



**UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE QUITO
CARRERA DE BIOMEDICINA**

**SISTEMA DE RECUPERACIÓN DE IMÁGENES HISTOPATOLÓGICAS BASADO
EN CONTENIDO VISUAL**

**Trabajo de titulación previo a la obtención del título de:
INGENIERO BIOMÉDICO**

AUTOR: ESTEBAN JAVIER GUALOTUÑA SUQUILLO

TUTOR: FABIÁN RODRIGO NARVÁEZ ESPINOZA, Ph.D

Quito-Ecuador

2024

**CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE
TITULACIÓN**

Yo, Esteban Javier Gualotuña Suquillo con documento de identificación No. 1727682963 manifiesto que:

Soy el autor y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Quito, 23 de febrero del año 2024

Atentamente,



Esteban Javier Gualotuña Suquillo
1727682963

**CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE
TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Yo, Esteban Javier Gualotuña Suquillo con documento de identificación No. 1727682963, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor del Proyecto Técnico: “Sistema de recuperación de imágenes histopatológicas basado en contenido visual”, el cual ha sido desarrollado para optar por el título de: Ingeniero Biomédico, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Quito, 23 de febrero del año 2024

Atentamente,



Esteban Javier Gualotuña Suquillo
1727682963

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Fabián Rodrigo Narvárez Espinoza con documento de identificación No. 0103674677, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: Sistema de recuperación de imágenes histopatológicas basado en contenido visual realizado por Esteban Javier Gualotuña Suquillo con documento de identificación No. 1727682963 obteniendo como resultado final el trabajo de titulación bajo la opción Proyecto Técnico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Quito, 23 de febrero del año 2024

Atentamente,



Fabián Rodrigo Narvárez Espinoza, Ph.D
0103674677

Dedicatoria

Dedico mi trabajo de titulación a Dios, quien ha sido mi guía y fortaleza en cada paso de mi camino. En especial a mis queridos padres, cuyo amor incondicional, apoyo constante y sacrificio han sido la fuerza motriz detrás de mis logros. A través de su ejemplo y aliento, han sido mi inspiración para alcanzar mis metas y perseguir mis sueños.

”El éxito es la suma de pequeños esfuerzos repetidos día tras día”

Robert Collier

Agradecimientos

A Dios, quien me ha brindado la fortaleza necesaria para alcanzar este logro y hacer realidad el sueño de convertirme en un profesional. Agradezco a mi padre, Aníbal Gualotuña, por su incansable esfuerzo y dedicación que han sido pilares fundamentales en este proceso universitario, su apoyo constante y sacrificio que hizo para que nada me faltara. A mi madre, Julia Suquillo, le doy gracias por su amor inquebrantable, comprensión y por permanecer a mi lado en los momentos más difíciles, brindándome el aliento necesario para seguir adelante. A mi hermano Alex, quien no solo ha sido mi hermano, sino también un amigo y un modelo a seguir. A todos ustedes, les expreso mi más profundo agradecimiento por su apoyo y guía en este viaje hacia el éxito.

Agradeciendo también a mi tutor, el Dr. Fabian Rodrigo Narvárez Espinoza Ph.D. Por su apoyo, orientación, valiosos consejos, arduo empeño y colaboración en el desarrollo de este trabajo por el cual hoy se convierte en un sueño hecho realidad.

Resumen

Este trabajo presenta el desarrollo de un sistema de recuperación de imágenes histopatológicas basado en su contenido visual, que sirva como una herramienta esencial para mejorar la accesibilidad en la búsqueda e interpretación de imágenes histopatológicas, acorde al principio de medicina basada en evidencia. El proceso de construcción y funcionamiento de este sistema consiste en cuatro etapas fundamentales.

Inicialmente, se procede a la construcción de un repositorio que contenga imágenes de diapositivas completas (WSI- Whole slide Images). Una vez las imágenes son digitalizadas, se establece una metodología que mejoran el contenido visual de las imágenes mediante técnicas adecuadas de procesamiento para la posterior extracción de características visuales de bajo nivel.

Seguido, se implementa un modelo de aprendizaje de máquina diseñado para aprender y extraer las características visuales relevantes. Este modelo desempeña un papel crucial en la fase de búsqueda del sistema CBIR basado en comparación de las características de bajo nivel, asegurando resultados más precisos y eficientes.

Finalmente, se evalúa el desempeño del sistema en funcionamiento, este análisis busca no solo mejorar la práctica de la histopatología, sino también contribuir al avance del conocimiento en medicina basada en evidencia, la investigación busca brindar un aporte significativo en el campo de la microscopía virtual.

Palabras clave: Microscopía Virtual, Whole Slide Imaging, Sistemas de recuperación de Imágenes, Redes Neuronales Convolucionales Inteligencia Artificial.

Abstract

This work presents the development of a histopathological image retrieval system based on its visual content, which serves as an essential tool to improve search accessibility, interpretation of histopathological images, which are guided by the principle of evidence-based medicine. The construction and operation process of this system consists of four fundamental stages.

Initially, a repository containing whole slide images (WSI - Whole slide Images) is built. Once the images are digitized, a methodology is established that enhances the visual content of the images by means of appropriate processing techniques for the subsequent extraction of low-level visual features.

Next, a machine learning model designed to learn and extract the relevant visual features is implemented. This model plays a crucial role in the search phase of the CBIR system based on low-level feature comparison, ensuring more accurate and efficient results.

Finally, the performance of the system in operation is evaluated. This analysis aims not only to improve the practice of histopathology, but also to contribute to the advancement of knowledge in evidence-based medicine. The research aims to make a significant contribution to the field of virtual microscopy.

Contenido

| | |
|--|-------------|
| Agradecimientos | vii |
| Resumen | viii |
| 1 Introducción | 1 |
| 1.1 Importancia y Alcance | 2 |
| 1.2 Objetivos | 3 |
| 1.2.1 Objetivo General | 3 |
| 1.2.2 Objetivos Específicos | 3 |
| 2 Capítulo 1 | 4 |
| 2.1 Muestras Histológicas | 4 |
| 2.2 La Microscopía Virtual | 5 |
| 2.3 (WSI) Whole Slide Imaging | 6 |
| 2.4 Sistemas de Recuperación basado en contenido (CBIR) | 7 |
| 2.4.1 Arquitectura de los Sistemas CBIR | 9 |
| 2.5 Aprendizaje de Máquina | 11 |
| 2.6 Modelos de Inteligencia Artificial que funcionan en los CBIR | 12 |
| 2.7 Medidas de Similitud | 15 |
| 3 Capítulo 2 | 17 |
| 3.1 Construcción de la base de datos con imágenes WSI | 17 |
| 3.2 Normalización de la base de datos | 20 |
| 3.3 Pre-procesamiento de las imágenes | 22 |
| 3.4 Extracción de Características | 24 |
| 3.5 Algoritmo K-NN | 25 |
| 3.6 Implementación del sistema de recuperación basado en contenido visual. | 27 |
| 4 Capítulo 3 | 32 |
| 4.1 Evaluación y Resultados | 32 |
| 4.1.1 Comparación del método | 36 |
| 5 Conclusiones y recomendaciones | 37 |
| 5.1 Conclusiones | 37 |
| 5.2 Recomendaciones | 37 |

Bibliografía

39

1 Introducción

El avance continuo en el campo de la histopatología demanda soluciones innovadoras que mejoren la accesibilidad y la interpretación de imágenes histopatológicas fundamentales para la toma de decisiones en medicina basada en evidencia. Este trabajo se encamina en abordar la problemática actual relacionada con la limitada presencia de desarrollos tecnológicos locales en histopatología, específicamente en el acceso remoto a imágenes de histología.

A pesar de la elevada demanda de tecnología para el acceso remoto a imágenes histológicas, se ha identificado una escasa presencia en la literatura de desarrollos locales. La mayoría de los estudios existentes se centran predominantemente en aplicaciones diagnósticas generales y con aplicaciones académicas, sin embargo, no hay reportes locales. Esta dificultad se atribuye a la falta de desarrollos tecnológicos en la región que permitan navegar de manera remota sobre este tipo de imágenes, un campo conocido como microscopía virtual, así como en las limitaciones de costos y de dependencia de accesibilidad a servidores remotos. Por esta razón, es importante desarrollar soluciones que aborden estas limitaciones, especialmente en áreas que no cuenten con expertos para la interpretación y diagnóstico de este tipo de imágenes [40].

Por otro lado, los desarrollos basados en modelos con Inteligencia Artificial (IA) actualmente son capaces de automatizar los procesos diagnósticos más aún cuando se trata de recuperar información entre grandes volúmenes de datos, específicamente basado en contenido visual. Es por eso que, en el campo de la microscopía virtual, ha generado una nueva área de investigación, evitando así los errores humanos y optimizando el tiempo aplicado para el desempeño de estos procesos.

Actualmente la investigación en el área de recuperación de imágenes por contenido se orienta al estudio de métodos y técnicas que facilitan el acceso mediante la información visual [5]. Este acceso no se restringe únicamente a un nivel conceptual por el dominio de aplicación, sino que también se amplía a un nivel perceptual mediante la aplicación de medidas objetivas para evaluar el contenido visual. Se han desarrollado diversas implementaciones de sistemas para la recuperación por contenido CBIR, sin embargo, el desempeño de estos sistemas es variable por la cantidad de imágenes procesadas bajo este esquema y por la variabilidad de las estructuras que se visualizan con estas imágenes, por lo que aún sigue en evaluación.

Frente a estos desafíos la investigación se enmarca en la realidad del análisis histológico ac-

tual, donde las limitaciones del enfoque manual y la necesidad de adaptarse a los cambios tecnológicos impulsan la búsqueda de un modelo basado en inteligencia artificial. Este modelo no solo busca resolver los desafíos existentes, sino también sentar las bases para una búsqueda, acceso e interpretación más eficiente y precisa de imágenes histológicas contenidas en repositorios, con impactos primordiales en la práctica e investigación biomédica.

Se plantea una necesidad fundamental para la construcción de un sistema de recuperación de imágenes basado contenido visual. Este sistema aprovechará algoritmos de inteligencia artificial para la búsqueda y recuperación de imágenes histopatológicas en grandes volúmenes. La implementación de CBIR permitirá a los profesionales de la salud buscar imágenes basadas en similitud visual, superando las limitaciones de los descriptores basados en texto. Se emplearán técnicas avanzadas de procesamiento de imágenes para mejorar la precisión de la búsqueda y la capacidad del sistema para reconocer patrones complejos en las imágenes histopatológicas.

Por otro lado, se sugiere complementar la solución con el desarrollo un modelo de inteligencia artificial basados en redes neuronales con propósitos similares. Es así que se busca comprender las prácticas exitosas que se han aplicado en dichos contextos. El proceso de revisión literaria no solo enriquecerá la comprensión teórica, sino que también proporcionará un marco sólido para ajustar los parámetros específicos de la investigación.

1.1. Importancia y Alcance

La demanda de la gran cantidad de imágenes histopatológicas en los centros hospitalarios exige la implementación de un esquema basado en modelos de aprendizaje máquina para la recuperación de este tipo de imágenes, siendo una herramienta importante para soportar el diagnóstico, posibilitando accesos más rápidos y precisos a grandes volúmenes de imágenes para los profesionales de la salud. Esto se traduce en una mejora general de los procesos diagnósticos al reducir los tiempos para la toma de decisiones soportados con un conjunto de casos previamente diagnosticados. Por otra parte, las técnicas de procesamiento de imágenes ayudan a identificar patrones complejos en las imágenes, que se asocian a la presencia de enfermedades, proporcionando a los especialistas de la salud información detallada y específica para una toma de decisiones más informada.

Por otro lado, este esquema de trabajo no solo contribuirá a los desarrollos locales en el área de la microscopia virtual, sino que también ampliará las posibilidades para entrenamiento de nuevos profesionales en esta misma área. Esto debido al uso de casos previamente diagnosticados y recuperados de una base de datos. Finalmente, fomentará la colaboración entre profesionales médicos e investigadores, mejorando la capacidad de predecir riesgos y definir

tratamientos de medicina basada en evidencia y análisis de datos.

1.2. Objetivos

1.2.1. Objetivo General

Implementar un sistema de recuperación de imágenes histopatológicas basado en contenido visual.

1.2.2. Objetivos Específicos

- Revisar la literatura acerca de los métodos y modelos de recuperación de imágenes basado en contenido visual.
- Construir una base de datos que contenga imágenes histológicas.
- Implementar técnicas de procesamiento de imágenes para describir el contenido visual de las mismas.
- Implementar un modelo de aprendizaje de máquina que permita recuperar imágenes histopatológicas basadas en su contenido visual.
- Evaluar el desempeño de los modelos implementados en el sistema de recuperación de imágenes basados en el contenido visual.

2 Capítulo 1

Marco teórico

En este capítulo se describe el fundamento teórico como un sustento para el desarrollo del sistema propuesto. El mismo que, brinda un conocimiento detallado acerca de los avances y tecnologías actualmente disponibles en el área de investigación. Básicamente, se enfoca en diversos aspectos fundamentales que proporcionan el contexto necesario para el desarrollo y comprensión del sistema.

2.1. Muestras Histológicas

Con las muestras histológicas se visualiza la estructura y composición de tejidos a nivel celular en el ámbito biológico. Se obtienen mediante el empleo de técnicas de coloración como la hematoxilina y eosina, técnicas usadas para el estudio de tejidos y células a nivel microscópico. Es por eso que estas muestras desempeñan un papel fundamental en el diagnóstico de diferentes patologías, permitiendo la observación detallada y la identificación de características morfológicas cruciales asociadas a la neoplasia bajo estudio mediante un microscopio óptico. Con esta técnica es posible llevar a cabo el análisis microscópico de tejidos biológicos, que cumplen una tarea muy importante en la comprensión de los procesos fisiológicos y patológicos a nivel celular. Sin embargo, la interpretación histopatológica requiere una formación avanzada por parte del patólogo y conlleva una carga laboral significativa durante la exploración e interpretación del contenido visual de este tipo de imágenes, tanto desde una perspectiva estructural como funcional [25].

