



Revista EIA
ISSN 1794-1237
e-ISSN 2463-0950
Año XVIII/ Volumen 18/ Edición N.35
Enero-Junio de 2021
Reia35008 pp. 1-12

Publicación científica semestral
Universidad EIA, Envigado, Colombia

**PARA CITAR ESTE ARTÍCULO /
TO REFERENCE THIS ARTICLE /**

Bravo Ortíz, M.A.; Arteaga Arteaga,
H.B.; Tabares Soto, R.; Padilla
Buriticá, J.I.; Orozco-Arias, S.
(2021). Clasificación de cáncer
cervical usando redes neuronales
convolucionales, transferencia de
aprendizaje y aumento de datos
Revista EIA, 18(35), Reia35008.
pp. 1-12.
<https://doi.org/10.24050/reia.v18i35.1462>

 *Autor de correspondencia:*

Tabares Soto, R. (Reinel):
Docente e investigador de la
Universidad Autónoma de Manizales
Correo electrónico:
rtabares@autonoma.edu.co

Recibido: 14-08-2020
Aceptado: 28-12-2020
Disponible online: 06-01-2021

Clasificación de cáncer cervical usando redes neuronales convolucionales, transferencia de aprendizaje y aumento de datos

MARIO ALEJANDRO BRAVO ORTÍZ¹

HAROLD BRAYAN ARTEAGA ARTEAGA¹

 REINEL TABARES SOTO¹

JORGE IVÁN PADILLA BURITICÁ¹

SIMÓN OROZCO-ARIAS¹

1. Universidad Autónoma de Manizales

Resumen

El cáncer cervical se forma en las células que revisten el cuello uterino y la parte inferior del útero. Debido a razones de costo y baja oferta de servicios destinados a la detección de este tipo de cáncer, muchas mujeres no tienen acceso a un diagnóstico pronto y preciso, ocasionando un inicio tardío del tratamiento. Para dar solución a este problema se implementó una metodología que clasifica de manera automática el tipo de cáncer cervical, entre leve (Tipo 1 y 2) y agresivo (Tipo 3), utilizando técnicas de procesamiento digital de imágenes y aprendizaje profundo. Se trabajó en la construcción de un modelo computacional con base en redes neuronales convolucionales, transferencia de aprendizaje y aumento de datos, obteniendo precisiones de clasificación de hasta **97,35%** sobre los datos de validación, asegurando la confiabilidad de los resultados. Con este trabajo se demostró que el diseño propuesto puede ser usado como un complemento para mejorar la eficiencia de las herramientas del diagnóstico asistido del cáncer.

Palabras Claves: aumento de datos, cáncer cervical, redes neuronales convolucionales, transferencia de aprendizaje.

Cervical cancer classification using convolutional neural networks, transfer learning and data augmentation

Abstract

Cervical cancer is formed in the cells that line the cervix and the lower part of uterus. Due to the cost and low reasons and low supply of services for the detection of this type of cancer many women do not have access to an early and accurate diagnosis. With the purpose of solving this issue it was created a certain method that helps us to automatically classify the different types of cervical cancer, such as mild type 1 and 2, and aggressive (type 3), using digital image processing techniques and deep learning. We have built a computational model based on convolutional neural networks, transfer learning and data increase, which help us obtain a classification accuracy up to 97.35% on the validation data, thus, we can ensure the reliability of the results. With this work it was demonstrated that the proposed design can be used as a complement to improve the tools of the assisted diagnosis of cancer.

Key Words: data augmentation, cervical cancer, convolutional neural networks, transfer learning.

