

PAPER

Automated Detection of Dental Caries from Oral Images using Deep Convolutional Neural Networks

Imane Lasri(✉), Naoufal
El-Marzouki, Anouar
Riadsolh, Mourad
Elbelkacemi

Mohammed V University in
Rabat, Rabat, Morocco

imane_lasri@um5.ac.ma

ABSTRACT

The urgent demand for accurate and efficient diagnostic methods to combat oral diseases, particularly dental caries, has led to the exploration of advanced techniques. Dental caries, caused by bacterial activities that weaken tooth enamel, can result in severe cavities and infections if not promptly treated. Despite existing imaging techniques, consistent and early diagnoses remain challenging. Traditional approaches, such as visual and tactile examinations, are prone to variations in expertise, necessitating more objective diagnostic tools. This study leverages deep learning to propose an explainable methodology for automated dental caries detection in images. Utilizing pre-trained convolutional neural networks (CNNs) including VGG-16, VGG-19, DenseNet-121, and Inception V3, we investigate different models and preprocessing techniques, such as histogram equalization and Sobel edge detection, to enhance the detection process. Our comprehensive experiments on a dataset of 884 oral images demonstrate the efficacy of the proposed approach in achieving accurate caries detection. Notably, the VGG-16 model achieves the best accuracy of 98.3% using the stochastic gradient descent (SGD) optimizer with Nesterov's momentum. This research contributes to the field by introducing an interpretable deep learning-based solution for automated dental caries detection, enhancing diagnostic accuracy, and offering potential insights for dental health assessment.

KEYWORDS

cavity, dental health, deep CNN, sobel edge detection

1 INTRODUCTION

The global prevalence of oral diseases, with dental caries as a major contributor, underscores the urgent need for accurate and efficient diagnostic methods. Dental caries, stemming from bacterial activities that weaken tooth enamel, can result in painful cavities and infections if untreated [1–3]. Despite the utilization of imaging techniques to aid in caries detection, challenges persist in achieving consistent and early diagnoses [4]. Early efforts in dental caries detection primarily relied on conventional methods. These methods included visual and tactile examination by dental professionals, which,

Lasri, I., El-Marzouki, N., Riadsolh, A., Elbelkacemi, M. (2023). Automated Detection of Dental Caries from Oral Images using Deep Convolutional Neural Networks. *International Journal of Online and Biomedical Engineering (iJOE)*, 19(18), pp. 53–70. <https://doi.org/10.3991/ijoe.v19i18.45133>

Article submitted 2023-09-19. Revision uploaded 2023-10-17. Final acceptance 2023-10-19.

© 2023 by the authors of this article. Published under CC-BY.

while valuable, were subject to variations in expertise and the ability to identify subtle lesions [5]. As a result, a pressing need emerged for more objective and reproducible diagnostic tools to complement the clinical judgment of dental experts. In response to this demand, researchers began exploring imaging technologies to enhance caries detection. Methods such as visible light transillumination [6] and calibrated diaphragm computed tomography [7] emerged as pioneering approaches. While these techniques showed promise, they often required specialized equipment and were limited in their ability to identify early caries lesions comprehensively.

Following the early works in dental caries detection, machine learning techniques [8] began to gain prominence as powerful tools for improving diagnostic accuracy. These techniques offered the potential to automate the process further and reduce the reliance on human expertise. Methods like artificial neural networks and traditional machine learning algorithms were employed to analyze dental images, contributing to the evolution of caries detection methodologies. Farhadian et al. [9] developed a support vector machine (SVM)-based system for diagnosing periodontal diseases. With a radial kernel function, the SVM model demonstrated effective performance in classifying periodontitis. Thanathornwong [10] developed a clinical decision support system for orthodontic treatment assessment, utilizing a Bayesian network model. The system demonstrated strong alignment with expert orthodontists' judgments, showcasing its accuracy in classifying patients' treatment needs. Another avenue of research delved into computer-assisted caries detection using X-ray images [11]. Studies, such as the one by Oliveira et al. [12], employed artificial neural networks to identify dental caries in panoramic dental X-ray images, achieving impressive accuracy rates. Tikhe et al. [13] explored algorithmic approaches for detecting enamel and interproximal caries in digital periapical radiography images.

After the emergence of deep learning, its successful application has transcended numerous domains, revolutionizing sectors such as education [14, 15, 16], healthcare [17], and finance [18]. In the realm of healthcare, deep learning techniques [19] have demonstrated their potential to transform traditional diagnostic and treatment methods. Notably, Gharaibeh et al. [20] employed Swin Transformer-Based Segmentation and Multi-Scale Feature Pyramid Fusion Module for Alzheimer's Disease with machine learning. Similarly, Al-hazaimeh et al. [21] demonstrated the power of combining artificial intelligence and image processing for the diagnosis of diabetic retinopathy in retinal fundus images, highlighting the capabilities of deep learning in medical diagnostics. In the domain of dental care, the utilization of CNNs and other deep learning architectures has enabled accurate detection of diabetic retinopathy in retinal images [22]. Bychkov et al. [23] combined convolutional and recurrent networks to predict colorectal cancer outcomes directly from tissue images, outperforming human histological assessment and highlighting deep learning's profound potential in medical diagnostics. Additionally, Oztekin et al. [24] introduced an explainable deep learning model, enhancing precise caries detection and offering visual insights to enhance diagnostic accuracy. Rao et al. [25] utilized Inception-V3, Inception-Resnet-V3, and DenseNet-121 to predict odontogenic keratocyst (OKC) recurrence from histopathology images of incisional biopsies. Their ensemble model improved accuracy and efficiency compared to traditional methods, demonstrating the transformative potential of deep learning in dental diagnostics.

In this context, our study contributes to the field by proposing an explainable deep learning-based approach for automatic caries detection in dental images, driven by the pressing need to provide accurate and efficient diagnostic methods for dental care. Leveraging pre-trained deep CNNs, such as VGG-16 [26], VGG-19 [26], DenseNet-121 [27], and Inception V3 [28], we explore various models and harness their capabilities for end-to-end caries detection. Before classification,

we incorporated preprocessing techniques such as histogram equalization and Sobel edge detection to enhance the robustness of our approach. We conducted comprehensive experiments on a collection of 884 oral images categorized as either “cavity” or “no_cavity”, validating the efficacy of our methodology.

The primary contributions of this research can be outlined as follows:

- Introduction of an explainable deep learning-based methodology for automated caries detection in dental images.
- Integration of preprocessing techniques, including histogram equalization and Sobel edge detection, to enhance the accuracy and reliability of the approach.
- Utilization of well-established pre-trained deep CNNs (VGG-16, VGG-19, DenseNet-121, and Inception V3) to create a robust end-to-end caries detection model.
- The performance of the model is evaluated by various evaluation metrics, affirming its accuracy and effectiveness in automated caries detection.

This research paper is represented as follows. Section 2 introduces our proposed deep learning-based approach for dental cavity classification. Section 3 analyzes the results, followed by the conclusion and future work in Section 4.