Por otro lado, la elevada producción de este tipo de muestras hace que los centros de salud requieran almacenar físicamente las láminas portaobjetos, sin aprovechar plenamente su contenido visual para mejorar los procesos diagnósticos clínicos. Además, el acceso a este tipo de información es un desafío mediante sistemas eficientes y precisos. Esto se debe a la variabilidad de patrones visuales contenidos en las muestras histológicas compuestos por características de diferente naturaleza, como el color, textura y bordes [5].

2.2. La Microscopía Virtual

En la actualidad, la microscopía virtual es una técnica de innovación en el campo de la patología, cuyo objetivo es la visualización y el despliegue de diferentes regiones de una placa histopatológica digitalizada en un computador o dispositivo de visualización. Al capturar imágenes detalladas de alta resolución de toda la placa histológica, esta tecnología brinda una representación interactiva que emula la experiencia de examinar las muestras con un microscopio convencional sobre una pantalla. Además, la microscopía virtual, actualmente es parte del flujo de trabajo en los laboratorios clínicos, posicionándose como una herramienta fundamental en educación, entrenamiento de nuevos patólogos, investigación y diagnóstico. A partir de estos desarrollos tecnológicos ha surgido la necesidad de utilizar métodos automáticos para los análisis histopatológicos de las imágenes [24]. En terminos generales, el proceso de microscopía virtual, inicia con la digitalización de las muestras de tejido, un procedimiento requiere el empleo de un microscopio equipado con una cámara de video digital, la cual escanea muestras completas. Posteriormente, estas muestras son sometidas a diversos procesos utilizando diferentes técnicas histológicas. La información visual resultante es entonces transmitida a una computadora con la capacidad de registrar la imagen obtenida en tiempo real [28], pero lo más importante esta relacionado al nivel de resolución de las imágenes obtenidas, que muchas veces llegan a requerir hasta Gigabytes. Es por eso que, esta área de investigación requiere constantemente de métodos eficaces para el almacenamiento, exploración y análisis de las muestras biológicas.

El proceso de microscopia virtual es ilustrado en Figura 2-1. Primero se prepara una muestra de tejido que se tiñe y se coloca en un portaobjetos de vidrio. Luego se utiliza un microscopio digital para tomar las capturas de diferentes regiones de la muestra. Seguidamente, las imágenes se transfieren a una computadora donde se procesan para mejorar su calidad y contraste.

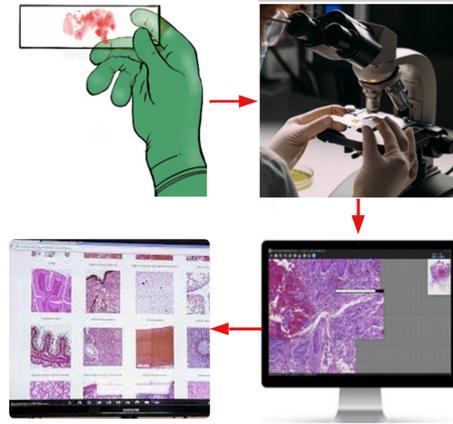


Figura 2-1: Ilustración del proceso de Microscopia Virtual, digitalización, almacenamiento y visualización de las imágenes

2.3. (WSI) Whole Slide Imaging

Whole Slide Imaging (WSI) es una tecnología que consiste en obtener imágenes digitales mediante el escaneo completo de porta objetos con la muestra histopatológica. Se han reportado avances significativos en el hardware y software para este propósito que han permitido el acceso rápido a este tipo de imágenes [42]. Esta tecnología ha pasado por una transición desde el uso de cámaras para la captura de imágenes estáticas hacia la obtención de imágenes completas digitalizadas por partes de las muestras histológicas contenidas sobre la placa portaobjetos. Las imágenes WSI no solo simplifica el almacenamiento y la gestión de grandes volúmenes de datos histológicos, sino que también abre nuevas posibilidades para la colaboración remota como: análisis computacional e interpretación histopatológico. Este avance tecnológico ha transformado la práctica en el campo de la anatomía patológica y ha impulsado la integración de herramientas digitales para una interpretación más precisa y eficiente de las muestras histológicas [14].

La obtención de imágenes de diapositivas completas (WSI) es un proceso avanzado que implica la captura de imágenes de alta resolución mediante un escaneo por partes sobre la totalidad del tejido muestreado, para lo cual, las imágenes son digitalizadas a diferente resolución. Estas imágenes son capturadas mediante cuadros pequeños de campos visuales, los que son conocidos como parches (Tiles), y ensamblados mediante estrategias conocidas como stitching (coser) para conformar los niveles de resolución, dando lugar a un esquema piramidal. Esto no solo permite la exploración detallada a nivel microscópico de la muestra, sino que también permite realizar acercamientos con diferente nivel de detalle visual sobre la imagen, generando imágenes de tamaño en Gigabytes. De esta manera, se logra una representación digital precisa y de gran calidad de imagen de la muestra biológica [3], tal como se

ilustra en la Figura 2-2 y su estructura piramidal capturando diferentes regiones en varios niveles de acercamiento. Por lo tanto, estas imágenes requieren formatos especiales para su almacenamiento, lo que obliga a que los visores sean de igual manera específicos para este tipo de modalidad.

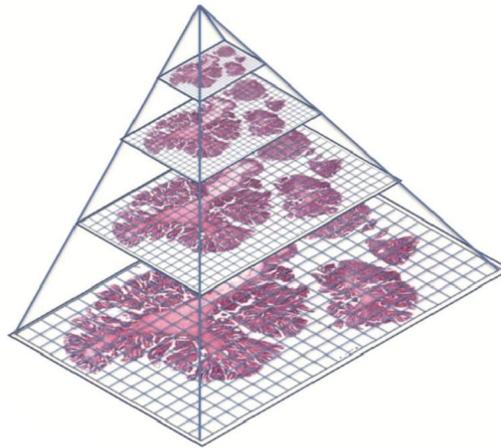


Figura 2-2: Representación del esquema piramidal de las imágenes WSI, tomado de [42].

Por otro lado, la aprobación reciente por parte de la FDA (Food and Drug Administration) de la tecnología de imágenes (WSI) destinado al diagnóstico de patología quirúrgica primaria presenta un hito significativo que ha propiciado una mayor aceptación y aplicación de esta tecnología en la práctica clínica convencional. Los avances tecnológicos continuos en escáneres digitales, métodos de visualización de imágenes y la integración de algoritmos derivados de la inteligencia artificial, con estos sistemas brindan oportunidades para mejorar los procesos diagnósticos con esta modalidad de imágenes. Entre los innumerables beneficios se incluyen: la facilidad de acceso a través de Internet, la reducción de espacio de almacenamiento físico y la ausencia de riesgo de deterioro de la calidad de la tinción o rotura de los portaobjetos. Por lo tanto, estos desarrollos tecnológicos actuales en el ámbito de imágenes (WSI) no solo tiene un impacto directo en la investigación y procesos diagnósticos sino que surge una nueva era de aplicación de modelos de Inteligencia Artificial para la interpretación y exploración del contenido visual en imágenes histológicas [22].

2.4. Sistemas de Recuperación basado en contenido (CBIR)

Un sistema CBIR (Content Based Image Retrieval) se define como, un sistema de recuperación de imágenes basado en contenido visual que buscan imágenes similares de acuerdo

a sus características visuales y son muchas las aplicaciones de sus usos [36]. Estos sistemas buscan imágenes similares comparando sus características de bajo nivel visuales mediante medidas de similitud que son implementadas de acuerdo a la modalidad de imágenes bajo estudio. Una aplicación importante dentro del área médica es el soporte a la interpretación y diagnóstico de imágenes médicas debido a que el especialista usa la información de los casos recuperados de grandes colecciones de imágenes para tomar decisiones. Durante los últimos años, se han desarrollado esquemas CBIR para soportar el diagnóstico de la presencia de cancer en imágenes de diferente modalidad [26], con diferentes estrategias para la caracterización de lesiones, como para la descripción y representación de las mismas. Actualmente, en la caracterización y representación de la información visual mediante enfoques con aprendizaje de máquina (deep learning) superan los modelos tradicionales, sin embargo aún es un tema de debate, al igual que las métricas de similitud usadas para comparar la información visual, lo que ha sido evaluado en trabajos previos, reportando una precisión y exactitud con tasas de 86,5 % y 86,8 % respectivamente [38].

Por otro lado, las medidas de similitud también definen el desempeño de los sistemas CBIR. Existen trabajos que se enfocan en el desarrollo de medidas óptimas para este propósito [37]. Entre estos métodos, algunos determinan la similitud entre imágenes histológicas mediante la teoría de grafos, la misma que es aplicada a regiones con mayor relevancia sobre la imagen. Esta técnica facilita la recuperación automática de imágenes estructuralmente similares respecto a una imagen dada como búsqueda (query) hacia una base de datos con grandes volúmenes de imágenes. Esta estrategia fue evaluada con imágenes de biopsia de mama. Además, se evaluó frente a una representación de imágenes mediante histogramas, para lo cual, los resultados obtenidos demostraron la eficacia y superioridad con una tasa del 80 % de mejora [37].

El desafío de encontrar imágenes similares se deriva de investigaciones realizadas en el área de visión artificial, que se basan en el análisis de reconocimiento de patrones y han resultado en el desarrollo de sistemas de recuperación. En el estudio mencionado [5], el sistema propuesto ha permitido a los usuarios buscar imágenes mediante una imagen objetivo o navegando por categorías de imágenes similares. Para evaluar la efectividad del sistema, se utilizaron medidas de desempeño como precisión, recall y rank, demostrando que la combinación de características puede superar una precisión de búsqueda del 70 %.

El esquema general del flujo de trabajo de un sistema CBIR es ilustrado en la Figura 2-3. Este proceso comienza por la selección manual de una imagen que sirve como consulta (query), la misma que es procesada mediante una etapa de extracción de características visuales mediante algoritmos específicos. Frecuentemente, estas características son determinadas por el color, textura y bordes. Posteriormente, el sistema realiza una búsqueda en la base de datos, estas imágenes fueron previamente caracterizadas y almacenadas. Entonces la búsqueda se

realiza mediante una medida de similitud que compara las características similares. Finalmente, se obtienen resultados recuperados de imágenes similares.

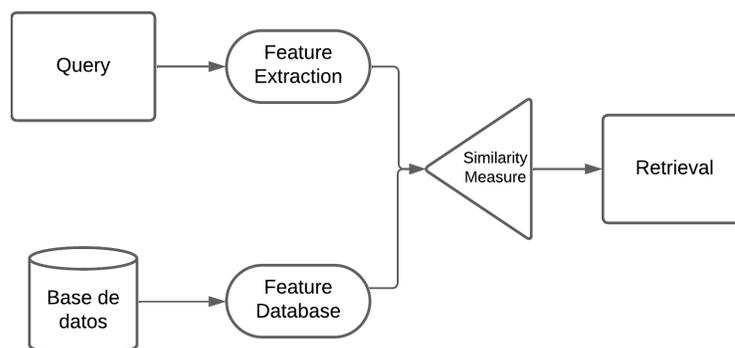


Figura 2-3: Ilustración de un esquema general del Sistema de Recuperación de imágenes por contenido visual [18]

Por otro lado, la integración de sistemas CBIR (Content-Based Image Retrieval), han generado un efecto considerable en la gestión eficiente de bases de datos, permitiendo a los usuarios recuperar imágenes según su similitud visual con una imagen de consulta. En este contexto, esta estrategia mejora el acceso y búsquedas sobre bases de datos con imágenes de patología digital, donde los descriptores basados en texto a menudo resultan insuficientes para describir con precisión el contenido visual complejo de las imágenes. Como resultado, los sistemas CBIR superan las limitaciones de los métodos basados en texto al centrarse en características visuales. Esto ofrece una herramienta de soporte a la toma de decisiones mediante la recuperación de imágenes relevantes y exploración visual de casos similares, contribuyendo a una comprensión más integral de las imágenes de entornos médicos [38].

Actualmente, la investigación en el área de recuperación de imágenes por contenido se orienta al estudio de métodos y técnicas que facilitan el acceso mediante la información visual [5]. Este acceso no se restringe únicamente a un nivel conceptual por el dominio de aplicación, sino que también se amplía a un nivel perceptual mediante la aplicación de medidas objetivas para evaluar el contenido visual. Se han desarrollado diversas implementaciones de sistemas para la recuperación por contenido CBIR, sin embargo, el desempeño de estos sistemas es variable por la cantidad de imágenes procesadas bajo este esquema y por la variabilidad de las estructuras que se visualizan con estas imágenes, por lo que aún sigue en evaluación.

2.4.1. Arquitectura de los Sistemas CBIR

Los sistemas CBIR siguen una estructura general que incluye cuatro módulos principales: interfaz de usuario, análisis de imágenes, almacenamiento y comparación [4]. La Figura 2-4

ilustra la arquitectura de los sistemas CBIR a través de diagramas de flujo, representando un proceso típico de la siguiente manera:

- **Interfaz de Usuario:** Permite el proceso de consultas y la visualización de los resultados. En ciertos casos, se puede incluir con nuevas imágenes o datos. Las consultas pueden llevarse a cabo mediante el uso de imágenes de referencia, gráficos, esquemas o dibujos en un editor de imágenes.
- **Procesamiento y análisis de Imágenes:** La tarea del análisis de imágenes consiste en procesar y extraer características utilizando técnicas de visión. Este proceso se aplica tanto a las imágenes almacenadas en la base de datos como a las imágenes de consulta.
- **Almacenamiento:** Gestionan la estructura de almacenamiento de información relacionada con las imágenes. Esto incluye la gestión de vectores de características utilizando técnicas eficientes de indexación.
- **Comparación:** El proceso de comparación se lleva a cabo mediante medidas de similitud entre los vectores de características. Se evalúan los vectores de características para determinar las imágenes que son más cercanas o similares a la consulta realizada.

