

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMÁTICA
POSTGRADO EN INFORMÁTICA



TESIS MAGISTER SCIENTIARUM

**PROGRAMA DE MAESTRIA EN GERENCIA ESTRATÉGICA DE
TECNOLOGÍAS Y SISTEMAS DE INFORMACIÓN
(GETSI) VERSIÓN 4 GESTIÓN 2010 - 2011**

**MODELO DE CLASIFICACIÓN DE NÓDULOS
TIROIDEOS A TRAVÉS DE IMÁGENES DE
CENTELLOGRAFÍA UTILIZANDO TECNOLOGÍAS
DE MACHINE LEARNING**

POR: Lic. Bladimir Rodolfo Calcina Choque

TUTOR: M. Sc. Rosa Flores Morales

La Paz – Bolivia

2021

Dedicatoria

A mi mamá y a mis hermanas Silvia y Jahel.

Agradecimientos

A toda mi familia por la paciencia que tuvieron y apoyo que me ofrecieron, a mi tutora Rosa Flores por todos los consejos y orientación que me brindo, a los amigos y amigas que me alentaron y a mi pequeño Patán.

RESUMEN

Los avances tecnológicos han permitido que el área de *machine learning* del campo de inteligencia artificial emerjan y permitan brindar nuevas soluciones a diferentes problemas. Siendo la medicina una de las ciencias que está impulsando estos avances.

Mediante información cedida por el Instituto Nacional de Medicina Nuclear (INAMEN), se ha desarrollado un modelo para clasificación de imágenes de la glándula tiroides tomadas a través de centellografía para coadyuvar en el diagnóstico de glándulas tiroides. La tiroides es una glándula muy importante para el funcionamiento del cuerpo, puesto que regula el funcionamiento de los diferentes órganos a través de la segregación de hormonas tiroideas. Para el desarrollo del presente trabajo se utilizó redes neuronales convolucionales con un total de 5.698 imágenes para el entrenamiento y pruebas, las imágenes fueron extraídas de los informes médicos y al final se ha llegado a un nivel de precisión del 70% en todos los casos como entrenamiento y evaluación. Este porcentaje fue considerado como aceptable por parte del especialista médico.

Keywords: Machine Learning, Redes Neuronales Convolucionales, Centellografía, clasificación de imágenes.

ÍNDICE GENERAL

ÍNDICE DE TABLAS.....	1
ÍNDICE DE FIGURAS.....	2
INTRODUCCIÓN	4
CAPITULO I.....	7
ASPECTOS GENERALES	7
1.1 ANTECEDENTES	7
1.2 PLANTEAMIENTO DEL PROBLEMA.....	9
1.3 FORMULACIÓN DEL PROBLEMA DE INVESTIGACIÓN	10
1.4 PLANTEAMIENTO DE OBJETIVOS.....	10
1.4.1. OBJETIVO GENERAL	10
1.4.2. OBJETIVOS ESPECÍFICOS	10
1.5 PLANTEAMIENTO DE LA HIPÓTESIS	10
1.5.1 VARIABLE.....	11
1.5.2 OPERACIONALIZACIÓN DE VARIABLES	11
1.6 DISEÑO METODOLÓGICO.....	12
TIPO DE INVESTIGACIÓN.....	13
TÉCNICAS DE INVESTIGACIÓN.....	13
POBLACIÓN Y MUESTRA DE ESTUDIO	15
CAPÍTULO II	16
MARCO TEÓRICO	16
2.1 CONCEPTOS MÉDICOS.....	16
2.1.1 GLÁNDULA TIROIDES	16
2.1.2 NÓDULO TIROIDEO.....	17
2.1.3 CENTELLOGRAFÍA.....	18
2.1.4 NÓDULO FRÍO	21
2.1.5 NÓDULO CALIENTE	22
2.1.6 NÓDULOS INDETERMINADOS.....	23
2.2 INTELIGENCIA ARTIFICIAL.....	23
2.2.1 APRENDIZAJE AUTOMÁTICO	24
2.2.2 APRENDIZAJE SUPERVISADO	24

2.2.3	REDES NEURONALES ARTIFICIALES.....	25
2.2.4	MODELO DE RED NEURONAL DE MCCULLOCH-PITTS.....	26
2.2.5	FUNCIÓN DE ACTIVACIÓN.....	27
2.2.6	REDES NEURONALES CONVOLUCIONALES.....	29
2.3	HERRAMIENTAS PARA DESARROLLO DEL MODELO.....	32
2.3.1	TENSORFLOW.....	32
2.3.2	KERAS.....	33
2.3.3	ANACONDA NAVIGATOR.....	33
2.3.4	MATRIZ DE CONFUSIÓN.....	33
2.4	PRUEBA DE HIPÓTESIS.....	35
2.4.1	ESTADÍSTICO DE ANDERSON-DARLING.....	36
CAPÍTULO III.....		37
MARCO PRÁCTICO.....		37
3.1	PROCEDIMIENTO PARA EL DIAGNÓSTICO DE NÓDULO TIROIDEO	37
3.1.1	PROCEDIMIENTO DE TOMA DE IMÁGENES.....	37
3.2	RECOPIACIÓN DE INFORMACIÓN.....	38
3.3	DESARROLLO DE MODELO DE CLASIFICACIÓN.....	43
3.3.1	PRE-PROCESAMIENTO.....	45
3.3.2	DESARROLLO DEL PROTOTIPO.....	47
3.3.2.1	ENTRENAMIENTO DEL PROTOTIPO.....	51
3.3.2.2	MODELO DE CLASIFICACIÓN DE IMÁGENES CON CNN.....	57
CAPÍTULO IV.....		59
RESULTADOS.....		59
4.1	EVALUACIÓN DEL MODELO.....	59
4.2	PREDICCIÓN.....	67
4.3	PRUEBA DE HIPÓTESIS.....	70
CAPÍTULO V.....		73
CONCLUSIONES Y RECOMENDACIONES.....		73
5.1	CONCLUSIONES.....	73
5.2	RECOMENDACIONES.....	74
ANEXOS.....		76
ANEXO A - OBTENCIÓN DE IMÁGENES DE DOCUMENTOS.....		77

ANEXO B - ORGANIZACIÓN DE IMÁGENES	78
ANEXO C - LIMPIEZA DE DATOS	81
ANEXO D - AJUSTE DE IMÁGENES	82
ANEXO E - PRUEBAS DE PREDICCIÓN	83
BIBLIOGRAFÍA.....	84

ÍNDICE DE TABLAS

Tabla 1 Valores de exactitud de modelos	8
Tabla 2 Valores de precisión de modelos	8
Tabla 3 Operacionalización de variable.....	11
Tabla 4 Descripción de indicadores de precisión en el diagnóstico	12
Tabla 5 Probabilidad que un nódulo tiroideo sea canceroso	20
Tabla 6 Diagnóstico de nódulos tiroideos.....	22
Tabla 7 Matriz de confusión	34
Tabla 8 Cantidad de imágenes utilizadas.....	47
Tabla 9 Comparación CNN vs SVM	48
Tabla 10 Precisión calculada por scikit-learn	56
Tabla 11 Cálculo de métricas con scikit-learn.....	57
Tabla 12 Etiquetas con cantidades de elementos.....	59
Tabla 13 Matriz de confusión	60
Tabla 14 Matriz de confusión obtenida scikit-learn	61
Tabla 15 Cálculo de falsos positivos	62
Tabla 16 Cálculo de falsos negativos	62
Tabla 17 Valores que no se tomaran en cuenta para bocio nodular	63
Tabla 18 Resultados de verdaderos negativos por clase.....	63
Tabla 19 Valores de sensibilidad, especificidad y precisión por etiqueta	64
Tabla 20 Comparación de predicciones utilizando imágenes.....	70

ÍNDICE DE FIGURAS

Figura 1. Glándula tiroidea	5
Figura 2. Glándula tiroidea	17
Figura 3. Nódulo tiroideo.....	18
Figura 4. Modelo aprendizaje supervisado	25
Figura 5. Neurona biológica	26
Figura 6. Estructura de una red neuronal	26
Figura 7. Neurona artificial.....	27
Figura 8. Función Threshold.....	28
Figura 9. Función sigmoide	28
Figura 10. Función rectificadora.....	29
Figura 11. Corteza visual del cerebro	30
Figura 12. Estructura de Red Neuronal Convolutiva.....	31
Figura 13. Ecuación de convolución.....	31
Figura 14. Convolución	31
Figura 15. Reducción o pooling.....	32
Figura 16. Grafica normal standard	35
Figura 17. Procedimiento de toma de imágenes de centellografía.	38
Figura 18. Informe de diagnóstico de nódulo frío	40
Figura 19. Informe de diagnóstico de nódulo frío	40
Figura 20. Informe de diagnóstico de bocio difuso	41
Figura 21. Informe de diagnóstico de bocio difuso	41
Figura 22. Tiroides normal	42
Figura 23. Metodología de construcción de modelo de machine learning	43
Figura 24. Modelo de clasificación de imágenes.....	44
Figura 25. Imagen descartada	45
Figura 26. Imagen con recuadro o borde	46
Figura 27. Nombre de radiofármaco en la imagen	46
Figura 28. Cantidad de nódulos procesados	47

Figura 29. Resultados de entrenamiento por época	51
Figura 30. Exactitud, pérdida de entrenamiento y validación de 50 épocas.....	52
Figura 31. Exactitud, pérdida de entrenamiento y validación de 80 épocas.....	52
Figura 32. Exactitud, pérdida de entrenamiento y validación de 150 épocas.....	53
Figura 33. Exactitud, pérdida de entrenamiento y validación de 250 épocas.....	53
Figura 34. Exactitud, pérdida de entrenamiento y validación de 500 épocas.....	54
Figura 35. Exactitud, pérdida de entrenamiento y validación de 1000 épocas.....	54
Figura 36. Exactitud, pérdida de entrenamiento y validación de 2000 épocas.....	55
Figura 37. Exactitud, pérdida de entrenamiento y validación de 3000 épocas.....	55
Figura 38. Modelo de clasificación de imágenes con CNN	58
Figura 39. Cálculo de precisión con la matriz de confusión.....	66
Figura 40. Cálculo de precisión con la matriz de confusión.....	67
Figura 41. Resultados de predicción del modelo	68
Figura 42. Tiroides con tiroiditis	68
Figura 43. Tiroides con tiroiditis	68
Figura 44. Tiroides con bocio nodular.....	69
Figura 45. Predicción con valores normalizados	71
Figura 46. Valores obtenidos con el método de Anderson-Darling	71
Figura 47. Comparación de valores críticos	72

INTRODUCCIÓN

Hoy en día la tecnología ha avanzado significativamente. Estos avances han permitido superar barreras que se tenían en tiempos pasados, respecto al hardware la capacidad de almacenamiento y procesamiento ya no son un problema. Esto ha generado que áreas como la inteligencia artificial hayan progresado mucho, dejando atrás el estancamiento que sufrió en los años 80 y ahora emergieron tecnologías como *Machine Learning* y *Deep Learning* que permiten automatizar tareas, aumentar la productividad de las empresas; las capacidades de Redes Neuronales para el procesamiento de imágenes, texto, video, etc., será cada vez más frecuente.

Una ciencia que está apostando por el uso de estas nuevas tecnologías es medicina, los pacientes generan diariamente una enorme cantidad de información, esto brinda la oportunidad de desarrollar algoritmos que analicen e interpreten imágenes médicas y otros datos, de esta manera las nuevas herramientas permiten colaborar a los profesionales del área de salud en el diagnóstico y estudio de diferentes enfermedades.

Las enfermedades tiroideas son un problema en la población, porque su presencia genera una disminución en la calidad de vida de las personas, manifestándose a través de cansancio excesivo, desgano, apatía, irritabilidad, subida de peso sin motivo, palpitaciones, dolor de garganta y otros síntomas. La tiroides es una glándula perteneciente al sistema endocrino en forma de mariposa situada en la base del cuello (ver Figura 1). Su función principal es liberar a la sangre hormonas tiroideas, como T3 (triyodotironina) y T4 (tiroxina o tetrayodotironina), muy parecidas entre ellas y cuyo compuesto básico es el aminoácido tirosina. Estas hormonas son vitales porque intervienen en el desarrollo del sistema nervioso y en los procesos cognitivos. Además, regulan el metabolismo y el ritmo que el cuerpo quema calorías para producir energía, por lo que influyen en casi todas las células del organismo y son necesarias para controlar las funciones de casi todos los órganos de nuestro cuerpo (AECAT, 2014).

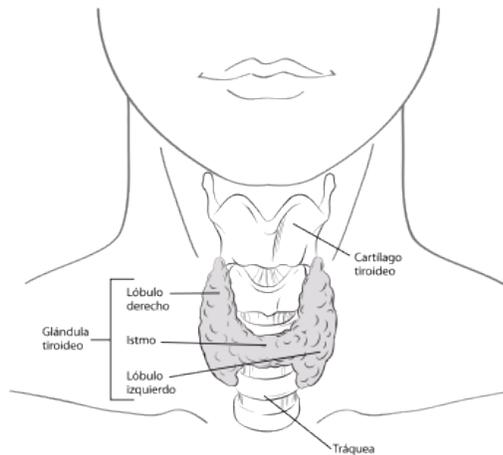


Figura 1. Glándula tiroidea

Fuente: (Brandan, Llanos, Horak, Tannuri, & Rodríguez, 2014)

Las patologías tiroideas más comunes son: el hipertiroidismo (la glándula tiroidea produce más hormonas de lo que el cuerpo necesita), hipotiroidismo (la tiroidea no produce las suficientes hormonas), bocio (agrandamiento de la glándula tiroidea) el cual puede ser difuso, nodular o multinodular y cáncer de tiroides. A las masas o protuberancias en la glándula tiroidea se les llama nódulos tiroideos, la mayoría son benignos, pero alrededor de 2 ó 3 de 20 casos son cancerosos (American Cancer Society, 2016).

En el año 2013 el 9% de la población española presentaba alguna enfermedad de tiroides, el cáncer de tiroides supone el 3,8% del total mundial de casos de cáncer, siendo responsable del 0,3% de la mortalidad, es decir, una de cada 1.000 personas fallecen por cáncer de tiroides (La voz de Galicia, 2014).

La Asociación Española de Cáncer de Tiroides publicó una recopilación de datos y valoración de las enfermedades tiroideas, donde se menciona que: “la incidencia anual estimada de aparición de nódulos tiroideos es del 0,1% y en España existe una población estimada de 46.116.779 personas, se espera que anualmente se desarrollen más de 46.000 nuevos nódulos tiroideos, los cuales deberán ser estudiados para descartar la presencia de malignidad en ellos” (AECAT, 2014)

En Bolivia el periódico “Los Tiempos” realizó una entrevista al Dr. Alberto Requena Urioste médico de la Caja Nacional de Salud, quien publicó un libro sobre la cirugía de tiroides. El galeno menciona entre los datos más destacados que: el 25% de las consultas en endocrinología son por enfermedades de la glándula tiroidea, el 30 y 40% de la población tiene nódulos en la tiroides y que de cada tres pacientes operados uno tiene cáncer (Los Tiempos, 2017).

CAPITULO I

ASPECTOS GENERALES

1.1 ANTECEDENTES

En la revisión bibliográfica efectuada se observó que cada vez se hace más extenso el uso de inteligencia artificial en la búsqueda de soluciones a diferentes problemas que no se pueden resolver con las técnicas de programación tradicional.

En Bolivia se ha encontrado una investigación para el diagnóstico de cáncer tiroideo basado en redes neuronales. En esta tesis se usa proposiciones que siguen un conjunto de reglas, que utilizan una base de conocimientos generada a partir de la experiencia de una especialista y así poder indicar si el paciente puede tener o no cáncer (Chambi, 2014).

En el Hospital Universitario Miguel Servet, Zaragoza – España, se realizó una evaluación de la capacidad de los árboles de clasificación y regresión (CART) para la valoración ecográfica de los nódulos tiroideos malignos, en dicha investigación se obtuvo un 87,5 % de sensibilidad, 98,8 % de especificidad y llegaron a la conclusión que los árboles de clasificación proveen una herramienta simple para la toma de decisión clínica para reducir las punciones de aspiración de aguja fina innecesarias, este es un procedimiento más invasivo (Franco, Pardo, Laborda, & Pérez, 2017).

En otra investigación se efectuaron evaluaciones con diferentes herramientas desarrolladas para el diagnóstico de enfermedades de la tiroides, realizando una clasificación de imágenes médicas mediante Redes Neuronales Artificiales, *Support Vector Machines*, lógica difusa, Árboles de decisión. Para el estudio se empleó imágenes de ultrasonido porque son baratas, no invasivas y fáciles de usar, el objetivo fue clasificar los nódulos en benignos/malignos y reducir operaciones invasivas como biopsias y punciones de aspiración de punta fina, en el estudio se pudo verificar que el uso de estas herramientas permitieron hacer el diagnóstico con una exactitud por encima del 80 por ciento (Agustin & Babu, 2013). De la evaluación se obtuvo los valores de exactitud que se muestra en la tabla 1:

Tabla 1
 Valores de exactitud de modelos
 Fuente: (Agustin & Babu, 2013)

Método de segmentación	Exactitud (%)
SVM	84,62
FCM-TDM	89,80
KNN (K-Nearest Neighbor)	98,90
Arboles de decisión	96,90

En el Departamento de Ingeniería Médica de la Universidad de Otto-von-Guericke (Alemania), se utilizaron imágenes de obtenidas por ultrasonido, las cuales fueron utilizadas para clasificar en tamaño y forma la tiroides mediante diferentes algoritmos como ser: *Active Countours without Edges (ACWE)*, *GraphCut (GC)*, *Pixel-based Classifier (PBC)*, *Random Forest Classifier (RFC)* y *Convolutional Neural Network (CNN)*, el uso de estos algoritmos permitió hacer una evaluación de la exactitud de la segmentación, tiempo de computo, número de interacciones y robustez, donde se pudo ver que los algoritmos supervisados (RFC y CNN) tienen mejor rendimiento como se muestra en la tabla 2 (Poudel, Illanes, Sheet, & Friebe, 2018).

Tabla 2
 Valores de precisión de modelos
 Fuente: (Poudel, Illanes, Sheet, & Friebe, 2018)

Algoritmo	Exactitud (%)
<i>Active contours without edges (ACWE)</i>	80,0
<i>Graph Cut (GC)</i>	76,5
<i>Pixel-based Classifier (PBC)</i>	67,0
<i>Random Forest Classifier (RFC)</i>	86,2
<i>Convolutional Neural Network (CNN)</i>	87,6

1.2 PLANTEAMIENTO DEL PROBLEMA

Algo muy importante en cualquier país es la salud, razón por la que se van creando centros especializados para el diagnóstico y tratamiento de diferentes enfermedades que aquejan la población en general. En Bolivia se cuenta con el Instituto Nacional de Medicina Nuclear (INAMEN), esta institución emplea material radioactivo con fines de diagnóstico, tratamiento e investigación médico nuclear para atender a la población boliviana.

Según los datos históricos del INAMEN (Murillo², 2018), desde el año 2011 hasta noviembre de la gestión 2018 se han registrado alrededor de 7.000 casos de enfermedades de la tiroides, actualmente se atienden en promedio de 6 a 8 casos nuevos por día, provenientes de cajas de salud, clínicas privadas y diferentes departamentos, lo que representa un carga de trabajo bastante pesada para el personal de la institución. Además, la unidad de endocrinología realiza atención a otras enfermedades y hace seguimiento a otros pacientes, esto genera en algunos casos demora en la entrega de los resultados pasadas las 48 horas (de acuerdo a procedimiento son 24 horas).

