Oral Diseases Recognition Based on Photographic Images

Mazin S. Mohammed[1], Salah Zrigui[1,3] and Mounir Zrigui[2]
[1]*University of Monastir, Research Laboratory in Algebra, Numbers Theory and Intelligent sys-tem (RLANTIS), Monstir 5019, Tunisia*
[2]*University of Al-Mosul, University of the Presidency, Nineveh, Iraq*
[3]*Lig Laboratory, Grenoble, France*

Keywords: Oral Disease Classification, Dental Caries Detection, Dental Images, Deep Learning.

Abstract: Recently, the automation diagnosis process of dental caries plays a critical role in medical applications. This paper presents a new dataset of photo-graphic images for six different types of oral diseases. The dataset is gathered and labelled by professional medical operators in the dentistry field. We use the collected dataset to train a binary classifier to determine whether the region of interests (ROI) needs detection or not inside the input image. Then, we train a detector to detect and localize the required ROI. Finally, we use the detected regions to train a CNN network by adopting transfer learning technique to classify various kinds of teeth diseases. With this model, we obtained an almost 93 % accuracy by modifying and re-training the pre-trained model VGG19.

1 INTRODUCTION

This increasing global vulnerability to diseases has left health care systems worldwide strained. To protect against the spread of disease, hospitals, clinics and nearly all types of medical facilities had to adhere to several protective guidelines. This led to a significant decrease in the number of patients that can be treated at any given moment. In response, researchers, more specifically researchers in the field of artificial intelligence have been innovating and proposing novel methods and technologies to ensure safe diagnosis and treatment with minimal direct contact. One of the most prominent fields for such innovation is the automated diagnosis of dental imagery (Araújo et al., 2023).

AI techniques have been used successfully in various types of disciplines such as nature language processing (Merhbene, Zouaghi and Zrigui, 2010; Mahmoud and Zrigui, 2019), computer vision (Mansouri, Charhad and Zrigui, 2017; Farhani, Terbeh and Zrigui, 2019; Daood, AL-Saegh and Mahmood, 2023), speech recognition (Bellagha and Zrigui, 2020; Slimi et al., 2020; Amari et. Al., 2022), biometrics, smart home applications (Alhafidh et al., 2018), medical imaging, healthcare, robotics, banking & finance, agriculture, military & defence, marketing & advertising, and even oil discovery & gas exploration. Lately, computer vision has been used as an efficient tool in medical applications to offer an accurate diagnosis and avoid errors in human judgement. The use of artificial intelligence in dentistry appears to has a great potential and it is expected to play a vital role in the future of dental health-care and oral diseases diagnosis.

Deep learning strategies have achieved remarkable progress in understanding and analysing dental images. Some neural network architectures such as Convolutional Neural Networks (CNNs) lend themselves naturally to exploit the availability of X-ray and photographic images dataset to perform teeth segmentation, classification, numbering, and lesions detection. With the long waiting time to receive dental care and the importance of an early diagnosis. We decided to build a tool that helps the average person get an early evaluation of his dental state. In this paper we build a dental care detection and classification system that can provide an early diagnosis from a simple picture taken via any smartphone. The system takes as an input a dataset that is comprised of photographic images collected from some local clinics with the help of medical team of specialist dentists. The remainder of the paper is structured as follows. Section 2 examines prior research works. Section 3 presents the material data and the proposed methodology. Section 4 showcases the experimental results. Lastly, we present

Mohammed, M., Zrigui, S. and Zrigui, M.
Oral Diseases Recognition Based on Photographic Images.
DOI: 10.5220/0012361500003636

In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024) - Volume 3*, pages 486-493
ISBN: 978-989-758-680-4; ISSN: 2184-433X

our conclusions in Section 5.