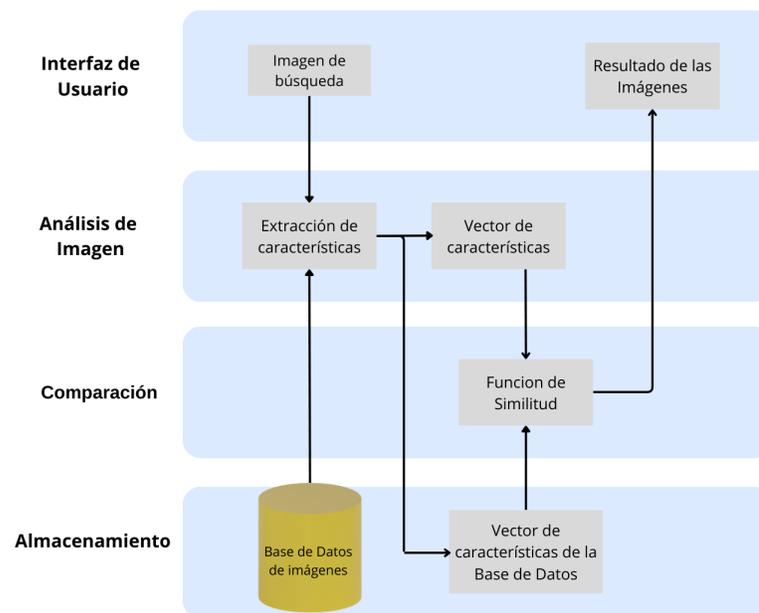


Figura 2-4: Módulos de Aquitectura CBIR [4].

2.5. Aprendizaje de Máquina

El aprendizaje de máquina ha experimentado un rápido crecimiento en las últimas décadas. Se busca imitar con las máquinas ciertas actividades que generalmente requiere de la toma de decisiones [35]. Este campo abarca diversas disciplinas, como el aprendizaje automático, la visión por computadora, el procesamiento del lenguaje natural y la toma de decisiones diagnósticas.

En los últimos años se han popularizado conceptos como Machine Learning y Deep Learning dentro del área de la salud [17]. Aunque es conocida desde hace décadas, no ha sido hasta hace pocos años que se ha incrementado su uso en la toma de decisiones con fines diagnósticos, para lo cual, se vienen desarrollando y aplicando librerías de aprendizaje automático muy conocidas como: TensorFlow ¹ y PyTorch ² respectivamente. El objetivo de estas bibliotecas es encontrar y representar patrones y correlaciones contenidas en información abstracta sobre bases de datos de partida, emulando así el aprendizaje y razonamiento humanos [39].

El aprendizaje de máquina ha demostrado gran capacidad y desempeño para tareas relacionadas con el procesamiento y análisis de imágenes y señales. En el contexto de esta investigación nos enfocaremos específicamente en el ámbito de procesamiento de imágenes médicas, donde proyectos existentes reportan los resultados de su uso para prever el riesgo de enfermedades y definir tratamientos más efectivos para los pacientes [39].

Especialmente en la radiología y patología, el aprendizaje de máquina ha mostrado su utilidad en el desarrollo de sistemas que ayudan al diagnóstico y toma de decisiones mediante algoritmos que aprenden y detecten patrones discriminatorios en las estructuras anatómicas contenidas en las imágenes. En el campo de la patología, la integración de la microscopía digital junto a modelos de aprendizaje de máquina configuran herramientas de soporte al diagnóstico [23]. Un estudio ha evidenciado la implementación de técnicas de aprendizaje profundo para mejorar el diagnóstico de tumores a través del análisis de imágenes histopatológicas y radiológicas, mediante el uso de redes neuronales convolucionales y otras técnicas avanzadas [17]. Estos desarrollos buscan contribuir con las limitaciones de exámenes subjetivos de los patólogos, aspirando perfeccionar la precisión del diagnóstico en diversos tipos de cáncer [17]. Al aprovechar algoritmos de inteligencia artificial estos métodos pueden identificar patrones y características sutiles en las imágenes médicas, lo que conduce a una detección temprana de tumores y una evaluación más precisa de su agresividad y extensión.

Por otro lado, en el campo de la cirugía vascular se ha aplicado técnicas de inteligencia artificial con el propósito de optimizar tanto el diagnóstico como el tratamiento de enfermedades

¹<https://www.tensorflow.org/>

²<https://pytorch.org/>

vasculares. El cual se basa en la automatización de la interpretación de imágenes médicas, la anticipación del crecimiento de lesiones y la planificación de procedimientos endovasculares, todo ello con el fin de elevar la precisión y eficiencia en el manejo de estas afecciones [2]. La aplicación de algoritmos de aprendizaje automático en este contexto puede ayudar a identificar signos tempranos de enfermedad vascular, permitiendo una intervención más oportuna y personalizada, lo que puede mejorar los resultados clínicos y reducir los riesgos asociados con las complicaciones vasculares.

Además, se han llevado a cabo investigaciones que exploran la aplicación de técnicas de aprendizaje automático para mejorar el proceso de diagnóstico en patología quirúrgica. Estas iniciativas buscan automatizar los procedimientos diagnósticos, reduciendo la carga laboral del personal médico y mejorando la precisión en la interpretación de imágenes médicas para diagnosticar enfermedades con mayor exactitud [19]. Al aprovechar grandes conjuntos de datos y algoritmos de aprendizaje automático, estos métodos pueden identificar patrones complejos en las imágenes médicas que pueden pasar desapercibidos para el ojo humano, lo que puede mejorar la precisión del diagnóstico y reducir la incidencia de errores médicos.

Por lo tanto, la aplicación de redes neuronales en estos campos demuestra la importancia y el potencial de estas tecnologías en la medicina moderna. Estas técnicas comparten un objetivo común, el análisis automatizado preciso de grandes volúmenes de datos médicos, como imágenes biomédicas y radiológicas de diferentes modalidades para la detección temprana y precisa de enfermedades, así como una herramienta de toma de decisiones para planificación y tratamiento.

2.6. Modelos de Inteligencia Artificial que funcionan en los CBIR

Redes Neuronales Convolucionales

Las redes neuronales convolucionales han sido utilizadas para la representación y caracterización del contenido visual de las imágenes en escenarios de reconocimiento de patrones y clasificación de imágenes médicas [7]. Se caracterizan por: en lugar de establecer conexiones entre todas las neuronas de capas consecutivas, las redes convolucionales presentan conexiones más específicas y limitadas entre neuronas. Además, se señala que las capas de estas redes pueden tener dos o más dimensiones, siendo típicamente multidimensionales para datos de imágenes y tridimensionales para la representación de objetos tridimensionales [10]. La Figura 2-5 ilustra una estructura básica de la red neuronal convolucional, conformada por dos capas convolucionales intercaladas con submuestreo (también conocido como capas de agrupamiento o pooling), seguidas por una capa de salida. Cada capa convolucional aplica filtros para extraer características de la entrada, mientras que las capas de submuestreo re-

ducen la dimensionalidad de las características extraídas. La capa de salida produce la salida final de la red.

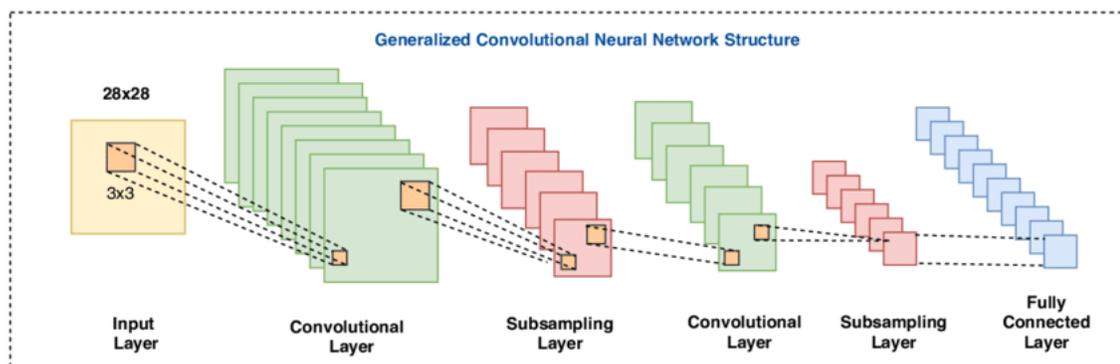


Figura 2-5: Estructura general de las redes neuronales Convolucionales, tomado de [27]

Los métodos utilizados para la obtención de descriptores globales en estudios previos [36], son adquiridos mediante el uso de redes neuronales convolucionales (CNN). Este método simplifica la extracción de vectores de características, que son necesarias para la evaluación de similitud entre imágenes durante el proceso de recuperación. Se utilizan métricas como la distancia de Sorensen y la distancia del coseno para calcular la similitud entre los vectores de características, permitiendo una recuperación más efectiva de imágenes similares en grandes conjuntos de datos.

Incluso se ha adaptado otros métodos para los sistemas CBIR en el diagnóstico de reconocimiento de cáncer de piel [7]. Con la implementación de una serie de modelos de aprendizaje profundo, que se centran a las redes neuronales convolucionales, las cuales aplican las arquitecturas MobileNetV3 y DenseNet121 [20][29] para la clasificación de imágenes de dermatoscopia, lo cual puede ser una ayuda a resolver problemas de desbalanceo de los conjuntos de datos. Además de poder mejorar estas técnicas adoptando varias redes a la vez, con el objetivo de hallar una estimación más acertada para el sistema.

Redes Generativas Adversarias (GANs)

Las Redes Generativas Adversarias consisten en dos modelos, generalmente que son entrenados de forma simultánea. En este proceso el generador, cuya función es producir datos, y el discriminador que se encarga de distinguir entre los datos generados y los ejemplos reales que evolucionan de manera conjunta. Al principio del proceso, la salida del generador es aleatoria, y el discriminador predice de manera arbitraria. Con el tiempo y el entrenamiento, ambos modelos mejoran: el generador para engañar al discriminador, y este último para discernir entre datos verdaderos y falsos generados por el primero [6].

Por otro lado, en el análisis automatizado de retinografías usada para el diagnóstico de enfermedades oculares [15], se emplean redes neuronales GANs junto con un modelo pre-entrenado VGG19 [9]. Implementado para mejorar la resolución de las imágenes del fondo de ojo. El propósito de esta red es aprender un mapeo de características que se utiliza para transferir las características texturales de una imagen. La evaluación de los resultados se lleva a cabo mediante medidas estadísticas y métricas de similitud, utilizando la distancia euclidiana para determinar el desempeño del modelo, minimizando el error de precisión y reducción en el número de imágenes no graduables en el sistema de recuperación.

Por ejemplo, este tipo red neuronal han contribuido significativamente a la preparación digital de muestras histológicas, han logrado generar imágenes microscópicas digitales que imitan las tinciones de hematoxilina-eosina, sin la necesidad de añadir colorantes, mediante una adaptación holográfica de microscopía cuantitativa de contrastes de fases [33]. La inclusión de estas aplicaciones en el trabajo del patólogo podría reducir la probabilidad de error humano en un 85 % [33].

Redes Siamesas

La red siamesa replica la estructura de una red neuronal convolucional y luego combina las salidas en capas compartidas. Este método simplifica la comparación de dos entradas, relacionando similitudes de imágenes de personas, extrayendo características de ambas y utilizando un clasificador para evaluar si la persona ha sido verificada correctamente. Esto produce un resultado claro, ya sea afirmativo o negativo fácil de interpretar [1].

Por otro lado, la implementación de redes neuronales convolucionales (CNN) siamesas ha contribuido en otros estudios [11], empleando un modelo de entrenamiento AlexNet. Esta red tiene una característica distintiva, por lo que comparten pesos y arquitectura en dos o más ramas. A diferencia de las redes convencionales, las CNN siamesas no se enfocan en clasificar o predecir clases específicas, sino que aprenden a medir la similitud o distancia entre pares de datos. Esta capacidad las hace especialmente útiles en aplicaciones de comparación y emparejamiento, como reconocimiento facial o búsqueda de imágenes similares.

En nuestro sistema propuesto, planeamos emplear la arquitectura de red neuronal convolucional conocida como SqueezeNet [16]. Esta elección se debe a las características específicas del modelo, diseñada para ser más liviana y eficiente en comparación con otras redes convolucionales. La estructura más delgada de SqueezeNet [16] se logra mediante el uso de filtros de convolución de tamaño reducido y técnicas de compresión, lo que resulta en una reducción significativa del número de parámetros y de la carga computacional requerida para su funcionamiento.

