired output: 0; Error: 0,4501
Input: 75,524; 10,8023; 0,8642; 0,1283; 0,3283; 12,2761; 0,9452; 0,0034; 6,0542; 7,898,0589; 435,6359; 0,0102; -0,0162; 0,9555; Output: 0,4501; Desired output: 0; Error: 0,4501
Input: 220,0444; 6,8861; -1,0248; 1,1667; 0,5555; 2,7358; 0,9708; 0,0114; 5,0212; -2,585,7277; 184,9699; 0,041; -0,0065; 0,9711; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 152,2897; 9,3657; 0,317; -0,5354; 0,31; 11,2292; 0,9358; 0,0028; 6,1548; 2,144,589; 338,7344; 0,0076; -0,007; 0,9379; Output: 0,4501; Desired output: 0; Error: 0,4501
Input: 154,6726; 8,2907; 0,1266; -0,3534; 0,3294; 9,6398; 0,9294; 0,0033; 6,0047; 580,1925; 263,5189; 0,0075; -0,0066; 0,9331; Output: 0,0259; Desired output: 0; Error: 0,0259
Input: 148,5788; 24,4708; -2,4184; 6,0868; 0,2842; 27,9629; 0,9741; 0,0027; 6,5223; -245,876,6616; 2,135,3931; 0,0099; -0,0101; 0,9755; Output: 0,2372; Desired output: 0; Error: 0,2372
Input: 175,4577; 9,4608; 0,027; -0,5815; 0,3119; 10,9676; 0,9388; 0,0027; 6,17; 263,7394; 347,6167; 0,0067; -0,0062; 0,9415; Output: 0,0213; Desired output: 0; Error: 0,0213
Input: 146,3604; 10,2862; -0,2086; -0,1305; 0,2486; 19,1855; 0,9082; 0,0019; 6,5165; -1,597,2708; 398,9516; 0,0056; -0,0063; 0,9185; Output: 0,2372; Desired output: 0; Error: 0,2372
Total Mean Square Error: 0.3213659900497219

```

Figura 35. Total MSE en Neuroph Studio - resultados 2.

Resultados 3:

Para el reconocimiento de imágenes se usó la Red Neuronal Perceptrón Multicapa con 14 neuronas en la capa de entrada, 300 neuronas en la capa oculta y una neurona en la capa de salida tomando 5000 iteraciones, un error máximo de 0.01, una tasa de aprendizaje de 0.5 y un momentum de 0.7 esto de prueba para el entrenamiento usando la función sigmoideal.

El gráfico total de error de la red muestra el siguiente aspecto:

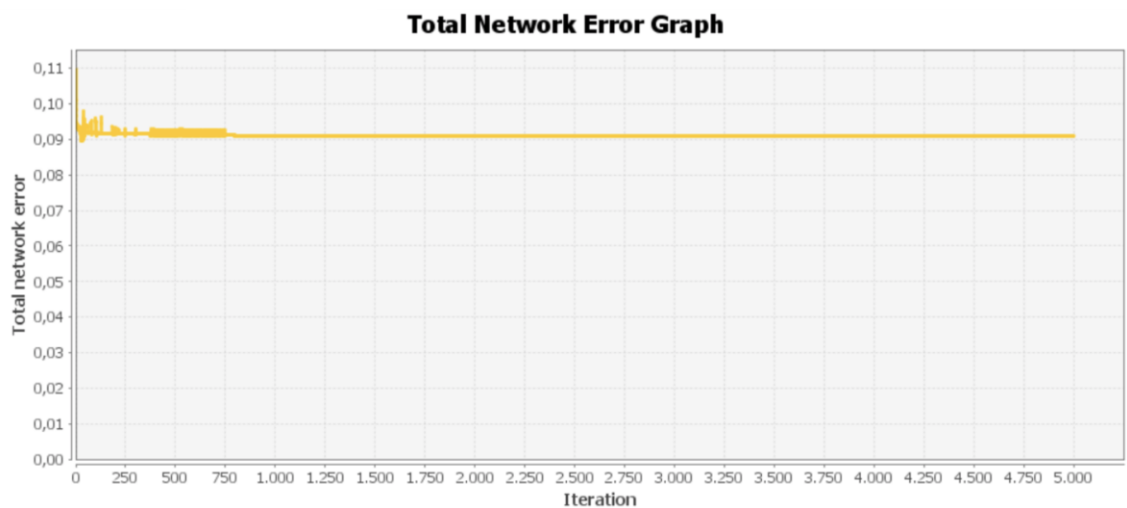


Figura 36. Total de error de la red en Neuroph Studio – resultados 3.

El total de error de la red es superior al máximo error indicado (0.01) el total de error de la red nos da un resultado de 0.091834753453475, obteniendo un resultado mejor a los anteriores.

Error cuadrático Medio (MSE): El error medio cuadrático tiene un resultado de 0.26813133912122853 teniendo un nivel bajo de error a diferencia de los resultados anteriores.

```

Input: 208,7906; 32,6246; -0,4233; -1,0115; 0,1026; 511,9646; 0,7436; 0,0031; 5,8422; -116,548,344; 3,481,9828; 0,0101; -0,0131; 0,9981; Output: 0,1858; Desired output: 1; Error: -0,1858;
Input: 205,8487; 46,6406; -1,0254; -0,0574; 0,1068; 801,5143; 0,7934; 0,0034; 5,8112; -709,814,0078; 6,958,1101; 0,0143; -0,013; 0,998; Output: 0,1858; Desired output: 1; Error: -0,1858;
Input: 193,66; 55,6753; -0,6811; -0,976; 0,0797; 806,2228; 0,8706; 0,0015; 6,7076; -1,011,570,6193; 11,652,0417; 0,0082; -0,0134; 0,9974; Output: 0,1858; Desired output: 1; Error: -0,1858;
Input: 193,5767; 55,9383; -0,6865; -0,9728; 0,0735; 820,6087; 0,8691; 0,0014; 6,7435; -1,024,633,176; 11,713,5728; 0,0072; -0,0137; 0,9977; Output: 0,1858; Desired output: 1; Error: -0,1858;
Input: 211,88; 44,1707; -1,4858; 2,289; 0,0769; 684,1086; 0,8068; 0,0021; 6,5575; -835,829,3906; 6,399,0748; 0,019; -0,0106; 0,9947; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 230,5357; 16,5691; -0,6942; 0,5737; 0,0757; 275,7702; 0,4587; 0,0011; 6,9788; -13,584,79; 743,144; 0,0048; -0,0044; 0,9319; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 205,9444; 30,9389; -0,7749; 0,0371; 0,0664; 324,3581; 0,8212; 0,0006; 7,5967; -160,591,966; 3,303,3187; 0,0018; -0,0083; 0,9836; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 208,4092; 23,4604; -0,3839; -0,5368; 0,0782; 215,3477; 0,7997; 0,0008; 7,2746; -32,842,3941; 1,935,2657; 0,0022; -0,0078; 0,9806; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 193,376; 19,9404; -0,0809; -0,0602; 0,0943; 188,7846; 0,7618; 0,0009; 7,1018; -5,294,4535; 1,396,1752; 0,0028; -0,008; 0,9776; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 211,1771; 29,6658; -1,6881; 3,9711; 0,0704; 379,775; 0,7562; 0,0006; 7,6407; -260,300,3656; 2,735,3342; 0,0024; -0,0064; 0,9669; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 185,1905; 21,9586; 0,0064; -0,9879; 0,0643; 280,8281; 0,7044; 0,0021; 6,1985; 13,209,2539; 1,619,1611; 0,0053; -0,0127; 0,9952; Output: 0,2137; Desired output: 0; Error: 0,2137;
Input: 194,5379; 25,0635; -1,9138; 5,762; 0,1038; 237,8144; 0,7792; 0,0006; 7,7333; -134,282,7897; 1,916,4589; 0,0018; -0,0051; 0,9301; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 236,1119; 16,3357; -2,8334; 12,7781; 0,1765; 90,3951; 0,8145; 0,0016; 6,8724; -72,719,4435; 884,1474; 0,0056; -0,003; 0,8691; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 201,6618; 39,2496; -1,1693; 1,7088; 0,1102; 297,4607; 0,8964; 0,0007; 7,5641; -489,083,2868; 5,443,2386; 0,0051; -0,0098; 0,9907; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 181,6683; 38,4009; -0,6301; -0,5839; 0,1112; 169,3416; 0,9408; 0,0009; 7,1872; -267,231,0388; 5,556,069; 0,0032; -0,0131; 0,9959; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 222,3977; 39,5223; -1,702; 1,5023; 0,2181; 121,1461; 0,9597; 0,0018; 6,9077; -805,761,1132; 5,883,9699; 0,0061; -0,0064; 0,9715; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 149,2978; 13,3166; 0,2393; -0,5501; 0,2616; 17,9911; 0,9479; 0,0015; 6,7136; 4,592,9613; 672,4594; 0,0044; -0,0079; 0,952; Output: 0,2137; Desired output: 0; Error: 0,2137;
Input: 113,9505; 21,534; 0,4886; 0,3482; 0,1277; 106,3603; 0,8854; 0,0006; 7,6306; 32,510,1432; 1,749,8471; 0,0031; -0,0109; 0,9577; Output: 0,2137; Desired output: 0; Error: 0,2137;
Input: 75,524; 10,8023; 0,8642; 0,1283; 0,3283; 12,2761; 0,9452; 0,0034; 6,0542; 7,898,0589; 435,6359; 0,0102; -0,0162; 0,9555; Output: 0,2137; Desired output: 0; Error: 0,2137;
Input: 220,0444; 6,8861; -1,0248; 1,1667; 0,5555; 2,7358; 0,9708; 0,0114; 5,0212; -2,585,7277; 184,9699; 0,041; -0,0065; 0,9711; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 152,2897; 9,3657; 0,317; -0,5354; 0,31; 11,2292; 0,9358; 0,0028; 6,1548; 2,144,589; 338,7344; 0,0076; -0,007; 0,9379; Output: 0,2137; Desired output: 0; Error: 0,2137;
Input: 154,6726; 8,2907; 0,1266; -0,3534; 0,3294; 9,6398; 0,9294; 0,0033; 6,0047; 580,1925; 263,5189; 0,0075; -0,0066; 0,9331; Output: 0,2137; Desired output: 0; Error: 0,2137;
Input: 148,5788; 24,4708; -2,4184; 6,0868; 0,2842; 27,9629; 0,9741; 0,0027; 6,5223; -245,876,6616; 2,135,3931; 0,0099; -0,0101; 0,9755; Output: 0,1858; Desired output: 0; Error: 0,1858;
Input: 175,4577; 9,4608; 0,027; -0,5815; 0,3119; 10,9676; 0,9388; 0,0027; 6,17; 263,7394; 347,6167; 0,0067; -0,0062; 0,9415; Output: 0,2137; Desired output: 0; Error: 0,2137;
Input: 146,3604; 10,7862; -0,7086; -0,1305; 0,2486; 19,1855; 0,9082; 0,0019; 6,5165; -1,597,2708; 398,9516; 0,0056; -0,0063; 0,9185; Output: 0,1858; Desired output: 0; Error: 0,1858;
Total Mean Square Error: 0.26813133912122853

```

Figura 37. Total MSE en Neuroph Studio - resultados 3.

Segundo Método

Combinación 1

Para el segundo método se han usado los mismos parámetros asignados de los resultados del método anterior.

Se han utilizado las siguientes técnicas:

Técnicas de Pre procesamiento: Dilatación y Escalado de Grises

Técnica de Segmentación: Thresholding

Extracción de Características: Para el reconocimiento de microcalcificaciones benignas o malignas se extrajeron los siguientes resultados: Estadísticos (Mean, desviación, skewness, kurtosis) y por GLCM(Contraste, correlación, máxima probabilidad, energía, entropía, homogeneidad, cluster shade, cluster tendency, correlación info 1, correlación info 2).

Resultados 1:

Para el reconocimiento de imágenes se usó la Red Neuronal Perceptrón Multicapa con 14 neuronas en la capa de entrada, 10 neuronas en la capa oculta y una neurona en la capa de salida tomando

45000 iteraciones, un error máximo de 0.01, una tasa de aprendizaje de 0.2 y un momentum de 0.7 esto de prueba para el entrenamiento usando la función sigmoïdal.

El gráfico del error total de la red presenta el siguiente aspecto:

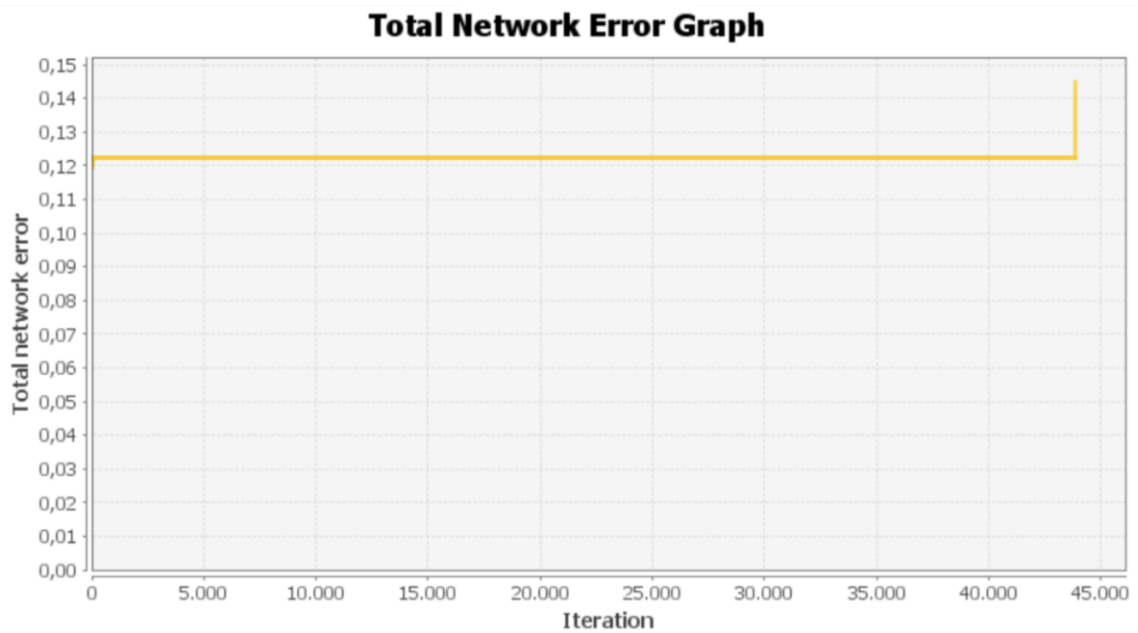


Figura 38. Método 2 - Gráfico Total de error de la red en Neuroph Studio.

El total de error de la red es superior al máximo error indicado (0.01) el total de error de la red nos da un resultado de 0.1452847344561.

Error cuadrático Medio (MSE): el error medio cuadrático tiene un resultado de 0.25576550814143656. Siendo esto un poco alto a diferencia del primero resultado en el primer método.