2 LITERATURE REVIEW

In this section we will review some of the methods that have been proposed to automate the diagnosis process of dental diseases. As we mentioned before, all the previous researches were either based on X-ray images or photographic images. For example, the authors in (Liu et al., 2019) proposed intelligent dental Health-IoT system which was implemented on smart hardware. Mobile phone was used as a terminal to capture images in order to perform the diagnosis. MASK R-CNN was used to perform the detection process by applying the training on 12,6000 collected images. The trained model achieved accuracy of 90% to detect and recognize 7 different types of dental diseases. The researchers in (Al Kheraif et al.,2019) collected 800 of X-ray images and then used adaptive histogram equalization which helped to divide the images into back-ground bones and foreground teeth. After that, they used hybrid graph cut to perform the segmentation process to separate the oral cavity and the tissues. Finally, deep learning networks were trained using the segmented images to perform the predication with accuracy of 97%. Orthopantomogram (OPD) images were collected in (Laishram and Thongam, 2020), and then pre-processing techniques were applied to prepare for the training process of faster-RCNN to perform the detection and the classification on the same time. The trained convolutional neural network achieved 90% in the detection process and 99% in the classification process. Mobile app (OralCam) was proposed in (Liang et al., 2020) to offer an end to end complete system with self- examination of five different diseases. 3,182 oral images were taken from 500 participants to train a conventional neural network which was tested to give on average detection sensitivity of 78.7%. 620 photographic images were captured of extracted molars using smartphone in (Duong et al., 2021). The collected images were labelled manually into three classes by four dentists. After that, a series of image pre-processing techniques were applied to enhance the gathered pictures and perform the segmentation process. Finally, the classification process was implemented using SVM classifier which was trained using colour intensity features of the collected dataset. 640 photographic images of different patients' oral cavities were captured using a smartphone in (Ding et al., 2021). Images enchantment and data augmentation were applied on the collected dataset. Data augmentation was used to increase the number of images to 3,990 to prepare the collected data for the training process. Then, transfer learning technique was used by retraining YOLOv3 CNN model to detect and recognize two types of caries.

The authors in (Zhu et al., 2022) presented a deep learning network as U-shape architecture to perform the segmentation process of 3127 panoramic radiograph images. The pro-posed network was called CariesNet to determine three different degrees of caries based on panoramic X-ray images. Additionally, they used full- scale axial attention module to enhance the segmentation process and improve the results. The proposed method achieved 93.61% of accuracy. The researchers in (Rashid et al., 2022) proposed A hybrid system to localize regions of caries by combining photographic and X-ray images. They used the collected dataset to train mask R-CNN deep learning model to perform the segmentation process to detect regions of cavities and oral diseases. The proposed system achieved about 92% of accuracy. 1902 photographic images were taken using a smartphone of 695 participants in (Thanh et al., 2022) to detect three different classes of caries. Four different deep learning architecture were re-trained to detect the oral lesions from the collected images. The trained models were Faster R-CNNs, YOLOv3, RetinaNet, and SSD. A retrospective study was presented in (Keser et al., 2023) by collecting photographic pictures of 65 healthy and 72 oral lesions. Inception V3 deep learning network was trained using the collected dataset to create a binary classifier. The trained architecture achieved accuracy of 100% for healthy and Oral lichen planus lesions cases. The re-searchers in (Gomes et al., 2023) collected 5069 images for six different types of oral mucosal lesions. The images were labelled and cropped manually by specialists. They trained four different convolutional neural networks using 70% of the collected dataset, the rest of the data was used to test the trained models. ResNet-50, VGG16, InceptionV3 and Xception were used as base classifiers to perform the learning process of the proposed models. A dataset of 470 Panoramic X-ray images was labelled and segmented in (Haghanifar et al., 2023). A genetic algorithm was proposed to perform the segmentation process with image processing operations to slice each tooth individually. Finally, capsule network was trained using the extracted features from different deep learning networks to achieve accuracy of 86.05%.

3 DATA MATERIAL AND METHODOLOGY

The first step of our project is data collecting and images gathering. The data collection process is carried out at some local clinics with the help of medical team of specialist dentists. We collect 1600 photographic images representing six common dental diseases. The collected dataset is assembled from both genders (male/female) with ages between 7 to 65 years. Since, the capturing process of the photographic images is an easy process, simply using a smartphone, we were able to collect images even from children. On the other hand, capturing X-ray images for such age is complicated task and prone to errors and mistakes. The collected dataset was obtained in an anonymous manner all recordings of any private information regarding the patients' names, ages, medical history, or even their status have been omitted. Figure 1 shows some samples of the collected images for the six cases of the oral diseases.