El INAMEN cuenta con cámaras GAMMA para la toma de imágenes con centellografía, estas imágenes son utilizadas por el especialista de medicina nuclear para realizar la valoración de manera visual del tipo de nódulo que tiene el paciente y elaborar el diagnóstico de lo observado. Sin embargo, la institución podría contar con una herramienta informática que permita reforzar la precisión de dicho diagnóstico, además de brindar información estadística que genere indicadores cuantitativos de tipologías de enfermedades tiroideas.

La importancia de diagnosticar correctamente los nódulos, es evitar que se vuelvan cancerígenos y que hagan metástasis², si llegan a este punto se habla de invasión del cáncer a otros órganos (por ejemplo las cuerdas vocales), lo que complicaría la salud del paciente.

²Extensión o propagación de la enfermedad

1.3 FORMULACIÓN DEL PROBLEMA DE INVESTIGACIÓN

De los aspectos mencionados, se formula el siguiente problema de investigación:

¿Con qué nivel de precisión un modelo basado en *Machine Learning* puede aportar en el diagnóstico de nódulos tiroideos a partir de imágenes de centellografía de tiroides?

1.4 PLANTEAMIENTO DE OBJETIVOS

1.4.1. OBJETIVO GENERAL

Desarrollar un modelo de clasificación basado en *Machine Learning* que permita coadyuvar en el diagnóstico de nódulos tiroideos a partir de imágenes de centellografía.

1.4.2. OBJETIVOS ESPECÍFICOS

Los objetivos específicos son los siguientes:

- Describir los procedimientos actuales de diagnóstico de nódulos tiroideos.
- Desarrollar un prototipo de software con *Machine Learning* para clasificar los nódulos tiroideos a partir de imágenes de centellografía.
- Evaluar el modelo propuesto para el diagnóstico de nódulos tiroideos.

1.5 PLANTEAMIENTO DE LA HIPÓTESIS

El modelo basado en *Machine Learning* diagnostica con una precisión mayor o igual al 70% el tipo de nódulo tiroideo utilizando imágenes de centellografía.

1.5.1 VARIABLE

La variable única de estudio será: “nivel de precisión en el diagnóstico a partir de las imágenes obtenidas por centellografía”.

1.5.2 OPERACIONALIZACIÓN DE VARIABLES

En la tabla 3 se presenta la operacionalización de la variable de estudio. El nivel de precisión es un concepto más complejo por lo que su evaluación se realizará mediante indicadores de precisión, sensibilidad, especificidad y exactitud.

Tabla 3
Operacionalización de variable

Variable	Concepto	Indicadores
Nivel de precisión en el diagnóstico a partir de las imágenes obtenidas por centellografía	Se considera informes de diagnósticos elaborados por el especialista	Precisión Sensibilidad Especificidad Exactitud

Para una mayor comprensión de los indicadores en la tabla 4 se definen con mayor detalle:

Tabla 4
Descripción de indicadores de precisión en el diagnóstico

Indicador	Definición	Operacionalización
Precisión	Es el grado de proximidad o cercanía de los resultados de diferentes mediciones entre sí (Zita, 2019).	Es la proporción de número de diagnósticos correctos y el total de diagnósticos
Exactitud	Es la cercanía de una medida al valor real (Zita, 2019)	Es la proporción entre los casos positivos reales y todos los casos positivos
Sensibilidad	En medicina se describe que tan bien una prueba puede detectar una enfermedad o afección (INC1, 2019).	Número de casos positivos correctamente identificados
Especificidad	En medicina se refiere al porcentaje de personas que obtienen un resultado negativo de una prueba para una prueba específica (INC2, 2019).	Número de casos negativos correctamente identificados

1.6 DISEÑO METODOLÓGICO

La presente investigación consiste en estimar el nivel de precisión del diagnóstico de nódulos tiroideos a partir de imágenes de centellografía utilizando un modelo de *Machine Learning*, por tal motivo se tomó un diseño no experimental porque no se hará variar la variable de estudio de manera intencional.

La investigación no experimental es sistemática y empírica en la que las variables no se manipulan porque ya han sucedido. Además es longitudinal, ya que se recolectó datos de diferentes momentos o periodos (Fernandez, Fernandez, & Baptista, 2014).

Esta es una investigación cuantitativa, porque seguirá un conjunto de pasos de manera secuencial y estructurada para recopilar y analizar datos, convirtiéndose en una investigación objetiva porque tiende a obtener resultados utilizando funciones matemáticas y estadísticas (Fernandez, Fernandez, & Baptista, 2014).

TIPO DE INVESTIGACIÓN

La investigación aplicada tiene por objetivo la resolución de un determinado problema, buscando aplicar o utilizar los conocimientos de una o varias áreas especializadas, con el propósito de implementar de forma práctica ofreciendo una solución, en este caso se utilizó *Machine Learning* para el desarrollo de una herramienta que coadyuva en el diagnóstico de nódulos tiroides mediante imágenes de centellografía y finalmente proponer un modelo para abordar trabajos similares.

TÉCNICAS DE INVESTIGACIÓN

Para la investigación se utilizó la técnica de observación, donde se revisó las imágenes para identificar información relevante y realizar el registro para un posterior análisis.

- Revisión de Material Bibliográfico, una importante fuente de información se encuentra en la documentación registrada a la fecha, que permite entender el trabajo que se desea realizar. Esta se clasifica de la siguiente manera:
 - Documentos: los artículos de investigaciones realizadas, sobre *Machine Learning*, nódulos tiroideos, libros de endocrinología, informes de

diagnósticos de imágenes tomadas por centellografía, de los cuales podemos mencionar a:

- Informes emitidos por los especialistas de endocrinología del INAMEN desde la gestión 2011 hasta noviembre 2018.
 - Fundamentos de Medicina – Endocrinología (Orrego, 2012).
 - Medicina Nuclear: Los requisitos (Ziessman, O'Malley, & Thrall, 2007).
 - Guía de Aplicaciones Clínicas de Medicina Nuclear (Barragan, 2004).
 - Tesis: Sistema experto para el diagnóstico de cáncer tiroideo basados en redes neuronales (Chambi, 2014).
- Para el desarrollo del prototipo se utilizó los siguientes procesos y herramientas:
 - Sistematización del proceso de diagnóstico.
 - Elección del algoritmo que se adecue de mejor manera a la investigación.
 - Anaconda Navigator, es una suite multiplataforma de código abierto que tiene integrado una serie de aplicaciones y librerías para el desarrollo de aplicaciones de ciencia de datos.
 - Tensorflow, es una plataforma para el aprendizaje automático, que facilita la compilación e implementación de aplicaciones.
 - Keras, es una librería de redes neuronales, que hace más sencillo el desarrollo de modelos de aprendizaje profundo.
 - MS-Paint, editor de imágenes, utilizado para remover textos o manchas que solo generan ruido para el entrenamiento de la red neuronal convolucional.
 - Aplicación de la Metodología. Esta fase se realizó siguiendo los siguientes pasos:
 - Establecer los criterios a ser medidos.
 - Analizar, procesar imágenes, informar y documentar los resultados.

POBLACIÓN Y MUESTRA DE ESTUDIO

La investigación se efectuó sobre 5.698 imágenes que fueron extraídas de los informes elaborados por el personal especializado de la Unidad de Endocrinología del INAMEN donde se tiene la valoración de los centellogramas. Se utilizó un 80% de las imágenes para el entrenamiento del algoritmo de clasificación y 20% para las pruebas de evaluación de la precisión. Para las pruebas de predicción se utilizaron 50 imágenes adicionales que fueron proporcionadas por la institución.

CAPÍTULO II

MARCO TEÓRICO

Para la presente investigación se tuvo que revisar algunos conceptos de medicina relacionados sobre la glándula tiroides, el estudio de imágenes con centellografía, estudios similares al presente trabajo y tecnologías como *machine learning* que permitieron el desarrollo del modelo de clasificación.

2.1 CONCEPTOS MÉDICOS

A continuación se describen los conceptos usados respecto a la glándula en estudio, las patologías que afectan a la tiroides y la técnica de toma de imágenes para el diagnóstico de nódulos tiroideos.

2.1.1 GLÁNDULA TIROIDES

La glándula tiroides recibe el nombre del griego *thureos* que significa escudo o protector, es un órgano con un peso total de 15 a 20 gramos (la mayor parte de él es aportado por el coloide). Posee dos lóbulos principales, localizados en ambos lados del tercio inferior de la tráquea, cada uno de aproximadamente de 2 a 2.5 cm x 2 a 2.5 cm x 4 cm. Tiene un istmo hacia la base y un lóbulo piramidal a partir de él, se localiza discretamente hacia el lóbulo tiroideo izquierdo (figura 2). El lóbulo derecho contiene más tejido vascular y esto explica su usual mayor tamaño y su tendencia a crecer más ante estímulos. La glándula posee una doble inervación automática, adrenérgica de los ganglios cervicales y colinérgica de los nervios vagos. Esta inervación cumple básicamente una función regulatoria del flujo sanguíneo, lo que regula el aporte de la hormona tiroestimulante o tirotrópina (TSH), del yodo y otros sustratos metabólicos para la adecuada función de la glándula. En cada célula tiroidea o unidad funcional celular del folículo (tirocito), se

cumplen diferentes funciones que conducen a la replicación celular y a la síntesis y secreción de las hormonas tiroideas. Este proceso se inicia con la captación de yodo inorgánico (I) y se le denomina hormogénesis intratiroidea (Orrego, 2012; Cooper & Ladenson, 2015).

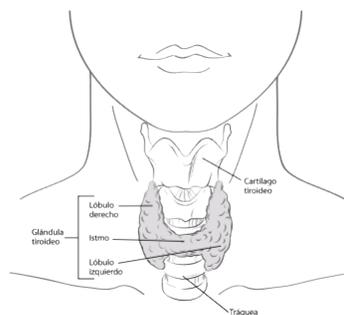


Figura 2. Glándula tiroidea

Fuente: (Brandan, Llanos, Horak, Tannuri, & Rodríguez, 2014)

2.1.2 NÓDULO TIROIDEO

Cuando se habla de un nódulo tiroideo se hace referencia a un crecimiento anormal de las células tiroideas que forman un bulto en la glándula (ver figura 3), algunos pueden crecer lo suficiente como para hacerse visibles, dificultando la respiración o la deglución, cambios en la voz o ronquera. La presencia de esta enfermedad es más común en mujeres que en hombres y se incrementa la probabilidad de que aparezcan con el pasar de los años. Hay pocos nódulos tiroideos cancerosos, pero algunos factores de riesgo (como haber estado expuesto a radiación o antecedentes familiares de cáncer de tiroides) pueden aumentar las probabilidades de que una persona sufra la forma cancerosa de esta condición (Merck, 2017; Clinic, 2017).

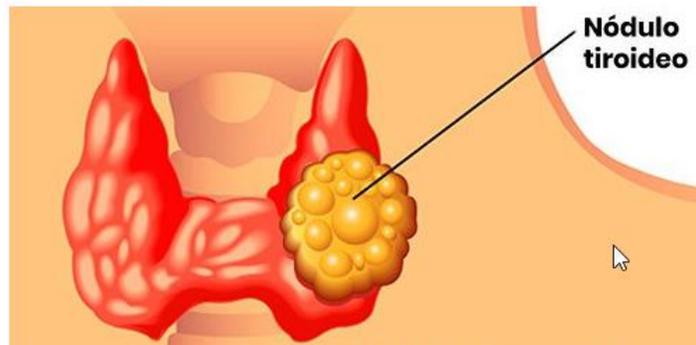


Figura 3. Nódulo tiroideo
Fuente: (Nasser, 2019)

En la mayoría de los casos los nódulos no causan síntoma alguno y a veces se presentan los síntomas mencionados anteriormente. Cuando las células del nódulo producen en exceso hormonas tiroideas, puede asociarse a hipertiroidismo esto genera cambios del estado de ánimo, aumento involuntario de peso, debilidad muscular, aumento de apetito, inflamación del rostro, intolerancia al frío, alteraciones del sueño y del ritmo de los latidos del corazón. Lo más normal es que el nódulo sea detectado por un médico durante un reconocimiento del cuello o al realizar las pruebas de imagen (Merck, 2017).

No se tiene definido exactamente el origen de los nódulos, aunque la falta de yodo en la dieta puede ser causa de su aparición, otro causante es la enfermedad de Hashimoto (afección causada por una reacción del sistema inmunitario contra la glándula tiroides), el desarrollo anormal del tejido de la glándula tiroides (adenoma tiroidea), quiste tiroideo (una cavidad que contiene fluidos o partes sólidas), que suele ser benigno; bocio multinodular y cáncer de tiroides (Barragan, 2004).

2.1.3 CENTELLOGRAFÍA

Para el diagnóstico de los nódulos, se utilizan diferentes técnicas de imagenología en medicina, en el INAMEN se usa centellografía mediante el cual obtienen las imágenes para el respectivo estudio por parte de los especialistas de la mencionada institución.

Según el Instituto Nacional del Cáncer de los Estados Unidos: “centellografía es un procedimiento mediante el cual se obtienen imágenes de las estructuras internas del cuerpo, incluso de las áreas donde hay células cancerosas, esta técnica se usa para diagnosticar, estadificar y vigilar una enfermedad”. Para su estudio se inyecta en una vena o se traga una cantidad pequeña de una sustancia química radioactiva (radionúclido). Los distintos radionúclidos se desplazan en el torrente sanguíneo hasta los diferentes órganos y a través de una máquina con una cámara especial se mueve alrededor de la persona acostada sobre una camilla y así detectar el tipo de radiación que emiten los radionúclidos, a través de una computadora se captura la imagen de las áreas donde se acumulan los radionúclidos, también se llama gammagrafía con radionúclido (Instituto del Cáncer, 2018).

En un centellograma de la tiroides se pueden ver (Barragan, 2004):

- La posición de la tiroides
- Su forma
- Su volumen
- Nódulos en la tiroides
- Funcionamiento de la glándula

Los tipos de centellografía o gammagrafía con radionúclidos son los siguientes:

- Centellografía tiroides con Iodo (I-123, I-131).
- Centellografía tiroides con tecnecio 99.

Una gammagrafía de la tiroides es más sensible que la exploración física para la detección de nódulos y puede determinar el estado funcional de un nódulo detectado durante la exploración física o mediante las técnicas de imagen anatómicas. Desde el punto de vista gammagráfico los nódulos tiroideos pueden clasificarse en nódulos fríos (hipofuncionantes respecto al tejido normal adyacente), calientes (hiperfuncionantes, con supresión del tejido tiroideo extranodular), templados (mayor captación que el tejido

normal adyacente pero sin supresión de tejido tiroideo extranodular) o indeterminados (palpables, pero no visibles en la gammagrafía). En la gammagrafía también es posible observar la existencia de nódulos múltiples. Esta clasificación puede proporcionar una idea de riesgo relativo de que el nódulo se trate de un proceso canceroso, tal como se observa en la tabla 5 (Ziessman, O'Malley, & Thrall, 2007).

Tabla 5
Probabilidad que un nódulo tiroideo sea canceroso
 Fuente: (Ziessman, O'Malley, & Thrall, 2007).

NÓDULO	PROBABILIDAD DE CÁNCER TIROIDEO
Frío	15 – 20 %
Indeterminado	15 – 20 %
Multinodular	5 %
Caliente	< 1 %

2.1.3.1 YODO RADIOACTIVO

El yodo radioactivo es un radiofármaco fisiológico ideal, es retenido selectivamente por el tiroides, donde se organifica e incorpora en las hormonas tiroideas. De este modo es posible obtener información clínica importante de la fisiología tiroidea, los radiofármacos empleados en la clínica son el I-123 y el I-131 (Ziessman, O'Malley, & Thrall, 2007).

Debido a la rápida absorción, captación y organificación del yodo, la radioactividad es detectable en la glándula tiroides a los pocos minutos, alcanzando la luz folicular tiroidea a los 20 – 30 minutos. La captación tiroidea aumenta de modo progresivo a lo largo de 24 horas (2 - 6 horas cuando se emplea I-123 o de 1 – 3 días para el I-131), el retraso entre la ingesta del yodo radioactivo y la detección gammagráfica depende del deseo que se produzca el aclaramiento del fondo y aumente la concentración en el órgano objetivo que de lenta captación glandular (Ziessman, O'Malley, & Thrall, 2007).

2.1.3.2 PERTECNETATO

El bajo coste del pertecnetato Tc-99m y la fácil disponibilidad a partir de los sistemas generadores de molibdeno-99/Tc-99m han contribuido a que este radiofármaco se haya empleado como alternativa al yodo radioactivo para la obtención de gammagrafías tiroideas (Ziessman, O'Malley, & Thrall, 2007).

A diferencia del yodo radioactivo, que se administra vía oral, el pertecnetato Tc-99m se administra por vía intravenosa. Al igual que el yoduro, el Tc-99m es captado por la tiroides, pero no es organificado, vale decir que no es retenido por la tiroides, por lo que la gammagrafía tiroidea con Tc-99m debe obtenerse cuando la captación es máxima, es decir, a los 20 – 30 minutos de su administración.

De los radiofármacos mencionados, la institución utiliza pertecnetato Tc-99m por el tiempo que necesita el personal para poder obtener las imágenes con centellografía.

2.1.4 NÓDULO FRÍO

Más del 85 – 90 % de los nódulos tiroideos gammagráficos son nódulos fríos (hipofuncionantes), es decir, presentan menor captación que el tejido tiroideo adyacente. Gran parte de este tipo de nódulos son de naturaleza benigna, como los quistes simples, los nódulos coloides, la tiroiditis, la hemorragia, la necrosis o los trastornos infiltrativos como la amiloidosis o la hemocromatosis (ver tabla 6). Sin embargo, en un subgrupo importante de pacientes los nódulos fríos son de origen canceroso (Ziessman, O'Malley, & Thrall, 2007).

Tabla 6

Diagnóstico de nódulos tiroideos

Fuente: (Ziessman, O'Malley, & Thrall, 2007).

Diagnóstico diferencial de los nódulos tiroideos
Nódulos fríos (no funcionantes)
Benignos
Nódulo coloide
Quiste simple
Quiste hemorrágico
Adenoma
Tiroiditis
Absceso
Quiste o adenoma paratiroideo
Malignos
Papilar
Folicular
Anaplásico
Medular
Células de Hurthle
Linfoma
Carcinoma metastático
Pulmón
Mama
Melanoma
Gastrointestinal
Renal
Nódulos funcionantes (templados o calientes)
Adenomas
Adenomas hiperfuncionantes

2.1.5 NÓDULO CALIENTE

La captación de yodo radioactivo por parte de un nódulo indica que es funcionante. Los nódulos funcionantes muy rara vez representan un proceso maligno, menos del 1 % de estos nódulos son malignos. El término nódulo caliente debe reservarse para aquellos nódulos que no sólo presentan una captación elevada en la gammagrafía, sino también se

acompaña de la supresión del tejido extranodular. Los nódulos templados son aquellos en los que el tejido extranodular no se encuentra suprimido (Ziessman, O'Malley, & Thrall, 2007).

2.1.6 NÓDULOS INDETERMINADOS

Un nódulo indeterminado es todo nódulo palpable o con un tamaño ecográfico mayor de un 1 cm que no puede diferenciarse en la gammagrafía tiroidea ni como nódulo frío ni como uno caliente cuando se compara con el tejido tiroideo normal adyacente. Con fines terapéuticos, en los nódulos indeterminados se deben seguir las pautas empleadas para los nódulos fríos (Ziessman, O'Malley, & Thrall, 2007).

2.2 INTELIGENCIA ARTIFICIAL

Esta tecnología está avanzando significativamente, ocasionando que diferentes instituciones y empresas comiencen a utilizarla para realizar alguna tarea, sobre inteligencia artificial (IA) podemos encontrar varias definiciones y todas válidas, pero se puede decir que: “La IA es la rama de la ciencia computacional que se encarga del estudio de la inteligencia humana en elementos artificiales y desde el punto de vista de la ingeniería, propone la creación de elementos que posean un comportamiento inteligente” (Romero, Dafonte, Gómez, & Penousal, 2007), en otras palabras es la ciencia que trata de imitar la inteligencia humana a través de diferentes técnicas y dar la capacidad a una computadora de pensar como humano, reconocer de manera visual los objetos, reconocer texto, etc., con el fin de alcanzar ese objetivo se fueron desarrollando tecnologías más específicas como redes neuronales artificiales, algoritmos genéticos, agentes inteligentes, lógica difusa, *Machine Learning*, *Deep Learning*, etc.

