



## PROYECTO DE TESIS

- ☐ Doctorado en Ingeniería Informática  
☒ Magíster en Ciencias de la Ingeniería Informática

|  |  |
|--|--|
| 1. Título del Proyecto de Tesis                  | A Content-Based Medical Image Retrieval for Pap Smear Images with Generative Adversarial Networks and Nearest Neighbor |
| 2. Nombre del Alumno                             | Camilo Esteban Núñez Fernández   |
| 3. Número de Teléfono - Celular                  | (+56) 9 42131666   |
| 4. Correo electrónico                            | camilo.nunezf@sansano.usm.cl   |
| 5. Fecha de Ingreso al Programa                  | 2021 - 1   |
| 6. Pregrado<br>(Título o Grado Institución, Año) | Licenciado en Ciencias de la Ingeniería Informática, Universidad Técnica Federico Santa María, 2021                    |
| 7. Profesor Guía de Tesis                        | Mauricio Solar   |
| 8. Fecha Presentación Tema de Tesis              |  |
| 9. Fecha Aprobación Tema de Tesis                |  |
| 10. Fecha Tentativa de Término                   | 30/07/2023   |
| 11. Comisión Interna de Graduación               |  |



## Resumen

El cáncer cervicouterino es el cuarto cáncer más común en mujeres en términos de incidencia y mortalidad en el mundo, y tercero en Chile. Debido a su actual prevalencia e impacto sobre la sociedad civil, es que se han realizado múltiples esfuerzos para su detección temprana por medio de tests. El test papanicolaou o pap smear test ha sido la metodología de detección temprana contra el cáncer cervicouterino más utilizado en el mundo.

Sin embargo, al tratarse de un test de análisis de citología manual, sufre altos falsos positivos, además de existir un bajo ratio de análisis efectivo por parte de los citotecnólogos debido a la complejidad que lleva el análisis celular. Para sopesar esto, múltiples instituciones han implementado sistemas de recuperación de imágenes (CBIR) para la recomendación asistida en el proceso de análisis. No obstante, debido a los nuevos avances en la calidad de las imágenes y al gran volumen con el cual se generan, es que los CBIR deben aplicar mejores algoritmos para mejorar su funcionamiento.

El presente trabajo busca investigar, desarrollar e implementar un CBIR eficiente y robusto para imágenes médicas de papanicolaou. Para lograr ello, se utilizará una red neural generativa para la extracción de factores representativos de las imágenes, los cuales serán indexados comparados por medio de la técnica de grafos de vecinos más cercanos.

Para validar y evaluar el CBIR propuesto se utilizarán los datasets ISBI Challenges 2014-2015, Herlev, SIPaKMeD y *liquid-based cytology pap smear dataset*, y se compararán con modelos de extracción de factores representativos del estado del arte aplicados en distintos CBIR.

Se espera que el CBIR propuesto pueda ser implementado en ambientes de producción clínicos como soporte en el análisis de citología para las muestras de papanicolaou en entidades nacionales e internacionales.

**Palabras Claves:** Cáncer cervicouterino, histopatología, pap smear, content-based image retrieval, deep learning, generative adversarial networks, segmentación células cervicouterinas, clasificación células cervicouterinas, grafos vecinos más cercanos.



## Abstract

Cervical cancer is the fourth cancer type most common and deadliest among women in the world, being the thirds in Chile. Due to its current prevalence and impact on the society, multiple efforts have been made for their early detection through tests. The pap smear test has been the most common screening method used for the early detection against cervical cancer in the world.

However, this test as it's a hand-operated method, suffers from high false positives cases, in addition to lowers effective rate of analysis by cytotechnologist due to the complexity involved in single-cell analysis. To solve this problem, multiple institutions have implemented content-based image retrieval (CBIR) systems for computer-aided in the analysis process. Nevertheless, due to new advances in the image quality and the large volume of data that generate this, CBIRs must apply better algorithms to improve the performance.

This thesis proposed it's focused on investigate, develop and implement an efficient and robust CBIR for pap smear images. To achieve this, a novel generative neural network is proposed for the feature extraction of images, which will be indexed a retrieval by a graph-based nearest neighbor algorithm.

The proposed CBIR will be evaluated and validated over the datasets ISBI Challenges 2014-2015, Herlev, SIPaKMeD and liquid-based cytology pap smear dataset, and will be compared with multiples features extraction methods availables in the current state of the arts.

The CBIR results are expected to assist medical specialists in the pap smear test analysis.

**Keywords:** Cervical cancer, cytopathology, pap smear, content-based image retrieval, deep learning, generative adversarial networks, cervical cell segmentation, cervical cell classification, nearest neighbor graph.

# 1. Formulación General de la Problemática y Propuesta Tesis

## 1.1. Antecedentes Generales del Problema

El cáncer cervicouterino es el cuarto cáncer más común en mujeres en términos de incidencia y mortalidad [WHO20]. Debido a esto, la World Health Organization (WHO) a fomentado de manera constante programas orientados a la prevención y detección por medio de prevalencia en la evacuación de contra el *human papillomavirus* (HPV) y la aplicación de métodos de *screening* en los sistemas de salud a lo largo del mundo [WHO20].

En Chile durante los años 2010 y 2016 se registraron 164.239 egresos por causas oncológicas en mujeres, de entre los cuales las principales tres causas más frecuentes son cáncer de mama (52.189), seguido del cáncer a la tiroides (20.488), y cáncer cervicouterino (16.946) [Sub18]. Frente a esto, como política de prevención, el Departamento de Manejo Integral del Cáncer del Minsal, determinó como examen de prevención temprana para el cáncer cervicouterino, la aplicación del examen Papanicolaou (*pap smear*) para grupos objetivos de mujeres entre 25 - 64 años [Sub18]. De este modo, y en base a las políticas de prevención tanto en Chile como en el mundo, es que el método de histopatología más utilizado es el *pap smear* o *liquid-based cytology method*, debido a su bajo costo y alta efectividad [SSL<sup>+</sup>12, DBI<sup>+</sup>06, Sub18].

