Retinal Lesion Segmentation Using Transfer Learning with an Encoder-Decoder CNN

Segmentación de Lesiones en la Retina Usando Transferencia de Conocimiento con CNN Encoder-Decoder

Rafael Ortiz-Feregrino ⓘ, *Saúl Tovar-Arriaga* ⓘ ✉, *Jesús Carlos Pedraza-Ortega* ⓘ, *Andras Takacs* ⓘ

Universidad Autónoma de Querétaro

ABSTRACT

Deep learning (DL) techniques achieve high performance in the detection of illnesses in retina images, but the majority of models are trained with different databases for solving one specific task. Consequently, there are currently no solutions that can be used for the detection/segmentation of a variety of illnesses in the retina in a single model. This research uses Transfer Learning (TL) to take advantage of previous knowledge generated during model training of illness detection to segment lesions with encoder-decoder Convolutional Neural Networks (CNN), where the encoders are classical models like VGG-16 and ResNet50 or variants with attention modules. This shows that it is possible to use a general methodology using a single fundus image database for the detection/segmentation of a variety of retinal diseases achieving state-of-the-art results. This model could be in practice more valuable since it can be trained with a more realistic database containing a broad spectrum of diseases to detect/segment illnesses without sacrificing performance. TL can help achieve fast convergence if the samples in the main task (Classification) and sub-tasks (Segmentation) are similar. If this requirement is not fulfilled, the parameters start from scratch.

KEYWORDS: Transfer learning, Encoder-decoder, Retinal images, Lesion segmentation, Deep learning

RESUMEN

Las técnicas de *Deep Learning* (DL) han demostrado un buen desempeño en la detección de anomalías en imágenes de retina, pero la mayoría de los modelos son entrenados en diferentes bases de datos para resolver una tarea en específico. Como consecuencia, actualmente no se cuenta con modelos que se puedan usar para la detección/segmentación de varias lesiones o anomalías con un solo modelo. En este artículo, se utiliza *Transfer Learning* (TL) con la cual se aprovecha el conocimiento adquirido para determinar si una imagen de retina tiene o no una lesión. Con este conocimiento se segmenta la imagen utilizando una red neuronal convolucional (CNN), donde los *encoders* o extractores de características son modelos clásicos como VGG-16 y ResNet50 o variantes con módulos de atención. Se demuestra así, que es posible utilizar una metodología general con bases de datos de retina para la detección/segmentación de lesiones en la retina alcanzando resultados como los que se muestran en el estado del arte. Este modelo puede ser entrenado con bases de datos más reales que contengan una gama de enfermedades para detectar/segmentar sin sacrificar rendimiento. TL puede ayudar a conseguir una convergencia rápida del modelo si la base de datos principal (Clasificación) se parece a la base de datos de las tareas secundarias (Segmentación), si esto no se cumple los parámetros básicamente comienzan a ajustarse desde cero.

Corresponding author

TO: **Saúl Tovar-Arriaga**

INSTITUTION: **Universidad Autónoma de Querétaro**

ADDRESS: **Cerro de las Campanas S/N, Col. Las Campanas, Centro, C. P. 76010, Santiago de Querétaro, Querétaro, México**

CORREO ELECTRÓNICO: **saul.tovar@uaq.mx**

INTRODUCTION

The retina plays an essential role in vision since it transforms the received optical signals into electrical and transfers them to the brain. It provides a clear window to blood vessels and other essential parts of the neural tissue[1]. Since it is an extension of the brain, it could also indicate possible mental health conditions [1].

Ophthalmologists diagnose retinal diseases by identifying specific signs on retinal images. Such signs could be related to diseases such as diabetes, diabetic retinopathy (DR), macular degeneration, glaucoma, and cardiovascular problems. Some of them are progressive and asymptomatic until advanced states [2]. Retinal image analysis turned into an essential matter in the medical area. Many kinds of research are being published to provide predictive information about many diseases even before clinical eye disease becomes detectable [3] [4].

Retinal disease detection has been widely studied using artificial intelligence (AI) methods, specifically machine learning (ML). The ML area has constantly grown in recent years because of the rapid increase in the performance of their methods [5]. They can detect if a retina presents DR and which grade of disease is [6], or the lesion location in the image by segmentation [7] [8]. Most approaches train the algorithms from scratch, meaning the model must learn the parameters without a reference.

This work presents a methodology using transfer learning (TL) to segment different lesions using prior knowledge of what a lesion looks like in the retinal image using an encoder-decoder network and a variant of a traditional convolutional neural network (CNN) [9]. The model's output is an image highlighting the target pixels (commonly referred to as segmentation). The proposed method uses the knowledge learned with a simple classification, whether an image has or not a lesion. Then, applying TL [10], a new model segmented three injuries: exudates, hemorrhages, or microaneurysm.

This process reduces the training time significantly, and the pre-trained model is less susceptible to overfitting. It could be implemented in different tasks, for example, a classification of the anatomical parts or segmentation of different diseases like DR with retinal images.

Contribution

1. The methodology helps to have a generalized model before focusing on a specific lesion or disease.

2. Using a pre-trained model as an encoder in segmentation tasks with similar datasets reduces the training time and helps avoid overfitting.