In the second stage of our project, we perform the annotation process by applying image labelling to separate our dataset into six categories of dental diseases. This process requires some manual effort to ensure accurate labelling. Unfortunately, the manual labouring of labelling process cannot be carried out by ordinary labellers. Hence, the manual annotation of the collected images is required to be performed by professionals with expertise in the field of dentistry. So, three dentists examined the collected dataset and categorized the images into six cases of oral diseases. Then we divide our dataset into two segments with the ratio of 80%:20%. The first part of our data is used for the training process to learn predication models to perform the detection and the recognition process. The second part of our data is used to test and evaluate the trained models and measure their performance. Convolutional Neural Networks (CNNs) have shown great success in dealing with image related learning tasks. This is due to their natural compatibility with the grid like structure of an image. Therefore, we adopt CNNs to implement the detection and the recognition operations. In our project, we propose a deep learning network to detect the region of teeth to localize the region of interest. Then, we use the detection model to crop the images of our dataset to train a CNN network to perform the classification of oral diseases. We use data augmentation to create more images and increase the dataset size. Data augmentation can be considered as a regularization technique by manipulating the original data to create more copies and synthesize a different version of the images through applying various types of transformation such as rotation, translation, scaling, and even light (brightness) changes. This technique is used to improve the performance and reduce overfitting by exposing the trained models to augmented versions of the original dataset which helps the models to generalize better and become more robust. The collected images are categorized into two kinds, as shown in figure 1. The first type contains only the teeth (which is our region of interest ROI). While the second type contains the teeth and some other parts of the face such as cheeks, nose, lips, and jaws. Therefore, we need a mechanism to separate the two types of images in our dataset. Manual splitting is not an option as it requires time and effort. More importantly, manual split will interfere with the automation process of the diagnosis because we need to detect our region of interest and then send the localized part for the classification process.First, we determine whether the detection process of ROI is required or not. After that, we localize our ROI to send the cropped parts of the images to training process. Finally, we use the training data to train deep learning models and apply the assessment and evaluation using the testing data to measure the accuracy of the trained models. We select 50 images for each case from our dataset, where the first 25 images require detection to localize the region of interest and the other 25 images do not require any detection. The purpose of this collection of images is to train a binary classifier to determine whether the tested image needs detection to find the region of interest or not, so we can use the important parts of the images and remove any unnecessary segments. To achieve the training process of this binary classifier, we need an efficient and accurate model. Therefore, we utilize the concept of the transfer learning by selecting a pre-trained network and modifying the chosen model to perform the binary classification.

Hence, we use MobileNetV2 network as our base model for the training process. Applying the transfer learning approach leverages the prior knowledge of MobileNetV2 network.

Firstly, we eliminate the last layer of the MobileNetV2 network and flatten the resulted features from the final layer. Subsequently, a fully connected layer with 64 nodes is attached to the model. Finally, we add the output layer with two nodes of SoftMax layer to classify the images into two classes (requires detection and does not require). After that, we perform the training process using the selected images.

Figure 1: Photographic samples of our dataset.

During this stage, we freeze all the layers of MobileNetV2 network (base layers) except for the additional layers we added explicitly for binary classification. This strategy allows us to optimize the parameters for newly added layers while retaining the original parameters of the MobileNetV2 network to keep the previous knowledge of the model. In the next phase, we need to train a fast and accurate detector to determine the region of interest in the contents of our images. For this particular task, we use a light deep learning network to achieve the training process of our detector model.

Adopting the transfer learning technique with light pre-trained model can give us an efficient model to implement the detection process. Hence, YOLO V4-tiny model is used for this particular task. YOLO V4-tiny is a single step object detector model which means it can accomplish both the detection and classification process in the same time in one step instead of performing the two operations in separate stages by applying initial detection in one step and subsequent classification in the next step.

YOLO V4-tiny is a scaled down version of the YOLO model has a smaller number of convolution layers (smaller number of parameters) than the ordinary YOLO. Therefore, adopting this model to apply the training process reduces the cost of the training time and the need of huge resources. Additionally, by using transfer learning the pre-trained YOLO V4-tiny returns the optimal values of small number of parameters in the selected network. Adopting this technique makes the learning process possible despite the small size of the dataset.

Before the training process, we need to select images from our data set to perform the detection process to localize the region of interest inside the images. Therefore, we selected 50 images, from each case, which need detection. Then, we need to label these images by providing a bounding box to determine the coordinates of the region of interest for each individual image. Since, we intend to train YOLO V4-tiny model to perform the detection process, the coordinates of bounding box for the labelling operation should match the format of YOLO network. Thus, we use Bbox-Label-Tool-Multi-Class of the Darknet-library for the labelling process. It is labelling tool that is completely compatible with the YOLO format. This tool offers a programmable configuration to initialize different modes of setting. BBox-Label-Tool can create a simple GUI window to input images and give the facility to label the region of interest by applying a bounding box manually by a user.

