



UNIVERSIDAD CÉSAR VALLEJO

**ESCUELA DE POSGRADO
PROGRAMA ACADÉMICO DE MAESTRÍA EN INGENIERÍA
DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA
INFORMACIÓN**

**Modelo predictivo en la detección de cáncer en mamografías, Lima
2024**

TESIS PARA OBTENER EL GRADO ACADÉMICO DE:
Maestro en Ingeniería de Sistemas con mención en Tecnologías de la Información

AUTOR:

Melendez Ruiz, Luis Jose (orcid.org/0000-0002-2179-4415)

ASESORES:

Dr. Vargas Huaman, Jhonatan Isaac (orcid.org/0000-0002-1433-7494)

Mg. Puente Zamora, Jonathan Alexis (orcid.org/0009-0007-1034-1617)

LÍNEA DE INVESTIGACIÓN:

Sistemas de Información y Comunicaciones

LÍNEA DE RESPONSABILIDAD SOCIAL UNIVERSITARIA:

Desarrollo económico, empleo y emprendimiento

LIMA – PERÚ

2024



UNIVERSIDAD CÉSAR VALLEJO

ESCUELA DE POSGRADO

MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN

Declaratoria de Autenticidad del Asesor

Yo, VARGAS HUAMAN JHONATAN ISAAC, docente de la ESCUELA DE POSGRADO MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA NORTE, asesor de Tesis titulada: "Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024", cuyo autor es MELENDEZ RUIZ LUIS JOSE, constato que la investigación tiene un índice de similitud de 18.00%, verificable en el reporte de originalidad del programa Turnitin, el cual ha sido realizado sin filtros, ni exclusiones.

He revisado dicho reporte y concluyo que cada una de las coincidencias detectadas no constituyen plagio. A mi leal saber y entender la Tesis cumple con todas las normas para el uso de citas y referencias establecidas por la Universidad César Vallejo.

En tal sentido, asumo la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión tanto de los documentos como de información aportada, por lo cual me someto a lo dispuesto en las normas académicas vigentes de la Universidad César Vallejo.

LIMA, 07 de Agosto del 2024

Apellidos y Nombres del Asesor:	Firma
JHONATAN ISAAC VARGAS HUAMAN DNI: 70430225 ORCID: 0000-0002-1433-7494	Firmado electrónicamente por: JIVARGASH el 07- 08-2024 23:31:21

Código documento Trilce: TRI - 0853970





UNIVERSIDAD CÉSAR VALLEJO

ESCUELA DE POSGRADO

MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN

Declaratoria de Originalidad del Autor

Yo, MELENDEZ RUIZ LUIS JOSE estudiante de la ESCUELA DE POSGRADO MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA NORTE, declaro bajo juramento que todos los datos e información que acompañan la Tesis titulada: "Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024", es de mi autoría, por lo tanto, declaro que la Tesis:

1. No ha sido plagiada ni total, ni parcialmente.
2. He mencionado todas las fuentes empleadas, identificando correctamente toda cita textual o de paráfrasis proveniente de otras fuentes.
3. No ha sido publicada, ni presentada anteriormente para la obtención de otro grado académico o título profesional.
4. Los datos presentados en los resultados no han sido falseados, ni duplicados, ni copiados.

En tal sentido asumo la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión tanto de los documentos como de la información aportada, por lo cual me someto a lo dispuesto en las normas académicas vigentes de la Universidad César Vallejo.

Nombres y Apellidos	Firma
LUIS JOSE MELENDEZ RUIZ DNI: 44180515 ORCID: 0000-0002-2179-4415	Firmado electrónicamente por: LMELENDEZRU el 07- 08-2024 22:30:08

Código documento Trilce: TRI - 0853971



Dedicatoria

A mis amados padres, Salvit y José Gabino, por todo el cariño, dedicación, tiempo y recursos que han invertido en mi formación personal y profesional.

A mis hijos: José Ladislao y Luis Daniel, quienes son mi mayor orgullo y motivación.

Agradecimiento

A Dios, de quien emana toda gracia, buena dádiva, conocimiento y sabiduría.

A mis asesores: Dr. Jhonatan Isaac Vargas Huaman y Mg. Jonathan Alexis Puente Zamora, por sus oportunas orientaciones durante la realización de esta investigación.

A la Universidad César Vallejo, por haberme acogido con los brazos abiertos durante estos años de estudio.

Índice de contenidos

	Pg.
Carátula	i
Declaratoria de autenticidad del asesor	ii
Declaratoria de originalidad del autor	iii
Dedicatoria.....	iv
Agradecimiento.....	v
Índice de contenidos	vi
Índice de tablas.....	vii
Índice de figuras.....	viii
Resumen	ix
Abstract.....	x
I. INTRODUCCIÓN	1
II. METODOLOGÍA	15
III. RESULTADOS	18
IV. DISCUSIÓN.....	25
V. CONCLUSIONES	30
VI. RECOMENDACIONES.....	31
REFERENCIAS	32
ANEXOS.....	38

Índice de tablas

	Pg.
Tabla 1 Estadísticos descriptivos de la dimensión sensibilidad	18
Tabla 2 Estadístico de normalidad de la dimensión sensibilidad	19
Tabla 3 Prueba de hipótesis de la dimensión sensibilidad	20
Tabla 4 Estadísticos descriptivos de la dimensión especificidad	21
Tabla 5 Estadístico de normalidad de la dimensión especificidad	21
Tabla 6 Prueba de hipótesis de la dimensión especificidad	22
Tabla 7 Estadísticos descriptivos de la dimensión exactitud	22
Tabla 8 Estadístico de normalidad de la dimensión exactitud	23
Tabla 9 Prueba de hipótesis de la dimensión exactitud.....	24
Tabla 10 Estadísticos descriptivos de la dimensión sensibilidad en test y retest	82
Tabla 11 Estadísticos descriptivos de la dimensión especificidad en test y retest	82
Tabla 12 Estadísticos descriptivos de la dimensión exactitud en test y retest	83
Tabla 13 Estadístico de normalidad de la dimensión sensibilidad en test y retest	84
Tabla 14 Estadístico de normalidad de la dimensión especificidad en test y retest.....	84
Tabla 15 Estadístico de normalidad de la dimensión exactitud en test y retest	84
Tabla 16 Coeficiente de correlación de la dimensión sensibilidad en test y retest.....	85
Tabla 17 Coeficiente de correlación de la dimensión especificidad en test y retest.....	85
Tabla 18 Coeficiente de correlación de la dimensión exactitud en test y retest	85
Tabla 19 Configuraciones para entrenamiento del modelo predictivo	120

Índice de figuras

	Pg.
Figura 1 Ciclo de vida del software con el enfoque incremental	107
Figura 2 Diagrama general de Casos de uso	108
Figura 3 Interfaz principal del software.....	112
Figura 4 Panel del módulo Análisis de Mamografías.....	113
Figura 5 Panel del módulo Evaluación de Lesiones	113
Figura 6 Panel del módulo Otras tareas.....	114
Figura 7 Código de red neuronal artificial tipo Perceptrón Multicapa.....	115
Figura 8 Diagrama de flujo del funcionamiento del algoritmo AIRS.....	117
Figura 9 Prueba de diferentes configuraciones para mejorar los indicadores.....	119

Resumen

Esta investigación se alinea con el ODS "Industria, Innovación e Infraestructura" y su propósito fue determinar la influencia de un modelo predictivo en la detección de cáncer en mamografías en Lima, 2024. La investigación fue de tipo aplicada, con un diseño preexperimental y un enfoque cuantitativo. La población consistió en 598 lesiones (143 masas y 455 calcificaciones) presentes en 544 mamografías, mientras que la muestra consistió en 235 lesiones (56 masas y 179 calcificaciones). Para la recolección de datos se utilizó como técnica el fichaje, empleando como instrumento una ficha de registro para los indicadores en el pre-test y post-test. En el análisis de datos, se aplicó el estadístico de normalidad de Kolmogórov-Smirnov y, en consecuencia, la prueba no paramétrica de Wilcoxon para muestras relacionadas para el contraste de hipótesis.

En cuanto a los resultados, se obtuvo un valor de $p < 0.05$ para la hipótesis general, lo que llevó a aceptar la hipótesis alternativa y, por tanto, a rechazar la hipótesis nula. En ese sentido, se concluyó que la influencia del modelo predictivo mejora significativamente la detección de cáncer en mamografías, Lima 2024.

Palabras clave: modelo de simulación, detección de enfermedades, cáncer.

Abstract

This research aligns with the SDG 'Industry, Innovation, and Infrastructure' and aimed to determine the influence of a predictive model on cancer detection in mammograms in Lima, 2024. The study was applied, with a pre-experimental design and a quantitative approach. The population consisted of 598 lesions (143 masses and 455 calcifications) present in 544 mammograms, while the sample consisted of 235 lesions (56 masses and 179 calcifications). Data collection used the technique of charting, employing a recording form for indicators in the pre-test and post-test. For data analysis, the Kolmogorov-Smirnov normality test was applied, and consequently, the Wilcoxon signed-rank test for related samples was used for hypothesis testing.

Regarding the results, a p-value <0.05 was obtained for the general hypothesis, leading to the acceptance of the alternative hypothesis and, therefore, the rejection of the null hypothesis. In this sense, it was concluded that the influence of the predictive model significantly improves cancer detection in mammograms, Lima 2024.

Keywords: cybernetic models, screening for disease, cancer.

I. INTRODUCCIÓN

La palabra Cáncer es generalmente utilizada para referirse a un gran conjunto de enfermedades que tienen como característica determinante la rápida proliferación de células anormales capaces de invadir áreas cercanas del cuerpo y, en ocasiones, otros órganos mediante un proceso conocido como metástasis.

El cáncer es una de las principales causas de muerte en todo el mundo: sólo en el año 2020 se atribuyeron a esta enfermedad casi 10 millones de defunciones (World Health Organization, 2022); siendo el cáncer de mama el tipo más común, con más de 2,2 millones de nuevos casos y 685 mil defunciones.

En el continente americano, se observó una tendencia similar en la prevalencia del cáncer de mama, siendo el más frecuente entre las mujeres con un 25,2%. Además, fue la segunda causa principal de muerte por tumores malignos en este grupo (15,1%), solo superada por el cáncer de pulmón (17,4%). Para el año 2030, la Organización Panamericana de la Salud estima que la región de América Latina y el Caribe enfrentará más de 572 mil nuevos casos y aproximadamente 130 mil fallecimientos por cáncer de mama (Pan American Health Organization, 2018).

En nuestro país, las cifras no son menos alarmantes, liderando la incidencia el cáncer de mama, junto al cáncer de cuello uterino. Entre enero y diciembre de 2022, se registraron 10,686 nuevos casos de cáncer en 62 hospitales, concentrándose el 86% de ellos en 25 hospitales. Las estadísticas de cáncer de mama para ese año, refieren 1,237 nuevos casos confirmados, de los cuales 1,224 corresponden a mujeres, y sólo 13 a hombres (RENACE, 2023).

En cuanto a la incidencia, no hay diferencias significativas entre la población de países desarrollados y en desarrollo. Sin embargo, las diferencias en la mortalidad son significativas: la mayoría de los fallecimientos ocurren en países de bajos ingresos, donde los diagnósticos suelen realizarse en etapas avanzadas debido a la falta de acceso a servicios de salud y la falta de conciencia sobre la detección temprana. En contraste, en países de altos ingresos, la mortalidad se ha reducido en un 40%. Sin embargo, aunque las estadísticas podrían sugerir lo contrario, el tratamiento del cáncer de mama puede ser muy efectivo, con tasas de supervivencia

del 90% o superiores, especialmente cuando se detecta en etapas tempranas (World Health Organization, 2024), es decir, cuando el cáncer es pequeño y localizado. Por lo tanto, promover la detección precoz sigue siendo crucial para mejorar el pronóstico y la supervivencia de las personas afectadas por esta enfermedad.

Antes de continuar, es oportuno destacar que la presente investigación se relaciona con el ODS “Industria, Innovación e Infraestructura”, en especial con la meta 9.b “Apoyar el desarrollo de tecnologías, la investigación y la innovación nacional en los países en desarrollo, incluso garantizando un entorno normativo propicio a la diversificación industrial y la adición de valor a los productos básicos, entre otras cosas”.

Actualmente la mamografía, o radiografía de las mamas, es la prueba de detección temprana más utilizada, tanto por su accesibilidad como por su efectividad. Los datos obtenidos de los programas de detección mediante mamografías indican, tras 11 años de seguimiento, una disminución del 20% en la mortalidad específica por cáncer de mama (World Health Organization, 2015). Otras pruebas que también generan imágenes, como la ecografía o la resonancia magnética, suelen emplearse para situaciones específicas, siendo consideradas complementarias a la mamografía. A pesar de su efectividad, la mamografía no es una prueba de detección perfecta. Debido a factores como: la edad y peso del paciente, la densidad del tejido mamario, la naturaleza de la lesión, la calidad de la imagen, la experiencia del radiólogo, etc.; existe el inconveniente de que la mamografía sea susceptible a una mala interpretación (National Cancer Institute, 2024). Por ejemplo, en el tejido mamario normal puede existir la posibilidad de que un cáncer de mama pase desapercibido en una mamografía, lo que se conoce como un falso negativo. Por otro lado, un radiólogo podría identificar una anomalía que parezca ser cáncer, pero en realidad no lo sea, lo que se denomina falso positivo. Además del estrés asociado con la preocupación por un diagnóstico de cáncer de mama, un falso positivo implica realizar más pruebas y consultas de seguimiento para los pacientes. Para mitigar estas limitaciones, a menudo se requiere realizar múltiples mamografías o utilizar otras técnicas de imagen mamaria. Sin embargo, para eliminar cualquier duda, los especialistas generalmente recomiendan realizar una biopsia (BreastCancer.Org, 2022).

Aunque los factores técnicos, del paciente y de la lesión son importantes, generalmente la causa de los errores está relacionada con las conductas de búsqueda, percepción y toma de decisiones de los especialistas. De hecho, se estima que los radiólogos todavía pasan por alto entre el 10-30% de los cánceres de mama, el 80% de las mujeres llamadas para vistas adicionales tienen resultados normales y el 40% de las lesiones evaluadas mediante biopsia son benignas. Las investigaciones muestran que la mayoría de los cánceres de mama que se pasan por alto son realmente visibles y se observan, pero pasan desapercibidos o se consideran benignos (Ekpo et al., 2018). Para minimizar las omisiones y errores causados por el factor humano, se ha recurrido al uso de algoritmos inteligentes, que pueden procesar la información de las imágenes, analizar las características de las lesiones y realizar una clasificación; convirtiéndose en parte del trabajo clínico de rutina para el diagnóstico de cáncer de mama. Los hallazgos efectuados por estos algoritmos se utilizan como segunda opinión y soporte en la toma de decisiones (Shiraishi et al., 2011), ayudando a los radiólogos a interpretar la información presente en las mamografías con un mayor grado de precisión, reduciendo la tasa de falsos positivos, falsos negativos y las intervenciones quirúrgicas innecesarias (Kooi et al., 2017).

Los algoritmos inteligentes utilizados en el diagnóstico de cáncer de mama tradicionalmente pertenecen al grupo denominado redes neuronales artificiales. Sin embargo, el almacenamiento y tratamiento de la información en estas redes suele representar una caja negra, y por ello su optimización resulta compleja. En la presente investigación se propone la implementación de un algoritmo más comprensible: un sistema inmune artificial de reconocimiento (AIRS). El AIRS fue presentado por Andrew Watkins y Lois Boggess en el Congreso de Computación Evolutiva '02 (Watkins & Boggess, 2002) y desde entonces ha sido aplicado con notable éxito a la resolución de diferentes problemas de las ciencias de la salud.

Respecto al cáncer de mama, el AIRS ha sido utilizado en el reconocimiento de células malignas en biopsias (Polat et al., 2005), en el diagnóstico de lesiones mediante una combinación de características observadas tanto en mamografía, ecografía y resonancia magnética (Katsis et al., 2013); así como, en la clasificación de masas en mamografías (Djamila & Belkhodja, 2019).

Sin embargo, no existen investigaciones relacionadas a la implementación del AIRS como un Modelo Predictivo en la Detección de cáncer en mamografías, por lo que es válido preguntar: ¿Cuál será la influencia de este Modelo Predictivo en la Detección de cáncer en mamografías?, y específicamente: 1) ¿Cuál será la influencia de este Modelo Predictivo en la sensibilidad de la Detección de cáncer en mamografías?, 2) ¿Cuál será la influencia de este Modelo Predictivo en la especificidad de la Detección de cáncer en mamografías? y, 3) ¿Cuál será la influencia de este Modelo Predictivo en la exactitud de la Detección de cáncer en mamografías?

Por tanto, la presente investigación tiene como justificación teórica: aportar nuevo conocimiento a la literatura especializada; como justificación práctica: desarrollar un modelo predictivo que ayude a los especialistas a lograr una detección más precisa, reduciendo la tasa de falsos positivos y falsos negativos en la aplicación de las pruebas de tamizaje; y como justificación social, facilitar a la población en riesgo, el acceso a una prueba de detección con mayor valor predictivo, lo que ayudaría a reducir las pruebas diagnósticas de confirmación.

A partir de las preguntas de investigación, se propuso como objetivo general: Determinar la influencia del Modelo Predictivo en la Detección de cáncer en mamografías, Lima 2024; y como objetivos específicos: 1) Determinar la influencia del Modelo Predictivo en la sensibilidad de la Detección de cáncer en mamografías, 2) Determinar la influencia del Modelo Predictivo en la especificidad de la Detección de cáncer en mamografías y, 3) Determinar la influencia del Modelo Predictivo en la exactitud de la Detección de cáncer en mamografías.

A continuación, se presentan como antecedentes los trabajos de investigación más recientes a nivel internacional, extraídos desde la National Library of Medicine, sitio oficial del Centro Nacional de Información Biotecnológica de los Estados Unidos de América, en cuanto a la detección de cáncer en mamografías utilizando diferentes algoritmos y técnicas de inteligencia artificial.

Arce et al. (2023), evaluaron la habilidad de un sistema de detección asistida por computador basado en redes neuronales convolucionales, para el carcinoma mamario lobular invasivo (un subtipo de cáncer de mama), utilizando mamografías

digitales de pacientes quienes habían sido diagnosticados mediante biopsia entre enero de 2017 y enero de 2022. Con un intervalo de confianza de 95%, el sistema mostró una sensibilidad del 80%, detectando el 100% de calcificaciones, 82% de masas irregulares y 86% de masas con márgenes espiculados. Sin embargo, el 88% de detecciones, también incluía al menos un área marcada como falso positivo, concluyendo que, aunque el sistema evaluado fue exitoso para marcar las áreas de malignidad en las mamografías digitales, los falsos positivos reducen su potencial uso en la vida real.

Pacilè et al. (2020), evaluaron los beneficios de utilizar una herramienta basada en Inteligencia Artificial para el proceso de detección de cáncer en mamografías tradicionales o bidimensionales. En dicho estudio multilectura, multicaso y retrospectivo, 14 radiólogos revisaron la mitad de un data set de 240 mamografías digitales adquiridas entre 2013 y 2016, utilizando la herramienta basada en IA y la otra mitad prescindiendo de ella, durante una primera sesión; y viceversa, durante una segunda sesión. Los resultados para el área bajo la curva, con un intervalo de confianza del 95%, fueron de 0.797 cuando los radiólogos utilizaron la herramienta y 0.769 cuando prescindieron de ella. Por otro lado, la sensibilidad fue superior en 0.033 cuando utilizaron la herramienta basada en IA. Respecto al flujo de trabajo, no se alteró significativamente, aunque el tiempo utilizado decreció ligeramente para la segunda sesión. En conclusión, esta investigación clínica demostró que utilizar una herramienta basada en IA mejoró el rendimiento de los radiólogos en la detección de cáncer de mama sin prolongar su flujo de trabajo.

Sasaki et al. (2020), compararon el rendimiento en la detección de cáncer en mamografías digitales de una población de 310 mujeres japonesas, entre un panel de tres lectores humanos y el sistema Transpara basado en Inteligencia Artificial. Estas mujeres se sometieron a exámenes de mamografía digital entre enero y octubre de 2018. Los panelistas utilizaron el sistema estándar BI-RADS para otorgar una puntuación, mientras que Transpara proporcionó una puntuación de apoyo a la decisión de los especialistas. Entonces, se compararon el área bajo la curva (AUC) y la sensibilidad en cada una de las mamografías leídas. El AUC para los lectores humanos fue de 0.816 y para Transpara de 0.706; mientras que la sensibilidad fue de 89% y 86% respectivamente. Sin embargo, aunque el rendimiento del sistema basado

de IA fue ligeramente menor que el de los especialistas, se espera que mejoras en los algoritmos utilizados reduzcan la diferencia entre ambos para la detección eficiente del cáncer de mama.

Rodríguez-Ruiz et al. (2019), compararon el desempeño en la detección del cáncer de mama de los radiólogos que leen exámenes mamográficos con el soporte de un sistema de inteligencia artificial versus aquellos que no recurren a dicho soporte. En el estudio, de tipo retrospectivo, totalmente cruzado, multicaso y multilectura, se incluyeron los exámenes de mamografía digital de 240 mujeres realizados entre 2013 y 2017, cuyo rango de edad variaba entre 39 y 89 años, siendo el promedio 62 años. Los 240 exámenes: 100 con cáncer confirmado. 40 falsos positivos y 100 normales, fueron interpretados por 14 radiólogos cualificados según la Ley de Estándares de Calidad en Mamografía, tanto teniendo el soporte del sistema de IA como prescindiendo de éste. El sistema de IA brindó a los radiólogos soporte interactivo para la toma de decisiones mostrando una probabilidad de cáncer local. Se midieron el AUC, la sensibilidad y la especificidad, así como el tiempo de lectura. Los resultados en promedio presentan un AUC de 0.89 con el soporte del sistema de IA y 0.87 sin éste. También aumentó la sensibilidad y la especificidad, siendo igual a 86% y 79% con el soporte del sistema de IA versus un 83% y 77% sin el soporte. En relación al tiempo, los radiólogos con el soporte del sistema de IA se demoraron en promedio 149 segundos y sin el soporte 146 segundos. En conclusión, los radiólogos mejoraron su desempeño en la detección de cáncer de mama con el soporte del sistema de IA, sin requerir un tiempo significativo de lectura adicional.

Adicionalmente, se encontraron algunos artículos en la base de datos bibliográfica SCOPUS, publicados entre 2014 y 2016, que tratan específicamente sobre el uso de sistemas inmunes artificiales en la detección y clasificación de cáncer de mama, y se presentan a continuación:

Magna et al. (2016), investigaron las propiedades de un clasificador basado en un conjunto de redes inmunes artificiales adaptativas denominado A²INET, aplicado a indicadores de imágenes de mamografías destinados a identificar anomalías de asimetría bilateral que se saben están correlacionadas a un mayor riesgo de cáncer de mama. A²INET demostró altos niveles de precisión en la clasificación, incluso al ser evaluado en dos conjuntos de datos distintos y diversos tanto en entrenamiento

como en pruebas. Los resultados fueron alentadores, alcanzando una precisión del 90%, una sensibilidad del 93% y una especificidad del 87%, lo cual sugiere que estas herramientas podrían ser útiles en aplicaciones clínicas prácticas.

Daoudi et al. (2014), propusieron varias mejoras del algoritmo CLONALG, uno de los algoritmos más populares entre los sistemas inmunes artificiales, en la forma de dos nuevos algoritmos denominados CCS-AIS y MF-AIS, que fueron aplicados para la clasificación del cáncer de mama utilizando descriptores obtenidos desde la Digital Database for Screening Mammography (DDSM). Los resultados obtenidos fueron de 98.31% para CCS-AIS, 97.74% para MF-AIS y 95.57% para el algoritmo original, probando la efectividad de las dos mejoras propuestas e implementadas en los descriptores utilizados.

Como nota final, respecto a los antecedentes, a nivel nacional y local no se encontraron investigaciones relacionadas con el uso de modelos predictivos en la detección de cáncer en mamografías.

A continuación, se presenta la conceptualización de la variable independiente: Modelo Predictivo, así como de la variable dependiente: Detección de cáncer en mamografías. Así mismo, se presentan las dimensiones e indicadores.

En relación a la variable independiente: Modelo Predictivo, Pérez (2022), explica que un modelo predictivo se desarrolla para adquirir conocimiento a partir de un conjunto de datos específico, los cuales se analizan durante una fase de entrenamiento con el objetivo de poder hacer predicciones sobre eventos futuros. Por su parte, Mora (2022), comenta que la medicina basada en evidencia utiliza modelos que parten de una hipótesis y luego buscan datos para respaldarla o refutarla. Dicho enfoque también se aplica en la creación de modelos predictivos confiables que pueden tener un impacto en la práctica clínica. El crecimiento en la cantidad de datos almacenados en registros clínicos electrónicos y el aumento en la capacidad computacional están impulsando el uso de técnicas de aprendizaje de máquina. Estas técnicas permiten desarrollar nuevos análisis predictivos y detectar patrones previamente desconocidos. Como resultado, están siendo cada vez más integradas en la práctica clínica diaria, proporcionando una toma de decisiones más precisa y rápida al cambiar la perspectiva de los datos hacia la información. Finalmente,

Sánchez R. et al. (2018), declaran que también es posible emplear modelos de pronóstico en línea en la atención oncológica para clasificar a los pacientes y determinar cuándo es apropiado utilizar terapias adicionales en personas con cáncer de mama en etapas tempranas. Este enfoque busca examinar la relación entre diferentes variables y anticipar la supervivencia en este grupo específico de pacientes.

La variable independiente: Modelo predictivo, se desarrolla en el contexto de la teoría del Procesamiento digital de imágenes y en la teoría sobre los Sistemas Inmunes Artificiales, las cuales se describen a continuación.

El Procesamiento digital de imágenes consiste en un conjunto de actividades que siguen un orden de ejecución y tienen por misión eliminar información que no sea relevante, y evidenciar cualquier área de interés, a fin de facilitar la búsqueda de información en las imágenes digitales (Davies, 2017). Estas actividades son: a) Adquisición de la imagen, mediante un dispositivo de captura. b) Preprocesamiento, que consiste en aplicar un grupo limitado de técnicas de corrección para adecuar las características presentes en la imagen hacia un estándar que permita reconocer la información importante con más precisión. c) Segmentación y Etiquetado, constituye la primera fase para el reconocimiento de figuras u objetos, al describir la imagen como una colección de áreas de interés que no se superponen unas sobre otras, por lo que es posible separar el fondo de los diferentes objetos presentes en la imagen, aplicar técnicas de filtrado para detectar píxeles, destacar o desvanecer líneas, bordes y regiones, realizar operaciones lógicas entre píxeles y etiquetarlos como parte de un objeto. d) Descripción y Extracción de características, que permite especificar cada objeto identificado, mediante sus características como altura, anchura, área, perímetro, etc., para poder reconocerlo. Cada objeto puede estar conformado por varios segmentos, o incluso, regiones. Las características extraídas son almacenadas en forma de un arreglo con una cierta cantidad de variables para representar una clase, por lo que es habitual que las características buscadas y extraídas no presenten variación con relación a su tamaño, posición, localización y orientación de la figura u objeto en cuestión. e) Reconocimiento y Clasificación, que permite destinar porciones del vector de características a diferentes clases o conjuntos de objetos, utilizando técnicas como el clustering, métodos estadísticos, funciones

discriminantes, redes neuronales, etc.; y f) Toma de decisiones, que dependerá del problema a resolver (Márquez, 2021).