2.2.1 APRENDIZAJE AUTOMÁTICO

El aprendizaje automático es un área de la inteligencia artificial, la traducción en inglés es *Machine Learning*, el objetivo de esta área es el estudio de reconocimiento de patrones usando ciencias como matemáticas, estadística, ciencias de la computación y procesos de ingeniería para que las computadoras puedan aprender.

El aprendizaje automático no es un concepto nuevo, a principios se utilizaba solo para clasificación de datos, ahora con los avances realizados ha tomado más fuerza y ya se puede hacer predicciones de tráfico y comportamientos de clientes (marketing), detección de cáncer, clasificación de contenido en buscadores web, recomendaciones en redes sociales, etc., el uso de esta tecnología va creciendo y representa una oportunidad para crear soluciones más complejas, los tipos de algoritmos que se manejan son: supervisado, no supervisado o por refuerzo.

2.2.2 APRENDIZAJE SUPERVISADO

En los algoritmos de aprendizaje supervisado se genera un modelo predictivo utilizando un conjunto de datos de entrada y de salida, se dice aprendizaje supervisado porque se tiene un conjunto de datos que han sido anteriormente etiquetados y supervisados, vale decir, que ya se sabe los datos a que grupo, valor o categoría pertenecen. Este grupo de datos será utilizado para el entrenamiento del modelo inicial planteado, de esta manera el algoritmo irá aprendiendo a clasificar las muestras de entrada comparando el resultado del modelo, creando una relación entre la entrada y la salida deseada, gráficamente podemos ver en la figura 4 (Contreras, 2016; Luna, 2018).

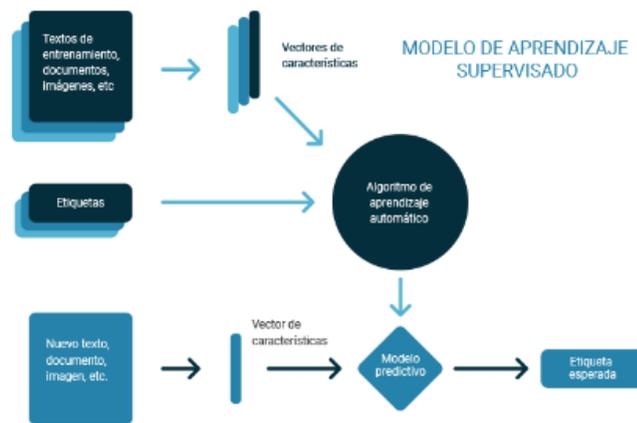


Figura 4. Modelo aprendizaje supervisado
Fuente: (Luna, 2018)

2.2.3 REDES NEURONALES ARTIFICIALES

Las redes neuronales son un intento de reproducir el funcionamiento del cerebro humano en una computadora, simulando las capacidades de memorizar y asociar hechos, lo que permite generar experiencia que posteriormente puede ser usada para resolver algún problema a través de esas capacidades (Matich, 2001).

Al buscar emular el comportamiento biológico de las neuronas humanas y como se organizan formando la estructura del cerebro, una neurona biológica (ver figura 5) está compuesta por las siguientes partes (Sancho, 2018):

- El cuerpo central, llamado soma, que contiene el núcleo celular.
- Una prolongación del soma, el axón.
- Un conjunto de ramificaciones terminales, las dendritas.
- Zonas de conexión entre una neurona y otra, conocidas como sinapsis.

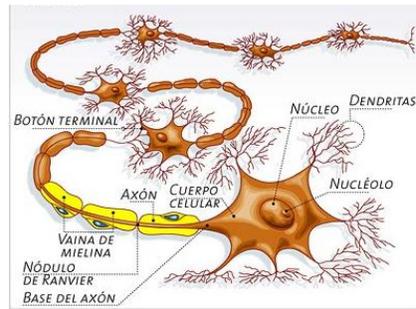


Figura 5. Neurona biológica
Fuente: (Sancho, 2018)

Las redes neuronales son modelos matemáticos que permiten hacer el tratamiento de información inspiradas en las neuronas humanas, las neuronas son las unidades básicas que generalmente están organizadas en capas (entrada, oculta y salida), como se muestra en la figura 6.

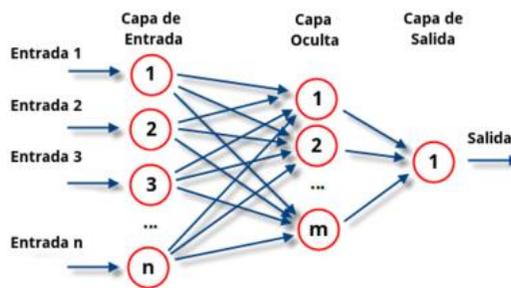


Figura 6. Estructura de una red neuronal
Fuente: (Matich, 2001)

En el presente trabajo se utilizó las neuronas Dense, estas son neuronas simples que conectan cada neurona en una capa con todas las salidas de la capa anterior (Rivera, 2020).

2.2.4 MODELO DE RED NEURONAL DE MCCULLOCH-PITTS

El primer modelo matemático de una neurona artificial, creado con el fin de llevar a cabo tareas simples, fue presentado en el año 1943 en un trabajo conjunto entre el psiquiatra y neuroanatomista Warren McCulloch y el matemático Walter Pitts.

La figura 7 muestra un ejemplo de modelo neuronal con N entradas (Sancho, 2018), que consta de:

- Un conjunto de entradas x_1, \dots, x_n .
- Los pesos sinápticos w_1, \dots, w_n , correspondientes a cada entrada.
- Una función de agregación, Σ .
- Una función de activación, f .
- Una salida, Y .

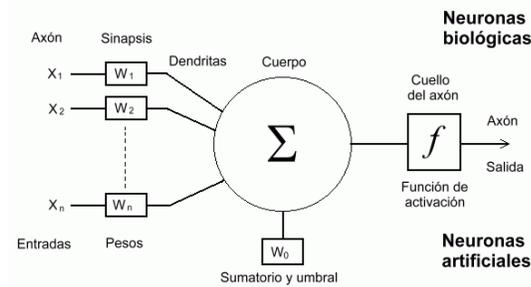


Figura 7. Neurona artificial
Fuente: (Sancho, 2018)

Las entradas son el estímulo que la neurona artificial recibe del entorno que la rodea y la salida es la respuesta a tal estímulo. La neurona puede adaptarse al medio y aprender para modificar el valor de sus pesos sinápticos y por ello son conocidos como los parámetros libres del modelo, ya que pueden ser modificados y adaptados para realizar una tarea determinada.

2.2.5 FUNCIÓN DE ACTIVACIÓN

Una función de activación permite determinar la salida de cada capa, es decir, que ella escoge el valor que pasa o se queda, estas funciones son muy importantes para el aprendizaje profundo usado en la clasificación. Entre las funciones más utilizadas se tienen las siguientes funciones:

La función *Threshold* (escalón) devuelve 1 cuando $x \geq 1$ y 0 en otro caso (ver figura 8), o sea que las salidas serán binarias, encendido o apagado. Esta es la función más rígida, no hay casos intermedios y sirve para clasificar de forma muy estricta.

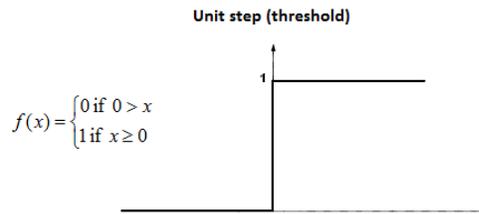


Figura 8. Función Threshold
Fuente: (Crispi, 2019)

La función *sigmoide*, es una curva más suave que va de 0 a 1 pudiendo tener valores intermedios, esta función es usada en la capa de salida cuando se quiere calcular una probabilidad (ver figura 9), no es tan estricta y es muy usada para regresión logística para *machine learning*.

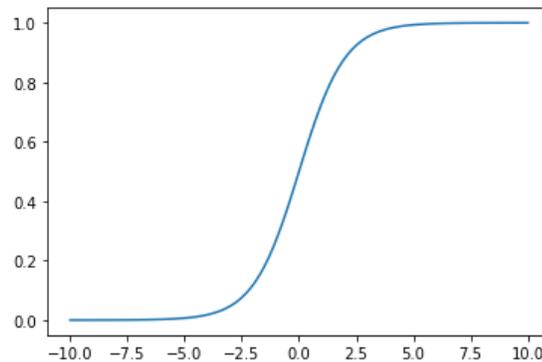


Figura 9. Función sigmoide
Fuente: (Crispi, 2019)

La función rectificadora (ReLU), es muy utilizada para clasificación de imágenes con buen desempeño en redes convolucionales, devuelve 1 si el valor es mayor a cero y en otro caso devuelve 0, como se muestra en la figura 10.

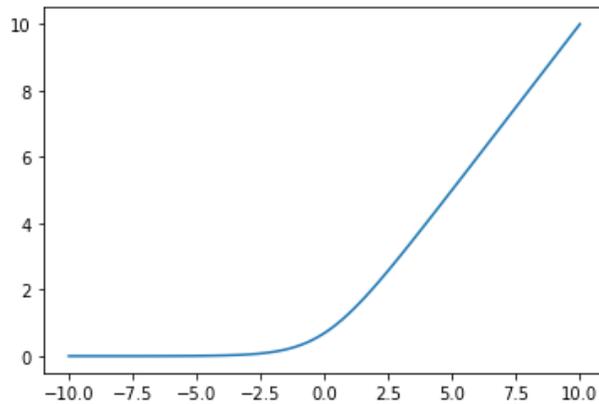


Figura 10. Función rectificadora
Fuente: (Crispi, 2019)

La función Softmax transforma las salidas a una representación en forma de probabilidades, de tal manera que el sumatorio de todas las probabilidades de las salidas de 1, utilizada para normalizar tipo multiclase y con buen rendimiento en las últimas capas (Calvo, 2017).

2.2.6 REDES NEURONALES CONVOLUCIONALES

Una de las cualidades de los seres humanos es la capacidad de reconocer objetos a través del sentido de la visión, porque en fracciones de segundos al observar algún objeto podemos identificar algo sin que sea necesario hacer mucho esfuerzo, la profundidad, distinguir contornos o colores y otras cosas más.

Esta tarea de recibir la información visual es procesada por la corteza visual del cerebro, la cual tiene diferentes neuronas que van realizando el procesamiento de reconocimiento. Las neuronas que se encuentran en la zona V1 (lóbulo occipital (3) y área visual primaria (5) en la figura 11) se ocupan de características visuales de bajo nivel, tales como pequeños segmentos de contorno, información básica de contraste, esta capa genera información para otras áreas como V2 que capta el color del entorno, en la zona V4 las neuronas participan en el procesamiento de la forma de los estímulos y en su reconocimiento (lóbulo occipital (3) y área visual primaria (5) en la figura 11). Cada una

de estas áreas se ocupa de aspectos específicos, es así como surgen las redes neuronales convolucionales (Castillero, 2019; Cooley, 2018).

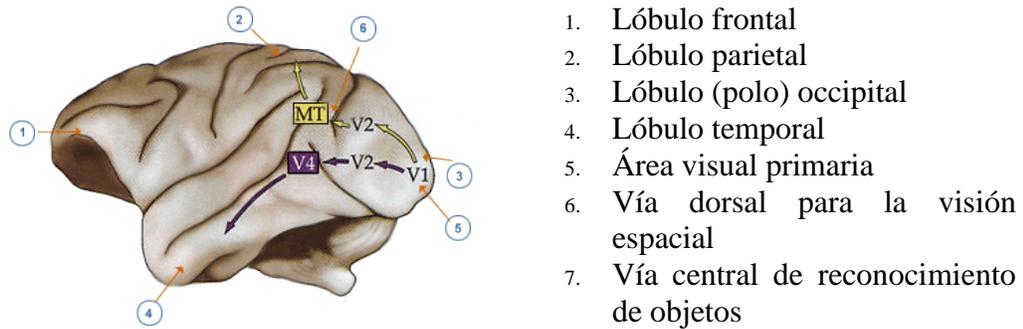


Figura 11. Corteza visual del cerebro

Fuente: (Lopez, 2016)

Las redes neuronales convolucionales (en inglés CNN - *Convolutional Neural Network*) son similares a las redes neuronales multicapas, con la principal ventaja que cada parte de la red se entrena para realizar una tarea específica, esta característica permite reducir el número de capas ocultas, generando que el entrenamiento sea más rápido.

Este tipo de redes son muy utilizadas para el reconocimiento y análisis de imágenes, porque son capaces de reconocer características como la detección de bordes, líneas, etc., haciendo continuas iteraciones (convoluciones). Entonces se hace más común el uso de los algoritmos en aplicaciones de detección facial en teléfonos móviles (snapchat), bancos, etc.

Una red neuronal convolucional es una red multicapa que está compuesta por varias capas convolucionales y de reducción, y al final tiene capas de conexión total que permite realizar la clasificación, como se ilustra en la figura 12 (Calvo, 2017).

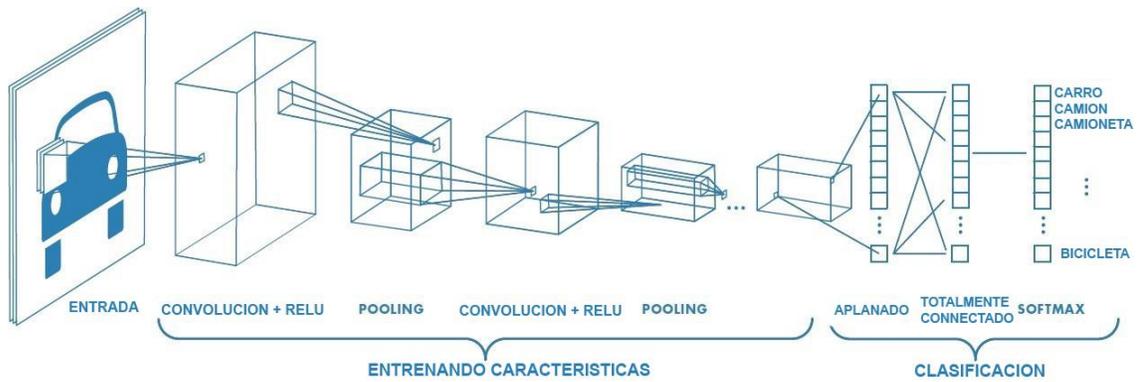


Figura 12. Estructura de Red Neuronal Convolutional
Fuente: (Saha, 2018)

2.2.6.1 CONVOLUCIÓN

Una convolución es una operación matemática en la que una función se aplica de alguna manera a otra función (i). El resultado se puede entender como una “mezcla” de las dos funciones, donde las convoluciones son buenas para detectar estructuras sencillas de una imagen (ver figura 14) y después se juntan esas sencillas funciones para construir funciones más complejas (Cooley, 2018).

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\gamma)g(t - \gamma)d\gamma \quad (1)$$

Figura 13. Ecuación de convolución

Fuente: (Crispi, 2019)

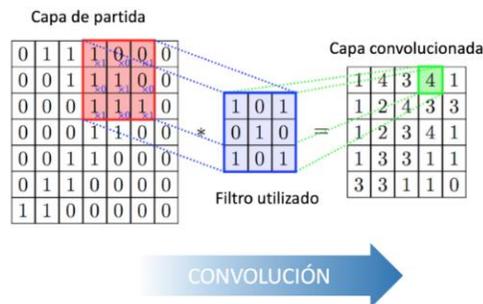


Figura 14. Convolución
Fuente: (Calvo, 2017)

2.2.6.2 REDUCCIÓN O *POOLING*

La reducción simplemente es la disminución de la cantidad de parámetros o dimensiones para quedarse con las características más comunes (ver figura 15), evitando la sobrecarga de cálculo para las próximas capas de la red.

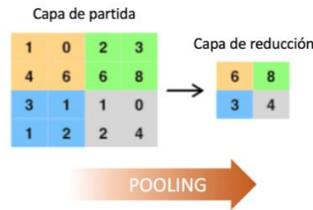


Figura 15. Reducción o pooling
Fuente: (Calvo, 2017)

2.3 HERRAMIENTAS PARA DESARROLLO DEL MODELO

Se han utilizado algunas herramientas para el desarrollo del modelo de *machine learning* y obtención de datos para la evaluación del modelo.

2.3.1 TENSORFLOW

Tensorflow es una plataforma de código abierto dirigida al aprendizaje automático (*Machine Learning*) a través de una serie de tareas. Ha sido desarrollado por Google para satisfacer las necesidades de sistemas capaces de construir y entrenar redes neuronales para detectar y descifrar patrones, similares al aprendizaje y razonamiento usados por los humanos, llegando a ser la herramienta más utilizada en el mundo de *Deep Learning*. Actualmente es utilizado tanto para la investigación como para la producción de productos de Google, reemplazando el rol de su predecesor de código cerrado DistBelief. *TensorFlow* fue originalmente desarrollado por el equipo de *Google Brain* para uso interno en Google antes de ser publicado bajo la licencia de código abierto Apache 2.0 el 9 de noviembre de 2015 y puede correr en múltiples CPUs y GPUs en diferentes plataformas.

En el sector salud es uno de los campos que más se están revolucionando y que mayor impacto tendrá para todos nosotros como sociedad en los próximos años, ya que permite crear herramientas que ayuden a los médicos.

2.3.2 KERAS

Keras es una librería de redes neuronales de código abierto escrita en Python, es capaz de ejecutarse sobre Tensorflow, Microsoft CognitiveToolkit o Theano 1. Esta herramienta fue diseñada especialmente para hacer posible la experimentación en más o menos poco tiempo con redes de aprendizaje profundo. Sus fuertes se centran en ser amigable para el usuario, modular y extensible. Desarrollada como parte de los esfuerzos de investigación del proyecto ONEIROS (*Open-ended Neuro-Electronic Intelligent Robot Operating System*), pensada como una interfaz para *machine learning*, haciendo más sencillo el desarrollo de modelos de aprendizaje profundo (Utrera, 2018).

2.3.3 ANACONDA NAVIGATOR

Anaconda es una suite multiplataforma de código abierto que abarca una serie de aplicaciones y librerías diseñadas para el desarrollo de Ciencia de datos en *Python*, asimismo permite adicionar más paquetes y dependencias gestionándolos de manera avanzada, proporcionando un entorno de trabajo adecuado que acelere la implementación de proyectos de ciencias de datos (Toro, 2017).

2.3.4 MATRIZ DE CONFUSIÓN

Una herramienta fundamental para evaluar si un algoritmo de clasificación funciona de manera adecuada o no, es la matriz de confusión, mediante ella podemos obtener métricas como la precisión de las predicciones del modelo (ver tabla 7).

La matriz de confusión de un problema de clase “n” es una matriz nxn en la que las filas se nombran según las clases reales y las columnas son las clases previstas por el modelo. Sirve para mostrar de forma explícita cuando una clase es confundida con otra (Recuero, 2018).