2 PROPOSED METHODOLOGY

To provide a better understanding of the proposed methodology, we present in the following subsections an overview of the approach designed for dental cavity classification. This study introduces a deep learning model for distinguishing between teeth with or without cavities in dental images. Popular pre-trained Deep CNN models are used to lower the expenses of model training from scratch. These models include pre-learned and optimized weights on large datasets. The single tooth image given as input to the deep learning model is estimated as “cavity” or “no_cavity” at the output. Figure 1 illustrates the architecture of the proposed approach.

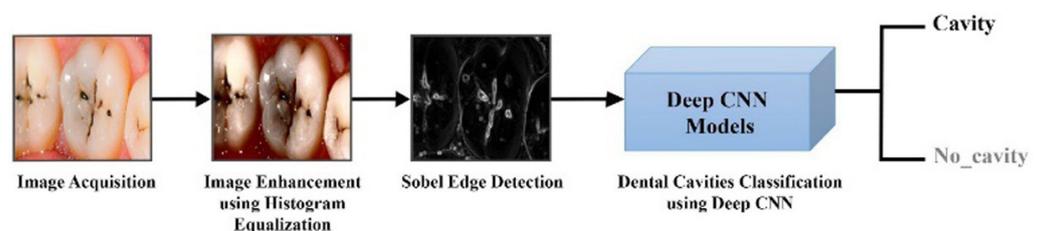


Fig. 1. Proposed architecture for dental cavity classification using Deep CNN

2.1 Contrast enhancement using histogram equalization

Upon the import of dental images, the need for contrast enhancement arises to mitigate noise presence. This enhancement process employs histogram equalization for each image. Histogram equalization is a point-based technique that effectively redistributes and fine-tunes the intensity values within the image, resulting in a uniform histogram representation. Achieving this involves an initial assessment of the image’s histogram as part of the contrast enhancement process. Subsequently, the process calculates the normalized addition of the image histogram, a step that facilitates the transformation of the original image into the output picture. It’s important to note that images with weaker intensity value distributions yield less information, whereas those with more optimal distributions provide richer data. Figure 2 illustrates the outcome of contrast enhancement through histogram equalization.

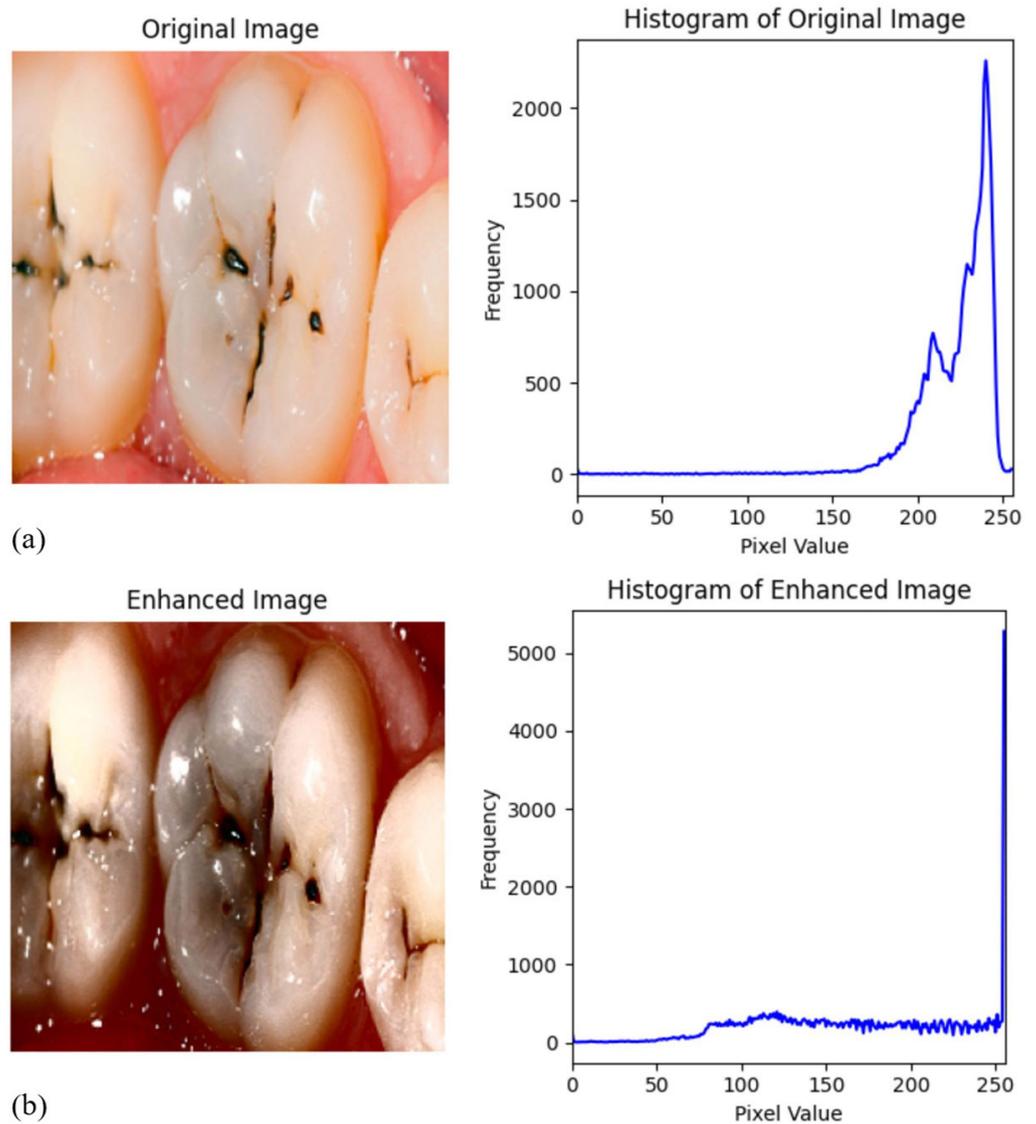


Fig. 2. Image enhancement through histogram equalization: (a) Original image with histogram and (b) Transformed image with histogram

2.2 Sobel edge detection

After performing histogram equalization and converting the image to grayscale, we further employed the Sobel edge detection technique [29]. This method involves detecting edges in two perpendicular directions - horizontally (G_x) and vertically (G_y). Instead of considering these directions in isolation, we made the deliberate choice to combine the results from both directions. This combination creates a comprehensive edge map that encompasses edge information in both the horizontal and vertical orientations, resulting in a more comprehensive representation of the image's edge features, as shown in Figure 3.

The algorithm for Sobel edge detection consists of several steps:

- Step 1: Input the image data.
- Step 2: Apply the G_x and G_y masks to the image data.
- Step 3: Execute the Sobel edge algorithm for gradient detection.

- Step 4: Independently configure the Gx and Gy masks on the image data.
 Step 5: Combine the results to calculate the absolute magnitude of the gradient.
 Step 6: The output represents the edges in terms of their absolute magnitude.