La aplicación tradicional del *pap smear* test consiste en la recolección de células desde la terminal *squamocolumnar* del cervix, las cuales son puestas en un portaobjetos (*glass slides*) para microscopio y examinadas por un citotecnólogo [Pap73, PT41]. Sin embargo, el proceso de análisis (*screening process*) por parte del especialista es la fase más costosa en tiempo y recursos al momentos de efectuar el test [MZCH20]; a modo de ejemplo, esta fase toma entre 5 a 10 minutos de analizar por muestra de acuerdo a la dificultad de análisis de las células, esto provoca que un citotecnólogo no pueda analizar más de 70 muestras por día [EAC<sup>+</sup>12].

Por lo anterior, y considerando además el aumento exponencial en los volúmenes de datos y en la mejor/mayor calidad de las imágenes que generan los nuevos instrumentos, es que se torna necesario el uso de sistemas de herramientas computacionales para la asistencia profesional al citotecnólogo durante la fase de diagnostico y análisis celular. Estos sistemas son identificados en la literatura como *computer-aided diagnosis* (CAD) [Doi07].

## 1.2. Definición del Problema

Los sistemas *computer-aided diagnosis* pueden encontrarse en múltiples soluciones profesionales, siendo las más características los *image-based diagnosis systems* [MMBG04]. Sin embargo, estos sistemas al contar con masivos volumen de datos almacenados, no permiten a los citotecnólogo, o tecnólogos en general, optimizar el estudio de los casos clínicos importantes al momento de referenciar regiones importantes en las imágenes de muestras, lo cual mantiene los tiempos de diagnósticos [MKCC<sup>+</sup>10]. Para sopesar esta situación, es que se vuelve menester el uso de sistemas recomendadores de imágenes, o *Content-Based Medical Image Retrieval* (CMBIR) durante el proceso de análisis de las células [XLA<sup>+</sup>08].

Los CMBIR, al ser una variante de los *Content-Based Image Retrieval* (CBIR), dependen de los *features vectors* que puedan describir la imagen médica, los cuales pueden ser obtenidos por medios de *handcraft-annotations* realizados por los profesionales, o por medio de técnicas supervisadas o no supervisadas de *machine learning* [SWS<sup>+</sup>00]. Sin embargo, problemas como el *overlapping* de células [ZSM<sup>+</sup>22], la semántica general de la imagen [RLW<sup>+</sup>20], la indexación eficiente y la ausencia de *feedback-based* [SRV21], suelen afectar el desempeño de los CMBIR; parte de estos problemas pueden ser tratados al construir *features* con descriptores más óptimos y robustos respecto a las propiedades de las imágenes.

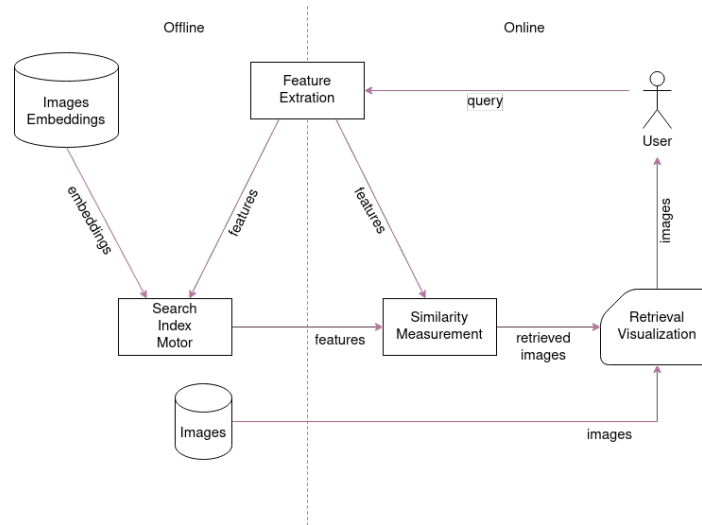


Figura 1: Workflow de funcionamiento de un CBIR. Se considera una parte offline, la cual corresponde a aquellos datos previamente calculados, como los *features vectors* de las imágenes, y una parte online que considera la *query image* ingresada por el usuario, la extracción de *features* de esta imagen, el calculo de distancia con respecto a los vectores ya cálculos, y el retorno y visualización de las imágenes más cercanas. Imagen modifica desde [KKC<sup>+</sup>13].

### 1.3. Marco Teórico

#### 1.3.1. Content-based image retrieval

Según lo mencionado anteriormente, los *CMBIR* son una variante de los *CBIR*, contando con la única diferencia en el tipo de imágenes que se indexan y se retorna al usuario, las cuales en este caso particular se tratan de *pap smear images*. A su vez, un *CBIR* es una técnica de búsqueda de imágenes, en donde dada una *query*, se retornan las imágenes más similares según una métrica definida. Las métricas que se utilizan miden esencialmente la similitud entre propiedades (*features*) de los elementos visuales que describen a la imagen, y que van desde el color, la forma, la distribución espacial de los elementos, o las regiones de interés (*ROI*) [DJLW08]. La extracción de estos *features* es determinante al momento de indexar las imágenes que utilizará el sistema recomendador.

El workflow de un *CBIR* se muestra en la Figura 1. La composición del workflow se centra en dos secciones, una offline en la cual *features* ya extraídos de las imágenes son almacenados e indexados por el motor de búsqueda; y una online donde la *image query* realizada por el usuario es procesa y sus *features* extraídos, los cuales son comparados con respecto a los *features* ya indexados por medio de un algoritmo de similitud.

