Use of Convolutional Neural Networks in Smartphones for the Identification of Oral Diseases Using a Small Dataset

Jesús-David González; Jormany Quintero-Rojas

Use of Convolutional Neural Networks in Smartphones for the Identification of Oral Diseases Using a Small Dataset

Jesús-David González[1]

Jormany Quintero-Rojas[2]

Abstract

Image recognition and processing is a suitable tool in systems using machine learning methods. The addition of smartphones as complementary tools in the health area for diagnosis is a fact nowadays due to the advantages they present. Following the trend of providing tools for diagnosis, this research aimed to develop a prototype mobile application for the identification of oral lesions, including potentially malignant lesions, based on convolutional neural networks, as early detection of indications of possible types of cancer in the oral cavity. A mobile application was developed for the Android operating system that implemented the TensorFlow library and the Mobilenet V2 convolutional neural network model. The training of the model was performed by transfer learning with a database of 500 images distributed in five classes for recognition (Leukoplakia, Herpes Simplex Virus Type 1, Aphthous stomatitis, Nicotinic stomatitis, and No lesion). The 80% of the images were used for training and 20% for validation. It was obtained that the application presented at least 80% precision in the recognition of four class. The f1-score and area under curve metrics were used to evaluate performance. The developed mobile application presented an acceptable performance with metrics higher than 75% for the

[1] Universidad de Los Andes (Mérida, Venezuela). ORCID: 0000-0002-9828-0594
[2] M. Sc. Universidad de Los Andes (Mérida, Venezuela). jormany@ula.ve. ORCID: 0000-0002-7180-4685

recognition of three lesions, on the other hand, it yielded an unfavorable performance lower than 70% for identifying nicotinic stomatitis cases with the chosen dataset.

Keywords: artificial intelligence; dentistry; machine learning; medical technology; oral diagnosis; oral disease identification; preventive medicine; smartphone.

Uso de redes neuronales convolucionales en teléfonos inteligentes para la identificación de enfermedades bucales empleando un pequeño conjunto de datos

Resumen

El reconocimiento y procesamiento de imágenes es una herramienta adecuada en los sistemas que usan métodos de aprendizaje automático. La adición de teléfonos inteligentes como herramientas complementarias en el área de la salud para el diagnóstico es un hecho hoy en día por las ventajas que presentan. Siguiendo la tendencia de proporcionar herramientas para el diagnóstico, esta investigación tuvo como objetivo desarrollar una aplicación móvil prototipo para la identificación de lesiones bucales, incluyendo lesiones potencialmente malignas, basado en redes neuronales convolucionales, como la detección temprana de indicios de posibles tipos de cáncer en la cavidad bucal. Se desarrolló una aplicación móvil para el sistema operativo Android que implementó la librería de *TensorFlow* y el modelo de redes neuronales convolucionales *Mobilenet V2*. El entrenamiento del modelo se realizó por transferencia de aprendizaje con una base de datos de 500 imágenes distribuidas en cinco clases para el reconocimiento (Leucoplasia, Herpes Simple Virus Tipo 1, Estomatitis aftosa, Estomatitis nicotínica y Sin lesión). Se utilizó el 80% de las imágenes para el entrenamiento y el 20% para la validación. Se obtuvo que la aplicación presentó al menos 80% de exactitud en el reconocimiento de cuatro clases. Se usaron las métricas de f1-valor y área bajo la curva para evaluar el desempeño. La aplicación móvil desarrollada presentó un comportamiento aceptable con métricas mayores al 75% para el reconocimiento de tres lesiones, por otro lado, arrojó un desempeño desfavorable menor al 70% para identificar los casos de estomatitis nicotínica con el conjunto de datos elegido.