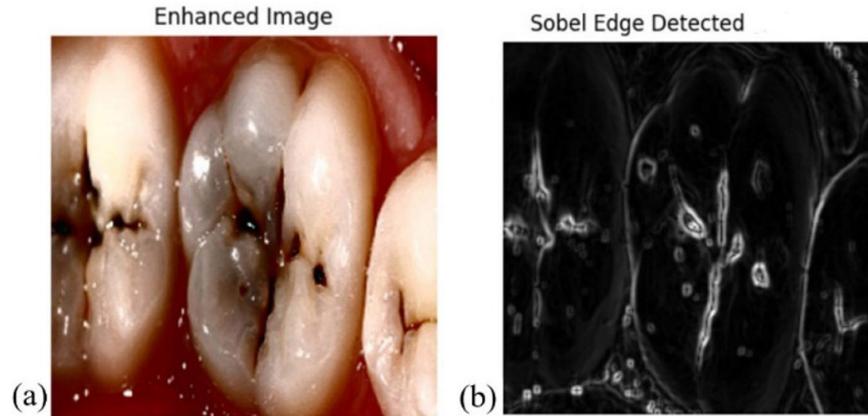


Fig. 3. Sobel edge detection: (a) Image after histogram equalization and before Sobel, (b) Image after Sobel edge detection

2.3 Image classification using Deep CNN models

In this study, we utilized CNN-based models, including VGG-16 [26], VGG-19 [26], DenseNet-121 [27], and Inception V3 [28], for image classification. The architecture of each model is illustrated in Figures 4–7, respectively. These models were pretrained on a large image dataset named ImageNet [30], and their weights optimized, making them suitable for transfer learning. To adapt them to our specific task of tooth image classification, we fine-tuned the last layers. We employed a common fine-tuning technique, freezing the Conv blocks in the pretrained models to retain their weights and replaced the final dense layers with new ones designed for classifying images into “cavity” and “no_cavity”. The new dense layers had channel sizes of 256 and 32, respectively, and were followed by RELU activation, batch normalization, and dropout layers (with $p = 0.25$) to prevent overfitting. For binary classification, we used the Sigmoid activation function and binary cross-entropy (BCE) loss function, as shown in Eq. (1) and Eq. (2).

$$\hat{y}_i = \frac{1}{1 + e^{-x}} \quad (1)$$

$$BCE = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (2)$$

Model optimization was performed using stochastic gradient descent (SGD) [31] with Nesterov’s momentum [32], defined in Eq. (3), set at a learning rate of 0.001 and Nesterov’s momentum of 0.9.

$$v_t = \gamma v_{t-1} + \eta \nabla_{\theta} J(\theta - \gamma v_{t-1}) \quad (3)$$

$$\theta = \theta - v_t$$

Where γ is usually set to 0.9.

Figures 4–7 provide an overview of the fine-tuning process applied to VGG-16, VGG-19, DenseNet-121, and Inception V3 models.

Layer (type)	Output Shape	Param #
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 256)	6422784
batch_normalization (Batch Normalization)	(None, 256)	1024
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 32)	8224
batch_normalization_1 (Batch Normalization)	(None, 32)	128
dropout_1 (Dropout)	(None, 32)	0
dense_2 (Dense)	(None, 2)	66
Total params: 21,146,914 Trainable params: 6,431,650 Non-trainable params: 14,715,264		

Fig. 4. Model summary of the proposed VGG-16 fine-tuning

Layer (type)	Output Shape	Param #
flatten (Flatten)	(None, 25088)	0
dense_3 (Dense)	(None, 256)	6422784
batch_normalization_2 (Batch Normalization)	(None, 256)	1024
dropout_2 (Dropout)	(None, 256)	0
dense_4 (Dense)	(None, 32)	8224
batch_normalization_3 (Batch Normalization)	(None, 32)	128
dropout_3 (Dropout)	(None, 32)	0
dense_5 (Dense)	(None, 2)	66
Total params: 26,456,610 Trainable params: 6,431,650 Non-trainable params: 20,024,960		

Fig. 5. Model summary of the proposed VGG-19 fine-tuning

Layer (type)	Output Shape	Param #	Connected to
flatten (Flatten)	(None, 50176)	0	['relu[0][0]']
dense (Dense)	(None, 256)	12845312	['flatten[0][0]']
batch_normalization (BatchNormalization)	(None, 256)	1024	['dense[0][0]']
dropout (Dropout)	(None, 256)	0	['batch_normalization[0][0]']
dense_1 (Dense)	(None, 32)	8224	['dropout[0][0]']
batch_normalization_1 (BatchNormalization)	(None, 32)	128	['dense_1[0][0]']
dropout_1 (Dropout)	(None, 32)	0	['batch_normalization_1[0][0]']
dense_2 (Dense)	(None, 2)	66	['dropout_1[0][0]']
=====			
Total params: 19,892,258			
Trainable params: 12,854,178			
Non-trainable params: 7,038,080			

Fig. 6. Model summary of the proposed DenseNet-121 fine-tuning

Layer (type)	Output Shape	Param #	Connected to
flatten (Flatten)	(None, 51200)	0	['mixed10[0][0]']
dense (Dense)	(None, 256)	13107456	['flatten[0][0]']
batch_normalization_94 (BatchNormalization)	(None, 256)	1024	['dense[0][0]']
dropout (Dropout)	(None, 256)	0	['batch_normalization_94[0][0]']
dense_1 (Dense)	(None, 32)	8224	['dropout[0][0]']
batch_normalization_95 (BatchNormalization)	(None, 32)	128	['dense_1[0][0]']
dropout_1 (Dropout)	(None, 32)	0	['batch_normalization_95[0][0]']
dense_2 (Dense)	(None, 2)	66	['dropout_1[0][0]']
=====			
Total params: 34,919,682			
Trainable params: 13,116,322			
Non-trainable params: 21,803,360			

Fig. 7. Model summary of the proposed Inception V3 fine-tuning

2.4 Model evaluation

In the realm of classification studies, the evaluation of deep learning models often relies on performance metrics derived from the confusion matrix. This matrix provides a visual representation of the relationship between the model's predicted class labels and the true labels of input images. We encounter four possible scenarios in class estimation, each carrying its own significance:

- True Positive (TP): This case arises when the deep learning classifier correctly predicts an image with a “cavity” label as having a cavity.
- False Positive (FP): In this situation, the classifier incorrectly predicts an image with “no_cavity” label as having a “cavity”.
- False Negative (FN): This scenario occurs when the classifier's deep learning model predicts an image with a “cavity” label as “no_cavity”.
- True Negative (TN): This case is observed when an image known to lack a “cavity” label is accurately predicted as “no_cavity” by the classifier's deep learning model.