Background

Retinal analysis is an area of much interest. It is not only used to know whether an eye has or not a disease but to segment areas of interest such as the optic disc, fovea area, veins, and arteries [11] [12] [13]. Those tasks have been evaluated with classical ML algorithms like support vector machines [14], or with more classic methods like mathematical morphological algorithms and naive Bayesian for image segmentation [15]. Classical algorithms have attractive advantages because the models do not need many examples to perform well, which is a significant advantage since datasets are typically small.

Deep learning (DL) has been the preferred method to detect, classify, and segment medical images because of its power to generalize data; the only inconvenience is the need for many examples for model training. Lately, researchers have found new methods to confront this problem. Contributions like [16] generate patches of the images to increase the number of examples of retina datasets. Another example implements the superpixel algorithm to generate patches [17].

Most researchers focus on classifying a specific lesion [7] [18] [19] [20]. Others use pre-trained models of different datasets before starting the parameters fine-tuning. It

may help, but in many cases, the fine-tuning dataset is not similar to the pre-trained one, making it almost the same as if the training started from scratch [21]. Many researchers widely used datasets like IDRiD, Drive, CHASE-B, MESSIDOR, KAGGLE, and the papers focus are almost the same, detecting lesions [4] [22] [23] [24], classifying DR [6], or segmenting some anatomical parts [25], or injuries [26]. Nevertheless, the models are trained from scratch or using different pre-trained parameters in a different original dataset like IMAGENET.

Different DL models' effort is to improve the metrics related to the task [22], but not many works try to improve the methodology to avoid starting it from scratch. Works like [27] [28] try to identify a relationship between the vessels caliber and cardiovascular risk problems; an example of exudates segmentation using variants of the classical UNET model [21], with almost the same idea [29], take the dataset and train a model from scratch to segment microaneurysms or some other lesions [24], that classify their presence or absence. All the models achieve high performance, but they will not achieve good results if we try to generalize for different lesions for a specific task.

Most researchers focused on segmenting anatomical areas such as the optic disk, macula [11], veins, and artery [30], training DL models from scratch. Other authors use TL with pre-trained models in a completely different dataset [8] that helps start, but the parameters need to fit the new dataset.

MATERIALS AND METHODS

This methodology consists of two principal stages: stage 1 trains DL models to classify whether or not a retinal image has a lesion, and in stage 2, the objective is to segment a specific lesion employing the knowledge obtained in stage 1 using an encoder-decoder model. The proposed encoder-decoder model has the classifier model's feature extractor as its encoder, and therefore only training the decoder parameters to seg-

ment retinal lesions such as exudates, hemorrhages, or microaneurysms. However, this central idea could be applied to other diseases or segment anatomical parts.

Preprocessing data

As we can see in Figure 1, the proposed methodology starts with image preprocessing, which increases the contrast and luminosity by equalizing the original images using CLAHE [31]. This treatment improves the contrast of some structures like red dots, vessels, and microaneurysms that are difficult to visualize. Then, a data augmentation process is applied to generate two new images of each example consisting of rotations, image slides, or zoom.
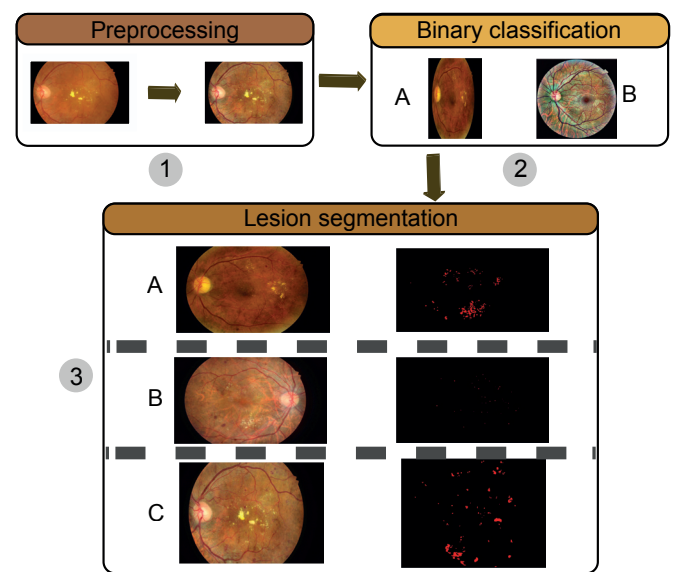

FIGURE 1. Methodology: 1) Preprocessing, 2) binary classification, where A is an image with a lesion and B is a healthy image, and 3) segmentation of injuries (A: exudates, B: microaneurysm, and C: hemorrhages).

The data sets used in stage 1 are the Messidor 1 [32], with 1054 images and Kaggle [33], with 36000. These data sets contain retinal images to classify DR grades. Our study uses them only to identify if an image has an injury, being a binary classification. We obtain 100046 training images and 11116 validation images by applying the preprocessing step.