#### 1.3.2. Feature Extraction

Un conjunto de tareas básicas que debe realizar un **computer-aided diagnosis** son la segmentación de la célula (incluyendo núcleo y citoplasma), la clasificación, y la extracción de *features* [WCH<sup>+</sup>18]. Para resolver estas tareas, es posible aplicar múltiples técnicas de deep learning durante el procesamiento de las imágenes. Estas técnicas pueden ser supervisadas, semi-supervisadas o no supervisadas. Un detalle resumido de las principales técnicas utilizadas en el procesamiento de células de *pap smear* se encuentra en la Figura 2. A partir de estas múltiples técnicas es que es posible crear distintos tipos de *features vectors* tanto de la imagen original, como de las mascarar de segmentación (i.e. *semantic mask* y *boundary mask*), los cuales pueden ser usados a su vez como *features vectors* para ser indexados.

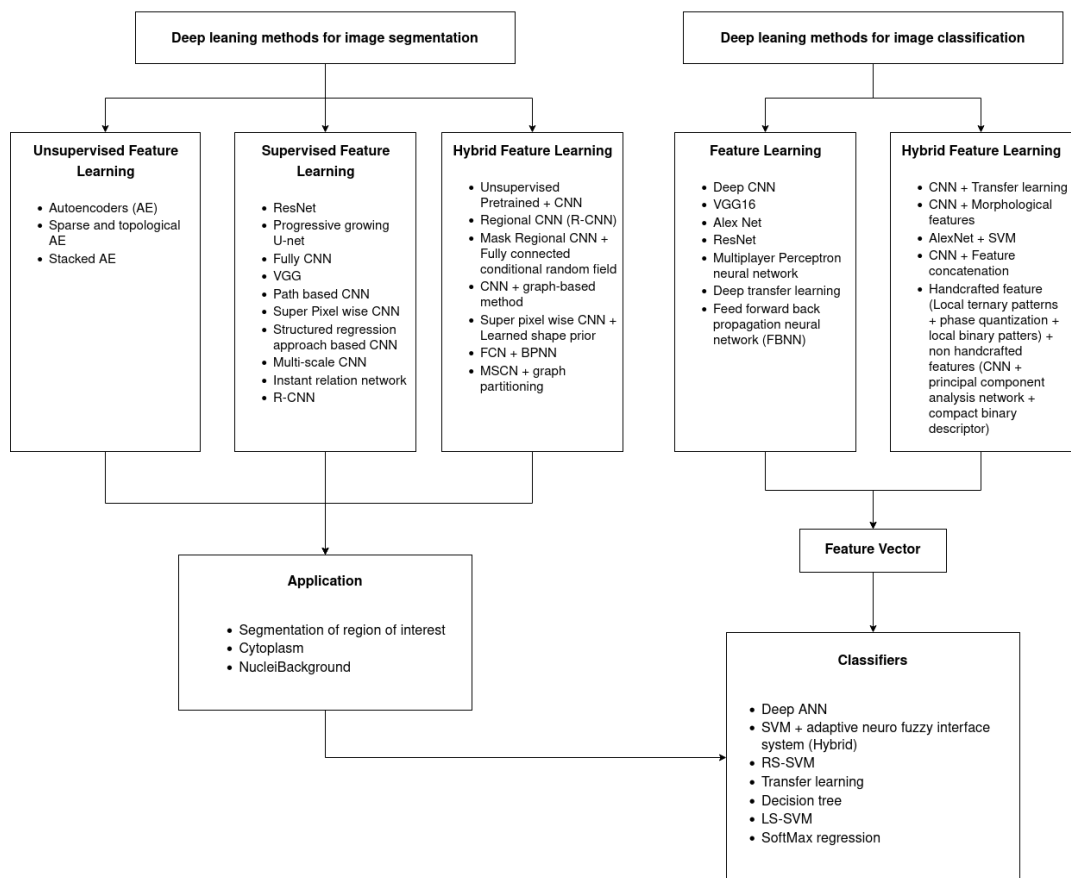


Figura 2: Principales algoritmos de deep learning usados en el análisis de imágenes de *pap smear*. Imagen modifica desde [RLW<sup>+</sup>20].

## 1.4. Trabajos Relacionados

### 1.4.1. Content-based medical image retrieval para cáncer cervicouterino

En materia de *CMBIR*, es posible encontrar en la literatura como primer trabajo usando *cervigram images* para el soporte en el diagnóstico del cáncer Cervicouterino el realizado por Xue Z. *et al.* en [XLA<sup>+</sup>08], donde se desarrolla un *framework web* para la selección de múltiples *ROI* de una *query* dada, los cuales son procesados para su extracción de *features* y comparados con valores pre-computados y almacenados en una base de datos. Los *features* seleccionados son el color, por medio de un histograma de intensidad de colores, la textura de la región obtenida por medio de la *Discrete Wavelet Transform*, y el tamaño de los *ROI*. Estos tres *features* son combinados y almacenados. Luego, al realizar una consulta, el usuario puede seleccionar la métrica de similaridad a usar, las cuales pueden ser *Minkowski-form*, distancia de histogramas, divergencia de *Jeffrey*, *quadratic-form* y *Earth mover's distance*. Finalmente, las métricas de evaluación usadas para el sistema son *precision* y *accuracy*.

Un enfoque en el uso de *CMBIR* para *pap smear images* se puede encontrar en el trabajo de Mera M. *et al.* [MGRMG15]. En este trabajo, la extracción de *features* se realizó por medio de *Tamura textures*, *Discrete Wavelet Transform* y el cálculo de la matriz de co-ocurrencias. En esta oportunidad se utilizó el algoritmo de *K-nearest neighbors* para retornar las *K* imágenes más similares de acuerdo a la *query* realizada por el usuario.