```

Input: 118,1262; 14,9794; 2,201; 11,7985; 0,1898; 165,9488; 0,6365; 0,0015; 6,8609; 28,744,7737; 747,1504; 0,0052; -0,0085; 0,9302; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 93,8905; 8,8879; 1,4918; 6,4288; 0,2543; 45,9432; 0,7168; 0,0045; 5,6953; 3,828,382; 278,5277; 0,0159; -0,0125; 0,9514; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 188,0107; 25,7714; 0,2362; -0,373; 0,1171; 137,2482; 0,8973; 0,0005; 7,7789; 31,189,6695; 2,535,2643; 0,0019; -0,0076; 0,9708; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 134,4862; 8,041; 2,2017; 12,3798; 0,1995; 57,487; 0,5795; 0,0039; 5,78; 3,582,3059; 215,9025; 0,0108; -0,0061; 0,899; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 170,2967; 15,9204; 1,3866; 4,686; 0,1543; 143,0616; 0,7226; 0,0019; 6,4248; 26,191,8573; 888,3896; 0,0091; -0,0087; 0,9745; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 122,6119; 23,0842; 0,1811; 0,0149; 0,1162; 197,48; 0,8133; 0,0005; 7,7584; 3,373,0944; 1,918,2943; 0,0018; -0,0103; 0,9593; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 154,0508; 12,1208; 3,625; 17,7625; 0,2135; 131,255; 0,5792; 0,0034; 5,9055; 34,572,154; 492,5869; 0,0117; -0,0056; 0,9072; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 131,5985; 12,199; 2,026; 4,0895; 0,1818; 113,7875; 0,637; 0,0038; 5,8947; 19,234,6953; 513,2082; 0,0125; -0,0078; 0,935; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 35,9711; 4,6135; 4,4648; 43,0124; 0,3924; 33,7; 0,2408; 0,0148; 4,6221; 860,256; 55,0777; 0,0364; -0,016; 0,8263; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 203,1515; 22,9026; -0,4717; -0,2954; 0,1317; 157,4231; 0,8437; 0,0011; 6,9357; -45,529,3551; 1,856,4911; 0,0046; -0,0089; 0,9869; Output: 0,4532; Desired output: 0; Error: 0,4681
Input: 116,6341; 14,0152; 1,6983; 8,0994; 0,1723; 157,4696; 0,6035; 0,0014; 6,8538; 14,271,3869; 636,8765; 0,0041; -0,008; 0,9203; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 171,4701; 18,6301; -0,1248; -0,3912; 0,1339; 128,7128; 0,8102; 0,0008; 7,2834; -14,868,7156; 1,227,6616; 0,0029; -0,0071; 0,9557; Output: 0,4532; Desired output: 0; Error: 0,4681
Input: 155,5165; 5,8604; 1,1101; 9,1953; 0,2412; 38,0136; 0,4503; 0,0054; 5,4675; 71,7568; 100,2822; 0,0136; -0,0041; 0,8481; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 82,0887; 12,0632; 8,0146; 91,5731; 0,2728; 231,913; 0,2535; 0,0055; 5,4329; 45,450,0044; 389,4178; 0,0151; -0,0093; 0,8842; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 70,6181; 7,0086; 3,7598; 21,8108; 0,4094; 48,8344; 0,5301; 0,0147; 4,8376; 5,205,6095; 159,0148; 0,0745; -0,0124; 0,9099; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 129,8876; 33,5458; -0,3024; -0,9245; 0,1248; 96,6759; 0,9552; 0,0007; 7,4795; -96,685,403; 4,220,1753; 0,0026; -0,0156; 0,9913; Output: 0,4532; Desired output: 0; Error: 0,4681
Input: 103,295; 19,0976; 0,2747; 0,4638; 0,2126; 80,1972; 0,8891; 0,001; 7,2333; 4,140,358; 1,366,0837; 0,0051; -0,0122; 0,9591; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 90,8741; 17,0078; 1,2874; 2,325; 0,2141; 90,868; 0,8447; 0,0014; 7,002; 37,059,0815; 1,079,647; 0,0062; -0,0127; 0,9491; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 70,9887; 10,7216; 2,6636; 20,2891; 0,3214; 136,5404; 0,4272; 0,0048; 5,6369; 3,031,0204; 340,1709; 0,0248; -0,017; 0,9541; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 67,5217; 5,1901; 5,5126; 53,0309; 0,4353; 37,5498; 0,3405; 0,0162; 4,6147; 2,350,4941; 76,3291; 0,0671; -0,0083; 0,8213; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 201,4512; 16,1343; 0,2831; 0,6849; 0,1313; 112,766; 0,7829; 0,0013; 6,8566; 804,9236; 926,2217; 0,0043; -0,0068; 0,9675; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 83,206; 7,9847; 5,1804; 44,0905; 0,2923; 93,8302; 0,3085; 0,0059; 5,3695; 7,548,2559; 177,532; 0,0135; -0,0072; 0,8359; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 97,3187; 4,5088; -0,0609; -0,8133; 0,3048; 11,9236; 0,7064; 0,0104; 4,7793; 33,7223; 69,3101; 0,0243; -0,0087; 0,9035; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 57,2593; 18,1458; 7,2415; 62,6891; 0,4062; 510,2108; 0,28; 0,015; 4,7877; 175,638,1315; 907,0734; 0,0462; -0,0123; 0,869; Output: 0,4681; Desired output: 0; Error: 0,4681
Input: 85,6694; 8,2956; 1,169; 5,404; 0,2196; 45,4795; 0,6681; 0,0044; 5,6936; 1,818,9242; 228,6032; 0,0136; -0,0109; 0,92; Output: 0,4681; Desired output: 0; Error: 0,4681
Total Mean Square Error: 0.25576550814143656

```

Figura 39. Método 2 - Total MSE en Neuroph Studio.

Resultados 2:

Para el reconocimiento de imágenes se usó la Red Neuronal Perceptrón Multicapa con 14 neuronas en la capa de entrada, 50 neuronas en la capa oculta y una neurona en la capa de salida tomando 9000 iteraciones, un error máximo de 0.01, una tasa de aprendizaje de 0.3 y un momentum de 0.6 esto de prueba para el entrenamiento usando la función sigmoideal.

El gráfico total de error de la red muestra el siguiente aspecto:

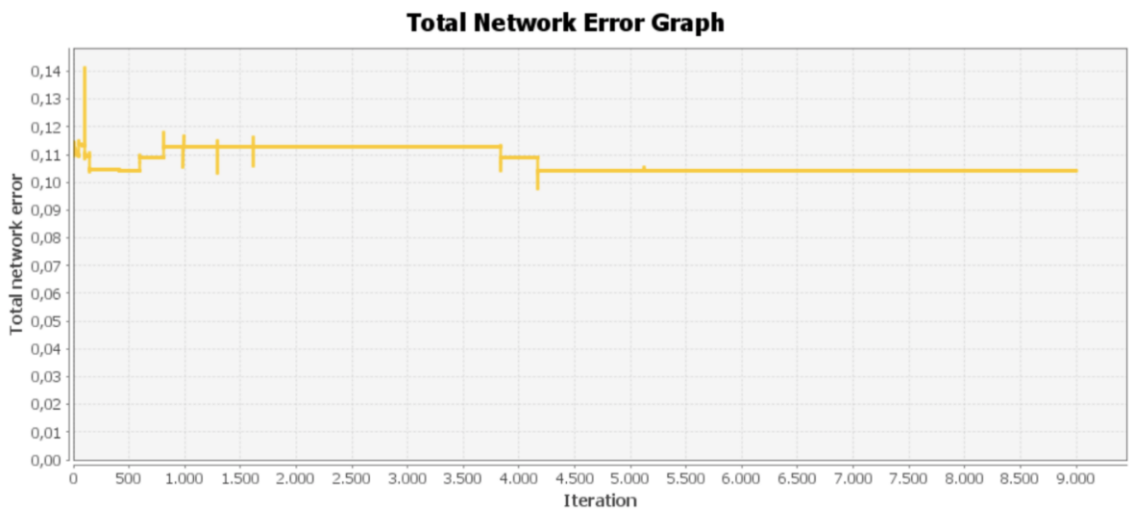


Figura 40. Método 2 - Total de error de la red en Neuroph Studio – resultados 2.

El total de error de la red es superior al máximo error indicado (0.01) el total de error de la red nos da un resultado de 0.1048353745573.

Error cuadrático Medio (MSE): El error medio cuadrático tiene un resultado de

0.3219766199933584, valor favorable.

```
Input: 118,1262; 14,9794; 2,201; 11,7985; 0,1898; 165,9488; 0,6365; 0,0015; 6,8609; 28,744,7737; 747,1504; 0,0052; -0,0085; 0,9302; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 93,8905; 8,8879; 1,4918; 6,4288; 0,2543; 45,9432; 0,7168; 0,0045; 5,6953; 3,828,382; 278,5277; 0,0159; -0,0125; 0,9514; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 188,0107; 25,7714; 0,2362; -0,373; 0,1171; 137,2482; 0,8973; 0,0005; 7,7789; 31,189,6695; 2,535,2643; 0,0019; -0,0076; 0,9708; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 134,4862; 8,041; 2,2017; 12,3798; 0,1995; 57,487; 0,5795; 0,0039; 5,78; 3,582,3059; 215,9025; 0,0108; -0,0061; 0,899; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 170,2967; 15,9204; 1,3866; 4,686; 0,1543; 143,0616; 0,7226; 0,0019; 6,4248; 26,191,8573; 888,3896; 0,0091; -0,0087; 0,9745; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 122,6119; 23,0842; 0,1811; 0,0149; 0,1162; 197,48; 0,8133; 0,0005; 7,7584; 3,373,0944; 1,918,2943; 0,0018; -0,0103; 0,9593; Output: 0,0032; Desired output: 0; Error: 0,0032;
Input: 154,0508; 12,1208; 3,625; 17,7625; 0,2135; 131,255; 0,5792; 0,0034; 5,9055; 34,572,154; 492,5869; 0,0117; -0,0056; 0,9072; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 131,5985; 12,199; 2,026; 4,0895; 0,1818; 113,7875; 0,637; 0,0038; 5,8947; 19,234,6953; 513,2082; 0,0125; -0,0078; 0,935; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 35,9711; 4,6135; 4,4648; 43,0124; 0,3924; 33,7; 0,2408; 0,0148; 4,6221; 860,256; 55,0777; 0,0364; -0,016; 0,8263; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 203,1515; 22,9026; -0,4717; -0,2954; 0,1317; 157,4231; 0,8437; 0,0011; 6,9357; -45,529,3551; 1,856,4911; 0,0046; -0,0089; 0,9869; Output: 0,399; Desired output: 0; Error: 0,399;
Input: 116,6341; 14,0152; 1,6983; 8,0994; 0,1723; 157,4696; 0,6035; 0,0014; 6,8538; 14,271,3869; 636,8765; 0,0041; -0,008; 0,9203; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 171,4701; 18,6301; -0,1248; -0,3912; 0,1339; 128,7128; 0,8102; 0,0008; 7,2834; -14,868,7156; 1,227,6616; 0,0029; -0,0071; 0,9557; Output: 0,399; Desired output: 0; Error: 0,399;
Input: 155,5165; 5,8604; 1,1101; 9,1953; 0,2412; 38,0136; 0,4503; 0,0054; 5,4675; 71,7568; 100,2822; 0,0136; -0,0041; 0,8481; Output: 0,0047; Desired output: 0; Error: 0,0047;
Input: 82,0887; 12,0632; 8,0146; 91,5731; 0,2728; 231,913; 0,2535; 0,0055; 5,4329; 45,450,0044; 389,4178; 0,0151; -0,0093; 0,8842; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 70,6181; 7,0086; 3,7598; 21,8108; 0,4094; 48,8344; 0,5301; 0,0147; 4,8376; 5,205,6095; 159,0148; 0,0745; -0,0124; 0,9099; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 129,8876; 33,5458; -0,3024; -0,9245; 0,1248; 96,6759; 0,9552; 0,0007; 7,4795; -96,685,403; 4,220,1753; 0,0026; -0,0156; 0,9913; Output: 0,399; Desired output: 0; Error: 0,399;
Input: 103,295; 19,0976; 0,2747; 0,4638; 0,2126; 80,1972; 0,8891; 0,001; 7,2333; 4,140,358; 1,366,0837; 0,0051; -0,0122; 0,9591; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 90,8741; 17,0078; 1,2874; 2,325; 0,2141; 90,868; 0,8447; 0,0014; 7,002; 37,059,0815; 1,079,647; 0,0062; -0,0127; 0,9499; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 70,9887; 10,7216; 2,6636; 20,2891; 0,3214; 136,5404; 0,4272; 0,0048; 5,6369; 3,031,0204; 340,1709; 0,0248; -0,017; 0,9541; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 67,5217; 5,1901; 5,5126; 53,0309; 0,4353; 37,5498; 0,3405; 0,0162; 4,6147; 2,350,4941; 76,3291; 0,0671; -0,0083; 0,8213; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 201,4512; 16,1343; 0,2831; 0,6849; 0,1313; 112,766; 0,7829; 0,0013; 6,8566; 804,9236; 926,2217; 0,0043; -0,0068; 0,9675; Output: 0,0032; Desired output: 0; Error: 0,0032;
Input: 83,206; 7,9847; 5,1804; 44,0905; 0,2923; 93,8302; 0,3085; 0,0059; 5,3695; 7,548,2559; 177,532; 0,0135; -0,0072; 0,8359; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 97,3187; 4,5088; -0,0609; -0,8133; 0,3048; 11,9236; 0,7064; 0,0104; 4,7793; 33,7223; 69,3101; 0,0243; -0,0087; 0,9035; Output: 0,0057; Desired output: 0; Error: 0,0057;
Input: 57,2593; 18,1458; 7,2415; 62,6891; 0,4062; 510,2108; 0,28; 0,015; 4,7877; 175,638,1315; 907,0734; 0,0462; -0,0123; 0,869; Output: 0,2751; Desired output: 0; Error: 0,2751;
Input: 85,6694; 8,2956; 1,169; 5,404; 0,2196; 45,4795; 0,6681; 0,0044; 5,6936; 1,818,9242; 228,6032; 0,0136; -0,0109; 0,92; Output: 0,2751; Desired output: 0; Error: 0,2751;
Total Mean Square Error: 0.3219766199933584
```

Figura 41. Método 2 - Total MSE en Neuroph Studio - resultados 2.

Resultados 3:

Para el reconocimiento de imágenes se usó la Red Neuronal Perceptrón Multicapa con 14 neuronas en la capa de entrada, 300 neuronas en la capa oculta y una neurona en la capa de salida tomando 1415 iteraciones, un error máximo de 0.01, una tasa de aprendizaje de 0.5 y un momentum de 0.7 esto de prueba para el entrenamiento usando la función sigmoideal.

El gráfico total de error de la red muestra el siguiente aspecto:

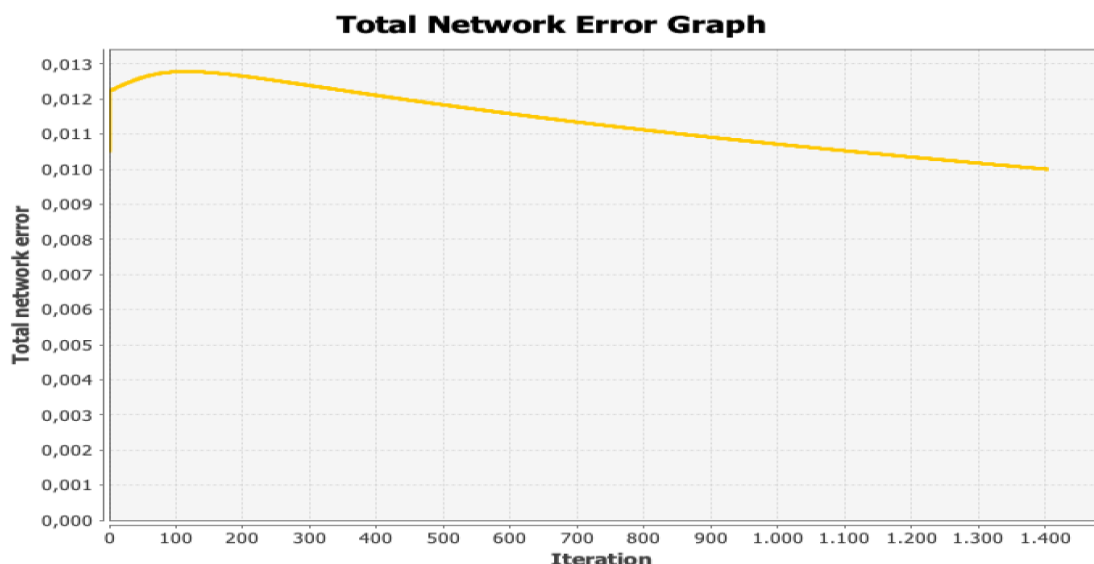


Figura 42. Método 2 - Total de error de la red en Neuroph Studio – resultados 3.

El total de error de la red es inferior al máximo error indicado (0.01) el total de error de la red nos da un resultado de 0.0099104 indicando que el entrenamiento se encuentra muy bien entrenado y completo.

Error cuadrático Medio (MSE): El error medio cuadrático tiene un resultado de 0.01097084082440842 muy favorable con los parámetros indicados así teniendo mejores resultados aplicando técnica de segmentación Thresholding.