The datasets used in stage 2 are 48 IDRiD [34] and 30 E-Optha [35], images. These datasets have exudates, hemorrhages, and microaneurysm annotations to execute a pixel-to-pixel classification. This stage uses patches of (160, 160, 3) to train the encoder-decoder model, which means that original images are divided in order to increase the number of examples to 36840 images. Figure 2 shows the two different sets of ima-ges used in stage 2.

Datasets

FIGURE 2. Three different datasets are used in stage 2. Each dataset has the original images patches and their mask.

Stage 1

As previously described, the main objective of stage 1 is to generate the knowledge to identify whether the mentioned lesions appear or not in the image; it does not matter where they are located. The models used for classification are shown in Figure 3. These models are the VGG-16 [36], ResNet50 [37], VGG-16 CBAM and ResNet50 CBAM. The last two models include CBAM [38], which pays attention to the channel and spatial axes. The models contain L2 regularizers [39], batch normalization layers [40], and dropouts [41], to prevent overfitting. The feature extractor of these models is used as an encoder in stage 2. The model's input is a tensor of shape (1, 480, 480, 3).

Stage 2

In stage 1, our models classify only whether an image has a lesion or not. However, now we want to know the lesion position in the image. The segmentation task
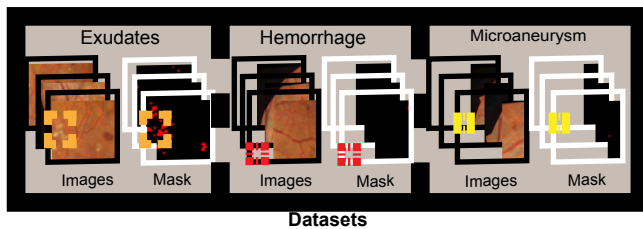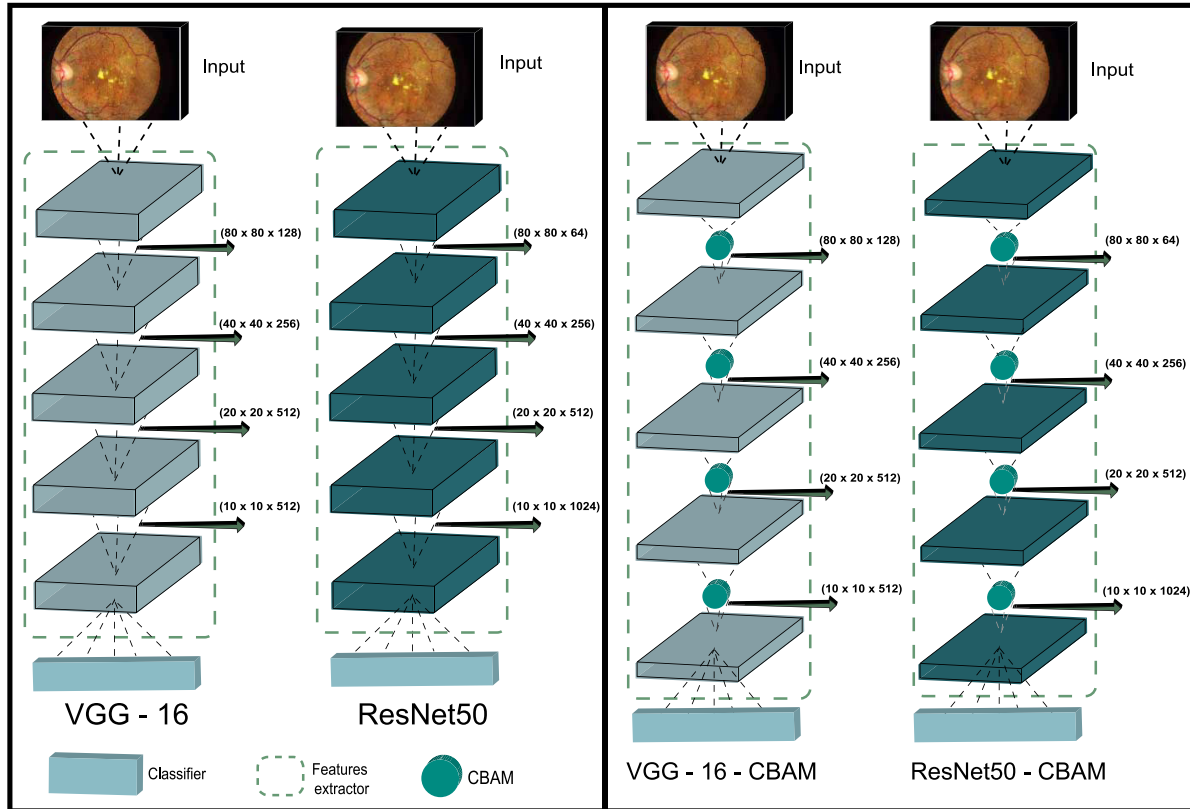

FIGURE 3. The four models were trained with the same hyperparameters: Adam optimizer, dropout, batch normalization, loss function, and activation functions.

consists of locating the required object pixels in the image. For this, we use an encoder-decoder model, the most classical architecture for image segmentation using DL. The segmentation model is not generated from scratch; it uses the feature extractor of the classifier model of stage 1 as the encoder. Typically, the proposed models in the literature are trained with random knowledge parameters meaning that the model must learn what feature extractor maps are helpful to classify the images correctly; it takes more training time and could be more challenging to achieve good results.