Respecto al uso de algoritmos de deep learning para la construcción de los *features vectors*, es importante destacar los trabajos de Praveena H. *et al.* [PGK<sup>+</sup>22] para *pap smear images* y de Ahmed A. *et al.* [AA<sup>+</sup>22] para *cervigram images*.

En el caso del trabajo de Praveena H. *et al.*, se realizó una propuesta híbrida de extracción de *features*, por medio de una red *ICNN* y los descriptores globales *HOG* [NSM18] y *LBP* [PSR18]; la recuperación de las imágenes se realizó por medio del cálculo de similaridad usando la distancia *chi-square*. Las métricas de evaluación utilizadas por los autores fueron *accuracy*, *precision*, *specificity*, *recall* y *F-score*.

Respecto al trabajo de Ahmed A. *et al.*, se utilizaron dos modelos para la extracción de *features*, *ResNet-18* [ZZC<sup>+</sup>17] y *SqueezeNet* [GKCA17]. Ambos modelos fueron pre-entrenados con el dataset *ImageNet* [DDS<sup>+</sup>09] para aplicar *transfer learning*. La métrica utilizada para la similaridad entre *features vectors* fue distancia euclidiana. Para evaluar la performance de los modelos y del *CMBIR*, se utilizó las métricas de *precision* y *recall*.

### 1.4.2. Feature Extraction para *pap smear images*

Los intereses en el desarrollo del presente proyecto de tesis se centrarán en dos técnicas *state of the art* para la extracción de *features* en *pap smear images*: (I) *Generative Adversarial Networks* y (II) *Feature Distillation Network* para segmentación y clasificación.

Huang J. *et al.* proponen la red *Cell-GAN* en su trabajo [HYL<sup>+</sup>21], para la segmentación de *pap smear images*, por medio de un generador tipo encoder-decoder con doble entrada, una para la imagen y otra para un factor de guía, el cual indica la locación de los núcleos (*guide factor*) de las células en una imagen completa, mientras la arquitectura usada en el discriminador corresponde a un modelo *inception* [VLBM08]. Dado que en este trabajo, el objetivo principal era generar la *boundary mask* de una imagen, se utilizó el *dice coefficient* para evaluar la calidad de la segmentación.

Por otro lado, Ilyas T. *et al.* proponen la red *Tissue specific feature distillation network (TSFD-Net)* [IMK<sup>+</sup>22], capaz de realizar las tareas de segmentación y clasificación simultáneamente. Esta red se compone de tres elementos principales: (I) un backbone primario para la extracción de *features* utilizando la combinación de las arquitecturas *Mobile-Net-v2* [SHZ<sup>+</sup>18] y *squeeze-excitation sub-network* [HSS18], (II) seguido de paths *cross-scale weighted feature fusion (CWFF)* para la combinación de *features*, reteniendo las propiedades espacial y locacional de la imagen; finalmente se considera un *interlinked decoder (ID)* que corresponde a dos decoders interconectados, uno para la *semantic mask* y otro para la *boundary mask*. La arquitectura descrita se puede observar en la Figura 3. Finalmente, los autores utilizaron la métrica *Panoptic quality (QA)* [GVR<sup>+</sup>19, GAKB<sup>+</sup>19] para evaluar el perfor-



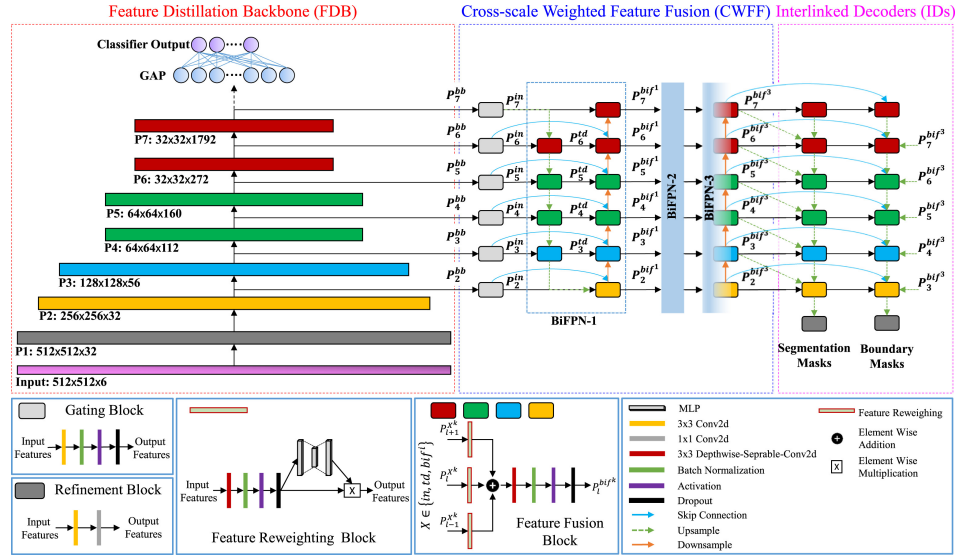


Figura 3: Arquitectura de la red *TSFD-Net* propuesta por *Talha Ilyas et al.* [IMK<sup>+</sup>22].

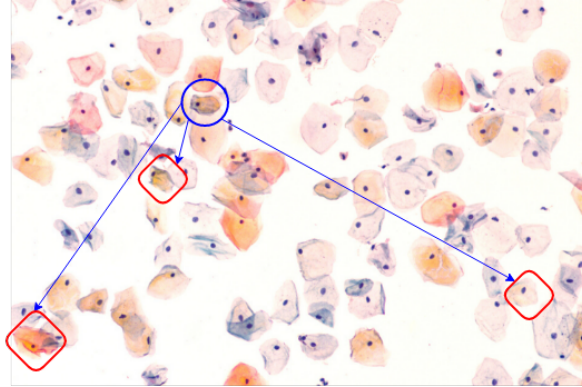


Figura 4: Consulta *single cell* respecto a las demás células de una imagen *pap smear*. El círculo en azul corresponde a la *query* y los rombos en rojo a las células más similares a la *query*.

mance del proceso de segmentación, y las métricas de *precision*, *recall*, y *F-score* para el proceso de clasificación.