Uso de redes neurais convolucionais em smartphones para identificação de doenças bucais por meio de um pequeno conjunto de dados

Resumo

O reconhecimento e processamento de imagens é uma ferramenta adequada em sistemas que usam métodos de aprendizado de máquina. A incorporação dos smartphones como ferramentas complementares na área da saúde para o diagnóstico é um fato hoje pelas vantagens que apresentam. Seguindo a tendência de disponibilizar ferramentas de diagnóstico, esta pesquisa teve como objetivo desenvolver um protótipo de aplicativo mobile para identificação de lesões bucais, inclusive lesões potencialmente malignas, com base em redes neurais convolucionais, como a detecção precoce de indícios de possíveis tipos de câncer na região. cavidade oral. Um aplicativo móvel foi desenvolvido para o sistema operacional Android que implementou a biblioteca TensorFlow e o modelo de rede neural convolucional Mobilenet V2. O treinamento do modelo foi realizado por transferência de aprendizagem com um banco de dados de 500 imagens distribuídas em cinco classes de reconhecimento (Leucoplasia, Herpes Simplex Vírus Tipo 1, Estomatite aftosa, Estomatite nicotínica e Sem lesão). 80% das imagens foram utilizadas para treinamento e 20% para validação. Obteve-se que o aplicativo apresentou acerto de pelo menos 80% no reconhecimento de quatro classes. As métricas do valor f1 e da área sob a curva foram usadas para avaliar o desempenho. O aplicativo mobile desenvolvido apresentou comportamento aceitável com métricas superiores a 75% para o reconhecimento de três lesões, por outro lado, apresentou desempenho desfavorável inferior a 70% para identificar casos de estomatite nicotínica com o conjunto de dados escolhido.

Palavras-chave: aprendizado de máquina; diagnóstico oral; identificação de doenças bucais; inteligência artificial; medicina preventiva; odontologia; tecnologia médica; telefones inteligentes.

I. INTRODUCTION

The diagnosis of diseases is based on the recognition and analysis of clinical findings, referring to the altered signs and symptoms that the individual presents at a given time [1]. A symptom is the patient's subjective perception related to the disease, while the sign is objectively recorded by the health professional. Some manifestations are characteristic of certain diseases; in others, complementary studies are indispensable to reach a definitive diagnosis [2]. At the time of the clinical examination, the observation of the oral mucosa and skin are indicators to guide the expert to a presumptive diagnosis.

During the clinical inspection, health professionals rely on manual, electronic, and software tools to obtain an accurate diagnosis. Among the existing software tools, those based on artificial intelligence (AI) methods to simplify and improve clinical activities stand out [3]. The continuous evolution of technology has made diagnostic tools evolve and include other means to develop these activities remotely. In this context, the use of mobile devices in medical diagnostic activities is increasing, an example of which is tele-dentistry, which has boosted the use of multiple communication platforms between patients and the dentist using the phone camera to capture the lesion and process it by mobile applications [4,5]. In the current market of apps, some support both the physician and the patient, providing solutions that include ruling out symptoms through multiple questions, obtaining a possible diagnosis based on the answers received [6].

Currently, there are cases of applications that use Artificial Neural Networks (ANN) to identify lesions using a database of more than 12000 images, applications that try to diagnose through some ANN method, taking a photograph to give a medical diagnosis with a high percentage of accuracy [7]. Convolutional Neural Networks (CNN) continue to pioneer Machine Learning (ML) methods because of their fault tolerance and ease of insertion with existing technology [8]. This CNN can be configured using TensorFlow, an open-source library for training and developing ML models developed by Google©. The base unit is the tensor, which can be viewed as

a multidimensional data matrix that includes image information such as width, length in pixels, and each color channel [9].

Current AI tools that support the dentist are based on the analysis of radiographic and computed tomography images [10]. However, evidence supporting the oral diagnosis of soft tissue diseases using computer vision, cell phones, and clinical images are scarce and limited. Having a mobile application that can diagnose in real-time with a simple image using smartphone hardware and software in real-time is advantageous, as it can be used to guide healthcare professionals in making a more accurate diagnosis. These tools can support those dentists who work and perform activities in rural areas where access to a specialist is difficult and complicated. In this sense, the purpose of this work was to develop a prototype mobile application for the identification of oral lesions based on convolutional neural networks.