The segmentation models, shown in Figure 4, are trained to segment a specific lesion, but not all the parameters are trained.
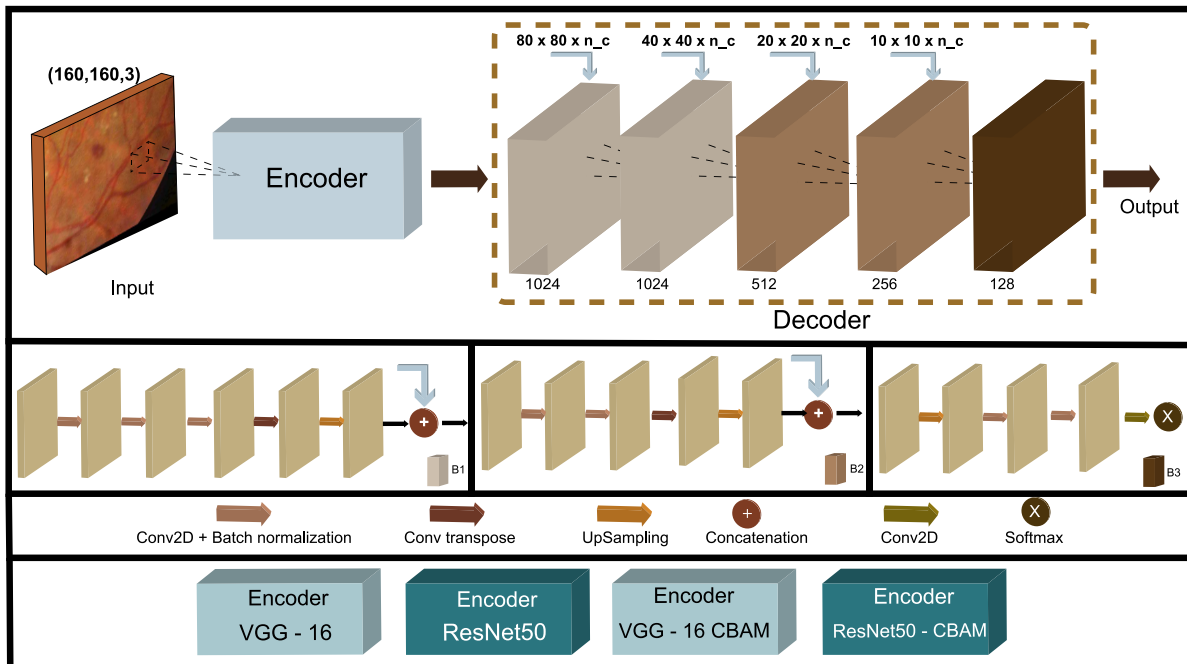

FIGURE 4. The proposed encoder-decoder model to segment the patch images consists of the feature extractor of the classifiers models as the encoder, with a residual operation at the end of each block.

As we mentioned earlier, in stage 1, the encoder generates the knowledge to identify how a lesion looks, so it is unnecessary to train the encoder parameters again, so we freeze them. The decoder parameters are trained to separately identify exudates, hemorrhages, and microaneurysms. We use three different datasets, one for each injury.

Software and Hardware

The models presented in this work were implemented in python 3 with TensorFlow and Keras libraries. OpenCV was used in the preprocessing part.

The classification and segmentation stages were trained in Colab with a GPU accelerator. The code is available in *https://github.com/MetaDown/RMIB-TL*.

RESULTS AND DISCUSSION

We divide the results into two parts: the binary classification model and the segmentation task. The second part has three different metrics for each injury (exudates, hemorrhages, and microaneurysm).

The classifier model obtains the metrics shown in Table 1.

TABLE 1. Binary classification results, showing whether an image has a lesion or not.

Model	Accuracy	Recall	Precision	Training time
VGG-16	87%	92%	80%	24 hrs.
VGG-16 CBAM	89%	81%	98%	26 hrs.
RESNET 50	86.6%	80.3%	90%	16 hrs.
RESNET 50 CBAM	87.1%	82%	90.02%	18 hrs.

This model does not pretend to achieve the best performance; the main task is to generate the knowledge to generalize how a lesion looks in the retinal images.

The classifier model can detect if an image has or does not have a lesion, and it pays attention to the lesion's shape, color, or composition. The knowledge is transferred to the segmentation model to classify pixel by pixel, and then the output is rebuilt to get the original image shape. Figure 5 compares the model segmentation with the ground truth and the original input image to have a better idea of the results; the metrics could show high performance, but it is challenging to observe the actual result due to most of the image being black.