After the user localizes the region of interest, BBOX tool will create the necessary files with the required information for the training process. Figure 2 shows samples of the image labelling. Once the labelling process of the selected images is done, we can utilize these images to train our customized detection model. As we mention before, we use for this particular task YOLO V4-tiny network. We retrain the pre-trained YOLO V4-tiny model with our dataset to implement the learning process of the detector. When the training process is completed, we use the trained detector to localize the region of interest to crop these regions. The cropped images are used to prepare the training data to perform the training of the disease classification. Figure 4 shows the pipeline of the proposed method.

As shown in figure 3, after applying the detection process, to determine the regions of interest, we create database of the training images to learn recognition models to classify 6 different types of dental caries. Transfer learning approach is adopted to achieve the training process of the classification models. The analysis of the transfer learning achieves the learning process by relying on the prior knowledge from a pre-trained model. So, instead of starting the training from scratch the learning process starts with trained parameters of a base model which has been trained using extensive amount of data. Currently, numerous numbers of pre-trained models are available to be utilized as base classifiers to perform the training.

To implement the training of the disease's recognizer, we use VGG16 network as a foundational classifier to exploit the prior knowledge of the selected model. We modify the architecture of the chosen network to achieve the training process with our own images. We replace the last layer of VGG16 (which is responsible for classification of 1000 classes) with new classification layers for our 6 classes.

As shown in Figure 4, we include flatten layer to make the size of vector features compatible with the new attached layers. Then, we add a fully connected layer of 512 nodes with a drop out layer of 0.5 dropping factor. The primary advantages of the drop out layer is to decrease the effect of the overfitting problem by skipping the update of the parameter's values during the training. Then, we add a second fully connected layer of 256 nodes and followed by another drop out layer. Finally, we wrap up the designed network with a soft-max layer with 6 output nodes to represent each individual disease.

After we complete the architecture of the proposed network, we need to retrain the designed model by applying fine-tune process to update the weights and the parameters of our network to adjust their values in a suitable configuration which allows the trained model to capture and learn the most relevant features from our dataset to achieve the diseases classification task.

In order to improve the results, we expand our set of experiments by testing different types of network architectures as a base classifier. Hence, we use additional pre-trained models to boost the accuracy of our classifier. We exploit the prior knowledge obtained from the following models: VGG16, VGG19, GoogleNet, Xception, InceptionV3, InceptionResNetV2, DenseNet201, MobileNetV2, and NASNetLarge. Practically, we repeat the same procedure of the fine-tune process to retrain these models.

First, we remove the classification layer from the selected network and we attach a fully connected layers of 512 nodes with drop out layer, and followed by another fully connected of 265 nodes with another drop out layer. At the end of the network, we add a soft max layer of 6 outputs nodes to classify 6 different diseases. All the results of these experiments will be shown in the next section.

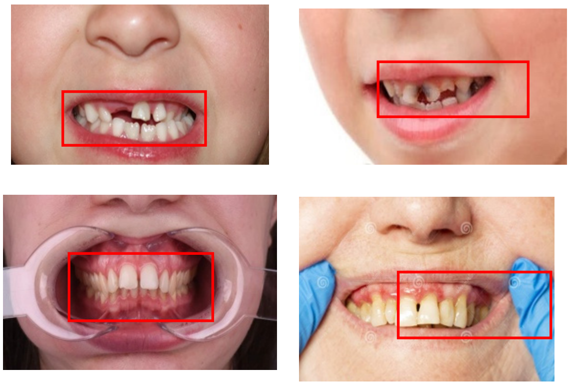

Figure 2: Data annotation to train the detector model.