2.7. Medidas de Similitud

Las medidas de similitud determinan la semejanza entre descriptores visuales de bajo nivel (vector de características) utilizando funciones de distancia o métricas, que son aplicadas a un espacio N-dimensional para calcular un valor numérico de las distancias cercanas y lejanas entre dos descriptores visuales en el espacio de características. En otras palabras, se relaciona este concepto de diferencias y similitudes [41]. Entre las medidas de similitud convencionales más usadas se encuentran:

- **Distancia Euclídeana**

La definición euclidiana de distancia se refiere a la medida de la distancia entre un par de descriptores visuales (vectores) x e y , que se expresa de la siguiente manera:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

- **Distancia Manhattan**

La distancia Manhattan es también conocida como Taxicab, se considera más eficiente en términos computacionales que la distancia euclidiana, ya que no requiere el cálculo de cuadrados ni de raíces cuadradas. Esto la convierte en una medida mucho más rápida de calcular.

$$d(x, y) = \sum_{i=1}^n |x_i - y_i|$$

- **Distancia Coseno**

Esta distancia mide el coseno del ángulo que forman los dos vectores. La definición de esta métrica es de la siguiente forma:

$$d(\bar{x}, \bar{y}) = 1 - \cos(\theta) = 1 - \frac{x \cdot y^T}{\sqrt{(xx^T)(yy^T)}}$$

Aunque el cálculo de esta métrica no es muy eficiente, proporciona información sobre la dirección relativa de los dos vectores de entrada. Por esta razón, a pesar de no cumplir con la condición del triángulo y no ser una métrica completa, la distancia Coseno es muy popular.

■ Distancia Hamming

La distancia de Hamming tuvo el primer rendimiento en la teoría de la información para detectar cuantitativamente los errores de corriente, recibiendo varios bits durante la transmisión. Se calcula teniendo en cuenta los cambios de bits entre el origen y el destino de la comunicación. La ecuación se expresa de la siguiente manera:

$$d_H(A, B) = \sum_{i=1}^n \delta(a_i, b_i)$$

donde $A = a_1a_2\dots a_n$ y $B = b_1b_2\dots b_n$, y $\delta(x, y)$ está definido como:

$$\delta(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{si } x \neq y \end{cases}$$

La medida de similitud utilizada en el estudio propuesto, se determinó mediante la distancia euclidiana, la cual es una métrica de precisión ampliamente aceptada para calcular la distancia directa entre dos puntos en un espacio. Esta opción resulta conveniente para comparar características numéricas extraídas de bases de datos de imágenes histopatológicas, tales como valores de píxeles o características de textura y color.

Es importante evaluar los beneficios como las limitaciones del sistema CBIR que se propone. Entre las ventajas del CBIR se encuentra su capacidad para buscar imágenes según su contenido visual, lo que facilita la identificación de patrones y características específicas en grandes conjuntos de datos de imágenes médicas [21]. Esta capacidad puede agilizar el proceso de diagnóstico, permitiendo a los médicos acceder rápidamente a imágenes relevantes y compararlas con casos similares. Sin embargo, algunas desventajas pueden surgir de la dependencia de la calidad y precisión de las características extraídas automáticamente, así como de la necesidad de una gran cantidad de datos de entrenamiento para desarrollar modelos efectivos. A pesar de estas limitaciones, se ha decidido implementar el sistema CBIR debido a su potencial para mejorar la precisión del diagnóstico, la eficiencia en la búsqueda de imágenes y la calidad general de la atención al paciente.

3 Capítulo 2

Metodología

En este capítulo se describen las estrategias implementadas para el sistema propuesto, el mismo que consiste en las siguientes fases de desarrollo, como se ilustra en la Figura 3-1. El flujo de trabajo comienza con la selección manual de una imagen de búsqueda, que previamente ha sido normalizada y procesada para su caracterización y comparación con las imágenes almacenadas en la base de datos. Siguiendo, se emplea un modelo de aprendizaje profundo (deep learning) para extraer características, generando descriptores visuales que se utilizan para calcular las medidas de similitud mediante un algoritmo K-NN (k-nearest neighbors) K-vecinos cercanos. Por último, el sistema despliega un conjunto de imágenes más similares, como se ilustra en la Figura 3-1.

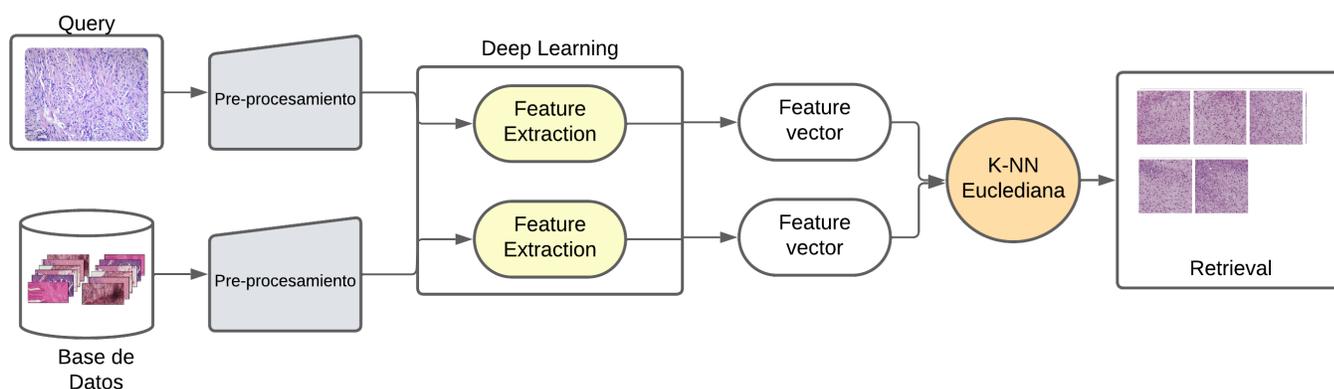


Figura 3-1: Módulos de Arquitectura CBIR [4].

3.1. Construcción de la base de datos con imágenes WSI

Para la construcción de la base de datos usada en este trabajo, se integraron dos conjuntos de imágenes histológicas WSI. El primer conjunto corresponde a un grupo de imágenes que fueron digitalizadas en la Universidad Politécnica Salesiana, específicamente de laboratorio de histología de la carrera de biomedicina. Para el proceso de digitalización se utilizó un escáner comercial de la marca OptraSCAN que ofrece una resolución óptica de X20 a X40

de aumento para el muestreo de diapositivas completas, procesando alrededor de 120 muestras de tejido que fueron digitalizadas y almacenadas en formato de compresión JPEG 2000 (JP2) para el almacenamiento de grandes volúmenes de información o metadatos, eficiente para guardar contenido de la imagen sin pérdidas, con mayores profundidades de color y resolución.

El proceso de digitalización consiste en capturar los campos de preparación de la muestra original para luego ir uniendo esos fragmentos en un solo fichero [34], el sistema de escaneo tiene la capacidad de digitalizar múltiples escalas para obtener una imagen perfectamente enfocada como se ilustra la Figura 3-2.

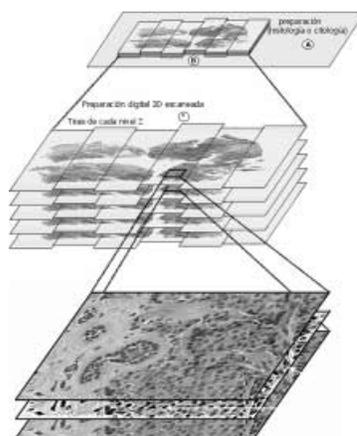


Figura 3-2: Digitalización de múltiples escalas de una preparación histológica [34].

Por lo tanto, este conjunto de imágenes consiste en 94 Whole Slide Imaging, con una resolución de 72 píxeles por pulgada (ppp) con un formato JP2000. Durante el proceso de digitalización se aseguró la buena calidad de las imágenes y la ausencia de artefactos dentro de la placa porta objetos. Finalmente, se consideró la inclusión de un total de 15 clases de tejidos para su análisis. Además, se optó por transformar el formato de las imágenes JPEG 2000 a DZI (Deep Zoom Images) creado por Microsoft®), ya que es un formato compatible con diferentes librerías de código abierto para la visualización estandar de este tipo de imágenes. Por ejemplo: la librería Libvips¹, usada en este trabajo.

El contenido de la Figura 3-3 ilustra la selección de las muestras más destacadas del repositorio universitario obtenido.

¹<https://www.libvips.org/>



Figura 3-3: Ilustración de la base de datos WSI

El segundo grupo de imágenes fueron seleccionadas del repositorio de dominio público perteneciente a BACH (Breast Cancer Histology)[3]. Esta colección de imágenes WSI se caracteriza por tener etiquetas asignadas a diferentes categorías de tipos de tejidos, proporcionando información detallada sobre el contenido de secciones específicas en el escaneo completo del tejido. Las imágenes se encuentran almacenadas a una resolución de 96 píxeles por pulgada (ppp) en formato TIFF (Tagged Image File Format), reconocido por su capacidad para preservar la calidad y la integridad de los datos de imagen. Este formato permite capturar con detalle la complejidad del tejido muestreado, presentando dimensiones de 2048x1536 píxeles y una escala de píxeles de $0,42, \mu\text{m} \times 0,42, \mu\text{m}$. Dichas imágenes corresponden a muestras histológicas de tejido mamario obtenidas de pacientes femeninas en el Hospital CUF Porto, Portugal, durante los años 2014, 2015 y 2017, utilizando un microscopio LED Leica DM 2000. La base de datos consta de 400 imágenes que abarcan tejidos sanos y cancerosos, específicamente, lesiones benignas, carcinoma in situ y carcinoma invasivo. Una característica destacada de este repositorio es la distribución de regiones extraídas de las imágenes completas, organizadas en cuatro carpetas que representan las clases mencionadas anteriormente.

Para la selección de las regiones de interés y su clasificación en carpetas, se segmentaron de manera automática las regiones, de acuerdo a trabajos previos reportados para la construcción de este tipo de repositorio [3]. Una vez segmentadas las regiones, fueron anotadas para identificar la correspondiente clase de patología, este proceso se ilustra en la Figura 3-4, en donde se presenta visualmente estas anotaciones, realizadas con una configuración de menor magnificación, donde la lesión benigna se representa en rojo, el carcinoma in situ en verde y el carcinoma invasivo en azul, proporcionando una perspectiva completa de la área patológica identificada.

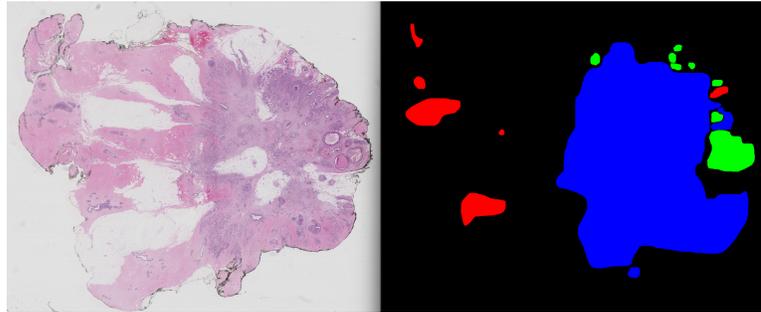


Figura 3-4: Ejemplo de una imagen de diapositiva completa anotada con diferente color de acuerdo a su segmentación [3]

3.2. Normalización de la base de datos

Una vez que las imágenes WSI fueron recolectadas y almacenadas de las diferentes bases de datos, un proceso de normalización del repositorio es llevado a cabo, esto con el propósito de homogenizar el tamaño de las imágenes para su procesamiento. Particularmente, el tamaño de las WSI es muy grande (en gigapíxeles) para la aplicación de técnicas de procesamiento de imágenes reportadas en la literatura. Por lo tanto, en este trabajo se aplicó un método para el recorte de las imágenes completas en regiones parciales (parches) y que se almacenen en el mismo formato para su procesamiento. Para tal efecto, se propone un método de recorte de las imágenes el mismo que consiste en la extracción de parches más pequeños de (512x512 píxeles) de acuerdo a su nivel de resolución. Este proceso se lleva a cabo con el paquete especial para el manejo de imágenes conocido como Libvips ², que es una librería de código abierto de lenguaje de programación C, permitiendo acciones como el redimensionamiento, recorte y división de los WSI en diferentes parches de acuerdo a su nivel de resolución (Zoom), tal como se ilustra en la Figura 3-5.

²<https://www.libvips.org/>

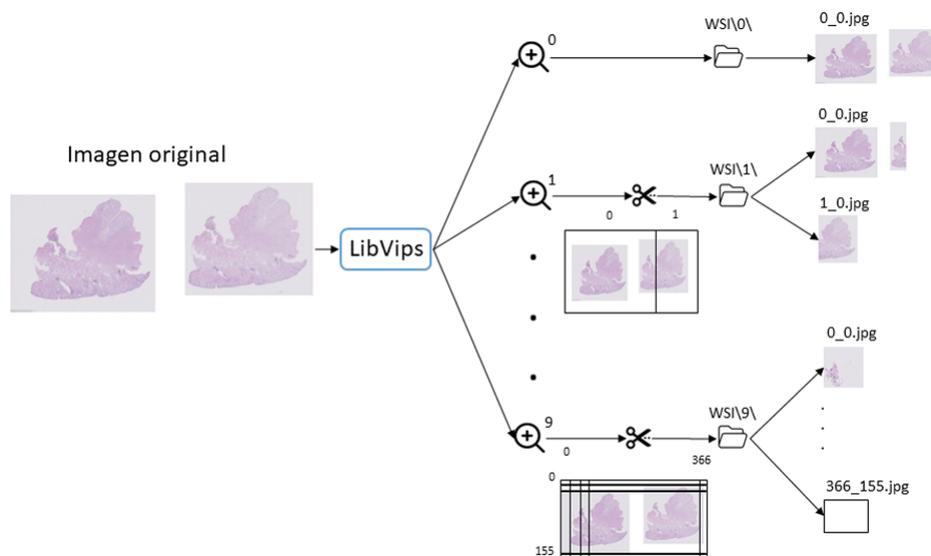


Figura 3-5: Distribución de de Ficheros generados con Libvips [31].