Por otro lado, para comprender la teoría sobre los Sistemas Inmunes Artificiales o AIS, debemos conocer su fundamento: el sistema inmunitario biológico, el cual se compone de una variedad de moléculas, células y órganos que trabajan en conjunto para preservar la salud del cuerpo, defendiéndolo de sustancias dañinas (antígenos) traídas por la invasión de agentes externos (patógenos). El sistema inmunitario biológico emplea un mecanismo de protección por niveles. El primer nivel es nuestra piel, que actúa como una barrera física que funciona como un amplio escudo. El segundo nivel lo componen diferentes condiciones fisiológicas tales como la saliva, el sudor, la temperatura, etc.; que resultan nada favorables para la supervivencia de ciertos patógenos. Los siguientes niveles son dos sistemas altamente relacionados: el Sistema Inmune Innato, llamado así porque nuestro organismo desde antes del nacimiento va desarrollando la habilidad de identificar y eliminar ciertos tipos de patógenos; tiene como papel principal activar la respuesta del Sistema Inmune Adaptativo, que muestra tres características muy interesantes: ante la presencia de un antígeno parecido a otro identificado con anterioridad, responderá de forma genérica para reprimir el ataque inicial (generalización), pero luego de un tiempo de exposición, iniciará un contraataque específico para anular y destruir completamente al nuevo antígeno (especificidad); además de que, una vez identificado el antígeno lo recordará (memoria), respondiendo de manera más efectiva y eficiente en futuras apariciones (Cruz & Coello, 2002).

Entonces, el sistema inmune biológico es extremadamente interesante para el procesamiento de información debido a su robustez, su estructura multicapa y diversa en componentes, su naturaleza distribuida y su capacidad de autoorganización. Además, de tolerar el ruido y ser muy eficaz en el reconocimiento de patrones, presenta identidad, al distinguir entre lo que pertenece y no pertenece al organismo, y un comportamiento emergente de gran complejidad, como consecuencia de interacciones locales elementales entre sus componentes (Cayzer & Aickelin, 2002)

Estas cualidades han inspirado la creación de diversos principios y teorías, así como el desarrollo de modelos computacionales conocidos como Sistemas Inmunes Artificiales. Estos sistemas adaptativos están fundamentados en la inmunología

teórica y se basan en las funciones, principios y modelos inmunológicos observados en los vertebrados, aplicándose a problemas complejos en diferentes dominios. Los AIS han sido utilizados en un sinnúmero de problemas, incluyendo el monitoreo de diferentes procesos industriales, la detección de virus, la clasificación de e-mail, el entrenamiento de robots, la detección de tráfico irregular en redes y el diagnóstico médico (Greensmith et al., 2010).

Entre los principios, teorías y modelos computacionales más representativos e importantes se encuentran: a) Principio de Discriminación entre lo Propio y lo Ajeno, es decir, entre lo que pertenece a un organismo y lo que viene del exterior. b) Modelo de Selección Negativa, basado en la eliminación de células que no son capaces de utilizar el principio anterior. De acuerdo a este modelo, las células inmunes que logran sobrevivir al proceso de aprender a discriminar entre lo propio y lo ajeno, son capaces de activarse solas y participar en una respuesta inmune ante la presencia de elementos extraños y nocivos. c) Modelo de Selección Clonal, fundamentado en la proliferación de los linfocitos B y en la mutación efectuada por estos para lograr la especificidad frente a antígenos similares o desconocidos. d) Teoría del Peligro, que cambia el enfoque entre la distinción entre propio y ajeno, por nocivo e inocuo, explicando así, por qué en ciertos casos los sistemas de defensa del organismo no son capaces de reaccionar ante la aparición de una extraña bacteria o virus, al afirmar que el sistema inmune biológico en realidad responde solamente ante el peligro; y e) Teoría de la Red Inmune, que propone que las células de nuestro organismo son capaces de reconocerse las unas a las otras a través del contacto continuo, lo que permite desarrollar un mecanismo de retroalimentación y estimular la habilidad de memoria (Cruz & Coello, 2002).

Respecto a la conceptualización de la variable dependiente: Detección de cáncer en mamografías, Aguilar-Torres et al. (2021), comentan que en estudios con asignación aleatoria, se ha observado que realizar tamizajes de manera bianual contribuye a reducir la mortalidad asociada al cáncer de mama. La mamografía presenta una ventaja significativa en la detección de lesiones no palpables en comparación con otros métodos como el examen clínico y la autoexploración. Sin embargo, alrededor del 10% de los resultados de mamografías resultan ser falsos positivos, lo que implica que aproximadamente la mitad de las mujeres

experimentarán un resultado falso positivo después de diez años de exámenes anuales. Por otro lado, según Pons-Rodríguez et al. (2022), aunque se estima que el tamizaje poblacional mediante mamografía reduce la mortalidad por cáncer en un 20% y previene una muerte por cáncer por cada 235 mujeres examinadas durante 20 años, también puede dar lugar a efectos secundarios no deseados, como los falsos positivos y el sobrediagnóstico de tumores que no habrían causado síntomas durante la vida de la mujer. Dado el considerable impacto emocional y psicológico del cáncer de mama en las mujeres diagnosticadas, el sobrediagnóstico y el tratamiento innecesario son considerados los efectos adversos más graves de la detección temprana, lo que ha generado un amplio debate. Asimismo, López-Muñoz & Salamanca-Gómez (2019), concuerdan que, aunque la detección de la enfermedad en etapas tempranas, independientemente de la edad, conduce a una disminución en la tasa de mortalidad, se han observado complicaciones asociadas tanto con el diagnóstico de una enfermedad grave, como la ansiedad y la depresión, así como con el tratamiento quirúrgico del cáncer de mama, incluyendo el dolor crónico postoperatorio y la hinchazón del tejido linfático. Por lo tanto, es crucial identificar estas complicaciones para poder brindar un tratamiento oportuno y potencialmente mejorar la calidad de vida de las pacientes.

La variable dependiente: Detección de cáncer en mamografías, se fundamenta en la teoría sobre las lesiones comúnmente observadas en las mamografías, siendo éstas, las calcificaciones y las masas; enfocándose en las características que permiten a los especialistas identificar su presencia con relativa precisión, y si son sospechosas de malignidad o no.

Las calcificaciones son depósitos de minerales que se observan como puntos o manchas blancas en las radiografías. La mayoría no son un signo de cáncer. Las calcificaciones redondeadas y grandes, denominadas macro calcificaciones suelen corresponder a tumores de tipo benigno, considerándose comunes en mujeres mayores de 50 años. Las muy pequeñas, denominadas microcalcificaciones son las que usualmente requieren vigilancia, pues su presencia indicaría cambios en la mama que deben controlarse, especialmente si aparecen aisladas o en grupos. La experiencia clínica sugiere rasgos de malignidad si las calcificaciones se presentan amorfas, con tamaños menores a 0.5 mm de diámetro, finas, ramificadas, lineales

pero discontinuas o agrupadas en forma de racimo. Por otro lado, las masas, conocidas como nódulos o protuberancias, son lesiones que ocupan espacio y son perceptibles en dos vistas radiográficas diferentes. Cuando una masa solo es visible desde una proyección, se suele denominar densidad hasta que se confirme su carácter tridimensional. Las masas pueden o no presentar sensibilidad, dolor o drenado. Muchas mujeres las desarrollan en sus años reproductivos, pero la mayoría resultan ser no cancerosas. Respecto a la forma, pueden presentarse redondas, ovales, lobulares o irregulares. También hay casos especiales, como la distorsión arquitectónica, que es la alteración del patrón del tejido mamario sin la presencia de una masa claramente definida, la cual puede manifestarse a través de espiculaciones que se irradian desde un punto, o retracción local en el borde del tejido mamario subcutáneo. Otro caso son las densidades asimétricas, que son áreas densas visibles en dos proyecciones pero que carecen de los bordes y características típicas de una masa verdadera. Estas áreas podrían representar zonas de tejido mamario normal que requieren una evaluación mediante imágenes adicionales para descartar la presencia de una masa real o una distorsión arquitectónica significativa. En todo caso, los rasgos de malignidad son frecuentemente asociados a márgenes micro lobulados, oscuros, mal definidos o espiculados. También, para el diagnóstico, se considera la atenuación esperada de los rayos X por la lesión, ya que la mayoría de los nódulos forman una región visible de igual o mayor densidad que el tejido mamario normal circundante. Por tanto, es poco común que un cáncer de mama sea menos denso que su entorno. Además, el cáncer de mama nunca es graso, ya que la grasa es radiolúcida. Sin embargo, el cáncer podría contener depósitos de grasa. Entonces, una mayor densidad en una lesión aumenta la probabilidad de que sea maligna (Responsable & Victoria, 1999).

En lo concerniente a las dimensiones de la variable dependiente: Detección de cáncer en mamografías, se medirán la sensibilidad, la especificidad y la exactitud. Es preciso recordar que los indicadores de estas dimensiones pueden ser afectados por diferentes condiciones del paciente, entre ellos: la edad (Aguilar-Torres et al., 2021).

La sensibilidad permite conocer qué tan eficaz es una prueba para demostrar que alguien realmente padece una enfermedad, es decir, que obtenga un resultado positivo estando enfermo (Torregroza-Diazgranados, 2021), y se medirá a través del

indicador: porcentaje de sensibilidad, cuya fórmula considera las regiones detectadas como lesiones que sí son lesiones, entre el total de regiones que son lesiones de cáncer de mama, multiplicado por cien. Según Abugattas Saba et al. (2015), el porcentaje de sensibilidad de la detección de cáncer en mamografías es aproximadamente de 63% en mamas con alta densidad y de 87% en mamas con mayor contenido de grasa, mientras que, según Octavio (2011), la mamografía a partir de los 40 o 50 años presenta un porcentaje de sensibilidad de 77 a 95%.

Por otro lado, la especificidad permite conocer qué tan eficaz es una prueba para demostrar que alguien realmente no padece una enfermedad, es decir, que obtenga un resultado negativo estando sano (Torregroza-Diazgranados, 2021), y se medirá a través del indicador: porcentaje de especificidad, cuya fórmula considera las regiones detectadas como sanas que no son lesiones, entre el total de regiones que no son lesiones de cáncer de mama, multiplicado por cien. Según Abugattas Saba et al. (2015), el porcentaje de especificidad de la detección de cáncer en mamografías es de alrededor de 87% en mamas muy densas y de 99% en mamas con mayor cantidad de grasa, mientras que, según Octavio (2011), la mamografía a partir de los 40 o 50 años presenta un porcentaje de especificidad de 94 a 97%.

Finalmente, la exactitud, también conocida como certeza diagnóstica, consiste en medir cuán bien se comporta la prueba, es decir, que reconozca a los enfermos como enfermos y a los sanos como sanos, frente al total de los individuos sometidos a evaluación (Torregroza-Diazgranados, 2021), por tanto, la exactitud se medirá a través del indicador: Porcentaje de exactitud, que considera las regiones detectadas como lesiones que sí son lesiones, más las regiones detectadas como sanas que no son lesiones, entre el total de regiones de la mama, multiplicado por cien. Sin embargo, Martínez-Pérez & Pérez-Martín (2022), consideran que el AUC es un mejor indicador de exactitud, pues toma en cuenta el punto de sensibilidad y especificidad más alto para conocer el rendimiento global de la prueba. De igual manera, Armesto (2011), menciona que la curva ROC proporciona un buen índice de discriminación entre diferentes estados de salud, pues representa la sensibilidad en función de los falsos positivos, es decir, de 1-especificidad.

Por último, una vez definidas las dimensiones de la variable dependiente y sus indicadores, se propuso como hipótesis general que: La influencia del Modelo

Predictivo mejora significativamente la Detección de cáncer en mamografías, estableciéndose como hipótesis específicas que: La influencia del Modelo Predictivo incrementa significativamente 1) la sensibilidad, 2) la especificidad y 3) la exactitud de la Detección de cáncer en mamografías.

II. METODOLOGÍA

2.1. Tipo, enfoque y diseño de investigación

Esta investigación fue de tipo Aplicada, pues se utilizaron conocimientos previamente investigados para lograr objetivos específicos, buscando validar la implementación potencial de productos, prototipos o modelos en los niveles de transferencia y madurez tecnológica (Castro-Maldonado et al, 2019); adoptando un enfoque Cuantitativo, al examinar la realidad de manera objetiva mediante mediciones numéricas, con el fin de identificar predicciones o patrones de comportamiento del fenómeno en cuestión (Hernández-Sampieri & Mendoza, 2018); y un diseño Preexperimental, al realizar como primer paso el establecimiento de evidencia a favor de una intervención sobre un grupo de la población a través de un pre-test y post-test (Ramos-Galarza, 2021).

El alcance de la investigación se centró en los siguientes aspectos: i) Utilizar el sistema inmune artificial de reconocimiento (AIRS), basado en el modelo de selección clonal e inspirado en el funcionamiento del sistema inmunitario biológico, para el desarrollo de un modelo predictivo. ii) Utilizar un conjunto de mamografías de dominio público, obtenidas desde el Repositorio Digital de Cáncer de Mama (BCDR) para el entrenamiento y prueba del sistema. Con antelación, las mamografías se procesaron digitalmente, extrayéndose las características de cada imagen. Las características tanto de regiones sanas, como de regiones con lesiones validadas previamente por los especialistas, fueron almacenadas en vectores, y sirvieron como antígenos para estimular el reconocimiento de patrones por parte del sistema inmune, desarrollando así el modelo predictivo. iii) Medir la influencia del modelo predictivo en la detección de cáncer en mamografías, a través de los indicadores: porcentaje de sensibilidad, porcentaje de especificidad y porcentaje de exactitud. Para ello, se compararon las regiones detectadas por el modelo versus las regiones segmentadas y validadas por los especialistas, obteniendo valores de verdaderos positivos, falsos negativos, falsos positivos y verdaderos negativos, útiles para el cálculo de los indicadores. Finalmente, se pusieron a prueba las hipótesis específicas, para deducir la aceptación o rechazo de la hipótesis general, cumpliéndose el objetivo general y los objetivos específicos establecidos en esta investigación.

2.2. Variables

La definición conceptual de la variable independiente y dependiente, así como de las dimensiones e indicadores empleados, se puede encontrar a detalle en el Anexo 1.

2.3. Población y muestra

En relación con la población, del conjunto de mamografías mencionado en el segundo párrafo, se extrajo el marco poblacional que consistió en aquellas que contenían masas o calcificaciones, considerando estas lesiones como las unidades de análisis. En total, el marco poblacional estuvo compuesto por 598 lesiones (143 masas y 455 calcificaciones). La muestra, seleccionada con un nivel de confianza del 95% y un margen de error del 5%, fue obtenida mediante muestreo probabilístico estratificado, resultando en un total de 235 lesiones (56 masas y 179 calcificaciones). El cálculo del tamaño de la muestra puede verse en el Anexo 7.

2.4. Técnicas e instrumentos de recolección de datos

Para la recolección de datos, se empleó la técnica de fichaje, utilizando como instrumento una ficha de registro para los tres indicadores. Las fichas de registro utilizadas en el pre-test y post-test están disponibles en el Anexo 2. Para asegurar la validez del instrumento, se realizó una validación mediante juicio de expertos. Los documentos de validación se muestran en el Anexo 3. Además, se evaluó la confiabilidad del instrumento mediante un análisis de consistencia interna, utilizando una correlación test-retest. El coeficiente de correlación de rangos de Spearman mostró una correlación positiva muy alta entre las dos mediciones. El análisis completo de la confiabilidad se encuentra en el Anexo 4.

2.5. Métodos para el análisis de datos

Los datos obtenidos fueron sometidos a la prueba de normalidad de Kolmogórov-Smirnov (muestra mayor a 50 unidades), descubriendo que tenían una distribución no normal, y al tener muestras relacionadas, se aplicó la prueba de tipo no paramétrica de Wilcoxon para evaluar las hipótesis específicas y llegar a una conclusión respecto a la hipótesis general. Para el tratamiento de los datos y la visualización de la información se utilizó el software estadístico IBM SPSS Statistics versión 25.0.

2.6. Aspectos éticos

En cuanto a las consideraciones éticas, esta investigación se llevó a cabo de acuerdo con los principios de beneficencia, no maleficencia, autonomía y justicia establecidos en el artículo 3° del Código de Ética en Investigación de la Universidad César Vallejo, aprobado por la RCU N° 0340-2021-UCV. Además, se adhirió rigurosamente a la Guía de Elaboración de Trabajos para Grados y Títulos, aprobada por la RVI N° 081-2024-VI-UCV el 01 de abril de 2024, que define la estructura y las rúbricas de evaluación para los trabajos académicos de la UCV. Asimismo, el contenido fue redactado, respetando la autoría según la Guía para la redacción de textos académicos estilo APA, del Programa de Formación Humanística UCV Lima Norte 2020, usando la séptima edición para las referencias y citas de contenido.

Finalmente, indicar que las mamografías obtenidas del Repositorio Digital, no atentaron ni lesionaron los derechos de los individuos sobre la privacidad de sus datos personales u otra información sensible, debido a que fue liberado al dominio público el 18 de abril de 2012, como un repositorio de amplio alcance, con dos objetivos fundamentales: 1) establecer una referencia novedosa para explorar métodos de detección y diagnóstico basados en computadora, y 2) capacitar a estudiantes de medicina, médicos formados y otros profesionales relacionados con la medicina; tal como se indica en la página oficial del repositorio: <https://bcdr.eu/information/about>. Además, al utilizar las imágenes de un repositorio público en esta investigación, se favorece la reproducción de la metodología utilizada y la validación de los hallazgos obtenidos.

III. RESULTADOS

A continuación, se presentan los resultados de la comparación de estadísticos descriptivos, prueba de normalidad y contraste de hipótesis de acuerdo al orden de los objetivos de esta investigación.

Para la prueba de normalidad de los datos, como se indicó en el capítulo II. Metodología, se utilizó el test de Kolmogórov-Smirnov debido a que la muestra para ambos tipos de lesiones (masas y calcificaciones) superaba las 50 unidades. El test se aplicó con un nivel de confianza del 95% y un nivel de significancia del 5%. En todos los casos, se estableció como Hipótesis Nula (H_0) que los datos seguían una distribución normal, y como Hipótesis Alternativa (H_a) que los datos no seguían una distribución normal. El criterio de decisión fue el siguiente: si $p < 0.05$, se rechazaba H_0 y se aceptaba H_a ; por el contrario, si $p \geq 0.05$, se aceptaba H_0 y se rechazaba H_a .

En función al resultado de la prueba de normalidad, en cada caso se procedió a realizar el contraste de hipótesis, teniendo en consideración que las muestras eran relacionadas.

Respecto al objetivo específico 1) Determinar la influencia del Modelo Predictivo en la sensibilidad de la Detección de cáncer en mamografías, se obtuvieron los siguientes resultados:

Tabla 1

Estadísticos descriptivos de la dimensión sensibilidad

Estadísticos descriptivos	Masas		Calcificaciones	
	Pre-test	Post-test	Pre-test	Post-test
<i>Elementos válidos</i>	56	56	179	179
<i>Elementos perdidos</i>	0	0	0	0
<i>Media</i>	83.0979	67.6370	95.0366	80.1160
<i>Mediana</i>	97.0700	66.8200	100.0000	85.7100
<i>Varianza</i>	692.408	359.684	198.132	526.956
<i>Desv. Desviación</i>	26.31364	18.96533	14.07593	22.95551
<i>Mínimo</i>	.00	20.00	25.00	.00

<i>Máximo</i>	100.00	100.00	100.00	100.00
<i>Desv. Error</i>	3.51631	2.53435	1.05208	1.71578

Fuente: autoría personal.

De la tabla 1, se observaron para ambos estratos en el pre-test, que la media y la mediana fueron superiores respecto al post-test. Sin embargo, no se encontró otro patrón en los datos que favoreciera la idea de que el modelo predictivo propuesto influyó negativamente la sensibilidad de la detección de cáncer en mamografías, comparado con los algoritmos inteligentes tradicionalmente utilizados.

Tabla 2

Estadístico de normalidad de la dimensión sensibilidad

Descripción	Masas		Calcificaciones	
	Pre-test	Post-test	Pre-test	Post-test
<i>Estadístico K-S</i>	.265	.101	.426	.243
<i>gl</i>	56	56	179	179
<i>Sig. (p)</i>	.000	.200	.000	.000

Fuente: autoría personal.

De la tabla 2, se observaron valores de p menores a 0.05 en ambos estratos, con excepción del post-test para las lesiones de tipo masa, lo que indicaba que los datos no presentaban una distribución normal, siendo prudente utilizar una prueba no paramétrica para el contraste de hipótesis, en este caso, la prueba de Wilcoxon para muestras relacionadas.

Para la aplicación de la prueba de Wilcoxon se definió como Hipótesis nula (H_0): La influencia del Modelo Predictivo no incrementa significativamente la sensibilidad de la Detección de cáncer en mamografías; y, como Hipótesis alternativa (H_1): La influencia del Modelo Predictivo incrementa significativamente la sensibilidad de la Detección de cáncer en mamografías; siendo el criterio de decisión, si $p \geq 0.05$, aceptar H_0 , caso contrario aceptar H_1 .

Tabla 3***Prueba de hipótesis de la dimensión sensibilidad***

	Masas	Calcificaciones
<i>Hipótesis nula</i>	<i>La mediana de las diferencias entre Pre-test y Post-test es igual a 0.</i>	<i>La mediana de las diferencias entre Pre-test y Post-test es igual a 0.</i>
<i>Prueba</i>	<i>Prueba de rangos con signo de Wilcoxon para muestras relacionadas.</i>	<i>Prueba de rangos con signo de Wilcoxon para muestras relacionadas.</i>
<i>Sig. (p)</i>	<i>.000</i>	<i>.000</i>
<i>Decisión</i>	<i>Rechazar la hipótesis nula.</i>	<i>Rechazar la hipótesis nula.</i>

Fuente: autoría personal.

De la tabla 3, se observó que los valores de p para ambos estratos fueron menores a 0.05, por tanto, ante los supuestos planteados, se rechazó H_0 y se aceptó H_1 , concluyendo que el modelo predictivo incrementa significativamente la sensibilidad de la detección de cáncer en mamografías.

La conclusión anterior, podría parecer contraproducente en primera instancia a los resultados mostrados en la tabla 1, donde la media y la mediana para ambos estratos fueron superiores en el pre-test respecto al post-test, lo que indicaría que el modelo predictivo tuvo una influencia negativa sobre la sensibilidad de la detección. Sin embargo, la superioridad de la media y la mediana en el pre-test no contradice necesariamente los resultados de la prueba de Wilcoxon. De hecho, se sabe que la mediana y sobre todo la media, pueden verse influenciadas por valores extremos, por lo cual, la prueba de Wilcoxon se enfoca en analizar cómo cambian los valores, considerando la mejora o empeoramiento de forma individual. Eso quiere decir que, aunque el valor promedio y la mediana para ambos estratos disminuyó en el post-test, Wilcoxon identificó una clara tendencia de mejora en la sensibilidad de la detección de cáncer en una cantidad significativa de mamografías con el modelo predictivo propuesto.

Respecto al objetivo específico 2) Determinar la influencia del Modelo Predictivo en la especificidad de la Detección de cáncer en mamografías, se obtuvieron los siguientes resultados:

Tabla 4***Estadísticos descriptivos de la dimensión especificidad***

Estadísticos descriptivos	Masas		Calcificaciones	
	Pre-test	Post-test	Pre-test	Post-test
<i>Elementos válidos</i>	56	56	179	179
<i>Elementos perdidos</i>	0	0	0	0
<i>Media</i>	50.7925	60.7382	45.0580	66.1277
<i>Mediana</i>	43.9250	59.6050	42.9300	65.9500
<i>Varianza</i>	498.799	99.904	488.837	96.904
<i>Desv. Desviación</i>	22.33382	9.99519	22.10966	9.84396
<i>Mínimo</i>	17.93	44.39	4.24	43.50
<i>Máximo</i>	93.90	91.61	99.85	96.24
<i>Desv. Error</i>	2.98448	1.33566	1.65255	.73577

Fuente: autoría personal.

De la tabla 4, se observó para ambos estratos en el post-test, que la media y el mínimo fueron superiores respecto al pre-test, además de una varianza y error más bajos, favoreciendo la idea de que el modelo predictivo propuesto influyó positivamente la especificidad de la detección del cáncer en mamografías, comparado con los algoritmos inteligentes tradicionalmente utilizados.

Tabla 5***Estadístico de normalidad de la dimensión especificidad***

Descripción	Masas		Calcificaciones	
	Pre-test	Post-test	Pre-test	Post-test
<i>Estadístico K-S</i>	.136	.095	.081	.043
<i>gl</i>	56	56	179	179
<i>Sig. (p)</i>	.011	.200	.007	.200

Fuente: autoría personal.

De la tabla 5, se observó para ambos estratos que p fue menor a 0.05 sólo en el pre-test. Por tanto, dado que en el pre-test los datos no presentaron una distribución

normal, se optó por una prueba no paramétrica para el contraste de hipótesis, en este caso, la prueba de Wilcoxon para muestras relacionadas.

Para la aplicación de la prueba de Wilcoxon se definió como Hipótesis nula (H_0): La influencia del Modelo Predictivo no incrementa significativamente la especificidad de la Detección de cáncer en mamografías; y, como Hipótesis alternativa (H_1): La influencia del Modelo Predictivo incrementa significativamente la especificidad de la Detección de cáncer en mamografías, siendo el criterio de decisión, si $p \geq 0.05$, aceptar H_0 , caso contrario aceptar H_1 .

Tabla 6

Prueba de hipótesis de la dimensión especificidad

	Masas	Calcificaciones
<i>Hipótesis nula</i>	<i>La mediana de las diferencias entre Pre-test y Post-test es igual a 0.</i>	<i>La mediana de las diferencias entre Pre-test y Post-test es igual a 0.</i>
<i>Prueba</i>	<i>Prueba de rangos con signo de Wilcoxon para muestras relacionadas.</i>	<i>Prueba de rangos con signo de Wilcoxon para muestras relacionadas.</i>
<i>Sig. (p)</i>	<i>.001</i>	<i>.000</i>
<i>Decisión</i>	<i>Rechazar la hipótesis nula.</i>	<i>Rechazar la hipótesis nula.</i>

Fuente: autoría personal.

De la tabla 6, se observó que los valores de p para ambos estratos fueron menores a 0.05, por tanto, ante los supuestos planteados, se rechazó H_0 y se aceptó H_1 , concluyendo que el modelo predictivo incrementa significativamente la especificidad de la detección de cáncer en mamografías. Esta conclusión fue acorde con el análisis realizado para la tabla 4.