I. Introducción

El cáncer cervical es la cuarta enfermedad con más muertes a nivel mundial [1] generando 528.000 casos al año, según se reportó en el 2018 [2], de los cuales 32.000 son en América latina[3]. En la actualidad, Este problema se enfrenta mediante intervenciones quirúrgicas, vacunas como la del virus del papiloma humano (VPH) y un chequeo frecuente mediante citologías[4], pero pese a estos esfuerzos la tasa de mortalidad sigue siendo muy alta, debido a que la identificación y clasificación del cáncer está dada por la experticia de ginecólogos que normalmente se encuentran en las grandes ciudades generando baja cobertura a nivel nacional en los países en vía de desarrollo[5]. Por lo tanto, se busca demostrar que las redes neuronales convolucionales (CNN por sus siglas en inglés) y la transferencia de aprendizaje (TL por sus siglas en inglés), son una buena alternativa de clasificación para esta patología. Las redes neuronales profundas han sido muy utilizadas en esta problemática, Akshaya y otros[2], tienen como objetivo crear una herramienta computacional que clasifique el cáncer cervical en leve y agresivo. Ellos proponen una CNN para clasificación de imágenes colposcópicas y obtienen precisiones entre el 92 - 96% sobre el conjunto de imágenes de validación. El problema principal de esta investigación es el número de imágenes, cuentan con 100 imágenes por cada nivel de cáncer cervical. Kaur y otros [5], proponen la detección de los tres tipos de cáncer cervical por medio de TL con las CNNs (VGG16/19), ver sección II. La base de datos que ellos utilizan se encuentra en [6]. El problema para Kaur y otros es la similitud entre el cáncer cervical tipo 1 y tipo 2 con lo cual obtienen precisión solo del 63% sobre el conjunto de validación. Estas dos investigaciones brindan el punto de partida y el objetivo de esta investigación. Basados en el trabajo de Akshaya y otros[2], se decidió separar la base de datos obtenida de [6] en cáncer cervical leve y agresivo. El objetivo de la investigación es demostrar que las CNNs, TL y aumento de datos son una buena alternativa para la clasificación del cáncer de cuello uterino entre leve y agresivo [5].

El algoritmo implementado utiliza las CNNs VGG16 [7] y VGG19 [8] previamente entrenadas con ImageNet[9] permitiendo disponer de un conjunto amplio de características que mejoran el aprendizaje de nuevos patrones en la base de datos objetivo[6]. La base de datos fue pre-procesada utilizando una segmentación semiautomática, aumento de datos y etiquetada con el tipo de cáncer (leve-agresivo). Para el TL se utilizó las capas convolucionales de las CNNs (VGG16/19) para re-entrenar la red y permitir que se adapte a la clasificación del tipo de cáncer. Los resultados obtenidos en la investigación muestran precisiones hasta del 97,35% sobre los datos de validación. El artículo está compuesto por las siguientes secciones: Sección I introducción. Sección II materiales y métodos, se define la metodología y los materiales. Sección III Resultados y discusión, Se describe la investigación desarrollada y resultados obtenidos. Sección IV Conclusiones.

II. Materiales y métodos

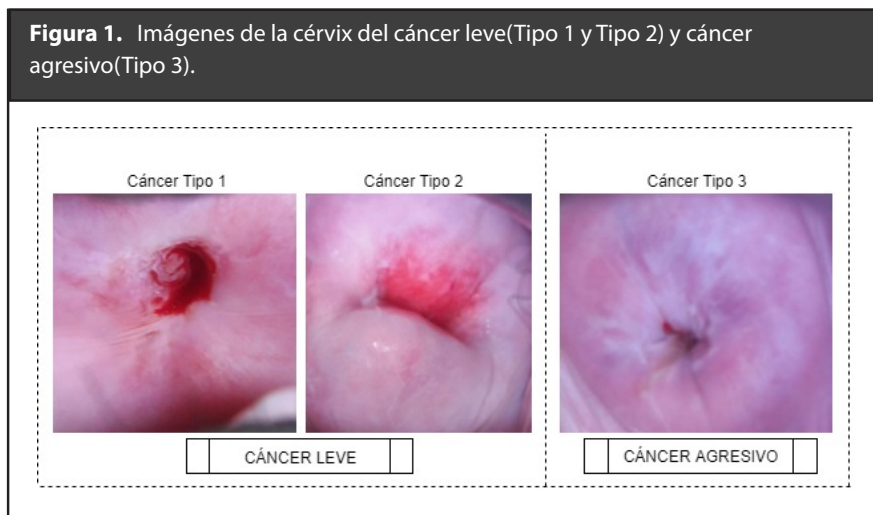
El presente estudio está basado en TL en el cual se utilizaron las CNNs VGG16[7] y la VGG19[8] que fueron construidas para clasificar 1.000 diferentes categorías de imágenes. Con el fin de lograr el objetivo propuesto en esta investigación, se desarrolló la siguiente metodología:

Paso 1: Revisión bibliográfica, el objetivo de esta revisión fue encontrar una base de datos [6] (ver sección II base de datos) que estuviera compuesta de imágenes del cáncer cervical debidamente etiquetadas, además de revisar los últimos experimentos y/o avances realizados para la detección del cáncer cervical[10][11][12].