```

Input: 0,3093; 0,2201; 0,511; 0,033; 0,1049; 0,0902; 0,7808; 0,1233; 0,390; 0,094; 0,0790; 0,1011; 0,0334; 0,8704; Output: 0,0404; Desired output: 0; Error: 0,0404;
Input: 0,5158; 0,0413; 0,5619; 0,3422; 0,4894; 0,0355; 0,3477; 0,3713; 0,1734; 0,0857; 0,0072; 0,2354; 0,8564; 0,1271; Output: 0,0005; Desired output: 0; Error: 0,0005;
Input: 0,405; 0,2222; 0,4161; 0,1375; 0,2306; 0,1715; 0,5467; 0,077; 0,5277; 0,0952; 0,068; 0,0592; 0,7245; 0,6251; Output: 0,0073; Desired output: 0; Error: 0,0073;
Input: 0,2855; 0,0929; 0,3448; 0,0795; 0,4328; 0,0379; 0,6576; 0,2676; 0,2539; 0,0859; 0,022; 0,2046; 0,5934; 0,7468; Output: 0,0213; Desired output: 0; Error: 0,0213;
Input: 0,7495; 0,4513; 0,2187; 0,006; 0,0028; 0,1395; 0,9069; 0,0175; 0,7434; 0,0961; 0,2437; 0,0138; 0,7538; 0,8579; Output: 0,0623; Desired output: 0; Error: 0,0623;
Input: 0,4857; 0,075; 0,4161; 0,1438; 0,261; 0,0507; 0,4678; 0,2281; 0,2738; 0,0858; 0,0158; 0,1356; 0,801; 0,4457; Output: 0,0002; Desired output: 0; Error: 0,0002;
Input: 0,6622; 0,2422; 0,3342; 0,0607; 0,1193; 0,146; 0,6656; 0,1072; 0,4252; 0,0943; 0,0819; 0,1116; 0,7161; 0,8792; Output: 0,0452; Desired output: 0; Error: 0,0452;
Input: 0,4271; 0,3943; 0,2132; 0,0102; 0,2066; 0,7909; 0,0182; 0,7385; 0,0857; 0,1831; 0,0132; 0,6657; 0,792; Output: 0,0005; Desired output: 0; Error: 0,0005;
Input: 0,5821; 0,1616; 0,5591; 0,202; 0,3049; 0,1329; 0,4675; 0,2004; 0,3032; 0,0974; 0,043; 0,147; 0,8178; 0,4931; Output: 0,1183; Desired output: 0; Error: 0,1183;
Input: 0,4714; 0,1632; 0,3985; 0,0542; 0,2057; 0,1134; 0,5474; 0,2242; 0,3007; 0,0916; 0,045; 0,1583; 0,7455; 0,6523; Output: 0,0036; Desired output: 0; Error: 0,0036;
Input: 0,0022; 0,6434; 0,475; 0,8656; 0,0242; 0,9097; 0,0017; 0,0847; 0,4821; 0,4809; 0,0285; Output: 0; Desired output: 0; Error: 0;
Input: 0,8242; 0,3904; 0,1476; 0,0068; 0,0487; 0,162; 0,8328; 0,0566; 0,5453; 0,0673; 0,177; 0,0517; 0,7095; 0,9503; Output: 0,1577; Desired output: 0; Error: 0,1577;
Input: 0,3977; 0,2018; 0,3656; 0,0976; 0,176; 0,1621; 0,5011; 0,0758; 0,526; 0,0898; 0,0572; 0,0439; 0,7388; 0,568; Output: 0,0002; Desired output: 0; Error: 0,0002;
Input: 0,668; 0,2997; 0,1824; 0,0058; 0,0556; 0,13; 0,7866; 0,0388; 0,627; 0,0788; 0,1152; 0,0282; 0,7696; 0,7713; Output: 0,0021; Desired output: 0; Error: 0,0021;
Input: 0,5894; 0,0287; 0,3065; 0,1094; 0,3916; 0,0291; 0,2894; 0,3226; 0,2003; 0,0844; 0,0044; 0,1737; 0,8671; 0,1534; Output: 0; Desired output: 0; Error: 0;
Input: 0,2274; 0,1603; 1; 1; 0,4908; 0,245; 0,0175; 0,3296; 0,1922; 0,1015; 0,0329; 0,1939; 0,698; 0,3611; Output: 0; Desired output: 0; Error: 0;
Input: 0,1708; 0,0531; 0,5726; 0,2458; 0,9186; 0,0411; 0,3997; 0,9065; 0,0524; 0,0864; 0,0102; 1; 0,5958; 0,5087; Output: 0,0003; Desired output: 0; Error: 0,0003;
Input: 0,463; 0,6163; 0,1646; 0,0271; 0,0944; 0,9869; 0,0272; 0,673; 0,0481; 0,4093; 0,0246; 0,4931; 0,9755; Output: 0,1375; Desired output: 0; Error: 0,1375;
Input: 0,3319; 0,3096; 0,2226; 0,015; 0,302; 0,076; 0,8956; 0,0509; 0,6152; 0,086; 0,1288; 0,058; 0,6048; 0,7905; Output: 0,1394; Desired output: 0; Error: 0,1394;
Input: 0,2707; 0,2653; 0,3243; 0,0351; 0,3067; 0,0879; 0,8343; 0,0745; 0,5609; 0,0983; 0,1007; 0,0723; 0,5891; 0,7335; Output: 0,1185; Desired output: 0; Error: 0,1185;
Input: 0,1726; 0,1319; 0,4625; 0,2293; 0,6431; 0,1388; 0,2574; 0,2843; 0,2401; 0,0856; 0,028; 0,3258; 0,4493; 0,7622; Output: 0,0062; Desired output: 0; Error: 0,0062;
Input: 0,1555; 0,0145; 0,7487; 0,5833; 1; 0,0285; 0,1378; 1; 0,0853; 0,0021; 0,8891; 0,7301; 0; Output: 0; Desired output: 0; Error: 0;
Input: 0,8158; 0,2467; 0,2234; 0,0174; 0,0474; 0,1123; 0,7489; 0,065; 0,5267; 0,0847; 0,0856; 0,0467; 0,7792; 0,8388; Output: 0,0056; Desired output: 0; Error: 0,0056;
Input: 0,2329; 0,0738; 0,7153; 0,4867; 0,5518; 0,0912; 0,0935; 0,3529; 0,1773; 0,0873; 0,012; 0,1713; 0,7654; 0,0834; Output: 0; Desired output: 0; Error: 0;
Input: 0,3024; 0,01889; 0,0012; 0,5911; 0,06433; 0,6345; 0,0387; 0,0844; 0,0014; 0,3185; 0,7161; 0,4719; Output: 0,0033; Desired output: 0; Error: 0,0033;
Input: 0,1049; 0,2894; 0,9224; 0,6877; 0,9088; 0,5548; 0,0542; 0,9227; 0,0406; 0,1504; 0,0837; 0,6149; 0,602; 0,2738; Output: 0,0544; Desired output: 0; Error: 0,0544;
Input: 0,245; 0,0804; 0,3124; 0,0684; 0,3242; 0,0374; 0,5903; 0,2587; 0,2535; 0,0851; 0,0171; 0,1737; 0,6472; 0,5663; Output: 0,0002; Desired output: 0; Error: 0,0002;
Total Mean Square Error: 0.01097084082440842

```

Figura 43. Método 2 - Total MSE en Neuroph Studio - resultados 3.

3.1.2. Tiempo de entrenamiento

Para el entrenamiento final utilizó 17000 milisegundos para su ejecución total.

Pruebas

MC = Microcalcificaciones

Imágenes con microcalcificaciones benignas	Negro
Imágenes con microcalcificaciones malignas	Rojo

A continuación se muestra la salida de la Red Neuronal Artificial(RNA) Perceptrón Multicapa. Si el resultado de la salida se acerca más a 1 quiere decir que encontró una microcalcificación benigna, si el resultado se acerca a 0 quiere decir que encontró una microcalcificación maligna,

a la vez estos resultados lo representaremos como porcentaje.

Método Propuesto dado los Resultados 3

Tabla 4. Método propuesto dado los resultados 3

Imagen MINI- MIAS	MC Salida de la RNA	Porcentaje
1	0.9712834734549382	97.12%
2	0.9826364345343495	98.26%
3	0.9637547562904584	96.37%
4	0.9711356472336756	97.11%
6	0.2892382966293451	28.92%
7	0.9684135745328435	96.84%
8	0.7854457843932364	78.54%
9	0.9673935763449458	96.73%
10	0.9723834763455673	97.23%
11	0.3053234234237542	30.53%
12	0.9720345783453583	97.20%
13	0.6582348573453901	65.82%
14	0.1727843753465106	17.27%
15	0.8521853678914021	85.21%
16	0.1013493457529057	10.13%
17	0.3147467832405682	31.47%
18	0.9105349685476328	91.05%
19	0.6738856640778262	67.38%

20	0.8248637583453457	82.48%
21	0.8932847625323045	89.32%
22	0.9564567839239875	95.64%
23	0.6873496845764195	68.73%
24	0.6593475849038574	65.93%
25	0.3413882720597381	34.13%
26	0.2946457689305873	29.46%
27	0.8723946874654839	87.23%
28	0.1753864756452673	17.53%
29	0.1468837564347078	14.68%
30	0.3728484927563011	37.28%
31	0.8257465734895679	82.57%
32	0.9325458675489433	93.25%
33	0.4017846752356967	40.17%
34	0.8763467854955374	87.63%
35	0.2384549867648930	23.84%
36	0.3656798476327839	36.56%

Matriz de Confusión 1: Aplicando Segmentación Piramidal Con Técnicas de Preprocesamiento

Tabla 5. Matriz de confusión: aplicando segmentación piramidal con técnicas de preprocesamiento.

	Predictivo
--	------------

		Positivo	Negativo	Totales
Actual	Positivo	16	8	24
	Negativo	4	8	12
	Total	20	16	36

Teniendo en cuenta el método:

Criterio Real Positivos (p): 24

Criterio Real Negativos (n): 12

Resultado Verdadero Positivo (Mvp): 16

Resultado Verdadero Negativo (Mvn): 8

Resultado Falso Positivo (Mfp): 4

Resultado Falso Negativo (Mfn): 8

Matriz de confusión 2: aplicando segmentación Thresholding con técnicas de Preprocesamiento (Método Propuesto)

Tabla 6. Matriz de confusión: aplicando Segmentación Thresholding con técnicas de Preprocesamiento (Método Propuesto)

		Predictivo		
		Positivo	Negativo	Totales
Actual	Positivo	20	4	24
	Negativo	3	9	12
	Total	23	13	36

Teniendo en cuenta con el método propuesto:

Criterio Real Positivos (p): 24

Criterio Real Negativos (n): 12

Resultado Verdadero Positivo (Mvp): 20

Resultado Verdadero Negativo (Mvn): 4

Resultado Falso Positivo (Mfp): 3

Resultado Falso Negativo (Mfn): 9

3.1.3. Propuesta de investigación

Se han implementado dos métodos con el objetivo de reconocer patrones tumorales en mamografías, cumpliendo así con el objetivo principal de esta investigación. Para la implementación de estos métodos, se desarrolló un software en los lenguajes de programación C++ y Java, utilizando los IDEs Qt Creator y Netbeans, respectivamente, junto con la biblioteca de visión artificial OpenCV y Neuroph para redes neuronales.

a. Adquisición

Esta estrategia se ha llevado a cabo mediante la separación de imágenes positivas y negativas, obtenidas de la base de datos MI-AS. Esta base de datos es de acceso libre y contiene un total de 332 imágenes en formato .pgm, disponibles en este enlace.

Para las pruebas experimentales, se utilizó una muestra de aproximadamente 36 imágenes. Para determinar cuáles presentan malignidad y cuáles no, se dispone de una lista detallada en la página mencionada.

Primera columna:

Número de referencia de base de datos MIAS.

Segunda columna:

Carácter del tejido de fondo:

F Grasos

G Graso-glandular

D-glandular denso

Tercera columna:

Clase de anormalidad presente:

CALC calcificación

Masas CIRC Well-defined/circumscribed

Masas SPIC espiculada

MISC Otros, masas mal definidas

Distorsión de la arquitectura ARCH

ASYM Asimetría

NORMA normal

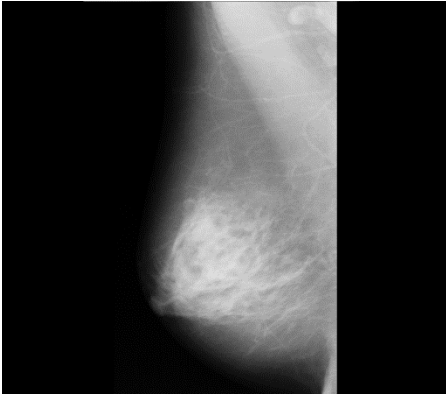
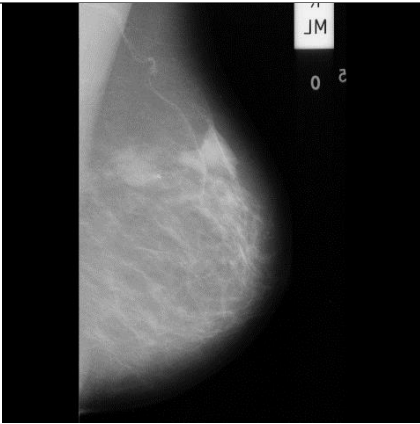
Cuarta columna:

La gravedad de la anomalía

B benigna

M maligno

De esta información detallada, nos valemos para escoger y clasificar las imágenes con las cuales trabajamos para realizar las respectivas pruebas.

Clasificación de la Imagen	Imágenes
<p style="text-align: center;">mdb211.pgm</p> <ul style="list-style-type: none"> - Tipo de tejido de fondo: Graso-Glandular - Tipo de anomalía presente: Calcificación - Severidad de la anomalía: Maligno 	
<p style="text-align: center;">mdb248.pgm</p> <ul style="list-style-type: none"> - Tipo de tejido de fondo: Graso - Tipo de anomalía presente: Calcificación - Severidad de la anomalía: Benigna 	

b. Preprocesamiento

Después de la entrada de la imagen mamográfica, se aplicará una transformación morfológica a partir de una región de interés (ROI). Primero, se utiliza un filtro morfológico de tipo TOP-HAT con un elemento estructural de un kernel de 8x8 para mejorar la ROI de la mamografía. Posteriormente, se convierte la ROI filtrada a una escala de grises de 8 bits, preparándola para la técnica de segmentación.

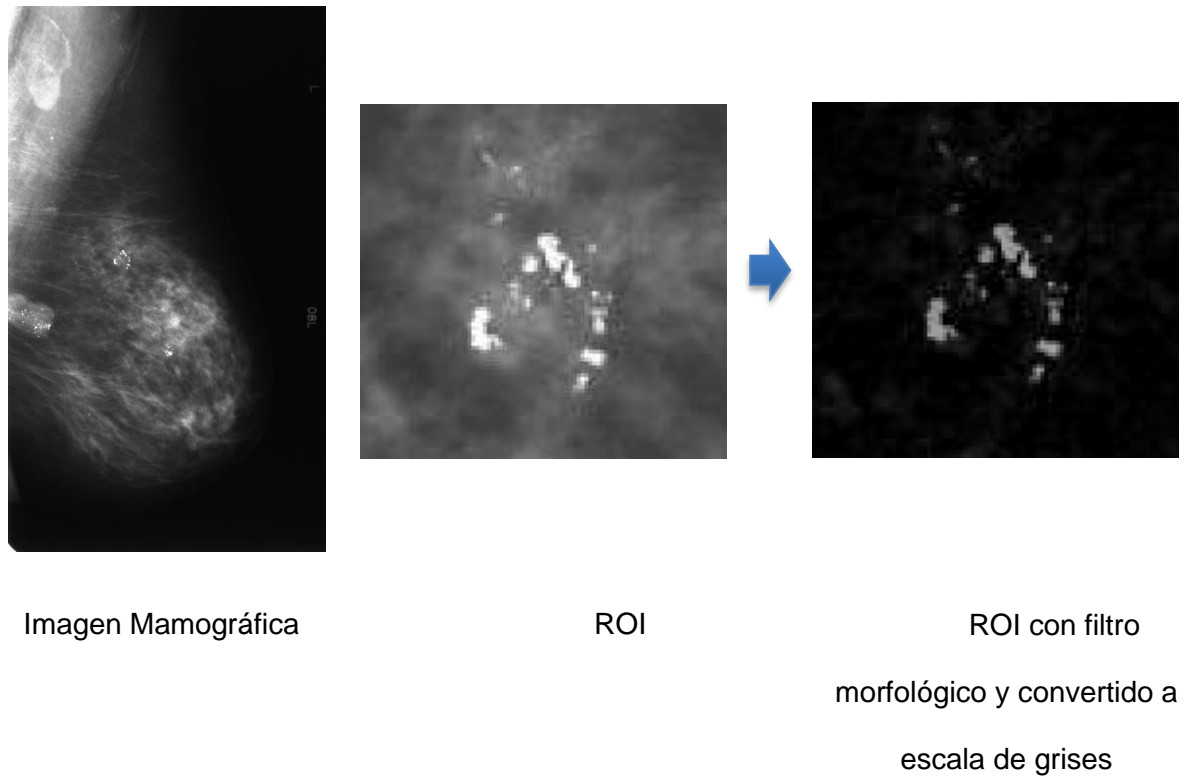
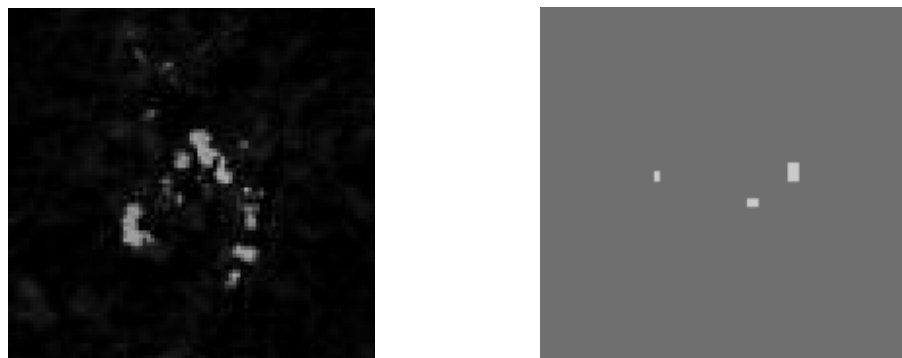


Figura 44. Preprocesamiento del método propuesto.

Una vez obtenida la imagen preprocesada, se procederá a aplicar la técnica de umbralización (thresholding), para lo cual fue necesario convertirla a escala de grises.

Segmentación por Umbral (Threshold)

Después de la etapa de preprocesamiento, se pasa a la etapa de segmentación. En este caso, se ha aplicado la segmentación por umbral, comparándola también con la técnica de segmentación piramidal. A continuación, se presentan los resultados obtenidos con la técnica de segmentación piramidal.

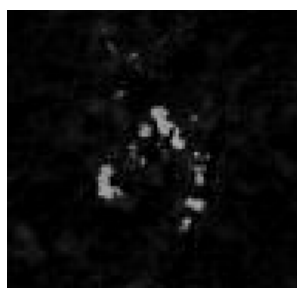


ROI con filtro morfológico y
convertido a escala de grises

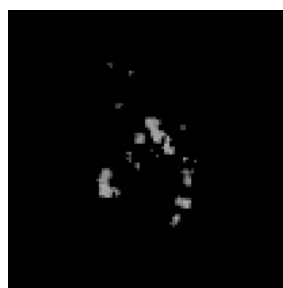
ROI Segmentada (Piramidal)

Figura 45. Segmentación piramidal.

Ahora examinaremos la técnica de Otsu. La binarización es un proceso que reduce la información de una imagen a dos valores: 0 (negro) y 255 (blanco). El método de Otsu determina el umbral óptimo para la segmentación de la imagen, basado en el criterio de máxima varianza entre el fondo y el objeto. Esta umbralización binaria se puede implementar utilizando la función `threshold` de OpenCV.



ROI con filtro morfológico y
convertido a escala de grises



ROI Segmentada (Otsu)

Figura 46. Resultado de la segmentación.

Observamos que la segmentación por umbral ofrece mejores resultados, ya que no solo elimina partes innecesarias para el procesamiento, sino que también mejora la segmentación de las microcalcificaciones adecuadas para nuestros fines. La imagen binaria resultante se dilató con un radio de 300 píxeles para aumentar el área de nuestro objeto y así determinar las ubicaciones exactas que muestran la mayor cantidad de microcalcificaciones a través de contornos.

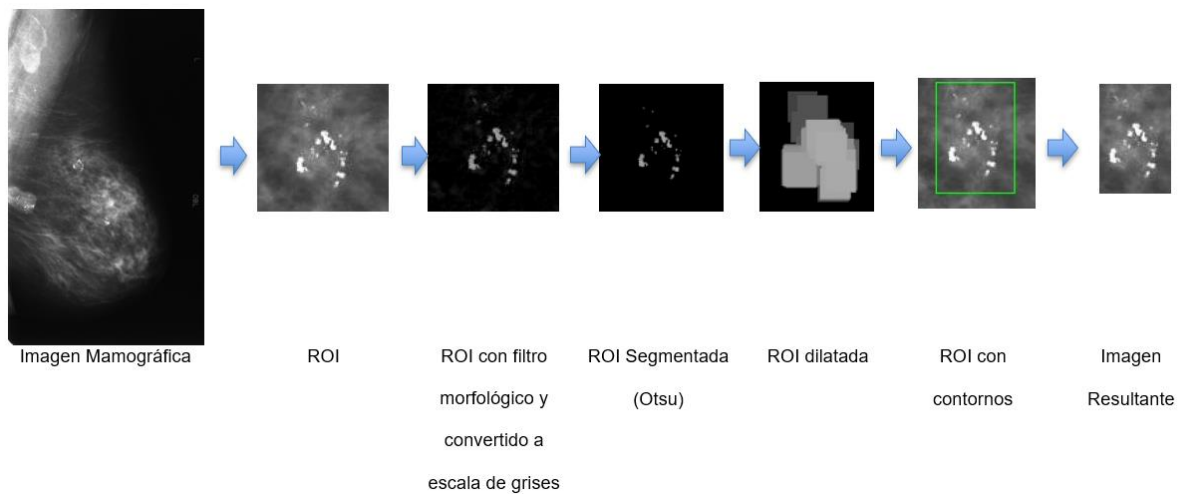


Figura 47. Diagrama de flujo de trabajo de la extracción de la ROI.

A partir de estos resultados, y con la imagen segmentada y la región de interés (ROI) extraída, obtenemos las imágenes resultantes para proceder con la extracción de características y, de este modo, realizar la clasificación.

c. Extracción de características

La extracción de características de imágenes tiene como objetivo proporcionar una interpretación numérica precisa. Todas estas características se definieron como atributos inherentes. En total, se obtuvieron 14 características, que incluyen métricas estadísticas (media, desviación estándar, asimetría, curtosis) y características derivadas de la matriz de co-ocurrencia de niveles de gris (GLCM), tales como contraste, correlación, máxima probabilidad, energía, entropía, homogeneidad, sombreado de conglomerados, tendencia de conglomerados, información de correlación 1 e información de correlación 2. Las características estadísticas consideradas son las siguientes:

Mean

$$\bar{x} = \frac{\sum x}{N}$$

Desviación

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}}, \quad (10)$$

Skewness

$$S = \frac{n \sqrt{n-1}}{n-2} \frac{\sum_{i=1}^n (X_i - X_{avg})^3}{(\sum_{i=1}^n (X_i - X_{avg})^2)^{3/2}} \quad (11)$$

Kurtosis

$$K = \frac{n(n+1)(n-1)}{(n-2)(n-3)} \frac{\sum_{i=1}^n (X_i - X_{avg})^4}{(\sum_{i=1}^n (X_i - X_{avg})^2)^2} \quad (12)$$

Se han tenido en cuenta las siguientes características GLCM:

Se utilizaron los factores de textura.

Contraste

$$Co = \sum_{i,j=0}^{N-1} P_{i,j} (i-j)^2 \quad (13)$$

Correlation

$$\sum_{i,j=0}^{N-1} P_{i,j} \frac{(i - \mu_i) - (j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \quad (14)$$

Cuando el área de una imagen es totalmente uniforme, la varianza GLCM es 0, a lo mismo que la imagen varianza de primer orden es 0. Se obtiene como producto, el denominador de la ecuación de correlación se convierte en 0, y la correlación se vuelve impreciso.

Máxima probabilidad

$$P_m = \max_{i,j} p(i,j) \quad (15)$$

La máxima probabilidad medida reconoce el valor P_{ij} más grande hallado a lo largo de la imagen.

Energía

$$E = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [p(i,j)] \quad (16)$$

En dicha formula se ve a los valores constantes máximos o uniformidad en la asignación periódica de nivel de gris. Dicha fórmula es 1 si la imagen es persistente.

Entropía

$$\sum_{i,j=0}^{N-1} p(i,j) \log(-\ln P_{i,j}) \quad (17)$$

La entropía mide el extravío de advertencia o mensaje en una señal transferida.

Homogeneidad

$$\sum_{i,j=0}^{N-1} \frac{p(i,j)}{1 + (i-j)^2} \quad (18)$$

Como resultado se obtiene el total de la suma de la división del valor del pixel en la posición (i,j) dividido entre la suma del cuadrado de la diferencia de i con j , sumado con 1.

Cluster Shade



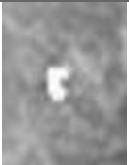
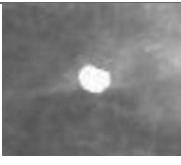
$$CS = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i + j - \mu_x - \mu_y\}^3 \quad (19)$$




Cluster Tendency


$$CP = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i + j - \mu_x - \mu_y\}^4 X \quad (20)$$

Los resultados obtenidos se muestran en el siguiente cuadro.

Tabla 7. Resultados de la extracción de características

ROI	Mean	Desviation	Skewness	Kurtosis	Homogeneidad	Contraste	Correlaci3n
	217.363095	46.965974	-0.956469	-0.640857	0.125120	1607.270833	0.592291
	208.339286	38.459755	-1.094330	0.331029	0.024352	2154.285714	0.146180
	226.125000	42.696128	-1.485931	0.789317	0.065181	3892.125000	-0.215427
	218.163636	56.940563	-1.663590	1.565029	0.110868	3585.272727	0.266566

ROI	Energia	Entropia	ClusterShade	Cluster Tendency	Maximum Probability	Correlation Info 1	Correlati on Info 2
	0.012418	5.065801	- 601302.285181	6277.123794	0.083333	-0.007963	0.98554 2
	0.013039	4.364803	- 227141.734910	2891.943311	0.023810	-0.013772	0.99858 8
	0.022569	3.813439	-97667.810041	2512.414931	0.041667	-0.008216	0.98853 3

	0.009491	4.825841	- 944326.237415	6191.407025	0.022727	-0.007390	0.98240 1
--	----------	----------	--------------------	-------------	----------	-----------	--------------

d. Clasificación

Para la clasificación de los resultados obtenidos en base a la extracción de características se ha escogido la red neuronal artificial (RNA) perceptrón multicapa como clasificador entrenado por el algoritmo BackPropagation desde el Framework Neuroph en Java.

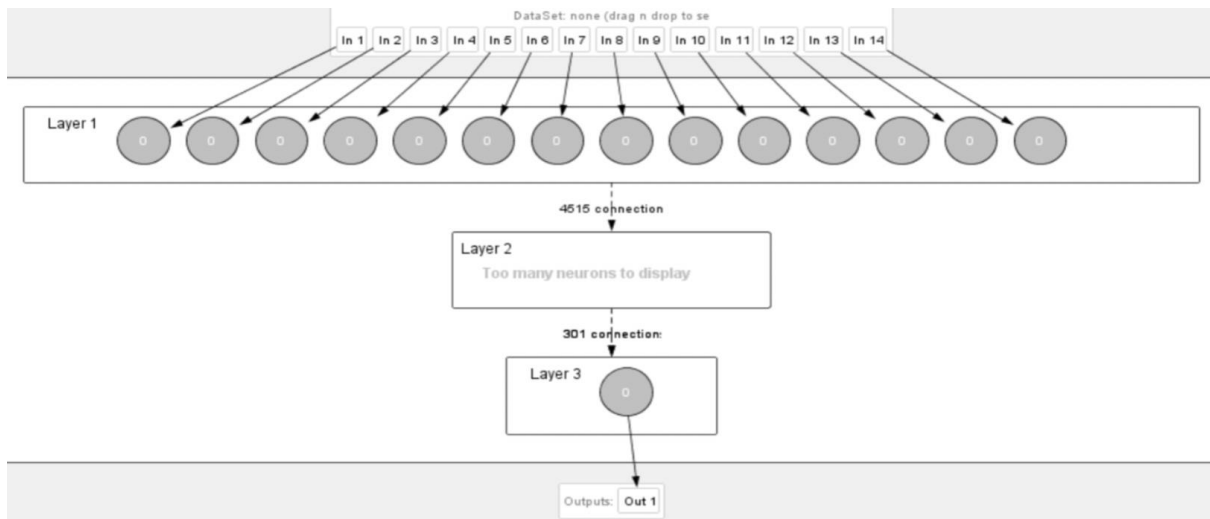


Figura 48. Diseño de la Red Neuronal bajo el Framework Neuroph.

Caracterización de la Red Neuronal Artificial (RNA)

Tipo de aprendizaje: Supervisado

El aprendizaje supervisado se caracteriza por un proceso de entrenamiento controlado por un agente externo (supervisor o maestro) que define la respuesta esperada de la red a partir de una entrada específica. El supervisor verifica la salida de la red y, si esta no coincide con la

respuesta deseada, se ajustan los pesos de las conexiones para lograr que la salida obtenida se aproxime a la esperada.

Algoritmo de aprendizaje: Backpropagation

El término backpropagation proviene del hecho de que el error se propaga hacia atrás a través de la red neuronal, comenzando desde la capa de salida. Este método permite que los pesos de las conexiones entre las neuronas en las capas ocultas se ajusten durante el entrenamiento. Los cambios en los pesos no solo afectan la entrada global, sino también la activación y, por ende, la salida de una neurona. Por lo tanto, es crucial considerar las variaciones en la función de activación cuando se modifican los pesos, fenómeno conocido como la sensibilidad de la función de activación.

Capa de entrada

Una neurona generalmente recibe múltiples entradas. En este caso, el número de entradas corresponde al número de características, que son 14 por imagen, resultando en una capa de entrada con 14 neuronas.

Capa Oculta

La capa oculta puede estar compuesta por varias subcapas internas, las cuales no tienen contacto directo con el entorno exterior. El número de capas ocultas puede variar desde cero hasta un número elevado. En este estudio, se decidió utilizar 300 neuronas en la capa oculta, ya que este número proporcionó mejores resultados.

Capa de Salida

La capa de salida expone la información procesada por la red neuronal al exterior. Se empleó una neurona en la capa de salida para clasificar las microcalcificaciones como benignas o malignas.

Construcción de la RNA

Para la construcción de la RNA, se utilizó el framework Neuroph de Java, implementando una red neuronal Perceptrón Multicapa entrenada mediante el algoritmo de Backpropagation. En este modelo, se introdujeron 14 neuronas en la capa de entrada, correspondientes a las características extraídas de cada imagen.

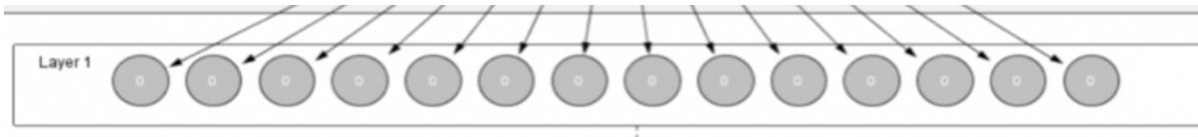


Figura 49. Diseño de la Red Neuronal bajo el Framework Neuroph – Capa de Entrada
Para la capa oculta se obtuvo 300 neuronas que se mencionó anteriormente y para la capa de salida tenemos 1 sola neurona que nos servirá para el resultado

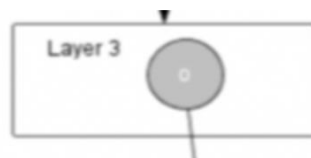


Figura 50. Diseño de la Red Neuronal bajo el Framework Neuroph – Capa de Salida.

El resultado obtenido es un número decimal que se puede aproximar a 1 o 0 según corresponda. Si el valor es cercano a 1, indica que se ha reconocido una microcalcificación benigna. Si el valor es cercano a 0, indica que se ha reconocido una microcalcificación maligna. La creación y entrenamiento de la red neuronal se puede llevar a cabo utilizando la plataforma Neuroph, ya sea en modo gráfico o mediante código Java.

Código java

Para implementar el proceso descrito en Java, primero se definen las 14 características extraídas por imagen en un archivo de texto, donde cada característica se separa por comas y se etiqueta al final con "positive" o "negative", que luego se convertirán en 1 (microcalcificación benigna) o 0 (microcalcificación maligna) respectivamente mediante código Java. Utilizando la

clase Scanner, se lee el archivo para obtener los datos.