4 THE RESULTS

In the current section, we provide the finding results of our experiments. We use the training data set to learn the optimal features of our images to create diseases classification model. Transfer learning method is used to perform the recognition process. First, we modify VGG16 network to implement the re-training process to reshape the weights and the parameters of the selected network to capture the optimal representation of the learnt features of our dataset. We utilize the testing segment of our dataset to assess the quality of the trained models. Therefore, we compute the accuracy, precision, and recall as metric measurements to conduct comprehensive

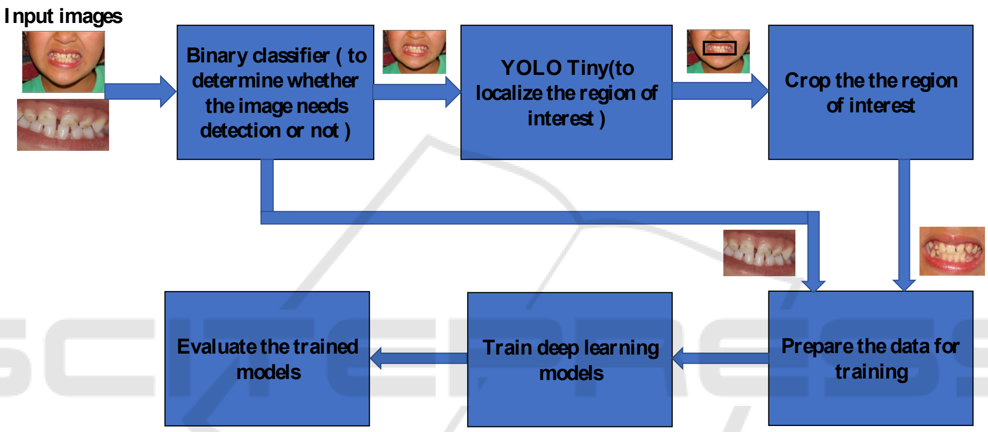

Figure 3: Flowchart of the proposed algorithm.

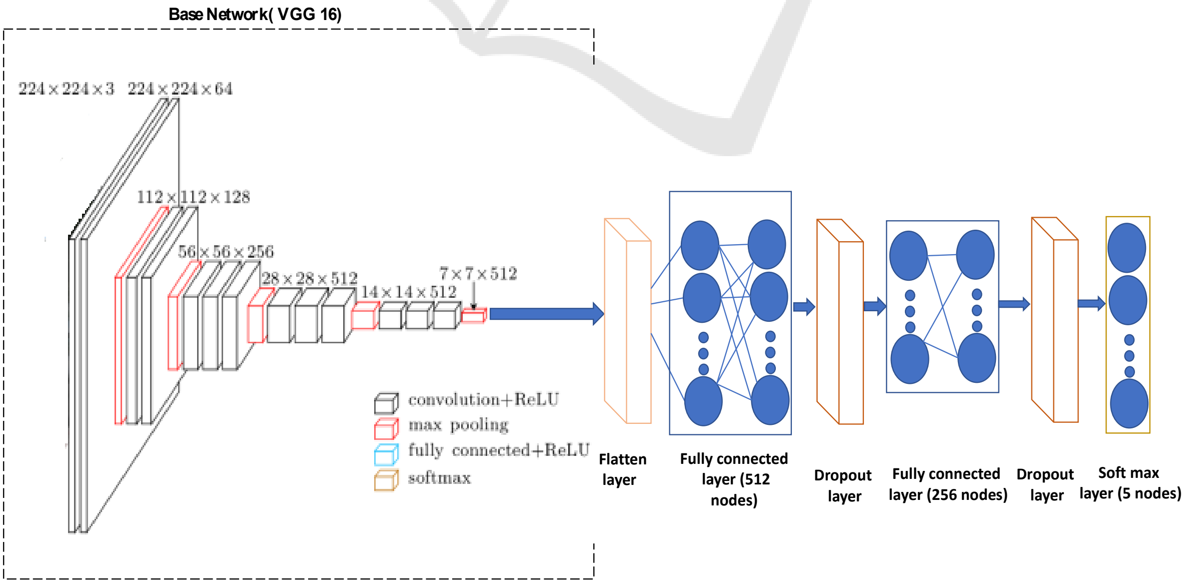

Figure 4: The architecture of the designed CNN network.

assessments, validations, and comparative analyses of the trained classifiers. By adopting VGG16, we obtain almost 92% of accuracy. Clearly, the prior knowledge obtained by training a model using extensive amount of dataset achieve reasonable performance by applying a fine-tuning process to learn the optimal features representation of recognizing 6 dental caries in our dataset.

Table 1: Accuracy of deep learning networks.