Tabla 7
Matriz de confusión
 Fuente: (Recuero, 2018)

	Predicción	
	N (modelo)	P (modelo)
N (real)	Negativos reales (a)	Falsos positivos (b)
P (real)	Falsos negativos (c)	Positivos reales (d)

Las métricas que se puede obtener a partir de la matriz de confusión son las siguientes (Recuero, 2018):

- Exactitud, representa la proporción entre el número de diagnósticos correctos (tanto positivos como negativos) y el total de diagnósticos:

$$AC = \frac{a + d}{a + b + c + d} \quad (2)$$

- Precisión, representa la proporción entre los positivos reales predichos por el algoritmo y todos los casos positivos:

$$P = \frac{d}{b + d} \quad (3)$$

- Sensibilidad, es la proporción de casos positivos que fueron correctamente identificadas por el algoritmo:

$$TP = \frac{d}{c + d} \quad (4)$$

- Especificidad, es la proporción de los casos negativos que el algoritmo ha clasificado correctamente:

$$TN = \frac{a}{a + b} \quad (5)$$

2.4 PRUEBA DE HIPÓTESIS

La prueba de una hipótesis es una regla que determina si se puede aceptar o rechazar una afirmación de una población dependiendo de la evidencia proporcionada por una muestra de datos. Esta prueba examina dos hipótesis opuestas sobre una población: la hipótesis nula y la hipótesis alternativa. Por lo general, la hipótesis nula es el enunciado que “no hay efecto” y que se pretende rechazar, siendo la hipótesis alternativa el enunciado que se quiere concluir como verdadero (Minitab, 2019).

Esta tarea se realiza usando los datos de muestra, la prueba determina si se puede rechazar la hipótesis nula. Se utiliza el valor de probabilidad (p) para tomar esa decisión, si el valor de p es menor que el nivel de significancia (α o alfa), entonces puede rechazar la hipótesis nula (Minitab, 2019).

El nivel de confianza representa un porcentaje de intervalos que incluirían el parámetro de población si se toma muestras de la misma población en más de una ocasión. Por lo general, un nivel de confianza del 95% funciona de manera adecuada (ver figura 16).

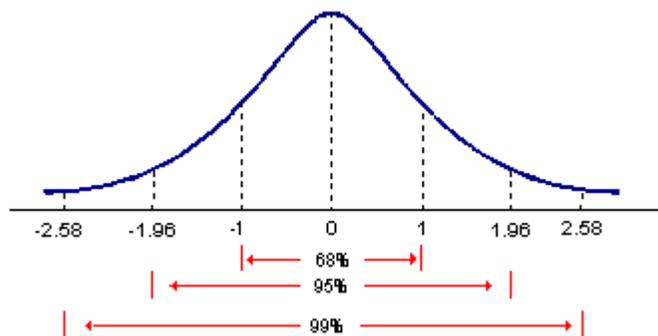


Figura 16. Grafica normal standard

El valor Z crítico, calculado a través de la ecuación (6), es el que determina si se rechaza o no la hipótesis nula. Si el nivel de significancia es $\alpha = 5$ y el nivel de confianza 95%, entonces las regiones de rechazo serán todos los valores menores a -1.96 y mayores a 1.96.

$$Z = \frac{\hat{\mu} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \quad (6)$$

Los valores de la ecuación son:

- $\hat{\mu}$, es la media de los valores.
- μ , es la media poblacional.
- σ , es la desviación estándar de la población.
- n , es la cantidad de elementos de la muestra.

2.4.1 ESTADÍSTICO DE ANDERSON-DARLING

El estadístico de Anderson-Darling mide que tan bien los datos siguen una distribución específica. Para un conjunto de datos y una distribución, mientras mejor se ajuste la distribución a los datos, menor será el estadístico (Minitab2, 2019).

El estadístico de Anderson-Darling es una alternativa a la prueba de chi-cuadrado (X^2) y también puede ser utilizado para comparar el ajuste de varias distribuciones con el fin de determinar cuál es el mejor. Sin embargo, para determinar que una distribución es la mejor, el estadístico debe ser sustancialmente menor que los demás. Cuando los estadísticos tengan valores cercanos, se debe usar criterios adicionales, como las gráficas de probabilidad (Minitab2, 2019).

El valor del Test estadístico de Anderson-Darling (Nist, 2018) es definido como:

$$A^2 = -N - S \quad (7)$$

Donde S

$$S = \sum_{i=1}^N \frac{(2i-1)}{N} [\ln F(Y_i) + \ln (1 - F(Y_{N+1-i}))] \quad (8)$$

CAPÍTULO III

MARCO PRÁCTICO

En esta sección se describe procedimiento que realiza la institución para la toma de imágenes y al mismo tiempo se desarrolla el modelo de clasificación con *machine learning*, probando diferentes estructuras hasta obtener el modelo que tenga mayor precisión.

3.1 PROCEDIMIENTO PARA EL DIAGNÓSTICO DE NÓDULO TIROIDEO

Antes de comenzar el desarrollo del prototipo del sistema basado en *Machine Learning*, se realizó previamente el análisis del procedimiento que efectúa la institución para la evaluación de enfermedades tiroideas.

Para el estudio de los nódulos tiroideos a través de imágenes con centellografía, el personal de la institución utiliza diferentes radiofármacos (yodo radioactivo y pertecnetato) los cuales son suministrados a los pacientes para que el equipo gammagráfico pueda captar la información que requieren los especialistas.

3.1.1 PROCEDIMIENTO DE TOMA DE IMÁGENES

Todos los pacientes que vienen a la institución para la toma de imágenes son referidos, vale decir, ellos llegan con una orden médica de alguna caja de salud o de un consultorio médico, donde posiblemente se ha identificado un problema o anomalía en la glándula tiroides y requieren información adicional para continuar con el diagnóstico. Cumplido este requisito, el procedimiento se realiza de la siguiente manera:

- a) Solicitar atención en INAMEN
- b) Recibir al paciente con la orden del seguro de salud o consultorio médico.
- c) Programar una fecha para realizar la toma.
- d) Aplicar el radiofármaco al paciente en la fecha y hora definida

- e) Esperar el tiempo establecido.
- f) Tomar la gammagrafía de la glándula tiroides.
- g) Entregar la imagen al especialista para que elabore el diagnóstico.
- h) Elaborar informe según lo observado en la imagen para la entrega al paciente o al personal del seguro de salud, según corresponda el caso.

En la presente investigación pretende colaborar al especialista en el diagnóstico entre los pasos g) y h) a través de la herramienta propuesta (ver figura 17).

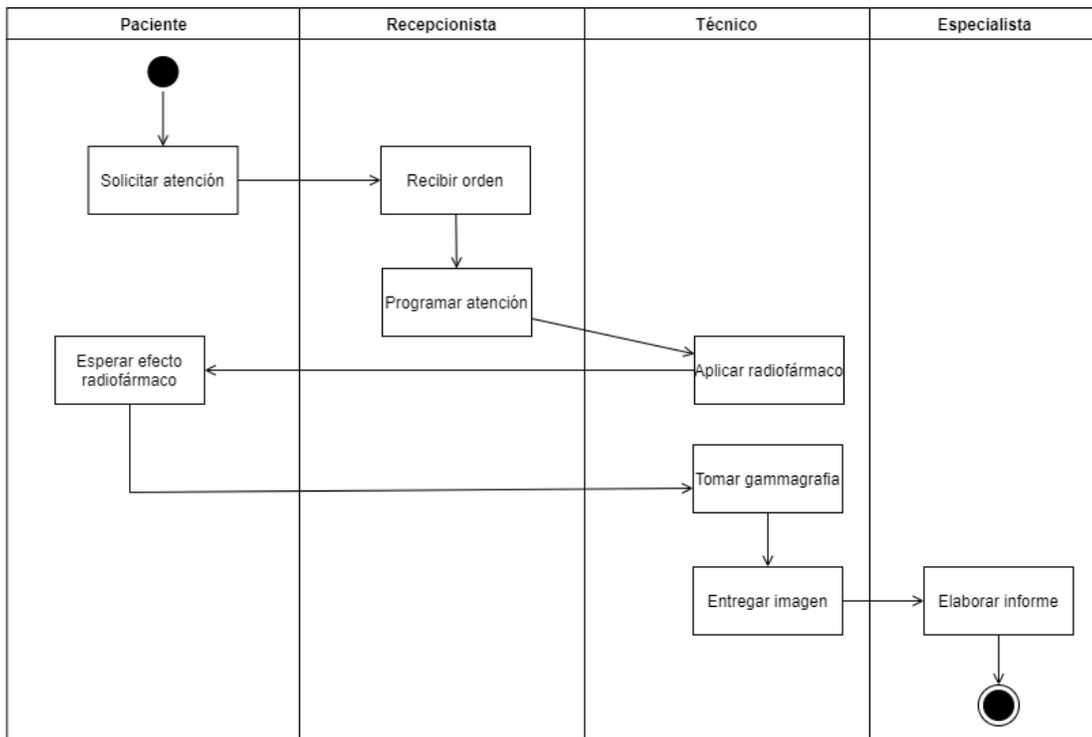


Figura 17. Procedimiento de toma de imágenes de centellografía.

3.2 RECOPIACIÓN DE INFORMACIÓN

El personal del INAMEN ha proporcionado alrededor de 7000 informes sobre el diagnóstico de la glándula tiroides, dichos documentos datan desde hace ocho años atrás, los cuales fueron elaborados por los diferentes especialistas que trabajaron o que trabajan

en la mencionada institución. Las imágenes se encuentran en documentos Word ordenados alfabéticamente en carpetas por la primera letra del apellido paterno del paciente.

Para obtener las imágenes se realizó el siguiente procedimiento:

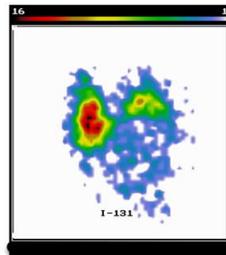
- Se ha desarrollado una aplicación en C# para leer todos los documentos y buscar las imágenes que contenga cada documento (ver anexo A), una vez identificada una imagen esta fue almacenada en una carpeta. Se utilizó C# porque el *framework* de .NET posee más métodos que han permitido extraer las imágenes de los documentos Word en comparación con lenguajes de programación como Java o Python.
- Para organizar las imágenes de acuerdo al tipo de nódulo que tienen, se utilizó la sección conclusión de los informes, en este lugar el especialista redacta el diagnóstico del paciente después de haber examinado las imágenes (ver figura 18 hasta la figura 22). Entonces se ha desarrollado otra aplicación en C# (ver anexo B) para que a partir de las conclusiones, organice las imágenes extraídas de acuerdo a los tipos que se necesita para el entrenamiento y validación de la red neuronal.

Luego de obtener las imágenes se evaluó con uno de los médicos especialistas de la institución, la Dra. Murillo durante una entrevista explicó algunas características que se toma en cuenta para identificar las enfermedades de la tiroides, como ser: la forma de la glándula tiroides, que tiene la forma de alas de una mariposa, un nódulo difuso pierde la forma mencionada mostrándose de diferente manera. El criterio que tienen los médicos para evaluar una enfermedad de la tiroides fue aprendida durante la formación académica y en base a experiencia (Murillo, 2019).

A continuación se muestran algunos casos de los informes proporcionados:

En la figura 18 se muestra un nódulo frío, donde el lóbulo izquierdo de la tiroides no se observa completamente los colores, esto quiere decir que este lugar el tejido tiroideo no capta o no retiene el radiofármaco suministrado al paciente.

Atn.: Dr. (a) [redacted] Fecha: [redacted]
Paciente: [redacted] No. [redacted]
Estudio: CENTELLOGRAFÍA TIROIDEA Radiofármaco: ^{131}I



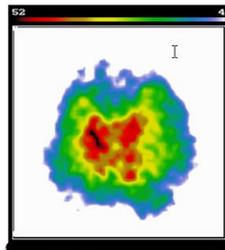
INFORME: La imagen nuclear tiroidea, obtenida luego de la administración de 250 uCi de ^{131}I , muestra una glándula de situación conservada, aumentada de volumen a expensas de lóbulo izquierdo que presenta una zona fría que ocupa los tercios inferiores del lóbulo, correspondiente a nódulo palpable: "nódulo frío". La concentración en el resto del parénquima se encuentra marcadamente disminuida.

Conclusión: Bocio nodular con nódulo frío en lóbulo izquierdo. Disminución global de la concentración del indicador radiactivo que sugiere ingesta o administración de sustancias yodadas (¿contraste?).

Figura 18. Informe de diagnóstico de nódulo frío
Fuente: INAMEN

En la figura 19 se observa otro nódulo frío donde hay una deformidad en la forma de tiroides, no hay captación correcta del radiofármaco en ambos lóbulos de la tiroides y además indica que son nódulos palpables.

Atn.: Dr. (a) [redacted] Fecha: [redacted]
Paciente: [redacted] No. [redacted]
Estudio: CENTELLOGRAFÍA DE TIROIDES Radiofármaco: $^{99\text{m}}\text{Tc}$



INFORME: La imagen nuclear tiroidea obtenida luego de la administración de 3mCi de $^{99\text{m}}\text{Tc}$, muestra una glándula de situación conservada, aumentada de volumen con heterogénea concentración del indicador radiactivo por presencia de zonas frías en ambos lóbulos tiroideos, correspondientes a nódulos palpables: "nódulos fríos".

Conclusión: Bocio nodular gigante con nódulos fríos en ambos lóbulos.

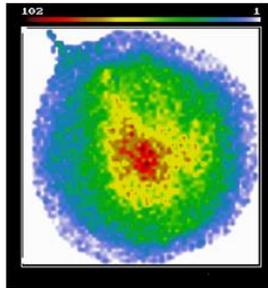
Figura 19. Informe de diagnóstico de nódulo frío
Fuente: INAMEN

En la figura 20 se observa que hay captación, pero tiene una deformidad respecto a la forma de la tiroides, como si fuera un solo lóbulo, esta característica permite indicar que es un bocio difuso.

Atn. : ██████████	No. : 106435
Paciente: ██████████	Fecha: ██████████
Estudio: CENTELLOGRAFIA TIROIDEA	Radiofármaco: ^{99m} TcO ₄

Estimado (a) Colega:

Le agradecemos el envío de su paciente ██████████, nacida el ██████████



INFORME: La imagen adquirida en proyección anterior de cuello, post administración de 4.83 mCi de ^{99m}TcO₄, muestra tiroides de posición y morfología conservados, con volumen aumentado. Captación parenquimatosa homogénea. Sin presencia de nódulos.

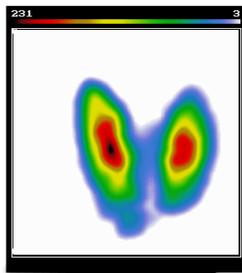
CONCLUSIÓN: BOCIO DIFUSO.

Figura 20. Informe de diagnóstico de bocio difuso

Fuente: INAMEN

En la figura 21 se observa que hay captación del radiofármaco suministrado, también se puede distinguir la forma de la tiroides, el médico por el aumento de volumen, determinó que era un bocio difuso.

Atn. : Dr. (a) ██████████	Fecha: ██████████
Paciente: ██████████	No. ██████████
Estudio: CENTELLOGRAFÍA DE TIROIDES	Radiofármaco: ^{99m} Tc



INFORME: La imagen nuclear tiroidea obtenida luego de la administración de 3mCi de ^{99m}Tc, muestra una glándula de situación conservada, discretamente aumentada de volumen con homogénea concentración del indicador radiactivo.

Conclusión: Bocio difuso.

Figura 21. Informe de diagnóstico de bocio difuso

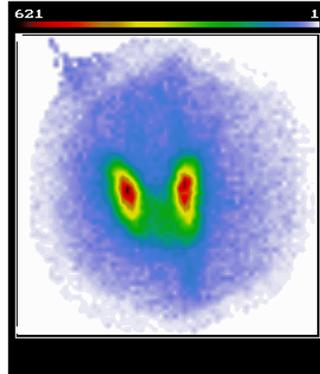
Fuente: INAMEN

La figura 22 se observa que hay captación homogénea y conserva la forma de una tiroides, por lo que el médico determinó que tiene una tiroides normal.

Atn. : ██████████	No. : ██████████
Paciente: ██████████	Fecha: ██████████
Estudio: CENTELLOGRAFÍA TIROIDEA	Radiofármaco: ^{99m} TcO ₄

Estimado (a) Colega:

Le agradecemos el envío de su paciente, ██████████



INFORME: La imagen adquirida en proyección anterior de cuello, post administración de 4.40 mCi de ^{99m}TcO₄ muestra tiroides de posición y morfología conservadas, con volumen normal. Captación parenquimatosa homogénea. Sin presencia de nódulos.

CONCLUSIÓN: GÁMMAGRAFIA TIROIDEA NORMAL.

Figura 22. Tiroides normal
Fuente: INAMEN

En las imágenes obtenidas, se observó que existe diferencia entre las enfermedades de la tiroides, también se vio que entre imágenes que pertenecen a un mismo grupo hay divergencias entre ellos como en las figuras 20 y 21, donde ambos son bocios difusos pero no siguen un mismo patrón, pero si se compara la figura 21 (bocio difuso) con la figura 22 (tiroides normal) estas tienen cierta similitud porque ambas conservan la forma de la tiroides. Al momento de establecer el resultado del diagnóstico el criterio de un médico puede variar respecto a otro, por lo que la experiencia del experto juega un papel muy importante.

Los tipos nódulos más comunes identificados son:

- Bocio nodular
- Bocio difuso
- Tiroides normal o sin alteraciones
- Nódulo frío
- Tiroiditis

Los nódulos fríos serán los que tenga mayor importancia de identificar, porque este nódulo tiene 15% de probabilidad de convertirse en cancerígeno (Ziessman, O'Malley, & Thrall, 2007).

3.3 DESARROLLO DE MODELO DE CLASIFICACIÓN

Para el desarrollo del modelo se consideró la metodología propuesta por Roman (2019), la que está compuesta por 4 procesos principales (figura 23):

- Pre-procesamiento, extracción de información (imágenes) para entrenar y validar el modelo.
- Aprendizaje, inicio de entrenamiento para poder seleccionar el algoritmo que permita clasificar con la mayor precisión.
- Evaluación, selección de la métrica para evaluar rendimiento del modelo.
- Predicción, se realizan pruebas para verificar la predicción.

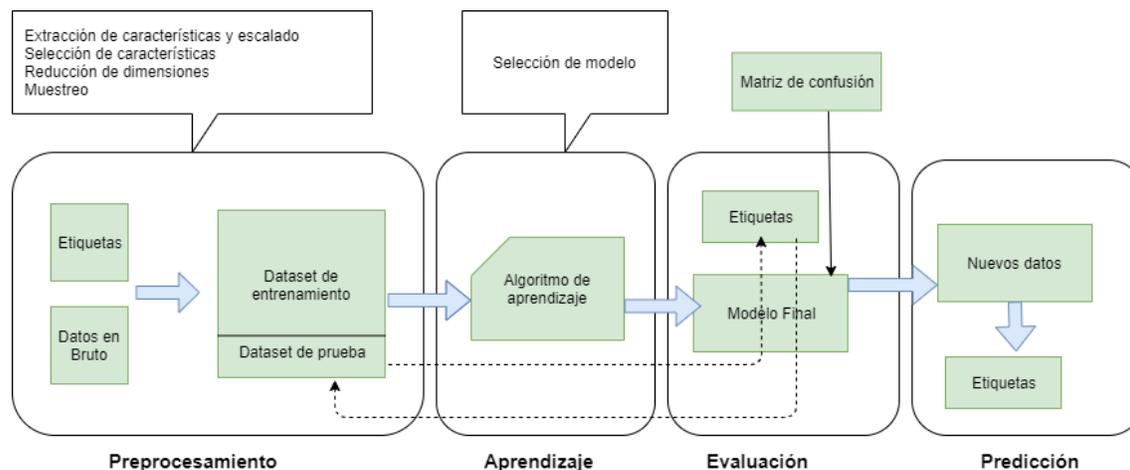


Figura 23. Metodología de construcción de modelo de machine learning
Fuente: (Roman, 2019) traducido por Bladimir Calcina

El modelo de clasificación busca coadyuvar a los profesionales médicos en el procedimiento de diagnóstico de nódulos tiroideos por imágenes de centellografía, para realizar esta tarea de clasificación, el modelo se organiza en 3 etapas o componentes que

se relacionan de manera secuencial (ver figura 24), las cuales son descritas en el siguiente detalle:

Pre-procesamiento, en esta etapa se realiza varias tareas como: la selección y depuración de imágenes, limpieza de datos (remover textos, bordes y manchas) que solo generan ruido y estandarización a un tamaño para que la red neuronal pueda entrenar.