### 1.5. Propuesta de Tesis

El desarrollo de esta tesis y su respectiva propuesta se centrará en la confección de un *framework* para *CMBIR* de imágenes *pap smear*, con un enfoque *computer-aided diagnosis* para las tareas de clasificación y segmentación. Este *framework* tendrá la capacidad de recomendar tanto imágenes *single-cell* como *multi-cell* de acuerdo al tipo de búsqueda que desee el usuario. Además de considerar a su vez la *interlocation* de células similares en un tejido a analizar según el tipo de clasificación de la célula consultada (ver Figura 4).

Para desarrollar esta propuesta, se considerará la experimentación de múltiples combinaciones de *features vectors* para indexar. Para ello se propone el análisis y estudio de un modelo generativo que incluye el entrenamiento de un generador de tres outputs: reconstrucción de una imagen *pap smear*, su *semantic mask* y *boundary mask*. Además del entrenamiento de tres discriminadores, uno para output del generador. Esta arquitectura se muestre en la Figura 5; a partir de este arquitectura, se busca estudiar la combinación de distintos tipos de *features*, como los entregados por el backbone del generador



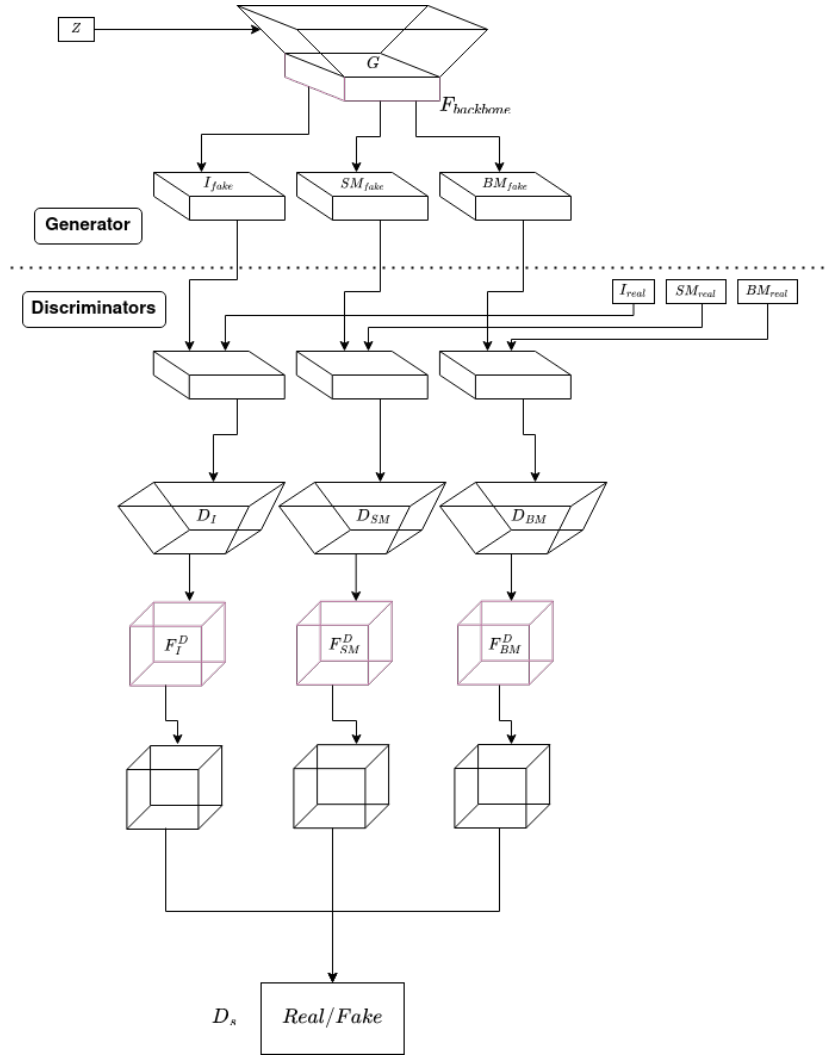


Figura 5: Arquitectura propuesta para un modelo generativo con un único generador con múltiples *outputs* y múltiples discriminadores para los *outputs*.

( $F_{backbone}$ ) o los entregados por los discriminadores ( $F_I^D, F_{SM}^D, F_{BM}^D$ ). Para el desarrollo del generador, se estudiara la factibilidad de modificación de la red *TSFD-Net* [IMK<sup>+</sup>22], particularmente el *FDB*, para considerar ahora una reconstrucción total de la imagen original.

Para el entrenamiento de la red, se planea utilizar los datasets ISBI Challenges 2014-2015 [LCB15][LCB<sup>+</sup>17], Herlev [MDJ09][DBJ<sup>+</sup>06][ADJ04][JD06], SIPaKMeD [PDS<sup>+</sup>18] y liquid-based cytology Pap smear dataset [HMBD20], los cuales en total suman más de 6800 imágenes *single-cell* y *multi-cells* de *pap smear*.

Finalmente, se estudiará el uso de un algoritmo *graph-based* de Nearest Neighbor para el cálculo de similitudes entre los *features vectors*. En particular se implementará el método *HNSW* [MY16, BN13], debido a sus excelentes resultados en el benchmark realizado por Aumüller M. et al. [ABF18].



## 2. Hipótesis de Trabajo

- Es posible generar una combinación de *features vectors* más robustos y de forma eficiente para un *CMBIR* de imágenes *pap smear* a partir de un red neural de clasificación y segmentación múltiple, entrenada de manera *adversarial*.
- El *CMBIR* propuesto es capaz de mantener o superar el tiempo y la eficiencia en la recuperación de imágenes *pap smear* respecto a otros *CMBIR* del *state of the art*.