Table 2 shows the performance of the four models, which have similar values, but the sensibility is the lower value of all. The accuracy or AUC metric could be
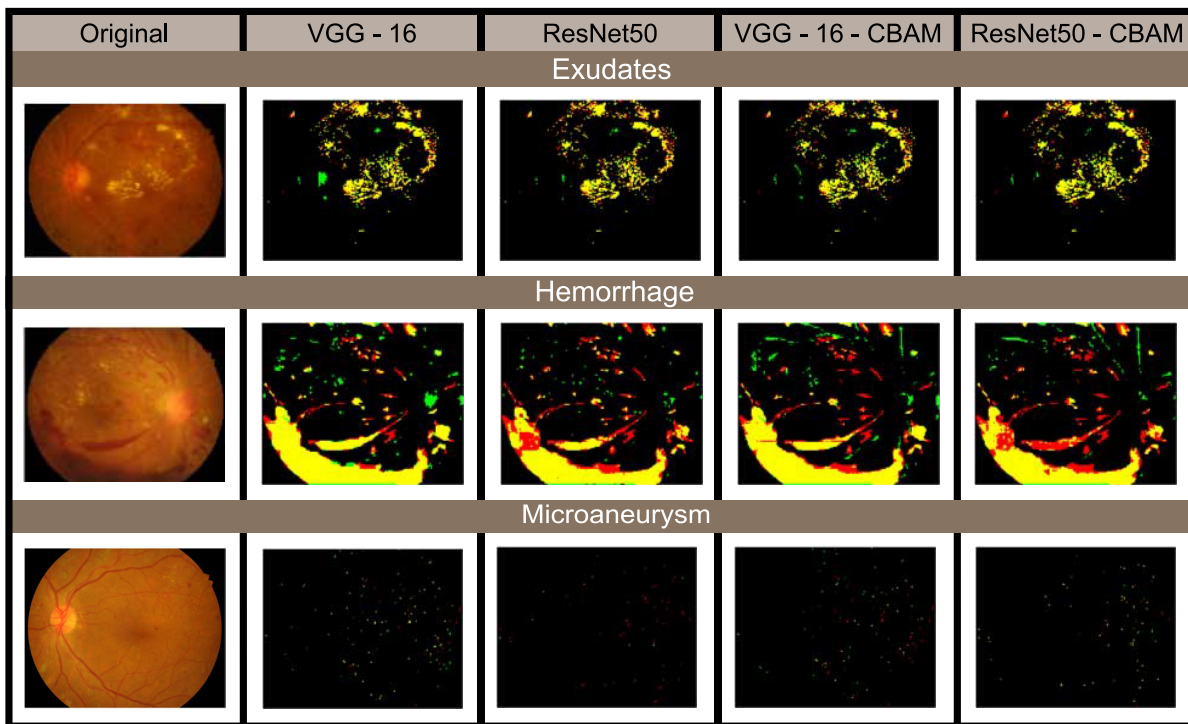

FIGURE 5. Exudates, hemorrhages, and microaneurysms segmentation results. Red pixels are false negatives, green pixels are false positives, yellow pixels represent true positives, and black pixels are true negatives.

misleading due to most of the pixels being black (No lesion presented). If we only pay attention to the accuracy or AUC metric, we are committing a mistake due to the pixel imbalance. We decide to compare the predictions directly with the ground truth to have a com-

plete idea of the model performance, as shown in figure 5. Microaneurysms are the most challenging lesions to segment; their shape and color could confuse other anatomical landmarks and other injuries. Metrics presented in other microaneurysms segmen-

TABLE 2. Metrics results with the validation data.

Author	Accuracy/ AUC/ F1 SCORE	Sensitivity	Specificity
Exudates			
Zong *et al.* [21]	Accuracy: 96.38%	96.14%	97.14%
Wisaeng *et al.* [42]	Accuracy: 98.35%	98.40%	98.13%
VGG-16	Accuracy: 98.1 %	88.12%	96.1 %
VGG-16 CBAM	Accuracy: 98.23%	89.46%	94.09%
RESNET 50	Accuracy: 98.09%	89.05%	96.05%
RESNET 50 CBAM	Accuracy: 97.04%	89.5%	97.2%
Hemorrhages			
Grinsven *et al.* [43]	AUC: 89.4%	91.9%	91.4%
Aziz *et al.* [19]	F1 Score 72.25%	74%	70%
VGG-16	Accuracy: 92.1 %	80.21%	94 %
VGG-16 CBAM	Accuracy: 92.09%	81.4%	94.9%
RESNET 50	Accuracy: 93.5%	81.6%	93.02%
RESNET 50 CBAM	Accuracy: 93.6%	82.01%	95.03%
Microaneurysms			
Long *et al.* [18]	AUC: 87 %	66.9 %	-
Kou *et al.* [29]	AUC: 99.99%	91.9%	93.6%
VGG-16	Accuracy: 93.1 %	73.13%	95.1 %
VGG-16 CBAM	Accuracy: 95.04%	74.6%	96.9%
RESNET 50	Accuracy: 95.69%	70.1%	97.8%
RESNET 50 CBAM	Accuracy: 96.2%	77.7%	98%

tation papers are high because the authors focused on improving the model in that specific task. Our proposed methodology focused on generalizing the model to achieve fast convergence in whatever assignment implies retinal images.