Paso 2: Segmentación del área afectada[13], con el fin de trabajar con una base de datos óptima, que permita un mejor aprendizaje por la CNN, se realizó un proceso de segmentación semiautomático, el cual inicialmente hacia una selección automática de la sección que tiene el cáncer. Luego se llevó a cabo una revisión detallada de todas las imágenes generadas para así hacerles los ajustes finales y/o selección y preprocesamiento manual.

Paso 3: Separación de los datos de entrenamiento y los datos de validación.

Paso 4: Aumento de datos[14][15], después de realizar una evaluación de la literatura y la base de datos, con la separación del cáncer cervical (Tipo 1 y Tipo 2) como leve y (Tipo 3) como agresivo ver Fig. 1. El número de imágenes para realizar el modelo en leve es el doble del fuerte, para resolver este desbalance se realiza un aumento de datos con rotaciones aleatorias(máximo 20°), desplazamientos tanto horizontales como verticales(máximo 20%), zoom(máximo 15%), recortes(máximo 15%), completación con pixeles vecinos de ser necesario, y espejos horizontales.



Paso 5: Diseño, implementación y entrenamiento del modelo de aprendizaje supervisado, se utiliza TL como método de inicialización de los filtros de las capas convolucionales (extracción de características) los cuales no son entrenables. Se diseña e implementa una nueva arquitectura para las capas densas (clasificación) los cuales son entrenables. Se realiza el entrenamiento de la CNN a partir de imágenes etiquetadas con el tipo de cáncer.

Paso 6: Clasificación del cáncer cervical entre tipo leve y agresivo sobre los datos de validación debidamente etiquetados.

A continuación, se procede a describir las herramientas usadas y/o configuraciones en la realización del proyecto de investigación:

A. Base de datos

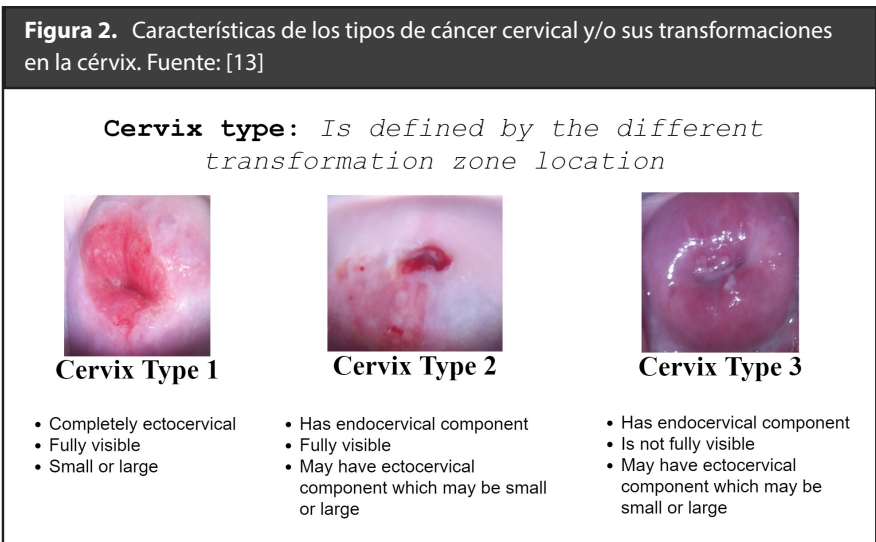
La base de datos fue obtenida de Intel y kaggle de la competencia Intel & Mobile ODT Cervical Cancer Screening [16], la cual proporciona dos bases de datos (principal y adicional). La base de datos principal y adicional contienen imágenes en formato JPG con tamaños de 3.096 x 4.128 y 2.448 x 3.264. En la base de datos se identifican las diferentes zonas de transformación, donde se localizan las lesiones cervicales, Estas son representadas en tres diferentes tipos de lesiones.

La base de datos principal tiene 1.993 y la adicional 10.240 imágenes y se dividen en entrenamiento (tipo 1, tipo 2 y tipo 3) y validación ver tabla 1. De estas bases de datos solamente fueron utilizadas 1.000 imágenes por cada tipo con el fin de disminuir el ruido [17] que se le ingresa a la red neuronal debido a la baja calidad de las imágenes.

Tabla 1. Base de datos Intel y kaggle [13].