### 3. Objetivos

#### 3.1. Objetivos Generales

El objetivo general de la presente propuesta es confeccionar e implementar un método eficiente de recuperación de imágenes médicas *pap smear* por medio de similitudes en los descriptores que entregan los *features vector* de las imágenes, extraídos por medio de redes neuronales.

#### 3.2. Objetivos Específicos

Los objetivos específicos de la presente propuesta son:

1. Identificar una arquitectura eficiente para generación de imágenes *pap smear* y sus respectivas *semantic mask* y *boundary mask*.
2. Identificar una arquitectura eficiente de discriminadores para la clasificación de imágenes *pap smear* y sus respectivas segmentaciones.
3. Implementar las arquitecturas descritas en el *framework PyTorch* [PGC<sup>+</sup>17] en una unidad de GPU.
4. Generar *features vectors* a partir de las arquitecturas implementadas de imágenes seleccionadas de los datasets a utilizar.
5. Identificar arquitecturas *state of the art* para sus respectivas comparaciones respecto a las arquitecturas diseñadas en la presente propuesta para la extracción de *features*.
6. Identificar un método eficiente para evitar el efecto *curse of dimensionality* sobre los *features vectors* generados.
7. Implementar el algoritmo *HNSW* para la generación de índices para los *features vectors* tanto de las arquitecturas propuestas como de las arquitecturas *state of the art*.
8. Comparar y analizar los resultados entregados por las implementaciones desarrolladas con respecto a otros métodos de *CMBIR* para imágenes *pap smear*.
9. Elaborar un análisis cuantitativo y comparativo de los resultados obtenidos.
10. Implementar la propuesta final en un *framework web* para su disponibilidad académica.



## 4. Metodología y Plan de Trabajo

### 4.1. Plan de Trabajo

El desarrollo del presente trabajo se estructura de acuerdo a la siguiente tabla de actividades:

| Actividad   | Tiempo [semana] |
|---|-----------------|
| Investigación y estudio del estado del arte                     | 2               |
| Implementación de los métodos estado del arte en <i>PyTorch</i> | 2               |
| Identificación de los métodos y arquitecturas a utilizar        | 2               |
| Implementación de los métodos identificados en <i>PyTorch</i>   | 2               |
| Identificación, recopilación y curación de datos                | 2               |
| Experimentación y evaluación de las implementaciones            | 2               |
| Validación y comparación respecto a otros métodos               | 1               |
| Redacción y confección del documento de tesis                   | 3               |
| Correcciones  | 2               |
| Redacción y confección del documento a publicar                 | 2               |
| <b>Total</b>  | 20              |

La proyección de desarrollo del trabajo en meses se proyecta de Marzo a Julio del semestre 2023-1.

## 5. Resultados

### 5.1. Aportes y Resultados Esperados

Los aportes y resultados realizados durante el presente trabajo se centraran en tres ejes: (I) reproducción del proyecto, (II) disponibilidad de datos procesados, (III) documentación *long-term support*.

#### 5.1.1. Reproducción del Proyecto

- El desarrollo del proyecto deberá contar con una contenerización de dos etapas: un contenedor para el modulo de computo el cual contendrá tanto el *enviroment* utilizado para ejecutar las arquitecturas y modelos, como su respectiva disponibilidad para la extracción de los *features* en ambientes de producción, además de un contenedor para acceso *web-based* por medio de una *API REST*.
- La contenerización descrita deberá contar con una integración para aprovisionamiento en la nube para su despliegue en caso de un entorno de producción.

#### 5.1.2. Disponibilidad de Datos Procesados

Tanto los *embeddings* generados por el motor de indexación como los *features vectors* obtenidos de los datos por medio de las redes serán dispuestos en un repositorio.

#### 5.1.3. Documentación *LTS*

- Se dispondrá de un repositorio con la documentación necesaria para replicar los experimentos realizados.
- Se documentará el proceso de puesta en marcha en producción en caso de uso en ambientes clínicos.
- Se versionará el proyecto de tesis realizado y se publicará un escrito en una conferencia o revista científica.

### 5.2. Formas de Validación

La validación de la presente propuesta considerará:

- **Validación del CMBIR:** Al igual que en los trabajos de *Ahmed A. et al.* [AA<sup>+</sup>22] y *Praveena H. et al.* [PGK<sup>+</sup>22], se considerarán las métricas de *precision* y *recall* para evaluar las imágenes retornas por el sistemas recomendador en los distintos casos propuestos.
- **Validación de imágenes generadas:** Se considerará el uso del *dice coefficient* para la evaluación de las mascararas generadas, similar a lo expuesto en el trabajo de *Huang J. et al.* [HYL<sup>+</sup>21], y además el *inception score (IS)* [SGZ<sup>+</sup>16] y el Fréchet inception distance (FID) [HRU<sup>+</sup>17] para las imágenes generadas a partir de la reconstrucción, similar a lo expuesto en el trabajo de *Zhao C. et al.* [ZSM<sup>+</sup>22].
- **Validación clínica:** Se tomará en consideración la evaluación del sistema recomendador en producción por medio de un grupo clínico de especialistas por definir.