Aprendizaje, a partir de la información obtenida se desarrolla el prototipo con la red neuronal convolucional, este prototipo fue entrenado para que aprenda a identificar a cual tipo de nódulo pertenece una imagen obtenida por centellografía. Se realizó varias iteraciones hasta encontrar la versión que obtuvo el mayor valor de precisión.

Evaluación, una vez encontrado el prototipo con mejor valor de precisión para la clasificación, se realiza la evaluación para comprobar el adecuado funcionamiento. Para esta tarea se utilizó la matriz de confusión que hace uso de cuatro indicadores: precisión, exactitud, sensibilidad y especificidad.

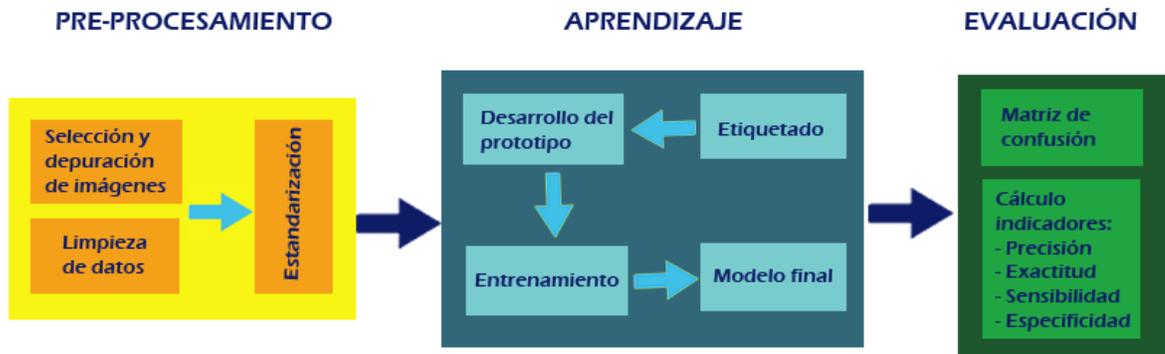


Figura 24. Modelo de clasificación de imágenes

A continuación se detalla las etapas para el desarrollo del modelo, de acuerdo a la metodología mencionada.

3.3.1 PRE-PROCESAMIENTO

El pre-procesamiento o *feature engineering*, es uno de los pasos más importantes para el desarrollo del modelo con *machine learning*, es el punto de inicio del trabajo de la investigación, donde se toman las imágenes obtenidas como fuente de datos para la construcción del modelo.

En esta etapa, se realizaron las siguientes actividades sobre las imágenes extraídas:

- Depuración de imágenes que no proporcionan información relevante. En más de una ocasión se observó que la imagen obtenida no permitía al especialista determinar el problema que tenía el paciente en la tiroides, razón por la cual se realizaba otra toma, se descartó esas imágenes que no estaban bien definidas. También se descartó los casos donde las imágenes no mostraban a la glándula tiroides (ver figura 25).

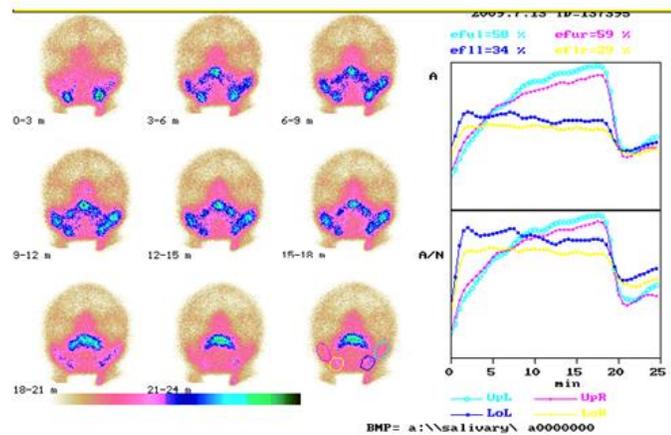


Figura 25. Imagen descartada
FUENTE: INAMEN

- Las imágenes tenían un recuadro (ver figura 26) el cual no forma parte de la tiroides, este recuadro no proporciona información relevante para el entrenamiento del algoritmo de clasificación. Mediante una aplicación en Python, se usó la librería PIL para el tratamiento de imágenes (ver anexo C), se recortó el recuadro de todas las imágenes.

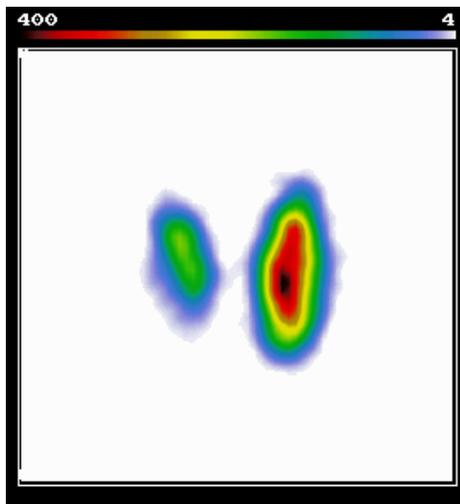


Figura 26. Imagen con recuadro o borde

- Estandarización de información, las imágenes obtenidas contenían información que solo generan ruido al momento de entrenar la red neuronal, es decir, existían textos en ciertos lugares de la imagen como el tipo de radiofármaco que se utilizó (ver figura 27) o el nombre del paciente, manchas adicionales a la imagen de la tiroides, en estos casos se tuvo que usar un editor de texto como MS-Paint para remover esa información que no era útil.

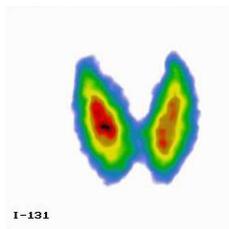


Figura 27. Nombre de radiofármaco en la imagen
Fuente: INAMEN

- Las imágenes se trabajaron como arreglos de datos y para hacer la comparación entre imágenes se requiere que tengan un mismo tamaño. Entonces se tuvo que desarrollar una aplicación en Python para que todas las imágenes tengan las mismas dimensiones (ver anexo D).

Después de haber realizado todas estas actividades de preparación de los datos, se obtuvo el grupo que se muestra en la tabla 8 y porcentualmente podemos observar en la figura 28.

Tabla 8
Cantidad de imágenes utilizadas

Tipo	Cantidad de imágenes
Tiroiditis	151
Bocio nodular	644
Tiroides normal	1.209
Bocio difuso	841
Nódulo frio	2.853
Total:	5.698

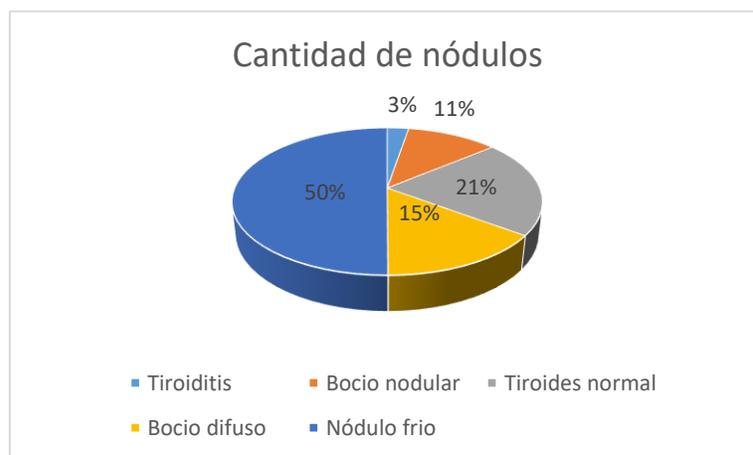


Figura 28. Cantidad de nódulos procesados

3.3.2 DESARROLLO DEL PROTOTIPO

Terminada la fase de organización de las imágenes se comenzó a probar el algoritmo de *machine learning* para el desarrollo del prototipo. Las redes neuronales convolucionales (CNN) es un modelo de *deep learning* que está diseñado para trabajar con imágenes,

porque a través de ella se pretende dar la capacidad de “ver” al ordenador. Se hizo prueba con el algoritmo de *support vector machine* (SVM), este método de clasificación ha sido extendido a problemas de clasificación con imágenes.

En la comparación que se tuvo entre los algoritmos (ver la tabla 9), la CNN fue la mejor opción para realizar el entrenamiento para la clasificación de imágenes por presentar los niveles más altos de precisión con datos de prueba.

Tabla 9
Comparación CNN vs SVM

Comparación	CNN	SVM
Precisión sin limpieza de datos	29 %	22 %
Precisión inicial con limpieza de datos	60 %	50 %
Tiempo de procesamiento por iteración sin GPU	60 segundos	75 segundos
Tiempo de procesamiento por iteración con GPU	2.5 segundos	3.5 segundos

Las características del equipo y hardware con el cual se trabajó, se mencionan a continuación y es el mínimo recomendable para el trabajo de procesamiento de imágenes:

- HP ENVY con procesador Core i7 de cuarta generación
- 16 GB en memoria RAM
- 1 TB en disco duro
- Sistema Operativo, se trabajó con Windows 10 de 64 bits y Ubuntu 18.04 de 64 bits, siendo Ubuntu el que tuvo mejor rendimiento, teniendo una media 30 segundos por cada iteración para el entrenamiento.
- Tarjeta gráfica GeForce GT 740M, de 2 GB en RAM

Al momento de realizar el entrenamiento se puede realizar por CPU (procesador de propósito general, para cualquier tipo de cálculo) o por GPU (procesador de propósito específico, optimizada para trabajar con grandes cantidades de datos y realizar las mismas operaciones varias veces).

La red neuronal convolucional se trabajó con python, tensorflow, keras y scikit-learn, este conjunto de tecnologías permiten armar un entorno de trabajo con el cual se puede realizar el entrenamiento y validación de la red neuronal. Dadas las características de una CNN, es recomendable que se implementen sobre hardware que tenga capacidad de procesar en paralelo, vale decir, que tenga una tarjeta de video que soporte GPU.

Para el procesamiento en paralelo con GPU, es buena opción tener instalado y configurado CUDA (*Compute Unified Device Architecture*), esta es una plataforma de computación en paralelo que incluye un compilador y un conjunto de herramientas de desarrollo creadas por la empresa NVIDIA que permiten a los programadores aprovechar la computación en paralelo para que ejecuten simultáneamente miles de tareas o subprocesos (Gomar, 2018).

Para la construcción de la red convolucional se procedió de la siguiente manera:

- Se utilizó el método de `Train_test_split` de la librería de scikit-learn, para crear arreglos con las imágenes cargadas en memoria.
- Se normalizaron los arreglos para que los valores se encuentren en el rango de 0 y 1, a través del método `to_categorical` de scikit-learn, convirtiendo las etiquetas en salidas con valores binarios para clasificar mejor la red neuronal, en la presente investigación se tienen 5 etiquetas de nódulos tiroideos: tiroiditis, bocio nodular, tiroides normal, bocio difuso y nódulo frío; la salida será un vector con 5 elementos, donde cada posición representa a la etiqueta en el orden mencionado, si por ejemplo tenemos `[0, 1, 0, 0, 0]` esto indica que es un bocio nodular al 100%.
- Para el entrenamiento y las pruebas se usaron los valores 80-20, es decir, 80% de las imágenes para el *training* y 20% para el *test*.
- Se definieron las siguientes constantes:
 - *Learning rate* para la tolerancia del error de 0,001.

- Cantidad de iteraciones (épocas), se comenzó desde 50 hasta 3000 épocas, para obtener la precisión más alta en el modelo.
 - Tamaño de lotes de las imágenes a procesar en memoria, se utilizó un valor de 64.
- Se creó una capa de neuronas convolucionales de 2 dimensiones de 120x120x3
- Se aplicaron 32 filtros (*kernel*) de tamaño 3x3, que detectan características de la imagen.
- Se utilizó la función LeakyReLU para la activación de las neuronas, esta función se comporta bien con imágenes, tiene buen desempeño en redes convolucionales, es parecida a la función ReLU, transforma los valores introducidos multiplicando los negativos por un coeficiente rectificado (que en este caso fue $\alpha = 0.1$), α es una constante para los números por debajo de cero, para que solo ingresen valores positivos.
- Para evitar el *overfitting* se añadió la técnica denominada Dropout, esta técnica permite seleccionar neuronas aleatoriamente e ir ignorando otras durante el entrenamiento. Según la información revisada sobre la aplicación de la función se recomienda usar valores de *dropout rate* que se encuentren entre 20% a 50%, una probabilidad demasiado baja tiene un efecto mínimo y un valor demasiado alto da como resultado un sub aprendizaje de la red (Geron, 2017; Brownlee, 2019). Cuando se utilizó una probabilidad de 20% se llegó a una precisión de 64% y con 50% se obtuvo una precisión de 70%.
- Posteriormente se aplanó las imágenes con 32 filtros, equivalente a 32 neuronas Dense.
- La última capa de salida se estableció con 5 neuronas, se utilizó la función de activación Softmax, para que normalice la capa de salida porque se maneja múltiples clases, en este caso son 5 etiquetas a clasificar.
- Y finalmente se compila la red neuronal con el optimizador del algoritmo de Adam, que adapta el ratio de aprendizaje a los parámetros establecidos.

3.3.2.1 ENTRENAMIENTO DEL PROTOTIPO

Durante el proceso de entrenamiento y para un correcto funcionamiento del prototipo y de la red neuronal convolucional se usó imágenes de 120 x 120 pixeles, porque a una mayor dimensión ocasionaba desbordamiento de memoria, limitándonos a realizar pocas iteraciones o épocas.

La librería keras durante el entrenamiento utiliza diferentes redes neuronales y por cada época genera 4 métricas, las cuales mejoran con cada iteración (ver figura 29), las métricas son: precisión (*accuracy*), porcentaje de pérdida (*loss*), validación de la precisión (*val_accuracy*) y validación de pérdida (*val_loss*). Todos estos datos generan un histórico los cuales pueden ser presentados a través de una gráfica.

```
Epoch 1/2000
57/57 [=====] - 2s 33ms/step - loss: 1.4598 - accuracy: 0.4397 - val_loss: 1.2571 - val_accuracy: 0.5181
Epoch 2/2000
57/57 [=====] - 2s 30ms/step - loss: 1.3567 - accuracy: 0.4817 - val_loss: 1.2298 - val_accuracy: 0.5181
Epoch 3/2000
57/57 [=====] - 2s 30ms/step - loss: 1.3381 - accuracy: 0.4837 - val_loss: 1.2060 - val_accuracy: 0.5181
Epoch 4/2000
57/57 [=====] - 2s 30ms/step - loss: 1.3096 - accuracy: 0.4886 - val_loss: 1.1659 - val_accuracy: 0.5862
Epoch 5/2000
57/57 [=====] - 2s 30ms/step - loss: 1.2908 - accuracy: 0.4944 - val_loss: 1.1451 - val_accuracy: 0.5203
Epoch 6/2000
```

Figura 29. Resultados de entrenamiento por época

Se realizó diferentes iteraciones para obtener el mayor valor de precisión: 50, 80, 150, 250, 500, 1000, 2000 y 3000 épocas. A continuación se detallan las iteraciones realizadas:

- Con 50 épocas, la precisión en el entrenamiento se acercó casi al 70% y en la validación la precisión fue inferior al 65%. En los valores de pérdida en el entrenamiento llegó casi al 0.75% y en la validación se mantuvo casi constante cerca al 1% (ver figura 30).

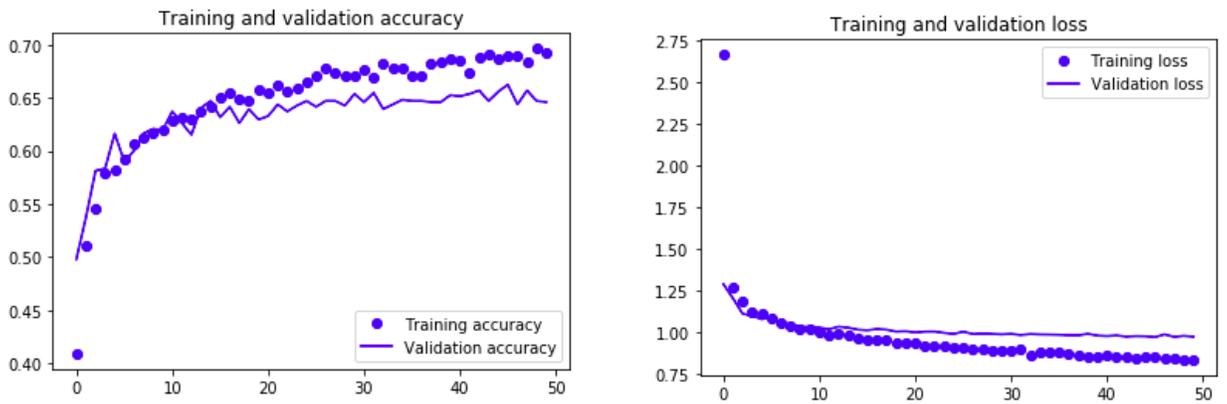


Figura 30. Exactitud, pérdida de entrenamiento y validación de 50 épocas

- Con 80 épocas, la precisión en el entrenamiento estuvo por encima del 70% y en la validación la precisión se mantuvo en 65%. En los valores de pérdida en el entrenamiento estuvo por debajo del 0.80% y en la validación estuvo por debajo del 1% (ver figura 31).

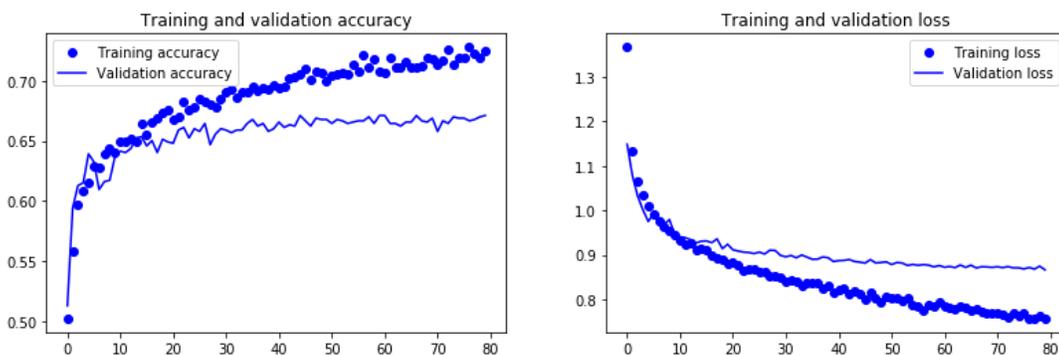


Figura 31. Exactitud, pérdida de entrenamiento y validación de 80 épocas

- Con 150 épocas, la precisión en el entrenamiento se acercó casi al 75% y en la validación la precisión estuvo por el 65%. En los valores de pérdida en el entrenamiento estuvo por debajo del 0.80% y en la validación se mantuvo casi constante cerca al 1% (ver figura 32).