```
94. 393604,28.049853,0.419667,-0.356561,0.193564,106.026931,0.932328,0.000624,7.676159,57094.672986,3027.535128,0.002854,-0.017540,0.981599,positive
115.208333,15.337370,0.591567,1.285217,0.110822,147.306159,0.669303,0.001544,6.611851,10202.818298,743.575821,0.007246,-0.012008,0.968864,positive
80.132586,20.201657,0.200669,0.174341,0.193787,90.785693,0.888976,0.000755,7.464425,2576.334919,1544.635484,0.001937,-0.015474,0.957300,positive
77.467391,14.383050,1.030516,2.899196,0.220637,78.395257,0.812262,0.002455,6.228912,15792.048511,756.758831,0.009881,-0.020455,0.979065,positive
111.338526,15.232728,0.843663,1.255459,0.173331,76.798889,0.821894,0.001776,6.568388,17920.107889,785.596636,0.006667,-0.012371,0.967833,positive
85.857438,10.938329,0.887005,1.979100,0.168948,105.504545,0.549754,0.002487,6.158389,4231.314748,363.148161,0.006818,-0.013071,0.945147,positive
53.468377,17.649704,0.922348,1.582944,0.199853,87.468347,0.857437,0.001293,6.944757,32775.672291,1139.615365,0.004482,-0.025658,0.967599,positive
82.268757,12.990681,0.191649,0.888021,0.221342,66.569565,0.795556,0.002668,6.132863,-397.494925,584.657353,0.010870,-0.018272,0.975038,positive
158.767157,32.438089,0.080715,-0.527675,0.125963,121.599147,0.940809,0.000394,8.053590,20070.968123,3987.081449,0.002008,-0.010002,0.979210,positive
54.006944,8.679875,10.707342,164.890170,0.355699,166.977273,-0.022891,0.012676,4.780674,12617.133508,159.503903,0.036932,-0.008586,0.777288,positive
105.750000,16.770051,0.599930,-0.527451,0.162199,71.403162,0.871136,0.001914,6.419257,20451.450985,1036.787280,0.006917,-0.015383,0.980255,positive
88.925769,11.167661,1.224386,2.979798,0.189017,85.102007,0.660575,0.001857,6.606017,7846.011492,416.346437,0.005295,-0.008840,0.890426,positive
44.030992,5.941612,2.597698,12.487061,0.368362,29.170455,0.605004,0.018573,4.971370,2796.497053,118.529499,0.038636,-0.016927,0.880598,positive
111.685950,14.628193,0.645830,0.460702,0.132866,165.593182,0.602519,0.001914,6.380313,8461.795857,667.619582,0.006818,-0.013594,0.976574,positive
97.718336,14.203026,0.743008,1.934605,0.141317,104.498965,0.736765,0.001908,6.401412,8015.634137,689.459109,0.006211,-0.014256,0.969538,positive
53.553719,6.135118,1.358624,8.570560,0.351165,28.520455,0.622125,0.007218,5.244574,589.262105,122.431193,0.025000,-0.016089,0.906643,positive
110.955300,27.194476,1.507507,1.636484,0.178767,168.973937,0.890377,0.001773,6.579615,221156.035787,2913.853041,0.009602,-0.016733,0.987969,positive
62.025974,10.583791,0.934446,0.946465,0.211903,48.858852,0.791404,0.003271,5.956600,5834.165935,419.595642,0.011962,-0.021057,0.963362,positive
94.352564,16.793067,1.468188,3.966012,0.214312,152.851399,0.741794,0.002411,6.265358,35840.121321,1031.096322,0.008741,-0.018007,0.983059,positive
65.541322,13.469753,1.601117,2.670840,0.231056,83.318182,0.766378,0.004631,5.786529,22386.312214,629.954029,0.018182,-0.021521,0.969731,positive
108.362949,18.325574,1.007163,0.973859,0.093216,163.360248,0.758696,0.001678,6.515400,38919.419485,1190.615777,0.005176,-0.015558,0.983369,positive
131.598485,12.199019,2.026813,4.089454,0.181836,113.787500,0.637039,0.003819,5.894736,19234.695316,513.208177,0.012500,-0.007837,0.934976,negative
134.486166,8.040953,2.201736,12.379789,0.199487,57.487013,0.579450,0.003882,5.779968,3582.305927,215.902480,0.010823,-0.006127,0.898982,negative
139.270321,15.162531,1.155478,2.316427,0.168811,92.904762,0.800600,0.002238,6.300146,25463.744557,864.877187,0.008282,-0.010675,0.974021,negative
93.890496,8.887869,1.491767,6.428800,0.254291,45.943182,0.716812,0.004514,5.695310,3828.381985,278.527681,0.015909,-0.012523,0.951436,negative
116.634146,14.015181,1.698283,0.099351,0.172339,157.469632,0.603524,0.001442,6.853777,14271.306861,636.876548,0.004065,-0.008044,0.928293,negative
165.158462,16.571183,2.210956,10.018086,0.155536,200.816054,0.650572,0.002301,6.299536,52585.793530,948.581806,0.007525,-0.007924,0.962937,negative
138.919067,14.932690,1.543822,3.651715,0.244317,104.028148,0.778388,0.002844,6.204004,31077.153742,834.805017,0.011852,-0.010101,0.969356,negative
118.126210,14.979439,2.201005,11.798493,0.189782,165.948820,0.636515,0.001461,6.860900,28744.773689,747.150361,0.005198,-0.008485,0.930238,negative
171.470100,18.630134,-0.124781,-0.391223,0.133920,128.712791,0.810211,0.000848,7.283418,-14868.715565,1227.661572,0.002907,-0.007095,0.955707,negative
201.451220,16.134271,0.283132,0.684886,0.131294,112.766026,0.782931,0.001269,6.856623,804.923567,926.221719,0.004274,-0.006799,0.967467,negative
129.887626,33.545762,-0.302435,-0.924452,0.124819,96.675926,0.955210,0.000662,7.479539,-96685.403018,4220.175331,0.002646,-0.015610,0.991295,negative
188.018600,25.771414,0.236223,-0.373048,0.117075,137.248243,0.897289,0.000508,7.778851,31189.669540,2535.264266,0.001850,-0.007583,0.970797,negative
203.151478,22.902566,-0.471710,-0.295424,0.131736,157.423077,0.843665,0.001133,6.935656,-45529.355118,1856.491133,0.004642,-0.008945,0.986901,negative
170.948235,12.587469,2.163238,5.962980,0.215338,102.131714,0.693134,0.003188,6.083167,22077.077140,563.511007,0.011509,-0.006125,0.936736,negative
183.772727,11.003287,1.675289,3.844471,0.186183,71.143182,0.719807,0.003458,5.924103,12861.627067,436.673218,0.011364,-0.006656,0.955822,negative
```

Figura 51. Formato txt de las características extraídas.

```
public static void main(String[] args) throws FileNotFoundException {
    Scanner lector = new Scanner(new File("D:\\caracteristicas.txt"));
```

A continuación, se configura un DataSet con las 14 características por imagen como entradas y una salida para definir si es positivo (microcalcificación benigna) o negativo (microcalcificación maligna).

```
DataSet trainingSet = new DataSet(14, 1);
```

Posteriormente, se lee el archivo de texto donde se encuentran estas características y se almacenan en el DataSet.

```
DataSet trainingSet = new DataSet(14, 1);

String n[];
int num1;

while (lector.hasNext()) {
    n = lector.next().split(",");
    if (n[14].equals("positive")) {
        num1 = 1;
        //num2 = 0;
    } else {
        num1 = 0;
        //num2 = 1;
    }

    trainingSet.addRow(new DataSetRow(new double[] {Double.parseDouble(String.valueOf(n[0])),
        Double.parseDouble(String.valueOf(n[1])), Double.parseDouble(String.valueOf(n[2])),
        Double.parseDouble(String.valueOf(n[3])), Double.parseDouble(String.valueOf(n[4])),
        Double.parseDouble(String.valueOf(n[5])), Double.parseDouble(String.valueOf(n[6])),
        Double.parseDouble(String.valueOf(n[7])), Double.parseDouble(String.valueOf(n[8])),
        Double.parseDouble(String.valueOf(n[9])), Double.parseDouble(String.valueOf(n[10])),
        Double.parseDouble(String.valueOf(n[11])), Double.parseDouble(String.valueOf(n[12])),
        Double.parseDouble(String.valueOf(n[13])),
        new double[] {num1}});
}
```

Figura 52. Código Java Aplicando Neuroph – Almacenando el DataSet.

Se define una red neuronal Perceptrón Multicapa, con 14 neuronas en la capa de entrada, 300 neuronas en la capa oculta y 1 neurona en la capa de salida. Para el entrenamiento de la red neuronal, se consideran 1415 épocas de ajuste utilizando la función sigmoideal, con un error máximo de 0.01, un momentum de 0.7 y una tasa de aprendizaje de 0.5. Los resultados del entrenamiento se guardan en un archivo .nnet, el cual se utiliza posteriormente para la ejecución de la red neuronal.

```

MultiLayerPerceptron myMLPerceptron = new MultiLayerPerceptron(TransferFunctionType.SIGMOID, 14, 300, 1);

MomentumBackpropagation backPropagation = new MomentumBackpropagation();

backPropagation.setMaxError(0.01);
backPropagation.setMomentum(0.7);
backPropagation.setLearningRate(0.5);

myMLPerceptron.learn(trainingSet, backPropagation);

testNeuralNetwork(myMLPerceptron, trainingSet);

System.out.println("Total Network Error: " + backPropagation.getTotalNetworkError());
System.out.println("Iteraciones(epocas): " + backPropagation.getMaxIterations());

myMLPerceptron.save("RM_v2_2_2.nnet");

```

Figura 53. Código Java Aplicando Neuroph – Especificando Parámetros para el entrenamiento de la RNA.

Creamos también una función testNeuralNetwork, esto nos mostrará las entradas de la red neuronal, sus respectivos pesos y salidas.