II. METHODOLOGY

This section describes the data, techniques and methodology used for the development of the project.

A. Dataset

The dataset used in this work was constructed with images found on the internet for each disease using web scraping. 500 different images distributed in 5 classes were used as follows: 97 leukoplakia images, 101 aphthous stomatitis images, 74 nicotinic stomatitis images, 110 HSV-1 images, and 118 images from the no lesion/no recognized class were included, decreasing false positives and improving performance when classifying. To avoid overtraining of the CNN, random linear transformations such as zooms and rotations were performed on each image, increasing the dataset at least eight times from the original set. In the particular case of HVS-1 images were used to test the feature extraction function performed by the mobile CNN used.

B. Model Used

The model used for the mobile application was Mobilenet V2, as it presents the best floating-point model for mobile vision applications. This model offers a lightweight deep neural network whose architecture uses depth separable convolutions, combining this convolution with batch normalization and linear rectifier functions, additionally, softmax was used as the final classifier function. The integrated TensorFlow lite API converter generated the model compatible with mobile applications.

C. Description and Requirements of the Application

The developed mobile application has two activities and one general service. The recognition service prepares the image, loads to the CNN model, analyzes the image, performs inferences, and displays the results. The recognition model transforms the received bitmap to a tensor that is understood by TensorFlow to run the model returning the inference probability through a dynamic list with confidence percentages. Additionally, it is allowed to choose which CPU performs the inference (CPU or neural network API as hardware accelerator). Information regarding frame size, cropping, rotation, and inference time is displayed in a tab (Figure 1). The Android version targeted by the application is Android 9, the minimum being Android 4.4.2 to run and the tests were performed on a Nokia 7.1 device running Android version 9.

Fig. 1. Screenshot of the running application.

D. Data Preparation and Training

The data set was randomly divided into two, 80% for training and 20% for testing and validation. The preprocessing of the images, included in addition to linear transformations, adapting the image to a 640x480px RGB object. To save training and computational time, the transfer learning method of the CNN ImageNet model was applied, using Google Collaboratory and Tensorflow.

III. RESULTS

The performance and validation of the mobile application were tested with 102 images. The inference time during testing did not exceed 500ms. To establish a measurable relationship between the actual and predicted values, the confusion matrix with unbalanced data shown in Figure 2 was performed. The results of the analysis of this matrix are described in Table 1, which summarizes the findings of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) for each class. A Kappa index of 0.739 was obtained to measure concordance.
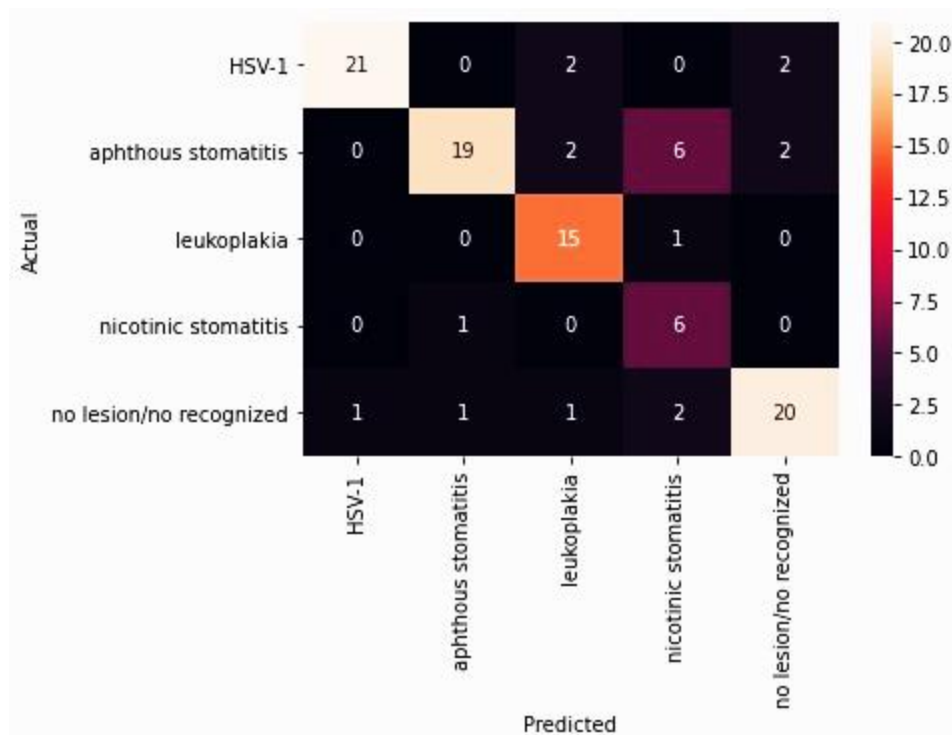
Fig. 2. Confusion matrix for the unbalanced data set.