Respecto al objetivo específico 3) Determinar la influencia del Modelo Predictivo en la exactitud de la Detección de cáncer en mamografías, se obtuvieron los siguientes resultados:

Tabla 7

Estadísticos descriptivos de la dimensión exactitud

Estadísticos descriptivos	Masas		Calcificaciones	
	Pre-test	Post-test	Pre-test	Post-test
<i>Elementos válidos</i>	56	56	179	179
<i>Elementos perdidos</i>	0	0	0	0
<i>Media</i>	51.4934	60.8854	45.4456	66.0730
<i>Mediana</i>	47.4600	60.1900	43.0700	65.6700
<i>Varianza</i>	472.236	92.175	478.572	94.173
<i>Desv. Desviación</i>	21.73100	9.60077	21.87629	9.70430
<i>Mínimo</i>	18.17	45.29	4.32	44.98
<i>Máximo</i>	93.72	87.65	99.83	96.19
<i>Desv. Error</i>	2.90393	1.28296	1.63511	.72533

Fuente: autoría personal.

De la tabla 7, se observó para ambos estratos en el post-test, que la media y el mínimo fueron superiores respecto al pre-test, además de una varianza y error más bajos, favoreciendo la idea de que el modelo predictivo propuesto influyó positivamente la exactitud de la detección del cáncer en mamografías, comparado con los algoritmos inteligentes tradicionalmente utilizados.

Tabla 8

Estadístico de normalidad de la dimensión exactitud

Descripción	Masas		Calcificaciones	
	Pre-test	Post-test	Pre-test	Post-test
<i>Estadístico K-S</i>	.135	.064	.077	.044
<i>gl</i>	56	56	179	179
<i>Sig. (p)</i>	.013	.200	.012	.200

Fuente: autoría personal.

De la tabla 8, se observó para ambos estratos que p fue menor a 0.05 sólo en el pre-test. Por tanto, dado que en el pre-test los datos no presentaron una distribución normal, se optó por una prueba no paramétrica para el contraste de hipótesis, en este caso, la prueba de Wilcoxon para muestras relacionadas.

Para la aplicación de la prueba de Wilcoxon se definió como Hipótesis nula (H_0): La influencia del Modelo Predictivo no incrementa significativamente la exactitud de la Detección de cáncer en mamografías; y, como Hipótesis alternativa (H_1): La influencia del Modelo Predictivo incrementa significativamente la exactitud de la Detección de cáncer en mamografías, siendo el criterio de decisión, si $p \geq 0.05$, aceptar H_0 , caso contrario aceptar H_1 .

Tabla 9

Prueba de hipótesis de la dimensión exactitud

	Masas	Calcificaciones
<i>Hipótesis nula</i>	<i>La mediana de las diferencias entre Pre-test y Post-test es igual a 0.</i>	<i>La mediana de las diferencias entre Pre-test y Post-test es igual a 0.</i>
<i>Prueba</i>	<i>Prueba de rangos con signo de Wilcoxon para muestras relacionadas.</i>	<i>Prueba de rangos con signo de Wilcoxon para muestras relacionadas.</i>
<i>Sig. (p)</i>	<i>.001</i>	<i>.000</i>
<i>Decisión</i>	<i>Rechazar la hipótesis nula.</i>	<i>Rechazar la hipótesis nula.</i>

Fuente: autoría personal.

De la tabla 9, los valores de p para ambos estratos fueron menores a 0.05, por tanto, ante los supuestos planteados, se rechazó H_0 y se aceptó H_1 , concluyendo que el modelo predictivo incrementa significativamente la exactitud de la detección de cáncer en mamografías. Esta conclusión fue acorde con el análisis realizado para la tabla 7.

Por tanto, aplicando el análisis deductivo, en función a las conclusiones alcanzadas en cada una de las hipótesis específicas respecto a las dimensiones de sensibilidad, especificidad y exactitud, se aceptó la hipótesis general de esta investigación, concluyendo que: La influencia del Modelo Predictivo mejora significativamente la Detección de cáncer en mamografías.

IV. DISCUSIÓN

A continuación, se discuten los resultados obtenidos en el capítulo anterior frente a los resultados presentados en los antecedentes.

En primer lugar, en relación a la dimensión sensibilidad, los valores alcanzados por el indicador porcentaje de sensibilidad con el modelo predictivo propuesto, tanto para las masas (67.6%), como para las calcificaciones (80.1%), resultaron consistentes con los valores mencionados por Abugattas Saba et al. (2015), ubicándose dentro del intervalo de 63 a 87 por ciento referenciado por dicho autor. De igual forma, otros investigadores como Arce et al. (2023), Sasaki et al. (2020), y Rodríguez-Ruiz et al. (2019), reportaron valores de sensibilidad dentro de dicho intervalo, los cuales fueron: 80% para el primero y 86% para los dos restantes. Sólo Magna et al. (2016) indicó haber alcanzado un porcentaje de sensibilidad de 93%.

Al respecto, es importante recordar que la sensibilidad permite conocer qué tan eficaz es una prueba para demostrar que alguien realmente padece una enfermedad, es decir, que una persona obtenga un resultado positivo estando enfermo (Torregroza-Diazgranados, 2021). Un alto porcentaje de sensibilidad representa una reducción en la aparición de falsos negativos, por tanto, una prueba sensible no pasará por alto los casos de cáncer que sí necesitan intervención médica, aumentando la confianza en la detección de cáncer en etapas tempranas, cuando el tratamiento suele ser más efectivo.

En segundo lugar, respecto a la dimensión especificidad, los valores alcanzados por el indicador porcentaje de especificidad con el modelo predictivo propuesto, tanto para las masas (60.7%), como para las calcificaciones (66.1%) no resultaron consistentes con los valores mencionados por Abugattas Saba et al. (2015), encontrándose por debajo del intervalo de 87 a 99 por ciento referenciado por dicho autor. De igual forma otros investigadores, como Rodríguez-Ruiz et al. (2019) y Magna et al. (2016), reportaron haber alcanzado valores de especificidad de 79% y 87%, respectivamente. De esto, se advierte que ninguno de ellos presentó valores de especificidad dentro del intervalo de referencia y aunque presentaron porcentajes inferiores, los tales fueron relativamente mayores a los obtenidos en esta investigación.

Respecto a esto, es importante recordar que la especificidad permite conocer qué tan eficaz es una prueba para demostrar que alguien realmente no padece una enfermedad (Torregroza-Diazgranados, 2021), es decir, que una persona obtenga un resultado negativo estando sano. Un alto porcentaje de especificidad representa una reducción en la aparición de falsos positivos, por tanto, una prueba con alta especificidad proporcionará tranquilidad, eliminando la necesidad de recurrir a procedimientos adicionales para confirmar el resultado negativo. Como lo indica Pons-Rodriguez et al. (2022) el impacto emocional y psicológico significativo del cáncer de mama en las mujeres diagnosticadas, el sobrediagnóstico y el tratamiento innecesario se consideran los efectos adversos más serios de la detección temprana, por lo que resulta necesario hacer uso de pruebas con alto porcentaje de especificidad.

Ahora bien, resulta que en muchas ocasiones los indicadores de las dimensiones sensibilidad y especificidad suelen encontrarse inversamente relacionados. Por ejemplo, en el caso de la detección de cáncer de mama, si una región verificada como una lesión es reconocida como tal, se incrementaría la cantidad de verdaderos positivos en 1, caso contrario, si fuese reconocida como una región sana, se incrementaría la cantidad de falsos negativos en 1. De forma análoga, si una región sana fuese reconocida como tal, se incrementaría la cantidad de verdaderos negativos en 1, caso contrario, si la región fuese reconocida como una lesión, se incrementaría la cantidad de falsos positivos en 1. Bajo esta premisa, si entendemos que los casos pueden ser o bien regiones sanas o bien lesiones, generalmente tendremos modelos altamente sensibles, pero con porcentajes bajos de especificidad o modelos que presenten una alta especificidad, pero que presenten una baja sensibilidad. Por tal motivo, es preciso recurrir a un indicador que pueda combinar los verdaderos positivos, los falsos negativos, los falsos positivos y los verdaderos negativos, de modo que obtengamos una medición menos sesgada, siendo este indicador el porcentaje de exactitud.

En este sentido, es importante recordar que la exactitud, también conocida como certeza diagnóstica, permite conocer cuán bien se comporta una prueba, es decir, que reconozca a los enfermos como enfermos y a los sanos como sanos, frente al total de los individuos sometidos a evaluación (Torregroza-Diazgranados, 2021).

Ahora bien, la fórmula tradicional para medir la exactitud está basada en la suma de verdaderos positivos y verdaderos negativos sobre el total de casos; sin embargo, hay muchas otras formas de medirla, por ejemplo, a través de la media geométrica, cuya fórmula corresponde a la raíz cuadrada de la multiplicación del porcentaje de sensibilidad y el porcentaje de especificidad; o utilizando el valor-F, que combina también ambos indicadores con el objeto de ponderar cuán lejos están de la utilidad teórica de la prueba. Asimismo, es posible utilizar el área bajo la curva (AUC), que no es otra cosa que una representación gráfica de la tasa de verdaderos positivos (medida de la sensibilidad) frente a la tasa de falsos positivos, es decir, 1 menos la tasa de verdaderos negativos (medida de la especificidad), sobre un plano cartesiano. En dicho plano, el eje X representa la tasa de falsos positivos, mientras que el eje Y representa la tasa de verdaderos positivos. El punto (0,1) corresponderá a una prueba perfecta, es decir, aquella que es capaz de clasificar todos los casos positivos y negativos correctamente. En tanto que el punto (0,0) corresponderá a una prueba que predice todos los casos como negativos y el punto (1,1) corresponderá a una prueba que predice todos los casos como positivos. Finalmente, el punto (1,0) corresponderá a una prueba que predice todos los casos de forma incorrecta. Algunos autores, como Martínez-Pérez & Pérez-Martín (2022) y Armesto (2011), consideran el AUC como un mejor indicador de exactitud pues elige el punto de sensibilidad y especificidad más altos para conocer el rendimiento global de una prueba.

Entonces, volviendo a los resultados de esta investigación, en relación a la dimensión exactitud, los valores alcanzados por el indicador porcentaje de exactitud con el modelo predictivo propuesto fueron de 60.9% para las masas y 66.1% para las calcificaciones, utilizando la fórmula tradicional de la exactitud. En tanto que, investigadores como Sasaki et al. (2020), Pacilè et al. (2020) y Rodríguez-Ruiz et al. (2019) utilizaron el AUC como indicador de exactitud, reportando valores de 70.6%, 79.7% y 89% respectivamente, los cuales se advierten relativamente altos comparados a los obtenidos en esta investigación.

En este punto, hay que destacar que, aunque se tratan de simples números, es relativamente complejo intentar homologar o equiparar los resultados obtenidos entre diferentes investigaciones, debido a la diversidad de los objetivos perseguidos en cada investigación. Por ejemplo, mientras Arce et al. (2023) trata de detectar específicamente un tipo de cáncer de mama, denominado carcinoma lobular invasivo

(relacionado a la presencia de microcalcificaciones), Magna et al. (2016) se enfoca en identificar anomalías de asimetría bilateral (relacionadas a la presencia de masas), en tanto que Sasaki et al. (2020) y Rodríguez-Ruiz et al. (2019) ejecutan evaluaciones más generales. Además, los algoritmos utilizados también difieren en cada una de estas investigaciones, abarcando desde implementaciones de redes neuronales y sistemas privados de detección asistidos por inteligencia artificial, hasta modificaciones de redes inmunes artificiales. Esta diversidad en el uso de métodos o técnicas en realidad no representa un problema en sí, ya que los resultados podrían ser contrastables siempre que se utilizara exactamente el mismo marco poblacional, sin embargo, el problema es que, en la realidad, el uso de un marco poblacional común es poco probable. Por tanto, los resultados de sensibilidad, especificidad y exactitud obtenidos de los antecedentes son referenciales, y no necesariamente deben tomarse como absolutos para decidir cual modelo es mejor en todos los casos.

Otro detalle a considerar es que la mayoría de investigaciones se decantan por el diagnóstico de cáncer de mama, y pocas tratan sobre la detección de cáncer. Aunque ambos términos suelen utilizarse indiscriminadamente como si se tratase de lo mismo, lo cierto es que, la detección se refiere a la identificación de la presencia o ausencia de una lesión, mientras que el diagnóstico se efectúa una vez identificada la lesión a fin de etiquetarla como benigna o maligna, o clasificarla dentro de algún tipo de cáncer.

A pesar de todas las diferencias mencionadas, los antecedentes citados en esta investigación si coincidieron en un aspecto, y es que todos utilizaron mamografías digitales. Sobre este punto es útil señalar, a modo de aclaración, que las mamografías actualmente pueden realizarse de forma digital, a través de un mamógrafo con características de comunicación e interface similar a una computadora, o de forma mecánica, a través de un material especial conocido como película de emulsión doble, que funciona como un rollo fotográfico, el cual luego de su uso en un mamógrafo tradicional, puede digitalizarse. Sin embargo, cuando se trata de probar nuevos métodos de detección de cáncer de mama asistidos por computadora, es preferible utilizar mamografías digitales para evitar la pérdida de información debido a la digitalización.

Como se indicó en el Capítulo I. Introducción, la mamografía no es una prueba de detección perfecta. Se han identificado múltiples factores que pueden afectar su correcta lectura e interpretación, entre ellos, la calidad de la imagen. Esto último, tiene sentido con lo explicado en el párrafo anterior sobre los tipos de mamografías, y también se puede deducir de la variedad de dispositivos electrónicos o mecánicos mediante los cuales dichas imágenes puedan obtenerse. No todos los mamógrafos instalados en los diferentes consultorios clínicos y centros de salud son de la misma marca o modelo, y aun siéndolo no presentarán exactamente las mismas características ni la misma configuración, lo que implica que las imágenes obtenidas en diferentes equipos tendrán algún grado de diferencia incluso si pertenecen a la misma paciente y fueron tomadas en la misma posición, con despreciable diferencia de tiempo.

En todo caso, siempre es posible minimizar las omisiones y errores causados por los diferentes factores gracias a la tecnología, siendo la presente investigación otra prueba más del uso de los algoritmos inteligentes y potenciales beneficios de los modelos predictivos, como herramientas de segunda opinión y soporte en la toma de decisiones de los especialistas, tal como mencionan Kooi et al. (2017), ayudando a la interpretación más precisa de la información presente en las mamografías, reduciendo la tasa de error, la necesidad de recurrir a pruebas de confirmación y las intervenciones quirúrgicas innecesarias.

V. CONCLUSIONES

Luego del tratamiento y análisis de los datos recopilados, se llegó a las siguientes conclusiones:

Respecto al objetivo específico 1) Determinar la influencia del Modelo Predictivo en la sensibilidad de la Detección de cáncer en mamografías, observándose un valor $p < 0.05$ en la prueba de rangos con signo de Wilcoxon para muestras relacionadas, tanto para lesiones de tipo masa como para las calcificaciones, se concluyó que el modelo predictivo incrementa significativamente la sensibilidad de la detección de cáncer en mamografías.

Respecto al objetivo específico 2) Determinar la influencia del Modelo Predictivo en la especificidad de la Detección de cáncer en mamografías, observándose un valor $p < 0.05$ en la prueba de rangos con signo de Wilcoxon para muestras relacionadas, tanto para lesiones de tipo masa como para las calcificaciones, se concluyó que el modelo predictivo incrementa significativamente la especificidad de la detección de cáncer en mamografías.

Respecto al objetivo específico 3) Determinar la influencia del Modelo Predictivo en la exactitud de la Detección de cáncer en mamografías, observándose un valor $p < 0.05$ en la prueba de rangos con signo de Wilcoxon para muestras relacionadas, tanto para lesiones de tipo masa como para las calcificaciones, se concluyó que el modelo predictivo incrementa significativamente la exactitud de la detección de cáncer en mamografías.

Por lo cual, en relación al objetivo general: Determinar la influencia del Modelo Predictivo en la Detección de cáncer en mamografías, aplicando el análisis deductivo en función a las conclusiones anteriores, se determinó que la influencia del modelo predictivo mejora significativamente la detección de cáncer en mamografías.

VI. RECOMENDACIONES

A continuación, se presentan las recomendaciones en función a cada dimensión:

En relación a mejorar el porcentaje de sensibilidad de la detección de cáncer en mamografías, se recomienda, como trabajo futuro, explorar nuevas combinaciones de características a extraer de las imágenes con el objeto de lograr una mayor discriminación de las lesiones. Como indica el National Cancer Institute (2024), la mamografía no es una prueba de detección perfecta. Múltiples factores, como la edad, el peso del paciente, la densidad del tejido mamario, la naturaleza de las lesiones e incluso la calidad de la imagen, la vuelven susceptible a una mala interpretación.

En relación a mejorar el porcentaje de especificidad de la detección de cáncer en mamografías, se recomienda, como trabajo futuro, la segmentación de la mama en zonas definidas y el entrenamiento de diferentes modelos predictivos que sean capaces de especializarse en el reconocimiento de estas. Aprender la ubicación y características de los nodos axilares, los lóbulos mamarios, el músculo pectoral mayor, el pezón y la piel que cubre la mama, ayudaría a marcar zonas con baja probabilidad de desarrollar cáncer de mama. La investigación realizada por Magna et al. (2016) es un buen ejemplo de esto.

En relación a mejorar el porcentaje de exactitud de la detección de cáncer en mamografías, se recomienda como trabajo futuro, explorar diferentes configuraciones del modelo predictivo o combinar el modelo con otras técnicas de aprendizaje. Realizar modificaciones al modelo puede abrir nuevas posibilidades de incrementar el porcentaje de exactitud. La investigación realizada por Daoudi et al. (2014), es un buen ejemplo de esto.

Finalmente, no se debe olvidar que el cáncer es una de las principales causas de muerte en el mundo, por tanto, investigaciones que promuevan la detección precoz, son cruciales para mejorar el pronóstico y la supervivencia de las personas afectadas por esta enfermedad.

REFERENCIAS

- Abugattas Saba, J., Manrique Hinojosa, J., & Vidaurre Rojas, T. (2015). *Mamografía como instrumento de tamizaje en cáncer de mama*. Revista Peruana de Ginecología y Obstetricia, 61(3), 311-319. ISSN 2304-5132
- Aguilar-Torres, C. R., Cisneros-Castolo, M., Stener-Lechuga, T., Pérez-Molinar, K. F., Parra-Acosta, H., Sáenz-Cabrales, I. P., & Ornelas-Flores, C. (2021). *Panorama actual del tamizaje para detección del cáncer de mama en el estado de Chihuahua, México*. Ginecología y obstetricia de México, 89(2), 91-99. <https://doi.org/10.24245/gom.v89i2.4466>
- Arce, S., Vijay, A., Yim, E., Spiguel, L. R., & Hanna, M. (2023). *Evaluation of an Artificial Intelligence System for Detection of Invasive Lobular Carcinoma on Digital Mammography*. Cureus, 15(5), e38770. <https://doi.org/10.7759/cureus.38770>
- Armesto, D. (2011). *Pruebas diagnósticas: curvas ROC*. Revista Electrónica de Biomedicina, 77-82. <https://biomed.uninet.edu/2011/n1/armesto.pdf>. ISSN: 1697-090X
- BreastCancer.Org. (2022, febrero). *Mammography: Benefits, Risks, What You Need to Know*. <https://www.breastcancer.org/screening-testing/mammograms/benefits-risks>
- Castro-Maldonado, J. J., Gómez-Macho, L. K., & Camargo-Casallas, E. (2023). *La investigación aplicada y el desarrollo experimental en el fortalecimiento de las competencias de la sociedad del siglo XXI*. Tecnura, 27(75). <https://doi.org/10.14483/22487638.19171>
- Cayzer, S., & Aickelin, U. (2002). *A recommender system based on the immune network*. Evolutionary Computation, 2002. CEC '02. Proceedings of the 2002 Congress on, 1, 807-812. <http://dx.doi.org/10.1109/CEC.2002.1007029>

- Cruz, N., & Coello, C. (2002). *El Sistema Inmune Artificial: Una Herramienta Computacional Moderna para Resolver Problemas Complejos*. Revista Digital Universitaria UNAM, 3(3). <https://ru.tic.unam.mx/handle/123456789/1119> <http://www.revista.unam.mx/vol.3/num3/art35/> ISSN: 1607-6079
- Daoudi, R., Djemal, K., & Benyettou, A. (2014). *Digital database for screening mammography classification using improved artificial immune system approaches*. In Proceedings of the International Conference on Evolutionary Computation Theory and Applications (IJCCI 2014) - ECTA. SciTePress, 244-250. ISBN 978-989-758-052-9. <https://doi.org/10.5220/0005079602440250>
- Davies, E. R., (2017). *Computer Vision Principles, Algorithms, Applications, Learning*. Academic Press. Fifth Edition. ISBN: 978-0-12-809284-2 <https://doi.org/10.1016/C2015-0-05563-0>
- Djamila, H., & Belkhodja, L. (2019). *IMCAD: Computer Aided System for Breast Masses Detection based on Immune Recognition*. International Journal of Interactive Multimedia and Artificial Intelligence, 5(5), 97-108. ISSN: 1989-1660 <https://doi.org/10.9781/ijimai.2018.12.006>
- Ekpo, E. U., Alakhras, M., & Brennan, P. (2018). *Errors in Mammography Cannot be Solved Through Technology Alone*. Asian Pacific Journal of Cancer Prevention, 19(2), 291-301. <https://doi.org/10.22034/APJCP.2018.19.2.291>
- Greensmith, J., Whitbrook, A., & Aickelin, U. (2010). *Artificial Immune Systems*. Handbook of Metaheuristics, 146. <https://doi.org/10.48550/arXiv.1006.4949>
- Hernández-Sampieri, R., & Mendoza, C. (2018). *Metodología de la Investigación: Las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill. Primera edición. 5-7. ISBN: 978-1-4562-6096-5
- Katsis, C., Gkogkou, I., Papadopoulos, C., Goletsis, Y., Boufounou, P., & Stylios, G. (2013). *Using artificial immune recognition systems in order to detect early breast cancer*. International Journal of Intelligent Systems and Applications (IJISA), 5(2), 34-40. <http://dx.doi.org/10.5815/ijisa.2013.02.04>

- Kooi, T., Litjens, G., Van Ginneken, B., Gubern-Mérida, A., Sánchez, C., Mann, R., Den Heeten, A., & Karssemeijer, N. (2017). *Large scale deep learning for computer aided detection of mammographic lesions*. *Medical Image Analysis*, 35(1), 303-312. ISSN: 1361-8415. <https://doi.org/10.1016/j.media.2016.07.007>
- López-Muñoz, E., & Salamanca-Gómez, F. (2019). *Cáncer de mama*. *Revista Médica del Instituto Mexicano del Seguro Social*, 58(Supl 1), S1–S3. <https://pubmed.ncbi.nlm.nih.gov/34695310/> <https://doi.org/10.24875/RMIMSS.M21000109>
- Magna, G., Casti, P., Jayaraman, S. V., Salmeri, M., Mencattini, A., Martinelli, E., & Natale, C. D. (2016). *Identification of mammography anomalies for breast cancer detection by an ensemble of classification models based on artificial immune system*. *Knowledge-Based Systems*, 101, 60-70. <https://doi.org/10.1016/j.knosys.2016.02.019>
- Márquez, M. (2021). *Procesamiento de Imágenes de mamografía digital para la caracterización de lesiones mamarias*. Universidad del Rosario, Bogotá D.C. <https://doi.org/10.26507/ponencia.1839>
- Martínez-Pérez, J. A., & Pérez-Martín, P. S. (2022). *La curva ROC*. *Medicina de Familia. Sociedad Española de Médicos de Atención Primaria*, 49(1), 101821. <https://doi.org/10.1016/j.semerg.2022.101821>
- Mora, J. (2022). *Modelos predictivos en salud basados en aprendizaje de máquina*. *Revista Médica Clínica Las Condes*, 15(6), 583-590. <https://doi.org/10.1016/j.rmclc.2022.11.002>
- National Cancer Institute. (2024, julio). *Breast Cancer Screening (PDQ®)–Health Professional Version*. <https://www.cancer.gov/types/breast/hp/breast-screening-pdq>
- Octavio, M. (2011) *Cáncer de mama: estrategias de prevención y vigilancia según nivel de riesgo*. *Revista Médica Clínica Las Condes*. 22(4), 436-443. [https://doi.org/10.1016/S0716-8640\(11\)70448-5](https://doi.org/10.1016/S0716-8640(11)70448-5)

- Pacilè, S., Lopez, J., Chone, P., Bertinotti, T., Grouin, J. M., & Fillard, P. (2020). *Improving Breast Cancer Detection Accuracy of Mammography with the Concurrent Use of an Artificial Intelligence Tool*. *Radiology. Artificial Intelligence*, 2(6), e190208. <https://doi.org/10.1148/ryai.2020190208>
- Pan American Health Organization. (2018). *Cancer de mama en las Américas*. Globocan. [https://www.paho.org/sites/default/files/Cancer-mama-Americas-factsheet-ES\(1\).pdf](https://www.paho.org/sites/default/files/Cancer-mama-Americas-factsheet-ES(1).pdf)
- Pérez, J. (2022). *Investigación sobre el método de predicción del cáncer de mama basado en aprendizaje automático*. *Revista de investigación de sistemas e informática*, 15(2), 5-12. <https://doi.org/10.15381/risi.v15i2.23402>
- Polat, K., Sahan, S., Kodaz, H., & Günes, S. (2005). *A new classification method for breast cancer diagnosis: Feature selection artificial immune recognition system (FS-AIRS)*. *International Conference on Natural Computation*, 830-838. https://doi.org/10.1007/11539117_117
- Pons-Rodriguez, A., Marzo-Castillejo, M., Cruz-Esteve, I., Galindo-Ortego, G., Hernández-Leal, M. J., & Rué, M. (2022). *Avances hacia el cribado personalizado del cáncer de mama: el papel de la Atención Primaria*. *Atención Primaria*, 54(5), 102288. <https://doi.org/10.1016/j.aprim.2022.102288>
- Ramos-Galarza, C. (2021). *Editorial: Diseños de Investigación Experimental*. *Revista CienciAmérica*, 10(1), 1-7. ISSN 1390-9592 <http://dx.doi.org/10.33210/ca.v10i1.356>
- RENACE (2023). *Boletín Epidemiológico del Perú, Semana Epidemiológica del 19 al 25 de febrero de 2023*. *Centro Nacional de Epidemiología*, 32(08), 181-185. https://www.dge.gob.pe/epipublic/uploads/boletin/boletin_20238_23_092046.pdf ISSN: 1816-8655
- Responsable, M., & Victoria, A. (1999). *Lenguaje en Imágenes del Seno*. *Revista Peruana de Radiología, Sección de Radiodiagnóstico*, 3(7). http://sisbib.unmsm.edu.pe/bvrevistas/radiologia/v03_n7/leng_image.htm ISSN: 1810-8423

- Rodríguez-Ruiz, A., Krupinski, E., Mordang, J.-J., Schilling, K., Heywang-Köbrunner, S. H., Sechopoulos, I., & Mann, R. M. (2019). *Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System*. *Radiology*, 290(2), 305-314. <https://doi.org/10.1148/radiol.2018181371>
- Sánchez R., C., Maldonado J., D., Jans B., J., Domínguez C., F., Galindo A., H., Camus A., M., Oddo B., D., Medina A., L., & Acevedo C., F. (2018). *Aplicación de tres modelos pronósticos en cáncer de mama precoz*. *Revista chilena de cirugía*, 70(3), 212-217. <https://dx.doi.org/10.4067/s0718-40262018000300212>
- Sasaki, M., Tozaki, M., Rodríguez-Ruiz, A., Yotsumoto, D., Ichiki, Y., Terawaki, A., Oosako, S., Sagara, Y., & Sagara, Y. (2020). *Artificial intelligence for breast cancer detection in mammography: Experience of use of the ScreenPoint Medical Transpara system in 310 Japanese women*. *Breast Cancer (Tokyo, Japan)*, 27(4), 642-651. <https://doi.org/10.1007/s12282-020-01061-8>
- Shiraishi, J., Li, Q., Appelbaum, D., & Doi, K. (2011). *Computer-Aided Diagnosis and Artificial Intelligence in Clinical Imaging*. *Seminars in Nuclear Medicine*, 41(Image Perception in Nuclear Medicine), 449-462.
- Torregroza-Diazgranados, E. J. (2021). *Pruebas diagnósticas: Fundamentos de los estudios diagnósticos, evaluación de la validez e interpretación clínica de sus resultados*. *Revista Colombiana de Cirugía*, 36(2), 193-204. <https://doi.org/10.30944/20117582.716>
- Watkins, A., & Boggess, L. (2002). *A new classifier based on resource limited artificial immune systems*. *Evolutionary Computation*, 2002. CEC '02. Proceedings of the 2002 Congress on, 2, 1546-1551. <http://dx.doi.org/10.1109/CEC.2002.1004472>
- World Health Organization. (2015). *WHO Position Paper on Mammography Screening*. <https://paho.org/hq/dmdocuments/2015/WHO-ENG-Mammography-Factsheet.pdf>
- World Health Organization. (2022, febrero). *Cancer*. <https://www.who.int/news-room/fact-sheets/detail/cancer>

World Health Organization. (2024, julio). *Breast Cancer*. <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>

ANEXOS

Anexo 1. Tabla de Operacionalización de Variables

Variable independiente: Modelo predictivo

Variable dependiente: Detección de Cáncer en mamografías

Dimensiones de la variable dependiente: Sensibilidad, Especificidad y Exactitud.