Finalmente, la aplicación de este proceso resulta en carpetas que contienen parches a diferente escalas (nivel de Zoom), por cada imagen [31]. La ejecución de este proceso es ilustrado en la Figura 3-6.

```

Símbolo del sistema
Microsoft Windows [Versión 10.0.22621.3007]
(c) Microsoft Corporation. Todos los derechos reservados.

C:\Users\Usuario>cd C:\Users\Usuario\Downloads\vips-dev-w64-web-8.15.0-rc1\vips-dev-8.15\bin

C:\Users\Usuario\Downloads\vips-dev-w64-web-8.15.0-rc1\vips-dev-8.15\bin>vips dzsave "C:\Users\Usuario\Desktop\trozo s\312438" 312438 --overlap 10 --tile-size 512

```

Figura 3-6: Comando de la librería Libvips

Para la ejecución del comando se consideran los siguientes argumentos: ruta de la imagen, la carpeta de destino, el solape, las dimensiones de recorte de la imagen y la calidad. Es así como se crea automáticamente un único fichero por cada imagen, con carpetas con las regiones recortadas a diferente nivel de resolución. Este esquema permite la reconstrucción completa de la imagen WSI original en esquemas de visualización para su exploración. Entonces, finalmente se obtienen 16 carpetas con ficheros anotados desde 0 a 15, donde la carpeta 0 contiene la imagen de entrada, pero con un tamaño de 512 x 512, de igual forma la carpeta 1 presenta dos divisiones de la imagen original con un tamaño de 512 x 512 pero con una característica distinta de ampliación del Zoom, realizando el mismo proceso

de forma incremental en el nivel de Zoom hasta la carpeta 15 que va a contener todos los parches producto de la división de la imagen original, tal como fue ilustrado en la Figura 3-5.

Finalmente, toda la base de datos es normalizada a regiones extraídas de las imágenes WSI del mismo tamaño y almacenadas en un único formato DZI.

3.3. Pre-procesamiento de las imágenes

Una vez que las imágenes fueron extraídas, normalizadas y almacenadas en la base de datos como regiones para su análisis, es necesario asegurar la calidad para el proceso de extracción de características, esto debido a la heterogeneidad y naturaleza de los diferentes tejidos contenidas en las regiones (parches) que presentan diferentes variaciones de color y contraste, como se ilustra en la Figura 3-7.

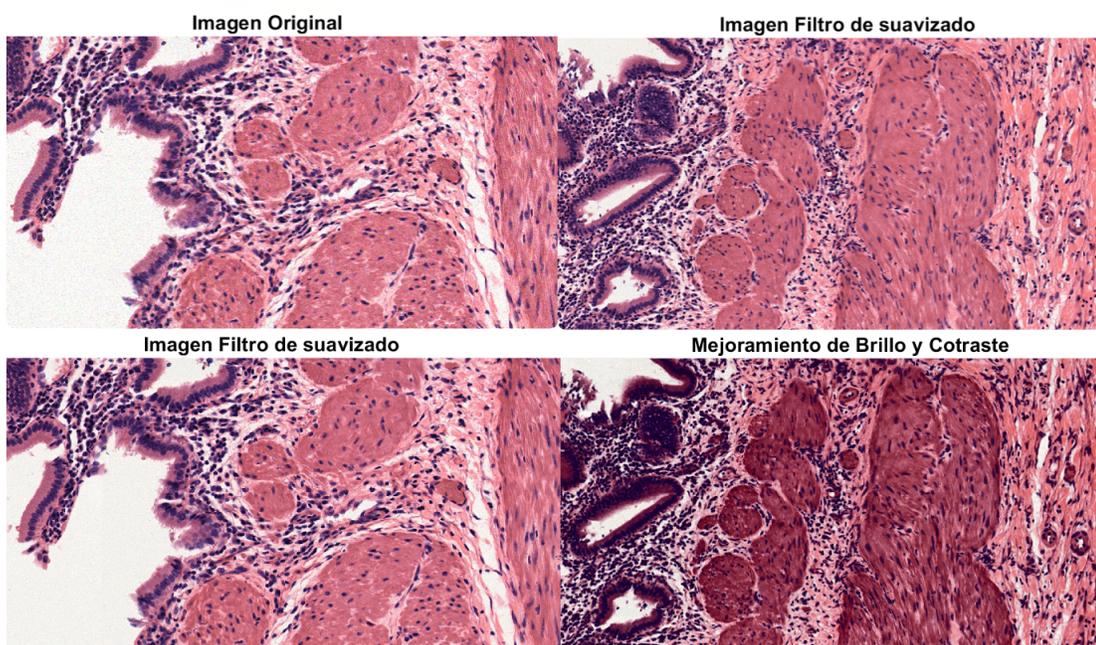


Figura 3-7: Imágenes de pre-procesamiento aplicado métodos de mejoramiento

Con el propósito de homogenizar el contraste y eliminar el ruido presentado en las imágenes, que se observa como pequeñas variaciones aleatorias en la intensidad de sus píxeles, a causa de artefactos presentes en los portaobjetos durante el proceso de digitalización, en este trabajo se aplicó un filtro de suavizado, el mismo que es definido por:

$$I_{\text{suavizado}}(x, y) = \frac{1}{N} \sum_{i=-k}^k \sum_{j=-k}^k I(x + i, y + j) \quad (3-1)$$

Donde:

- $I_{\text{suavizado}}(x, y)$: El valor suavizado de la intensidad de la imagen en las coordenadas (x, y) .
- x, y : Las coordenadas espaciales dentro de la imagen, donde x representa la posición horizontal y y la posición vertical.
- N : El tamaño del vecindario considerado para el suavizado, que indica el número total de píxeles en el vecindario alrededor del punto (x, y) que se están promediando.
- i, j : Variables de índice que indican la posición relativa dentro del vecindario alrededor del punto (x, y) . Estas variables se utilizan para iterar sobre el vecindario, con i que varía desde un valor negativo hasta uno positivo, y lo mismo para j .
- $I(x + i, y + j)$: El valor de intensidad del píxel en la imagen desplazada i unidades horizontalmente desde x y j unidades verticalmente desde y . Es decir, $I(x + i, y + j)$ es la intensidad del píxel en la posición relativa (i, j) respecto al punto (x, y) .

Este filtro es conocido como promedio, y sirve para homogenizar los detalles finos y el ruido presente en las estructuras de la imagen. El resultado de este proceso se observa en el espacio en blanco que delimita los bordes del tejido histológico, donde los detalles ahora son más discernibles y la textura es más uniforme.

Posteriormente, es necesario resaltar los colores de tñido de (H&E) de la muestra, redistribuyendo los niveles de intensidad que aumenten la diferencia entre los valores de píxeles de diferente color. Para tal efecto, es ecualizado el histograma de color mediante:

$$L_{\text{ajustado}} = \text{histeq}(L) \times \text{máxima_luminosidad} \quad (3-2)$$

Donde:

- L_{ajustado} : Representa el resultado de ajustar la luminancia de una imagen mediante la aplicación de una transformación específica, como la ecualización del histograma, seguida de un posible escalado o modificación adicional.
- $\text{histeq}(L)$: Es una función que opera sobre la luminancia original L , transformándola de alguna manera para mejorar ciertas características visuales de la imagen, como el contraste o la distribución de intensidades.
- $\text{máxima_luminosidad}$: Representa el valor máximo posible de luminancia en la imagen. Este valor puede depender del formato de imagen utilizado y del rango de valores de píxeles disponibles.

El proceso se realiza en el espacio de color LAB. Inicialmente, se ajusta el canal de luminosidad (L) para mejorar el contraste. Luego se transforma nuevamente la imagen al espacio de color RGB. Con este proceso resalta las áreas de interés haciendo que los detalles sean mas visibles y definidos, ajustando la intensidad de luz de la imagen.

3.4. Extracción de Características

Para la extracción de características se utilizó un modelo de aprendizaje profundo conocido como "SqueezeNet" [16], que es una red neuronal convolucional diseñada para aprender automáticamente características relevantes de las imágenes. Este modelo se selecciono debido a su capacidad reducir la cantidad de parámetros necesario durante el reconocimiento de patrones visuales complejos en las imágenes. La Figura 3-8 ilustra la arquitectura del modelo squeezeNet, comenzando con la capa de convolución Conv1 que utiliza 96 filtros, esta red emplea agrupamiento máximo (maxpool) para reducir la dimensionalidad de las características antes de pasar a través de una serie de bloques de fuego (Fire2 a Fire9). Estos bloques consisten en capas de compresión y expansión que utilizan convoluciones 1x1 para reducir el número de canales de características y convoluciones 3x3 para capturar características más complejas. Después de cada bloque de fuego se realiza otro agrupamiento máximo para reducir aún más la dimensionalidad. La red culmina con una capa de convolución final (Conv10) con 1000 filtros, seguida de un agrupamiento promedio global (global AvgPool) para generar una representación promedio de las características. La salida final consta de una función de activación Softmax, que proporciona las probabilidades de clasificación para las diferentes clases en el conjunto de datos.

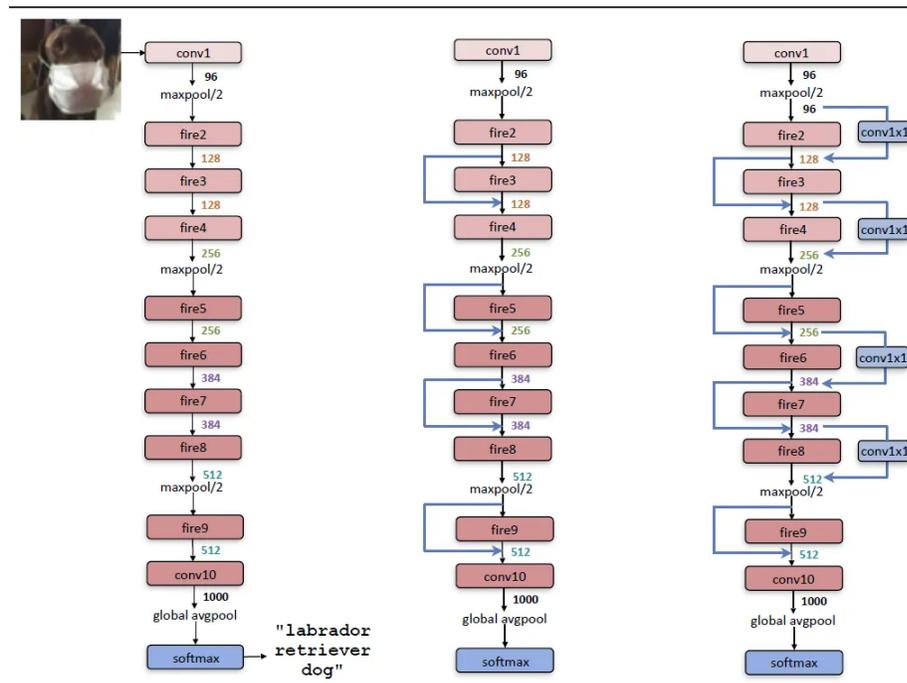


Figura 3-8: arquitectura del modelo SqueezeNet [16]

El aprendizaje profundo (DL) posibilita la formación de modelos computacionales mediante capas múltiples de procesamiento, se facilita que estos modelos adquieran conocimientos sobre las representaciones de datos. De acuerdo con [12] indica que con estos modelos muestran la capacidad de generar imágenes originales que representan una amplia variedad de elementos, como objetos, animales, paisajes y personas, utilizando simplemente indicaciones breves expresadas en lenguaje natural.

3.5. Algoritmo K-NN

El algoritmo de los vecinos más cercanos (K-NN) es una técnica en el ámbito del aprendizaje automático y el reconocimiento de patrones [13]. Se fundamenta en la premisa de que puntos de datos similares tienden a agruparse en un espacio de características. En términos prácticos, K-NN evalúa la distancia entre un punto de consulta y todos los demás puntos en un conjunto de datos. Luego, clasifica el punto de consulta según la mayoría de las clases presentes en sus vecinos más cercanos (los K puntos más próximos). La distancia euclidiana presentada por la Ecuación 3-3 es una métrica implementada en este proceso, calculando la distancia entre los descriptores visuales obtenidos de la red neuronal y llevadas a un espacio Euclidiano. Este enfoque es particularmente útil en problemas de clasificación donde las clases están bien diferenciadas en el espacio de características.

La Figura 3-9 muestra un ejemplo ilustrativo de cómo el algoritmo de los vecinos más cercanos (K-NN) clasifica una muestra de prueba en dos clases distintas, cuadrados azules y triángulos rojos. El parámetro k , que indica la cantidad de vecinos más cercanos considerados para la clasificación, se ilustra con dos casos: cuando $k = 3$ y $k = 5$. En el primer caso ($k = 3$), el punto de prueba se clasifica como un triángulo rojo debido a que hay 2 triángulos y solo 1 cuadrado dentro del círculo definido por los 3 vecinos más cercanos. En el segundo caso ($k = 5$), el punto de prueba se clasifica como un cuadrado azul, ya que hay 3 cuadrados y 2 triángulos dentro del área definida por los 5 vecinos más cercanos. Este ejemplo ilustra cómo varía la clasificación de una muestra de prueba según el valor K y cómo clasifica la clase a la que pertenece en función de la mayoría de vecinos más cercanos dentro de un área específica del espacio de características.

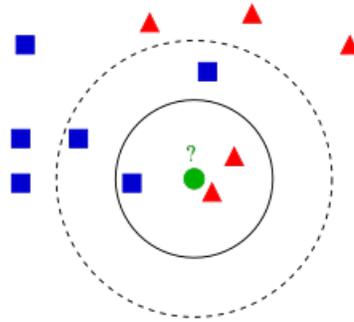


Figura 3-9: Ilustración del funcionamiento del algoritmo K-vecinos cercanos [30]

Por lo tanto, la medida de similitud determina la semejanza entre descriptores visuales (vectores de características) que son proyectados sobre el espacio N -dimensional, entregando un valor numérico de las distancias cercanas y lejanas entre las imágenes [41]. Finalmente, para este trabajo se implementó la distancia euclidiana, tal como sigue:

■ Distancia Euclidiana

La definición de distancia euclidiana se refiere a la medida de la distancia entre un par de puntos x e y , y se expresa de la siguiente manera:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3-3)$$

donde x_i es el descriptor de la imagen de prueba, y y_i es el descriptor de la imagen contenida en la base de datos y n es la dimensionalidad o cantidad de características del descriptor visual.