The training time in the segmentation stage by lesion was about one epoch or 28 min, achieving the results shown previously We have excellent time training thanks to the TF applied in stage 1. The model doesn't need to fit the parameters from scratch.

If we only assess the segmenting accuracy, we could conclude that the metrics are very high. Nevertheless, this metric could be misleading since the imbalance present in retinal images is high. For example, the number of microaneurysms' pixels compared to the background pixels is very low. By showing the contrast between the actual and the predicted image, we can better understand the model's performance, Figure 5.

CONCLUSIONS

Training a model from scratch takes more time to fit the parameters until the excellent performance. TL is a powerful alternative to training models since it helps generalize faster than training from scratch. In this paper, using the knowledge learned in a binary classification, we proved that it is possible to perform a specific task like segmentation. The single requirement is a similar dataset as the main problem to solve.

Public datasets with segmenting masks are very limited because of the effort involved in preparing a single instance. This work proved that it is possible to take advantage of datasets created for detection/classification purposes to pre-train DL models to achieve better performance in segmentation tasks. The metrics achieved in our experiments are comparable to the state-of-the-art models and have the advantage of deploying a general methodology using a single fundus image database for the detection/segmentation of various retinal diseases achieving state-of-the-art results. This model could be in practice more valuable since it can be trained with a more realistic database containing a broad spectrum of conditions to detect/segment illnesses without sacrificing performance.

Future work would be to apply the same methodology presented in this work to the new architectures named Transformers, given that they require much more examples than CNN's, and tasks like medical image prediction or classification need a large number of samples, as we could see in the presented work. The advantage of the Transformers is the attention that its model could pay to some specific image regions, converting the Transformers into a strong candidate to replace the conventional CNN's. But we can ask, could we apply TF from CNN to a Transformer achieving fast convergence?

AUTHOR CONTRIBUTIONS

R.O.F. conceptualized the project, collected, and gathered raw data, performed data curation, proposed software for modelling and the methodology to segment and classify the lesions, designed the software for the implementation of models for the analyses, and contributed to writing the original draft. S.T.A. conceptualized the project, proposed software for modelling and the methodology to segment and classify the lesions and contributed to writing the original draft. J.C.P.O. performed data curation and contributed to the reviewing and editing of final version of the manuscript. A.T. proposed software for modelling and the methodology to segment and classify the lesions and contributed to the reviewing and editing of final version of the manuscript. All authors reviewed and approved the final version of the manuscript.

ETHICAL STATEMENT

The databases used in this work are public; therefore, all ethical considerations are met.

REFERENCES

[1] Trucco E, MacGillivray T, Xu Y (eds). Computational Retinal Image Analysis [Internet]. Cambridge, United States: Elsevier; 2019. 481p. Available from: https**://doi.org/10.1016/B978-0-08-102816-2.09994-9**

[2] Prado-Serrano A, Guido-Jiménez MA, Camas- Benítez JT. Prevalencia de retinopatía diabética en población mexicana. Rev Mex Oftalmol [Internet]. 2009; 83(5):261-266. Available: *https://www.medigraphic.com/pdfs/revmexoft/rmo-2009/rmo095c.pdf*

[3] Ricci E, Perfetti R. Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification. IEEE Trans Med Imaging [Internet]. 2007;26(10):1357-1365. Available from: **https://doi.org/10.1109/TMI.2007.898551**

[4] Sarhan MH, Nasseri MA, Zapp D, Maier M, et al. Machine Learning Techniques for Ophthalmic Data Processing: A Review. IEEE J Biomed Health Inform [Internet]. 2020;24(12):3338-3350. Available from: **http://dx.doi.org/10.1109/JBHI.2020.3012134**

[5] Ammu R, Sinha N. Small Segment Emphasized Performance Evaluation Metric for Medical Images. 2020 International Conference on Signal Processing and Communications (SPCOM) [Internet]. Bangalore : IEEE;2020; 1–5. Available from: **http://dx.doi.org/10.1109/SPCOM50965.2020.9179617**

[6] Quellec G, Charrière K, Boudi Y, Cochener B, et al. Deep image mining for diabetic retinopathy screening. Med Image Anal [Internet]. 2017;39:178-193. Available from: **https://doi.org/10.1016/j.media.2017.04.012**

[7] Zhang X, Thibault G, Decencière E, Marcotegui B, et al. Exudate detection in color retinal images for mass screening of diabetic retinopathy. Med Image Anal [Internet]. 2014;18(7):1026-1043. Available from: **https://doi.org/10.1016/j.media.2014.05.004**

[8] Feng Z, Yang J, Yao L, Qiao Y, et al. Deep Retinal Image Segmentation: A FCN-Based Architecture with Short and Long Skip Connections for Retinal Image Segmentation. In: Liu D, Xie S, Li Y, Zhao D, et al. (eds). Neural Information Processing ICONIP 2017. Lecture Notes in Computer Science, vol. 10637 [Internet]. Cham: Springer; 2017; 713-722. Available from: **https://doi.org/10.1007/978-3-319-70093-9_76**