Indicador	Definición	Instrumento	Escala de medición	Unidad	Fórmula
Porcentaje de Sensibilidad	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para demostrar que alguien realmente padece una enfermedad, es decir, que obtenga un resultado positivo estando enfermo.	Ficha de Registro	Razón	Porcentaje	$VP / (VP + FN) * 100,$ siendo: VP = Número de regiones detectadas como lesiones que SÍ son lesiones, y FN = Número de regiones detectadas como sanas pero que SÍ son lesiones.
Porcentaje de Especificidad	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para demostrar que alguien realmente no padece una enfermedad, es decir, que obtenga un resultado negativo estando sano.	Ficha de Registro	Razón	Porcentaje	$VN / (VN + FP) * 100,$ siendo: VN = Número de regiones detectadas como sanas que NO son lesiones, y FP = Número de regiones detectadas como lesiones pero que NO son lesiones.
Porcentaje de Exactitud	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para reconocer a los enfermos como enfermos y a los sanos como sanos frente al total de los individuos sometidos a evaluación.	Ficha de Registro	Razón	Porcentaje	$(VP + VN) / (VP + FN + FP + VN) * 100$

Anexo 2. Instrumentos de recolección de datos

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
MASAS (1-20 de 56)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_105_143_1_RCC.tif	357	0	5167	3444	100.00	40.00	42.38
img_105_143_1_RO.tif	393	1	6175	3693	99.75	37.42	39.82
img_108_146_1_RCC.tif	31	0	3880	1072	100.00	21.65	22.14
img_108_146_1_RO.tif	20	0	4800	2227	100.00	31.69	31.89
img_12_20_1_RO.tif	655	89	1934	1600	88.04	45.27	52.71
img_134_180_1_RCC.tif	406	38	1939	1346	91.44	40.97	46.98
img_139_188_1_LO.tif	566	12	4062	4117	97.92	50.34	53.48
img_174_233_1_LCC.tif	41	18	1731	2467	69.49	58.77	58.92
img_177_237_1_LO.tif	661	26	1900	6765	96.22	78.07	79.41
img_186_251_1_LCC.tif	38	0	3728	949	100.00	20.29	20.93
img_18_27_1_LCC.tif	41	5	5556	3060	89.13	35.52	35.80
img_18_28_1_LO.tif	7	0	3508	1377	100.00	28.19	28.29
img_194_262_1_LCC.tif	10	0	2445	2628	100.00	51.80	51.90
img_194_262_1_LO.tif	10	0	2614	3168	100.00	54.79	54.87
img_196_264_1_RCC.tif	22	0	1161	1292	100.00	52.67	53.09
img_206_455_1_LCC.tif	1	8	257	3953	11.11	93.90	93.72
img_239_320_1_LO.tif	2	3	2569	3071	40.00	54.45	54.44
img_241_449_1_RCC.tif	29	0	8242	1801	100.00	17.93	18.17
img_261_346_1_LCC.tif	276	39	1365	2101	87.62	60.62	62.87
img_273_360_1_LCC.tif	29	0	2134	1480	100.00	40.95	41.42

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
MASAS (21-40 de 56)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_287_379_1_RCC.tif	63	0	1446	1069	100.00	42.51	43.91
img_287_379_1_RO.tif	54	0	2334	2147	100.00	47.91	48.53
img_306_405_1_RO.tif	169	15	4284	2862	91.85	40.05	41.35
img_323_430_1_RCC.tif	7	6	2093	2012	53.85	49.01	49.03
img_323_430_1_RO.tif	8	4	2879	2647	66.67	47.90	47.94
img_353_487_1_RCC.tif	95	1	4123	2525	98.96	37.98	38.85
img_3_4_1_LO.tif	120	0	3683	1741	100.00	32.10	33.57
img_408_567_1_RCC.tif	146	0	1587	1124	100.00	41.46	44.45
img_429_596_1_RCC.tif	193	4	3606	2721	97.97	43.01	44.67
img_448_622_1_RCC.tif	49	104	751	3765	32.03	83.37	81.69
img_448_622_1_RO.tif	127	301	1018	4896	29.67	82.79	79.20
img_449_625_1_RCC.tif	67	3	1341	1916	95.71	58.83	59.60
img_449_625_1_RO.tif	67	14	2122	2336	82.72	52.40	52.94
img_459_645_1_LO.tif	5	44	264	3949	10.20	93.73	92.77
img_46_65_1_LO.tif	8	1	3037	1750	88.89	36.56	36.66
img_476_672_1_RO.tif	55	39	1095	7570	58.51	87.36	87.05
img_477_674_1_RO.tif	0	27	1304	7035	0.00	84.36	84.09
img_488_691_1_LO.tif	27	7	1038	4217	79.41	80.25	80.24
img_491_696_1_LO.tif	72	28	2402	6031	72.00	71.52	71.52
img_4_8_1_LCC.tif	23	0	4001	2131	100.00	34.75	35.00

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (1-20 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_100_135_1_RCC.tif	116	0	5958	2160	100.00	26.61	27.64
img_105_143_1_RCC.tif	357	0	3734	4877	100.00	56.64	58.36
img_105_143_1_RO.tif	392	2	3863	6005	99.49	60.85	62.34
img_106_144_1_RCC.tif	2	0	3243	526	100.00	13.96	14.00
img_106_144_1_RO.tif	4	0	3849	920	100.00	19.29	19.36
img_114_152_1_LO.tif	14	0	3751	1185	100.00	24.01	24.22
img_115_154_1_LCC.tif	5	0	2384	1368	100.00	36.46	36.55
img_115_155_1_LO.tif	6	0	4553	1185	100.00	20.65	20.74
img_117_158_1_LCC.tif	3	0	3522	156	100.00	4.24	4.32
img_117_158_1_LO.tif	2	0	3609	396	100.00	9.89	9.93
img_118_159_1_LCC.tif	4	0	3504	292	100.00	7.69	7.79
img_118_159_1_RCC.tif	9	0	1868	601	100.00	24.34	24.62
img_118_159_1_RO.tif	6	0	2174	376	100.00	14.75	14.95
img_123_164_1_RO.tif	222	1	2934	782	99.55	21.04	25.49
img_124_165_1_LCC.tif	1	0	1205	477	100.00	28.36	28.40
img_127_170_1_RCC.tif	4	0	4803	2001	100.00	29.41	29.45
img_127_170_1_RO.tif	7	0	6329	2474	100.00	28.10	28.16
img_138_186_1_RCC.tif	8	0	2049	1637	100.00	44.41	44.53
img_138_187_1_RCC.tif	9	0	1252	2700	100.00	68.32	68.39
img_140_190_1_RO.tif	2	0	6312	1377	100.00	17.91	17.93

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (21-40 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_145_197_1_LCC.tif	2	0	2460	5170	100.00	67.76	67.77
img_145_198_1_RO.tif	4	0	3978	3059	100.00	43.47	43.50
img_150_203_1_RCC.tif	4	0	4902	1081	100.00	18.07	18.12
img_150_203_1_RO.tif	6	0	6034	1117	100.00	15.62	15.69
img_153_206_1_LCC.tif	2	0	1390	1614	100.00	53.73	53.76
img_153_206_1_LO.tif	1	0	3078	1209	100.00	28.20	28.22
img_153_207_1_LO.tif	2	0	2266	1057	100.00	31.81	31.85
img_155_209_1_RO.tif	4	0	2187	4156	100.00	65.52	65.54
img_161_216_1_RCC.tif	4	0	3046	1202	100.00	28.30	28.36
img_164_219_1_LCC.tif	155	0	4482	207	100.00	4.42	7.47
img_164_219_1_LO.tif	175	0	5710	313	100.00	5.20	7.87
img_165_220_1_LCC.tif	4	2	876	3031	66.67	77.58	77.56
img_169_224_1_LCC.tif	2	1	1792	5575	66.67	75.68	75.67
img_169_225_1_LO.tif	1	0	5140	5000	100.00	49.31	49.32
img_16_24_1_LO.tif	5	0	2663	1958	100.00	42.37	42.43
img_172_231_1_RO.tif	3	1	4122	4528	75.00	52.35	52.36
img_175_234_1_RCC.tif	2	0	4343	3052	100.00	41.27	41.29
img_177_237_1_LO.tif	680	7	2843	5822	98.98	67.19	69.53
img_181_242_1_LCC.tif	1	0	2158	981	100.00	31.25	31.27
img_181_242_1_LO.tif	2	0	4466	1673	100.00	27.25	27.28

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (41-60 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_183_245_1_RCC.tif	4	0	2530	2362	100.00	48.28	48.33
img_183_246_1_RCC.tif	2	0	1171	3197	100.00	73.19	73.20
img_183_246_1_RO.tif	1	0	4114	3528	100.00	46.17	46.17
img_188_253_1_RCC.tif	4	0	2166	186	100.00	7.91	8.07
img_189_254_1_RO.tif	6	0	3469	2974	100.00	46.16	46.21
img_18_28_1_LCC.tif	8	0	1091	937	100.00	46.20	46.42
img_191_258_1_RCC.tif	552	35	2056	829	94.04	28.74	39.78
img_194_262_1_RCC.tif	6	0	3556	1851	100.00	34.23	34.31
img_194_262_1_RO.tif	4	0	2688	1763	100.00	39.61	39.66
img_196_264_1_RCC.tif	14	0	1322	1139	100.00	46.28	46.59
img_196_264_1_RO.tif	19	0	2403	1596	100.00	39.91	40.19
img_199_268_1_RCC.tif	4	0	2668	1234	100.00	31.63	31.70
img_199_269_1_RCC.tif	4	0	2602	1311	100.00	33.50	33.57
img_199_269_1_RO.tif	3	0	3477	1561	100.00	30.99	31.03
img_205_275_1_LO.tif	224	0	4118	3600	100.00	46.64	48.15
img_210_282_1_LCC.tif	6	0	5149	1593	100.00	23.63	23.70
img_210_283_1_LCC.tif	5	0	4898	2580	100.00	34.50	34.55
img_210_283_1_LO.tif	4	0	5106	2979	100.00	36.85	36.88
img_212_286_1_LCC.tif	8	0	4823	3729	100.00	43.60	43.66
img_212_286_1_RO.tif	4	0	5560	3849	100.00	40.91	40.93

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (61-80 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_218_294_1_LCC.tif	7	0	3517	538	100.00	13.27	13.42
img_229_308_1_LCC.tif	134	0	3279	1298	100.00	28.36	30.40
img_246_328_1_LCC.tif	55	122	297	8332	31.07	96.56	95.24
img_249_332_1_RCC.tif	249	16	2430	3447	93.96	58.65	60.18
img_259_344_1_LCC.tif	142	0	1318	2340	100.00	63.97	65.32
img_259_344_1_LO.tif	179	5	1957	3116	97.28	61.42	62.68
img_262_450_1_RCC.tif	4	0	1104	1193	100.00	51.94	52.02
img_262_450_1_RO.tif	5	0	1707	2025	100.00	54.26	54.32
img_26_37_1_LCC.tif	5	0	1379	1772	100.00	56.24	56.31
img_273_360_1_LCC.tif	29	0	2283	1331	100.00	36.83	37.33
img_275_363_1_LCC.tif	4	0	3084	667	100.00	17.78	17.87
img_276_365_1_RO.tif	6	0	4188	1543	100.00	26.92	27.00
img_293_386_1_LCC.tif	8	0	1995	1150	100.00	36.57	36.73
img_293_386_1_LO.tif	9	0	2892	1776	100.00	38.05	38.17
img_297_392_1_LCC.tif	3	1	1427	3525	75.00	71.18	71.19
img_297_392_1_LO.tif	4	0	2659	3323	100.00	55.55	55.58
img_2_3_1_LCC.tif	5	0	4681	4176	100.00	47.15	47.18
img_2_3_1_RCC.tif	5	0	6951	2082	100.00	23.05	23.09
img_300_396_1_RCC.tif	2	0	2031	2992	100.00	59.57	59.58
img_301_398_1_RCC.tif	335	0	3491	1363	100.00	28.08	32.72

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (81-100 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_301_398_1_RO.tif	395	0	4787	2242	100.00	31.90	35.52
img_305_403_1_LO.tif	343	10	1048	2441	97.17	69.96	72.46
img_306_405_1_RCC.tif	186	40	3437	2335	82.30	40.45	42.03
img_306_405_1_RO.tif	162	22	4078	3068	88.04	42.93	44.07
img_30_41_1_LCC.tif	2	0	1253	1007	100.00	44.56	44.61
img_30_41_1_LO.tif	2	0	2007	1855	100.00	48.03	48.06
img_315_416_1_RO.tif	4	0	3828	2075	100.00	35.15	35.20
img_319_422_1_LCC.tif	12	0	3896	2764	100.00	41.50	41.61
img_319_423_1_RCC.tif	48	0	2735	2353	100.00	46.25	46.75
img_31_43_1_RO.tif	4	0	5910	1184	100.00	16.69	16.74
img_323_430_1_RCC.tif	3	0	2929	1186	100.00	28.82	28.87
img_327_437_1_RCC.tif	9	0	2875	3777	100.00	56.78	56.84
img_328_438_1_LCC.tif	15	0	3740	4983	100.00	57.13	57.20
img_328_438_1_LO.tif	14	0	4594	4502	100.00	49.49	49.57
img_332_442_1_LCC.tif	4	0	645	2947	100.00	82.04	82.06
img_33_46_1_LO.tif	2	0	6594	1943	100.00	22.76	22.78
img_341_469_1_LCC.tif	2	0	1788	838	100.00	31.91	31.96
img_354_489_1_LO.tif	4	0	5106	1555	100.00	23.35	23.39
img_360_496_1_RO.tif	4	0	5272	2125	100.00	28.73	28.77
img_364_501_1_RO.tif	2	0	3566	3595	100.00	50.20	50.22

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (101-120 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_365_503_1_RCC.tif	7	0	3268	1646	100.00	33.50	33.59
img_365_503_1_RO.tif	21	0	3573	1989	100.00	35.76	36.00
img_374_518_1_LCC.tif	6	0	3236	2359	100.00	42.16	42.23
img_375_519_1_LCC.tif	11	0	1956	2704	100.00	58.03	58.13
img_37_51_1_RO.tif	1	0	2038	2218	100.00	52.12	52.13
img_385_534_1_LCC.tif	4	0	1220	4224	100.00	77.59	77.61
img_385_534_1_LO.tif	3	0	3548	4505	100.00	55.94	55.96
img_386_535_1_RCC.tif	4	0	1527	2714	100.00	63.99	64.03
img_386_535_1_RO.tif	4	0	2044	4338	100.00	67.97	67.99
img_38_53_1_RCC.tif	14	0	2034	876	100.00	30.10	30.44
img_38_54_1_RCC.tif	14	0	2281	546	100.00	19.31	19.71
img_38_54_1_RO.tif	12	0	3573	752	100.00	17.39	17.62
img_390_539_1_LO.tif	31	3	2803	1378	91.18	32.96	33.43
img_392_543_1_LCC.tif	5	0	539	1726	100.00	76.20	76.26
img_392_543_1_LO.tif	4	0	1268	2117	100.00	62.54	62.59
img_392_543_1_RCC.tif	2	0	1249	1998	100.00	61.53	61.56
img_395_548_1_LCC.tif	4	0	6504	2996	100.00	31.54	31.57
img_395_548_1_LO.tif	4	0	7542	3768	100.00	33.32	33.34
img_396_549_1_RCC.tif	95	1	2510	1189	98.96	32.14	33.83
img_396_549_1_RO.tif	67	2	2184	2421	97.10	52.57	53.23

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (121-140 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_399_553_1_LCC.tif	2	3	395	3932	40.00	90.87	90.81
img_399_553_1_LO.tif	1	3	2389	4477	25.00	65.21	65.18
img_401_555_1_LO.tif	25	12	2995	2531	67.57	45.80	45.95
img_402_556_1_LO.tif	1	0	5658	3992	100.00	41.37	41.37
img_406_563_1_RCC.tif	4	0	1740	404	100.00	18.84	18.99
img_412_573_1_LCC.tif	6	0	1221	1845	100.00	60.18	60.25
img_412_573_1_LO.tif	5	0	3296	1488	100.00	31.10	31.18
img_418_581_1_LCC.tif	4	0	4692	4703	100.00	50.06	50.08
img_418_581_1_LO.tif	6	0	5944	6126	100.00	50.75	50.78
img_418_581_1_RCC.tif	27	1	4223	4206	96.43	49.90	50.05
img_419_582_1_RO.tif	303	29	3590	7773	91.27	68.41	69.06
img_433_600_1_RCC.tif	32	1	3320	1865	96.97	35.97	36.36
img_433_600_1_RO.tif	32	0	4104	2325	100.00	36.16	36.48
img_438_606_1_RO.tif	12	0	4272	3994	100.00	48.32	48.39
img_447_619_1_LCC.tif	2	1	350	5083	66.67	93.56	93.54
img_448_622_1_RCC.tif	116	37	1723	2793	75.82	61.85	62.31
img_452_631_1_RCC.tif	5	3	286	4268	62.50	93.72	93.67
img_456_638_1_LCC.tif	6	0	1882	1883	100.00	50.01	50.09
img_469_663_1_RCC.tif	7	0	1561	1174	100.00	42.93	43.07
img_475_671_1_LCC.tif	4	0	1734	1832	100.00	51.37	51.43

FICHA DE REGISTRO							
Tipo de Prueba:	PRE (X) / POST () TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (141-160 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_475_671_1_LO.tif	4	0	2755	2336	100.00	45.89	45.93
img_475_671_1_RCC.tif	4	0	1669	1940	100.00	53.76	53.81
img_477_674_1_LO.tif	4	0	2304	6457	100.00	73.70	73.71
img_481_678_1_RCC.tif	2	0	705	5411	100.00	88.47	88.48
img_481_678_1_RO.tif	1	0	864	6718	100.00	88.61	88.61
img_483_681_1_RCC.tif	16	0	2230	3152	100.00	58.57	58.69
img_487_688_1_LCC.tif	5	1	30	5051	83.33	99.41	99.39
img_487_688_1_LO.tif	5	0	947	5300	100.00	84.84	84.85
img_488_691_1_LO.tif	34	0	2215	3040	100.00	57.85	58.12
img_49_69_1_LCC.tif	3	0	2195	2639	100.00	54.59	54.62
img_4_7_1_LCC.tif	3	0	2929	2254	100.00	43.49	43.52
img_4_7_1_LO.tif	4	0	3427	4127	100.00	54.63	54.66
img_504_718_1_LCC.tif	5	1	7	4605	83.33	99.85	99.83
img_504_719_1_LCC.tif	1	1	8	4582	50.00	99.83	99.80
img_507_725_1_RCC.tif	64	13	260	9282	83.12	97.28	97.16
img_507_725_1_RO.tif	78	21	350	10082	78.79	96.65	96.48
img_508_727_1_RCC.tif	5	0	1360	2513	100.00	64.89	64.93
img_508_727_1_RO.tif	4	0	2088	3280	100.00	61.10	61.13
img_51_73_1_LCC.tif	4	0	1756	2774	100.00	61.24	61.27
img_51_73_1_LO.tif	6	0	2078	3079	100.00	59.71	59.75

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
MASAS (1-20 de 56)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_105_143_1_RCC.tif	325	32	3376	5235	91.04	60.79	62.00
img_105_143_1_RO.tif	361	33	3336	6532	91.62	66.19	67.17
img_108_146_1_RCC.tif	18	13	1764	3188	58.07	64.38	64.34
img_108_146_1_RO.tif	12	8	3120	3907	60.00	55.60	55.61
img_12_20_1_RO.tif	575	169	1466	2068	77.29	58.52	61.78
img_134_180_1_RCC.tif	322	122	1623	1662	72.52	50.59	53.21
img_139_188_1_LO.tif	456	122	3186	4993	78.89	61.05	62.23
img_174_233_1_LCC.tif	21	38	1049	3149	35.59	75.01	74.47
img_177_237_1_LO.tif	468	219	2491	6174	68.12	71.25	71.02
img_186_251_1_LCC.tif	34	4	2191	2486	89.47	53.15	53.45
img_18_27_1_LCC.tif	24	22	1246	7370	52.17	85.54	85.36
img_18_28_1_LO.tif	7	0	2478	2407	100.00	49.27	49.35
img_194_262_1_LCC.tif	6	4	1763	3310	60.00	65.25	65.24
img_194_262_1_LO.tif	6	4	2060	3722	60.00	64.37	64.37
img_196_264_1_RCC.tif	18	4	972	1481	81.82	60.38	60.57
img_206_455_1_LCC.tif	2	7	1783	2427	22.22	57.65	57.57
img_239_320_1_LO.tif	1	4	2578	3062	20.00	54.29	54.26
img_241_449_1_RCC.tif	24	5	3584	6459	82.76	64.31	64.37
img_261_346_1_LCC.tif	205	110	1079	2387	65.08	68.87	68.55
img_273_360_1_LCC.tif	22	7	602	3012	75.86	83.34	83.28

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
MASAS (21-40 de 56)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_287_379_1_RCC.tif	55	8	1108	1407	87.30	55.94	56.71
img_287_379_1_RO.tif	45	9	1969	2512	83.33	56.06	56.38
img_306_405_1_RO.tif	135	49	3466	3680	73.37	51.50	52.05
img_323_430_1_RCC.tif	7	6	1649	2456	53.85	59.83	59.81
img_323_430_1_RO.tif	7	5	2337	3189	58.33	57.71	57.71
img_353_487_1_RCC.tif	56	40	2715	3933	58.33	59.16	59.15
img_3_4_1_LO.tif	96	24	2762	2662	80.00	49.08	49.75
img_408_567_1_RCC.tif	117	29	734	1977	80.14	72.93	73.29
img_429_596_1_RCC.tif	159	38	2484	3843	80.71	60.74	61.34
img_448_622_1_RCC.tif	86	67	1587	2929	56.21	64.86	64.58
img_448_622_1_RO.tif	241	187	1918	3996	56.31	67.57	66.81
img_449_625_1_RCC.tif	45	25	1432	1825	64.29	56.03	56.21
img_449_625_1_RO.tif	48	33	2190	2268	59.26	50.88	51.02
img_459_645_1_LO.tif	30	19	1231	2982	61.22	70.78	70.67
img_46_65_1_LO.tif	8	1	1628	3159	88.89	65.99	66.03
img_476_672_1_RO.tif	66	28	3642	5023	70.21	57.97	58.10
img_477_674_1_RO.tif	15	12	3387	4952	55.56	59.38	59.37
img_488_691_1_LO.tif	20	14	1897	3358	58.82	63.90	63.87
img_491_696_1_LO.tif	59	41	3140	5293	59.00	62.77	62.72
img_4_8_1_LCC.tif	19	4	2875	3257	82.61	53.12	53.23

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (1-20 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_100_135_1_RCC.tif	92	24	3397	4721	79.31	58.16	58.45
img_105_143_1_RCC.tif	270	87	2493	6118	75.63	71.05	71.23
img_105_143_1_RO.tif	304	90	2468	7400	77.16	74.99	75.07
img_106_144_1_RCC.tif	2	0	1730	2039	100.00	54.10	54.12
img_106_144_1_RO.tif	4	0	2096	2673	100.00	56.05	56.09
img_114_152_1_LO.tif	14	0	2063	2873	100.00	58.21	58.32
img_115_154_1_LCC.tif	4	1	1412	2340	80.00	62.37	62.39
img_115_155_1_LO.tif	5	1	2705	3033	83.33	52.86	52.89
img_117_158_1_LCC.tif	2	1	1898	1780	66.67	48.40	48.41
img_117_158_1_LO.tif	2	0	1921	2084	100.00	52.04	52.06
img_118_159_1_LCC.tif	4	0	1953	1843	100.00	48.55	48.61
img_118_159_1_RCC.tif	8	1	1120	1349	88.89	54.64	54.76
img_118_159_1_RO.tif	6	0	1335	1215	100.00	47.65	47.77
img_123_164_1_RO.tif	173	50	1766	1950	77.58	52.48	53.90
img_124_165_1_LCC.tif	1	0	743	939	100.00	55.83	55.85
img_127_170_1_RCC.tif	4	0	3192	3612	100.00	53.09	53.11
img_127_170_1_RO.tif	7	0	3620	5183	100.00	58.88	58.91
img_138_186_1_RCC.tif	8	0	1318	2368	100.00	64.24	64.32
img_138_187_1_RCC.tif	9	0	1032	2920	100.00	73.89	73.95
img_140_190_1_RO.tif	2	0	3402	4287	100.00	55.76	55.77