3.6. Implementación del sistema de recuperación basado en contenido visual.

En esta sección se describe la implementación del sistema completo propuesto para recuperación de imágenes histopatológicas basada en su contenido visual. Para tal efecto, se utilizó Orange Data Mining³. Esta es una herramienta ampliamente recomendada para este propósito, ya que permite la creación de flujos de trabajo para el análisis de minería de datos de manera intuitiva y eficiente. Por lo que ofrece una interfaz gráfica altamente intuitiva que facilita su uso. Esta herramienta incluye una variedad de modelos de aprendizaje de máquina, deep learning y muchas herramientas usadas en minería de datos [32]. Entonces, permite construir arquitecturas y modelos de aprendizaje de manera gráfica y secuencial. Este proceso se lleva a cabo mediante la conexión de sus módulos implementados de manera dinámica. Además de ser una herramienta para redes neuronales y retroalimentación de aprendizaje automático, también permite manejar bases de datos, procesar, analizar y configurar las especificaciones necesarias para ejecutar modelos de aprendizaje.

En la Figura 3-10 se presenta el diagrama general del flujo de trabajo propuesto en este trabajo, con sus correspondientes conexiones de los módulos widget utilizados en esta propuesta.

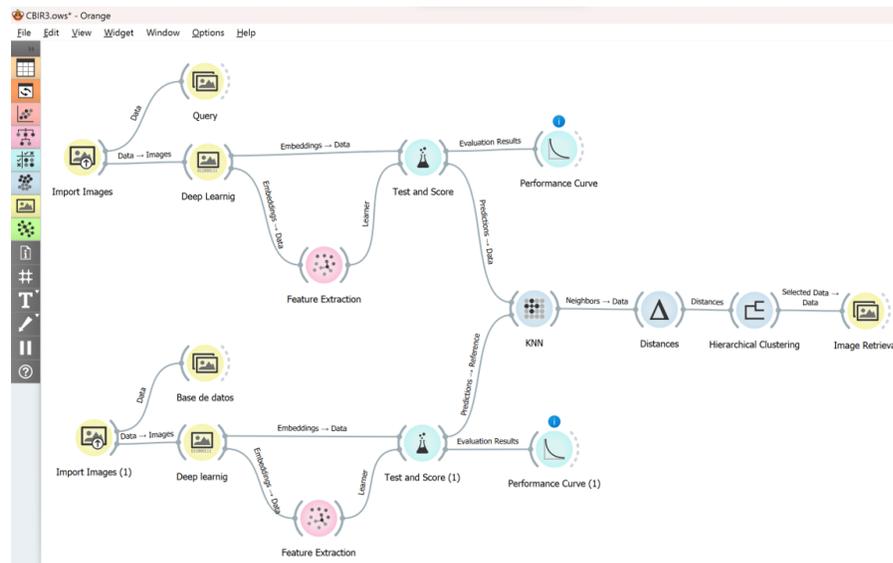


Figura 3-10: Esquema de recuperación basado en Contenido visual

La facilidad de manipular bases de datos en la minería de Orange permite la preparación y organización de los datos, ya que permite clasificar y aceptar información clasificada. El widget responsable de cargar la base de datos es una (Import Image), y fue usado para

³<https://orangedatamining.com/>

almacenar los parches VIPS y la recopilación de los dos conjuntos de datos. Para la evaluación del sistema, se ocupo 1984 imágenes de entrenamiento y 400 imágenes de pruebas, es decir, un total de 2384 imágenes histopatológicas divididas en un total de 19 categorías, como se muestra en la Figura 3-11.

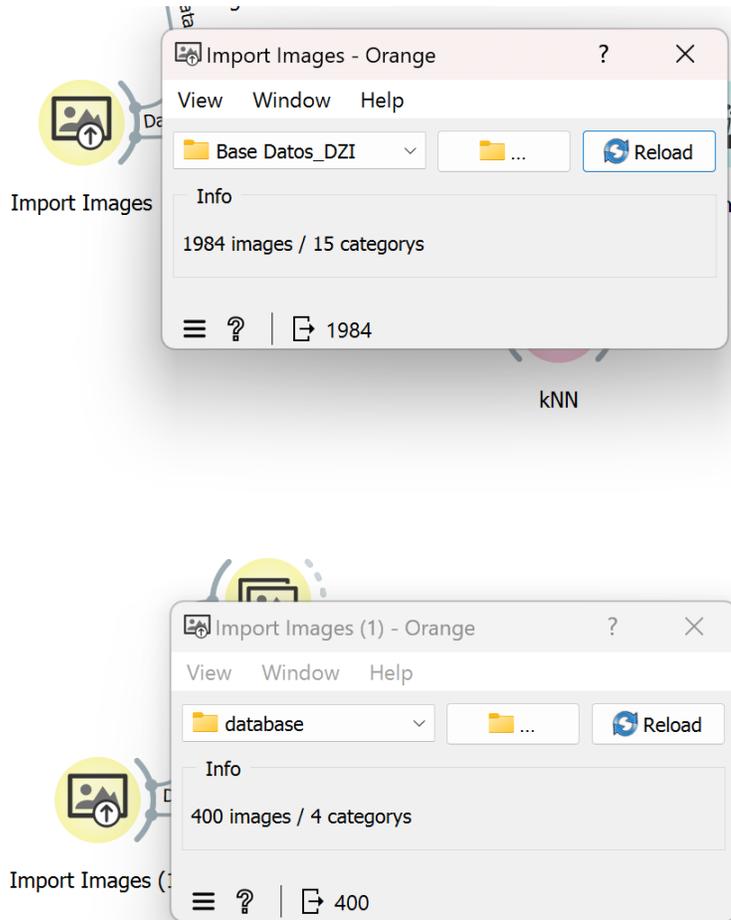


Figura 3-11: Base de Datos de Imágenes Histológicas

Para el procesamiento de las imágenes se utiliza el widget conocido como (Image Embedding) para el uso de algoritmos de aprendizaje profundo (Deep Learning), convirtiendo las imágenes cargadas en descriptores visuales (vectores de características). Además el modelo entrenó procesa un total de 1984 imágenes de muestras histológicas para la clasificación que se observan en el Figura 3-12.

La arquitectura de la red neuronal convolucionada usada en este trabajo es "SqueezeNet", seleccionada por ser un modelo ligero en términos de parámetros, haciéndola mucho más ágil en recursos computacionales. Adecuado para aplicaciones o sistemas integrados que requieren un equilibrio entre precisión y eficiencia. Útil para extraer patrones y características las mismas que son aprendidas por las capas convolucionales, mientras se tenga una gran

cantidad de imágenes hace mucho más eficiente el modelo de recuperación.

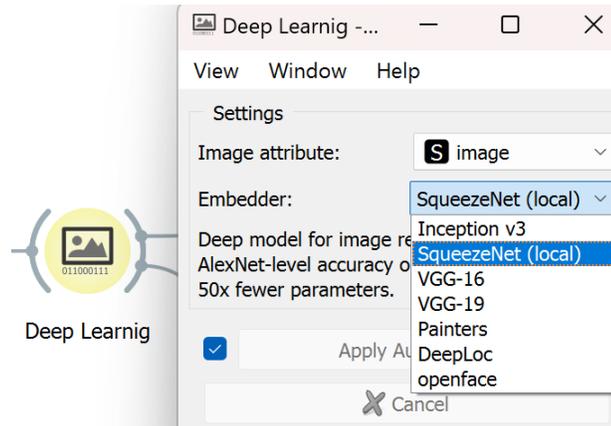


Figura 3-12: Modelo de entrenamiento SqueezeNet

Posteriormente, con el uso del modelo SqueezeNet se extraen las características de las imágenes (parches). Así los vectores que resultan de estas características son almacenados en un widget específico denominado (Feature Extraction) como de muestra en la Figura 3-13. Estos vectores encapsulan las representaciones semánticas de las imágenes, es decir, las características visuales aprendidas por el modelo, entre estas: bordes, texturas y patrones específicos, importantes para la tarea de clasificación.

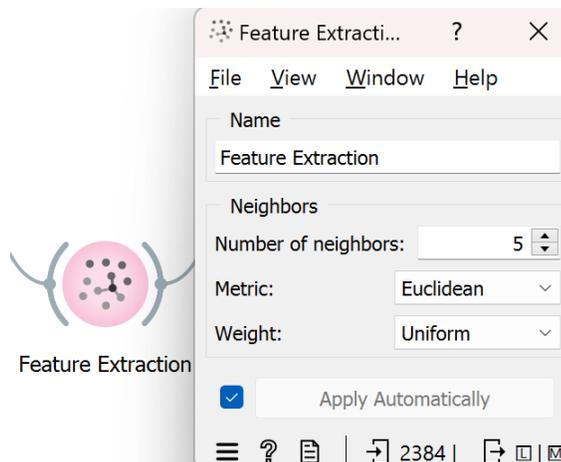


Figura 3-13: Extracción de Características

Una vez que las imágenes son caracterizadas mediante sus vectores de características, se aplica el algoritmo de los K -vecinos cercanos, conocido como "K-NN", el cual permite clasificar las imágenes en función de sus vecinos cercanos, al igual que recupera las imágenes con

mayor similitud visual.

La Figura 3-14 ilustra el proceso donde se realizan operaciones adyacentes entre las imágenes del conjunto de datos. Este algoritmo se utiliza para clasificar o predecir imágenes, agrupando puntos de datos individuales en función de su proximidad.

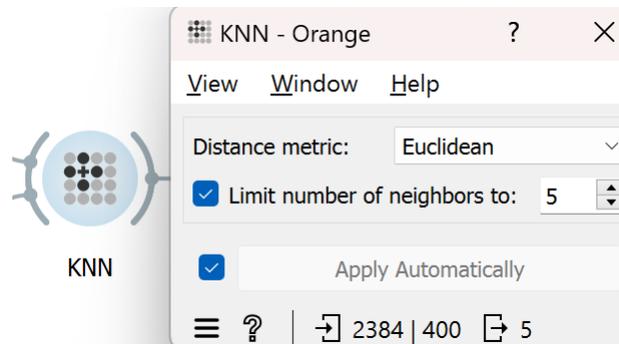


Figura 3-14: Métricas de Similitud de K-NN vecinos

Para comparar las medidas de similitud, se utilizó el widget (Distances), como se ilustra en la Figura 3-15. Durante el cálculo de la distancia, esta métrica evalúa la similitud entre dos puntos: uno perteneciente a la imagen de consulta y otro a la base de datos, permitiendo así medir la similitud o diferencia entre los vectores numéricos de las imágenes. Como resultado, se clasifican las imágenes más similares a la imagen de consulta.

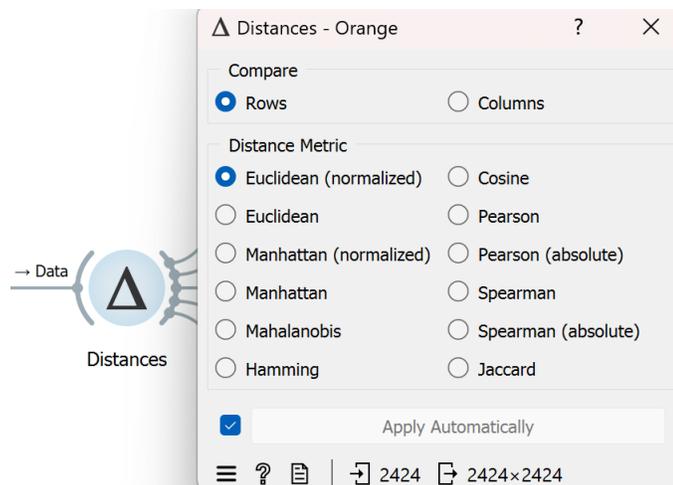


Figura 3-15: Medida de Similitud (Euclidiana)

La Figura 3-16 muestra el widget de (Hierarchical Clustering) utilizado para agrupar las categorías del modelo entrenado, en este caso imágenes histopatológicas de la base de datos, clasificando según sus niveles de similitud. Esta división se presenta en forma de ramificaciones para que se pueda elegir desde las agrupaciones más pequeñas hasta las más grandes

según las especificaciones que se necesite.