```

public static void testNeuralNetwork(NeuralNetwork nnet, DataSet testSet) {

    for (DataSetRow dataRow : testSet.getRows()) {
        nnet.setInput(dataRow.getInput());
        nnet.calculate();
        double[] networkOutput = nnet.getOutput();
        Double[] pesos = nnet.getWeights();
        System.out.print("Input: " + Arrays.toString(dataRow.getInput()) + "\n");
        System.out.print("Pesos: " + Arrays.toString(pesos) + "\n");
        System.out.println("Output: " + Arrays.toString(networkOutput));
    }
}

```

Figura 54. Código Java Aplicando Neuroph – Arrojando Resultados del Entrenamiento.

Ejecución de algoritmo de aprendizaje

Elección del conjunto inicial de Pesos

Caracterización:

Estadísticos: Mean, desviation, skewness, kurtosis

Textura: Contraste, correlation, máxima probabilidad, energía, entropía, homogeneidad, cluster shade, cluster tendency, correlation info 1, correlation info 2.

Capa de entrada

Se usó 1415 épocas, se pone solo 3 iteraciones de los pesos finales en la capa de entrada la dimensión del total global de datos.

Tabla 8. Capa de entrada – época 1

Entradas	Pesos
Mean	7.928691961378421
Desviation	3.935662828884283
Skewness	-0.24689971598700836
Kurtosis	-0.7501227659488651
Contraste	-0.5796049103320604
Correlation	178.7990815711513
Máxima Probabilidad	-0.6875580955050974
Energía	0.07168418429576004
Entropía	0.40648650476957904
Homogeneidad	-78962.05467827854
Cluster Shade	876.4316803359756
Cluster Tendency	-0.5890356626338249
Correlation Info 1	-0.5540561266002232
Correlation Info 2	-0.12299798163465352

Tabla 9. Capa de entrada – época 2

Entradas	Pesos
Mean	-0.4072601479240161

Desviation	6.111542083256386
Skewness	1.4948011116568591
Kurtosis	-0.3436075966576778
Contraste	-0.29965316447502566
Correlation	-0.697091756414161
Máxima Probabilidad	85.28817281477568
Energia	-0.25232957068166634
Entropia	-0.13520326265140578
Homogeneidad	0.018639837447568596
Cluster Shade	-19749.869875393186
Cluster Tendency	179.53990408541537
Correlation Info 1	-0.46741786634410576
Correlation Info 2	-0.497624525650733

Tabla 10. Capa de entrada – época 3

Entradas	Pesos
Mean	-0.23199577034834132
Desviation	0.06389685766458374
Skewness	4.080745984217373
Kurtosis	-0.8759359155815029
Contraste	0.24217853665933584
Correlation	-0.28703281212555026
Máxima Probabilidad	-0.6623073737262734
Energia	20.139110376756015
Entropia	0.07384027879213764
Homogeneidad	-0.39299826911850827
Cluster Shade	0.23701375301259953
Cluster Tendency	46722.549835488324
Correlation Info 1	-466.3949857313429

Correlation Info 2	-0.23353922159801568
--------------------	----------------------

Capa Oculta

Se usó 1415 épocas, se pone solo 3 iteraciones de los pesos finales en la capa de entrada la dimensión del total global de datos.

Tabla 11. Capa oculta – época 1

Entradas	Pesos
Mean	2.1707362841936595
Desviation	-1.7033958468043393
Skewness	0.041795657798761175
Kurtosis	0.07879053069382214
Contraste	0.5786717365793669
Correlation	-28.908841124552488
Máxima Probabilidad	0.21068871253606858
Energia	0.3121351863611659
Entropia	0.46428563938465084
Homogeneidad	88745.37298548406
Cluster Shade	-923.7287698957937
Cluster Tendency	-0.1406425424276331
Correlation Info 1	0.5004462941064838
Correlation Info 2	-0.6767256908823059

Tabla 12. Capa oculta – época 2

Entradas	Pesos
Mean	-0.21483747364366315
Desviation	7.5384765850197075
Skewness	2.6218285520319875
Kurtosis	0.16643396500546778

Contraste	-0.2673000831659656
Correlation	-0.6954314848501717
Máxima Probabilidad	155.7605496545189
Energia	-0.5559341008171782
Entropia	-0.3066496115839405
Homogeneidad	0.5628690268212629
Cluster Shade	-56558.13190843344
Cluster Tendency	648.8304681058352
Correlation Info 1	-0.2828632715774175
Correlation Info 2	0.2129613360649614

Tabla 13. Capa oculta – época 3

Entradas	Pesos
Mean	0.13312900499585484
Desviation	-0.19524730938323143
Skewness	1.161296067200792
Kurtosis	-5.2205610474362665
Contraste	0.3759746639555594
Correlation	-0.22417180417195295
Máxima Probabilidad	-0.2105516733211866
Energia	-140.46916809168408
Entropia	-0.5469625142686481
Homogeneidad	0.22996869936425868
Cluster Shade	-0.49712132556044225
Cluster Tendency	157206.7589336661
Correlation Info 1	-1738.739894476607
Correlation Info 2	-0.3384938141567687

Capa de Salida

Se usó 1415 épocas, se pone solo 3 iteraciones de los pesos finales en la capa de entrada la dimensión del total global de datos.

Tabla 14. Capa de salida – época 1

Entradas	Pesos
Mean	0.0020850764625241873
Desviation	-0.017179054148175162
Skewness	-0.0022869976475443188
Kurtosis	-0.01663876095055096
Contraste	-0.013592415605393038
Correlation	-0.012495985472688875
Máxima Probabilidad	-0.017418591090376398
Energia	-0.009060954232794665
Entropia	-0.018887831278943095
Homogeneidad	-0.008036503214300276
Cluster Shade	-0.04212386320051037
Cluster Tendency	-0.016320125021469116
Correlation Info 1	-0.019684284168401074
Correlation Info 2	-0.02069600151615388

Tabla 15. Capa de salida – época 2

Entradas	Pesos
Mean	-0.015985594502438312
Desviation	-0.00634456402637858
Skewness	-0.025498840310404957
Kurtosis	-0.042592180528907425
Contraste	-0.04076443765805135
Correlation	-0.022305904753445592
Máxima Probabilidad	-0.003085259321733761

Energia	-0.032332983771389816
Entropia	-0.008451936431420906
Homogeneidad	-0.041745604144043935
Cluster Shade	-0.01653512393027263
Cluster Tendency	-0.009927925976578551
Correlation Info 1	-0.02129474115220129
Correlation Info 2	-0.003005356420989192

Tabla 16. Capa de salida – época 3

Entradas	Pesos
Mean	-0.030149795280500456
Desviation	-0.022879895805607023
Skewness	-0.01540200306509349
Kurtosis	-0.02732708059629174
Contraste	-0.0025872857989252714
Correlation	-0.029271682205741487
Máxima Probabilidad	-0.009323711649757715
Energia	-0.050509922929233786
Entropia	-0.025573174593223544
Homogeneidad	-0.02768634315903874
Cluster Shade	-0.02515145790971454
Cluster Tendency	-0.01221826969531337
Correlation Info 1	-0.0251063287299746
Correlation Info 2	-0.002775274535480491

Función de Activación

Se emplea la función sigmoïdal debido a su amplia utilización y su capacidad para adecuar eficazmente los pesos de la red neuronal.

$$f(x) = \frac{1}{1 + e^{-gx}}, \text{ con } x = gin_i - \theta_i$$

Esta función opera en un intervalo continuo de valores comprendidos entre 0 y 1.

Épocas

Se realizaron 1415 épocas, las cuales representan las iteraciones ejecutadas durante el proceso de aprendizaje del modelo.

Error Máximo

Se estableció un umbral máximo de error de 0.01 para garantizar la precisión del entrenamiento del modelo.

Tasa de Aprendizaje

Se utilizó una tasa de aprendizaje de 0.5, la cual demostró ser efectiva en el proceso de ajuste de los pesos de la red neuronal.

Momentum

Se aplicó un valor de momentum de 0.7 para mejorar la convergencia del algoritmo durante el entrenamiento de la red neuronal.

3.2. Discusión

3.2.1. Sensibilidad

Este indicador denota la probabilidad de clasificar correctamente a un caso positivo o enfermo.

Tabla 17. Sensibilidad

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
---	--

$P_s = \frac{Mvp}{Mvp + Mfn}$ $P_s = \frac{16}{16 + 4}$ $P_s = 0.6666$	$P_s = \frac{Mvp}{Mvp + Mfn}$ $P_s = \frac{20}{20 + 9}$ $P_s = 0.6896$
--	--

El método propuesto obtiene un 68.96% de probabilidad de encontrar microcalcificaciones en mamografías a diferencia del otro método.

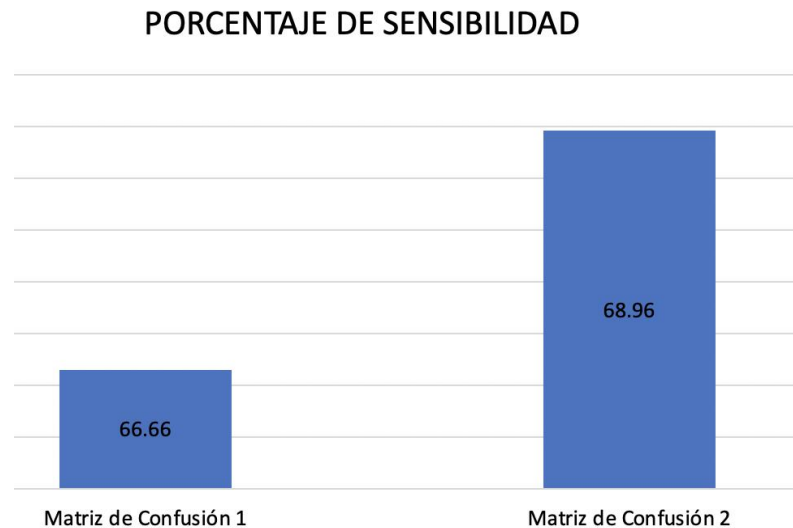


Figura 55. Porcentaje de sensibilidad.

3.2.2. Especificidad

Este indicador denota la probabilidad de clasificar correctamente a un caso negativo o sano.

Tabla 18. Especificidad

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
$P_e = \frac{Mvn}{Mvn + Mfp}$ $P_e = \frac{8}{8 + 12}$ $P_e = 0.6666$	$P_e = \frac{Mvn}{Mvn + Mfp}$ $P_e = \frac{4}{4 + 3}$ $P_e = 0.5714$

El método propuesto obtiene un 57.14% de probabilidad de encontrar microcalcificaciones en mamografías a diferencia del otro método.

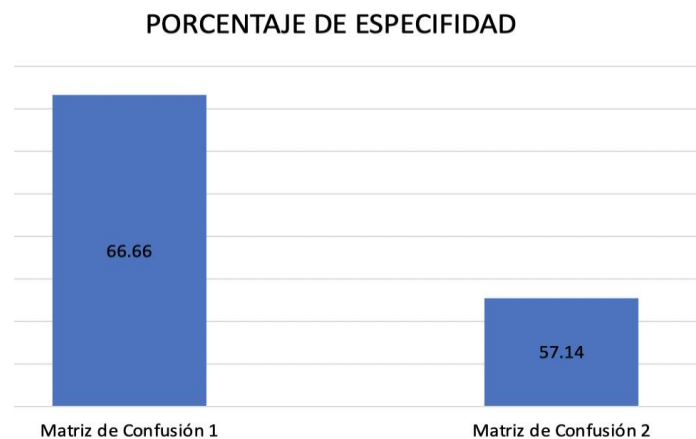


Figura 56. Porcentaje de especificidad.

3.2.3. Valor predictivo positivo

Este indicador denota la probabilidad de ser realmente positivo la muestra que resultó positivo después de aplicado el método.

Tabla 19. Valor predictivo positivo

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
$V_{pp} = \frac{M_{vp}}{M_{vp} + M_{fp}}$ $V_{pp} = \frac{16}{16 + 4}$ $V_{pp} = 0.8$	$V_{pp} = \frac{M_{vp}}{M_{vp} + M_{fp}}$ $V_{pp} = \frac{20}{20 + 3}$ $V_{pp} = 0.8695$

El método propuesto obtiene un 86.95% de probabilidad de encontrar microcalcificaciones en mamografías a diferencia del otro método.

VALOR PREDICTIVO POSITIVO

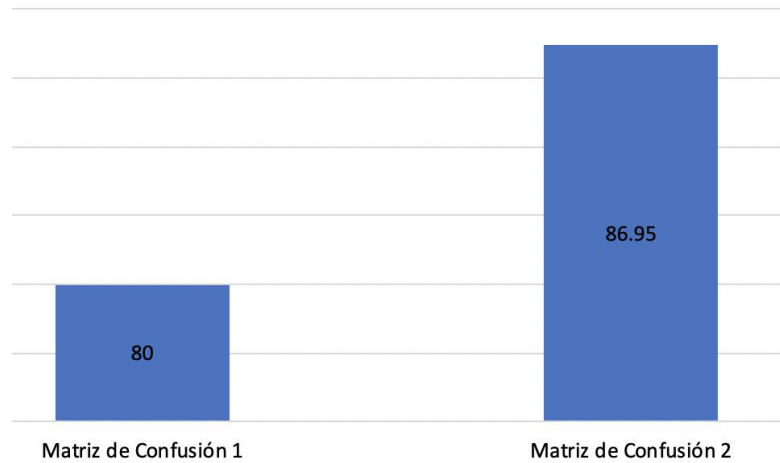


Figura 57. Valor predictivo positivo.

3.2.4. Valor predictivo negativo

Este indicador denota la probabilidad de ser realmente negativo la muestra que resultó negativa después de aplicado el método.

Tabla 20. Valor predictivo negativo

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
$V_{pn} = \frac{M_{vn}}{M_{fn} + M_{vn}}$ $V_{pn} = \frac{8}{8 + 8}$ $V_{pn} = 0.5$	$V_{pn} = \frac{M_{vn}}{M_{fn} + M_{vn}}$ $V_{pn} = \frac{4}{9 + 4}$ $V_{pn} = 0.3076$

El método propuesto obtiene un 30.76% de probabilidad de encontrar microcalcificaciones en mamografías a diferencia del otro método.

VALOR PREDICTIVO NEGATIVO

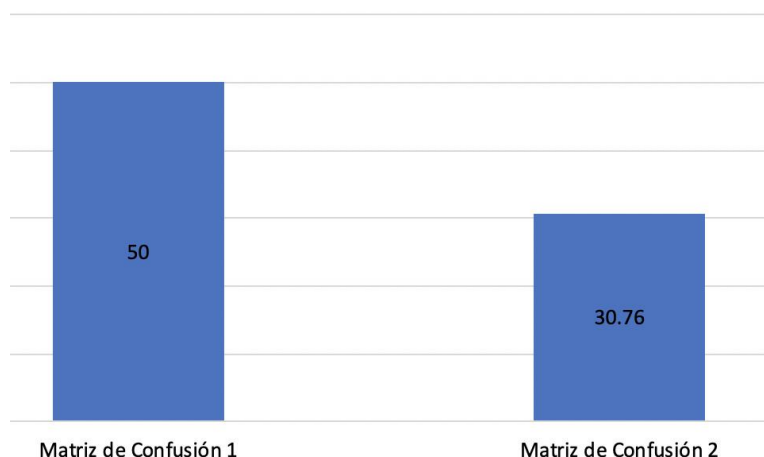


Figura 58. Valor predictivo negativo.

3.2.5. Razón de verosimilitudes positivas

Este indicador denota la probabilidad de que sean totalmente ciertos los positivos en cuanto a los resultados de las muestras de verdaderos positivos.

Tabla 21. Razón de verosimilitudes positivas

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
$Rvp = \frac{Sensibilidad}{1 - Especificidad}$ $Rvp = \frac{0.6666}{1 - 0.6666}$ $Rvp = 0.3332$	$Rvp = \frac{Sensibilidad}{1 - Especificidad}$ $Rvp = \frac{0.6896}{1 - 0.5714}$ $Rvp = 0.6089$

El método propuesto obtiene un 60.89% de probabilidad de encontrar microcalcificaciones en mamografías a diferencia del otro método.

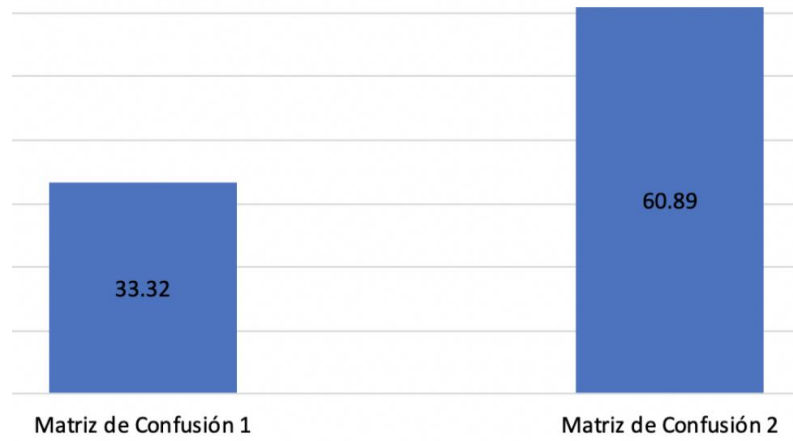


Figura 59. Razón de verosimilitudes positivas.

3.2.6. Razón de verosimilitudes negativas

Este indicador denota la probabilidad de que sean totalmente ciertos los falsos positivos en cuanto a los resultados de las muestras de positivos.

Tabla 22. Razón de verosimilitudes negativas

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
$Rvp = \frac{Sensibilidad}{1 - Especificidad}$ $Rvp = \frac{0.6666}{1 - 0.6666}$ $Rvp = 0.3332$	$Rvp = \frac{Sensibilidad}{1 - Especificidad}$ $Rvp = \frac{0.6896}{1 - 0.5714}$ $Rvp = 0.5414$

En este caso el método propuesto obtiene un 54,14 %.

RAZON DE VEROSIMILITUDES NEGATIVAS

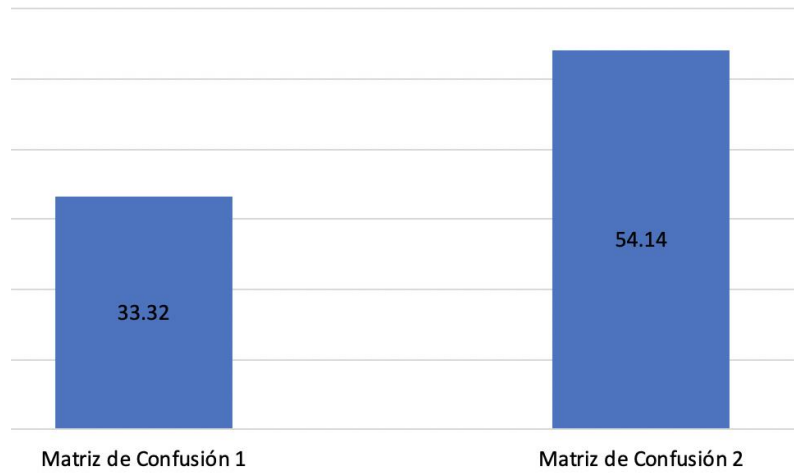


Figura 60. Razón de verosimilitudes negativas.

3.2.7. Precisión

Este indicador denota la exactitud del método.

Tabla 23. Precisión

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
$P = \frac{Mvp + M}{p + n}$ $P = \frac{16 + 8}{24 + 12}$ $P = 0.6666$	$P = \frac{Mvp + M}{p + n}$ $P = \frac{20 + 9}{24 + 12}$ $P = 0.8055$

El método propuesto obtiene un 80.55% a diferencia de los otros métodos, el cual demuestra un proceso exitoso porque está dándole un grado de confiabilidad superior al 80%.

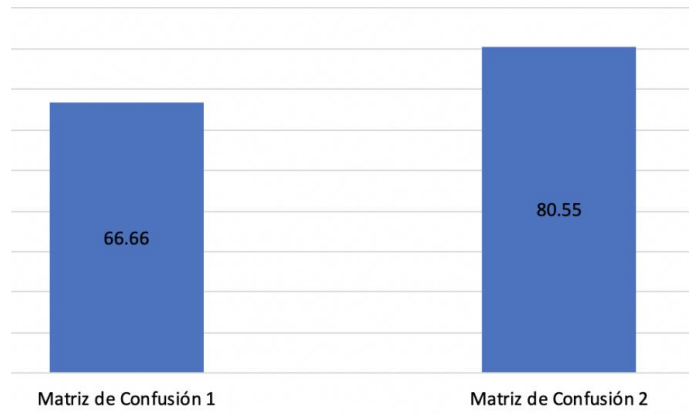


Figura 61. Precisión.

3.2.8. Ratio o razón de falsos descubrimientos

Este indicador denota la razón de encontrar negativos errados.

Tabla 24. Ratio o razón de falsos descubrimientos

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
$Rfd = \frac{Mfp}{Mfp + Mvp}$ $Rfd = 0.2$	$Rfd = \frac{Mfp}{Mfp + Mvp}$ $Rfd = 0.1304$

El método obtiene un 13.04% de encontrar negativos errados.

RATIO O RAZON DE FALSOS DESCUBRIMIENTOS

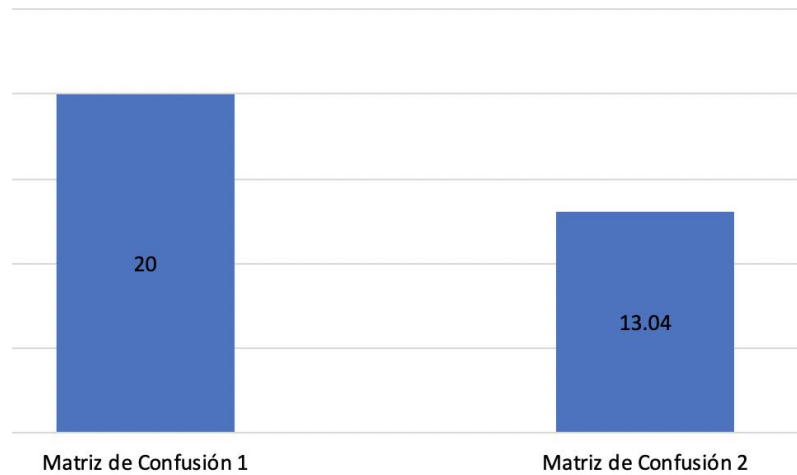


Figura 62. Ratio o razón de falsos descubrimientos.

3.2.9. Ratio o razón de falsos positivos