Base de datos	Entrenamiento			Validación
	Tipo 1	Tipo 2	Tipo 3	
Principal	250	781	450	512
Adicional	1.191	3.567	1.976	3.506

Después de realizar pruebas de clasificación con cada tipo de cáncer cervical ver Fig. 1, se concluyó que el Tipo 1 y Tipo 2 eran muy similares, debido a que éste se encuentra en la ectocérvix (parte externa de la cérvix) y la Tipo 3 se encuentra en la endocérvix (parte interna de la cérvix) [18], por lo tanto se unificaron los Tipo 1 y Tipo 2 en una clase llamada leve debido a sus semejanzas (Fig. 2) y por ser menos agresivos que el Tipo 3 (llamado en este trabajo como tipo agresivo).



B. Aumento de datos.

Es una técnica utilizada para balancear el número de datos, ayudar a la generalización del modelo y disminuir el sobre ajuste [19]. El aumento de datos se desarrolló usando la librería de Keras[20], la cual permite realizar distintas operaciones sobre una imagen como: rotaciones, desplazamientos, zoom, recortes, completación con pixeles vecinos, y espejos horizontales ver Fig. 3.



C. Configuración del software y hardware

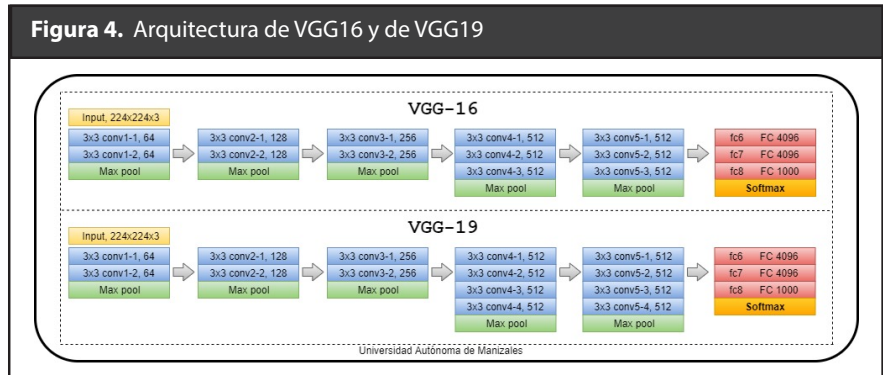
Los experimentos realizados se hicieron en un computador con 64 GB de RAM, 1.2 TB de disco duro, tarjeta gráfica NVIDIA Geforce RTX 2080 Ti 12G y un procesador Intel Core i9-9900k de 8 núcleos y 16 hilos. El sistema operativo fue Ubuntu 18.04 con Python 3.7, Keras 2.2.4[20] y Tensor-Flow 1.14[21]

D. VGG-16

Es una CNN propuesta en la universidad de Oxford por K. Simonyan y A. Zisserman [7]. Es una de las redes más utilizadas para clasificar imágenes, esta arquitectura tiene la capacidad de clasificar 1.000 tipos de imágenes. La VGG16 es una mejora de Alex-net [22], la cual implementa el primer filtro de 11x11 y el segundo filtro de 5x5 mientras la VGG16 implementa múltiples filtros de 3x3. Esta última red es pre-entrenada en ImageNet [9] con un conjunto de datos de más de 15 millones de imágenes que pertenecen a 22.000 categorías. la CNN cuenta con 16 capas, de las cuales 13 son filtros y 3 densas completamente conectadas. El tamaño de las imágenes de entrada para esta red neuronal es de 224 x 224x3 ver Fig. 4.