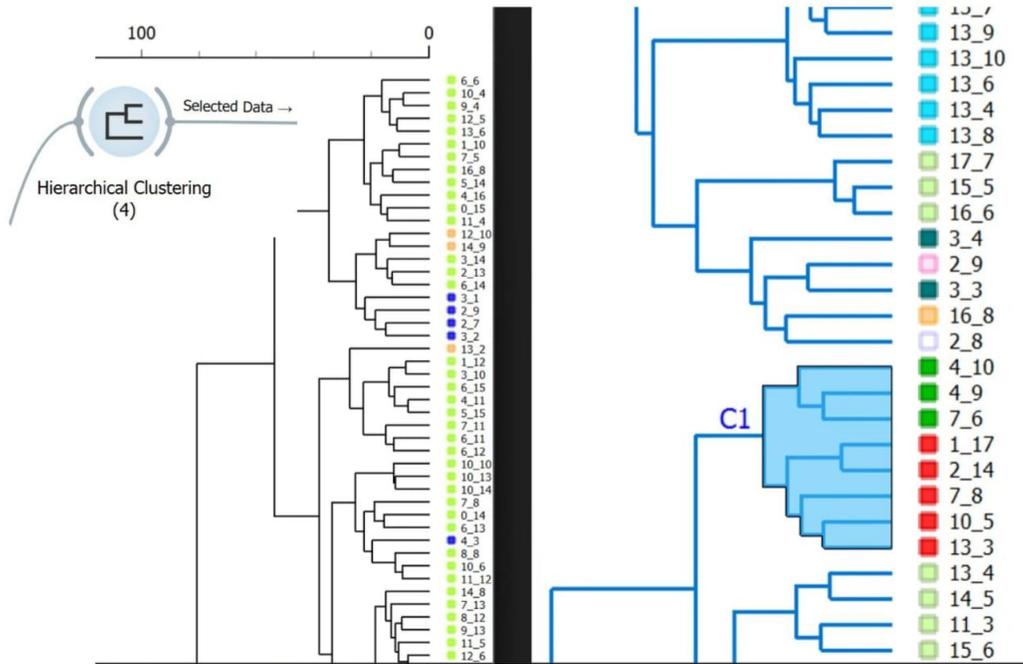


Figura 3-16: Clasificación del modelo

4 Capítulo 3

4.1. Evaluación y Resultados

- **Precisión**

Para evaluar la efectividad del sistema CBIR propuesto, la precisión es usada como método de medición del desempeño para recuperar imágenes similares, la misma que es determinada por la Ecuación 4-1, en donde se calcula este valor dividiendo el número de imágenes relevantes recuperadas por el sistema entre el número total de imágenes recuperadas. Matemáticamente, se expresa como:

$$\text{Precisión} = \frac{\text{Número de imágenes relevantes recuperadas}}{\text{Número total de imágenes recuperadas}} \times 100 \% \quad (4-1)$$

Una precisión alta indica que el sistema está devolviendo principalmente resultados relevantes, mientras que una precisión baja sugiere una alta proporción de resultados irrelevantes en la recuperación.

- **Recall**

Es conocido como la sensibilidad que representa el porcentaje de casos positivos detectados por nuestro modelo en relación con todos los casos positivos reales presentes en la base de datos. La Ecuación 4-2 de recall calcula este valor dividiendo el número de imágenes relevantes recuperadas entre el número total de imágenes relevantes en la base de datos. Matemáticamente, se expresa como:

$$\text{Recall} = \frac{\text{Número de imágenes relevantes recuperadas}}{\text{Número total de imágenes relevantes en la base de datos}} \times 100 \% \quad (4-2)$$

Un recall alto indica que el sistema está recuperando la mayoría de las imágenes relevantes, mientras que un recall bajo sugiere que se están perdiendo imágenes relevantes en la recuperación.

En este trabajo propuesto, se realizó una evaluación del desempeño del sistema mediante estas métricas para determinar la eficacia y la precisión de la búsqueda para encontrar imágenes similares. Además, se comparó el rendimiento del sistema con otros trabajos similares

presentados en la Tabla 4-1. Esta comparación proporcionó una base sólida para identificar aspectos que podrían mejorar el sistema CBIR.

Por lo tanto, estos valores fueron estimados mediante las curvas de rendimiento, las mismas que están implementadas en el widget (Performance Curve) de la interfaz de Orange. Esta herramienta permite visualizar de manera efectiva la relación entre la precisión y el recall del modelo.

La Figura 4-1 presenta los resultados obtenidos mediante la gráfica del rendimiento del modelo. En estas gráficas se observa cómo varían estos parámetros en diferentes configuraciones. Es importante destacar que los valores específicos de precisión y recall para este conjunto de datos se han determinado cuidadosamente durante el proceso de entrenamiento del modelo, logrando una precisión del 92.21% y un recall del 99.2% en los datos de entrenamiento. Estos resultados son indicativos del alto nivel de rendimiento alcanzado por el modelo en la tarea de clasificación significativa.

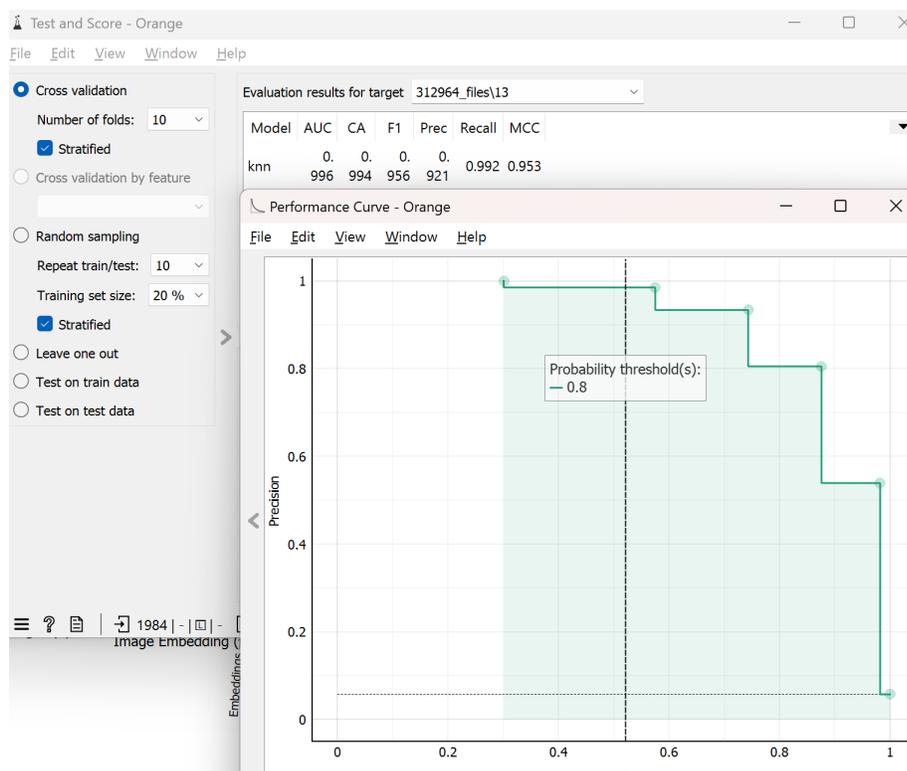


Figura 4-1: Gráfica en función Precisión-Recall conjunto de entrenamiento

Por otro lado, la Figura 4-2 muestra una gráfica con los valores de porcentaje de precisión de 0.615 y recall de 0.595, correspondientes de los datos de prueba. Al comparar los resultados obtenidos en los datos de entrenamiento con los de prueba, se evidencia la capacidad del modelo para generalizar a datos nuevos. La significativa diferencia entre los altos niveles

de precisión y recall en los datos de entrenamiento y los valores más bajos en los datos de prueba indica la posible presencia de sobreajuste o falta de generalización del modelo.

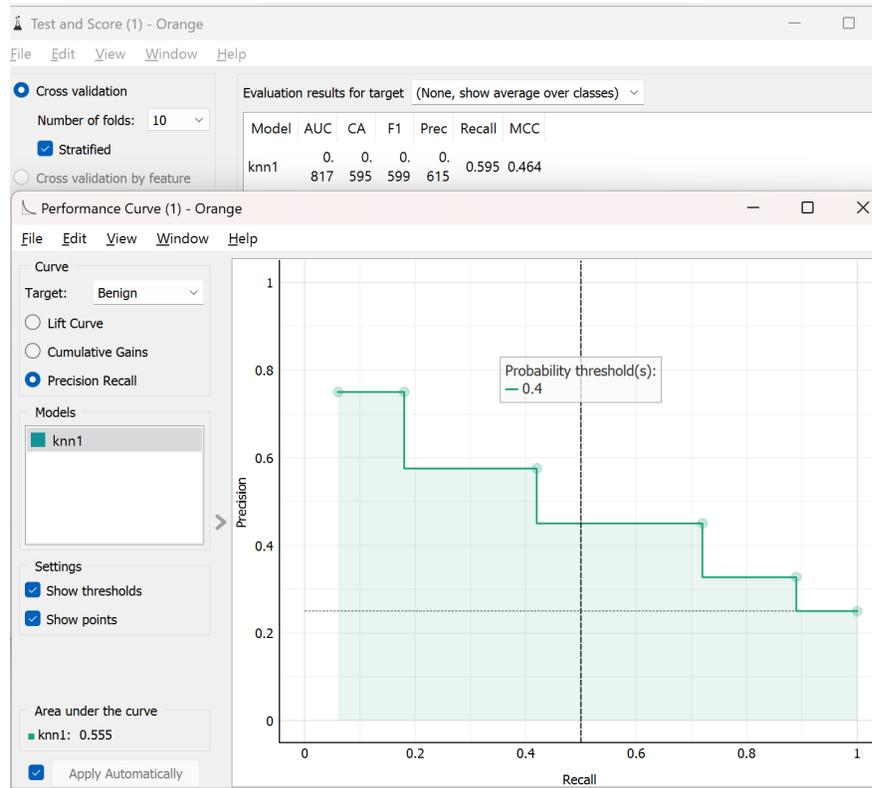
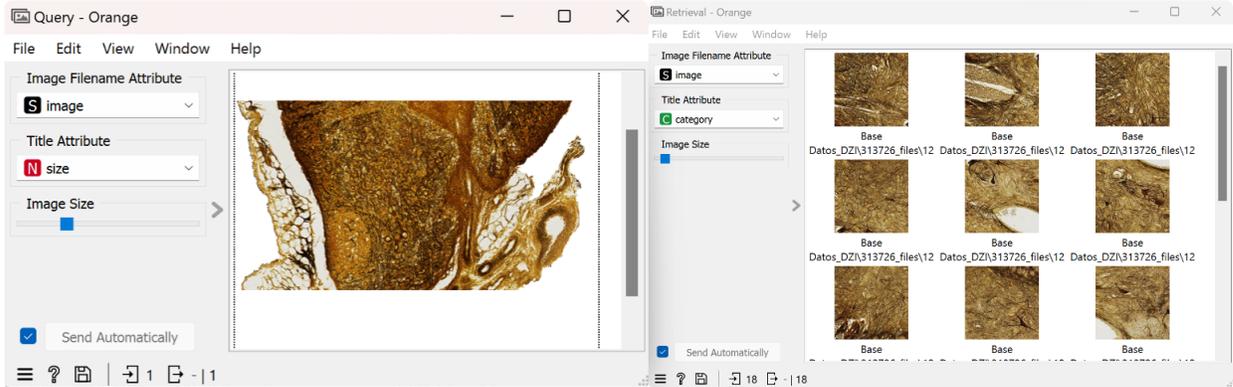


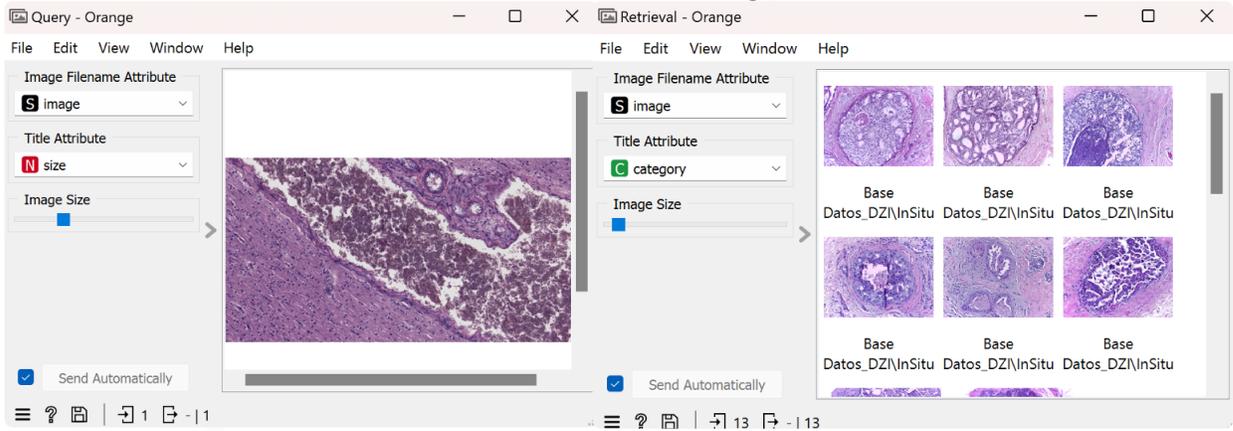
Figura 4-2: Gráfica en función Precisión-Recall conjunto de Pruebas

Finalmente, se presenta los resultados obtenidos para la tarea de recuperación de imágenes similares, la misma que evidencia una variedad de imágenes recuperadas basadas por su contenido visual, tras llevar a cabo una búsqueda en el conjunto de datos. Como se ilustra en la Figura 4-3, se observa la interfaz del visor proporcionada por el widget (Image Viewer), que ofrece una perspectiva detallada de las imágenes recuperadas. Este widget presenta una disposición que incluye una imagen de consulta (query) y una serie de imágenes recuperadas (retrieval). Como se puede observar, para el primer resultado se presentan 18 imágenes, para el segundo resultado se muestran 13 imágenes y para el tercer resultado 7 clases de imágenes de muestras histológicas similares. Cada conjunto de imágenes recuperadas representa una selección cuidadosamente evaluada que refleja la relevancia y la similitud con respecto a la imagen de consulta original. Este enfoque permite una exploración eficiente y detallada de las imágenes relacionadas, facilitando así el análisis y la interpretación de los resultados.