By organizing TP, FP, TN, and FN values into a 2×2 matrix, we create what is commonly referred to as the confusion matrix. To quantify and standardize the model's performance further, we employ four performance metrics, represented in Eq. (4), Eq. (5), Eq. (6), and Eq. (7), including:

- Accuracy: this metric gauge the proportion of correct predictions (TP and TN) relative to the total number of predictions made.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

- Precision: evaluates the model's ability to make accurate positive predictions and is calculated as TP divided by the sum of TP and FP.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

- Recall: measures the model's capacity to identify all relevant instances and is computed as TP divided by the sum of TP and FN.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

- F1 Score: is the harmonic mean of precision and recall, offering a balanced measure of a model's performance.

$$F1 = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (7)$$

3 EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Dental dataset

The dataset employed in this research is sourced from Kaggle [33] and focuses on dental data. It consists of a combined total of 884 images, partitioned into 708 images

for training and an extra 176 images for testing purposes. Each individual image in this dataset is categorized into either “cavity” or “no_cavity”, representing the presence or absence of dental cavities, respectively. Table 1 provides the number of labeled images in both the training and testing subsets of the dataset. Figure 8 visually presents tooth samples labeled as either carious or non-carious. Data augmentation techniques have been employed to mitigate any potential drawbacks arising from the small number of images. The data augmentation procedure included the application of rotations within a range of ($\pm 10^\circ$) and a zoom of ($\pm 20\%$). These augmentation techniques were exclusively employed on the training set and were not applied to the test set.

Table 1. Distribution of images in training and tests sets

Phase	Cavity	No_Cavity
Train	389	319
Test	97	79
Total	486	398

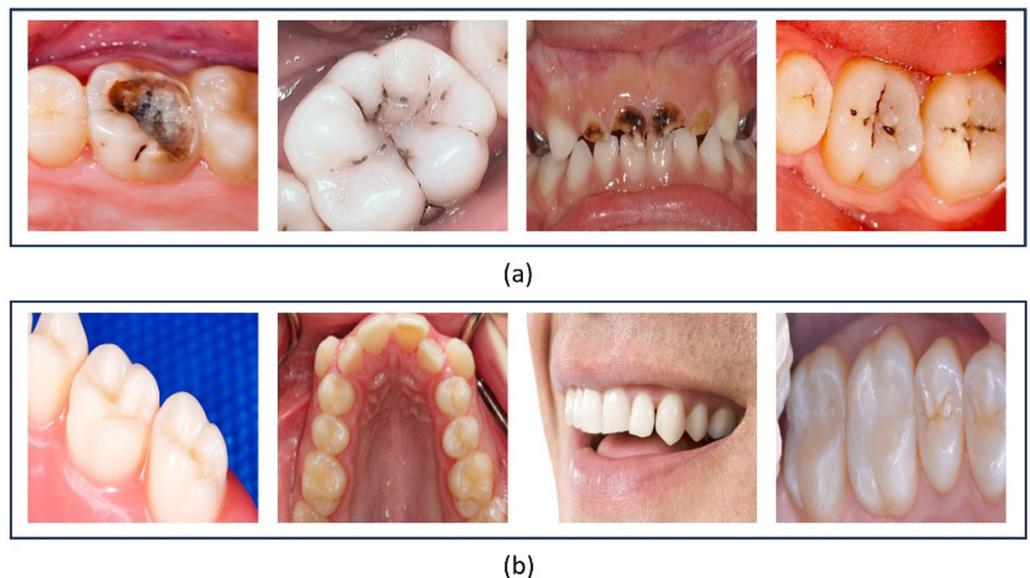


Fig. 8. Illustration of tooth samples utilized in our study – (a) teeth with cavities and (b) teeth without cavities

3.2 Experimental setup

The training process for the deep learning model involved using the VGG-16, VGG-19, DenseNet-121, and Inception V3 architectures. The input images were resized to dimensions of 224×224 pixels before being fed into the models. Sigmoid activation and binary cross-entropy loss functions were utilized for distinguishing between images with cavities and those without. The Python-based Keras library [34] was employed to facilitate the integration of models with varying layer counts into the training workflow. Instead of commencing training with randomly initialized weights, we opted to employ a process known as “fine-tuning”, utilizing pre-trained ImageNet weights as the initial foundation for our training models. It’s important to

highlight that the final layers of the models underwent a meticulous customization process during this fine-tuning procedure. These adapted final layers were methodically tailored to effectively differentiate between images featuring dental cavities and those that did not exhibit any signs of cavities.

The training was carried out over 100 epochs using the stochastic gradient descent (SGD) optimizer with Nesterov momentum. A learning rate of 0.001 and a batch size of 16 were chosen for the training process. The entire execution was conducted within the Google Colab environment, using GPUs to ensure consistent conditions for all four models. Throughout the training phases, model weights were tracked to capture the highest validation accuracy. Table 2 provides a summary of the optimal hyperparameters utilized during the training process. Following the training, various performance metrics were evaluated to assess the models' effectiveness, and the best-performing classifier model was identified based on these measures.

Table 2. Optimal hyperparameters for deep learning dental cavity detection model

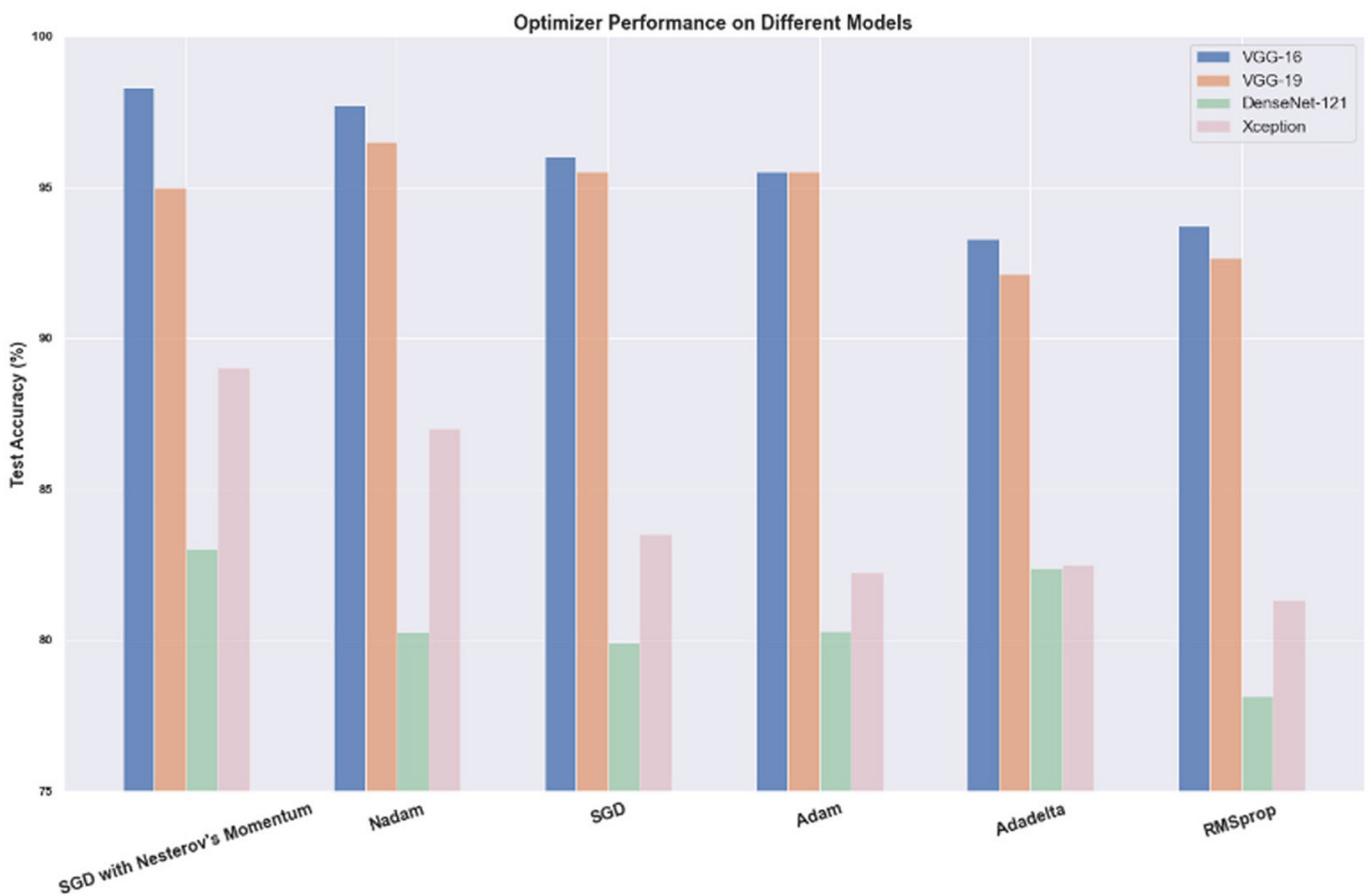
Hyperparameter	Value
Input image dimension	(224, 224, 3)
Activation function	Sigmoid
Loss function	Binary Cross-Entropy
Optimizer	SGD with Nesterov momentum
Learning rate	0.001
Momentum	0.9
Decay	1e-4
Batch size	16
Batch normalization	Yes
Number of epochs	100