Table 1. Results of the confusion matrix analysis.

Class	Truth Overall	Classification Overall	TP	FP	FN	TN
Aphthous stomatitis	21	29	19	10	2	71
HSV-1	22	25	21	4	1	76
Leukoplakia	20	16	15	1	5	81
Nicotinic stomatitis	15	7	6	1	9	86
No lesión	24	25	20	5	4	73

TP: true positives, FP: false positives, FN: false negatives, TN: true negatives

The results of the confusion matrix analysis allowed the generation of metrics to compare the performance of each of the classes of the trained model (Table 2). The metrics used were accuracy, precision, sensitivity, specificity, f1-score and area under the curve (AUC) shown in Figure 3. It was observed that the HSV-1 class presented better performance in accuracy and precision, the leukoplakia class was

the most sensitive in this set of tests and the nicotinic stomatitis class the most specific. To compare the overall performance between them, the AUC and f1-score metrics were used, resulting in the HSV-1 class with better performance and the nicotinic stomatitis class with unfavorable performance.

Table 2. Performance metrics for each class.

Class	Accuracy	Precision	Recall	Specificity	F1-score	AUC
Aphthous stomatitis	0.8824	0.6552	0.9048	0.8765	0.7600	0.8907
HSV-1	**0.9510**	0.8400	**0.9545**	0.9500	**0.8936**	**0.9523**
Leukoplakia	0.9412	**0.9375**	0.7500	0.9878	0.8333	0.8689
Nicotinic stomatitis	0.9020	0.8571	0.4000	**0.9885**	0.5455	0.6943
No lesión	0.9118	0.8000	0.8333	0.9359	0.8163	0.8846
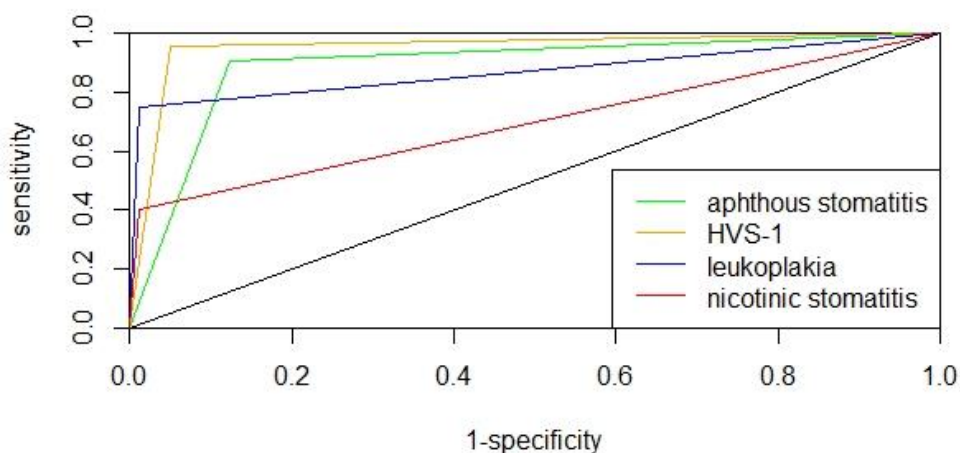
Fig. 3. ROC curve for each disease.