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (21-40 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_145_197_1_LCC.tif	2	0	2072	5558	100.00	72.84	72.85
img_145_198_1_RO.tif	4	0	2593	4444	100.00	63.15	63.17
img_150_203_1_RCC.tif	4	0	2201	3782	100.00	63.21	63.24
img_150_203_1_RO.tif	6	0	2514	4637	100.00	64.84	64.87
img_153_206_1_LCC.tif	2	0	1395	1609	100.00	53.56	53.59
img_153_206_1_LO.tif	1	0	2054	2233	100.00	52.09	52.10
img_153_207_1_LO.tif	2	0	1715	1608	100.00	48.39	48.42
img_155_209_1_RO.tif	4	0	1356	4987	100.00	78.62	78.64
img_161_216_1_RCC.tif	3	1	1547	2701	75.00	63.58	63.59
img_164_219_1_LCC.tif	92	63	1991	2698	59.36	57.54	57.60
img_164_219_1_LO.tif	129	46	3190	2833	73.71	47.04	47.79
img_165_220_1_LCC.tif	4	2	847	3060	66.67	78.32	78.30
img_169_224_1_LCC.tif	2	1	1349	6018	66.67	81.69	81.68
img_169_225_1_LO.tif	1	0	3084	7056	100.00	69.59	69.59
img_16_24_1_LO.tif	4	1	1345	3276	80.00	70.89	70.90
img_172_231_1_RO.tif	3	1	2551	6099	75.00	70.51	70.51
img_175_234_1_RCC.tif	2	0	2755	4640	100.00	62.75	62.76
img_177_237_1_LO.tif	350	337	2145	6520	50.95	75.25	73.46
img_181_242_1_LCC.tif	1	0	1466	1673	100.00	53.30	53.31
img_181_242_1_LO.tif	2	0	2465	3674	100.00	59.85	59.86

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (41-60 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_183_245_1_RCC.tif	3	1	1114	3778	75.00	77.23	77.23
img_183_246_1_RCC.tif	2	0	1622	2746	100.00	62.87	62.88
img_183_246_1_RO.tif	1	0	2968	4674	100.00	61.16	61.17
img_188_253_1_RCC.tif	4	0	855	1497	100.00	63.65	63.71
img_189_254_1_RO.tif	5	1	2187	4256	83.33	66.06	66.07
img_18_28_1_LCC.tif	7	1	698	1330	87.50	65.58	65.67
img_191_258_1_RCC.tif	268	319	1034	1851	45.66	64.16	61.03
img_194_262_1_RCC.tif	6	0	1951	3456	100.00	63.92	63.96
img_194_262_1_RO.tif	4	0	1471	2980	100.00	66.95	66.98
img_196_264_1_RCC.tif	9	5	813	1648	64.29	66.97	66.95
img_196_264_1_RO.tif	13	6	1458	2541	68.42	63.54	63.56
img_199_268_1_RCC.tif	4	0	1751	2151	100.00	55.13	55.17
img_199_269_1_RCC.tif	2	2	1770	2143	50.00	54.77	54.76
img_199_269_1_RO.tif	3	0	2076	2962	100.00	58.79	58.82
img_205_275_1_LO.tif	177	47	2402	5316	79.02	68.88	69.16
img_210_282_1_LCC.tif	6	0	2296	4446	100.00	65.95	65.98
img_210_283_1_LCC.tif	4	1	2279	5199	80.00	69.52	69.53
img_210_283_1_LO.tif	4	0	2487	5598	100.00	69.24	69.26
img_212_286_1_LCC.tif	7	1	3387	5165	87.50	60.40	60.42
img_212_286_1_RO.tif	4	0	3463	5946	100.00	63.20	63.21

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (61-80 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_218_294_1_LCC.tif	6	1	1804	2251	85.71	55.51	55.56
img_229_308_1_LCC.tif	78	56	1646	2931	58.21	64.04	63.87
img_246_328_1_LCC.tif	49	128	1696	6933	27.68	80.35	79.29
img_249_332_1_RCC.tif	126	139	1618	4259	47.55	72.47	71.39
img_259_344_1_LCC.tif	89	53	1264	2394	62.68	65.45	65.34
img_259_344_1_LO.tif	114	70	1601	3472	61.96	68.44	68.21
img_262_450_1_RCC.tif	4	0	744	1553	100.00	67.61	67.67
img_262_450_1_RO.tif	4	1	1096	2636	80.00	70.63	70.65
img_26_37_1_LCC.tif	2	3	961	2190	40.00	69.50	69.46
img_273_360_1_LCC.tif	16	13	1020	2594	55.17	71.78	71.64
img_275_363_1_LCC.tif	2	2	1561	2190	50.00	58.38	58.38
img_276_365_1_RO.tif	6	0	2372	3359	100.00	58.61	58.65
img_293_386_1_LCC.tif	8	0	1158	1987	100.00	63.18	63.27
img_293_386_1_LO.tif	8	1	1703	2965	88.89	63.52	63.57
img_297_392_1_LCC.tif	2	2	807	4145	50.00	83.70	83.68
img_297_392_1_LO.tif	4	0	1732	4250	100.00	71.05	71.07
img_2_3_1_LCC.tif	5	0	2210	6647	100.00	75.05	75.06
img_2_3_1_RCC.tif	4	1	3849	5184	80.00	57.39	57.40
img_300_396_1_RCC.tif	2	0	1912	3111	100.00	61.94	61.95
img_301_398_1_RCC.tif	183	152	1700	3154	54.63	64.98	64.31

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (81-100 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_301_398_1_RO.tif	224	171	2391	4638	56.71	65.98	65.49
img_305_403_1_LO.tif	247	106	812	2677	69.97	76.73	76.11
img_306_405_1_RCC.tif	96	130	1861	3911	42.48	67.76	66.81
img_306_405_1_RO.tif	95	89	2150	4996	51.63	69.91	69.45
img_30_41_1_LCC.tif	2	0	597	1663	100.00	73.58	73.61
img_30_41_1_LO.tif	2	0	1370	2492	100.00	64.53	64.55
img_315_416_1_RO.tif	3	1	2001	3902	75.00	66.10	66.11
img_319_422_1_LCC.tif	8	4	1285	5375	66.67	80.71	80.68
img_319_423_1_RCC.tif	30	18	1649	3439	62.50	67.59	67.54
img_31_43_1_RO.tif	3	1	2565	4529	75.00	63.84	63.85
img_323_430_1_RCC.tif	3	0	1763	2352	100.00	57.16	57.19
img_327_437_1_RCC.tif	7	2	1925	4727	77.78	71.06	71.07
img_328_438_1_LCC.tif	15	0	3377	5346	100.00	61.29	61.35
img_328_438_1_LO.tif	12	2	3403	5693	85.71	62.59	62.62
img_332_442_1_LCC.tif	1	3	942	2650	25.00	73.78	73.72
img_33_46_1_LO.tif	2	0	3225	5312	100.00	62.22	62.23
img_341_469_1_LCC.tif	1	1	951	1675	50.00	63.79	63.78
img_354_489_1_LO.tif	3	1	2862	3799	75.00	57.03	57.04
img_360_496_1_RO.tif	4	0	3131	4266	100.00	57.67	57.70
img_364_501_1_RO.tif	1	1	1939	5222	50.00	72.92	72.92

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (101-120 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_365_503_1_RCC.tif	7	0	1817	3097	100.00	63.02	63.08
img_365_503_1_RO.tif	14	7	2123	3439	66.67	61.83	61.85
img_374_518_1_LCC.tif	5	1	1959	3636	83.33	64.99	65.01
img_375_519_1_LCC.tif	8	3	1284	3376	72.73	72.45	72.45
img_37_51_1_RO.tif	1	0	1296	2960	100.00	69.55	69.56
img_385_534_1_LCC.tif	4	0	1225	4219	100.00	77.50	77.52
img_385_534_1_LO.tif	1	2	2013	6040	33.33	75.00	74.99
img_386_535_1_RCC.tif	2	2	1196	3045	50.00	71.80	71.78
img_386_535_1_RO.tif	3	1	1827	4555	75.00	71.37	71.38
img_38_53_1_RCC.tif	9	5	1321	1589	64.29	54.61	54.65
img_38_54_1_RCC.tif	13	1	1314	1513	92.86	53.52	53.71
img_38_54_1_RO.tif	11	1	2017	2308	91.67	53.36	53.47
img_390_539_1_LO.tif	29	5	1398	2783	85.29	66.56	66.71
img_392_543_1_LCC.tif	5	0	464	1801	100.00	79.51	79.56
img_392_543_1_LO.tif	4	0	929	2456	100.00	72.56	72.59
img_392_543_1_RCC.tif	2	0	935	2312	100.00	71.20	71.22
img_395_548_1_LCC.tif	4	0	2578	6922	100.00	72.86	72.88
img_395_548_1_LO.tif	4	0	3041	8269	100.00	73.11	73.12
img_396_549_1_RCC.tif	61	35	1116	2583	63.54	69.83	69.67
img_396_549_1_RO.tif	47	22	1452	3153	68.12	68.47	68.46

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (121-140 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_399_553_1_LCC.tif	2	3	883	3444	40.00	79.59	79.55
img_399_553_1_LO.tif	1	3	1341	5525	25.00	80.47	80.44
img_401_555_1_LO.tif	12	25	1764	3762	32.43	68.08	67.84
img_402_556_1_LO.tif	1	0	3675	5975	100.00	61.92	61.92
img_406_563_1_RCC.tif	4	0	770	1374	100.00	64.09	64.15
img_412_573_1_LCC.tif	5	1	1377	1689	83.33	55.09	55.14
img_412_573_1_LO.tif	5	0	2005	2779	100.00	58.09	58.13
img_418_581_1_LCC.tif	4	0	2716	6679	100.00	71.09	71.10
img_418_581_1_LO.tif	5	1	3356	8714	83.33	72.20	72.20
img_418_581_1_RCC.tif	19	9	2214	6215	67.86	73.73	73.71
img_419_582_1_RO.tif	166	166	2766	8597	50.00	75.66	74.93
img_433_600_1_RCC.tif	31	2	2499	2686	93.94	51.80	52.07
img_433_600_1_RO.tif	30	2	2632	3797	93.75	59.06	59.23
img_438_606_1_RO.tif	10	2	2547	5719	83.33	69.19	69.21
img_447_619_1_LCC.tif	3	0	1411	4022	100.00	74.03	74.04
img_448_622_1_RCC.tif	60	93	1118	3398	39.22	75.24	74.06
img_452_631_1_RCC.tif	5	3	451	4103	62.50	90.10	90.05
img_456_638_1_LCC.tif	4	2	1160	2605	66.67	69.19	69.19
img_469_663_1_RCC.tif	6	1	994	1741	85.71	63.66	63.71
img_475_671_1_LCC.tif	3	1	868	2698	75.00	75.66	75.66

FICHA DE REGISTRO							
Tipo de Prueba:	PRE () / POST (X) TEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (141-160 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_475_671_1_LO.tif	4	0	1302	3789	100.00	74.43	74.45
img_475_671_1_RCC.tif	4	0	889	2720	100.00	75.37	75.39
img_477_674_1_LO.tif	4	0	2605	6156	100.00	70.27	70.28
img_481_678_1_RCC.tif	2	0	2006	4110	100.00	67.20	67.21
img_481_678_1_RO.tif	1	0	2178	5404	100.00	71.27	71.28
img_483_681_1_RCC.tif	12	4	1164	4218	75.00	78.37	78.36
img_487_688_1_LCC.tif	3	3	191	4890	50.00	96.24	96.19
img_487_688_1_LO.tif	4	1	895	5352	80.00	85.67	85.67
img_488_691_1_LO.tif	24	10	1382	3873	70.59	73.70	73.68
img_49_69_1_LCC.tif	3	0	1267	3567	100.00	73.79	73.81
img_4_7_1_LCC.tif	3	0	1756	3427	100.00	66.12	66.14
img_4_7_1_LO.tif	4	0	2526	5028	100.00	66.56	66.58
img_504_718_1_LCC.tif	5	1	604	4008	83.33	86.90	86.90
img_504_719_1_LCC.tif	0	2	624	3966	0.00	86.41	86.37
img_507_725_1_RCC.tif	39	38	1473	8069	50.65	84.56	84.29
img_507_725_1_RO.tif	38	61	2076	8356	38.38	80.10	79.71
img_508_727_1_RCC.tif	3	2	932	2941	60.00	75.94	75.92
img_508_727_1_RO.tif	4	0	922	4446	100.00	82.82	82.84
img_51_73_1_LCC.tif	4	0	1100	3430	100.00	75.72	75.74
img_51_73_1_LO.tif	6	0	1323	3834	100.00	74.35	74.38

Anexo 3. Fichas de validación de instrumentos para la recolección de datos



UNIVERSIDAD CÉSAR VALLEJO

CARTA DE PRESENTACIÓN

Señor:

BRIGS FASABI VASQUEZ

Presente

Asunto: **VALIDACIÓN DE INSTRUMENTOS A TRAVÉS DE JUICIO DE EXPERTO.**

Me es muy grato comunicarme con usted para expresarle mis saludos y así mismo, hacer de su conocimiento que, siendo estudiante del **Programa Académico de Maestría en Ingeniería de Sistemas con mención en Tecnologías de la Información**, de la Escuela de Posgrado de la UCV, campus Lima Norte, ciclo 2024 - I, requiero validar el instrumento con el cual se recogerá la información necesaria para poder desarrollar mi investigación y con la que sustentaré mis competencias investigativas en la Experiencia curricular de **Diseño y Desarrollo del Trabajo de Investigación**.

El nombre de mi variable dependiente es: **Detección de cáncer en mamografías** y siendo imprescindible contar con la aprobación de profesionales especializados para poder aplicar el instrumento en mención, he considerado conveniente recurrir a usted, ante su connotada experiencia en temas relacionados a la Ingeniería de Sistemas, Tecnologías de la Información e Investigación Científica.

El expediente de validación, que le hago llegar contiene:

- Matriz de operacionalización de la variable.
- Ficha de Registro para las dimensiones de la variable.
- Ficha de validación de contenido para el instrumento.
- Matriz de validación de indicadores para la variable.
- Ficha de validación de juicio de experto.

Expresándole mis sentimientos de respeto y consideración me despido de usted, no sin antes agradecerle por la atención que dispense a la presente.

Atentamente.

.....
LUIS JOSÉ MELÉNDEZ RUIZ

D.N.I. 44180515



UNIVERSIDAD CÉSAR VALLEJO

MATRIZ DE OPERACIONALIZACIÓN DE LA VARIABLE DEPENDIENTE: DETECCIÓN DE CÁNCER EN MAMOGRAFÍAS

Indicador	Definición	Instrumento	Escala de medición	Unidad	Fórmula
Porcentaje de Sensibilidad	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para demostrar que alguien realmente padece una enfermedad, es decir, que obtenga un resultado positivo estando enfermo.	Ficha de Registro	Razón	Porcentaje	$VP / (VP + FN) * 100$, siendo: VP = Número de regiones detectadas como lesiones que Sí son lesiones, y FN = Número de regiones detectadas como sanas pero que Sí son lesiones.
Porcentaje de Especificidad	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para demostrar que alguien realmente no padece una enfermedad, es decir, que obtenga un resultado negativo estando sano.	Ficha de Registro	Razón	Porcentaje	$VN / (VN + FP) * 100$, siendo: VN = Número de regiones detectadas como sanas que NO son lesiones, y FP = Número de regiones detectadas como lesiones pero que NO son lesiones.
Porcentaje de Exactitud	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para reconocer a los enfermos como enfermos y a los sanos como sanos frente al total de los individuos sometidos a evaluación.	Ficha de Registro	Razón	Porcentaje	$(VP + VN) / (VP + FN + FP + VN) * 100$



UNIVERSIDAD CÉSAR VALLEJO

FICHA DE VALIDACIÓN DE CONTENIDO PARA EL INSTRUMENTO

INSTRUCCIÓN: A continuación, se le hace llegar el instrumento de recolección de datos del fichaje que permitirá recoger la información en la investigación que lleva por título: **Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024.**

Por lo que se le solicita que tenga a bien evaluar el instrumento haciendo, de ser caso, las sugerencias para realizar las correcciones pertinentes. Los criterios de validación de contenido son:

Criterios	Detalle	Calificación
Suficiencia	El indicador pertenece a la dimensión y basta para obtener la medición de ésta	1: de acuerdo 0: en desacuerdo
Claridad	El indicador se comprende fácilmente, es decir, su sintáctica y semántica son adecuadas.	1: de acuerdo 0: en desacuerdo
Coherencia	El indicador tiene relación lógica con la dimensión que está midiendo.	1: de acuerdo 0: en desacuerdo
Relevancia	El indicador es esencial o importante, es decir, debe ser incluido.	1: de acuerdo 0: en desacuerdo

Nota. Criterios adaptados de la propuesta de Escobar y Cuervo (2008).

f

f



UNIVERSIDAD CÉSAR VALLEJO

MATRIZ DE VALIDACIÓN DE INDICADORES PARA LA VARIABLE:

DETECCIÓN DE CÁNCER EN MAMOGRAFÍAS

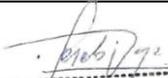
Definición de la variable: Proceso de examinar las mamas de una mujer para identificar la presencia de lesiones probablemente cancerosas a través de una radiografía, antes de que aparezcan los primeros síntomas o signos de la enfermedad.

Dimensiones	Indicadores	Suficiencia	Claridad	Coherencia	Relevancia	Observación
Sensibilidad	Porcentaje de Sensibilidad	1	1	1	1	NINGUNA
Especificidad	Porcentaje de Especificidad	1	1	1	1	NINGUNA
Exactitud	Porcentaje de Exactitud	1	1	1	1	NINGUNA



UNIVERSIDAD CÉSAR VALLEJO

FICHA DE VALIDACIÓN DE JUICIO DE EXPERTO

Nombre del instrumento	Ficha de Registro
Objetivo del instrumento	Medir las dimensiones de Sensibilidad, Especificidad y Exactitud en la Detección de Cáncer en Mamografías.
Nombres y apellidos del experto	BRIGS FASABI VASQUEZ
Documento de identidad	41221873
Años de experiencia en el área	13 AÑOS
Máximo Grado Académico	MAESTRO
Nacionalidad	PERUANA
Institución	CONTRALORIA GENERAL DE LA REPUBLICA
Cargo	JEFE DE OCI
Número telefónico	999402127
Firma	 ----- Brigs Fasabi Vásquez Maestro en Ingeniería de Sistemas con mención en Tecnologías de la Información
Fecha	15/JUN/2024

f



UNIVERSIDAD CÉSAR VALLEJO

CARTA DE PRESENTACIÓN

Señor:

JOSE EDGAR GARCIA DIAZ

Presente

Asunto: **VALIDACIÓN DE INSTRUMENTOS A TRAVÉS DE JUICIO DE EXPERTO.**

Me es muy grato comunicarme con usted para expresarle mis saludos y así mismo, hacer de su conocimiento que, siendo estudiante del **Programa Académico de Maestría en Ingeniería de Sistemas con mención en Tecnologías de la Información**, de la Escuela de Posgrado de la UCV, campus Lima Norte, ciclo 2024 - I, requiero validar el instrumento con el cual se recogerá la información necesaria para poder desarrollar mi investigación y con la que sustentaré mis competencias investigativas en la Experiencia curricular de **Diseño y Desarrollo del Trabajo de Investigación**.

El nombre de mi variable dependiente es: **Detección de cáncer en mamografías** y siendo imprescindible contar con la aprobación de profesionales especializados para poder aplicar el instrumento en mención, he considerado conveniente recurrir a usted, ante su connotada experiencia en temas relacionados a la Ingeniería de Sistemas, Tecnologías de la Información e Investigación Científica.

El expediente de validación, que le hago llegar contiene:

- Matriz de operacionalización de la variable.
- Ficha de Registro para las dimensiones de la variable.
- Ficha de validación de contenido para el instrumento.
- Matriz de validación de indicadores para la variable.
- Ficha de validación de juicio de experto.

Expresándole mis sentimientos de respeto y consideración me despido de usted, no sin antes agradecerle por la atención que dispense a la presente.

Atentamente.

.....
LUIS JOSÉ MELÉNDEZ RUIZ

D.N.I. 44180515



UNIVERSIDAD CÉSAR VALLEJO

MATRIZ DE OPERACIONALIZACIÓN DE LA VARIABLE DEPENDIENTE: DETECCIÓN DE CÁNCER EN MAMOGRAFÍAS

Indicador	Definición	Instrumento	Escala de medición	Unidad	Fórmula
Porcentaje de Sensibilidad	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para demostrar que alguien realmente padece una enfermedad, es decir, que obtenga un resultado positivo estando enfermo.	Ficha de Registro	Razón	Porcentaje	$VP / (VP + FN) * 100$, siendo: VP = Número de regiones detectadas como lesiones que Sí son lesiones, y FN = Número de regiones detectadas como sanas pero que Sí son lesiones.
Porcentaje de Especificidad	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para demostrar que alguien realmente no padece una enfermedad, es decir, que obtenga un resultado negativo estando sano.	Ficha de Registro	Razón	Porcentaje	$VN / (VN + FP) * 100$, siendo: VN = Número de regiones detectadas como sanas que NO son lesiones, y FP = Número de regiones detectadas como lesiones pero que NO son lesiones.
Porcentaje de Exactitud	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para reconocer a los enfermos como enfermos y a los sanos como sanos frente al total de los individuos sometidos a evaluación.	Ficha de Registro	Razón	Porcentaje	$(VP + VN) / (VP + FN + FP + VN) * 100$



UNIVERSIDAD CÉSAR VALLEJO

FICHA DE VALIDACIÓN DE CONTENIDO PARA EL INSTRUMENTO

INSTRUCCIÓN: A continuación, se le hace llegar el instrumento de recolección de datos del fichaje que permitirá recoger la información en la investigación que lleva por título: **Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024.**

Por lo que se le solicita que tenga a bien evaluar el instrumento haciendo, de ser caso, las sugerencias para realizar las correcciones pertinentes. Los criterios de validación de contenido son:

Criterios	Detalle	Calificación
Suficiencia	El indicador pertenece a la dimensión y basta para obtener la medición de ésta	1: de acuerdo 0: en desacuerdo
Claridad	El indicador se comprende fácilmente, es decir, su sintáctica y semántica son adecuadas.	1: de acuerdo 0: en desacuerdo
Coherencia	El indicador tiene relación lógica con la dimensión que está midiendo.	1: de acuerdo 0: en desacuerdo
Relevancia	El indicador es esencial o importante, es decir, debe ser incluido.	1: de acuerdo 0: en desacuerdo

Nota. Criterios adaptados de la propuesta de Escobar y Cuervo (2008).



UNIVERSIDAD CÉSAR VALLEJO

MATRIZ DE VALIDACIÓN DE INDICADORES PARA LA VARIABLE:

DETECCIÓN DE CÁNCER EN MAMOGRAFÍAS

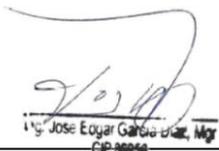
Definición de la variable: Proceso de examinar las mamas de una mujer para identificar la presencia de lesiones probablemente cancerosas a través de una radiografía, antes de que aparezcan los primeros síntomas o signos de la enfermedad.

Dimensiones	Indicadores	Suficiencia	Claridad	Coherencia	Relevancia	Observación
Sensibilidad	Porcentaje de Sensibilidad	1	1	1	1	NINGUNA
Especificidad	Porcentaje de Especificidad	1	1	1	1	NINGUNA
Exactitud	Porcentaje de Exactitud	1	1	1	1	NINGUNA



UNIVERSIDAD CÉSAR VALLEJO

FICHA DE VALIDACIÓN DE JUICIO DE EXPERTO

Nombre del instrumento	Ficha de Registro
Objetivo del instrumento	Medir las dimensiones de Sensibilidad, Especificidad y Exactitud en la Detección de Cáncer en Mamografías.
Nombres y apellidos del experto	JOSE EDGAR GARCIA DIAZ
Documento de identidad	41208549
Años de experiencia en el área	18 AÑOS
Máximo Grado Académico	MAESTRO
Nacionalidad	PERUANA
Institución	UNIV. NAC. DE LA AMAZONIA PERUANA
Cargo	DOCENTE ASOCIADO
Número telefónico	977275421
Firma	 Ing. Jose Edgar Garcia Diaz, Mg. CIP 98858
Fecha	14 DE JUNIO DE 2024

f



UNIVERSIDAD CÉSAR VALLEJO

CARTA DE PRESENTACIÓN

Señor:

RONALD PERCY MELCHOR INFANTES

Presente

Asunto: **VALIDACIÓN DE INSTRUMENTOS A TRAVÉS DE JUICIO DE EXPERTO.**

Me es muy grato comunicarme con usted para expresarle mis saludos y así mismo, hacer de su conocimiento que, siendo estudiante del **Programa Académico de Maestría en Ingeniería de Sistemas con mención en Tecnologías de la Información**, de la Escuela de Posgrado de la UCV, campus Lima Norte, ciclo 2024 - I, requiero validar el instrumento con el cual se recogerá la información necesaria para poder desarrollar mi investigación y con la que sustentaré mis competencias investigativas en la Experiencia curricular de **Diseño y Desarrollo del Trabajo de Investigación**.

El nombre de mi variable dependiente es: **Detección de cáncer en mamografías** y siendo imprescindible contar con la aprobación de profesionales especializados para poder aplicar el instrumento en mención, he considerado conveniente recurrir a usted, ante su connotada experiencia en temas relacionados a la Ingeniería de Sistemas, Tecnologías de la Información e Investigación Científica.

El expediente de validación, que le hago llegar contiene:

- Matriz de operacionalización de la variable.
- Ficha de Registro para las dimensiones de la variable.
- Ficha de validación de contenido para el instrumento.
- Matriz de validación de indicadores para la variable.
- Ficha de validación de juicio de experto.

Expresándole mis sentimientos de respeto y consideración me despido de usted, no sin antes agradecerle por la atención que dispense a la presente.

Atentamente.

.....
LUIS JOSÉ MELÉNDEZ RUIZ

D.N.I. 44180515



UNIVERSIDAD CÉSAR VALLEJO

MATRIZ DE OPERACIONALIZACIÓN DE LA VARIABLE DEPENDIENTE: DETECCIÓN DE CÁNCER EN MAMOGRAFÍAS

Indicador	Definición	Instrumento	Escala de medición	Unidad	Fórmula
Porcentaje de Sensibilidad	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para demostrar que alguien realmente padece una enfermedad, es decir, que obtenga un resultado positivo estando enfermo.	Ficha de Registro	Razón	Porcentaje	$VP / (VP + FN) * 100$, siendo: VP = Número de regiones detectadas como lesiones que Sí son lesiones, y FN = Número de regiones detectadas como sanas pero que Sí son lesiones.
Porcentaje de Especificidad	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para demostrar que alguien realmente no padece una enfermedad, es decir, que obtenga un resultado negativo estando sano.	Ficha de Registro	Razón	Porcentaje	$VN / (VN + FP) * 100$, siendo: VN = Número de regiones detectadas como sanas que NO son lesiones, y FP = Número de regiones detectadas como lesiones pero que NO son lesiones.
Porcentaje de Exactitud	Según (Torregroza-Diazgranados, 2021), es la medida de la capacidad de una prueba para reconocer a los enfermos como enfermos y a los sanos como sanos frente al total de los individuos sometidos a evaluación.	Ficha de Registro	Razón	Porcentaje	$(VP + VN) / (VP + FN + FP + VN) * 100$



UNIVERSIDAD CÉSAR VALLEJO

FICHA DE VALIDACIÓN DE CONTENIDO PARA EL INSTRUMENTO

INSTRUCCIÓN: A continuación, se le hace llegar el instrumento de recolección de datos del fichaje que permitirá recoger la información en la investigación que lleva por título: **Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024.**

Por lo que se le solicita que tenga a bien evaluar el instrumento haciendo, de ser caso, las sugerencias para realizar las correcciones pertinentes. Los criterios de validación de contenido son:

Criterios	Detalle	Calificación
Suficiencia	El indicador pertenece a la dimensión y basta para obtener la medición de ésta	1: de acuerdo 0: en desacuerdo
Claridad	El indicador se comprende fácilmente, es decir, su sintáctica y semántica son adecuadas.	1: de acuerdo 0: en desacuerdo
Coherencia	El indicador tiene relación lógica con la dimensión que está midiendo.	1: de acuerdo 0: en desacuerdo
Relevancia	El indicador es esencial o importante, es decir, debe ser incluido.	1: de acuerdo 0: en desacuerdo

Nota. Criterios adaptados de la propuesta de Escobar y Cuervo (2008).