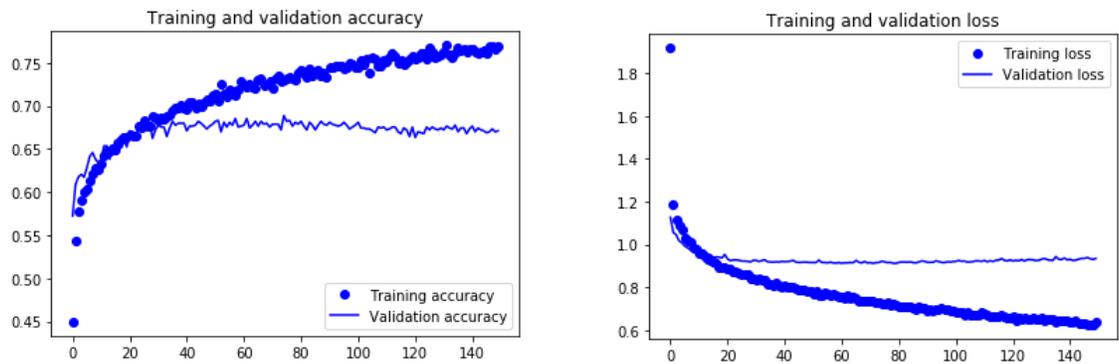


Figura 32. Exactitud, pérdida de entrenamiento y validación de 150 épocas

- Con 250 épocas, la precisión en el entrenamiento superó el 75% y en la validación la precisión estuvo por encima del 65%. En los valores de pérdida en el entrenamiento estuvo cerca del 0.60% y en la validación se mantuvo casi constante por debajo del 1% (ver figura 33).

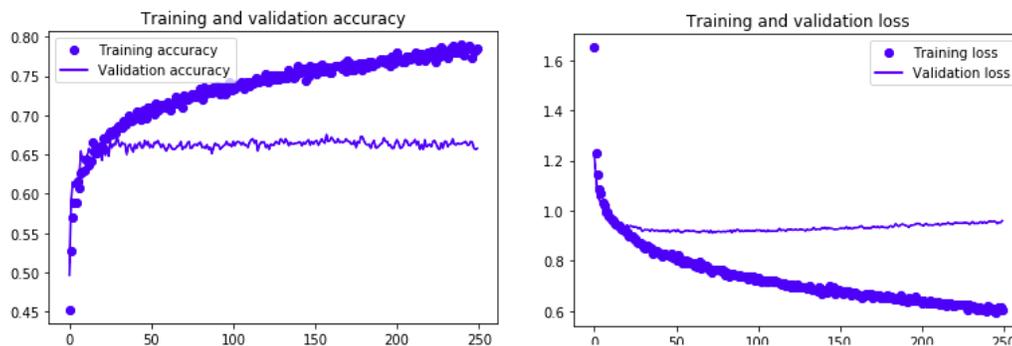


Figura 33. Exactitud, pérdida de entrenamiento y validación de 250 épocas

- Con 500 épocas, la precisión en el entrenamiento llego al 80% y en la validación la precisión se mantuvo por encima del 65%. En los valores de pérdida en el entrenamiento estuvo cerca del 0.60% y en la validación se mantuvo casi constante cerca al 1% (ver figura 34).

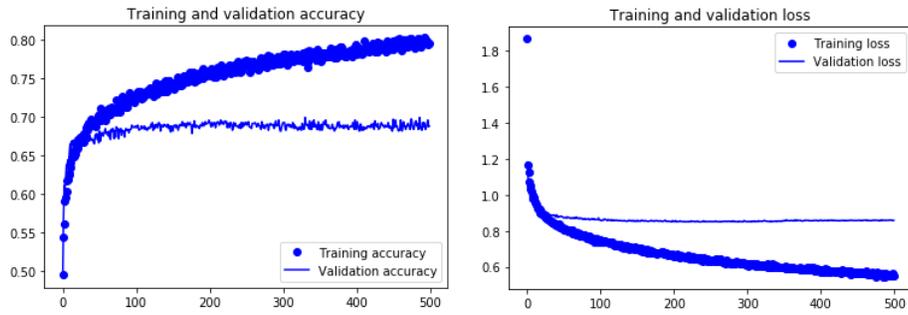


Figura 34. Exactitud, pérdida de entrenamiento y validación de 500 épocas

- Con 1000 épocas, la precisión en el entrenamiento llego al 85% y en la validación la precisión se mantuvo cerca al 65%. En los valores de pérdida en el entrenamiento estuvo por debajo del 0.60% y en la validación subió un poco más del 1% (ver figura 35).

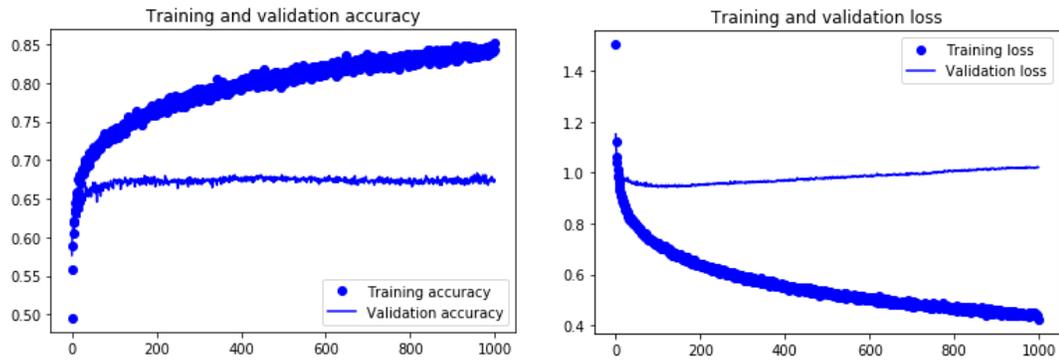


Figura 35. Exactitud, pérdida de entrenamiento y validación de 1000 épocas

- Con 2000 épocas, la precisión llegó hasta el 80% y la validación de la precisión estuvo casi hasta el 70%. En los valores de pérdida en el entrenamiento estuvo por debajo del 0.60% y en la validación subió un poco más del 1% (ver figura 36).

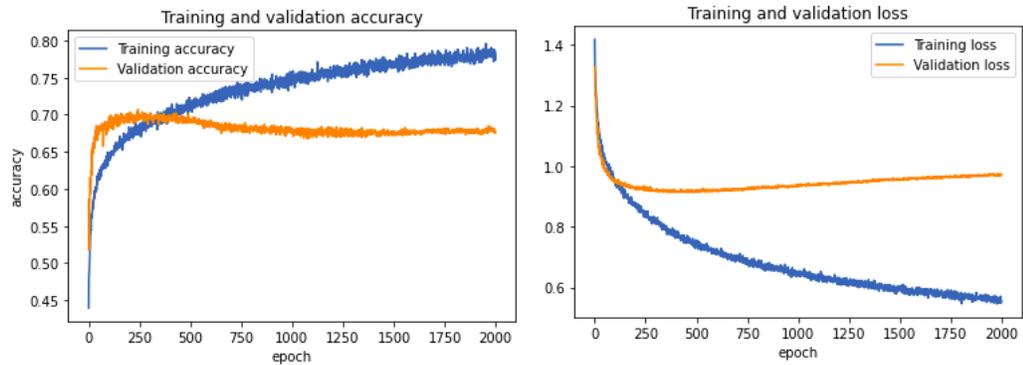


Figura 36. Exactitud, pérdida de entrenamiento y validación de 2000 épocas

- Con 3000 épocas, la precisión llegó hasta el 80% y la validación de la precisión estuvo por el 65%. En los valores de pérdida en el entrenamiento estuvo por debajo del 0.60% y en la validación subió un poco más del 1% (ver figura 37).

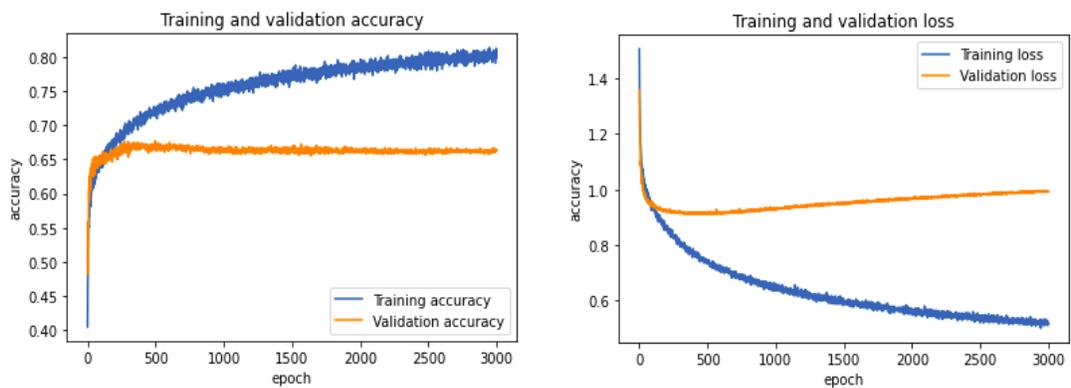


Figura 37. Exactitud, pérdida de entrenamiento y validación de 3000 épocas

De las iteraciones realizadas, a través de scikit-learn, se hizo una evaluación de cada modelo obtenido (ver tabla 10), que a un inicio se tenía un nivel de precisión muy bajo (29%) pero la precisión mejoró con la limpieza de las imágenes, donde la herramienta obtiene la precisión como se muestra en el siguiente listado (Thiest, 2017):

- Micro promedio, es el promedio de los positivos verdaderos totales, falsos negativos y falsos positivos, este valor recibe contribuciones de todas las clases para calcular la métrica promedio.

- Macro promedio, es el promedio de la media no ponderada por etiqueta (tiroiditis, bocio nodular, tiroides normal, bocio difuso, nódulo frío), que calcula la precisión independientemente para cada etiqueta, tratando a cada una de las etiquetas por igual.
- Promedio ponderado, es el promedio de la media ponderada por soporte por etiqueta, en el cálculo del promedio el puntaje de cada clase se pondera por su presencia en la muestra de datos reales o número de instancias verdaderas por cada etiqueta, este dato altera el macro promedio.

Tabla 10
Precisión calculada por scikit-learn

Nro. de épocas	Precisión		
	Micro Promedio	Macro Promedio	Promedio Ponderado
50	64	68	64
80	67	64	66
150	67	66	67
250	66	68	68
500	67	68	67
1000	69	69	69
2000	67	69	67
3000	67	69	66

De acuerdo a los datos que se tienen en la tabla 10, se decidió utilizar el modelo entrenado por 1000 épocas, porque tiene los valores de precisión más altos, es decir, que la precisión a nivel de cada clase fue el más alto (macro promedio), la precisión con contribución de otras clases también fue el más alto (micro promedio) y en la precisión de acuerdo al número de instancias por etiqueta también fue el más alto (promedio ponderado). El detalle de los valores de precisión que se obtuvo por cada clase se muestra en la tabla 11.

Tabla 11
Cálculo de métricas con scikit-learn

	Precisión	Cantidad de imágenes
Clase 0 (tiroiditis)	0.75	33
Clase 1 (bocio nodular)	0.69	136
Clase 2 (tiroides normal)	0.70	261
Clase 3 (bocio difuso)	0.59	164
Clase 4 (nódulo frío)	0.71	546
Micro promedio	0.69	1140
Macro promedio	0.69	1140
Promedio ponderado	0.69	1140

No se realizó más iteraciones porque el incremento en la cantidad de épocas no ayudaba a mejorar el valor de la precisión. Otro problema que se vio, fue que algunas imágenes de enfermedades tiroideas tienen cierto parecido y forman parte de otros grupos, esto también comenzaba a saturar a la red neuronal. La calidad de las imágenes podría ser otro factor que no permite mejorar la precisión durante la fase del entrenamiento del modelo.

3.3.2.2 MODELO DE CLASIFICACIÓN DE IMÁGENES CON CNN

Después del proceso de entrenamiento, el modelo más representativo que se obtuvo fue con mil épocas y con un valor de precisión de 70 %, donde el valor máximo alcanzado de precisión durante el entrenamiento estuvo cerca al 85%, a mayores iteraciones no se logró mejorar la precisión. Una vez finalizado el proceso de limpieza se cargaron las imágenes categorizándolas y convirtiendo los datos en matrices para las respectivas operaciones.

En la primera convolución se utilizó 32 filtros de 3x3 con función de activación LeakyRelu por tener mejor desempeño con las CNN, donde se obtiene información básica

de una imagen con bordes y sombras, cabe mencionar que la extracción de estas características son propias de la forma de la imagen y de los bordes extraídos en el proceso de limpieza.

La segunda convolución trabajó con 32 filtros de 3x3 con función de activación LeakyRelu, busca obtener características más específicas como texturas y contrastes. Debido a la calidad de las imágenes adicionar más capas de convolución no ayudó a extraer nuevas características para el proceso de entrenamiento.

Posteriormente se realizó el aplanamiento de las imágenes con 32 filtros equivalentes a 32 neuronas *dense* para poder utilizarlo como un vector de características en esta capa y como última capa se usó la función de activación softmax con 5 neuronas *dense* por las 5 etiquetas que se tienen, para normalizar la salida y este de acuerdo al “*hot encoding*” que se realizó en la configuración, gráficamente podemos representar el modelo como se observa en la figura 38.

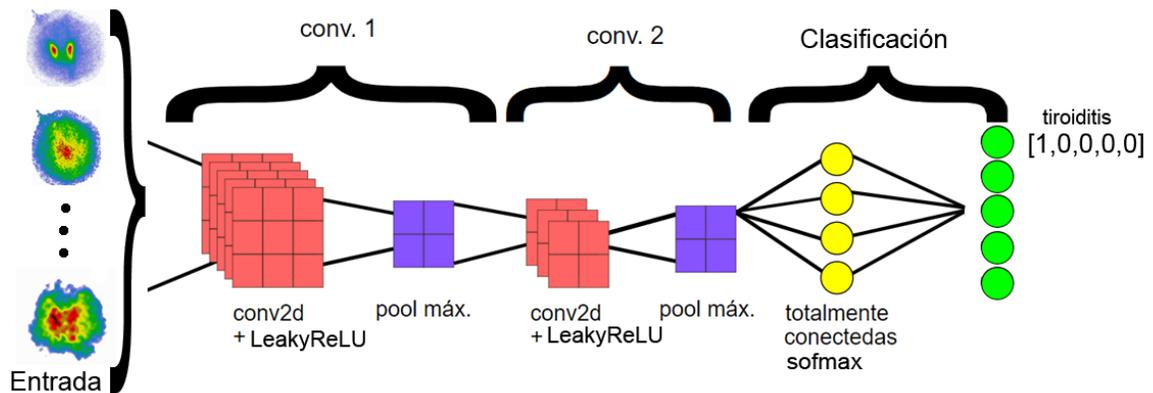


Figura 38. Modelo de clasificación de imágenes con CNN

Este modelo se utilizó para la etapa de evaluación con la matriz de confusión a través de sus respectivos indicadores: precisión, exactitud, sensibilidad y especificidad. Este modelo también fue usado para las pruebas de predicción comparando contra otras imágenes que no fueron tomadas en cuenta durante la fase de entrenamiento y pruebas.

CAPÍTULO IV

RESULTADOS

Una vez que se ha obtenido el modelo con el valor de precisión más alto, se realiza la evaluación del modelo a través del cálculo de los indicadores que se han definido: precisión, exactitud, especificidad y sensibilidad. Posteriormente se verifica su comportamiento al momento de hacer predicciones con imágenes nuevas.

4.1 EVALUACIÓN DEL MODELO

Una vez construido el modelo con el prototipo que obtuvo el mejor promedio de precisión, se decidió utilizar la matriz de confusión para ver el desempeño del algoritmo que se emplea en el aprendizaje supervisado.

Durante la fase del entrenamiento, las imágenes se organizaron en 5 grupos siguiendo la regla 80 – 20. De las 5.698 imágenes se utilizó el 20% para la evaluación, vale decir 1.140 imágenes, la cantidad de elementos por cada clase fue determinada por keras como se muestra en la tabla 12:

Tabla 12
Etiquetas con cantidades de elementos

Etiqueta	Nombre de etiqueta o clases	Cantidad imágenes
0	Tiroiditis	33
1	Bocio nodular	136
2	Tiroides normal	261
3	Bocio difuso	164
4	Nódulo frío	546

En la matriz de confusión las predicciones son las columnas y las filas son los valores reales (ver tabla 13), donde tenemos los siguientes valores:

- Positivos reales (*True positives* - TP), cuando el valor real coincide con la predicción, por ejemplo una imagen puede ser un nódulo frío y el modelo efectivamente predice que es un nódulo frío.
- Negativos reales (*True negatives* - TN), cuando se indica que un valor no pertenece a una clase y el predicción indica que no pertenece a esa clase, por ejemplo una imagen no es nódulo frío y el modelo predice que no es nódulo frío.
- Falsos negativos (*False negatives* - FN), cuando el valor pertenece a una clase pero el modelo predice que no pertenece, por ejemplo una imagen es un nódulo frío y el modelo predice que no es un nódulo frío, este error se traduce como una falta de detección temprana de la enfermedad.
- Falsos positivos (*False positives* - FP), cuando el valor no pertenece a una clase pero el modelo predice que si, por ejemplo una imagen no es un nódulo frío y el modelo predice que si es un nódulo frío, se harán pruebas adicionales que terminaran descartando el diagnóstico.

Tabla 13
Matriz de confusión

	Predicción	
	Positivo (modelo)	Negativo (modelo)
Positivo (real)	Positivos reales	Falsos negativos
Negativo (real)	Falsos positivos	Negativos reales

Mediante scikit-learn se ha realizado el cálculo de los respectivos valores para la matriz de confusión y el resultado obtenido se puede ver en la tabla 14:

Tabla 14
Matriz de confusión obtenida scikit-learn

	Tiroiditis	Bocio nodular	Tiroides normal	Bocio difuso	Nódulo frío
Tiroiditis	14	0	2	2	15
Bocio nodular	0	59	6	7	64
Tiroides normal	0	2	194	19	46
Bocio difuso	1	3	30	81	49
Nódulo frío	2	17	44	28	455

En esta matriz (tabla 14), la diagonal representa los casos efectivamente clasificados por cada etiqueta o clase (bocio nodular, tiroides normal, tiroiditis, bocio difuso y nódulo frío), estos valores representan los positivos reales (TP).

Por otro lado, para identificar los demás valores (negativos reales, falsos positivos y falsos negativos) se tiene que realizar los cálculos por cada una de las etiquetas o clases creadas.

Los valores de falsos positivos (FP), donde los casos pertenecen a una clase, pero el modelo predijo que no; se representan mediante la suma de las columnas sin contar los valores de las diagonales (ver tabla 15).

Tabla 15
Cálculo de falsos positivos

	Tiroiditis	Bocio nodular	Tiroides normal	Bocio difuso	Nódulo frío
Tiroiditis	14	0	2	2	15
Bocio Nodular	0	59	6	7	64
Tiroides normal	0	2	194	19	46
Bocio difuso	1	3	30	81	49
Nódulo frío	2	17	44	28	455
Falsos positivos	3	22	82	56	174

Los valores de falsos negativos (FN), donde los casos no pertenecen a una clase, pero el modelo predijo que sí; se representan mediante la suma de las filas sin contar los valores de las diagonales (ver tabla 16).

Tabla 16
Cálculo de falsos positivos

	Tiroiditis	Bocio nodular	Tiroides normal	Bocio difuso	Nódulo frío	Falsos negativos
Tiroiditis	14	0	2	2	15	19
Bocio Nodular	0	59	6	7	64	77
Tiroides Normal	0	2	194	19	46	67
Bocio Difuso	1	3	30	81	49	83
Nódulo Frío	2	17	44	28	455	91

Los valores de negativos reales (TN), cuando no pertenecen a una clase y el modelo predice de la misma manera, estos se calcularán mediante la sumatoria de valores. Para el caso de bocio nodular no se tomará en cuenta en los cálculos los valores de la fila y

columna de la clase (ver tabla 17), el ejercicio se debe repetir de la misma manera para las demás clases. El cálculo de los valores de negativos reales de todas las clases se observa en la tabla 18.