3.3 Experimental results

We conducted a comprehensive evaluation of various optimizers across multiple models, including VGG-16, VGG-19, DenseNet-121, and Inception V3, as illustrated in Table 3 and depicted in Figure 9. Among the optimizers tested, SGD with Nesterov's Momentum consistently delivered outstanding results, emerging as the top performer. It achieved an impressive test accuracy of 98.3% for VGG-16, 95% for VGG-19, 83% for DenseNet-121, and 89% for Inception V3. Other optimizers also demonstrated strong performance. Nadam achieved accuracy rates of 97.7%, 96.5%, 80.25%, and 87% for the respective models. Additionally, SGD, Adam, Adadelta, and RMSprop showcased competitive results with varying degrees of accuracy across the models. These findings highlight the nuanced impact of optimizer selection on model accuracy and underscore SGD with Nesterov's Momentum as a standout choice for enhancing performance across the evaluated models. Considering these results, we will select SGD with Nesterov's Momentum as the preferred optimizer for cavity detection.

Table 3. Performance comparison of optimizers across different deep CNN models

Optimizer	Test Accuracy			
	VGG-16	VGG-19	DenseNet-121	Xception
SGD with Nesterov's Momentum	98.3%	95%	83%	89%
Nadam	97.7%	96.5%	80.25%	87%
SGD	96%	95.5%	79.9%	83.5%
Adam	95.5%	95.5%	80.3%	82.22%
Adadelta	93.3%	92.11%	82.36%	82.47%
RMSprop	93.7%	92.66%	78.14%	81.3%

**Fig. 9.** Comparison of optimizer performance on different deep CNN models

Then, we thoroughly assessed the performance of four distinct models: VGG-16, VGG-19, DenseNet-121, and Inception V3, employing a diverse range of performance metrics encompassing accuracy, precision, recall, and F1-score. This comprehensive evaluation is depicted in both Table 4 and Figure 10. Particularly noteworthy is the exceptional prowess of VGG-16, which emerges as the foremost performer, boasting an extraordinary accuracy rate of 98.3% using the SGD with Nesterov's Momentum optimizer. Notably, VGG-16 demonstrates precision, recall,

and F1-score values of 98.3% each, affirming its exceptional all-round performance. While the other models also exhibit robust performance, VGG-19 attains a solid accuracy of 95%, with corresponding precision, recall, and F1-score values of 95%, 94%, and 94%. Inception V3 maintains an impressive accuracy level of 89%, along with precision, recall, and F1-score values of 88%, 87%, and 88%, respectively. DenseNet-121 achieves a commendable 83% accuracy, and similar values for precision, recall, and F1-score. In this comparative analysis, VGG-16 distinctly showcases the highest accuracy, precision, recall, and F1-score using SGD with Nesterov’s Momentum, unequivocally designating it as the favored choice amongst these models.

Table 4. Performance metrics of different deep learning models using SGD with Nesterov’s momentum

Model	Accuracy	Precision	Recall	F1-score
VGG-16	98.3%	98.3%	98.3%	98.3%
VGG-19	95%	95%	94%	94%
DenseNet-121	83%	83%	83%	83%
Xception	89%	88%	87%	88%

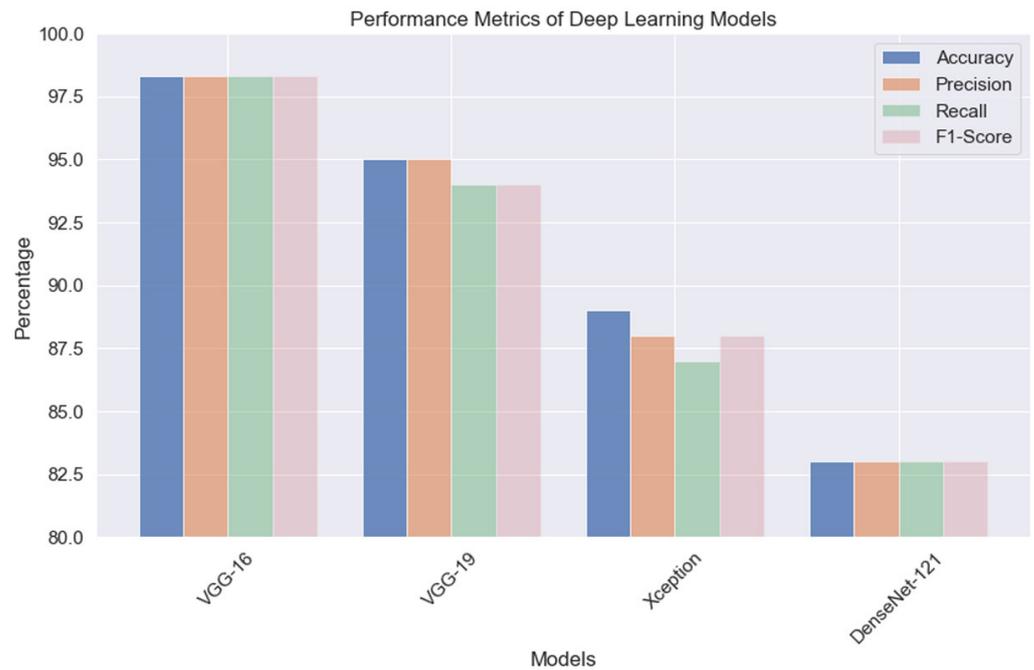


Fig. 10. Bar chart comparison of performance metrics for four deep learning models using SGD with Nesterov’s momentum