[Handwritten mark]

[Handwritten mark]



UNIVERSIDAD CÉSAR VALLEJO

MATRIZ DE VALIDACIÓN DE INDICADORES PARA LA VARIABLE:

DETECCIÓN DE CÁNCER EN MAMOGRAFÍAS

Definición de la variable: Proceso de examinar las mamas de una mujer para identificar la presencia de lesiones probablemente cancerosas a través de una radiografía, antes de que aparezcan los primeros síntomas o signos de la enfermedad.

Dimensiones	Indicadores	Suficiencia	Claridad	Coherencia	Relevancia	Observación
Sensibilidad	Porcentaje de Sensibilidad	1	1	1	1	NINGUNA
Especificidad	Porcentaje de Especificidad	1	1	1	1	NINGUNA
Exactitud	Porcentaje de Exactitud	1	1	1	1	NINGUNA



UNIVERSIDAD CÉSAR VALLEJO

FICHA DE VALIDACIÓN DE JUICIO DE EXPERTO

Nombre del instrumento	Ficha de Registro
Objetivo del instrumento	Medir las dimensiones de Sensibilidad, Especificidad y Exactitud en la Detección de Cáncer en Mamografías.
Nombres y apellidos del experto	RONALD PERCY MELCHOR INFANTES
Documento de identidad	09846627
Años de experiencia en el área	29 AÑOS
Máximo Grado Académico	MAESTRO
Nacionalidad	PERUANA
Institución	UNIV. NAC. DE LA AMAZONIA PERUANA
Cargo	DOCENTE AUXILIAR
Número telefónico	968007618
Firma	 ing. Ronald Percy Melchor Infantes Ingeniero de Sistemas y Computo Reg. CIP. 18861.20
Fecha	14 DE JUNIO DEL 2024

f

Anexo 4. Resultados del análisis de consistencia interna

El análisis de confiabilidad del Instrumento de recolección de datos, se realizó a través del método de test-retest. Se realizaron mediciones adicionales (retest) de las dimensiones, a través de los indicadores de sensibilidad, especificidad y exactitud, antes de la aplicación del modelo predictivo (variable independiente) en la detección de cáncer en mamografías (variable dependiente). Tales mediciones adicionales fueron sometidas a evaluación, para determinar el grado de correlación respecto a las mediciones del test (pre-test).

A continuación, se muestran los estadísticos descriptivos por cada dimensión:

Tabla 10

Estadísticos descriptivos de la dimensión sensibilidad en test y retest

Estadísticos descriptivos	Masas		Calcificaciones	
	test	retest	test	retest
<i>Elementos válidos</i>	56	56	179	179
<i>Elementos perdidos</i>	0	0	0	0
<i>Media</i>	83.0979	83.7711	95.0366	95.6488
<i>Mediana</i>	97.0700	98.0350	100.0000	100.0000
<i>Varianza</i>	692.408	682.499	198.132	171.986
<i>Desv. Desviación</i>	26.31364	26.12468	14.07593	13.11435
<i>Mínimo</i>	.00	.00	25.00	25.00
<i>Máximo</i>	100.00	100.00	100.00	100.00
<i>Desv. Error</i>	3.51631	3.49106	1.05208	.98021

Fuente: autoría personal.

Tabla 11

Estadísticos descriptivos de la dimensión especificidad en test y retest

Estadísticos descriptivos	Masas		Calcificaciones	
	test	retest	test	retest
<i>Elementos válidos</i>	56	56	179	179

<i>Elementos perdidos</i>	0	0	0	0
<i>Media</i>	50.7925	48.1871	45.0580	44.2168
<i>Mediana</i>	43.9250	40.5100	42.9300	43.2400
<i>Varianza</i>	498.799	536.458	488.837	486.126
<i>Desv. Desviación</i>	22.33382	23.16156	22.10966	22.04827
<i>Mínimo</i>	17.93	16.92	4.24	3.51
<i>Máximo</i>	93.90	93.68	99.85	99.65
<i>Desv. Error</i>	2.98448	3.09509	1.65255	1.64796

Fuente: autoría personal.

Tabla 12

Estadísticos descriptivos de la dimensión exactitud en test y retest

Estadísticos descriptivos	Masas		Calcificaciones	
	test	retest	test	retest
<i>Elementos válidos</i>	56	56	179	179
<i>Elementos perdidos</i>	0	0	0	0
<i>Media</i>	51.4934	48.9925	45.4456	44.6255
<i>Mediana</i>	47.4600	46.8100	43.0700	43.3100
<i>Varianza</i>	472.236	510.128	478.572	476.486
<i>Desv. Desviación</i>	21.73100	22.58602	21.87629	21.82856
<i>Mínimo</i>	18.17	17.16	4.32	3.59
<i>Máximo</i>	93.72	93.51	99.83	99.63
<i>Desv. Error</i>	2.90393	3.01818	1.63511	1.63154

Fuente: autoría personal.

De las tablas 10, 11 y 12, se observó que los valores de los estadísticos descriptivos entre test y retest eran muy similares, tanto en las lesiones de tipo masa como en las calcificaciones, lo que en primera instancia indicaba que el instrumento presentaba un comportamiento coherente entre mediciones.

Seguidamente, mediante la prueba de normalidad de Kolmogórov-Smirnov se evaluó si las mediciones presentaban una distribución normal o no:

Tabla 13*Estadístico de normalidad de la dimensión sensibilidad en test y retest*

Descripción	Masas		Calcificaciones	
	test	retest	test	retest
Estadístico K-S	.265	.274	.426	.434
gl	56	56	179	179
Sig. (p)	.000	.000	.000	.000

*Fuente: autoría personal.***Tabla 14***Estadístico de normalidad de la dimensión especificidad en test y retest*

Descripción	Masas		Calcificaciones	
	test	retest	test	retest
Estadístico K-S	.136	.147	.081	.068
gl	56	56	179	179
Sig. (p)	.011	.004	.007	.045

*Fuente: autoría personal.***Tabla 15***Estadístico de normalidad de la dimensión exactitud en test y retest*

Descripción	Masas		Calcificaciones	
	test	retest	test	retest
Estadístico K-S	.135	.139	.077	.063
gl	56	56	179	179
Sig. (p)	.013	.009	.012	.077

Fuente: autoría personal.

De las tablas 13, 14 y 15, se observó un valor p menor a 0.05 en todos los casos, con excepción del retest en lesiones de tipo calcificaciones de la dimensión exactitud.

Por tanto, ya que los datos en su mayoría no presentaban una distribución normal, para la medida de la correlación entre test y retest, se aplicó una prueba no paramétrica, en este caso la prueba de correlación de Spearman:

Tabla 16

Coefficiente de correlación de la dimensión sensibilidad en test y retest

	Masas	Calcificaciones
Descripción	test vs retest	test vs retest
<i>Estadístico Rho de Spearman</i>	.930	.953
<i>Sig. (bilateral)</i>	.000	.000
<i>N</i>	56	179

Fuente: autoría personal.

Tabla 17

Coefficiente de correlación de la dimensión especificidad en test y retest

	Masas	Calcificaciones
Descripción	test vs retest	test vs retest
<i>Estadístico Rho de Spearman</i>	.982	.993
<i>Sig. (bilateral)</i>	.000	.000
<i>N</i>	56	179

Fuente: autoría personal.

Tabla 18

Coefficiente de correlación de la dimensión exactitud en test y retest

	Masas	Calcificaciones
Descripción	test vs retest	test vs retest
<i>Estadístico Rho de Spearman</i>	.985	.993
<i>Sig. (bilateral)</i>	.000	.000
<i>N</i>	56	179

Fuente: autoría personal.

De las tablas 16, 17 y 18, se observó coeficientes de correlación con un valor de Rho, en el intervalo 0.9 a 1, expresando una correlación positiva muy alta entre test y retest. Este resultado es acorde con lo observado en los estadísticos descriptivos presentados en las tablas 10, 11 y 12, cuyo análisis en primera instancia indicaba que el instrumento presentaba un comportamiento coherente entre mediciones. La alta consistencia y fiabilidad entre ambas mediciones, indican sin lugar a dudas, que el instrumento de recolección de datos es estable y produce resultados consistentes.

Como constancia, se adjuntan en las siguientes páginas, las fichas de registro con los datos recolectados del retest.

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
MASAS (1-20 de 56)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_105_143_1_RCC.tif	357	0	5399	3212	100.00	37.30	39.80
img_105_143_1_RO.tif	394	0	6473	3395	100.00	34.40	36.92
img_108_146_1_RCC.tif	31	0	3948	1004	100.00	20.28	20.77
img_108_146_1_RO.tif	20	0	4883	2144	100.00	30.51	30.71
img_12_20_1_RO.tif	667	77	2006	1528	89.65	43.24	51.31
img_134_180_1_RCC.tif	409	35	1956	1329	92.12	40.46	46.61
img_139_188_1_LO.tif	567	11	4192	3987	98.10	48.75	52.00
img_174_233_1_LCC.tif	45	14	2159	2039	76.27	48.57	48.96
img_177_237_1_LO.tif	661	26	1959	6706	96.22	77.39	78.78
img_186_251_1_LCC.tif	38	0	3841	836	100.00	17.88	18.54
img_18_27_1_LCC.tif	46	0	6785	1831	100.00	21.25	21.67
img_18_28_1_LO.tif	7	0	3646	1239	100.00	25.36	25.47
img_194_262_1_LCC.tif	10	0	2631	2442	100.00	48.14	48.24
img_194_262_1_LO.tif	10	0	2765	3017	100.00	52.18	52.26
img_196_264_1_RCC.tif	22	0	1195	1258	100.00	51.28	51.72
img_206_455_1_LCC.tif	1	8	266	3944	11.11	93.68	93.51
img_239_320_1_LO.tif	2	3	2614	3026	40.00	53.65	53.64
img_241_449_1_RCC.tif	29	0	8344	1699	100.00	16.92	17.16
img_261_346_1_LCC.tif	294	21	1634	1832	93.33	52.86	56.23
img_273_360_1_LCC.tif	29	0	2604	1010	100.00	27.95	28.52

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
MASAS (21-40 de 56)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_287_379_1_RCC.tif	63	0	1514	1001	100.00	39.80	41.27
img_287_379_1_RO.tif	54	0	2403	2078	100.00	46.37	47.01
img_306_405_1_RO.tif	174	10	4393	2753	94.57	38.53	39.93
img_323_430_1_RCC.tif	7	6	2135	1970	53.85	47.99	48.01
img_323_430_1_RO.tif	8	4	2902	2624	66.67	47.49	47.53
img_353_487_1_RCC.tif	96	0	4401	2247	100.00	33.80	34.74
img_3_4_1_LO.tif	120	0	3711	1713	100.00	31.58	33.06
img_408_567_1_RCC.tif	146	0	1671	1040	100.00	38.36	41.51
img_429_596_1_RCC.tif	193	4	3761	2566	97.97	40.56	42.29
img_448_622_1_RCC.tif	50	103	772	3744	32.68	82.91	81.26
img_448_622_1_RO.tif	130	298	1047	4867	30.37	82.30	78.79
img_449_625_1_RCC.tif	67	3	1456	1801	95.71	55.30	56.15
img_449_625_1_RO.tif	72	9	2275	2183	88.89	48.97	49.68
img_459_645_1_LO.tif	4	45	293	3920	8.16	93.05	92.07
img_46_65_1_LO.tif	8	1	3105	1682	88.89	35.14	35.24
img_476_672_1_RO.tif	55	39	1116	7549	58.51	87.12	86.81
img_477_674_1_RO.tif	0	27	1410	6929	0.00	83.09	82.82
img_488_691_1_LO.tif	26	8	1029	4226	76.47	80.42	80.39
img_491_696_1_LO.tif	73	27	2381	6052	73.00	71.77	71.78
img_4_8_1_LCC.tif	23	0	4104	2028	100.00	33.07	33.32

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (1-20 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_100_135_1_RCC.tif	116	0	5949	2169	100.00	26.72	27.75
img_105_143_1_RCC.tif	357	0	3661	4950	100.00	57.49	59.18
img_105_143_1_RO.tif	392	2	3740	6128	99.49	62.10	63.54
img_106_144_1_RCC.tif	2	0	3290	479	100.00	12.71	12.76
img_106_144_1_RO.tif	4	0	3932	837	100.00	17.55	17.62
img_114_152_1_LO.tif	14	0	3842	1094	100.00	22.16	22.38
img_115_154_1_LCC.tif	5	0	2459	1293	100.00	34.46	34.55
img_115_155_1_LO.tif	6	0	4617	1121	100.00	19.54	19.62
img_117_158_1_LCC.tif	3	0	3549	129	100.00	3.51	3.59
img_117_158_1_LO.tif	2	0	3668	337	100.00	8.41	8.46
img_118_159_1_LCC.tif	4	0	3509	287	100.00	7.56	7.66
img_118_159_1_RCC.tif	9	0	1834	635	100.00	25.72	25.99
img_118_159_1_RO.tif	6	0	2149	401	100.00	15.73	15.92
img_123_164_1_RO.tif	223	0	2950	766	100.00	20.61	25.11
img_124_165_1_LCC.tif	1	0	1253	429	100.00	25.51	25.55
img_127_170_1_RCC.tif	4	0	4808	1996	100.00	29.34	29.38
img_127_170_1_RO.tif	7	0	6340	2463	100.00	27.98	28.04
img_138_186_1_RCC.tif	8	0	2048	1638	100.00	44.44	44.56
img_138_187_1_RCC.tif	9	0	1269	2683	100.00	67.89	67.96
img_140_190_1_RO.tif	2	0	6309	1380	100.00	17.95	17.97

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (21-40 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_145_197_1_LCC.tif	2	0	2526	5104	100.00	66.89	66.90
img_145_198_1_RO.tif	4	0	4024	3013	100.00	42.82	42.85
img_150_203_1_RCC.tif	4	0	5075	908	100.00	15.18	15.23
img_150_203_1_RO.tif	6	0	6191	960	100.00	13.43	13.50
img_153_206_1_LCC.tif	2	0	1568	1436	100.00	47.80	47.84
img_153_206_1_LO.tif	1	0	3083	1204	100.00	28.09	28.10
img_153_207_1_LO.tif	2	0	2322	1001	100.00	30.12	30.17
img_155_209_1_RO.tif	4	0	2158	4185	100.00	65.98	66.00
img_161_216_1_RCC.tif	4	0	3018	1230	100.00	28.96	29.02
img_164_219_1_LCC.tif	155	0	4486	203	100.00	4.33	7.39
img_164_219_1_LO.tif	175	0	5723	300	100.00	4.98	7.66
img_165_220_1_LCC.tif	4	2	965	2942	66.67	75.30	75.29
img_169_224_1_LCC.tif	3	0	1963	5404	100.00	73.35	73.37
img_169_225_1_LO.tif	1	0	5283	4857	100.00	47.90	47.91
img_16_24_1_LO.tif	5	0	2532	2089	100.00	45.21	45.27
img_172_231_1_RO.tif	3	1	4141	4509	75.00	52.13	52.14
img_175_234_1_RCC.tif	2	0	4172	3223	100.00	43.58	43.60
img_177_237_1_LO.tif	679	8	2865	5800	98.84	66.94	69.28
img_181_242_1_LCC.tif	1	0	2220	919	100.00	29.28	29.30
img_181_242_1_LO.tif	2	0	4492	1647	100.00	26.83	26.85

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (41-60 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_183_245_1_RCC.tif	4	0	2155	2737	100.00	55.95	55.98
img_183_246_1_RCC.tif	2	0	1083	3285	100.00	75.21	75.22
img_183_246_1_RO.tif	1	0	3924	3718	100.00	48.65	48.66
img_188_253_1_RCC.tif	4	0	2188	164	100.00	6.97	7.13
img_189_254_1_RO.tif	6	0	3366	3077	100.00	47.76	47.81
img_18_28_1_LCC.tif	8	0	1147	881	100.00	43.44	43.66
img_191_258_1_RCC.tif	577	10	2114	771	98.30	26.72	38.83
img_194_262_1_RCC.tif	6	0	3569	1838	100.00	33.99	34.07
img_194_262_1_RO.tif	4	0	2736	1715	100.00	38.53	38.59
img_196_264_1_RCC.tif	14	0	1355	1106	100.00	44.94	45.25
img_196_264_1_RO.tif	19	0	2449	1550	100.00	38.76	39.05
img_199_268_1_RCC.tif	4	0	2755	1147	100.00	29.40	29.47
img_199_269_1_RCC.tif	4	0	2705	1208	100.00	30.87	30.94
img_199_269_1_RO.tif	3	0	3602	1436	100.00	28.50	28.55
img_205_275_1_LO.tif	224	0	4250	3468	100.00	44.93	46.49
img_210_282_1_LCC.tif	6	0	5148	1594	100.00	23.64	23.71
img_210_283_1_LCC.tif	5	0	4911	2567	100.00	34.33	34.37
img_210_283_1_LO.tif	4	0	5137	2948	100.00	36.46	36.49
img_212_286_1_LCC.tif	8	0	5100	3452	100.00	40.37	40.42
img_212_286_1_RO.tif	4	0	5853	3556	100.00	37.79	37.82

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (61-80 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_218_294_1_LCC.tif	7	0	3640	415	100.00	10.23	10.39
img_229_308_1_LCC.tif	134	0	3396	1181	100.00	25.80	27.91
img_246_328_1_LCC.tif	60	117	367	8262	33.90	95.75	94.50
img_249_332_1_RCC.tif	243	22	2537	3340	91.70	56.83	58.34
img_259_344_1_LCC.tif	142	0	1405	2253	100.00	61.59	63.03
img_259_344_1_LO.tif	181	3	2018	3055	98.37	60.22	61.56
img_262_450_1_RCC.tif	4	0	1144	1153	100.00	50.20	50.28
img_262_450_1_RO.tif	5	0	1763	1969	100.00	52.76	52.82
img_26_37_1_LCC.tif	5	0	1442	1709	100.00	54.24	54.31
img_273_360_1_LCC.tif	29	0	1988	1626	100.00	44.99	45.43
img_275_363_1_LCC.tif	4	0	3261	490	100.00	13.06	13.16
img_276_365_1_RO.tif	6	0	4225	1506	100.00	26.28	26.36
img_293_386_1_LCC.tif	8	0	2108	1037	100.00	32.97	33.14
img_293_386_1_LO.tif	9	0	2982	1686	100.00	36.12	36.24
img_297_392_1_LCC.tif	3	1	1539	3413	75.00	68.92	68.93
img_297_392_1_LO.tif	4	0	2659	3323	100.00	55.55	55.58
img_2_3_1_LCC.tif	5	0	4690	4167	100.00	47.05	47.08
img_2_3_1_RCC.tif	5	0	6917	2116	100.00	23.43	23.47
img_300_396_1_RCC.tif	2	0	2116	2907	100.00	57.87	57.89
img_301_398_1_RCC.tif	335	0	3492	1362	100.00	28.06	32.70

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (81-100 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_301_398_1_RO.tif	395	0	4777	2252	100.00	32.04	35.66
img_305_403_1_LO.tif	343	10	1076	2413	97.17	69.16	71.73
img_306_405_1_RCC.tif	185	41	3418	2354	81.86	40.78	42.33
img_306_405_1_RO.tif	164	20	4056	3090	89.13	43.24	44.39
img_30_41_1_LCC.tif	2	0	1240	1020	100.00	45.13	45.18
img_30_41_1_LO.tif	2	0	2036	1826	100.00	47.28	47.31
img_315_416_1_RO.tif	4	0	3815	2088	100.00	35.37	35.42
img_319_422_1_LCC.tif	12	0	3628	3032	100.00	45.53	45.62
img_319_423_1_RCC.tif	48	0	2695	2393	100.00	47.03	47.53
img_31_43_1_RO.tif	4	0	5972	1122	100.00	15.82	15.86
img_323_430_1_RCC.tif	3	0	2918	1197	100.00	29.09	29.14
img_327_437_1_RCC.tif	9	0	2557	4095	100.00	61.56	61.61
img_328_438_1_LCC.tif	15	0	3585	5138	100.00	58.90	58.97
img_328_438_1_LO.tif	14	0	4421	4675	100.00	51.40	51.47
img_332_442_1_LCC.tif	4	0	718	2874	100.00	80.01	80.03
img_33_46_1_LO.tif	2	0	6785	1752	100.00	20.52	20.54
img_341_469_1_LCC.tif	2	0	1856	770	100.00	29.32	29.38
img_354_489_1_LO.tif	4	0	5125	1536	100.00	23.06	23.11
img_360_496_1_RO.tif	4	0	5075	2322	100.00	31.39	31.43
img_364_501_1_RO.tif	2	0	3594	3567	100.00	49.81	49.83

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (101-120 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_365_503_1_RCC.tif	7	0	3289	1625	100.00	33.07	33.16
img_365_503_1_RO.tif	21	0	3589	1973	100.00	35.47	35.72
img_374_518_1_LCC.tif	6	0	3175	2420	100.00	43.25	43.31
img_375_519_1_LCC.tif	11	0	2080	2580	100.00	55.37	55.47
img_37_51_1_RO.tif	1	0	2146	2110	100.00	49.58	49.59
img_385_534_1_LCC.tif	4	0	1198	4246	100.00	77.99	78.01
img_385_534_1_LO.tif	3	0	3498	4555	100.00	56.56	56.58
img_386_535_1_RCC.tif	4	0	1541	2700	100.00	63.66	63.70
img_386_535_1_RO.tif	4	0	2075	4307	100.00	67.49	67.51
img_38_53_1_RCC.tif	14	0	2106	804	100.00	27.63	27.98
img_38_54_1_RCC.tif	14	0	2362	465	100.00	16.45	16.86
img_38_54_1_RO.tif	12	0	3690	635	100.00	14.68	14.92
img_390_539_1_LO.tif	31	3	2671	1510	91.18	36.12	36.56
img_392_543_1_LCC.tif	5	0	627	1638	100.00	72.32	72.38
img_392_543_1_LO.tif	4	0	1279	2106	100.00	62.22	62.26
img_392_543_1_RCC.tif	2	0	1299	1948	100.00	59.99	60.02
img_395_548_1_LCC.tif	4	0	6409	3091	100.00	32.54	32.57
img_395_548_1_LO.tif	4	0	7517	3793	100.00	33.54	33.56
img_396_549_1_RCC.tif	93	3	2230	1469	96.88	39.71	41.16
img_396_549_1_RO.tif	66	3	2171	2434	95.65	52.86	53.49

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (121-140 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_399_553_1_LCC.tif	2	3	412	3915	40.00	90.48	90.42
img_399_553_1_LO.tif	1	3	2380	4486	25.00	65.34	65.31
img_401_555_1_LO.tif	27	10	3013	2513	72.97	45.48	45.66
img_402_556_1_LO.tif	1	0	5562	4088	100.00	42.36	42.37
img_406_563_1_RCC.tif	4	0	1828	316	100.00	14.74	14.90
img_412_573_1_LCC.tif	6	0	1357	1709	100.00	55.74	55.83
img_412_573_1_LO.tif	5	0	3330	1454	100.00	30.39	30.47
img_418_581_1_LCC.tif	4	0	4506	4889	100.00	52.04	52.06
img_418_581_1_LO.tif	6	0	5849	6221	100.00	51.54	51.57
img_418_581_1_RCC.tif	27	1	3872	4557	96.43	54.06	54.20
img_419_582_1_RO.tif	301	31	3487	7876	90.66	69.31	69.92
img_433_600_1_RCC.tif	32	1	3213	1972	96.97	38.03	38.41
img_433_600_1_RO.tif	32	0	4061	2368	100.00	36.83	37.15
img_438_606_1_RO.tif	12	0	3976	4290	100.00	51.90	51.97
img_447_619_1_LCC.tif	2	1	374	5059	66.67	93.12	93.10
img_448_622_1_RCC.tif	128	25	1937	2579	83.66	57.11	57.98
img_452_631_1_RCC.tif	5	3	470	4084	62.50	89.68	89.63
img_456_638_1_LCC.tif	6	0	1934	1831	100.00	48.63	48.71
img_469_663_1_RCC.tif	7	0	1656	1079	100.00	39.45	39.61
img_475_671_1_LCC.tif	4	0	1899	1667	100.00	46.75	46.81

FICHA DE REGISTRO							
Tipo de Prueba:	RETEST						
Variable:	Detección de Cáncer en Mamografías						
Dimensiones:	SENSIBILIDAD, ESPECIFICIDAD Y EXACTITUD						
Indicadores:	Porcentaje de Sensibilidad (P_SENS) Porcentaje de Especificidad (P_ESPE) Porcentaje de Exactitud (P_EXAC)						
Escala:	Razón						
Fórmulas:	$P_SENS = VP / (VP + FN) * 100$ $P_ESPE = VN / (VN + FP) * 100$ $P_EXAC = (VP + VN) / (VP + FN + FP + VN) * 100$ <p>Siendo:</p> <ul style="list-style-type: none"> • VP = Número de regiones detectadas como lesiones que SÍ son lesiones, • FN = Número de regiones detectadas como sanas pero que SÍ son lesiones, • FP = Número de regiones detectadas como lesiones pero que NO son lesiones, y • VN = Número de regiones detectadas como sanas que NO son lesiones. 						
CALCIFICACIONES (141-160 de 179)							
Mamografía	VP	FN	FP	VN	P_SENS	P_ESPE	P_EXAC
img_475_671_1_LO.tif	4	0	3059	2032	100.00	39.91	39.96
img_475_671_1_RCC.tif	4	0	1805	1804	100.00	49.99	50.04
img_477_674_1_LO.tif	4	0	2381	6380	100.00	72.82	72.84
img_481_678_1_RCC.tif	2	0	888	5228	100.00	85.48	85.49
img_481_678_1_RO.tif	1	0	1146	6436	100.00	84.89	84.89
img_483_681_1_RCC.tif	16	0	2602	2780	100.00	51.65	51.80
img_487_688_1_LCC.tif	6	0	39	5042	100.00	99.23	99.23
img_487_688_1_LO.tif	5	0	1097	5150	100.00	82.44	82.45
img_488_691_1_LO.tif	34	0	2345	2910	100.00	55.38	55.66
img_49_69_1_LCC.tif	3	0	2244	2590	100.00	53.58	53.61
img_4_7_1_LCC.tif	3	0	2974	2209	100.00	42.62	42.65
img_4_7_1_LO.tif	4	0	3645	3909	100.00	51.75	51.77
img_504_718_1_LCC.tif	5	1	23	4589	83.33	99.50	99.48
img_504_719_1_LCC.tif	1	1	16	4574	50.00	99.65	99.63
img_507_725_1_RCC.tif	67	10	332	9210	87.01	96.52	96.45
img_507_725_1_RO.tif	84	15	414	10018	84.85	96.03	95.93
img_508_727_1_RCC.tif	5	0	1640	2233	100.00	57.66	57.71
img_508_727_1_RO.tif	4	0	2331	3037	100.00	56.58	56.61
img_51_73_1_LCC.tif	4	0	1757	2773	100.00	61.21	61.25
img_51_73_1_LO.tif	6	0	2114	3043	100.00	59.01	59.06

Anexo 5. Consentimiento o asentimiento informado UCV

Consentimiento Informado

Título de la investigación: Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024.