Este indicador denota la razón de encontrar falsos positivos.

Tabla 25. Ratio o razón de falsos positivos

Matriz de confusión 1: Segmentación piramidal con técnicas de preprocesamiento	Matriz de confusión 2: Segmentación thresholding con técnicas de preprocesamiento
$Rfp = \frac{Mfp}{Mfp + Mvn}$ $Rfp = \frac{4}{4 + 8}$ $Rfp = 0.3333$	$Rfp = \frac{Mfp}{Mfp + Mvn}$ $Rfp = \frac{3}{3 + 4}$ $Rfp = 0.4285$

El método propuesto obtiene un 42.85% de encontrar falsos positivos en casos negativos a diferencia de los otros métodos.

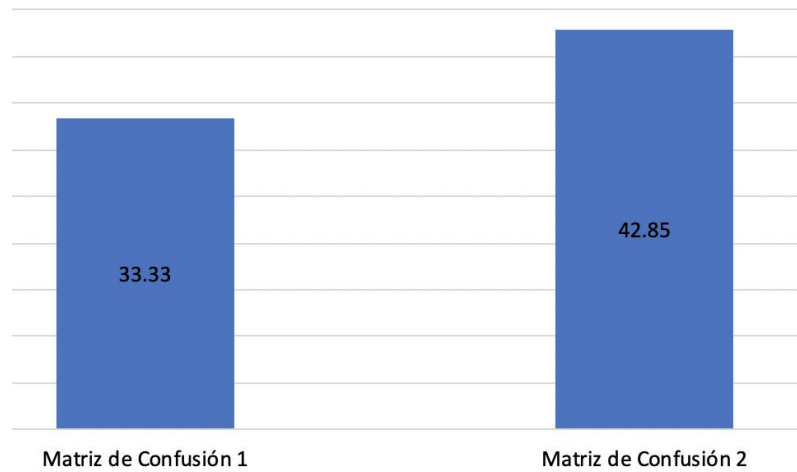


Figura 63. Ratio o razón de falsos positivos.

IV. CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones:

- Se logró cumplir satisfactoriamente el objetivo principal con una precisión superior al 80% en la detección de microcalcificaciones asociadas a tumores en la glándula mamaria.
- Las mamografías fueron caracterizadas de manera adecuada mediante el uso de una lista de observación estructurada y aprovechando la precisión de la base de datos "The mini-MIAS database of mammograms".
- Se aplicaron técnicas de preprocesamiento como morfología externa y escalado de grises para mejorar las características de las imágenes mamográficas.
- Se implementaron algoritmos de segmentación, específicamente thresholding, para extraer la región de interés de las mamografías preprocesadas.
- Se emplearon algoritmos estadísticos (Media, Desviación, Asimetría, Curtosis) y de textura (GLCM) para la extracción automatizada de características de las imágenes médicas.
- La clasificación se llevó a cabo utilizando una red neuronal Perceptrón Multicapa entrenada con Neuroph y el algoritmo de Backpropagation, haciendo uso de la función sigmoideal. Los

resultados preliminares, basados en 36 imágenes con 24 positivas y 12 negativas, mostraron los siguientes valores: Verdaderos Positivos (Mvp): 20, Verdaderos Negativos (Mvn): 4, Falsos Positivos (Mfp): 3, Falsos Negativos (Mfn): 9.

4.2. Recomendaciones:

- Enfocarse exclusivamente y de manera individual en cada descriptor o patrón durante el trabajo, ya que el desarrollo puede volverse tedioso y complicado, resultando en posibles errores al extraer características.
- Evitar el uso de técnicas de preprocesamiento como la reducción de ruido y el realzado de contraste en imágenes mamográficas normalizadas, dado que pueden afectar su integridad. Si se dispone de imágenes normalizadas, es recomendable abstenerse de aplicar y evaluar estas técnicas.
- Realizar comparaciones de los resultados obtenidos con bibliotecas de visión artificial como OpenCV para la segmentación de imágenes.
- Explorar la clasificación utilizando otras técnicas de inteligencia artificial como Máquinas de Vectores de Soporte (SVM) y Redes Bayesianas, además de la red neuronal utilizada, para evaluar su rendimiento y eficacia en el contexto de detección de microcalcificaciones en mamografías.

REFERENCIAS

- [1] V. Debien *et al.*, “Real-world clinical outcomes of patients with stage I HER2-positive breast cancer treated with adjuvant paclitaxel and trastuzumab,” *Critical Reviews in Oncology/Hematology*, vol. 190, p. 104089, 2023, doi: <https://doi.org/10.1016/j.critrevonc.2023.104089>.
- [2] F. Cardoso *et al.*, “Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†,” *Annals of Oncology*, vol. 30, no. 8, pp. 1194–1220, 2019, doi: <https://doi.org/10.1093/annonc/mdz173>.
- [3] C. Kaushal, S. Bhat, D. Koundal, and A. Singla, “Recent Trends in Computer Assisted Diagnosis (CAD) System for Breast Cancer Diagnosis Using Histopathological Images,” *IRBM*, vol. 40, no. 4, pp. 211–227, 2019, doi: <https://doi.org/10.1016/j.irbm.2019.06.001>.
- [4] F. J. Escudero, “Historical evolution of breast reconstruction,” *Anales del Sistema Sanitario de Navarra*, vol. 28, pp. 7–18, 2005.
- [5] Quinters, *Detección automática de lesiones en mamografías usando técnica de análisis de imágenes*. Ambato, Ecuador, 2015.
- [6] Shan Alam and Gara, “Computer-Aided Diagnosis for Breast Ultrasound Using Computerized BI-RADS Features and Machine Learning Methods,” 2015.
- [7] Sudarshan, *Application of wavelet techniques for cancer diagnosis using ultrasound images*. Singapore, 2015.
- [8] R. Zemouri *et al.*, “Constructive Deep Neural Network for Breast Cancer Diagnosis,” *IFAC-PapersOnLine*, vol. 51, no. 27, pp. 98–103, 2018, doi: <https://doi.org/10.1016/j.ifacol.2018.11.660>.
- [9] S. N. Kathija and D. M. M. Sathik, “Breast Cancer Data Classification Using Neural Network Approach of MLP Algorithm,” 2017.
- [10] P. J. G. Laencina, *Mejora en la detección de microcalcificaciones en mamografías digitalizadas mediante la aplicación de arquitecturas neuronales*. Cartagena, España, 2004.
- [11] Gomes, *Sistema de auxílio na avaliação de calcificações mamárias por processamento digital de imagens e inteligência artificial*. Brasil, 2012.
- [12] García González, Gallardo and Macías, “Consistent performance measurement of a system to detect masses in mammograms based on blind feature extraction,” 2013.
- [13] H. Wang, B. Zheng, S. W. Yoon, and H. S. Ko, “A support vector machine-based ensemble algorithm for breast cancer diagnosis,” *European Journal of Operational Research*, vol. 267, no. 2, pp. 687–699, 2018, doi: <https://doi.org/10.1016/j.ejor.2017.12.001>.
- [14] F. Gao *et al.*, “SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis,” *Computerized Medical Imaging and Graphics*, vol. 70, pp. 53–62, 2018, doi: <https://doi.org/10.1016/j.compmedimag.2018.09.004>.
- [15] Y. Feng *et al.*, “Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis,” *Genes and Diseases*, vol. 5, no. 2, pp. 77–106, 2018. doi: [10.1016/j.gendis.2018.05.001](https://doi.org/10.1016/j.gendis.2018.05.001).
- [16] M. A. Arrechea Irigoyen, F. Vicente García, A. Córdoba Iturriagagoitia, B. Ibañez Beroiz, M. Santamaría Martínez, and F. Guillén Grima, “Molecular subtypes of breast cancer: prognostic implications and clinical and immunohistochemical characteristics,” *Anales del Sistema Sanitario de Navarra*, vol. 34, pp. 219–233, Aug. 2011.
- [17] P. K. Shahi and G. P. Manga, “HER2-negative metastatic breast cancer: Analysis of the efficacy of chemotherapy in different groups,” 2007.
- [18] E. A. Sickles, C. J. D’Orsi, L. W. Bassett, and others, “ACR BI-RADS Mammography,” in *ACR BI-RADS Atlas, Breast Imaging Reporting and Data System*, 5th ed., Reston, VA: American College of Radiology, 2013, pp. 134–136.
- [19] R. Medina and J. Bellera, *Bases de Procesamiento de Imágenes Médicas*. Venezuela: Universidad de Los Andes, 2001.
- [20] M. Jourlin, J. Breugnot, B. Abdallah, J. Corvo, E. Couka, and M. Carré, “Chapter One - Image Segmentation in the Field of the Logarithmic Image Processing Model: Special Focus on the Hierarchical Ascendant Classification Techniques,” in *Advances in Imaging and Electron Physics*,

vol. 177, P. W. Hawkes, Ed., Elsevier, 2013, pp. 1–44. doi: 10.1016/B978-0-12-407702-7.00001-2.

- [21] L. Ding and A. Goshtasby, “On the Canny edge detector,” *Pattern Recognition*, vol. 34, no. 3, pp. 721–725, 2001, doi: [https://doi.org/10.1016/S0031-3203\(00\)00023-6](https://doi.org/10.1016/S0031-3203(00)00023-6).
- [22] Teknomo, “Analysis of Simple K-Means with Multiple Dimensions using WEKA,” 2015.
- [23] A. Belič, I. Škrjanc, D. Zupančič-Božič, and F. Vrečer, “9 - Optimisation of compression parameters with AI-based mathematical models,” in *Formulation Tools for Pharmaceutical Development*, J. E. Aguilar, Ed., in Woodhead Publishing Series in Biomedicine. , Woodhead Publishing, 2013, pp. 229–261. doi: <https://doi.org/10.1533/9781908818508.229>.
- [24] Ramos, “Discovering Mammography-based machine Learning Classifiers for Breast Cancer Diagnosis,” 2011.

ANEXOS:




ANEXO 01: DECLARACIÓN JURADA DE ORIGINALIDAD

Quien suscribe la DECLARACIÓN JURADA, soy Wildor Alexander Flores Vasquez. Del Programa de Estudios de Ingeniería de sistemas de la Universidad Señor de Sipán S.A.C, declaro bajo juramento que soy autor del trabajo titulado:

RECONOCIMIENTO DE MICROCALCIFICACIONES EN MAMOGRAFIAS HACIENDO USO DE UNA MAQUINA DE APRENDIZAJE

El texto de mi trabajo de investigación responde y respeta lo indicado en el Código de Ética del Comité Institucional de Ética en Investigación de la Universidad Señor de Sipán, conforme a los principios y lineamientos detallados en dicho documento, en relación con las citas y referencias bibliográficas, respetando el derecho de propiedad intelectual, por lo cual informo que la investigación cumple con ser inédito, original y autentico.

En virtud de lo antes mencionado, firman:

Wildor Alexander Flores Vasquez	DNI: 75342738	
---------------------------------	---------------	---

Pimentel, 18 de julio de 2024.

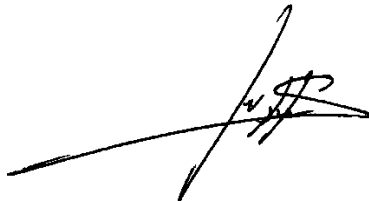
ANEXO 02: ACTA DE REVISIÓN DE SIMILITUD DE LA INVESTIGACIÓN

Yo **Heber Ivan Mejia Cabrera** docente del curso de **Investigación II** del Programa de Estudios de **Ingeniería de Sistemas** y revisor de la investigación del estudiante, **Flores Vasquez Wildor Alexander**, titulada:

RECONOCIMIENTO DE MICROCALCIFICACIONES EN MAMOGRAFIAS HACIENDO USO DE UNA MAQUINA DE APRENDIZAJE

Se deja constancia que la investigación antes indicada tiene un índice de similitud del **22%**, verificable en el reporte final del análisis de originalidad mediante el software de similitud TURNITIN. Por lo que se concluye que cada una de las coincidencias detectadas no constituyen plagio y cumple con lo establecido en la Directiva sobre índice de similitud de los productos académicos y de investigación en la Universidad Señor de Sipán S.A.C., aprobada mediante Resolución de Directorio N° 0758-2022/FIAU-USS.

Pimentel, 18 de Julio del 2024



Mg. Heber Ivan Mejia Cabrera

DNI N° 41639565


ANEXO 03: Acta de aprobación de asesor:



ACTA DE APROBACIÓN DEL ASESOR

Yo **Jorge Luis Gutierrez Gutierrez**, quien suscribe como asesor designado mediante Resolución de Facultad N° **Número de resolución**, del proyecto de investigación titulado **Reconocimiento de microcalcificaciones en mamografías haciendo uso de una máquina de aprendizaje**, desarrollado por el(los) estudiante(s): **Flores Vasquez Wildor Alexander**, del programa de estudios de **Ingeniería de Sistemas**, acredito haber revisado, y declaro expedito para que continúe con el trámite pertinentes.

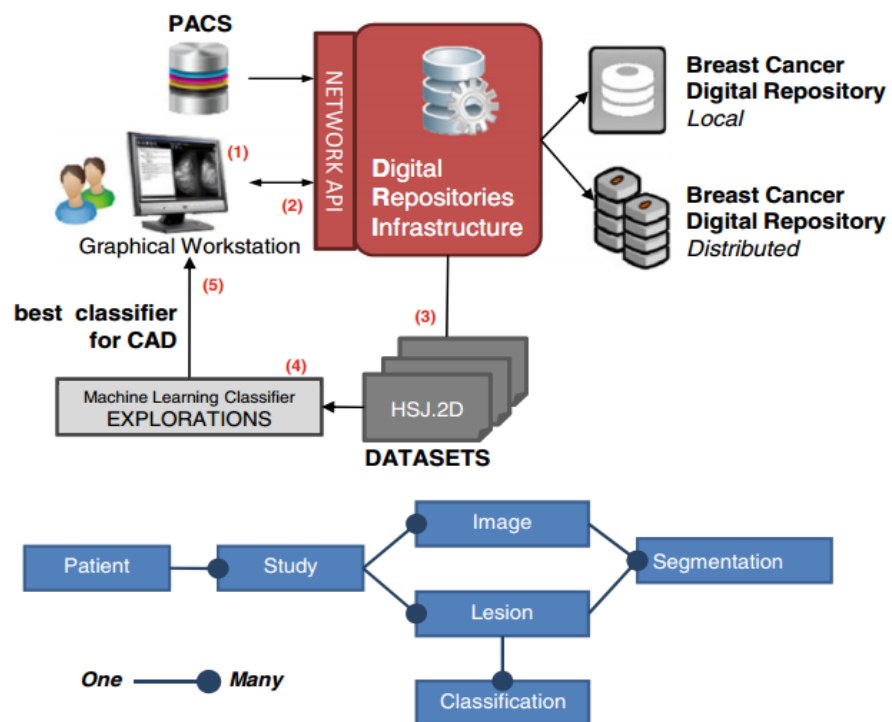
En virtud de lo antes mencionado, firman:

Gutierrez Gutierrez Jorge Luis	29429722	
--------------------------------	----------	--

Pimentel, 18 de Julio del 2024

Anexo 04:

1. La mamografía las imágenes son preprocesadas a través una estación de trabajo gráfica
2. Radiólogos especializados extraen imágenes
3. Se extrae datos de las anotaciones
4. MLC configuraciones son masivamente buscado y seleccionados
5. Se integran de nuevo en la estación de trabajo proporcionando segunda opinión automatizado el diagnóstico de los médicos



Fuente: Discovering Mammography-based Machine Learning Classifiers for Breast Cancer

Diagnosis [24]

NOMBRE DEL TRABAJO

**Informe Tesis Flores-Vasquez -
turnitin.docx**RECUENTO DE
PALABRAS**11149 Words**

RECUENTO DE CARACTERES

66728 CharactersRECUENTO DE
PÁGINAS**99 Pages**

TAMAÑO DEL ARCHIVO

24.7MB

FECHA DE ENTREGA

Sep 3, 2024 10:50 AM GMT-5

FECHA DEL INFORME

Sep 3, 2024 10:51 AM GMT-5**● 22% de similitud general**

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos.

- 21% Base de datos de Internet
- Base de datos de Crossref
- 11% Base de datos de trabajos entregados
- 5% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

● Excluir del Reporte de Similitud

- Material bibliográfico
- Coincidencia baja (menos de 8 palabras)
- Material citado

Código Fuente

PanelDibujo.java

En esta sección lo que hacemos es dibujar una ROI sobre la imagen mamográfica