IV. Discussion

The implemented model was based on CNN for the recognition of a group of oral lesions, using the Mobilenet V2 network pre-trained with ImageNet, which is characterized by being a large database with several categories including plants, flowers, animals, objects, among others, with excellent results at the time of classification [11,12]. This is the reason why this model was considered to perform the learning transfer. Lesion recognition systems using learning transfer have been

described in the literature with the AlexNet [11], VGGNet [13,14], and ResNet [12,15-18] network models, whose performance is similar to the model used for this work. Regarding the training of the model, the use of TensorFlow for mobile applications of disease recognition by artificial vision has been evidenced in different research with good results [12, 16, 20, 21].

When working with images of the oral cavity, it is difficult to find many images for each case study. One of the limitations is the patients' consent for the use of their images and the uncertainty about the privacy of their data [22]. The nature of training CNN models for lesion identification requires a large image bank. In this work, the number of items per class was small, around 100 images, which involved using the method of linear transformations, zooming, and random rotations to expand the database at the time of training [9], this technique was also used by Jae-Hong et al. in their research for lesion classification in radiographic images achieving greater than 90% recognition in their classes [23].

During testing and validation, when using Tensorflow Lite, an inference time of less than 500ms was obtained, which is excellent when used as a real-time diagnostic tool. This time is considered fast for identification, as is the percentage change in certainty when the lesion is identified. The recognition of the images is not direct but made based on the probability of belonging to one of the classes programmed in the model.

The application developed presented an acceptable performance for the identification of HSV-1, aphthous stomatitis, and leukoplakia type lesions and an unfavorable performance for the recognition of nicotinic stomatitis type lesions. Unlike other applications, this one is based on clinical recognition of various oral lesions considering observation and computer vision from a smartphone. It has been evidenced that similar lesion recognition systems present a lower performance than the one obtained in this research, around 75% [14, 24-27] using image preprocessing and a higher dataset than the one used in this work. However, studies have described that good prediction performance should be greater than 90% [12, 16, 23, 28, 29]. A similar study used recognition of aphthous stomatitis-like lesions

and HSV-1 with results greater than 90% using a dataset of 200 images and the random forest method as the recognition method [8].

Based on the metrics obtained the best overall performing class was HSV-1, however, in terms of accuracy and specificity attributes other classes performed better. Although the nicotinic stomatitis class obtained unfavorable and unacceptable metrics, it had the best performance in terms of specificity, this is due to the set of images that were used for training the network, despite having a sensitivity of 40% its specificity was 99%. The low sensitivity of this class was based on the similarity of the training images with other classes, which reduces the true positives recognition rate. However, when comparing these metrics with the study of A. Rana et al. it can be seen that the results of this work were superior despite having a similar data set [21].

V. CONCLUSIONS

Lesion recognition systems implemented using machine learning methods must present a good performance, to minimize the false positive and false negative rate. The developed application showed acceptable performance for three chosen lesions, however, the results reflect that this model needs improvements in the network training process. Being able to perform the recognition locally in real time using smartphones presents a great advantage for the use of this application as a diagnostic tool in remote areas without internet access.

This application would serve as an ally for dentists in early clinical diagnosis activities of potentially malignant lesions that can evolve into more complicated pathologies such as oral cancer. For this reason, further research in the area is needed to improve the features of the CNN model, such as increasing the number of images for training, separation of the clinical characteristics of the lesion to increase the rate of true positives and allowing the incorporation of other data to improve the diagnosis using additional data.

Author's Contribution

Jesús-David González: Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Visualization, Writing – original draft.

Jormany Quintero-Rojas: Conceptualization, Data curation, Investigation, Project administration, Resources, Supervision, Validation, Visualization, Writing-review and editing.