[9] Ye JC, Sung WK. Understanding Geometry of Encoder-Decoder CNNs. 36th International Conference on Machine Learning, ICML 2019 [Internet]. Long Beach: Proceedings of Machine Learning Research; 2019;97:12245-12254. Available from: *https://proceedings.mlr.press/v97/ye19a.html*

[10] Tan C, Sun F, Kong T, Zhang W, et al. A Survey on Deep Transfer Learning. In: Manolopoulos Y, Hammer B, Iliadis L, Maglogiannis I (eds). Artificial Neural Networks and Machine Learning - ICANN 2018. ICANN 2018. Lecture Notes in Computer Science [Internet]. Cham: Springer; 2018. 11041: 270-279. Available from: **http://dx.doi.org/10.1007/978-3-030-01424-7_27**

[11] Tang S, Qi Z, Granley J, Beyeler M. U-Net with Hierarchical Bottleneck Attention for Landmark Detection in Fundus Images of the Degenerated Retina. In: Fu H, Garvin MK, MacGillivray T, Xu Y, et al (eds). Ophthalmic Medical Image Analysis. OMIA 2021. Lecture Notes in Computer Science [Internet]. Cham: Springer; 2021. 12970:62-71. Available from: **https://doi.org/10.1007/978-3-030-87000-3_7**

[12] Welikala RA, Foster PJ, Whincup PH, Rudnicka AR, et al. Automated arteriole and venule classification using deep learning for retinal images from the UK Biobank cohort. Comput Biol Med [Internet]. 2017;90:23-32. Available from: **https://doi.org/10.1016/j.compbiomed.2017.09.005**

[13] Fraz MM, Rudnicka AR, Owen CG, Barman SA. Delineation of blood vessels in pediatric retinal images using decision trees-based ensemble classification. Int J Comput Assist Radiol Surg [Internet]. 2014;9:795-811. Available from: **https://doi.org/10.1007/s11548-013-0965-9**

[14] Adel A, Soliman MM, Khalifa NEM, Mostafa K. Automatic Classification of Retinal Eye Diseases from Optical Coherence Tomography using Transfer Learning. 2020 16th International Computer Engineering Conference (ICENCO) [Internet]. Cairo: IEEE; 2020: 37-42. Available from: **https://doi.org/10.1109/ICENCO49778.2020.9357324**

[15] Xiao Z, Adel M, Bourennane S. Bayesian Method with Spatial Constraint for Retinal Vessel Segmentation. Comput Math Methods Med [Internet]. 2013:401413. Available from: **https://doi.org/10.1155/2013/401413**

[16] Lam C, Yu C, Huang L, Rubin D. Retinal Lesion Detection with Deep Learning Using Image Patches. Invest Ophthalmol Vis Sci [Internet]. 2018;59(1):590-596. Available from: **https://doi.org/10.1167/iovs.17-22721**

[17] Li Q, Feng B, Xie L, Liang P, et al. A Cross- Modality Learning Approach for Vessel Segmentation in Retinal Images. IEEE Trans Med Imaging [Internet]. 2016;35(1):109-118. Available from: **https://doi.org/10.1109/TMI.2015.2457891**

[18] Long S, Chen J, Hu A, Liu H, et al. Microaneurysms detection in color fundus images using machine learning based on directional local contrast. Biomed Eng Online [Internet]. 2020;19(1):21. Available from: **https://doi.org/10.1186/s12938-020-00766-3**

[19] Aziz T, Ilesanmi AE, Charoenlarpnopparut C. Efficient and Accurate Hemorrhages Detection in Retinal Fundus Images Using Smart Window Features. Appl Sci [Internet]. 2021;11(14):6391. Available from: **https://doi.org/10.3390/app11146391**

[20] Liu Q, Liu H, Zhao Y, Liang Y. Dual-Branch Network with Dual-Sampling Modulated Dice Loss for Hard Exudate Segmentation in Colour Fundus Images. IEEE J Biomed Health Inform [Internet]. 2022;26(3):1091-1102. Available from: **https://doi.org/10.1109/jbhi.2021.3108169**

[21] Zong Y, Chen J, Yang L, Tao S, et al. U-net Based Method for Automatic Hard Exudates Segmentation in Fundus Images Using Inception Module and Residual Connection. IEEE Access [Internet]. 2020;8:167225-35. Available from: **http://dx.doi.org/10.1109/ACCESS.2020.3023273**

[22] Tan JH, Fujita H, Sivaprasad S, Bhandary SV, et al. Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. Inf Sci [Internet]. 2017;420:66-76. Available from: **https://doi.org/10.1016/j.ins.2017.08.050**

[23] Kou C, Li W, Yu Z, Yuan L. An Enhanced Residual U-Net for Microaneurysms and Exudates Segmentation in Fundus Images. IEEE Access [Internet]. 2020;8:185514-185525. Available from: **http://dx.doi.org/10.1109/ACCESS.2020.3029117**

[24] Gondal WM, Köhler JM, Grzeszick R, Fink GA, et al. Weakly-supervised localization of diabetic retinopathy lesions in retinal fundus images. 2017 IEEE International Conference on Image Processing (ICIP) [Internet]. Beijing: IEEE; 2017; 2069-2073. Available from: **https://doi.org/10.1109/ICIP.2017.8296646**