## 6. Recursos

### 6.1. Recursos Disponibles

No Aplica

### 6.2. Recursos Solicitados

No Aplica

## Referencias

- [AA<sup>+</sup>22] Ali Ahmed, , Alaa Omran Almagrabi, Ahmed Hamza Osman, and and. Pre-trained convolution neural networks models for content-based medical image retrieval. *International Journal of ADVANCED AND APPLIED SCIENCES*, 9(12):11–24, December 2022. 7, 13
- [ABF18] Martin Aumüller, Erik Bernhardsson, and Alexander Faithfull. Ann-benchmarks: A benchmarking tool for approximate nearest neighbor algorithms, 2018. 9
- [ADJ04] Nicholas Ampazis, George Dounias, and Jan Jantzen. Efficient second order neural network training algorithms for the construction of a pap-smear classifier. volume 3025, 01 2004. 9
- [BN13] Leonid Boytsov and Bilegsaikhan Naidan. Engineering efficient and effective non-metric space library. In *Similarity Search and Applications*, 2013. 9
- [DBI<sup>+</sup>06] Elizabeth Davey, Alexandra Barratt, Les Irwig, Siew F Chan, Petra Macaskill, Patricia Mannes, and A Marion Saville. Effect of study design and quality on unsatisfactory rates, cytology classifications, and accuracy in liquid-based versus conventional cervical cytology: a systematic review. *The Lancet*, 367(9505):122–132, January 2006. 4
- [DBJ<sup>+</sup>06] Georgios D. Dounias, Beth Bjerregaard, Jan Jantzen, Athanasios Tsakonas, Nikolaos Ampazis, Georgia Panagi, and Evangelia Panourgias. Automated identification of cancerous smears using various competitive intelligent techniques. *Oncology reports*, 15 Spec no.:1001–6, 2006. 9
- [DDS<sup>+</sup>09] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009. 7
- [DJLW08] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z. Wang. Image retrieval. *ACM Computing Surveys*, 40(2):1–60, April 2008. 5
- [Doi07] Kunio Doi. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31(4):198–211, 2007. Computer-aided Diagnosis (CAD) and Image-guided Decision Support. 4
- [EAC<sup>+</sup>12] Tarik M. Elsheikh, R. Marshall Austin, David F. Chhieng, Fern S. Miller, Ann T. Moriarty, and Andrew A. Renshaw. American society of cytopathology workload recommendations for automated pap test screening: Developed by the productivity and quality assurance in the era of automated screening task force. *Diagnostic Cytopathology*, 41(2):174–178, February 2012. 4
- [GAKB<sup>+</sup>19] Jevgenij Gamper, Navid Alemi Koohbanani, Ksenija Benet, Ali Khuram, and Nasir Rajpoot. Pannuke: An open pan-cancer histology dataset for nuclei instance segmentation and classification. In Constantino Carlos Reyes-Aldasoro, Andrew Janowczyk, Mitko Veta, Peter Bankhead, and Korsuk Sirinukunwattana, editors, *Digital Pathology*, pages 11–19, Cham, 2019. Springer International Publishing. 7
- [GKCA17] Kasthurirangan Gopalakrishnan, Siddhartha K. Khaitan, Alok Choudhary, and Ankit Agrawal. Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*, 157:322–330, 2017. 7
- [GVR<sup>+</sup>19] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis*, 58:101563, 2019. 7



- [HMBD20] Elima Hussain, Lipi B. Mahanta, Himakshi Borah, and Chandana Ray Das. Liquid based-cytology pap smear dataset for automated multi-class diagnosis of pre-cancerous and cervical cancer lesions. *Data in Brief*, 30:105589, June 2020. [9](#)
- [HRU<sup>+</sup>17] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. 2017. [13](#)
- [HSS18] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7132–7141, 2018. [7](#)
- [HYL<sup>+</sup>21] Jinjie Huang, Guihua Yang, Biao Li, Yongjun He, and Yani Liang. Segmentation of cervical cell images based on generative adversarial networks. *IEEE Access*, 9:115415–115428, 2021. [7](#), [13](#)
- [IMK<sup>+</sup>22] Talha Ilyas, Zubaer Ibna Mannan, Abbas Khan, Sami Azam, Hyongsuk Kim, and Friso De Boer. TSFD-net: Tissue specific feature distillation network for nuclei segmentation and classification. *Neural Networks*, 151:1–15, July 2022. [7](#), [8](#), [9](#)
- [JD06] Jan Jantzen and George Dounias. Analysis of pap-smear image data. In *Proceedings of the Nature-Inspired Smart Information Systems 2nd Annual Symposium*. NiSIS, 2006. Nature-Inspired Smart Information Systems 2nd Annual Symposium ; Conference date: 01-01-2006. [9](#)
- [KKC<sup>+</sup>13] 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(6):1025–1039, July 2013. [5](#)
- [LCB15] Zhi Lu, Gustavo Carneiro, and Andrew P. Bradley. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *IEEE Transactions on Image Processing*, 24(4):1261–1272, 2015. [9](#)
- [LCB<sup>+</sup>17] Zhi Lu, Gustavo Carneiro, Andrew P. Bradley, Daniela Ushizima, Masoud S. Nosrati, Andrea G. C. Bianchi, Claudia M. Carneiro, and Ghassan Hamarneh. Evaluation of three algorithms for the segmentation of overlapping cervical cells. *IEEE Journal of Biomedical and Health Informatics*, 21(2):441–450, March 2017. [9](#)
- [MDJ09] Yannis Marinakis, Georgios Dounias, and Jan Jantzen. Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification. *Computers in Biology and Medicine*, 39(1):69–78, January 2009. [9](#)
- [MGRMG15] Maritza Mera-Gaona, Sandra Roa Martínez, and Carolina González. Content-based image retrieval system to support the diagnosis of human papillomavirus. *Health and Technology*, 5, 08 2015. [7](#)
- [MKCC<sup>+</sup>10] Henning Müller, Jayashree Kalpathy-Cramer, Barbara Caputo, Tanveer Syeda-Mahmood, and Fei Wang. Overview of the first workshop on medical content-based retrieval for clinical decision support at miccai 2009. In Barbara Caputo, Henning Müller, Tanveer Syeda-Mahmood, James S. Duncan, Fei Wang, and Jayashree Kalpathy-Cramer, editors, *Medical Content-Based Retrieval for Clinical Decision Support*, pages 1–17, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. [4](#)
- [MMBG04] Henning Müller, Nicolas Michoux, David Bandon, and Antoine Geissbuhler. A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. *International Journal of Medical Informatics*, 73(1):1–23, February 2004. [4](#)
-