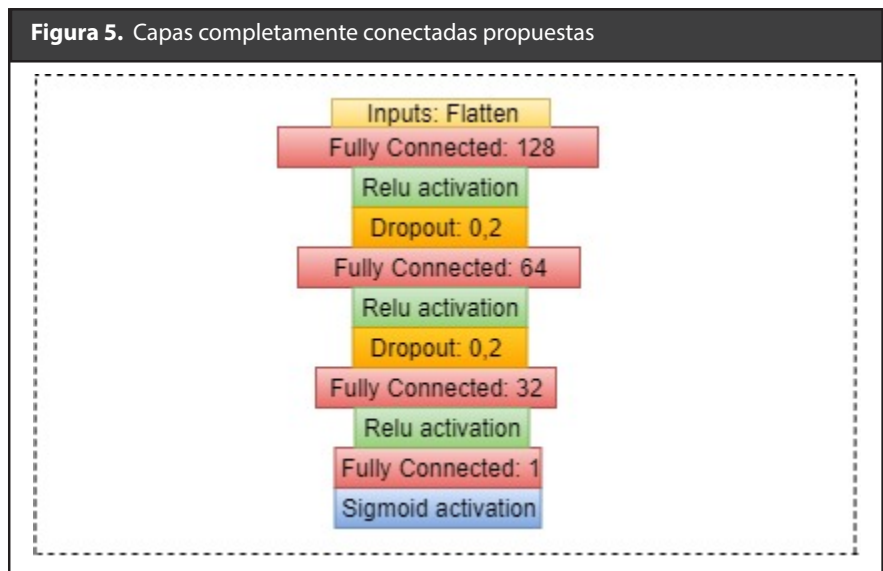
E. VGG-19

La VGG19 es una red neuronal convolucional de la familia VGG16 y la diferencia global es que esta red cuenta con 19 capas [8], de las cuales 16 capas son filtros convolucionales y 3 densas completamente conectadas ver Fig. 4. El tamaño de las imágenes de entrada de esta CNN son las mismas de la VGG16.



F. Fully connected

Las VGG16/19 se adaptaron a la clasificación del cáncer cervical leve y agresivo, reemplazando las 3 capas densas por una arquitectura propia de capas completamente conectadas ver Fig. 5. la cual está compuesta por 4 capas densas de 128 neuronas, 64 neuronas, 32 neuronas y una neurona. En las tres primeras capas densas la función de activación es Relu [23], en la última es Sigmoide [24]. La función de pérdida que se utilizó es binary crossentropy [25] y el optimizador es stochastic gradient descent (SGD) [26], con learning rate (lr) 0,001 y un momentum de 0,9, para disminuir el sobre ajuste [27] se aplicó dropout [28] de 0,2 entre la primera y segunda capa y entre la segunda y la tercera capa.



III. Resultados y discusión

Se propuso una arquitectura de CNNs que utiliza TL y aumento de datos, con una etapa de pre-procesamiento para permitir la convergencia de la red, esta arquitectura se puede observar en Fig. 6 y se explica a continuación:

Etapa de preprocesamiento

Segmentación semiautomática y aumento de datos: se diseñó un algoritmo para realizar segmentación el cual está compuesto por distintas operaciones sobre la imagen, con el fin de enfocar la información en la lesión donde está el cáncer. Debido al ruido y la calidad de las imágenes de la base de datos [6], el algoritmo para segmentar no era suficiente por lo tanto se decidió realizar como segundo paso, la segmentación manual de las imágenes que no fueron segmentadas por el algoritmo. Con el fin de no ingresar ruido a la arquitectura propuesta (Fig. 6), se seleccionaron las mejores 1.000 imágenes por cada tipo de cáncer cervical. El cáncer cervical leve cuenta con 2.000 imágenes y el agresivo con 1.000, para que la base de datos quede balanceada se utilizó aumento de datos para equilibrar el número de imágenes (ver sección II, aumento de datos).

División del conjunto de datos de entrenamiento y validación

Después de obtener la base de datos con cáncer cervical leve y agresivo, se procede a separar la base de datos en 80% para entrenamiento y 20 % de las imágenes para validación.

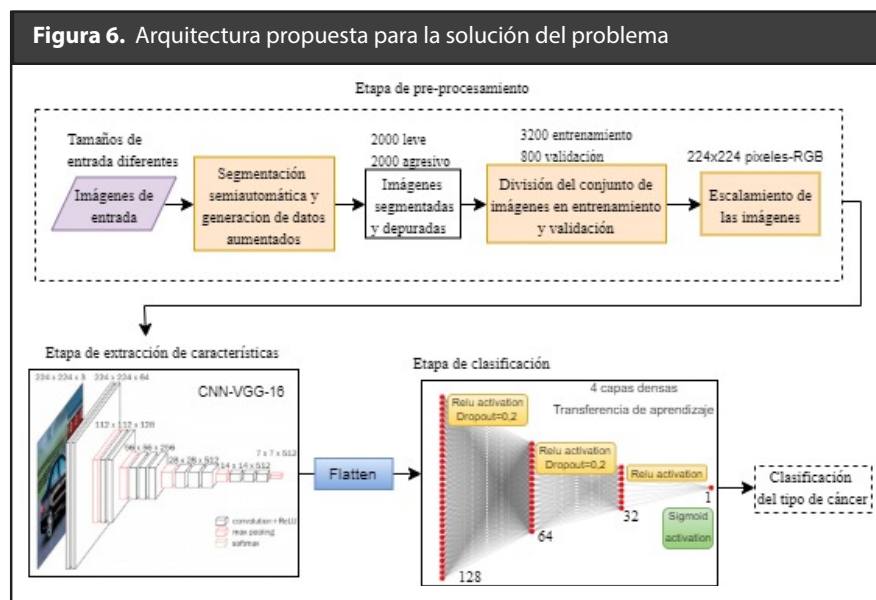
Escalamiento de las imágenes: La VGG16/19 tiene un tamaño de imagen definido de 224x224x3 [7][8], por lo tanto, las imágenes de entrenamiento y validación son escaladas utilizando el método de interpolación bicúbica.