Base network model	Accuracy	Precision	Recall
VGG16	92.11%	92.42%	92.99%
VGG19	**93.55%**	**93.27 %**	**93.13%**
AlexNet	89.19%	88.99%	89.41%
Resnet50	92.74%	92.02%	92.83%
GoogleNet	91.25%	91.14%	90.79%
NasNet-Mobile	85.86%	86.32%	85.95%
DenseNet201	92.76%	92.58%	92.39%
MobileNetV2	87.82 %	88.01%	88.91%
InceptionResNet V2	90.68%	90.49%	90.18%
Xception	90.45%	90.61%	90.71%
InceptionV3	90.41%	90.97%	90.39%

It is important to highlight that all the experiment and the obtained results are conducted by a personal laptop type Lenovo where the processor is Core I7 with memory of 16 G RAM. The re-trained models which are used in our experiments exhibit diversity in their characteristics presenting variations in the architecture design, connection mappings, layer depth, and parameters quantities. Therefore, they offer different performance and efficiency based on their variations and properties. Basically, these models may respond differently to the new given task with unseen dataset to their previous knowledge. Upon examining the results of Table 1, the comparisons clearly demonstrate that VGG19 outshines as the most prominent base model which offered an almost 93% of accuracy. The conducted experiments show that the modified version of VGG19 stands out by learning the best features of our dataset to encapsulate the optimal representation of different patterns for the oral diseases. Furthermore, measured the confusion matrix of the modified version of VGG19, these results are shown in figure 5.

Figure 5: The confusion matrix of the re-trained model VGG19.

5 CONCLUSIONS

In this paper, we propose a new dataset of photographic images to train a predication model to diagnose 6 different kinds of oral diseases. The gathered images are annotated by expert dentists. The collected images are used to train a binary recognizer to determine whether detection is necessary inside the dental images to find the region of interest (ROI). After that, we deploy a modified version of YOLO V4-tiny network to perform the detection process of ROI. The detected parts of ROI within our data are cropped to prepare our dataset for the classification process. Finally, we adopt the transfer learning strategy to train multiple pre-trained models to implement the recognition process. The modification of these models allows us to exploit their previous knowledge and achieve 93% accuracy to classify six different types of oral diseases.

REFERENCES

Araújo, A. L. D., da Silva, V. M., Kudo, M. S., de Souza, E. S. C., Saldivia-Siracusa, C., Giraldo-Roldán, D., ... & Moraes, M. C. (2023). Machine learning concepts applied to oral pathology and oral medicine: A convolutional neural networks' approach. *Journal of Oral Pathology & Medicine*, 52(2), 109-118.

Amari, R., Noubigh, Z., Zrigui, S., Berchech, D., Nicolas, H., & Zrigui, M. (2022, September). Deep Convolutional Neural Network for Arabic Speech Recognition. *In International Conference on Computational Collective Intelligence* (pp. 120-134). Cham: Springer International Publishing.

Alhafidh, B. M. H., Daood, A. I., Alawad, M. M., & Allen, W. (2018). FPGA hardware implementation of smart home autonomous system based on deep learning. *In Internet of Things–ICIOT 2018: Third International Conference, Held as Part of the Services Conference Federation, SCF 2018, Seattle, WA, USA, June 25-30, 2018, Proceedings* 3 (pp. 121-133). Springer International Publishing.

Al Kheraif, A. A., Wahba, A. A., & Fouad, H. (2019). Detection of dental diseases from radiographic 2d dental image using hybrid graph-cut technique and convolutional neural network. *Measurement,* 146, 333-342.

Bellagha, M. L., & Zrigui, M. (2020, November). Speaker naming in tv programs based on speaker role recognition. *In 2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA)* (pp. 1-8). IEEE.

DAOOD, A., AL-SAEGH, A. L. I., & MAHMOOD, A. F. (2023). HANDWRITING DETECTION AND RECOGNITION OF ARABIC NUMBERS AND CHARACTERS USING DEEP LEARNING METHODS. *Journal of Engineering Science and Technology*, 18(3), 1581-1598.

Ding, B., Zhang, Z., Liang, Y., Wang, W., Hao, S., Meng, Z., ... & Lv, Y. (2021). Detection of dental caries in oral photographs taken by mobile phones based on the YOLOv3 algorithm. *Annals of Translational Medicine*, 9(21).