The normalized confusion matrix, as depicted in Figure 11, encapsulates the comprehensive performance evaluation of the VGG-16 model designed specifically for distinguishing between “cavity” and “no_cavity” images. Within this matrix, each value is represented as a percentage, offering an encompassing viewpoint into the model’s predictive abilities. It is worth noting that the model achieved an impressive overall accuracy of 98.3%, thereby amplifying the profound significance of the matrix. Noteworthy values within the matrix provide insightful nuances: for instance, the value of 100% in the top-left cell indicates that all instances predicted

as “cavity” were accurately classified as such. Furthermore, the value of 90.91% in the bottom-right cell underscores the high accuracy achieved in predicting “no_cavity” images. Intriguingly, the presence of 2.10% in the bottom-left cell signifies a relatively modest fraction of instances where “cavity” images were inadvertently misclassified as “no_cavity”.

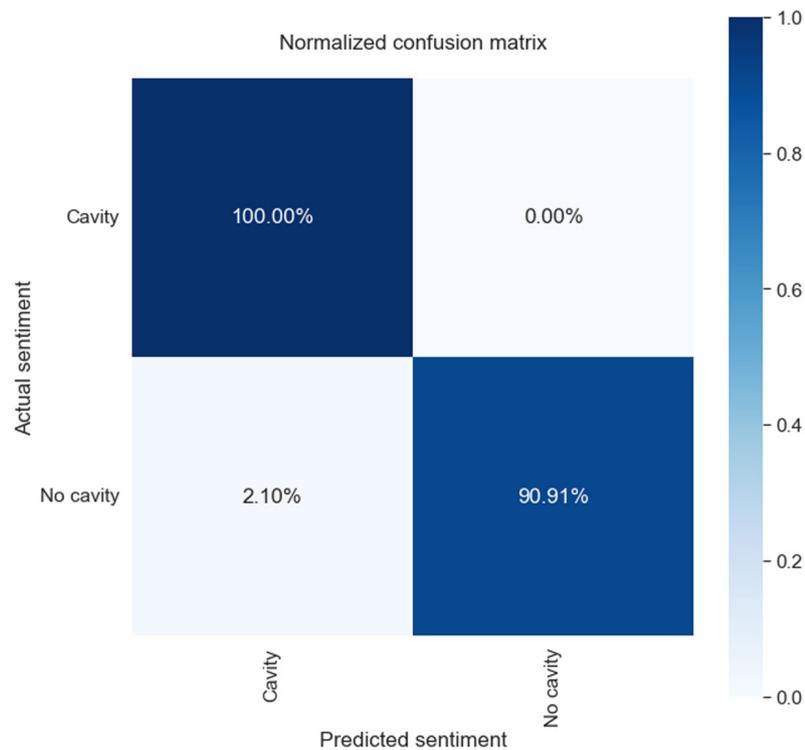


Fig. 11. Confusion matrix of the fine-tuned VGG-16 model

Figure 12 illustrates the learning curve of the VGG-16 model over the course of 100 epochs, showcasing the evolution of training and test accuracy in (a), as well as the corresponding training and test losses in (b).

Within Table 5, a comprehensive overview of various classification methodologies on dental datasets is provided. Each of these methods undergoes rigorous evaluation, encompassing parameters such as the total sample count, the number of distinct classes, and the accuracy achieved in classification. Notably, the ResNet-50 model, as implemented by [35], was applied to a robust dataset comprising 13870 samples, with a focus on 2 classes (Caries and Non-caries). Impressively, this approach garnered a commendable accuracy rate of 92%. The efficacy of the AlexNet architecture was scrutinized by [35], who conducted experiments on a dataset of 1900 samples categorized into 2 classes: Cavity and No cavity. The outcome was a remarkable accuracy of 96.08%. In a distinct approach, [36] leveraged a Convolutional Neural Network (CNN) to analyze a relatively compact dataset comprising 81 samples distributed across 3 distinct categories: Non-caries, Enamel, and Dentin. The meticulous application of this CNN yielded a substantial accuracy score of 90.75%. Finally, our novel proposed approach, utilizing the powerful VGG-16 architecture, underwent meticulous scrutiny on a dataset containing 884 samples, delineated into 2 categorical distinctions: Cavity and No_cavity. The results were striking, with our approach achieving the pinnacle of performance with an impressive accuracy of 98.3%.