RESULTADO DE BUSQUEDA 1



RESULTADO DE BUSQUEDA 2



RESULTADO DE BUSQUEDA 3

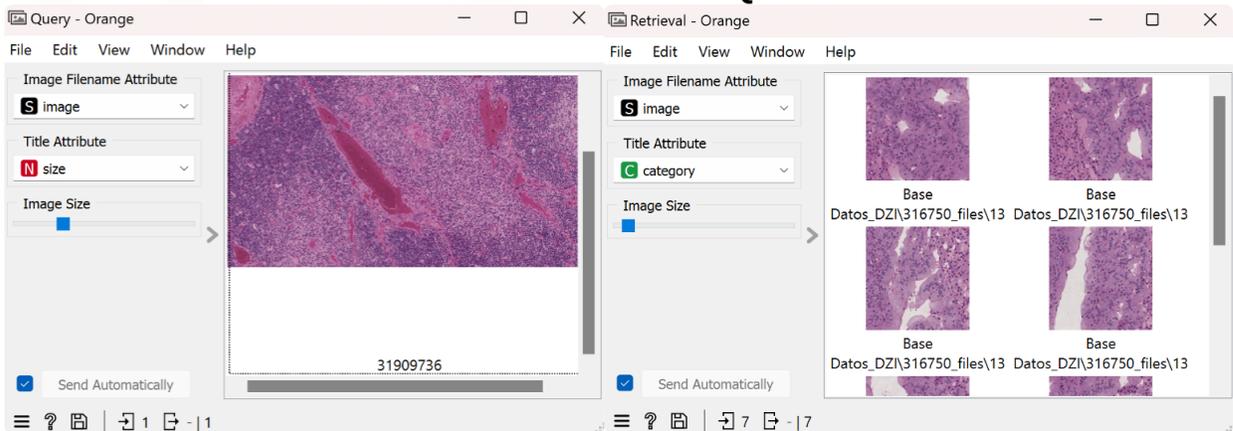


Figura 4-3: Esquema del Sistema de Recuperación de Imágenes Histológicas

4.1.1. Comparación del método

Para evaluar el rendimiento del modelo, se compararon dos métodos utilizando una base de datos de entrenamiento. La Tabla 4-1 presenta los resultados obtenidos. El método MIL-CNN empleado en un estudio sobre detección de cancer de mama [8], el cual alcanzó una precisión del 88.83% y un recall del 88.89%. Estos resultados indican un nivel aceptable de efectividad, aunque con un margen de error mínimo. En contraste, el método propuesto utilizando SqueezeNet demostró una mayor precisión, alcanzando un 92.10%. Además, logró un recall del 99.20%, lo que indica una capacidad destacada para identificar instancias positivas sin generar muchos falsos positivos. Estos hallazgos sugieren que el modelo basado en SqueezeNet es significativamente más efectivo en términos de precisión y capacidad de identificación de instancias positivas, lo que lo convierte en una opción sólida para aplicaciones de clasificación de imágenes. Sin embargo no se puede asegurar que la capacidad sea significativamente exacta debido a que son bases de datos diferentes.

| Método | Precision | Recall |
|------------|-----------|--------|
| MIL-CNN | 88.83 | 88.89 |
| SqueezeNet | 92.10 | 99.92 |

Tabla 4-1: Método comparativo

5 Conclusiones y recomendaciones

5.1. Conclusiones

Los resultados presentados en la sección anterior muestran el desempeño del sistema propuesto y su capacidad de recuperar imágenes histopatológicas al variar el valor de K vecinos cercanos mediante agrupamientos de las imágenes. Este sistema permite al usuario seleccionar grupos específicos de imágenes para análisis, desde pequeños conjuntos altamente precisos hasta agrupaciones más amplias con características similares y diferentes vecinos cercanos. Así se han demostrado que la representación de las características visuales de las imágenes tienen una correspondencia visual con las búsquedas desde la base de datos, lo que determina que la medida de similitud se ajusta a los descriptores extraídos mediante el aprendizaje profundo.

Además, se estableció un repositorio que contiene una amplia variedad de imágenes histológicas, las cuales han sido clasificadas en varias categorías. Este repositorio representa un aporte para futuras evaluaciones de sistemas de recuperación de imágenes basados en contenido (CBIR), mientras exista gran cantidad de imágenes se puede optimizar al modelo entrenado con la red neuronal.

También se aplicarán técnicas de mejoramiento y procesamiento de imágenes con el objetivo de favorecer su calidad y prepararlas adecuadamente para su análisis en modelos de aprendizaje automático.

Finalmente, se ha implementado y evaluado el modelo en una interfaz gráfica, que permite a los usuarios modificar los flujos de trabajo utilizando Widgets interactivos diseñados para realizar tareas específicas, de almacenamiento de bases de datos, procesamiento, modelado y una evaluación del rendimiento del esquema propuesto.

5.2. Recomendaciones

Una recomendación técnica para enfrentar el desafío de manejar grandes volúmenes de imágenes histológicas de diapositivas completa (WSI), es utilizar herramientas que permitan la

división de las imágenes en parches con una resolución estandarizada. Esta estrategia tiene como objetivo resolver el problema de almacenamiento de metadatos o giga píxeles.

Al dividir las imágenes se reducirá significativamente la carga computacional y los requisitos de almacenamiento. Esto conducirá a un manejo más eficiente de las imágenes histológicas y acelerará su procesamiento. Además, al trabajar con parches más pequeños, será posible aplicar técnicas de análisis y procesamiento de manera más focalizada, lo que mejorará a los algoritmos de aprendizaje de máquina.

Bibliografía

- [1] Asier Alcaide. Redes neuronales convolucionales siamesas aplicadas a la verificación facial, 2020.
- [2] F Alvarez, N Alonso, and J De Haro. Inteligencia artificial y modelado avanzado en cirugía vascular. *Implicaciones para la práctica clínica. Rev Angiolol*, 73(2):65–75, 2021.
- [3] Guilherme Aresta, Teresa Araújo, Scotty Kwok, Sai Saketh Chennamsetty, Mohammed Safwan, Varghese Alex, Bahram Marami, Marcel Prastawa, Monica Chan, Michael Donovan, et al. Bach: Grand challenge on breast cancer histology images. *Medical image analysis*, 56:122–139, 2019.
- [4] José Orlando Maldonado Bautista. Estudio de métodos de indexación y recuperación en bases de datos de imágenes. *Universidad del País Vasco*, 2008.
- [5] J Caicedo, F Gonzáles, E Romero, and Y Sánchez. Evaluación de características de bajo nivel para recuperación de imágenes de histología por contenido. In *Congreso Colombiano de Computación*, 2007.
- [6] Laura Calcagni. Redes generativas antagónicas y sus aplicaciones, 2020.
- [7] David Castillo Berná et al. Redes neuronales convolucionales para el reconocimiento de enfermedades de la piel. 2022.
- [8] Kausik Das, Sailesh Conjeti, Jyotirmoy Chatterjee, and Debdoot Sheet. Detection of breast cancer from whole slide histopathological images using deep multiple instance cnn. *IEEE Access*, 8:213502–213511, 2020.
- [9] Nilanjan Dey, Yu-Dong Zhang, V Rajinikanth, R Pugalenth, and N Sri Madhava Raja. Customized vgg19 architecture for pneumonia detection in chest x-rays. *Pattern Recognition Letters*, 143:67–74, 2021.
- [10] S Donís. Desarrollo de una red generativa antagónica para la generación de imágenes de flores falsas, 2022.
- [11] Atoany Fierro, Mariko Nakano, Keiji Yanai, and Héctor Pérez. Redes convolucionales siamesas y tripletas para la recuperación de imágenes similares en contenido. *Información tecnológica*, 30(6):243–254, 2019.

-
- [12] J Franganillo. La inteligencia artificial generativa y su impacto en la creación. *Revista de ciencias sociales*, 11(2):15, 2023.
- [13] GeeksforGeeks. K-nearest Neighbours (KNN). 2024.
- [14] M Hanna, A Parwani, and S Sirintrapun. Whole slide imaging: Technology and applications. *Anatomic Pathology*, 2020.
- [15] María Herrero Tudela et al. Procesado de retinografías basado en deep learning para la ayuda al diagnóstico de la retinopatía diabética. 2020.
- [16] Saeed Iqbal, Adnan N. Qureshi, Jianqiang Li, and Tariq Mahmood. On the analyses of medical images using traditional machine learning techniques and convolutional neural networks. *Archives of Computational Methods in Engineering*, 30(5):3173–3233, 2023.
- [17] Hyun-Jong Jang, In Hye Song, and Sung Hak Lee. Generalizability of deep learning system for the pathologic diagnosis of various cancers. *Applied Sciences*, 11(2):808, 2021.
- [18] Amandeep Khokher and Rajneesh Talwar. Content-based image retrieval: Feature extraction techniques and applications. In *International conference on recent advances and future trends in information technology (iRAFIT2012)*, pages 9–14, 2012.
- [19] Daichi Kitaguchi, Nobuyoshi Takeshita, Hiroki Matsuzaki, Hiro Hasegawa, Takahiro Igaki, Tatsuya Oda, and Masaaki Ito. Deep learning-based automatic surgical step recognition in intraoperative videos for transanal total mesorectal excision. *Surgical Endoscopy*, pages 1–9, 2022.
- [20] Brett Koonce and Brett Koonce. Mobilenetv3. *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*, pages 125–144, 2021.
- [21] Ashnil Kumar, Jinman Kim, Weidong Cai, Michael Fulham, and Dagan Feng. Content-based medical image retrieval: a survey of applications to multidimensional and multi-modality data. *Journal of digital imaging*, 26:1025–1039, 2013.
- [22] N Kumar, R Gupta, and S Gupta. Whole slide imaging (wsi) in pathology: Current perspectives and future directions. *Journal of digital imaging*, 33(4):1034–1040, 2020.
- [23] D Lanzagorta, D Carrillo, and R Carrillo. Inteligencia artificial en medicina: presente y futuro. *Gaceta médica de México*, 158:17–21, 2022.
- [24] D Marín and E Romero. Sistemas de microscopía virtual: análisis y perspectivas. *Biomédica*, 31(1):144–155, 2011.
- [25] O Mera, J Pérez, and M López. Conocimientos de histología con relevancia en la práctica del médico en ecuador. *Conrado*, 14(63):20–24, 2018.

- [26] Fabián Rodrigo Narváez Espinoza. *Sistema de Anotación para Apoyo en el Seguimiento y Diagnóstico de Cáncer de Seno*. PhD thesis, 2010.
- [27] Ahmet Murat Ozbayoglu, Mehmet Ugur Gudelek, and Omer Berat Sezer. Deep learning for financial applications: A survey. *Applied soft computing*, 93:106384, 2020.
- [28] Andrea Pachamé and Enrique Portiansky. Microscopía virtual: una nueva herramienta tecnológica para la enseñanza de la histología y la patología. 2017.
- [29] Raj Kumar Pattanaik, Satyasis Mishra, Mohammed Siddique, Tiruveedula Gopikrishna, Sunita Satapathy, et al. Breast cancer classification from mammogram images using extreme learning machine-based densenet121 model. *Journal of Sensors*, 2022, 2022.
- [30] VB Prasath, Haneen Arafat Abu Alfeilat, Ahmad Hassanat, Omar Lasassmeh, Ahmad S Tarawneh, Mahmoud Bashir Alhasanat, and Hamzeh S Eyal Salman. Distance and similarity measures effect on the performance of k-nearest neighbor classifier—a review. *arXiv preprint arXiv:1708.04321*, 2017.
- [31] Cristian Camilo Pulgarín Ospina. Diseño, desarrollo e implementación de un servicio web para la anotación de imágenes de alta resolución: Integración de modelos de deep learning para la segmentación automática de regiones. 2022.
- [32] Yaine Quintanilla Cruz and Roger Julio González Hernández. Herramienta para la realización de flujos de trabajo de minería de datos usando r. B.S. thesis, Universidad de las Ciencias Informáticas. Facultad 6, 2015.
- [33] Marcial García Rojo. Aplicaciones de deep learning, un método de inteligencia artificial, en anatomía patológica. *Libro Blanco 2021 de la Anatomía Patológica en España*, page 33.
- [34] Marcial García Rojo, Gloria Bueno García, Carlos Peces Mateos, Jesús González García, and Manuel Carbajo Vicente. Preparaciones digitales en los servicios de anatomía patológica (ii). análisis de soluciones existentes. *Rev Esp Patología*, 38(4):207–220, 2005.
- [35] Diego Sanchez. La inteligencia artificial: una revolución al alcance de todos. *Pontificia Universidad Católica de Perú*, 2023.
- [36] S Sánchez. Content-based image retrieval using descriptors generated by convolutional neural networks. *Revista Cubana de Ciencias Informáticas*, 12(4):78–90, 2018.
- [37] Harshita Sharma, Alexander Alekseychuk, Peter Leskovsky, Olaf Hellwich, Radhey Shyam Anand, Norman Zerbe, and Peter Hufnagl. Determining similarity in histological images using graph-theoretic description and matching methods for content-based image retrieval in medical diagnostics. *Diagnostic pathology*, 7(1):1–20, 2012.

-
- [38] A Sridhar, S Doyle, and A Madabhushi. Content-based image retrieval of digitized histopathology in boosted spectrally embedded spaces. *Journal of pathology informatics*, 6(1):41, 2015.
- [39] Alejandro Suárez. Aplicación de técnicas de aprendizaje profundo (deep learning) a clasificación de imágenes histológicas. *Escuela Técnica Superior de Ingeniería*, 2018.
- [40] Paula Toro. Construcción de un banco de imágenes virtuales de dermatopatología para la docencia en posgrado. *Universidad Nacional de Colombia*, 2019.
- [41] Sergio Rodríguez Vaamonde. *Detección y categorización de objetos invariante y multivista en imágenes digitales mediante visión artificial bioinspirada*. PhD thesis, Universidad del País Vasco-Euskal Herriko Unibertsitatea, 2016.
- [42] Aleksandra Zuraw and Famke Aeffner. Whole-slide imaging, tissue image analysis, and artificial intelligence in veterinary pathology: An updated introduction and review. *Veterinary Pathology*, 59(1):6–25, 2022.