Etapa de extracción de características

Para extraer características se utilizaron CNN pre-entrenadas (ver sección II, VGG16 y VGG19).

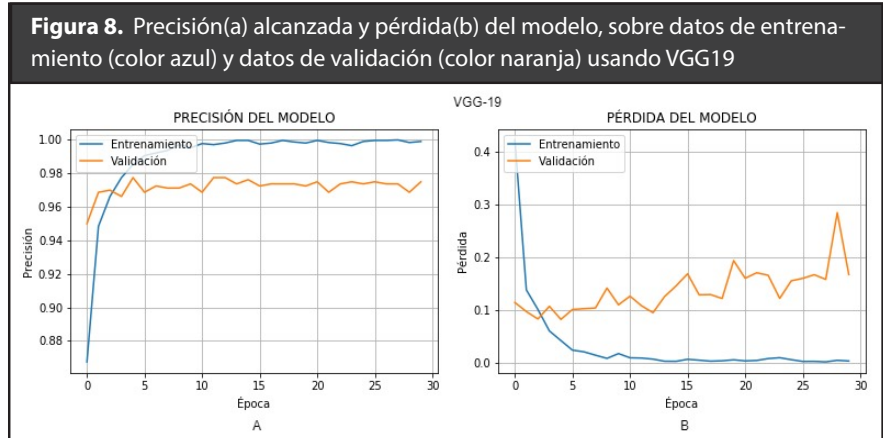
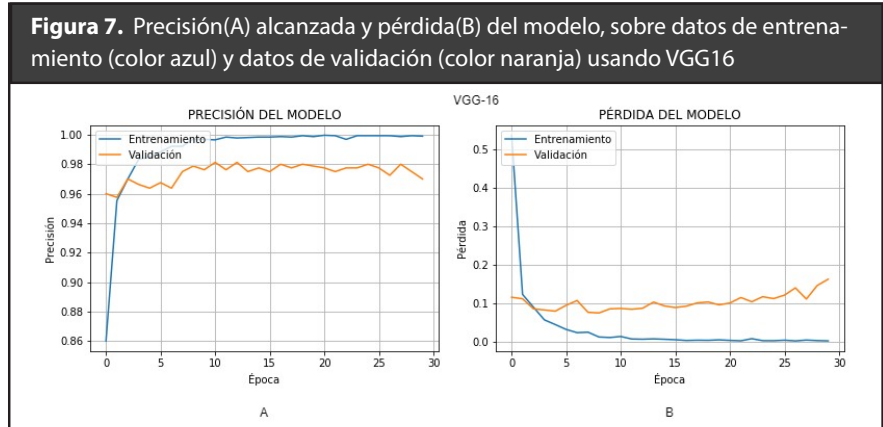
Etapa de clasificación

Para realizar la etapa de clasificación se eliminan las tres capas densas de las CNN VGG16/19 y se agrega 4 capas densas propias (ver Fig.5) con el fin de adaptarlas a esta investigación, en la salida de las redes se vectorizan con un flatten, luego se adapta la capa completamente conectada (ver sección II, Fully connected)(Fig. 5)



Para validar el proceso de aprendizaje, se ha graficado la pérdida por épocas. En Fig. 8b, se observa un crecimiento de la pérdida a medida que aumenta el número de épocas lo que significa que existe sobre ajuste, lo que no ocurre según se observa en Fig. 7b, la pérdida se mantiene estable a medida que aumenta el número de épocas, por lo tanto, no existe sobre ajuste.