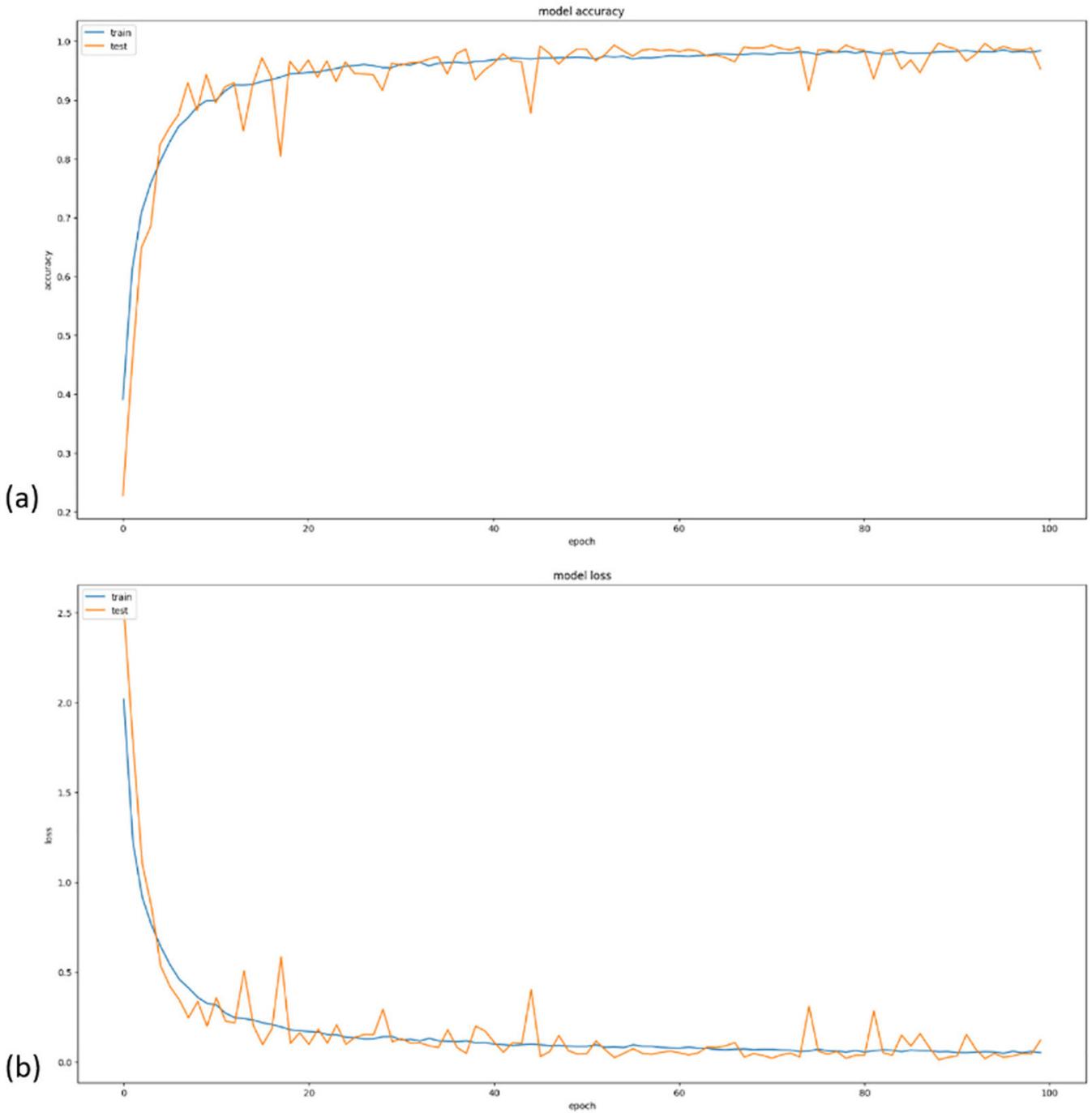


Fig. 12. Training and testing accuracy (a) and loss (b) of the VGG-16 model

Table 5. Comparison of our work with some state-of-the-art methods in the context of automated caries detection

Methods	Total of Samples	Number of Class	Accuracy
ResNet-50 [35]	13870	2 (Caries, Non-caries)	92%
AlexNet [35]	1900	2 (Cavity, No cavity)	96.08%
CNN [36]	81	3 (Non-caries, Enamel, Dentin)	90.75%
Our proposed model	884	2 (Cavity, No_cavity)	98.3%

4 CONCLUSION AND FUTURE WORK

In this research, an explainable deep learning-based method was introduced for automated dental caries detection using dental images. The study employed four well-established pre-trained deep CNN: VGG-16, VGG-19, DenseNet-121, and Inception V3 to distinguish between images with and without cavities. By integrating preprocessing techniques like histogram equalization and Sobel edge detection, the proposed approach achieved remarkable accuracy of 98.3% using the VGG-16 model with the stochastic gradient descent (SGD) optimizer and Nesterov's momentum. However, the main limitation of this work is that we used only 884 images, which could potentially be expanded for further validation. This method's success demonstrates its potential to revolutionize dental diagnostics, enabling early and accurate caries detection and offering insights into decision-making processes, ultimately leading to improved patient outcomes and enhancing overall dental health assessment. For future works, the methodology could be extended to include larger and more diverse datasets, encompassing variations in age and dental conditions. Furthermore, this deep learning approach could be adapted for the automated detection of skin cancer from dermatological images.

5 REFERENCES

- [1] National Institute of Dental and Craniofacial Research (NIDCR), "TMJ Disorders," 2015. Retrieved from <https://www.nidcr.nih.gov/sites/default/files/2017-12/tmj-disorders.pdf>. [Accessed: December 14, 2022].
- [2] N. B. Pitts, *et al.*, "Dental caries," *Nat. Rev. Dis. Prim.*, vol. 3, p. 17030, 2017. <https://doi.org/10.1038/nrdp.2017.30>
- [3] Centers for Disease Control and Prevention, "National Center for Health Statistics." Retrieved from <https://www.cdc.gov/nchs/>. [Accessed: December 15, 2022].
- [4] S. Michou, C. Vannahme, A. Bakhshandeh, K. R. Ekstrand, and A. R. Benetti, "Intraoral scanner featuring transillumination for proximal caries detection: An in vitro validation study on permanent posterior teeth," *J. Dent.*, vol. 116, p. 103841, 2022. <https://doi.org/10.1016/j.jdent.2021.103841>
- [5] J. Gomez, "Detection and diagnosis of the early caries lesion," *BMC Oral Health*, vol. 15, p. S3, 2015. <https://doi.org/10.1186/1472-6831-15-S1-S3>
- [6] S. Datta, N. Chaki, and B. Modak, "A novel technique to detect caries lesion using isophote concepts," *IRBM*, vol. 40, pp. 174–182, 2019. <https://doi.org/10.1016/j.irbm.2019.04.001>
- [7] J. Sinton, R. Wood, M. Pharoah, and D. Lewis, "Influence of the addition of restorations on the diagnosis of caries from digitized bitewing radiographs," *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.*, vol. 84, pp. 443–448, 1997. [https://doi.org/10.1016/S1079-2104\(97\)90046-9](https://doi.org/10.1016/S1079-2104(97)90046-9)
- [8] I. Lasri, A. Riadsolh, and M. Elbelkacemi, "Real-time Twitter sentiment analysis for Moroccan Universities using machine learning and big data technologies," *International Journal of Emerging Technologies in Learning (IJET)*, vol. 18, no. 5, pp. 42–61, 2023. <https://doi.org/10.3991/ijet.v18i05.35959>
- [9] M. Farhadian, P. Shokouhi, and P. Torkzaban, "A decision support system based on support vector machine for diagnosis of periodontal disease," *BMC Res. Notes*, vol. 13, no. 1, p. 337, 2020. <https://doi.org/10.1186/s13104-020-05180-5>
- [10] B. Thanathornwong, "Bayesian-based decision support system for assessing the needs for orthodontic treatment," *Healthc. Inform. Res.*, vol. 24, no. 1, pp. 22–28, 2018. <https://doi.org/10.4258/hir.2018.24.1.22>