[25] Joshi GD, Sivaswamy J, Krishnadas SR. Optic Disk and Cup Segmentation From Monocular Color Retinal Images for Glaucoma Assessment. IEEE Trans Med Imaging [Internet]. 2011;30(6):1192-1205. Available from: **https://doi.org/10.1109/TMI.2011.2106509**

[26] Harangi B, Antal B, Hajdu A. Automatic exudate detection with improved naïve-Bayes classifier. 2012 25th IEEE International Symposium on Computer- Based Medical Systems (CBMS) [Internet]. Rome: IEEE; 2012; 1-4. Available from: **https://doi.org/10.1109/CBMS.2012.6266341**

[27] Cheung CY, Xu D, Cheng CY, Sabanayagam C, et al. A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel caliber. Nat Biomed Eng [Internet]. 2021;5(6):498-508. Available from: **https://doi.org/10.1038/s41551-020-00626-4**

[28] Dai L, Wu L, Li H, Cai C, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nat Commun [Internet]. 2021;12(1):3242. Available from: **https://doi.org/10.1038/s41467-021-23458-5**

[29] Kou C, Li W, Liang W, Yu Z, et al. Microaneurysms segmentation with a U-Net based on recurrent residual convolutional neural network. J Med Imaging [Internet]. 2019;6(2):025008. Available from: **https://dx.doi.org/10.1117%2F1.JMI.6.2.025008**

[30] Xu X, Tan T, Xu F. An Improved U-Net Architecture for Simultaneous Arteriole and Venule Segmentation in Fundus Image. In: Nixon M, Mahmoodi S, Zwiggelaar R (eds). Medical Image Understanding and Analysis. MIUA 2018. Communications in Computer and Information Science [Internet]. Cham: Springer; 2018; 333-340. Available from: **https://doi.org/10.1007/978-3-319-95921-4_31**

[31] Yadav G, Maheshwari S, Agarwal A. Contrast limited adaptive histogram equalization based enhancement for real time video system. 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI) [Internet]. Delhi: IEEE; 2014; 2392-2397. Available from: **https://doi.org/10.1109/ICACCI.2014.6968381**

[32] Decencière E, Zhang X, Cazuguel G, Lay B, et al. Feedback on a publicly distributed image database: The Messidor database. Image Anal Stereol [Internet]. 2014;33(3):231. Available from: **https://doi.org/10.5566/ias.1155**

[33] Kaggle. Diabetic Retinopathy Detection. [Internet] Kaggle. 2015. Available from: *https://www.kaggle.com/c/diabetic-retinopathy-detection*

[34] Porwal P, Pachade S, Kamble R, Kokare M, et al. Indian Diabetic Retinopathy Image Dataset (IDRiD) [Internet]. IEEE Dataport; 2018. Available from: **https://dx.doi.org/10.21227/H25W98**

[35] Decencière E, Cazuguel G, Zhang X, Thibault G, et al. TeleOphta: Machine learning and image processing methods for teleophthalmology. IRBM [Internet]. 2013;34(2):196-203. Available from: **https://doi.org/10.1016/j.irbm.2013.01.010**

[36] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: Bengio Y, LeCun Y (eds). 3rd International Conference on Learning Representations, ICLR 2015 [Internet]. San Diego: arXiv;2015; 1-14. Available from: **https://doi.org/10.48550/arXiv.1409.1556**

[37] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Las Vegas: IEEE; 2016; 770-778. Available from: **https://doi.org/10.1109/CVPR.2016.90**

[38] Woo S, Park J, Lee JY, Kweon IS. CBAM: Convolutional Block Attention Module. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds). Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science [Internet]. Cham: Springer; 2018; 3-19. Available from: **https://doi.org/10.1007/978-3-030-01234-2_1**

[39] Cortes C, Research G, Mohri M, Rostamizadeh A. L2 Regularization for Learning Kernels. 25th Conference on Uncertainty in Artificial Intelligence (UAI 2009) [Internet]. Montreal: Association for Uncertainty in Artificial Intelligence (AUAI); 2009; 109-116. Available from: **https://dl.acm.org/doi/pdf/10.5555/1795114.1795128**

[40] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37 (ICML'15) [Internet]. Lille: JMLR; 2015; 448-456. Available from: **http://proceedings.mlr.press/v37/ioffe15.pdf**

[41] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J Mach Learn Res [Internet]. 2014; 15(56):1929-1958. Available from: **http://jmlr.org/papers/v15/srivastava14a.html**

[42] Wisaeng K, Sa-Ngiamvibool W. Exudates Detection Using Morphology Mean Shift Algorithm in Retinal Images. IEEE Access [Internet]. 2019;7:11946-11958. Available from: **http://dx.doi.org/10.1109/ACCESS.2018.2890426**

[43] van Grinsven MJJP, van Ginneken B, Hoyng CB, Theelen T, et al. Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection in Color Fundus Images. IEEE Trans Med Imaging [Internet]. 2016;35(5):1273-1284Available from: **https://doi.org/10.1109/TMI.2016.2526689**