- [MY16] Yu A. Malkov and Dmitry A. Yashunin. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42:824–836, 2016. 9
- [MZCH20] Baoyan Ma, Jian Zhang, Feng Cao, and Yongjun He. Macd r-cnn: An abnormal cell nucleus detection method. *IEEE Access*, 8:166658–166669, 2020. 4
- [NSM18] Swati Nigam, Rajiv Singh, and A. K. Misra. Efficient facial expression recognition using histogram of oriented gradients in wavelet domain. *Multimedia Tools and Applications*, 77(21):28725–28747, May 2018. 7
- [Pap73] George N. Papanicolaou. New cancer diagnosis. *CA: A Cancer Journal for Clinicians*, 23(3):174–179, 1973. 4
- [PDS<sup>+</sup>18] Marina E. Plissiti, P. Dimitrakopoulos, G. Sfikas, Christophoros Nikou, O. Krikoni, and A. Charchanti. Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3144–3148, 2018. 9
- [PGC<sup>+</sup>17] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017. 11
- [PGK<sup>+</sup>22] Hiral Dwaraka Praveena, Nirmala S. Guptha, Afsaneh Kazemzadeh, B. D. Parameshachari, and K. L. Hemalatha. Effective CBMIR system using hybrid features-based independent condensed nearest neighbor model. *Journal of Healthcare Engineering*, 2022:1–9, March 2022. 7, 13
- [PSR18] TV Priya, G Vinitha Sanchez, and NR Raaan. Facial recognition system using local binary patterns (lbp). *International Journal of Pure and Applied Mathematics*, 119(15):1895–1899, 2018. 7
- [PT41] George N. Papanicolaou and Herbert F. Traut. The diagnostic value of vaginal smears in carcinoma of the uterus\*\*this study has been aided by the commonwealth fund. presented before the new york obstetrical society, march 11, 1941. *American Journal of Obstetrics and Gynecology*, 42(2):193–206, 1941. 4
- [RLW<sup>+</sup>20] Md Mamunur Rahaman, Chen Li, Xiangchen Wu, Yudong Yao, Zhijie Hu, Tao Jiang, Xiaoyan Li, and Shouliang Qi. A survey for cervical cytopathology image analysis using deep learning. *IEEE Access*, 8:61687–61710, 2020. 4, 6
- [SGZ<sup>+</sup>16] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, and Xi Chen. Improved techniques for training gans. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. 13
- [SHZ<sup>+</sup>18] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. 2018. 7
- [SRV21] Ana Filipa Sampaio, Luís Rosado, and Maria João M. Vasconcelos. Towards the mobile detection of cervical lesions: A region-based approach for the analysis of microscopic images. *IEEE Access*, 9:152188–152205, 2021. 4
- [SSL<sup>+</sup>12] Debbie Saslow, Diane Solomon, Herschel W. Lawson, Maureen Killackey, Shalini L. Kulasingham, Joanna Cain, Francisco A. R. Garcia, Ann T. Moriarty, Alan G. Waxman, David C. Wilbur, Nicolas Wentzensen, Levi S. Downs, Mark Spitzer, Anna-Barbara Moscicki, Eduardo L. Franco, Mark H. Stoler, Mark Schiffman, Philip E. Castle, and Evan R. Myers and. American cancer society, american society for colposcopy and cervical pathology, and american society for clinical pathology screening guidelines for the

- prevention and early detection of cervical cancer. *CA: A Cancer Journal for Clinicians*, 62(3):147–172, March 2012. 4
- [Sub18] Subsecretaría de Salud Pública, Ministerio de Salud. Plan nacional del cáncer 2018-2028. [https://cdn.digital.gob.cl/filer\\_public/d3/0a/d30a1f5e-53d9-4a31-a4fe-e90d8d9a2348/documento\\_plan\\_nacional\\_de\\_cancer.pdf](https://cdn.digital.gob.cl/filer_public/d3/0a/d30a1f5e-53d9-4a31-a4fe-e90d8d9a2348/documento_plan_nacional_de_cancer.pdf), 2018. 4
- [SWS<sup>+</sup>00] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000. 4
- [VLBM08] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning - ICML '08*. ACM Press, 2008. 7
- [WCH<sup>+</sup>18] Khin Yadanar Win, Somsak Choomchuay, Kazuhiko Hamamoto, Manasanan Raveesunthornkiat, Likit Rangsirattanakul, and Suriya Pongsawat. Computer aided diagnosis system for detection of cancer cells on cytological pleural effusion images. *BioMed Research International*, 2018:1–21, November 2018. 5
- [WHO20] WHO. World cancer report: Cancer research for cancer prevention. <https://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports/World-Cancer-Report-Cancer-Research-For-Cancer-Prevention-2020>, 2020. 4
- [XLA<sup>+</sup>08] Zhiyun Xue, L. Long, Sameer Antani, Jose Jeronimo, and George Thoma. A web-accessible content-based cervicographic image retrieval system. *Proceedings of SPIE - The International Society for Optical Engineering*, 6919, 04 2008. 4, 7
- [ZSM<sup>+</sup>22] Chen Zhao, Renjun Shuai, Li Ma, Wenjia Liu, and Menglin Wu. Improving cervical cancer classification with imbalanced datasets combining taming transformers with t2t-ViT. *Multimedia Tools and Applications*, 81(17):24265–24300, March 2022. 4, 13
- [ZZC<sup>+</sup>17] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017. 7
-