References

[1]   C. Carvajal, "El impacto del diagnóstico médico como experiencia traumática. Algunas reflexiones," *Revista Médica Clínica Las Condes,* vol. 28, no. 6, pp. 841-848, 2017. https://doi.org/10.1016/j.rmclc.2017.10.010

[2]   A. Goic, "Sobre el origen y desarrollo del libro Semiología Médica," *Revista Médica Chile*, vol. 146, no. 3, pp. 387-390, 2018. https://doi.org/10.4067/s0034-98872018000300387

[3]   M. Del Río, J. M. López, C. Vaquero, "La inteligencia artificial en el ámbito médico," *Revista Española de Investigaciones Quirúrgicas*, vol. 21, no. 3, pp. 113-116, 2018.

[4]   R. Karthikayan, A. Sukumaran, M. Parangimalai, V. Raj, "Accuracy of smartphone based photography in screening for potentially malignant lesions among a rural population in Tamil Nadu: A cross-sectional study," *Digital Medicine*, vol. 5, no. 2, pp. 56, 2019. https://doi.org/10.4103/digm.digm_29_18

[5]   A. Pereira, S. A. Lazaro, C. G. Molina-Bastos, V. L. Oliveira, R. Nunes, M. Rodrigues, V. Coelho, "Teledentistry in the diagnosis of oral lesions: A systematic review of the literature," *Journal of the American Medical Informatics Association*, vol. 27, no. 7, pp. 1166-1172, 2020. https://doi.org/10.1093/jamia/ocaa069

[6]   M. Estai, Y. Kanagasingam, D. Xiao, J. Vignarajan, B. Huang, E. Kruger, M. Tennant, "A proof-of-concept evaluation of a cloud-based store-and-forward telemedicine app for screening for oral diseases," *Journal of telemedicine and telecare*, vol. 22, no. 6, pp. 319-325, 2016. https://doi.org/10.1177/1357633X15604554

[7]   J. González, "Diseño e implementación de una aplicación móvil inteligente en Android para reconocimiento de lesiones y enfermedades cutáneas y en la mucosa bucal," Grade Thesis, Universidad de Los Andes, Mérida, Venezuela, 2019.

[8]   R. Anantharaman, M. Velazquez, Y. Lee, "Utilizing Mask R-CNN for detection and segmentation of oral diseases," in *IEEE International Conference on Bioinformatics and Biomedicine*, Spain, 2018, pp. 2197-2204. https://doi.org/10.1109/BIBM.2018.8621112

[9]   A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv Preprint*, 2017.

[10]   A. Leite, K. Vasconceslos, H. Willems, R. Jacobs, "Radiomics and machine learning in oral healthcare," *Proteomics Clinical Applications*, vol. 14, e1900040, 2020. https://doi.org/10.1002/prca.201900040

[11]   A. Bhattacharya, A. Young, A. Wong, S. Stalling, M. Wei, D. Hadley, "Precision Diagnosis Of Melanoma And Other Skin Lesions From Digital Images," in *AMIA Joint Summits on Translational Science proceedings*, 2017, pp. 220-226

[12]   J. Velasco, C. Pascion, J. W. Alberio, J. Apuang, J. S. Cruz, M. A. Gomez, B. Jr. Molina, L. Tuala, A. Thio-ac, R. Jr. Jorda, "A Smartphone-Based Skin Disease Classification Using MobileNet CNN," *International Journal of Advanced Trends in Computer Science and Engineering,* vol. 8, no. 5, pp. 2632-2637, 2019. https://doi.org/10.30534/ijatcse/2019/116852019

[13]   A. Romero-Lopez, X. Giro-i-Nieto, J. Burdick, O. Marques, "Skin Lesion Classification from Dermoscopic Images Using Deep Learning Techniques," in *Biomedical Engineering*, Canada, 2017, pp. 49-54. https://doi.org/10.2316/P.2017.852-053

[14]   J. Burdick, O. Marques, J. Weinthal, B. Furht, "Rethinking Skin Lesion Segmentation in a Convolutional Classifier," *Journal of Digital Imaging*, vol. 31, no. 4, pp. 4354440, 2018. https://doi.org/10.1007/s10278-017-0026-y

[15]   S. Han, M. Kim, W. Lim, G. Park, I. Park, S. Chang, "Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm," *Journal of Investigative Dermatology*, vol. 138, no. 7, pp. 1529-1538, 2018. https://doi.org/10.1016/j.jid.2018.01.028

[16]   E. Ziaten-Cerezo, "Clasificación de lesiones en la piel usando aprendizaje profundo," Grade Thesis, Universidad de Málaga, Málaga, Spain, 2019.