Tabla 17

Valores que no se tomaran en cuenta para bocio nodular

	Tiroiditis	Bocio nodular	Tiroides normal	Bocio difuso	Nódulo frío
Tiroiditis	14	0	2	2	15
Bocio Nodular	0	59	6	7	64
Tiroides normal	0	2	194	19	46
Bocio difuso	1	3	30	81	49
Nódulo frío	2	17	44	28	455

Tabla 18

Resultado de verdaderos negativos por clase

Clase o etiqueta	TN
Tiroiditis	1104
Bocio nodular	982
Tiroides normal	797
Bocio difuso	920
Nódulo frío	420

El indicador de exactitud del modelo se calcula mediante la suma de los valores correctamente identificados (TP) entre la suma del total de elementos de la matriz (ecuaciones 9 y 10)

$$\text{Exactitud} = \frac{\sum c_{ii}}{\sum \sum c_{ij}} = \frac{803}{1140} \quad (9)$$

$$\text{Exactitud} = 70,43 \quad (10)$$

Para el cálculo de los indicadores de sensibilidad (también llamada *recall*), especificidad y precisión del modelo, se tuvo que calcular por cada etiqueta y promediar los valores (ver tabla 19):

Tabla 19

Valores de sensibilidad, especificidad y precisión por etiqueta.

Clase o etiqueta	Sensibilidad (TP / (TP + FN))	Especificidad (TN / (TN + FP))	Precisión (TP / (TP + FP))
Tiroiditis (TI)	0,4242	0,9830	0,8335
Bocio nodular (BN)	0,4338	0,4897	0,7283
Tiroides normal (TN)	0,7432	0,9224	0,7028
Bocio difuso (BD)	0,4939	0,9172	0,5912
Nódulo frío (NF)	0,8333	0,6709	0,7233

Los indicadores anteriormente conseguidos, no se pueden promediar de forma común: sumar todos los valores y dividirlo entre el número de elementos, porque se tomaron diferentes cantidades de imágenes por clase (ver tabla 11) que no son proporcionales en cantidad por clase, para este caso se utiliza el promedio ponderado (PP) que toma cada valor obtenido por el número de instancias de la clase, los cálculos corresponden desde la ecuación 11 hasta la ecuación 19.

$$\text{Precisión PP} = (\text{instancias TI} * \text{precisión TI}) + (\text{instancias BN} \quad (11)$$

$$* \text{precisión BN}) + (\text{instancias TN} * \text{precisión TN}) + (\text{instancias BD} * \text{precisión bocio BD}) + (\text{instancias NF} * \text{precisión NF})$$

$$\text{Precisión PP} = \left(\frac{\# \text{ imágenes TI}}{\text{Total imágenes}} * \text{precisión TI} \right) + \quad (12)$$

$$\left(\frac{\# \text{ imágenes BN}}{\text{Total imágenes}} * \text{precisión BN} \right) + \left(\frac{\# \text{ imágenes TN}}{\text{Total imágenes}} * \right.$$

$$\text{precisión } TN) + \left(\frac{\# \text{ imágenes } BD}{\text{Total imágenes}} * \text{precisión } BD\right) +$$

$$\left(\frac{\# \text{ imágenes } NF}{\text{Total imágenes}} * \text{precisión } NF\right)$$

$$\text{Precisión PP} = \left(\frac{33}{1140} * 0,8335\right) + \left(\frac{136}{1140} * 0,7283\right) + \left(\frac{261}{1140} * 0,7028\right) + \left(\frac{164}{1140} * 0,5912\right) + \left(\frac{546}{1140} * 0,7233\right) \quad (13)$$

$$\text{Precisión PP} = 0,7031$$

$$\text{Sensibilidad PP} = \left(\frac{\# \text{ imágenes } TI}{\text{Total imágenes}} * \text{sensibilidad } TI\right) +$$

$$\left(\frac{\# \text{ imágenes } BN}{\text{Total imágenes}} * \text{sensibilidad } BN\right) + \left(\frac{\# \text{ imágenes } TN}{\text{Total imágenes}} * \text{sensibilidad } TN\right) +$$

$$\left(\frac{\# \text{ imágenes } BD}{\text{Total imágenes}} * \text{sensibilidad } BD\right) +$$

$$\left(\frac{\# \text{ imágenes } NF}{\text{Total imágenes}} * \text{sensibilidad } NF\right) \quad (14)$$

$$\text{Sensibilidad PP} = \left(\frac{33}{1140} * 0,4242\right) + \left(\frac{136}{1140} * 0,4338\right) +$$

$$\left(\frac{261}{1140} * 0,7432\right) + \left(\frac{164}{1140} * 0,4939\right) + \left(\frac{546}{1140} * 0,8333\right) \quad (15)$$

$$\text{Sensibilidad PP} = 0,7031 \quad (16)$$

$$\text{Especificidad PP} = \left(\frac{\# \text{ imágenes } TI}{\text{Total imágenes}} * \text{especificidad } TI\right) +$$

$$\left(\frac{\# \text{ imágenes } BN}{\text{Total imágenes}} * \text{especificidad } BN\right) + \left(\frac{\# \text{ imágenes } TN}{\text{Total imágenes}} * \text{especificidad } TN\right) +$$

$$\left(\frac{\# \text{ imágenes } BD}{\text{Total imágenes}} * \text{especificidad } BD\right) + \left(\frac{\# \text{ imágenes } NF}{\text{Total imágenes}} * \text{especificidad } NF\right) \quad (17)$$

$$\text{Especificidad PP} = \left(\frac{33}{1140} * 0,9830\right) + \left(\frac{136}{1140} * 0,4897\right) +$$

$$\left(\frac{261}{1140} * 0,9224\right) + \left(\frac{164}{1140} * 0,9172\right) + \left(\frac{546}{1140} * 0,6709\right) \quad (18)$$

Especificidad PP = 0,7514

(19)

Entonces a través de la matriz de confusión se obtuvieron los indicadores mencionados en la operacionalización de variables:

- Precisión, el modelo llegó a un 70,31 %.
- Sensibilidad, se obtuvo un 70,43 % de casos correctamente identificados.
- Especificidad, se obtuvo un 75,14 % de casos negativos clasificados correctamente.

Con la finalidad de comparar los datos que se obtuvo con la matriz de confusión, se fue buscando algunos recursos en línea, donde se encontraron 2 sitios web, los cuales permiten ingresar los valores de la matriz de confusión y determinar la precisión. En el primer sitio la precisión obtenida es de 70,44% (Matrix, 2018), en este caso el valor precisión es el mismo que el calculado con la matriz de confusión (ver figura 39).

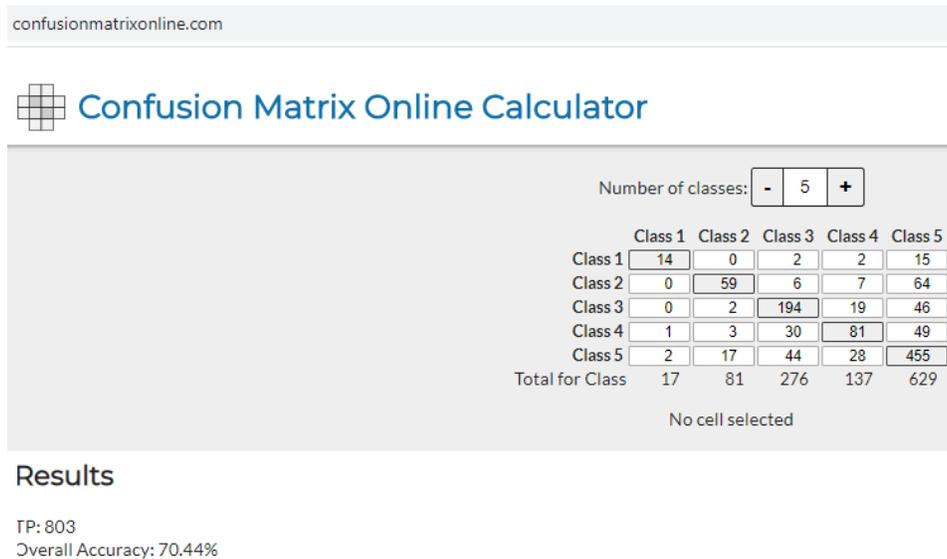


Figura 39. Cálculo de precisión con la matriz de confusión
Fuente: (Matrix, 2018)

Se encontró otra página del investigador en inteligencia artificial llamado Marco Vanetti, quien desarrolló una calculadora de matriz de confusión (Vanetti, 2007), donde el valor de la precisión que se obtuvo fue de 70,44% (ver figura 40), al igual que la anterior página y la precisión calculada mediante la matriz de confusión, el valor es el mismo.

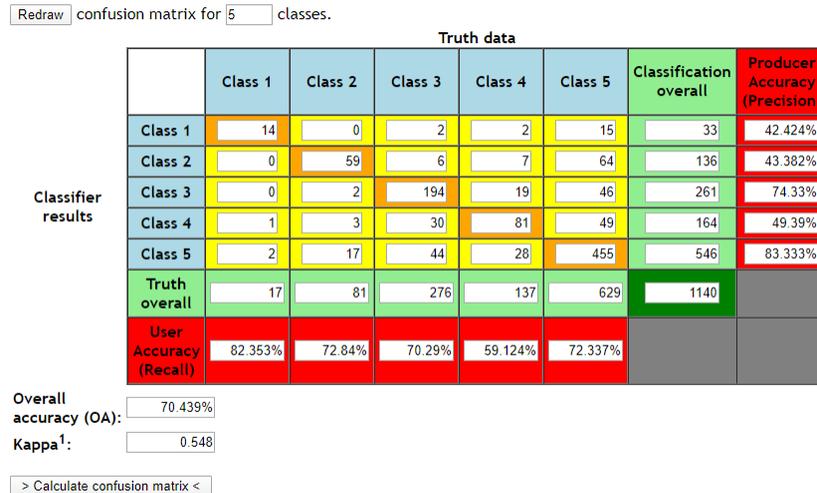


Figura 40. Cálculo de precisión con la matriz de confusión
Fuente: (Vanetti, 2007)

4.2 PREDICCIÓN

Una vez que se entrenó y se ajustó el modelo de la red neuronal, se utilizó el modelo generado durante la fase de entrenamiento, el cual a través de una aplicación desarrollada con Python (ver Anexo E), se proporcionó diferentes imágenes al modelo para que realice la predicción a cuál de las etiquetas pertenece. En la figura 41 se observa diferentes casos, donde las ubicaciones del arreglo de salida son los mismos de las etiquetas que se utilizó durante la fase de entrenamiento: tiroiditis (0), bocio nodular (1), tiroides normal (2), bocio difuso (3) y nódulo frío (4):

```
train_data_dir = os.path.join(main_dir, "prediccion3")
prediction_images(train_data_dir)

/home/bladimiro/Documents/maestria/python/inamen3/prediccion3/tiroiditis/
result [0.0000000e+00 0.0000000e+00 0.0000000e+00 1.0000000e+00 3.2533174e-24]
answer 3 (A)
/home/bladimiro/Documents/maestria/python/inamen3/prediccion3/tiroiditis/
result [0. 0. 0. 0. 1.]
answer 4 (B)
/home/bladimiro/Documents/maestria/python/inamen3/prediccion3/nodular/
result [0. 0. 0. 0. 1.]
answer 4 (C)
/home/bladimiro/Documents/maestria/python/inamen3/prediccion3/nodular/
result [0. 1. 0. 0. 0.]
answer 1
/home/bladimiro/Documents/maestria/python/inamen3/prediccion3/nodular/
result [0. 1. 0. 0. 0.]
answer 1
/home/bladimiro/Documents/maestria/python/inamen3/prediccion3/nodular/
result [0. 1. 0. 0. 0.]
answer 1
```

Figura 41. Resultados de predicción del modelo

La imagen que tiene la letra (A) de la figura 41, corresponde a una tiroides que tenía tiroiditis, pero en la predicción del modelo se identificó en un 100% como bocio difuso, en pequeños porcentajes se encontró similar a las otras etiquetas, esta predicción fue errada (ver figura 42).

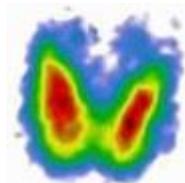


Figura 42. Tiroides con tiroiditis
Fuente: INAMEN

La imagen que tiene la letra (B), tampoco se predijo correctamente porque la glándula tenía tiroiditis pero el modelo indicó que era un nódulo frío (ver figura 43).

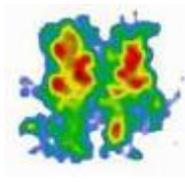


Figura 43. Tiroides con tiroiditis
Fuente: INAMEN

La imagen que tiene la letra (C), la predicción fue correcta porque la tiroides si tenía un bocio nodular (ver figura 44).

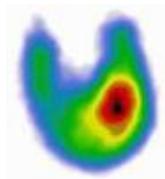


Figura 44. Tiroides con bocio nodular

Fuente: INAMEN

Adicionalmente a las 5.698 usadas en el entrenamiento y pruebas, se utilizó 50 imágenes para las pruebas de predicción (ver tabla 20), donde se obtuvo los siguientes resultados:

- En el caso de bocio nodular, solo 5 de las 10 imágenes fueron predicciones acertadas, en las otras 5 la predicción indicó que era un nódulo frío.
- En el caso de una tiroides normal, 8 imágenes fueron predicciones acertadas, en las otras 2 predicciones el modelo indicó que era un nódulo frío.
- En el caso de bocio difuso, debido a las características de las imágenes, 2 predicciones indicaron que era una tiroides normal, 2 predicciones fueron acertadas y las otras 6 la predicción indicaban nódulo frío.
- En el caso de los nódulos fríos, 15 de las predicciones fueron acertadas, 3 indicaron que era una tiroides normal.
- Para la tiroiditis ninguna predicción fue acertada.

Tabla 20
Comparación de predicciones utilizando imágenes

Datos Reales		Predicciones			
Diagnóstico	Total de casos	Bocio nodular	Tiroides normal	Bocio difuso	Nódulo frío
Bocio nodular	10	5			5
Tiroides normal	10		8		2
Bocio difuso	10		2	2	6
Nódulo frío	18		3		15
Tiroiditis	2			1	1

De los 50 casos de predicción, 30 fueron predicciones acertadas, lo que equivale a decir que 60 % fueron predicciones correctas.

De la información proporcionada se pudo ver que:

- El 50 % de los pacientes tienen nódulos fríos y necesitan más estudios para descartar que los nódulos no sean cancerígenos.
- El 21 % son pacientes con tiroides normal.
- El 11 % de los pacientes tienen bocio nodular
- El 15 % de los pacientes tienen nódulos difusos.
- El 3 % de los pacientes tienen tiroiditis.

4.3 PRUEBA DE HIPÓTESIS

Se utilizó la hipótesis nula, donde H_0 :

- H₀: El modelo basado en *Machine Learning* diagnostica con una precisión menor al 70% el tipo de nódulo tiroideo utilizando imágenes de centellografía.

Para el caso de *machine learning*, en lugar de la ecuación para el cálculo del valor crítico, se han utilizado las librerías de *keras* y *scipy* con el método de Anderson-Darling. Con el modelo obtenido mediante el método de predicción de *keras* se obtuvo un arreglo con los valores normalizados como se muestra en la figura 45.

```

▶ predictions = thyroid_model.predict(valid_X, verbose=0).flatten()
print(predictions.shape)
print(predictions)

(4555,)
[0.11970718 0.27637592 0.00384649 ... 0.00201238 0.30614656 0.01712583]

```

Figura 45. Predicción con valores normalizados

Con estos datos generados, el método Anderson-Darling utiliza por defecto la distribución normal, extrae una población, obtiene el estadístico de Anderson-Darling y realiza el cálculo para los valores críticos con cinco valores de significancia: 1, 2.5, 5, 10 y 15 (ver figura 46).

```

[21] result = anderson(predictions)
print('result', result)
print('statistic: %.3f', result.statistic)

result AndersonResult(statistic=512.7155967993476, critical_values=array([0.575, 0.655, 0.786, 0.917, 1.
statistic: %.3f 512.7155967993476

```

Figura 46. Valores obtenidos con el método de Anderson-Darling

Como se puede observar en la figura 47, a través del método de Anderson-Darling para los 5 valores de significancia el estadístico del método es mayor a los valores críticos, de esta manera se ha rechazado la hipótesis nula, por lo tanto la hipótesis planteada es verdadera.

```
[ ] for i in range(len(result.critical_values)):
    sl, cv = result.significance_level[i], result.critical_values[i]
    if result.statistic < result.critical_values[i]:
        print('%.3f: %.3f, los datos parecen normales (no se rechazan H0)' % (sl, cv))
    else:
        print('%.3f: %.3f, los datos no parecen normales (rechazar H0)' % (sl, cv))

15.000: 0.575, los datos no parecen normales (rechazar H0)
10.000: 0.655, los datos no parecen normales (rechazar H0)
5.000: 0.786, los datos no parecen normales (rechazar H0)
2.500: 0.917, los datos no parecen normales (rechazar H0)
1.000: 1.091, los datos no parecen normales (rechazar H0)
```

Figura 47. Comparación de valores críticos

CONCLUSIONES Y RECOMENDACIONES

Finalizado el desarrollo del presente trabajo, se describe el cumplimiento de los objetivos e hipótesis, asimismo se establece las conclusiones a las que se llegó y recomendaciones que se sugiere tomar en cuenta para investigaciones similares.

5.1 CONCLUSIONES

Respecto a los objetivos planteados en el presente trabajo, se concluye lo siguiente:

- Se logró describir el procedimiento que realiza el INAMEN para la obtención de imágenes con centellografía para que el médico especialista pueda analizar la imagen y elaborar el respectivo informe con el diagnóstico respectivo. Existen imágenes con un patrón claro que permite el diagnóstico del nódulo, sin embargo, algunas imágenes presentan una forma poco definida por lo que se requiere la experiencia del experto para un diagnóstico acertado.
- Se desarrolló el prototipo con el algoritmo de *machine learning* para que el modelo pueda clasificar las imágenes. Se hizo muchas pruebas hasta lograr un prototipo estable con una red neuronal del tipo CNN.
- Se desarrolló el modelo de clasificación de imágenes, buscando el mayor valor de precisión. Este fue evaluado y se obtuvo un valor de precisión del 70 % en general. Cabe hacer notar que debido a que la cantidad de imágenes de nódulos fríos es mayor, el modelo tiene mayor capacidad de identificar un nódulo frío; durante la fase de entrenamiento el valor de la precisión llegó al 85 % en estos casos.
- Respecto a la hipótesis planteada en la presente investigación, podemos indicar que el nivel de precisión al que se llegó con el modelo de diagnóstico para la

identificación de tipos de nódulos tiroideos utilizando *Machine Learning* fue del 70% considerando todas las imágenes de estudio.

- La centellografía es una técnica de imagenología que no proporciona una buena calidad en las imágenes para el entrenamiento del modelo con *machine learning*, esto representó una limitante al momento de intentar mejorar el nivel de precisión en el modelo de clasificación de imágenes. Si bien es una limitante, en nuestro medio es una de las técnicas de diagnóstico de uso frecuente en centros hospitalarios estatales.
- De la información recopilada y procesada se pudo observar que existen varios pacientes que tienen nódulos fríos, alrededor del 50%, que tienen tendencia a tener cáncer de tiroides, los cuales tendrían que ser revisados con más profundidad y evitar la invasión a otros órganos, lo que complicaría seriamente el estado de salud del paciente.

5.2 RECOMENDACIONES

- Debido a la complejidad del procesamiento de imágenes se recomienda tener cuidado respecto al estado en el que se encuentran estas antes de comenzar el trabajo. En algunos casos se requiere un proceso previo para obtener imágenes apropiadas. Esta tarea puede consistir en eliminar información que solo genera ruido (manchas, líneas, texto, bordes), de esta manera el procesamiento de las imágenes se mejora al momento de realizar el entrenamiento automático.
- El modelo obtenido puede ser utilizado como una herramienta para comparar los diagnósticos que realicen los médicos practicantes de la institución, puesto que aún se necesita la opinión profesional del especialista médico para determinar con mayor precisión el tipo de nódulo tiroideo que tenga el paciente.