Además de la pérdida, se ha graficado la precisión para las imágenes de entrenamiento y validación. En las figuras Fig. 7a y Fig. 8a se observan las curvas de precisión en entrenamiento y validación tanto para VGG16/19 respectivamente, las curvas de precisión en validación aumentan hasta la época 10 aproximadamente, y seguido a esto se estabilizan logrando una precisión en validación de 0,9735 para VGG16 y 0,9710 con VGG19.



La precisión en validación para la VGG16 es mayor que la de la VGG19, como se muestra en las figuras Fig. 7a y Fig. 8a, debido a que estas hacen parte de la misma familia [7][8], la diferencia significativa es que la VGG16 mantiene una pérdida estable (Fig. 7b).

IV. Conclusión

La arquitectura propuesta alcanza precisiones hasta del **97,35%** en la clasificación de cáncer cervical (leve y agresivo) utilizando segmentación semiautomática, aumento de datos y TL a partir de la VGG16/19. La diferencia en los resultados entre estas dos CNNs es el sobre ajuste. La VGG16 obtiene 97,35% de precisión en validación con pérdida estable ver Fig. 7, mientras la VGG19 obtiene 97,10% de precisión en validación con pérdida creciente ver Fig. 8. Teniendo en cuenta lo anterior, se observó que, haciendo una correcta segmentación de las imágenes, un buen balance de clases (aumentó de datos) y TL sobre la extracción de características se obtienen porcentajes que ayudan a tener un diagnóstico más preciso en la clasificación del cáncer cervical.

Trabajo a futuro:

Se desea crear una herramienta computacional de libre acceso que permita clasificar los tipos de cáncer cervical (tipo 1, tipo 2 y tipo 3 independientemente) para ayudar en la precisión del diagnóstico generado por los ginecólogos.

Liberar un repositorio de imágenes segmentadas y etiquetadas con los 3 tipos de cáncer para generar una línea base de conocimiento en este tipo de problemas.

Agradecimientos:

Los investigadores agradecen a la Universidad Autónoma de Manizales por brindar las herramientas computacionales y el presupuesto para desarrollar esta investigación enmarcado en el proyecto con código 589-089.

Agradecimiento al semillero de investigación en bioinformática e inteligencia artificial, el cual proporcionó los conocimientos necesarios para obtener los resultados de esta publicación.

Referencias

- McGuire S. World cancer report 2014. Geneva, Switzerland: World Health Organization, international agency for research on cancer, WHO Press, 2015. *Advances in Nutrition: An International Review Journal*, 7, 418-419, 2016.
- Akshaya R., Manie R., Monisha B., Ranichadra V. Convolutional Neural Networks Aiding Colposcopy Image Classification. *International Journal of Trend in Research and development*, 5, 270-274, 2018.
- Almonte M., Sánchez G.I., Jerónimo J., Ferrección C., Lazcano E., Herrera R. Nuevos Paradigmas en la Prevención y Control de Cáncer de Cuello Uterino en América Latina. *Salud Pública de México*, 52, No 6, 2010.
- Lorena M., Villate S., Jiménez D., Conduct in regard to the papanicolaou test: The voice of the patients in face of abnormal growth in the cervix, *Revista Colombiana de Enfermería*, Vol. 18, páginas 1-13, 2019
- Kaur N., Panagrahi N., Mittal A. Automated Cervical Cancer Screening Using Transfer Learning. *International Journal Of Advanced Research in Science and Engineering*, 6, 2110-2119, 2017.
- Intel & MobileODT, Cervical Cancer Screening, 2017, [Online]. Available: <https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening/data>

- Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large Scale Image Recognition. Published as a conference paper at ICLR 2015. San Diego, California, Estados Unidos, abril, 2015.
- Park Chansung, Transfer Learning in Tensorflow (VGG19 on CIFAR-10): Part 1, 2018, 10 Octubre 2019, [Online]. Available: <https://towardsdatascience.com/transfer-learning-in-tensorflow-9e4f7eae3bb4>
- Stanford University, Princeton University, ImageNet, 2016, 10 Octubre 2019, [Online]. Available: <http://www.image-net.org/>
- Zhang XQ, Zhao S-G, Cervical image classification based on image segmentation preprocessing and a CapsNet network model, Wiley, páginas 19-28, 2019, [Online]. Available: <https://doi.org/10.1002/ima.22291>
- Fernandes K., Cardoso J., Fernandes J., Automated Methods for the Decision Support of Cervical Cancer Screening Using Digital Colposcopies, IEEE Xplore, Vol. 6, páginas 33910-33927, 2018, [Online]. Available: <https://ieeexplore.ieee.org/document/8362613>
- Vasudha, Mittal A., Juneja M., Cervix Cancer Classification using Colposcopy Images by Deep Learning Method, IJETS, Vol. 5, páginas 426-432, 2018, [Online]. Available: <https://pdfs.semanticscholar.org/f099/0cd17037129f7a55fcdf279ea6e9d613e8fe.pdf>
- Caraiman S., Vasile I., Histogram-based segmentation of quantum images, ELSEVIER, Vol. 529, páginas 46-60, 2014, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0304397513005835>
- Adrian Rosebrock, pyimagesearch, Keras ImageDataGenerator and Data Augmentation (Julio 8, 2019), consultado por última vez el 10 de octubre del 2019 en: https://www.pyimagesearch.com/2019/07/08/keras-imagedatagenerator-and-data-augmentation/?utm_source=facebook&utm_medium=ad-08-07-2019&utm_campaign=8+July+2019+BP++Traffic&utm_content=Default+name++Traffic&fbid_campaign=6116019415846&fbid_adset=6116019416246&utm_adset=1+July+2019+BP++All+Visitors+90+Days++Worldwide++18%2B&fbid_ad=6116019417246
- Mikolajczyk A. Grochowski M, Data augmentation for improving deep learning in image classification problem, IEEE Xplore, Poland, 2018, 21 Junio 2018, [Online]. Available: <https://ieeexplore.ieee.org/document/8388338>
- Intel & MobileODT, Cervical Cancer Screening, 2017, [Online]. Available: <https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening>
- Tiago S. Nazar'e, Gabriel B. Paranhos da Costa, Welinton A. Contato, and Moacir Ponti, Deep Convolutional Neural Networks and Noisy Images, ResearchGate, paginas 416-424, 2018, [Online]. Available: https://www.researchgate.net/publication/322915518_Deep_Convolutional_Neural_Networks_and_Noisy_Images
- Nawal M. Nour, Cervical Cancer: A Preventable Death, Obstet Gynecol, Vol. 2, páginas 240-244, 2009, [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2812875/>
- Ayan E, H. Muray Ü, Data augmentation importance for classification of skin lesions via deep learning, IEEE Xplore, páginas 1-5, 2019, [Online]. Available: <https://ieeexplore.ieee.org/document/8391469/citations?tabFilter=papers#citations>
- Keras Documentation, Keras, [Online]. Available: <https://keras.io/why-use-keras/>
- TensorFlow Core r1.14, Tensorflow, [Online]. Available: https://www.tensorflow.org/versions/r1.14/api_docs/python/tf
- Krizhevsky A., Sutskever I., Hinton G.E. ImageNet Classification with Deep Convolutional Neural Networks. In NIPS, 2012.
- Abien Fred M. Agarap, Cornell University, Deep Learning using Rectified Linear Units (ReLU), 2019, 7 febrero 2019, [Online]. Available: <https://arxiv.org/abs/1803.08375>
- Sridhar Narayan, The generalized sigmoid activation function: Competitive supervised learning, ScienceDirect, Vol. 99, páginas 69-82, 1997, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0020025596002009>

- Daniel Godoy, Towards Data Science, Understanding binary cross-entropy / log loss: a visual explanation, 2018, 10 octubre 2019, [Online]. Available: <https://towardsdatascience.com/understanding-binary-cross-entropy-log-loss-a-visual-explanation-a3ac6025181a>
- Zhang S., Choromanska A., and LeCun Y. Deep learning with Elastic Averaging SGD. Neural Information Processing Systems Conference (NIPS 2015), Vol. 28, páginas 1–24, 2015, [Online]. Available: <https://papers.nips.cc/paper/5761-deep-learning-with-elastic-averaging-sgd>
- Piotr Skalski, Towards Data Science, Preventing Deep Neural Network from Overfitting, 2018, 10 Octubre 2019, [Online]. Available: <https://towardsdatascience.com/preventing-deep-neural-network-from-overfitting-953458db800a>
- Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R., Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research, Vol. 15, páginas. 1929-1958, 2014, [Online]. Available: <http://jmlr.org/papers/v15/srivastava14a.html>