Farhani, N., Terbeh, N., & Zrigui, M. (2019). Object recognition approach based on generalized hough transform and color distribution serving in generating arabic sentences. *International Journal of Computer and Information Engineering*, 13(6), 339-344.

Duong, D. L., Kabir, M. H., & Kuo, R. F. (2021). Automated caries detection with smartphone color photography using machine learning. Health Informatics Journal, 27(2), 14604582211007530.

Gomes, R. F. T., Schmith, J., Figueiredo, R. M. D., Freitas, S. A., Machado, G. N., Romanini, J., & Carrard, V. C. (2023). Use of artificial intelligence in the classification of elementary oral lesions from clinical images. *International Journal of Environmental Research and Public Health*, 20(5), 3894.

Haghanifar, A., Majdabadi, M. M., Haghanifar, S., Choi, Y., & Ko, S. B. (2023). PaXNet: Tooth segmentation and dental caries detection in panoramic X-ray using ensemble transfer learning and capsule classifier. *Multimedia Tools and Applications*, 1-21.

Karaoglu, A., Ozcan, C., Pekince, A., & Yasa, Y. (2023). Numbering teeth in panoramic images: A novel method based on deep learning and heuristic algorithm. *Engineering Science and Technology, an International Journal*, 37, 101316.

Keser, G., Bayrakdar, İ. Ş., Pekiner, F. N., Çelik, Ö., & Orhan, K. (2023). A deep learning algorithm for classification of oral lichen planus lesions from photographic images: A retrospective study. *Journal of Stomatology, Oral and Maxillofacial Surgery*, 124(1), 101264.

Kim, H. N., Kim, K., & Lee, Y. (2023). Intra-oral photograph analysis for gingivitis screening in orthodontic patients. *International Journal of Environmental Research and Public Health*, 20(4), 3705.

Laishram, A., & Thongam, K. (2020, February). Detection and classification of dental pathologies using faster-RCNN in orthopantomogram radiography image. *In 2020 7th international conference on signal processing and integrated networks (SPIN)* (pp. 423-428). IEEE.

Liang, Y., Fan, H. W., Fang, Z., Miao, L., Li, W., Zhang, X., ... & Chen, X. A. (2020, April). OralCam: enabling self-examination and awareness of oral health using a smartphone camera. *In Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-13).

Liu, L., Xu, J., Huan, Y., Zou, Z., Yeh, S. C., & Zheng, L. R. (2019). A smart dental health-IoT platform based on intelligent hardware, deep learning, and mobile terminal. *IEEE journal of biomedical and health informatics*, 24(3), 898-906.

Mansouri, S., Charhad, M., & Zrigui, M. (2017). Arabic Text Detection in News Video Based on Line Segment Detector. *Res. Comput. Sci*., 132, 97-106.

Mahmoud, A., & Zrigui, M. (2019). Deep neural network models for paraphrased text classification in the Arabic language. *In Natural Language Processing and Information Systems: 24th International Conference on Applications of Natural Language to Information Systems, NLDB 2019, Salford, UK, June 26–28, 2019, Proceedings* 24 (pp. 3-16). Springer International Publishing.

Merhbene, L., Zouaghi, A., & Zrigui, M. (2010, June). Ambiguous Arabic words disambiguation. *In 2010 11th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing* (pp. 157-164). IEEE.

Rashid, U., Javid, A., Khan, A. R., Liu, L., Ahmed, A., Khalid, O., ... & Nawaz, R. (2022). A hybrid mask RCNN-based tool to localize dental cavities from real-time mixed photographic images. *PeerJ Computer Science*, 8, e888.

Slimi, A., Hamroun, M., Zrigui, M., & Nicolas, H. (2020, November). Emotion recognition from speech using spectrograms and shallow neural networks. *In Proceedings of the 18th International Conference on Advances in Mobile Computing & Multimedia* (pp. 35-39).

Thanh, M. T. G., Van Toan, N., Ngoc, V. T. N., Tra, N. T., Giap, C. N., & Nguyen, D. M. (2022). Deep learning application in dental caries detection using intraoral photos taken by smartphones. *Applied Sciences*, 12(11), 5504.

Zhu, H., Cao, Z., Lian, L., Ye, G., Gao, H., & Wu, J. (2022). CariesNet: a deep learning approach for segmentation of multi-stage caries lesion from oral panoramic X-ray image. *Neural Computing and Applications*, 1-9.