```
package microbreastcancer;

import java.awt.*;
import java.awt.RenderingHints;
import java.awt.event.MouseEvent;
import java.awt.event.MouseListener;
import java.awt.event.MouseMotionListener;
import java.awt.geom.Rectangle2D;
import java.awt.image.BufferedImage;
import java.io.File;
import java.io.IOException;
import javax.imageio.ImageIO;
import javax.swing.JOptionPane;
import javax.swing.JPanel;

public class PanelDibujo extends JPanel implements MouseMotionListener, MouseListener {

    Image img;
    BufferedImage Imagememoria;
    BufferedImage imgrecortada;

    Graphics2D g2D;

    float x = 0;
    float y = 0;
    float ancho = 0;
    float alto = 0;

    public PanelDibujo(BufferedImage f) {
        this.img = f;
        this.setSize(f.getWidth(), f.getHeight());
        this.setVisible(true);
        addMouseMotionListener(this);
        addMouseListener(this);
    }

    @Override
    protected void paintComponent(Graphics g) {
        Graphics2D g2 = (Graphics2D) g;
        if (img != null) {
            Imagememoria = new BufferedImage(this.getWidth(), this.getHeight(), BufferedImage.TYPE_INT_RGB);
            g2D = Imagememoria.createGraphics();

            g2D.setRenderingHint(RenderingHints.KEY_ANTIALIASING, RenderingHints.VALUE_ANTIALIAS_ON);
            g2D.drawImage(img, 0, 0, img.getWidth(this), img.getHeight(this), this);
            g2D.setStroke(new BasicStroke(2f));
            g2D.setColor(Color.WHITE);
            Rectangle2D r2 = new Rectangle2D.Float(x, y, ancho, alto);
            g2D.draw(r2);
            g2D.drawImage(Imagememoria, 0, 0, this);
        }
    }

    public void guardar_imagen(String nombearchivo, String seleccion_op) {
        imgrecortada = ((BufferedImage) img).getSubimage((int) x, (int) y, (int) ancho, (int) alto);
        try {
            String ruta;
            if (seleccion_op.equals("Si")) {
                ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Cancer\\MicroBreastCancer\\imagenes_micro\\positive\\";
            } else {
                ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Cancer\\MicroBreastCancer\\imagenes_micro\\negative\\";
            }
            ImageIO.write(imgrecortada, "jpg", new File(ruta + nombearchivo));
            JOptionPane.showMessageDialog(null, "Se ha guardado Correctamente la Region de Interes");
        } catch (IOException e) {
            JOptionPane.showMessageDialog(null, "Error, Trate nuevamente, " + e.getMessage());
        }
    }

    public void mouseDragged(MouseEvent e) {
        ancho = e.getX() - x;
        alto = e.getY() - y;

        if (ancho < 0) {
            ancho = 0;
        }
        if (alto < 0) {
            alto = 0;
        }
        if (x > this.getWidth()) {
            x = this.getWidth() - ancho;
        }
        if (y > this.getHeight()) {
            y = this.getHeight() - alto;
        }
    }
}
```

```

84     }
85
86     this.repaint();
87 }
88
89 public void mouseMoved(MouseEvent e) {
90 }
91
92 public void mouseClicked(MouseEvent e) {
93 }
94
95 }
96
97 public void mousePressed(MouseEvent e) {
98     x = e.getX();
99     y = e.getY();
100 }
101
102 public void mouseReleased(MouseEvent e) {
103 }
104
105 }
106
107 public void mouseEntered(MouseEvent e) {
108 }
109
110 public void mouseExited(MouseEvent e) {
111 }
112
113 }
114 }
115

```

jfMicrocalcificacion.java

Esta interfaz nos servirá para ver la imagen mamográfica, segmentarla

```

package microbreastcancer;

import java.awt.Image;
import java.io.File;
import java.io.IOException;
import java.util.ArrayList;
import java.util.HashMap;
import java.util.List;
import java.util.Map;
import java.util.Random;
import java.util.Vector;
import javax.imageio.ImageIO;
import javax.swing.Icon;
import javax.swing.ImageIcon;
import javax.swing.JComboBox;
import javax.swing.JFileChooser;
import javax.swing.JOptionPane;
import org.neuroph.core.NeuralNetwork;
import org.neuroph.core.exceptions.VectorSizeMismatchException;
import org.neuroph.imgregc.ImageRecognitionPlugin;
import org.opencv.core.Core;
import org.opencv.core.CvType;
import org.opencv.core.Mat;
import org.opencv.core.MatOfFloat;
import org.opencv.core.MatOfInt;
import org.opencv.core.MatOfPoint;
import org.opencv.core.MatOfPoint2f;
import org.opencv.core.Point;
import org.opencv.core.Rect;
import org.opencv.core.Scalar;
import org.opencv.core.Size;
import org.opencv.highgui.Highgui;
import org.opencv.imgproc.*;

/*
 * To change this license header, choose License Headers in Project Properties.
 * To change this template file, choose Tools | Templates
 * and open the template in the editor.
 */
/**
 *
 * @author ALEX
 */
public class jfMicrocalcificacion extends javax.swing.JFrame {

    PanelDibujo pd;
    PanelDibujo1 pd1;
    String nombreArchivo = "";
    Object seleccion_op;
    double suma;

    public jfMicrocalcificacion() {
        initComponents();
        setLocationRelativeTo(null);
        System.loadLibrary(Core.NATIVE_LIBRARY_NAME);
    }

    /**
     * This method is called from within the constructor to initialize the form.
     * WARNING: Do NOT modify this code. The content of this method is always
     * regenerated by the Form Editor.
     */
    @SuppressWarnings("unchecked")
    Generated Code

    private void jMenuItem2ActionPerformed(java.awt.event.ActionEvent evt) {

        JFileChooser selector = new JFileChooser();
        int resultado = selector.showOpenDialog(null);
        if (resultado == JFileChooser.APPROVE_OPTION) {
            panelPrincipal.removeAll();
            panelPrincipal.repaint();
            try {
                nombreArchivo = selector.getSelectedFile().getName();
                pd = new PanelDibujo(ImageIO.read(selector.getSelectedFile()));
                pd.repaint();
            }
        }
    }
}

```

```

        panelPrincipal.add(pd);
    } catch (IOException e) {
        // TODO Auto-generated catch block
        e.printStackTrace();
    }
}

private void jMenuItem3ActionPerformed(java.awt.event.ActionEvent evt) {
    /*
    String[] op = {"Si", "No"};
    final JComboBox<String> combo = new JComboBox<>(op);
    String[] options = {"Aceptar"};

    String title = "?Sospecha de que es una microcalcificacion?";
    int seleccion = JOptionPane.showOptionDialog(null, combo, title, JOptionPane.DEFAULT_OPTION, JOptionPane.PLAIN_MESSAGE, null,
    if (seleccion > 0) {
        System.out.println("seleccion is: " + options[seleccion]);
    }

    seleccion_op = combo.getSelectedItem();
    */
    //if (seleccion_op != null) {
    // System.out.println(seleccion_op);

    pd.guardar_imagen(nombreArchivo, "");

    //)
    String ruta;
    ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\p_roi\\";
    /*
    if (seleccion_op.equals("Si")) {
        ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\positive\\";
    } else {
        ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\negative\\";
    }
    */

    System.out.println(ruta + nombreArchivo);
    ImageIcon icon = new ImageIcon(ruta + nombreArchivo);
    Icon icono = new ImageIcon(icon.getImage().getScaledInstance(lblRoi.getWidth(), lblRoi.getHeight(), Image.SCALE_DEFAULT));
    lblRoi.setText("");
    lblRoi.setIcon(icono);

    /*
    lblRoi.removeAll();
    lblRoi.repaint();
    try {
        pd1 = new PanelDibujo1(ImageIO.read(new File(ruta + nombreArchivo)));
        pd1.repaint();
        pd1.guardar_imagen(nombreArchivo, "");
        repaint();
        lblRoi.add(pd1);
    } catch (IOException e) {
        // TODO Auto-generated catch block
        e.printStackTrace();
    }
    */
}

private void jMenuItem4ActionPerformed(java.awt.event.ActionEvent evt) {
    HistROI();
    HistSeg();
}

private void PreProcSegActionPerformed(java.awt.event.ActionEvent evt) {
    String ruta_roi, ruta_seg;
    ruta_roi = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\p_roi\\";
    ruta_seg = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\p_tresh\\";
    /*
    if (seleccion_op.toString().equals("Si")) {
        ruta_roi = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\positive\\";
        ruta_seg = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\threshold_positiv";
    } else {
        ruta_roi = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\negative\\";
    }
    */
}

```

```

}
}
*/
Mat src = Highgui.imread(ruta_roi + nombreArchivo);
Mat src_1 = new Mat();
Mat gray = new Mat();
Mat morph = new Mat();
Mat dest = new Mat();

String ruta_carp_seg = ruta_seg + nombreArchivo;
File file_carp_seg = new File(ruta_carp_seg);
if (file_carp_seg.exists()) {
    file_carp_seg.delete();
    file_carp_seg.mkdirs();
} else {
    file_carp_seg.mkdirs();
}
}
Mat kernel = Mat.ones(8, 8, CvType.CV_8U);
Imgproc.morphologyEx(src, morph, Imgproc.MORPH_TOPHAT, kernel);

Imgproc.cvtColor(morph, gray, Imgproc.COLOR_RGB2GRAY); // Convertimos la imagen a escala de grises

Imgproc.threshold(gray, dest, 50, 255, 3); // Segmentamos la imagen con la tecnica segmentacion por umbral(otsu)

Highgui.imwrite(ruta_carp_seg + "/thresh_" + nombreArchivo, dest);

Size kernelSize = new Size(8, 100);

Mat kernel_t = Imgproc.getStructuringElement(Imgproc.MORPH_RECT, kernelSize);

Imgproc.dilate(dest, dest, kernel, new Point(-1, -1), 3);

String ruta_rec_seg = ruta_seg + nombreArchivo + "/recortes";
File file_rec_seg = new File(ruta_rec_seg);
file_rec_seg.mkdirs();

List<MatOfPoint> contourList = new ArrayList<>();
Mat hierarchy = new Mat();

Imgproc.findContours(dest, contourList, hierarchy, Imgproc.RETR_CCOMP, Imgproc.CHAIN_APPROX_SIMPLE);

```

```

//Drawing contours on a new image
//Mat contours = new Mat();
//contours.create(src.rows(), src.cols(), CvType.CV_8UC3);
Random r = new Random();
Rect rectCrop = null;
System.out.println("Contornos: " + contourList.size());
//MatOfPoint temp_contour = contourList.get(0);

if (contourList.size() > 0) {
    for (int i = 0; i < contourList.size(); i++) {
        /*
        MatOfPoint2f new_mat = new MatOfPoint2f( temp_contour.toArray() );
        int contourSize = (int)temp_contour.total();
        MatOfPoint2f approxCurve_temp = new MatOfPoint2f();
        Imgproc.approxPolyDP(new_mat, approxCurve_temp, contourSize*0.05, true);
        MatOfPoint points = new MatOfPoint( approxCurve_temp.toArray() );

        Rect rect = Imgproc.boundingRect(points);
        //Core.rectangle(src, new Point(rect.x,rect.y), new Point(rect.x+rect.width,rect.y+rect.height), new Scalar(170,0,150,0), 5
        //Core.rectangle(src, rect.tl(), rect.br(), Scalar.all(255), 2, 8, 0);
        */
        Rect rect = Imgproc.boundingRect(contourList.get(i));

        rectCrop = new Rect(rect.x, rect.y, rect.width, rect.height);
        // Imgproc.drawContours(src, contourList, i, new Scalar(r.nextInt(255), r.nextInt(255), r.nextInt(255)), 1);

        Mat image_roi = new Mat(src, rectCrop);
        Highgui.imwrite(ruta_rec_seg + "\\\" + "micro_" + (i + 1) + ".jpg", image_roi);

        //Core.rectangle(src, new Point(rect.x, rect.y), new Point(rect.x + rect.width, rect.y + rect.height), new Scalar(0, 255, 255), 2);

        //Imgproc.drawContours(src, contourList, i, new Scalar(r.nextInt(255), r.nextInt(255), r.nextInt(255)), 1);

    }
}
for (int i = 0; i < contourList.size(); i++) {
    double area = Imgproc.contourArea(contourList.get(i));
}

```

```

        if (area > 1 && area < 10000) {
            Rect rect = Imgproc.boundingRect(contourList.get(i));

            //Imgproc.drawContours(src, contourList, i, new Scalar(r.nextInt(255), r.nextInt(255), r.nextInt(255)), 1);
            Core.rectangle(src, new Point(rect.x, rect.y), new Point(rect.x + rect.width, rect.y + rect.height), new Scalar(0,
        )
    }
    Highgui.imwrite(ruta_carp_seg + "\\seg_" + nombreArchivo, src);
} else {
    //Highgui.imwrite(ruta_carp_seg + "\\seg_" + nombreArchivo, src);
}

/* ----- */
//Highgui.imwrite(ruta_carp_seg + nombreArchivo, src);
JOptionPane.showMessageDialog(this, "Segmentacion Guardada");

// Thresh
ImageIcon icon = new ImageIcon(ruta_carp_seg + "\\thresh_" + nombreArchivo);
Icon icono = new ImageIcon(icon.getImage().getScaledInstance(lblRoi.getWidth(), lblRoi.getHeight(), Image.SCALE_DEFAULT));
lblSegmentada.setText("");
lblSegmentada.setIcon(icono);

// Segmentacion con Rect
ImageIcon icon_rect = new ImageIcon(ruta_carp_seg + "\\seg_" + nombreArchivo);
Icon icono_rect = new ImageIcon(icon_rect.getImage().getScaledInstance(lblRoi.getWidth(), lblRoi.getHeight(), Image.SCALE_DEFA
lblSegmentada_rect.setText("");
lblSegmentada_rect.setIcon(icono_rect);

txtImagen.setText(ruta_carp_seg + "\\ " + "recortes");
System.out.println("Ruta Carpeta Imagen: " + ruta_carp_seg + "\\ " + "recortes");
}

/**
 * @param args the command line arguments
 */
public static void main(String args[]) {

    /* Set the Nimbus look and feel */
    LookAndFeel lookAndFeel = new NimbusLookAndFeel();
    lookAndFeel.install();

    /* Create and display the form */
    java.awt.EventQueue.invokeLater(new Runnable() {
        public void run() {
            new jfMicrocalcificacion().setVisible(true);
        }
    });
}

public void HistROI() {
    String ruta;
    ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\p_roi\\";
    /*
    if (seleccion_op.equals("Si")) {
        ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\positive\\";
    } else {
        ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\negative\\";
    }
    */
    Mat image = Highgui.imread(ruta + nombreArchivo);

    Mat src = new Mat(image.height(), image.width(), CvType.CV_8UC2);

    Imgproc.cvtColor(image, src, Imgproc.COLOR_RGB2GRAY);

    Vector<Mat> bgr_planes = new Vector<>();
    Core.split(src, bgr_planes);

    MatOfInt histSize = new MatOfInt(256);

    final MatOfFloat histRange = new MatOfFloat(0f, 256f);

    boolean accumulate = false;

    Mat b_hist = new Mat();

```



```

    Imgproc.calcHist(bgr_planes, new MatOfInt(0), new Mat(), b_hist, histSize, histRange, accumulate);

    int hist_w = 512;
    int hist_h = 600;
    long bin_w;
    bin_w = Math.round((double) (hist_w / 256));

    Mat histImage = new Mat(hist_h, hist_w, CvType.CV_8UC1);

    Core.normalize(b_hist, b_hist, 3, histImage.rows(), Core.NORM_MINMAX);

    for (int i = 1; i < 256; i++) {
        Core.line(histImage, new Point(bin_w * (i - 1), hist_h - Math.round(b_hist.get(i - 1, 0)[0])),
            new Point(bin_w * (i), hist_h - Math.round(Math.round(b_hist.get(i, 0)[0]))),
            new Scalar(255, 0, 0), 2, 8, 0);
    }

    Highgui.imwrite("C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\hist\\roi\\
    ImageIcon icon = new ImageIcon("C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_mic
    Icon icono = new ImageIcon(icon.getImage().getScaledInstance(lblRoi.getWidth(), lblRoi.getHeight(), Image.SCALE_DEFAULT));
    lblHistRoi.setText("");
    lblHistRoi.setIcon(icono);
}

public void HistSeg() {
    String ruta;
    ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\p_tresh\\";
    /*
    if (seleccion_op.equals("Si")) {
        ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\treshhold_positive\\"
    } else {
        ruta = "C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\treshhold_negative\\"
    }
    */
}

```

```

    Mat image = Highgui.imread(ruta + nombreArchivo);

    Mat src = new Mat(image.height(), image.width(), CvType.CV_8UC2);

    Imgproc.cvtColor(image, src, Imgproc.COLOR_RGB2GRAY);