- [11] D. Osterloh and S. Viriri, "Caries detection in non-standardized periapical dental X-rays," in *Computer Aided Intervention and Diagnostics in Clinical and Medical Images*, 2019, pp. 143–152. https://doi.org/10.1007/978-3-030-04061-1_14
- [12] J. Oliveira and H. Proença, "Caries detection in panoramic dental x-ray images," in *Computational Vision and Medical Image Processing*, 2011, pp. 175–190. https://doi.org/10.1007/978-94-007-0011-6_10
- [13] S. V. Tikhe, A. M. Naik, S. D. Bhide, T. Saravanan, and K. Kaliyamurthie, "Algorithm to identify enamel caries and interproximal caries using dental digital radiographs," in *Proceedings of the 2016 IEEE 6th International Conference on Advanced Computing (IACC)*, Bhimavaram, India, 2016, pp. 225–228. <https://doi.org/10.1109/IACC.2016.50>
- [14] I. Lasri, A. Riadsolh, and M. Elbelkacemi, "Facial emotion recognition of deaf and hard-of-hearing students for engagement detection using deep learning," *Education and Information Technologies*, vol. 28, no. 4, pp. 4069–4092, 2023. <https://doi.org/10.1007/s10639-022-11370-4>
- [15] I. Lasri, A. Riadsolh, and M. Elbelkacemi, "Self-attention-based Bi-LSTM model for sentiment analysis on tweets about distance learning in higher education," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 12, pp. 119–141, 2023. <https://doi.org/10.3991/ijet.v18i12.38071>
- [16] M. Al-Nawashi, O. M. Al-Hazaimeh, and M. Saraee, "A novel framework for intelligent surveillance system based on abnormal human activity detection in academic environments," *Neural Computing & Applications*, vol. 28, no. Suppl 1, pp. 565–572, 2017. <https://doi.org/10.1007/s00521-016-2363-z>
- [17] G. Litjens, et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017. <https://doi.org/10.1016/j.media.2017.07.005>
- [18] I. Lasri, A. Riadsolh, and M. El Belkacemi, "Toward an effective analysis of COVID-19 Moroccan business survey data using machine learning techniques," in *Proceedings of the 13th International Conference on Machine Learning and Computing (ICMLC 2021)*, Association for Computing Machinery, New York, NY, USA, 2021, pp. 50–58. <https://doi.org/10.1145/3457682.3457690>
- [19] O. M. Al-Hazaimeh, M. Al-Nawashi, and M. Saraee, "Geometrical-based approach for robust human image detection," *Multimedia Tools and Applications*, vol. 78, pp. 7029–7053, 2019. <https://doi.org/10.1007/s11042-018-6401-y>
- [20] N. Gharaibeh, A. A. Abu-Ein, O. M. Al-hazaimeh, K. M. Nahar, W. A. Abu-Ain, and M. M. Al-Nawashi, "Swin transformer-based segmentation and multi-scale feature pyramid fusion module for Alzheimer's disease with machine learning," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 19, no. 4, pp. 22–50, 2023. <https://doi.org/10.3991/ijoe.v19i04.37677>
- [21] O. M. Al-hazaimeh, A. Abu-Ein, N. Tahat, M. Al-Smadi, and M. Al-Nawashi, "Combining artificial intelligence and image processing for diagnosing diabetic retinopathy in retinal fundus images," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 18, no. 13, pp. 131–151, 2022. <https://doi.org/10.3991/ijoe.v18i13.33985>
- [22] A. Malhi, R. Grewal, and H. S. Pannu, "Detection and diabetic retinopathy grading using digital retinal images," *Int J Intell Robot Appl*, vol. 7, pp. 426–458, 2023. <https://doi.org/10.1007/s41315-022-00269-5>
- [23] D. Bychkov, et al., "Deep learning based tissue analysis predicts outcome in colorectal cancer," *Sci. Rep.*, vol. 8, no. 1, p. 3395, 2018. <https://doi.org/10.1038/s41598-018-21758-3>
- [24] F. Oztekin, et al., "An explainable deep learning model to prediction dental caries using panoramic radiograph images," *Diagnostics*, vol. 13, p. 226, 2023. <https://doi.org/10.3390/diagnostics13020226>

- [25] R. S. Rao, *et al.*, “Ensemble deep-learning-based prognostic and prediction for recurrence of sporadic odontogenic keratocysts on hematoxylin and eosin stained pathological images of incisional biopsies,” *Journal of Personalized Medicine*, vol. 12, no. 8, p. 1220, 2022. <https://doi.org/10.3390/jpm12081220>
- [26] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv:1409.1556, 2015.
- [27] G. Huang, Z. Liu, L. Van Der Maaten, and K. Weinberger, “Densely connected convolutional networks,” in *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>
- [28] C. Szegedy, *et al.*, “Going deeper with convolutions,” in *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, 2015, pp. 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [29] N. Kanopoulos, N. Vasanthavada, and R. L. Baker, “Design of an image edge detection filter using the Sobel operator,” *IEEE Journal of Solid-State Circuits*, vol. 23, no. 2, pp. 358–367, 1988. <https://doi.org/10.1109/4.996>
- [30] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>
- [31] H. Robbins and S. Monro, “A stochastic approximation method,” *Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, 1951. <https://doi.org/10.1214/aoms/1177729586>
- [32] Y. Nesterov, “A method of solving a convex programming problem with convergence rate $o(1/k^2)$,” *Soviet Mathematics Doklady*, vol. 27, no. 2, pp. 372–376, 1983.
- [33] Dental dataset. <https://www.kaggle.com/datasets/nabeel1921/dental-cavity>. [Accessed: June 15, 2023].
- [34] F. Chollet, *et al.*, Keras, 2015. <https://github.com/fchollet/keras>
- [35] M. M. Lakshmi and P. Chitra, “Classification of dental cavities from X-ray images using Deep CNN algorithm,” in *Proceedings of the 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, 2020, pp. 774–779. <https://doi.org/10.1109/ICOEI48184.2020.9143013>
- [36] H. S. Salehi, M. Barchini, Q. Chen, and M. Mahdian, “Toward development of an automated grading system for carious lesions classification using deep learning and OCT imaging,” in *Proceedings of the Medical Imaging 2021: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 2021, pp. 245–252. <https://doi.org/10.1117/12.2581318>

6 AUTHORS

Imane Lasri is a PhD student at the Laboratory of Conception and Systems (Electronics, Signals, and Informatics), Faculty of Sciences Rabat, University Mohammed V in Rabat, Morocco. She received a Master’s degree in Big Data Engineering from the Faculty of Sciences Rabat. She received the awards of excellence of the major winner from the Mohammed V University in Rabat in 2019. Her current field of research is pattern recognition applied to higher education using machine learning and deep learning algorithms (E-mail: imane_lasri@um5.ac.ma).

Naoufal El-Marzouki is a PhD student at the Faculty of Sciences Rabat, University Mohammed V in Rabat, Morocco. He is a member of the Laboratory of Conception and Systems (Electronics, Signals and Informatics). He received a Master’s degree in Mathematics from the Faculty of Sciences Kenitra. His current field of research is artificial intelligence applied to education (E-mail: naoufal_elmarzouki@um5.ac.ma).

Anouar Riadsolh is a Professor at the Faculty of Sciences Rabat, University Mohammed V in Rabat, Morocco. He received his PhD in Computer Science from the Faculty of Sciences Rabat. He is a member of the Laboratory of Conception and Systems (Electronics, Signals, and Informatics). His current research interests are focused on data mining, big data, and machine learning (E-mail: a.riadsolh@um5r.ac.ma).

Mourad Elbelkacemi is a Professor at the Faculty of Sciences Rabat, University Mohammed V in Rabat, Morocco. He received his PhD in Computer Science. He is a member of the Laboratory of Conception and Systems (Electronics, Signals, and Informatics). His main research interests are focused on electronics, education, and data mining (E-mail: mourad_prof@yahoo.fr).