- Se recomienda continuar con la investigación de clasificación de imágenes con *Machine Learning* con imágenes de ultrasonido, a fin de desarrollar herramientas que permitan apoyar a los médicos en el diagnóstico de enfermedades tiroideas y otras, especialmente a aquellas que resultan ser tan dañinas para la salud como el cáncer o para poder afrontar nuevas enfermedades como el coronavirus, donde no se pudo detectar fácilmente en los diagnósticos y saber cuáles eran los patrones de las anomalías que presentaban los diferentes órganos.
- Se recomienda ampliar la investigación usando algoritmos no supervisados, porque en el presente trabajo se utilizó imágenes que ya han sido etiquetadas y el modelo solo podrá predecir en base a estas etiquetas existentes, pero no podrá predecir que hay una nueva enfermedad tiroidea, en ese caso se podría utilizar un algoritmo de *clustering* para indicar que hay un nuevo grupo que no ha sido etiquetado.

ANEXOS

ANEXO A

OBTENCIÓN DE IMÁGENES DE DOCUMENTOS

A continuación se muestra parte del código que se utilizó para obtener las imágenes de los documentos Word.

```
protected void CopyFromClipbordInlineShape(String directory, String name)
{
    InlineShape inlineShape = m_word.ActiveDocument.InlineShapes[m_i];
    inlineShape.Select();
    m_word.Selection.Copy();
    Computer computer = new Computer();
    Image img = computer.Clipboard.GetImage();
    if (computer.Clipboard.GetDataObject() != null)
    {
        System.Windows.Forms.IDataObject data =
computer.Clipboard.GetDataObject();
        if (data.GetDataPresent(System.Windows.Forms.DataFormats.Bitmap))
        {
            System.Drawing.Image image =
(System.Drawing.Image)data.GetData(System.Windows.Forms.DataFormats.Bitmap,
true);
            if (!Directory.Exists(Server.MapPath("~/Images/") + directory)) {
                DirectoryInfo di =
Directory.CreateDirectory(Server.MapPath("~/Images/") + directory);
            }
            image.Save(Server.MapPath("~/Images/" + directory + "/" + name),
System.Drawing.Imaging.ImageFormat.Jpeg);
        }
        else
        {
            LabelMessage.Text = "The Data In Clipboard is not as image format";
        }
    }
    else
    {
        LabelMessage.Text = "The Clipboard was empty";
    }
}
```

ANEXO B

ORGANIZACIÓN DE IMÁGENES

A continuación se muestra parte del código utiliza para organizar imágenes de acuerdo a la conclusión del informe médico.

```
Microsoft.Office.Interop.Word.Document docs = m_word.Documents.Open(ref
FileName,
                                ref missing, ref readOnly, ref missing, ref missing,
                                ref missing, ref missing, ref missing, ref missing,
                                ref missing, ref missing, ref missing, ref missing, ref
missing, ref missing, ref missing);
    try
    {
        String conclu = string.Empty;
        String nompaci = string.Empty;
        String infor = string.Empty;

        int index1 = 0;
        for (int i = 0; i < docs.Paragraphs.Count; i++)
        {
            string temp = docs.Paragraphs[i + 1].Range.Text.Trim();
            if (temp != string.Empty)
            {
                index1 = temp.IndexOf("Conclusi");
                if (index1 >= 0)
                {
                    string sourcePath = Server.MapPath("Images") + "/" + fi.Name;
// + "/" + sfi.Name;

                    string targetPath = Server.MapPath("Images");
                    conclu = temp;
                    if (temp.IndexOf("nodulo frio") > 0 || temp.IndexOf("nódulo
frío") > 0 || temp.IndexOf("nódulos fríos") > 0 ||
                        temp.IndexOf("Nodulo Frio") > 0 || temp.IndexOf("Nódulo
Frío") > 0 || temp.IndexOf("Nódulos Fríos") > 0)
                    {
                        targetPath = targetPath + "\\nodulofrio";
                    }
                }
            }
        }
    }
}
```

```

else
{
    if (temp.IndexOf("bocio difuso") > 0 || temp.IndexOf("Bocio
Difuso") > 0 || temp.IndexOf("Bocio difuso") > 0)
    {
        targetPath = targetPath + "\\bociodifuso";
    }
    else
    {
        if (temp.IndexOf("bocio nodular") > 0 ||
temp.IndexOf("Bocio Nodular") > 0 || temp.IndexOf("Bocio nodular") > 0 ||
temp.IndexOf("bocio Nodular") > 0)
        {
            targetPath = targetPath + "\\bocionodular";
        }
        else
        {
            if (temp.IndexOf("tiroiditis") > 0 ||
temp.IndexOf("Tiroiditis") > 0)
            {
                targetPath = targetPath + "\\tiroiditis";
            }
            else
            {
                if (temp.IndexOf("sin alteraciones") > 0 ||
temp.IndexOf("Sin Alteraciones") > 0)
                {
                    targetPath = targetPath + "\\sinalteraciones";
                }
                else
                {
                    if (temp.IndexOf("minima concentracion") > 0 ||
temp.IndexOf("Mínima concentración") > 0 ||
temp.IndexOf("Mínima concentración") > 0 ||
temp.IndexOf("mínima concentración") > 0)
                    {
                        targetPath = targetPath +
"\\minimaconcentracion";
                    }
                    else
                    {

```

```

        if (temp.IndexOf("multinodular") > 0 ||
temp.IndexOf("Multinodular") > 0)
        {
            targetPath = targetPath + "\\multinodular";
        }
    }
}
}
}
}
int index = sfi.Name.IndexOf(".doc");
Console.Write(sfi.Name);
String namejpg = sfi.Name.Substring(0, index) + ".jpg";
Thread thread = new Thread(() =>
CopyFileNewAddress(sourcePath, targetPath, namejpg));
    thread.SetApartmentState(ApartmentState.STA);
    thread.Start();
    thread.Join();
}
data = data + temp;
}
}
}
finally
{
    object save = false;
    m_word.Quit(ref save, ref missing, ref missing);
    m_word = null;
}

```

ANEXO C

LIMPIEZA DE DATOS

A continuación se muestra el código para recortar las imágenes, quitando los bordes que no generan información relevante:

```
import cv2
from PIL import Image
from pathlib import Path

def cut_images(data_directory):
    dirs = [d for d in os.listdir(data_directory)
            if os.path.isdir(os.path.join(data_directory,d))]

    labels = []
    images = []
    for d in dirs:
        label_dir = os.path.join(data_directory, d)
        file_names = [os.path.join(label_dir, f)
                      for f in os.listdir(label_dir)
                      if f.endswith(".jpg")]
        for f in file_names:
            imgfile = Path(f)
            img1 = Image.open(imgfile)
            width = img1.size[0]
            height = img1.size[1]
            img3 = img1.crop((10,30,width-15,height-25) )
            img3.save(data_directory + '2\\' + d + '\\ ' + f[f.rfind("\\"):
])
```

ANEXO D

AJUSTE DE IMÁGENES

```
def resizeimage(data_directory):
    dirs = [d for d in os.listdir(data_directory)
            if os.path.isdir(os.path.join(data_directory,d))]
    for d in dirs:
        label_dir = os.path.join(data_directory, d)
        file_names = [os.path.join(label_dir, f)
                      for f in os.listdir(label_dir)
                      if f.endswith(".jpg")]
        counter = 0
        for f in file_names:
            imgfile = Path(f)
            img = Image.open(f)
            new_img = img.resize((120,120))
            url = data_directory + '2/' + d + '/' + f[f.rfind("/")+:]
            new_img.save(url)
            counter = counter + 1
    print('counter', counter)
```

ANEXO E

PRUEBAS DE PREDICCIÓN

A continuación se muestra el código que se utilizó para probar las predicciones con las nuevas imágenes.

```
import numpy as np
from keras.preprocessing.image import load_img, img_to_array
from keras.models import load_model
import tensorflow as tf
```

```
longitud, altura = 120, 120
modelo = './modelo/modelo1.h5'
pesos_modelo = './modelo/pesos1.h5'
cnn = tf.keras.models.load_model(modelo)
cnn.load_weights(pesos_modelo)
```

```
def predict(file):
    x = load_img(file, target_size=(longitud, altura))
    x = img_to_array(x)
    x = np.expand_dims(x, axis=0)
    array = cnn.predict(x)
    result = array[0]
    answer = np.argmax(result)
    print("answer", answer)
```

```
def prediction_images(path):
    file_names = [os.path.join(path, f)
                  for f in os.listdir(path)
                  if f.endswith(".jpg")]
    for f in file_names:
        predict(f)
```

BIBLIOGRAFÍA

- AECAT. (2014). *Impacto Social de las enfermedades tiroideas en España*. Obtenido de Asociación Española de Cáncer de Tiroides: <https://www.aecat.net/wp-content/uploads/2014/07/Analisis-Impacto-Social-de-las-Enfermedades-Tiroideas-.pdf>
- Agustin, S., & Babu, S. (2013). *A Review of Thyroid Disorder Detection, Segmentation and Classification on Medical Images*. Obtenido de Semantic Scholar: <https://pdfs.semanticscholar.org/e904/30636e49cc3172caa1ead71dfacac7737ae7.pdf>
- American Cancer Society. (31 de Marzo de 2016). *Qué es cáncer de tiroides*. Obtenido de American Cancer Society: <https://www.cancer.org/es/cancer/cancer-de-tiroides/acerca/que-es-cancer-de-tiroides.html>
- Barragan, L. M. (2004). *Guia de Aplicaciones Clínicas de Medicina Nuclear*. La Paz - Bolivia.
- Brandan, N., Llanos, I., Horak, F., Tannuri, H., & Rodríguez, A. (2014). *Hormonas Tiroideas*. Obtenido de Universidad Nacional del Nordeste: [https://med.unne.edu.ar/sitio/multimedia/imagenes/ckfinder/files/files/Carrera-Medicina/BIOQUIMICA/hormona%20tiroidea%202014\(1\).pdf](https://med.unne.edu.ar/sitio/multimedia/imagenes/ckfinder/files/files/Carrera-Medicina/BIOQUIMICA/hormona%20tiroidea%202014(1).pdf)
- Brownlee, J. (13 de septiembre de 2019). *Drop Regularization in Deep Learning Models with Keras*. Obtenido de Machine Learning Mastery: <https://machinelearningmastery.com/dropout-regularization-deep-learning-models-keras/>
- Calvo, D. (20 de julio de 2017). *Red Neuronal Convolutacional*. Obtenido de Red Neuronal Convolutacional: <http://www.diegocalvo.es/red-neuronal-convolutacional/>
- Cáncer. (27 de Abril de 2018). *Estadísticas del Cáncer*. Obtenido de Instituto Nacional del Cáncer: <https://www.cancer.gov/espanol/cancer/naturaleza/estadisticas>

- Castillero, O. (15 de Noviembre de 2019). *Corteza visual del cerebro*. Obtenido de Psicología y mente: <https://psicologiymente.com/neurociencias/corteza-visual>
- Chambi, M. C. (2014). *Sistema experto par el diagnóstico de cáncer tiroideo basados en redes neuronales*. La Paz - Bolivia.
- Clinic. (18 de febrero de 2017). *Nódulos Tiroideos*. Obtenido de Mayo Clinic: <https://www.mayoclinic.org/es-es/diseases-conditions/thyroid-nodules/symptoms-causes/syc-20355262>
- Contreras, F. (2016). *Introduccion a Machine Learning*. Obtenido de zemsania: https://www.zemsania.com/recursos-zemsania/whitepapers/DTS/Machine_learning.pdf
- Cooley, J. (7 de diciembre de 2018). *Redes neuronales convolucionales*. Obtenido de Redes neuronales convolucionales: Redes neuronales convolucionales
- Cooper, D., & Ladenson, P. (2015). *Access Medicina*. Obtenido de Access Medicina: <https://accessmedicina.mhmedical.com/content.aspx?bookid=1497§ionid=92779964>
- Crispi, H. (1 de junio de 2019). *Redes Neuronales Convolucionales*. Obtenido de Medium: https://medium.com/@jcrispis56/introducci%C3%B3n-al-deep-learning-parte-2-redes-neuronales-convolucionales-f743266d22a0?postPublishedType=initial&fbclid=IwAR3RcbURx9Sjwua5Glg_RJZqt8AncCqt336Ml6Yynp_rQ9PLA_AhyhbD5Ns
- FAO. (2004). *FAO*. Obtenido de Examen: <http://www.fao.org/3/y4666s0c.htm>
- Fernandez, R. S., Fernandez, C. C., & Baptista, P. L. (2014). *Metodologia de la Investigación*. Mexico DF: McGraw Hill.
- Franco, C., Pardo, F., Laborda, R., & Pérez, C. (2017). Aplicación de la técnica de árboles de clasificación y regresión en la valoración ecográfica de los nódulos tiroideos. *Revista Argentina de Radiología*, 11.

- Geron, A. (2017). *Hands-On Machine Learning with Scikit-learn & Tensorflow*. O'Reilly.
- Gomar, J. (9 de 10 de 2018). *Nvidia CUDA Cores*. Obtenido de Profesional Review: <https://www.profesionalreview.com/2018/10/09/que-son-nvidia-cuda-core/>
- Gonzales, A. (24 de Enero de 2019). *Conceptos básicos de Machine Learning*. Obtenido de Cleverdata: <https://cleverdata.io/conceptos-basicos-machine-learning/>
- Hurwitz, J., & Kirsch, D. (2018). *Machine Learning for dummies*. United States: IBM.
- INC1. (2019). *Diccionario del Cáncer*. Obtenido de Instituto Nacional de Cáncer: <https://www.cancer.gov/espanol/publicaciones/diccionario/def/sensibilidad>
- INC2. (2019). *Especificidad*. Obtenido de Instituto Nacional de Cáncer: <https://www.cancer.gov/espanol/publicaciones/diccionario/def/especificidad>
- Instituto del Cáncer. (2018). *Instituto Nacional de Cáncer*. Obtenido de Diccionario del cáncer: <https://www.cancer.gov/espanol/publicaciones/diccionario/def/centellografia>
- La voz de Galicia. (2014). *Bocio afecta la población y habitantes de zonas montañosas*. Obtenido de La voz de Galicia: <https://www.lavozdegalicia.es/noticia/lugo/lugo/2014/09/03/bocio-afecta-128-poblacion-30-habitantes-zonas-montanosas-deficit-yodo/00031409768617772166240.htm>
- Lopez, B. R. (2 de agosto de 2016). *Redes neuronales convolucionales*. Obtenido de Redes neuronales convolucionales: <https://relopezbriega.github.io/blog/2016/08/02/redes-neuronales-convolucionales-con-tensorflow/>
- Los Tiempos. (27 de Marzo de 2017). *Preguntas sobre la tiroides*. Obtenido de Los tiempos: <http://www.lostiempos.com/oh/tendencias/20170327/10-preguntas-tiroides>

- Luna, J. G. (8 de febrero de 2018). *Medium*. Obtenido de Tipos de aprendizaje automático: <https://medium.com/soldai/tipos-de-aprendizaje-autom%C3%A1tico-6413e3c615e2>
- Matich, D. (2001). *Redes Neuronales: Conceptos Básicos y Aplicaciones*. 55.
- Matrix. (2018). *Confusion Matrix Online*. Obtenido de Confusion Matrix Online: confusionmatrixonline.com
- Merck. (29 de junio de 2017). *Conoce síntomas del nódulo tiroideo*. Obtenido de Cuida tu tiroides: <https://www.cuidatutiroides.com/t/sintomas-nodulo-tiroideo/>
- Minitab. (2019). *¿Qué es una prueba de hipótesis?* Obtenido de minitab.com: <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/basics/what-is-a-hypothesis-test/>
- Minitab2. (2019). *¿Qué es el estadístico de Anderson-Darling?* Obtenido de Minitab.com: <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/normality/the-anderson-darling-statistic/>
- Murillo, B. (10 de julio de 2019). Diagnóstico de nódulos tiroideos. (B. Calcina, Entrevistador)
- Murillo2, B. (20 de diciembre de 2018). Recopilación de imágenes. (B. Calcina, Entrevistador)
- Nasser, J. (18 de 6 de 2019). *Conocer nódulos tiroideos*. Obtenido de Temas sobre salud: <https://temassobresalud.com/nodulos-tiroideos/>
- Nist. (2018). *Anderson-Darling Test*. Obtenido de nist.gov: <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htm#:~:text=The%20test%20is%20a%20one,the%20sample%20size%2C%20n>.
- Orrego, A. (2012). *Fundamentos de Medicina Endocrinología*. Medellín - Colombia: Corporación de Investigaciones Biológicas.

- Poudel, P., Illanes, A., Sheet, D., & Friebe, M. (2018). Evaluation of Commonly Used Algorithms for Thyroid Ultrasound Images Segmentation and Improvement Using Machine Learning Approaches. *Journal of Healthcare Engineering*, 13.
- Recuero, P. (23 de Enero de 2018). *Machine Learning a tu alcance: La matriz de confusión*. Obtenido de Luca AI: <https://data-speaks.luca-d3.com/2018/01/ML-a-tu-alcance-matriz-confusion.html>
- Rivera, M. (1 de agosto de 2020). *Centro de Investigación de Matemáticas*. Obtenido de Red Neuronal Multicapa en Keras: http://personal.cimat.mx:8181/~mrivera/cursos/aprendizaje_profundo/nn_multicapa/nn_multicapa.html
- Rodriguez, J. M. (17 de junio de 2013). *Clasificación de Texto: medidas de rendimiento y desempeño*. Obtenido de Blogspot: <http://pdln.blogspot.com/2013/06/clasificacion-de-texto-medidas-de.html>
- Roman, V. (6 de 2 de 2019). *Introduccion Machine Learning*. Obtenido de Medium: <https://medium.com/datos-y-ciencia/introduccion-al-machine-learning-una-gu%C3%ADa-desde-cero-b696a2ead359>
- Romero, J., Dafonte, C., Gómez, Á., & Penousal, F. (2007). *Inteligencia Artificial y Computación Avanzada*. Santiago de Compostela.
- Saha, S. (15 de diciembre de 2018). *A Comprehensive Guide to Convolutional Neural Networks*. Obtenido de Towards Data Science: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- Sajama, J. (1 de junio de 2012). *Samajeras*. Obtenido de Proceso de Investigacion: <https://samajeras.files.wordpress.com/2012/06/proceso-de-investigacion3b3n.pdf>
- Sancho, F. C. (26 de diciembre de 2018). *Redes Neuronales: una visión supervicial*. Obtenido de Blog investigación: <http://www.cs.us.es/~fsancho/?e=72>

- Scikit-learn. (2019). *Model evaluation: quantifying the quality of predictions*. Obtenido de Scikit-learn: https://scikit-learn.org/stable/modules/model_evaluation.html#classification-report
- Thiest. (15 de Octubre de 2017). *Data Science Stack exchange*. Obtenido de Micro Average vs Macro average Performance in a Multiclass classification: <https://datascience.stackexchange.com/questions/15989/micro-average-vs-macro-average-performance-in-a-multiclass-classification-settin>
- Toro, L. (14 de septiembre de 2017). *Anaconda Distribution*. Obtenido de Desde Linux: <https://blog.desdelinux.net/ciencia-de-datos-con-python/>
- Utrera, J. B. (20 de Junio de 2018). *Deep Learning básico con Keras*. Obtenido de En mi local funciona: <https://enmilocalfunciona.io/deep-learning-basico-con-keras-parte-1/>
- Vanetti, M. (2007). *Confusion matrix online calculator*. Obtenido de Marco Vanetti: <http://www.marcovanetti.com/pages/cfmatrix/?noc=5>
- Ziessman, H., O'Malley, J., & Thrall, J. (2007). *Medicina Nuclear - Los Requisitos*. Elsevier Mosby.
- Zita, A. (2019). *Exactitud y precisión*. Obtenido de Diferenciador: <https://www.diferenciador.com/diferencia-entre-exactitud-y-precision/>