    Vector<Mat> bgr_planes = new Vector<>();
    Core.split(src, bgr_planes);

    MatOfInt histSize = new MatOfInt(256);

    final MatOfFloat histRange = new MatOfFloat(0f, 256f);

    boolean accumulate = false;

    Mat b_hist = new Mat();

    Imgproc.calcHist(bgr_planes, new MatOfInt(0), new Mat(), b_hist, histSize, histRange, accumulate);

    int hist_w = 512;
    int hist_h = 600;
    long bin_w;
    bin_w = Math.round((double) (hist_w / 256));

    Mat histImage = new Mat(hist_h, hist_w, CvType.CV_8UC1);

    Core.normalize(b_hist, b_hist, 3, histImage.rows(), Core.NORM_MINMAX);

    for (int i = 1; i < 256; i++) {
        Core.line(histImage, new Point(bin_w * (i - 1), hist_h - Math.round(b_hist.get(i - 1, 0)[0])),
            new Point(bin_w * (i), hist_h - Math.round(Math.round(b_hist.get(i, 0)[0]))),
            new Scalar(255, 0, 0), 2, 8, 0);
    }

    Highgui.imwrite("C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_micro\\hist\\segme
    ImageIcon icon = new ImageIcon("C:\\Users\\ALEX\\Desktop\\Desktop\\Tesis Breast Cancer - Alex\\MicroBreastCancer\\imagenes_mic
    Icon icono = new ImageIcon(icon.getImage().getScaledInstance(lblRoi.getWidth(), lblRoi.getHeight(), Image.SCALE_DEFAULT));
}

```

```

        lblHistSeg.setText("");
        lblHistSeg.setIcon(icono);
    }

    // Variables declaration - do not modify
    private javax.swing.JMenuItem PreProcSeg;
    private javax.swing.JLabel jLabel1;
    private javax.swing.JMenu jMenu1;
    private javax.swing.JMenu jMenu2;
    private javax.swing.JMenuBar jMenuBar1;
    private javax.swing.JMenuItem jMenuItem1;
    private javax.swing.JMenuItem jMenuItem2;
    private javax.swing.JMenuItem jMenuItem3;
    private javax.swing.JMenuItem jMenuItem4;
    private javax.swing.JPanel jPanel1;
    private javax.swing.JPanel jPanel2;
    private javax.swing.JPanel jPanel3;
    private javax.swing.JTabbedPane jTabbedPane1;
    private javax.swing.JTabbedPane jTabbedPane2;
    private javax.swing.JTabbedPane jTabbedPane3;
    private javax.swing.JLabel lblHistRoi;
    private javax.swing.JLabel lblHistSeg;
    private javax.swing.JLabel lblMicro;
    private javax.swing.JLabel lblRoi;
    private javax.swing.JLabel lblSegmentada;
    private javax.swing.JLabel lblSegmentada_rect;
    private javax.swing.JPanel panelPrincipal;
    private javax.swing.JTextField txtImagen;
    // End of variables declaration
}

```

```

#include "caracteristicas.h"

Caracteristicas::Caracteristicas()
{
}

double Caracteristicas::mean(Mat src){
    cv::Scalar mean, stddev;
    meanStdDev(src, mean, stddev, cv::Mat());
    return mean.val[0];
}

double Caracteristicas::desv(Mat src){
    cv::Scalar mean, stddev;
    meanStdDev(src, mean, stddev, cv::Mat());
    return stddev.val[0];
}

double Caracteristicas::skewness(Mat src){
    this->suma = 0.0;
    double meanValue = mean(src);
    double desvValue = desv(src);

    for (int j = 0; j < src.rows; j++) {
        uchar* data = src.ptr<uchar>(j);
        for (int i = 0; i < src.cols; i++) {
            int valor = data[i];
            double superior = pow((valor - meanValue), 3);
            double inferior = pow(desvValue, 3);
            double div = superior / inferior;
            this->suma += div;
        }
    }
    return this->suma / (src.rows * src.cols);
}

double Caracteristicas::skewness(Mat src, double mean, double desv){
    this->suma = 0.0;

    for (int j = 0; j < src.rows; j++) {
        uchar* data = src.ptr<uchar>(j);
        for (int i = 0; i < src.cols; i++) {
            int valor = data[i];
            double superior = pow((valor - mean), 3);
            double inferior = pow(desv, 3);
            double div = superior / inferior;
            this->suma += div;
        }
    }
    return this->suma / (src.rows * src.cols);
}

```

```

double Caracteristicas::kurtosis(Mat src){
    this->suma = 0.0;
    double meanValue = mean(src);
    double desvValue = desv(src);

    for (int j = 0; j < src.rows; j++) {
        uchar* data = src.ptr<uchar>(j);
        for (int i = 0; i < src.cols; i++) {
            int valor = data[i];
            double superior = pow((valor - meanValue), 4);
            double inferior = pow(desvValue, 4);
            double div = (superior / inferior) - 3;
            this->suma += div;
        }
    }
    return this->suma / (src.rows * src.cols);
}

double Caracteristicas::kurtosis(Mat src, double mean, double desv){
    this->suma = 0.0;

    for (int j = 0; j < src.rows; j++) {
        uchar* data = src.ptr<uchar>(j);
        for (int i = 0; i < src.cols; i++) {
            int valor = data[i];
            double superior = pow((valor - mean), 4);
            double inferior = pow(desv, 4);
            double div = (superior / inferior) - 3;
            this->suma += div;
        }
    }
    return this->suma / (src.rows * src.cols);
}

vector<double> Caracteristicas::glcmCV(Mat src){
    vector<double> retorno;
    IplImage copy = src;
    IplImage* nueva_imagen = &copy;
    IplImage* src = cvCloneImage(nueva_imagen);
    CvGLCM* glcm;
    glcm = cvCreateGLCM(src, 2, NULL, 4, CV_GLCM_OPTIMIZATION_NONE);
    cvCreateGLCMDescriptors(glcm, CV_GLCMDESC_OPTIMIZATION_ALLOWDOUBLENES);
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_HOMOGENITY));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_CONTRAST));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_CORRELATION));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_ENERGY));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_ENTROPY));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_CLUSTERSHADE));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_CLUSTER_TENDENCY));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_MAXIMUM_PROBABILITY));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_CORRELATION_INFO1));
    retorno.push_back(cvGetGLCMDescriptor(glcm, 0, CV_GLCMDESC_CORRELATION_INFO2));
    cvReleaseImage(&src);
    return retorno;
}

```

```

#include "mainwindow.h"
#include "ui_mainwindow.h"
#include "procesamiento.h"
#include "caracteristicas.h"
#include <QMessageBox>
#include <QInputDialog>
#include <QLineEdit>
#include <QProcess>
#include <ctime>
#include <sys/timeb.h>
#include <iostream>
#include <fstream>
#include <QDirIterator>

MainWindow::MainWindow(QWidget *parent) :
    QMainWindow(parent),
    ui(new Ui::MainWindow)
{
    ui->setupUi(this);
}

MainWindow::~MainWindow()
{
    delete ui;
}

using namespace std;

void MainWindow::on_btnCaracteristicas_clicked()
{
    Caracteristicas c;

    double mean, contraste, desv, correlacion, energy, kurtosis, skewness, homo;
    vector<double> res;

    ui->txtCaracteristicas->clear();
    QString ruta_carpeta = ui->txtRutaCarpeta->text().trimmed();

    QDirIterator dirIt(ruta_carpeta, QDirIterator::Subdirectories);
    QString carac_q = "";
    String carac_s = "";
    while (dirIt.hasNext()) {
        dirIt.next();

        if (QFileInfo(dirIt.filePath()).isFile()){
            try {

                Mat output_img = imread(dirIt.filePath().toStdString());

                cvtColor( output_img, output_img, CV_RGB2GRAY);

                mean = c.mean(output_img);
                desv = c.desv(output_img);

                skewness = c.skewness(output_img, mean, desv);
                kurtosis = c.kurtosis(output_img, mean, desv);

                // contraste = c.contrast(img);

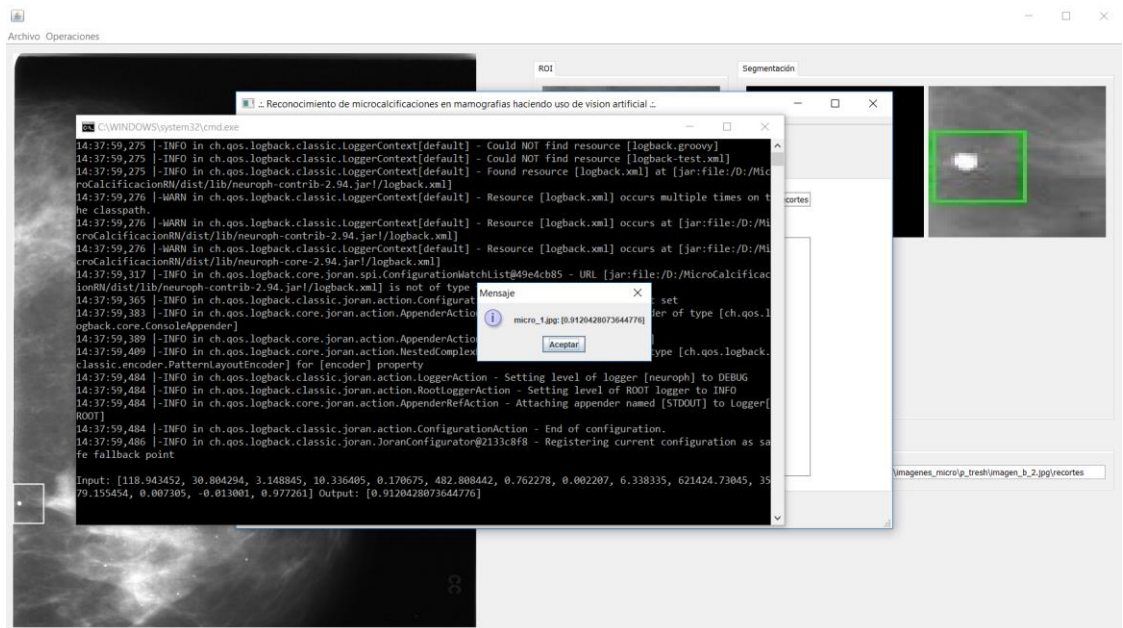
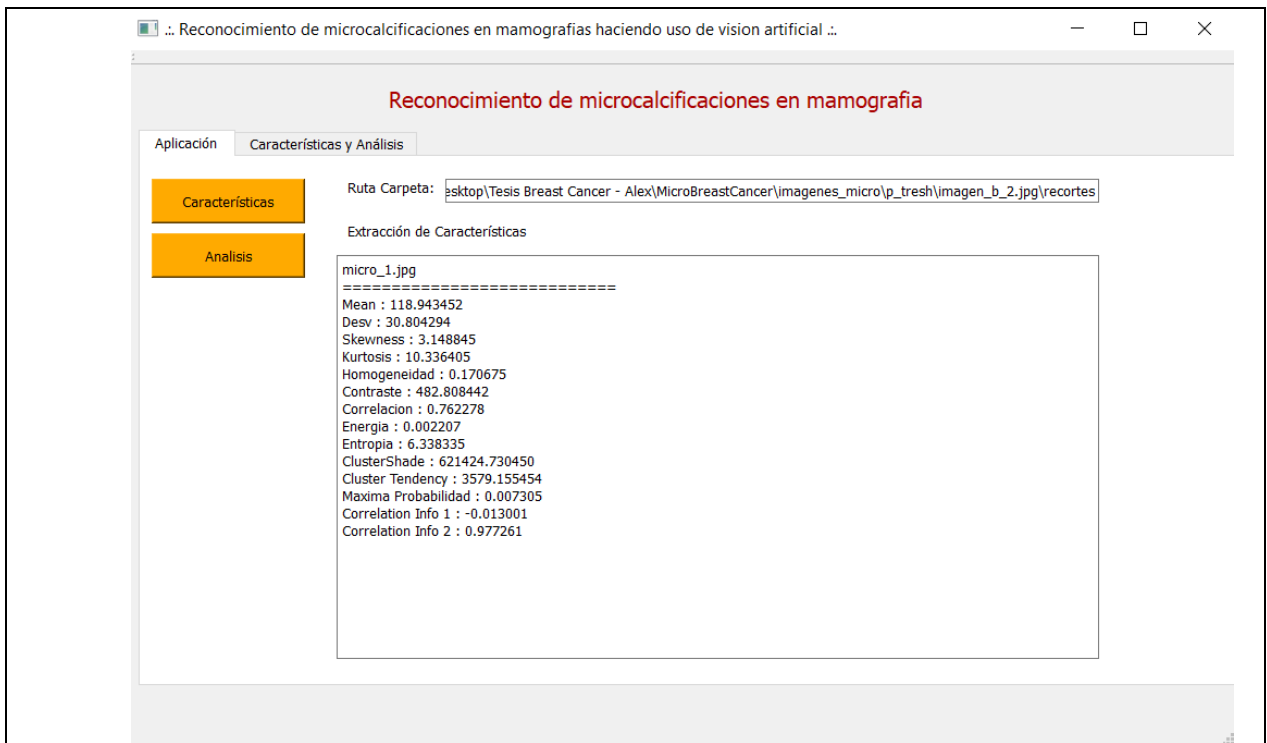
                res = c.glcmCV(output_img);

                cout << "Mean : " << std::to_string(mean)<<endl;
                cout << "Desv : " << std::to_string(desv)<<endl;
                cout << "Skewness : " << std::to_string(skewness)<<endl;
                cout << "Kurtosis : " << std::to_string(kurtosis)<<endl;
                cout << "Homogeneidad : " << std::to_string(res.at(0))<<endl;
                cout << "Contraste : " << std::to_string(res.at(1))<<endl;
                cout << "Correlacion : " << std::to_string(res.at(2))<<endl;
                cout << "Energia : " << std::to_string(res.at(3))<<endl;
                cout << "Entropia : " << std::to_string(res.at(4))<<endl;
                cout << "ClusterShade : " << std::to_string(res.at(5))<<endl;
                cout << "Cluster Tendency : " + std::to_string(res.at(6))<<endl;
                cout << "Maxima Probabilidad : " << std::to_string(res.at(7))<<endl;
                cout << "Correlation Info 1 : " << std::to_string(res.at(8))<<endl;
                cout << "Correlation Info 2 : " << std::to_string(res.at(9))<<endl;

                carac_s = "";
                carac_q = "";

                carac_s += dirIt.fileName().toStdString() + "\n";
                carac_s += "=====\n";
                carac_s += "Mean : " + std::to_string(mean) + "\n";
                carac_s += "Desv : " + std::to_string(desv) + "\n";
                carac_s += "Skewness : " + std::to_string(skewness) + "\n";
                carac_s += "Kurtosis : " + std::to_string(kurtosis) + "\n";
                carac_s += "Homogeneidad : " + std::to_string(res.at(0)) + "\n";
                carac_s += "Contraste : " + std::to_string(res.at(1)) + "\n";
                carac_s += "Correlacion : " + std::to_string(res.at(2)) + "\n";
            }
        }
    }
}

```

Visualizamos la interfaz creada y vemos que reconoce satisfactoriamente que hay una microcalcificación benigna en dicha mamografía

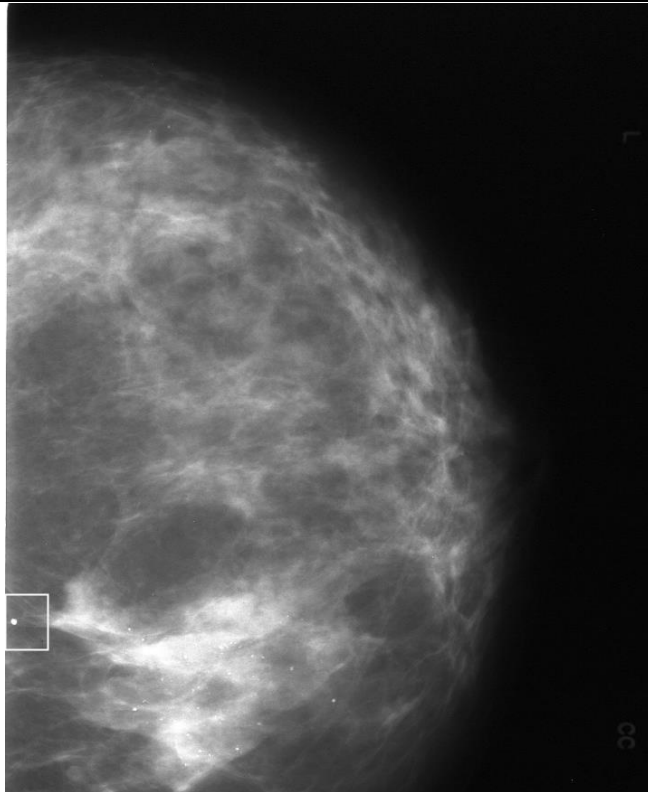
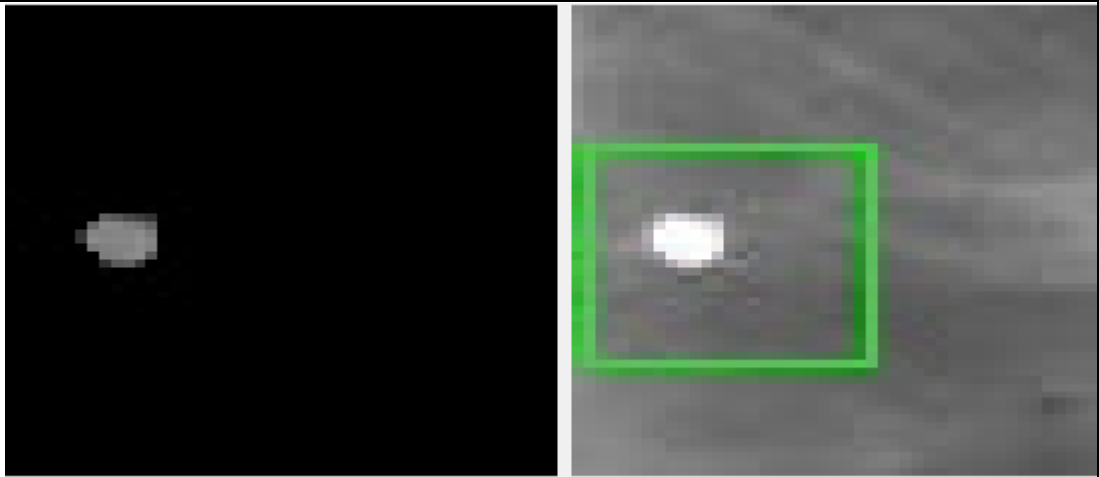


Imagen Mamográfica con la ROI seleccionada



ROI apartada de toda la imagen mamográfica



ROI segmentada usando threshold y aplicando contornos para así poder extraer las características