Investigador: Luis José Meléndez Ruiz

Propósito del estudio

Le invitamos a participar en la investigación titulada "Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024", cuyo objetivo es: Determinar la influencia de un Modelo Predictivo en la Detección de Cáncer en Mamografías. Esta investigación es desarrollada por un estudiante del Programa Académico de Maestría en Ingeniería de Sistemas con mención en Tecnologías de la Información, de la Universidad César Vallejo del campus Lima Norte, aprobado por la autoridad correspondiente de la Universidad y con el permiso de la institución

Según la OMS, el cáncer es una de las principales causas de muerte en todo el mundo: sólo en el año 2020 se atribuyeron a esta enfermedad casi 10 millones de defunciones; siendo el cáncer de mama el tipo más común, con más de 2,2 millones de nuevos casos y 685 000 defunciones. Actualmente la mamografía, o radiografía de las mamas, es la prueba de detección temprana más utilizada, tanto por su accesibilidad como por su efectividad. Los datos obtenidos de los programas de detección mediante mamografías indican que, tras 11 años de seguimiento, se observa una disminución del 20% en la mortalidad específica por cáncer de mama. Aunque las estadísticas podrían sugerir lo contrario, el tratamiento del cáncer de mama puede ser muy efectivo, con tasas de supervivencia del 90% o superiores, especialmente cuando se detecta en etapas tempranas, es decir, cuando el cáncer es pequeño y localizado. Por lo tanto, promover la detección precoz sigue siendo crucial para mejorar el pronóstico y la supervivencia de las personas afectadas por esta enfermedad.

Procedimiento

Si usted decide participar en la investigación se realizará lo siguiente:

1. Se realizará una encuesta o entrevista donde se recogerán datos personales y algunas preguntas.
2. Esta encuesta o entrevista tendrá un tiempo aproximado de minutos y se realizará en el ambiente de de la institución

Las respuestas al cuestionario o guía de entrevista serán codificadas usando un número de identificación y, por lo tanto, serán anónimas.

Participación voluntaria (principio de autonomía):

Puede hacer todas las preguntas para aclarar sus dudas antes de decidir si desea participar o no, y su decisión será respetada. Si posterior a la aceptación no desea continuar puede hacerlo sin ningún problema.

Riesgo (principio de No maleficencia):

NO existe riesgo o daño al participar en la investigación. Sin embargo, en el caso que existan preguntas que le puedan generar incomodidad usted tiene la libertad de responderlas o no.

Beneficios (principio de beneficencia):

Se le informa que los resultados de la investigación serán alcanzados a la institución al término de la investigación. No recibirá ningún beneficio económico ni de ninguna otra índole. El estudio no aportará a la salud individual de su persona, sin embargo, los resultados del estudio podrán convertirse en beneficio de la salud pública.

Confidencialidad (principio de justicia):

Los datos recolectados serán anónimos y no existirá ninguna forma de identificarlo como participante. Garantizamos que la información que usted nos brinde es totalmente Confidencial y no será usada para ningún otro propósito fuera de la investigación. Los datos permanecerán bajo custodia del investigador principal y pasado un tiempo determinado serán eliminados convenientemente.

Problemas o preguntas:

Si tiene preguntas sobre la investigación puede contactar con el investigador: Luis José Meléndez Ruiz, al email: lmelendezru@ucvvirtual.edu.pe y asesor: Jhonatan Isaac Vargas Huaman, al email: jivargash@ucvvirtual.edu.pe

Consentimiento

Después de haber leído los propósitos de la investigación autorizo mi participación en la investigación antes mencionada.

Nombre y apellidos:

Fecha y hora:

Firma:

.....

Asentimiento Informado

Título de la investigación: Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024.

Investigador: Luis José Meléndez Ruiz

Propósito del estudio

Le invitamos a participar en la investigación titulada "Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima 2024", cuyo objetivo es: Determinar la influencia de un Modelo Predictivo en la Detección de Cáncer en Mamografías. Esta investigación es desarrollada por un estudiante del Programa Académico de Maestría en Ingeniería de Sistemas con mención en Tecnologías de la Información, de la Universidad César Vallejo del campus Lima Norte, aprobado por la autoridad correspondiente de la Universidad y con el permiso de la Institución

Según la OMS, el cáncer es una de las principales causas de muerte en todo el mundo: sólo en el año 2020 se atribuyeron a esta enfermedad casi 10 millones de defunciones; siendo el cáncer de mama el tipo más común, con más de 2,2 millones de nuevos casos y 685 000 defunciones. Actualmente la mamografía, o radiografía de las mamas, es la prueba de detección temprana más utilizada, tanto por su accesibilidad como por su efectividad. Los datos obtenidos de los programas de detección mediante mamografías indican que, tras 11 años de seguimiento, se observa una disminución del 20% en la mortalidad específica por cáncer de mama. Aunque las estadísticas podrían sugerir lo contrario, el tratamiento del cáncer de mama puede ser muy efectivo, con tasas de supervivencia del 90% o superiores, especialmente cuando se detecta en etapas tempranas, es decir, cuando el cáncer es pequeño y localizado. Por lo tanto, promover la detección precoz sigue siendo crucial para mejorar el pronóstico y la supervivencia de las personas afectadas por esta enfermedad.

Procedimiento

Si usted decide participar en la investigación se realizará lo siguiente:

1. Se realizará una encuesta o entrevista donde se recogerán datos personales y algunas preguntas.
2. Esta encuesta o entrevista tendrá un tiempo aproximado de minutos y se realizará en el ambiente de de la institución

Las respuestas al cuestionario o guía de entrevista serán codificadas usando un número de identificación y, por lo tanto, serán anónimas.

Participación voluntaria (principio de autonomía):

Su menor hijo(a)/representado puede hacer todas las preguntas para aclarar sus dudas antes de decidir si desea participar o no, y su decisión será respetada. Posterior a que su menor hijo(a)/representado haya aceptado participar, puede dejar de participar sin ningún problema.

Riesgo (principio de No maleficencia):

NO existe riesgo o daño para su menor hijo(a)/representado al participar en la investigación. Sin embargo, en el caso que existan preguntas que le puedan generar incomodidad a su menor su hijo(a)/representado tiene la libertad de responderlas o no.

Beneficios (principio de beneficencia):

Se le informa que los resultados de la investigación serán alcanzados a la institución al término de la investigación. No recibirá ningún beneficio económico ni de ninguna otra índole. El estudio no aportará a la salud individual de su menor hijo(a)/representado, sin embargo, los resultados del estudio podrán convertirse en beneficio de la salud pública.

Confidencialidad (principio de justicia):

Los datos recolectados serán anónimos y no existirá ninguna forma de identificar al participante. Garantizamos que la información recogida en la encuesta o entrevista a su hijo(a)/representado es totalmente confidencial y no será usada para ningún otro propósito fuera de la investigación. Los datos permanecerán bajo custodia del investigador principal y pasado un tiempo determinado serán eliminados convenientemente.

Problemas o preguntas:

Si tiene preguntas sobre la investigación puede contactar con el investigador: Luis José Meléndez Ruiz, al email: lmelendezru@ucvvirtual.edu.pe y asesor: Jhonatan Isaac Vargas Huaman, al email: jvargash@ucvvirtual.edu.pe

Asentimiento

Después de haber leído los propósitos de la investigación autorizo que mi hijo(a)/representado participe en la investigación.

Nombre y apellidos:

Fecha y hora:

Firma:

.....

Anexo 6. Reporte de similitud en software Turnitin

LUIS JOSE MELENDEZ RUIZ | Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima, 2024

UNIVERSIDAD CÉSAR VALLEJO

ESCUELA DE POSGRADO
PROGRAMA ACADÉMICO DE MAESTRIA EN INGENIERIA
DE SISTEMAS CON MENCION EN TECNOLOGIAS DE LA
INFORMACION

**Modelo Predictivo en la Detección de Cáncer en Mamografías, Lima
2024**

TESIS PARA OBTENER EL GRADO ACADÉMICO DE:
Maestro en Ingeniería de Sistemas

AUTOR:
Meléndez Ruiz, Luis José (orcid.org/0000-0002-2179-4415)

ASESORES:
Dr. Vargas Huaman, Jhonatan Isaac (orcid.org/0000-0002-1433-7494)
Mg. Puente Zamora, Jonathan Alexis (orcid.org/0009-0007-1034-1617)

LÍNEA DE INVESTIGACIÓN:
Sistemas de Información y Comunicaciones

LÍNEA DE RESPONSABILIDAD SOCIAL UNIVERSITARIA:
Desarrollo económico, empleo y emprendimiento

LIMA – PERÚ
2024

18 %

Se están viendo fuentes estándar

Ver fuentes en inglés

Concidencias	Porcentaje
1 hdi.baniplo.net Fuente de Internet	3 %
2 Entregado a Universidad... Trabajo del estudiante	2 %
3 repositorio.uoc.edu.pe Fuente de Internet	1 %
4 Entregado a Universidad... Trabajo del estudiante	1 %
5 library.co Fuente de Internet	1 %
6 www.researchgate.net Fuente de Internet	1 %
7 www.cine.gov.do Fuente de Internet	<1 %
8 revista.areasindia.edu.co Fuente de Internet	<1 %
9 doaj.org Fuente de Internet	<1 %
10 repositorio.uletech.edu... Fuente de Internet	<1 %
11 docplayer.es Fuente de Internet	<1 %
12 repositorio.unamba.edu... Fuente de Internet	<1 %
13 www.cancerquest.org Fuente de Internet	<1 %
14 www.medilija.com Fuente de Internet	<1 %
15 Entregado a Universidad... Trabajo del estudiante	<1 %

Feedback Studio - Google Chrome

Feedback studio

Número de palabras: 9425

Página: 1 de 30

Alta resolución

14:54

8/09/2024

Anexo 7. Análisis complementario

CÁLCULO DEL TAMAÑO DE LA MUESTRA

Como se indicó en el capítulo II. Metodología, la población de la presente investigación estuvo conformada por un conjunto de mamografías que contenían dos tipos de lesiones de cáncer de mama: masas y calcificaciones, siendo estas lesiones las unidades de análisis. El marco poblacional correspondió a 598 lesiones (143 masas y 455 calcificaciones).

Para calcular el tamaño de la muestra, se utilizó la fórmula general:

$$n = \frac{Z^2 \cdot p \cdot (1 - p)}{E^2}$$

Donde:

- n = Tamaño de la muestra
- Z = Valor crítico de Z correspondiente al nivel de confianza. Para un alfa de 0.05, el nivel de confianza es 95%, y el valor crítico de Z es 1.96
- p = Estimación de la proporción de la población. Si no se conoce, se usa 0.5 como una estimación conservadora.
- E = Margen de error tolerable (alfa = 0.05)

Sustituyendo los datos en la fórmula (para el valor final se usa redondeo hacia arriba):

$$n = \frac{(1.96)^2 \cdot 0.5 \cdot (1 - 0.5)}{(0.05)^2} = \frac{3.8416 \cdot 0.5 \cdot (1 - 0.5)}{0.0025} = \frac{0.9604}{0.0025} \approx 384.16 = 385$$

Como la población es finita, se ajusta el tamaño de la muestra con la fórmula de corrección para poblaciones finitas:

$$n_{adj} = \frac{n}{1 + \frac{n - 1}{N}}$$

Donde:

- n_{adj} = Tamaño de muestra ajustado

- n = Tamaño de la muestra
- N = Tamaño total de la población

Entonces, para la población de 598 lesiones, la muestra fue de:

$$n_{adj} = \frac{385}{1 + \frac{385 - 1}{598}} = \frac{385}{1 + \frac{384}{598}} = \frac{385}{1 + 0.642} = \frac{385}{1.642} \approx 234.47 = 235$$

Una vez conocido el tamaño de la muestra, para asegurar una correcta representación de ambos tipos de lesiones (estratos), se realizó un muestro aleatorio estratificado, de tipo proporcional, evitando así cualquier sesgo en la selección de la muestra.

La fórmula para calcular el tamaño de la muestra de cada estrato es:

$$n_h = \frac{N_h}{N} \times n$$

Donde:

- n_h = Tamaño de la muestra en el estrato h
- N_h = Tamaño del estrato h en la población
- N = Tamaño total de la población
- n = Tamaño total de la muestra

Entonces, para las lesiones de tipo masas:

$$n_{masas} = \frac{143}{598} \times 235 = 0.239 \times 235 \approx 56.17 = 56$$

Mientras que, para las lesiones de tipo calcificaciones:

$$n_{masas} = \frac{455}{598} \times 235 = 0.761 \times 235 \approx 178.83 = 179$$

A efectos de conservar la cantidad de la muestra, se usa el redondeo por defecto.

Anexo 8. Autorizaciones para el desarrollo del proyecto de investigación



DECLARACIÓN JURADA: USO DE DATOS PÚBLICOS

Apellidos y nombres	Meléndez Ruiz, Luis José
DNI	44180515
Código de estudiante	7003066422
Filial	Lima Norte
Programa	Maestría en Ingeniería de Sistemas con mención en Tecnologías de la Información
Modalidad	Presencial
Grupo	A1
Docente asesor	Dr. Jhonatan Isaac Vargas Huaman

Declaro que la información que utilizaré para el desarrollo de mi trabajo de investigación titulado "Modelo predictivo en la detección de cáncer en mamografías, Lima 2024", **son imágenes obtenidas de casos anónimos de archivos médicos, cumpliendo con las regulaciones de privacidad**; por tanto, no requiero tener la autorización de las personas ni institución correspondiente. Asumo la responsabilidad de la veracidad de lo expuesto.

Lima, 6 de abril del 2024

Firma : 
DNI : 44180515



Huella digital

Anexo 9. Otras evidencias

METODOLOGÍA DE IMPLEMENTACIÓN

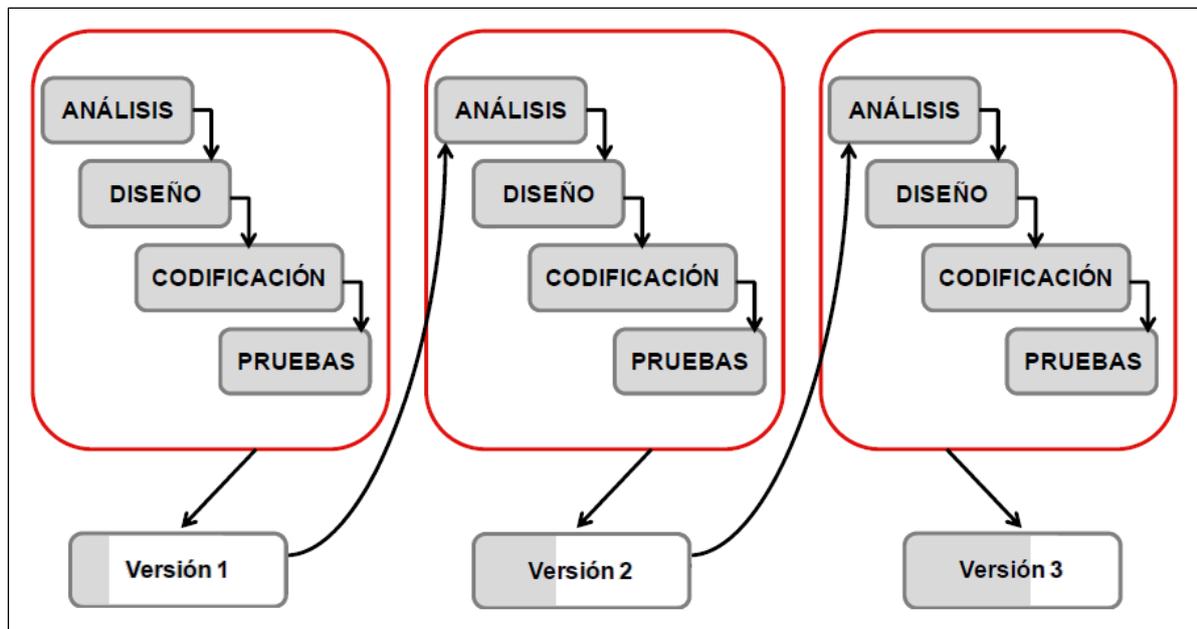
Aunque, explícitamente no forma parte de los objetivos de esta investigación, se desarrolló un prototipo software, denominado “MammoAnalyzer”, para la detección y diagnóstico de cáncer en mamografías.

El prototipo se compone de tres (3) módulos, el primero dedicado al análisis de las mamografías (búsqueda y extracción de características de las imágenes), el segundo dedicado a la evaluación de lesiones (entrenamiento del algoritmo AIRS para el reconocimiento de las características extraídas en el módulo anterior), y el último, para otras tareas (utilidades como manejo de listas, copiado de archivos, etc.).

El prototipo fue codificado en lenguaje Java (utilizando específicamente Java™ SE Development Kit 20.0.2 y la plataforma Apache Netbeans IDE 21), y aplicando un enfoque Incremental, es decir, en cada iteración se fueron agregando funcionalidades o mejoras, para obtener una versión parcial y cada vez más refinada del software en el tiempo.

Figura 1

Ciclo de vida del software con el enfoque incremental



Fuente: autoría personal.

El enfoque incremental divide el ciclo de desarrollo en 4 partes: Análisis, Diseño, Codificación y Pruebas, las cuales permiten generar software operativo de forma rápida y admisible, siendo particularmente útil cuando no se cuenta con una dotación de personal suficiente.

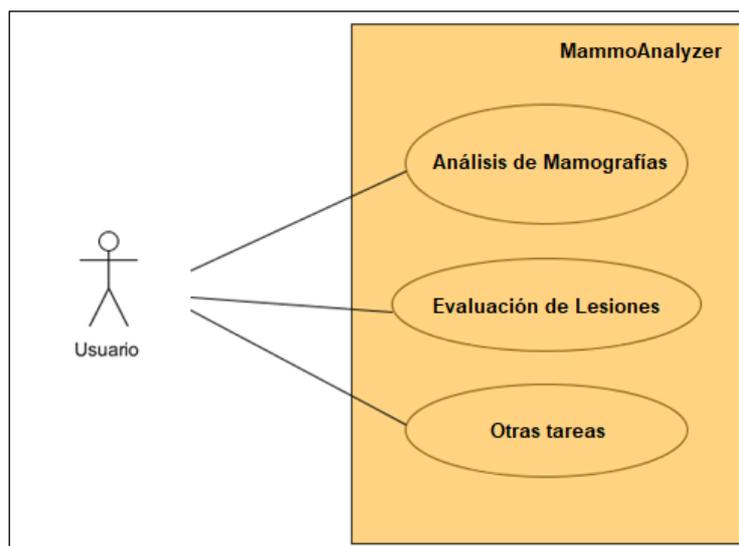
ANÁLISIS

Durante la etapa de Análisis, se utilizó el lenguaje de modelado unificado (UML) para la elaboración de diferentes diagramas que proporcionen una mejor comprensión del diseño, arquitectura del código y la implementación del software.

- Inicialmente se requerían dos módulos, uno para la búsqueda y extracción de información desde las mamografías, y otro para la evaluación de dicha información. Durante el desarrollo del prototipo, la idea de tener un tercer módulo para “utilidades” fue tomando forma, por lo cual, también fue implementado.
- Por otro lado, el prototipo debía adaptarse exclusivamente para procesar la información contenida en las mamografías del Repositorio Digital de Cáncer de Mama (BCDR), esto es, al formato de las imágenes, estructura del archivo de segmentaciones de las lesiones realizadas por los especialistas, etc.

Figura 2

Diagrama general de Casos de uso



Fuente: autoría personal.

Especificación de Casos de Uso

Nombre:	Análisis de Mamografías (módulo 1)
Autor:	Luis José Meléndez Ruiz
Fecha:	01 de abril del 2024
Descripción:	<p>Este caso de uso describe la ejecución del análisis de mamografías, que consiste en la extracción de características de forma y textura del tejido sano y lesiones existentes en las imágenes del BCDR.</p>
Actores:	Usuario
Precondiciones:	<p>El usuario debe tener acceso al archivo de segmentación de lesiones y a las mamografías del BCDR.</p>
Flujo Normal:	<ol style="list-style-type: none">1. El usuario hace clic en el botón “examinar” de la opción “1. Archivo de segmentación de lesiones (Outlines)” para buscar y seleccionar el citado archivo, en formato csv.2. El usuario hace clic en el botón “examinar” de la opción “2. Carpeta con mamografías” para buscar y seleccionar la carpeta correspondiente.3. El usuario hace clic en el botón “Siguiente >” para continuar al segundo panel, donde verifica que la información de cada segmentación realizada por el radiólogo, se muestre correctamente en columnas y filas. Caso contrario, puede modificar los parámetros disponibles en el panel.4. Verificada la información, el usuario hace clic en el botón “Siguiente >” para continuar al tercer panel, donde relaciona las columnas que corresponden a los campos: nombre de archivo, coordenadas en X, coordenadas en Y, así como, la clasificación de las lesiones.5. El usuario hace clic en el botón “Siguiente >” para continuar al cuarto y último panel, donde visualiza las mamografías e inicia la extracción de características.6. Finalizada la extracción, las características se almacenan en un archivo llamado “dataset.txt” en la carpeta con mamografías.7. Para la detección, el usuario tendrá que activar el checkbox “Habilitar detección”, y proceder a hacer clic en el botón “examinar” de la opción “3. Resultado de análisis previo” para buscar y seleccionar el archivo “dataset.txt”. Una vez hecho esto, el flujo es el mismo desde el punto 4, con la excepción de que no solo se extraerán las características de las regiones, sino que el sistema asignará una clasificación, previo entrenamiento del modelo con la información del dataset.
Flujo Alternativo:	<p>El usuario puede cambiar el paso 1 por el 2, sin alterar los resultados.</p>

Postcondiciones:

Ninguna.

Nombre:	Evaluación de Lesiones (módulo 2)
Autor:	Luis José Meléndez Ruiz
Fecha:	01 de mayo del 2024

Descripción:

Este caso de uso describe la ejecución de la evaluación de lesiones. Se debe entrenar al algoritmo AIRS con la información previamente obtenida del análisis de mamografías, para que aprenda del pronóstico dado por el radiólogo. La implementación permite realizar una validación cruzada con K iteraciones, o bien disponer el porcentaje destinado al entrenamiento del AIRS y prueba del modelo predictivo resultante.

Actores:

Usuario

Precondiciones:

El usuario debe tener acceso al archivo "dataset.txt", resultado del análisis de las mamografías.

Flujo Normal:

1. El usuario hace clic en el botón "Examinar", de la opción "1. Resultado del Análisis Previo" y selecciona el archivo "dataset.txt".
2. El usuario elige una de dos opciones disponibles para ejecutar la evaluación de lesiones:
 - "Mediante Validación Cruzada" con K=10 iteraciones.
 - "Utilizando la información obtenida del análisis en un..." 70%
3. Para la opción "Mediante Validación Cruzada", K por defecto es igual a 10, pero el usuario puede escoger valores entre 2 y 20; mientras que para la opción "Utilizando la información obtenida del análisis en un...", 70% es el porcentaje por defecto, pero el usuario puede escoger entre 10 y 90 por ciento.
4. Al hacer clic en el botón "Siguiendo >" el usuario inicia la evaluación de las lesiones dependiendo de la opción escogida.
5. Finalmente, el software muestra al usuario un resumen general de las estadísticas en términos de lesiones clasificadas correcta o incorrectamente; o en su defecto una lista

<p>detallada lesión por lesión comparando el pronóstico del radiólogo con la clasificación dada por el algoritmo inteligente.</p>
<p>Flujo Alternativo: No aplica.</p>
<p>Postcondiciones: Ninguna.</p>

Nombre:	Otras tareas (módulo 3)
Autor:	Luis José Meléndez Ruiz
Fecha:	15 de mayo del 2024
<p>Descripción: Este caso de uso describe la ejecución de tareas secundarias que se pueden realizar con el conjunto de mamografías del BCDR.</p>	
<p>Actores: Usuario</p>	
<p>Precondiciones: El usuario debe tener acceso a las mamografías del BCDR.</p>	
<p>Flujo Normal:</p> <ol style="list-style-type: none"> 1. El usuario selecciona entre las opciones: <ul style="list-style-type: none"> - “Crear lista de mamografías existentes en una carpeta” - “Extraer mamografías de una carpeta mediante una lista [recreando la estructura de directorios]” - “Dividir contenido de la lista en N partes [de forma aleatoria]” - “Dividir aleatoriamente el contenido de la lista. Una parte debe ser el X%”. 2. El usuario hace clic en el botón “Examinar”, en la opción “2. Seleccione la carpeta o archivo, según corresponda”, para buscar la carpeta de las imágenes o la lista, en caso ésta ya exista. 3. El usuario inicia la tarea respectiva al hacer clic en el botón “Siguiente”. 	

4. El usuario espera hasta que reciba un mensaje en pantalla avisándole que la tarea ha finalizado.

Flujo Alternativo:

No aplica.

Postcondiciones:

El usuario obtiene una lista de las mamografías o en su defecto una o varias carpetas con mamografías seleccionadas según la opción que haya elegido.