[17]   J. Yap, W. Yolland, P. Tschandl, "Multimodal skin lesion classification using deep learning," *Experimental Dermatology*, vol. 27, no. 11, pp. 1261-1267, 2018. https://doi.org/10.1111/exd.13777

[18]   T. Domènech, "Clasificación de imágenes dermatoscópicas utilizando Redes Neuronales Convolucionales e información de metadatos," Grade Thesis, Universidad Politécnica de Catalunya, Barcelona, Spain, 2019.

[19]   C. S. Cheng, P. Shueng, C. Chang, C. Kuo, "Adapting an Evidence-based Diagnostic Model for Predicting Recurrence Risk Factors of Oral Cancer", *Journal of Universal Computer Science*, vol. 24, no. 6, pp. 742-752, 2018. https://doi.org/10.3217/jucs-024-06-0742

[20]   R. Méndez-Hernández, "Aprendizaje profundo para la segmentación de lesiones pigmentadas de la piel," Grade Thesis, Universidad de Sevilla, Sevilla, Spain, 2019.

[21]   A. Rana, G. Yauney, L. C. Wong, O. Gupta, A. Muftu, P. Shah, "Automated Segmentation of Gingival Diseases from Oral Images," in *IEEE Healthcare Innovations and Point of Care Technologies,* United States, 2017, pp. 144-147. https://doi.org/10.1109/HIC.2017.8227605

[22]   M. Estai, Y. Kanagasingam, B. Huang, J. Shiikha, E. Kruger, S. Bunt, M. Tenant, "Comparison of a smartphone-based photographic method with face-to-face caries assessment: a mobile teledentistry model," *Telemedicine and e-Health*, vol. 23, no. 5, pp. 435-440, 2017. https://doi.org/10.1089/tmj.2016.0122

[23]   L. Jae-Hong, K. Do-Hyung, J. Seong-Nyum, "Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network," *Oral Diseases*, vol. 26, pp. 152-158, 2019. https://doi.org/10.1111/odi.13223

[24] J. Yang, X. Sun, J. Liang, P. L. Rosin, "Clinical Skin Lesion Diagnosis Using Representations Inspired by Dermatologist Criteria," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, United States, 2018, pp. 1258-1266. https://doi.org/10.1109/CVPR.2018.00137

[25] J. Serrano-Fernández, "Sistema de ayuda al diagnóstico para la detección temprana de melanomas," Grade Thesis, Universidad Carlos III, Madrid, Spain, 2017.

[26] Y. Gal, R. Islam, Z. Ghahramani, "Deep Bayesian Active Learning with Image Data," in *Proceedings 34th International Conference on Machine Learning*, Australia, 2017, pp. 1183-1192. https://doi.org/10.17863/CAM.11070

[27] J. Sánchez-Hernández, D. Hernández-Rabadán, "Comparación de métodos de clasificación aplicados al diagnóstico de melanomas malignos mediante asimetría," *Programación Matemática y Software*, vol. 6, no. 2, pp. 51-56, 2014.

[28] J. Kawahara, A. BenTaieb, G. Hamarneh "Deep features to classify skin lesions," in *IEEE 13th International Symposium on Biomedical Imaging*, Praga, 2016, pp. 1397-1400. https://doi.org/10.1109/ISBI.2016.7493528

[29] U. Kalwa, C. Legner, T. Kong, S.Pandey, "Skin Cancer Diagnostics with an All-Inclusive Smartphone Application," *Symmetry*, vol. 11, no. 6, pp. 790, 2019. https://doi.org/10.3390/sym11060790