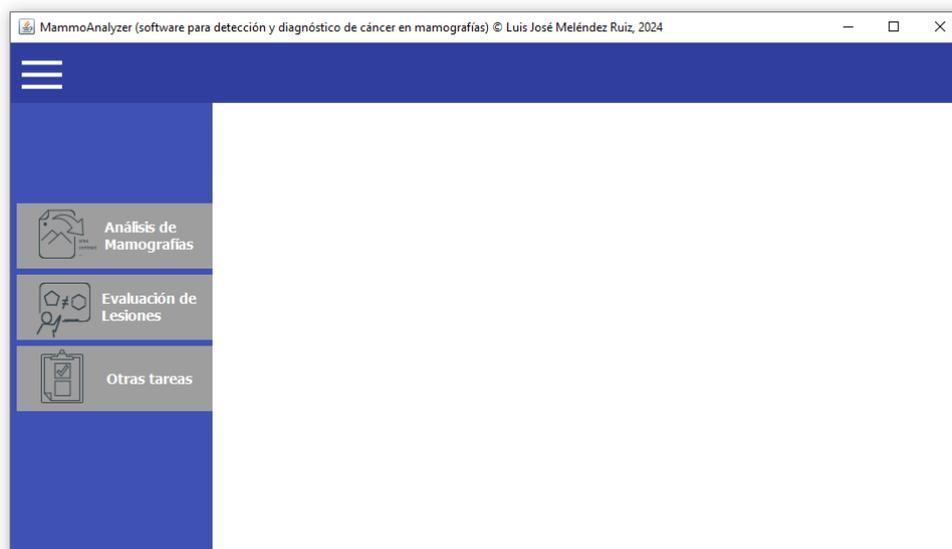
Cabe indicar que el prototipo software no presenta conexión a base de datos, no tiene control de usuarios y tampoco se relaciona con otros sistemas a nivel de red.

DISEÑO

- La interfaz principal contempló un menú desplegable ubicado al lado derecho para los botones de acceso a cada módulo, mientras que la zona ubicada a la izquierda debía contener los paneles para los diferentes componentes (campos y áreas de texto, cajas de listas, entre otros) organizados según las necesidades emergentes.

Figura 3

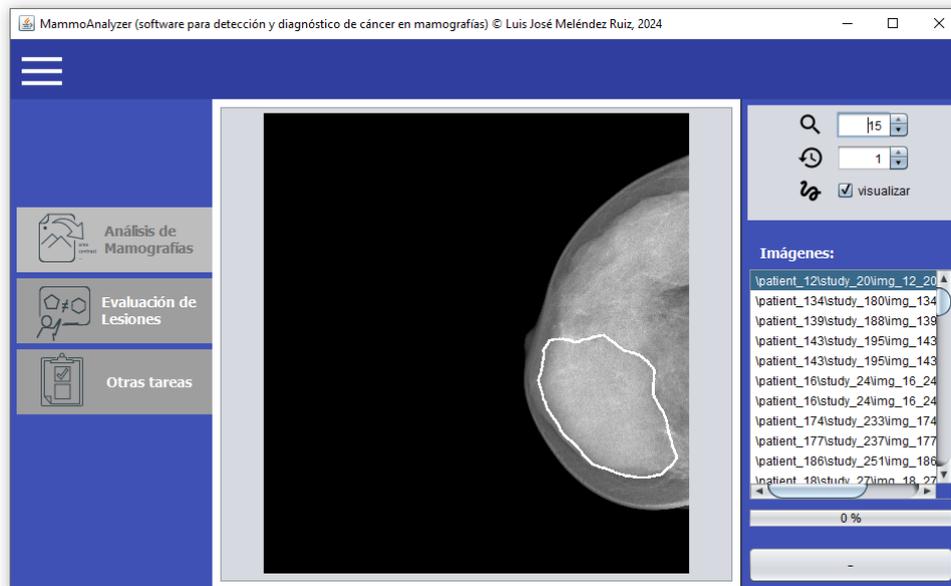
Interfaz principal del software



Fuente: autoría personal

Figura 4

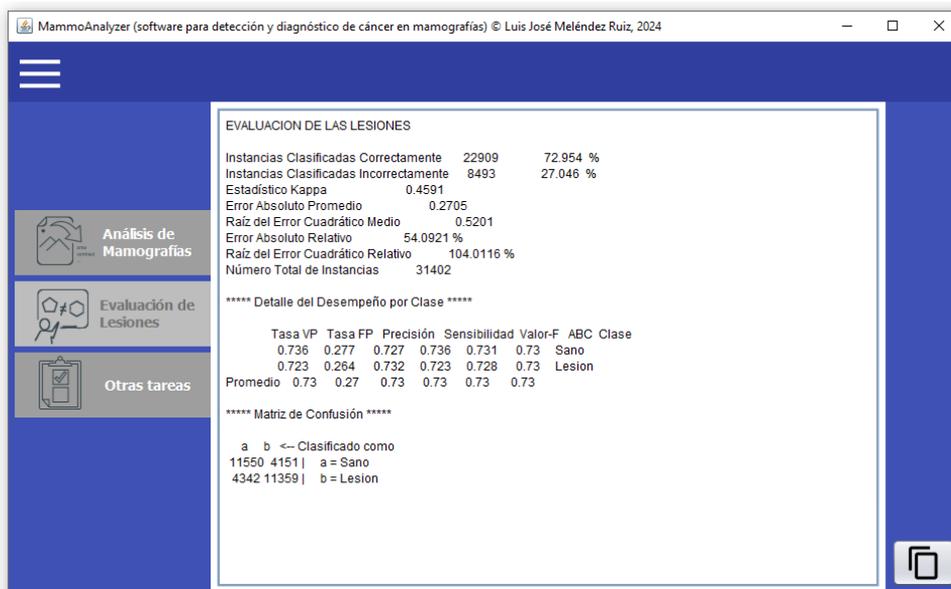
Panel del módulo Análisis de Mamografías



Fuente: autoría personal

Figura 5

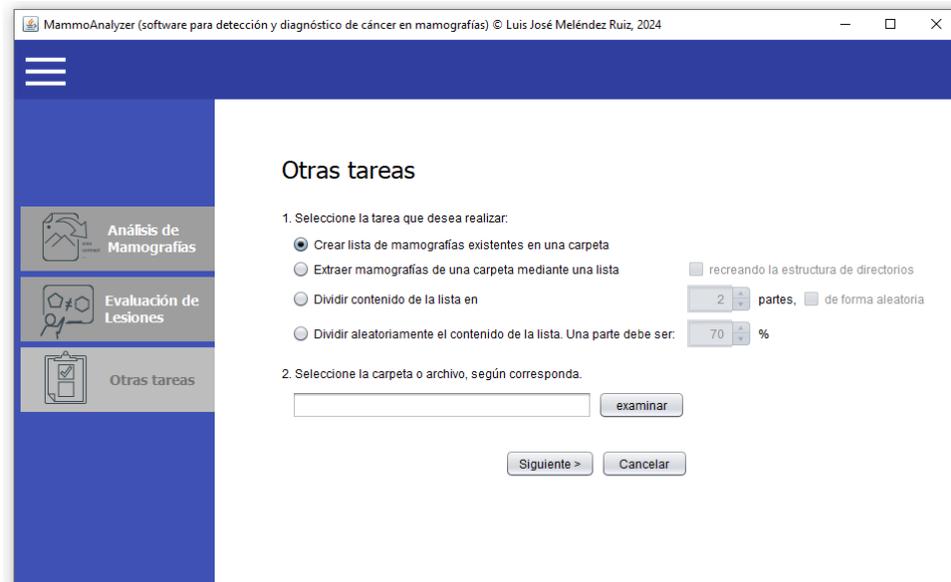
Panel del módulo Evaluación de Lesiones



Fuente: autoría personal

Figura 6

Panel del módulo Otras tareas



Fuente: autoría personal.

CODIFICACIÓN

Para efectos de recolectar los datos del pretest, se desarrolló una red neuronal artificial, como algoritmo tradicionalmente utilizado en la práctica clínica para el diagnóstico de cáncer de mama.

La red neuronal fue codificada en lenguaje Python versión 3.10.13, utilizando el intérprete Spyder versión 5.5.1, de la plataforma Anaconda. Se crearon funciones para la carga del dataset, la configuración de la red neuronal, el entrenamiento del modelo predictivo, carga y descarga del modelo, cálculo de indicadores a nivel de todo el conjunto de imágenes y también, por mamografía.

La arquitectura de la red corresponde a un Perceptrón multicapa, con todas sus neuronas conectadas, estableciéndose 13 neuronas como entradas, una por cada característica extraída de las lesiones y 1 neurona para la asignación de la clase; además de una capa oculta con 64 neuronas con activación de tipo relu, y sólo 1 neurona en la capa de salida con activación de tipo sigmoide. La configuración de los pesos de la red se estableció como aleatoria, utilizando el algoritmo de optimización Adam para el control de la tasa de error y aprendizaje.

Figura 7

Código de red neuronal artificial tipo Perceptrón Multicapa

```
ANN.py x
1 import pandas as pd
2
3 import tensorflow as tf
4 from tensorflow.keras.models import Sequential
5 from tensorflow.keras.layers import Dense
6
7 from sklearn.preprocessing import StandardScaler
8 from sklearn.metrics import confusion_matrix, accuracy_score
9
10 import matplotlib.pyplot as plt
11 import seaborn as sns
12
13 from collections import defaultdict
14
15 def cargar_datos(csv_file, separador=',', tiene_encabezado=True):
16     if tiene_encabezado:
17         data = pd.read_csv(csv_file, sep=separador)
18     else:
19         # Si no hay encabezados, especificar los nombres de las columnas
20         columnas = ['nombre_imagen', 'característica_1', 'característica_2', 'característica_3',
21                    'característica_4', 'característica_5', 'característica_6', 'característica_7',
22                    'característica_8', 'característica_9', 'característica_10', 'característica_11',
23                    'característica_12', 'característica_13', 'clase']
24         data = pd.read_csv(csv_file, sep=separador, names=columnas)
25
26     # Separar los datos en características (X) y etiquetas (y)
27     X = data.drop(['nombre_imagen', 'clase'], axis=1) # características sin clase y nombre_imagen
28     y = data['clase'] # etiquetas
29     nombre_imagen = data['nombre_imagen'] # nombres de las imágenes
30
31     # Normalizar características
32     scaler = StandardScaler()
33     X_normalized = scaler.fit_transform(X)
34
35     return X_normalized, y, nombre_imagen
36
37 def configurar_red_neuronal(input_shape):
38     model = Sequential([
39         Dense(64, activation='relu', input_shape=(input_shape,)),
40         Dense(1, activation='sigmoid')
41     ])
42     model.compile(optimizer='adam',
43                  loss='binary_crossentropy',
44                  metrics=['accuracy'])
45     return model
46
47 def entrenar_modelo(X_train, y_train, epochs=10, batch_size=32, guardar_modelo='modelo.h5'):
48     model = configurar_red_neuronal(X_train.shape[1])
49     history = model.fit(X_train, y_train, epochs=epochs, batch_size=batch_size, verbose=1)
50
51     acc = history.history['accuracy']
52     loss = history.history['loss']
53     epochs = range(len(acc))
54     plt.plot(epochs, acc, 'r', 'Training Accuracy')
55     plt.plot(epochs, loss, 'b', 'Training Loss')
56     plt.title('Training Accuracy')
57     plt.show()
58
59     model.save(guardar_modelo)
60     return model
61
62 def cargar_modelo(ruta_modelo):
63     model = tf.keras.models.load_model(ruta_modelo)
64     return model
65
66 def evaluar_modelo(model, X_test, y_test):
67     y_pred = model.predict(X_test)
68     y_pred = (y_pred > 0.5) # convertir probabilidades a etiquetas binarias
69     cm = confusion_matrix(y_test, y_pred)
70     accuracy = accuracy_score(y_test, y_pred)
71     return cm, accuracy
72
```

```

72
73 def evaluar_modelo_por_imagen(model, X_test, y_test, nombre_imagen):
74     # Obtener las predicciones del modelo
75     y_pred = model.predict(X_test)
76     y_pred = (y_pred > 0.5).astype(int) # convertir probabilidades a etiquetas binarias
77     cm = confusion_matrix(y_test, y_pred)
78     accuracy = accuracy_score(y_test, y_pred)
79
80     # Agrupar predicciones y etiquetas por nombre imagen
81     resultados_por_imagen = defaultdict(lambda: {'vp': 0, 'vn': 0, 'fp': 0, 'fn': 0})
82     for i in range(len(X_test)):
83         imagen = nombre_imagen[i]
84         prediccion = y_pred[i]
85         etiqueta = y_test[i]
86
87         if prediccion == 1 and etiqueta == 1:
88             resultados_por_imagen[imagen]['vp'] += 1
89         elif prediccion == 0 and etiqueta == 0:
90             resultados_por_imagen[imagen]['vn'] += 1
91         elif prediccion == 1 and etiqueta == 0:
92             resultados_por_imagen[imagen]['fp'] += 1
93         elif prediccion == 0 and etiqueta == 1:
94             resultados_por_imagen[imagen]['fn'] += 1
95
96     return dict(resultados_por_imagen), cm, accuracy
97
98 def visualizar_matriz_confusion(matriz_confusion):
99     plt.figure(figsize=(6, 6))
100     sns.heatmap(matriz_confusion, annot=True, cmap='Blues', fmt='g', cbar=False,
101                annot_kws={'size': 14}, linewidths=0.5, linecolor='black')
102     plt.xlabel('Predicted Labels')
103     plt.ylabel('True Labels')
104     plt.title('Confusion Matrix')
105     plt.show()
106

```

```

106
107 def guardar_indicadores(cm, accuracy, guardar_indicadores='indicadores.txt'):
108     vp, fp, fn, vn = cm.ravel()
109     error_rate = 1 - accuracy
110
111     with open(guardar_indicadores, 'w') as f:
112         f.write(f'Verdaderos Positivos: {vp}\n')
113         f.write(f'Falsos Negativos: {fn}\n')
114         f.write(f'Falsos Positivos: {fp}\n')
115         f.write(f'Verdaderos Negativos: {vn}\n')
116         f.write(f'Exactitud: {accuracy:.3f}\n')
117         f.write(f'Tasa de error: {error_rate:.3f}\n')
118
119     print('Indicadores guardados correctamente en', guardar_indicadores)
120
121 def guardar_indicadores_por_imagen(resultados_por_imagen, archivo_salida='indicadores_por_imagen.txt'):
122     with open(archivo_salida, 'w') as f:
123         f.write('Mamografía;VP;FN;FP;VN;Sensibilidad;Especificidad;Exactitud\n')
124         for nombre_imagen, resultados in resultados_por_imagen.items():
125             vp = resultados['vp']
126             fn = resultados['fn']
127             fp = resultados['fp']
128             vn = resultados['vn']
129             sensibilidad = vp/(vp+fn)*100
130             especificidad = vn/(vn+fp)*100
131             exactitud = (vp+vn)/(vp+fn+fp+vn)*100
132             f.write(f'{nombre_imagen};{vp};{fn};{fp};{vn};{sensibilidad:.3f};{especificidad:.3f};{exactitud:.3f}\n')
133
134     print('Indicadores por imagen guardados correctamente en', archivo_salida)
135

```

```

# Paso 1: Cargar datos de entrenamiento y prueba por separado
# Para archivo sin encabezado
X_train, y_train, nombre_imagenes_train = cargar_datos('Datasets/d02_ann/calificaciones-entrenamiento-25.txt', tiene_encabezado=False)
X_test, y_test, nombre_imagenes_test = cargar_datos('Datasets/d02_ann/calificaciones-muestra-25.txt', tiene_encabezado=False)

# Paso 2 y 3: Entrenar y guardar el modelo
modelo_entrenado = entrenar_modelo(X_train, y_train, epochs=20, batch_size=32, guardar_modelo='calificaciones_modelo_retest_25.h5')

# Paso 4: Cargar un modelo previamente guardado
modelo_cargado = cargar_modelo('calificaciones_modelo_retest_25.h5')

# Paso 5: Evaluar el modelo cargado
# cm, accuracy = evaluar_modelo(modelo_cargado, X_test, y_test)

# Evaluar modelo y Guardar indicadores por imagen
resultados_por_imagen, cm, accuracy = evaluar_modelo_por_imagen(modelo_cargado, X_test, y_test, nombre_imagenes_test)
guardar_indicadores_por_imagen(resultados_por_imagen, archivo_salida='indicadores_calificaciones_retest-25.txt')

# Paso 6: Visualizar la matriz de confusión y guardar indicadores
visualizar_matriz_confusion(cm)
guardar_indicadores(cm, accuracy, guardar_indicadores='indicadores_globales_calificaciones_retest-25.txt')

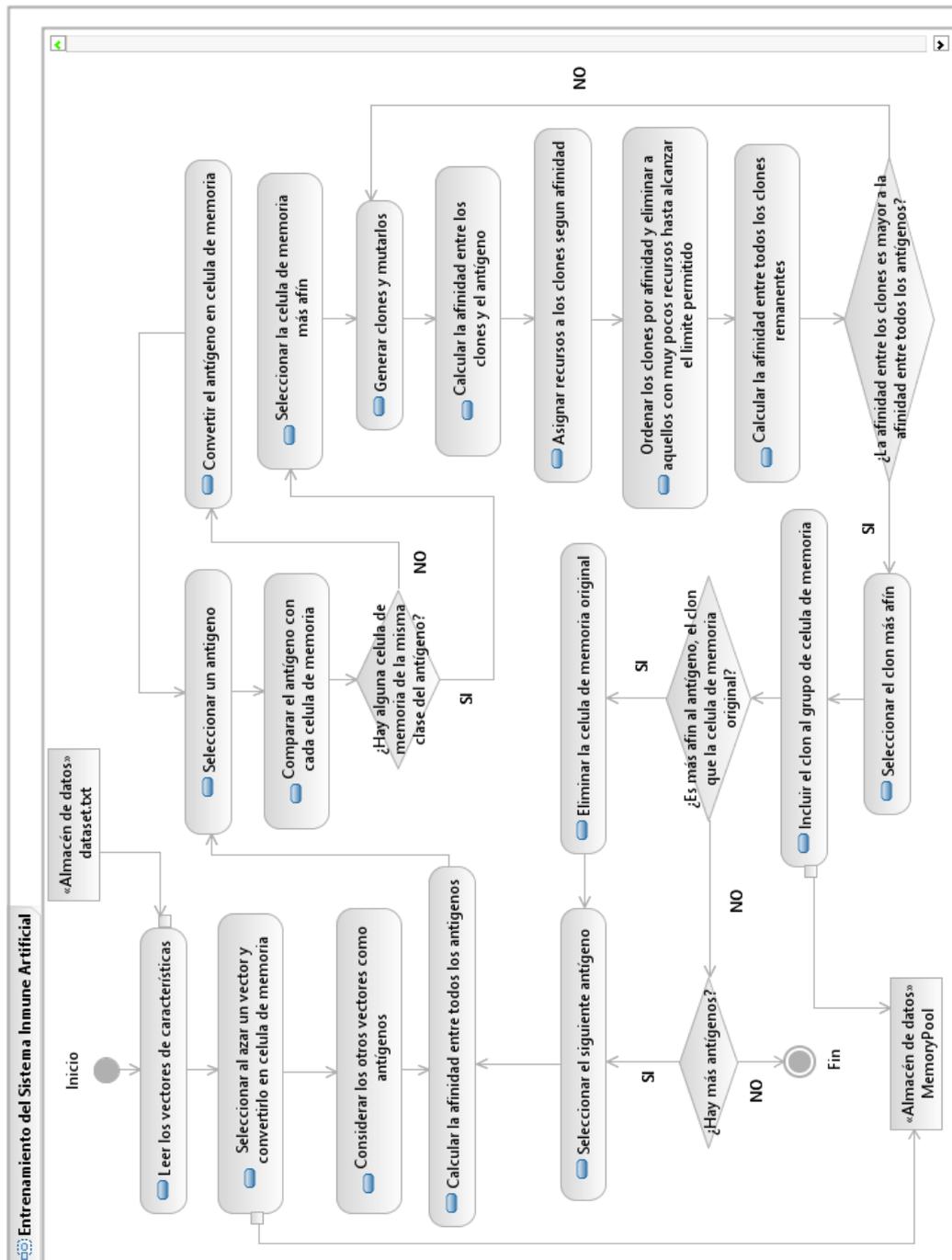
```

Fuente: autoría personal.

Por otro lado, para el post-test, se utilizó el algoritmo AIRS, basado en los principios, modelos y teorías del sistema inmune biológico y cuyo flujo de funcionamiento se presenta en el siguiente diagrama:

Figura 8

Diagrama de flujo del funcionamiento del algoritmo AIRS



Fuente: autoría personal.

PRUEBAS

La idea básica para desarrollar el software fue el siguiente:

Para poder clasificar algo, hay que reconocerlo, y para reconocerlo hay que caracterizarlo. Bien, esas características no son otra cosa más que mediciones. En relación a las imágenes, básicamente se pueden extraer o medir dos tipos de características, aquellas que determinan la forma, y aquellas que determinan la textura.

Por otro lado, en los casos de cáncer de mama, como ya se explicó en el contenido del documento, existen dos tareas posibles de realizar, estas son, la detección y el diagnóstico. En el diagnóstico, las lesiones ya están reconocidas o identificadas visualmente por un especialista, quien define una forma fácil de segmentar y, es por eso, que la mayoría de estudios relacionados a diagnosticar cáncer en imágenes, se enfocan en aprender a clasificar las lesiones utilizando las características de forma. En el caso de la detección, se debe entender que todavía no se ha identificado ninguna lesión y, en consecuencia, no existe forma definida de la cual extraer las características. Por tanto, lo más práctico es recorrer toda la imagen, extrayendo las características de textura presentes en determinadas áreas o regiones.

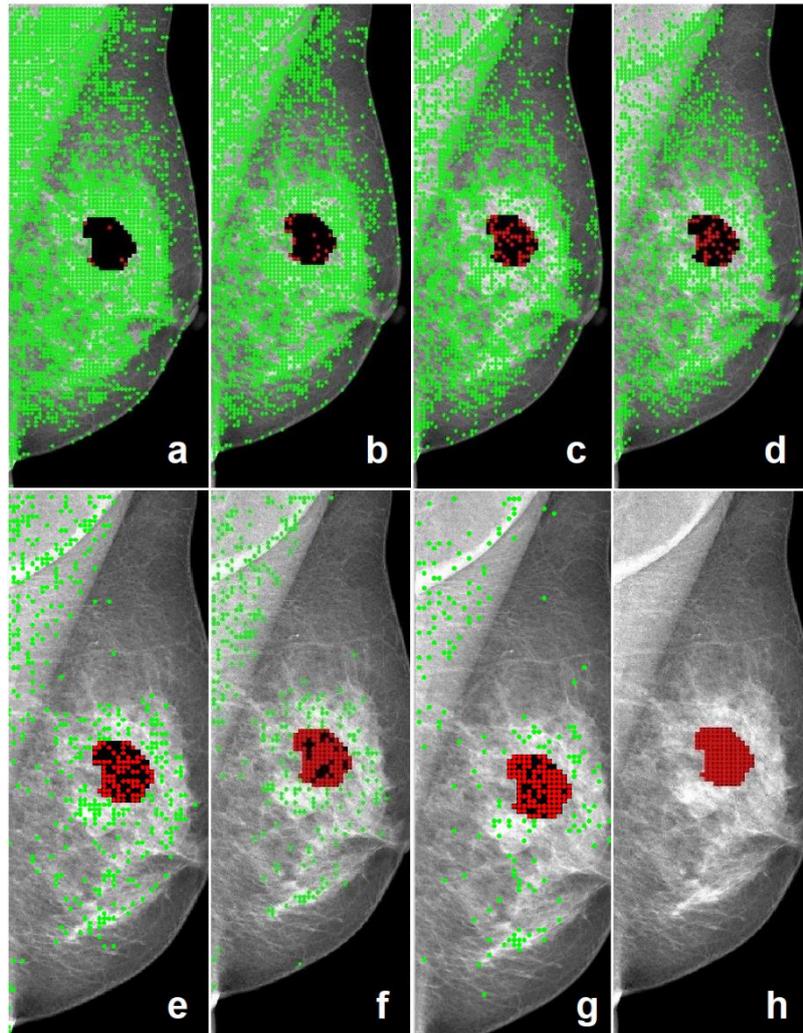
En ese sentido, en esta investigación se procedió a aplicar una rejilla o malla que dividió cada imagen en pequeñas áreas o regiones a fin de que sea más fácil de recorrer y no tome demasiado tiempo calcular las características. El problema con este método fue definir qué tan pequeño o que tan grande debía ser cada región, para que las medidas realizadas permitieran realmente clasificarlas correctamente como regiones sanas o como lesiones.

Por tanto, las primeras pruebas que se ejecutaron, estuvieron relacionadas a extraer las características de las regiones, según diferentes tamaños, partiendo desde áreas de 100 x 100 píxeles hasta llegar a áreas más pequeñas de 20 x 20 píxeles. De esta exploración, los mejores resultados se obtuvieron con regiones de 25 x 25 píxeles. Una vez establecido el tamaño ideal del área para clasificar, se procedió a tratar de mejorar la detección, buscando disminuir o eliminar por completo aquellas

áreas marcadas como falsos positivos y falsos negativos, e incrementar los verdaderos positivos y verdaderos negativos.

Figura 9

Prueba de diferentes configuraciones para mejorar los indicadores



Fuente: autoría personal.

En la figura 9, se aprecian los falsos positivos (en verde), los falsos negativos (en rojo), los verdaderos positivos (en negro) y los verdaderos negativos (color original).

Cada captura corresponde a una configuración de diferentes proporciones de regiones sanas y lesiones utilizadas durante el entrenamiento del algoritmo AIRS. Lo que quiere decir que, para el entrenamiento, cada imagen fue dividida en regiones de

25 x 25 píxeles y a cada región se le asignó una clasificación como sano o lesión, dependiendo de si estaban dentro del perímetro de una lesión identificada y segmentada por un especialista.

Como puede deducirse, la cantidad de regiones sanas superó con creces la cantidad de regiones localizadas dentro del perímetro de una lesión, produciendo un desbalance en el aprendizaje del modelo. Aunque un modelo desbalanceado tiende a clasificar mejor aquello con lo que está más familiarizado, lo más recomendable es optar por una configuración con el menor sesgo posible.

Tabla 19

Configuraciones para entrenamiento del modelo predictivo

Configuración	Entrenamiento			Prueba		
	Orden y forma de carga al Sistema	Regiones Sanas	Lesiones	Sensibilidad (capacidad de clasificar lesiones)	Especificidad (capacidad de clasificar regiones sanas)	Exactitud (capacidad para clasificar correctamente)
A	<i>En grupos, primero las lesiones</i>	X	X	67.64	87.95	87.82
B	<i>En grupos, primero los sanos</i>	X	X	62.85	89.29	89.10
C	<i>En solitario, intercalado</i>	X	X	58.16	90.04	89.83
D	<i>En grupos, primero las lesiones</i>	2X	X	50.71	91.72	91.43
E	<i>En grupos, primero las lesiones</i>	30%	100%	17.49	98.90	98.34
F	<i>En grupos, primero las lesiones</i>	50%	100%	11.31	99.28	98.66
G	<i>En grupos, primero las lesiones</i>	75%	100%	6.02	99.33	97.11
H	<i>En grupos, primero las lesiones</i>	100%	100%	0.00	100.00	99.24

Fuente: autoría personal.

Por tanto, para la obtención del modelo final, se optó finalmente por utilizar la configuración A, es decir, se utilizó un dataset balanceado, con el mismo número de regiones sanas y lesiones, las que fueron cargadas al algoritmo AIRS en grupos, ingresando primero al entrenamiento del modelo, el grupo de lesiones.

Una vez entrenado el algoritmo AIRS, se obtuvo el modelo predictivo propuesto y se ejecutó el post-test. Los resultados se encuentran en las fichas de registro del Anexo 2.