Original paper

Digital recognition of color alteration in gingiva using a convolutional neural network

Ali Raad AbdulAzeez[1,A–D], Hashim Mueen Hussein[2,E,F], Mahmood Ali Jumaah[3,A–D], Athraa Ali Mahmood[4,E,F]

[1] Department of Periodontics, College of Dentistry, University of Al-Mashreq, Baghdad, Iraq
[2] Department of Conservative Dentistry, College of Dentistry, Mustansiriyah University, Baghdad, Iraq
[3] Department of Artificial Intelligence, College of Computer Science, University of Technology, Baghdad, Iraq
[4] Department of Periodontics, College of Dentistry, Mustansiriyah University, Baghdad, Iraq

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Dental and Medical Problems, ISSN 1644-387X (print), ISSN 2300-9020 (online) *Dent Med Probl*. 2026;63(3):655–665

Address for correspondence
Hashim Mueen Hussein
E-mail: hashimmueenhussein@uomustansiriyah.edu.iq

Funding sources
None declared

Conflict of interest
None declared

Acknowledgements
The authors would like to thank the University of Al-Mashreq, Baghdad, Iraq, and the Dean of the College of Dentistry, Dr. Omar A. Abdulqader. Additionally, the authors would like to express their gratitude to the College of Dentistry at Mustansiriyah University, Baghdad, Iraq.

Received on June 2, 2024
Reviewed on July 8, 2024
Accepted on July 11, 2024

Published online on June 30, 2026

Cite as
AbdulAzeez AR, Hussein HM, Jumaah MA, Mahmood AA. Digital recognition of color alteration in gingiva using a convolutional neural network. *Dent Med Probl.* 2026;63(3):655–665. doi:10.17219/dmp/191092

DOI
10.17219/dmp/191092

Abstract

Background. Gingivitis is a common condition in individuals with inadequate oral hygiene. Although dental indices are widely used to assess periodontal status and educate patients, digital technologies have recently been incorporated into this field.

Objectives. The aim of the study was to provide an objective digital method to resolve the ambiguity between the first two scores of the gingival index (GI): absence of color alteration (score 0); and presence of color alteration in the gingiva (score 1), by designing a convolutional neural network (CNN) algorithm capable of detecting gingival color changes with high accuracy.

Material and methods. In this cross-sectional study, 10 CNN models were developed and trained to distinguish between normal and abnormal gingival color. A total of 6,660 augmented, pre-classified frontal-view images of the patients' gingiva were included. The dataset was divided into a training set (4,640 images; 70%) to train the CNN models and a test set (2,020 images; 30%) to test them. All images were previously classified by periodontists as representing either normal (0) or abnormal (1) gingival color. During training, the CNN models learned to classify gingival color based on these reference labels. Model performance was subsequently evaluated in the test phase by comparing the models' predictions with the reference diagnoses provided by periodontists.

Results. To evaluate inter-rater reliability, model performance was examined using Cohen's kappa coefficient and classification accuracy relative to the reference diagnoses established by the periodontists. Model 4 demonstrated excellent inter-rater reliability ($\kappa = 1.00$) and achieved 99.9% accuracy, followed by Model 8 (99.7%) and Model 5 (99.3%). Models 7 and 3 showed lower accuracies of 81.1% and 88.1%, respectively. All models demonstrated statistically significant agreement with the reference diagnoses ($p < 0.001$).

Conclusions. Differentiation between GI scores of 0 and 1 using CNNs provides an objective method for detecting gingival color alterations. This approach has potential for improving patients' awareness of periodontal health, enhancing motivation to maintain oral hygiene, and encouraging dental visits.

Keywords: artificial intelligence, convolutional neural network, gingivitis, digital dentistry, gingival index

Highlights

- A convolutional neural network was trained to detect gingival color changes and distinguish normal from abnormal gingival color.
- The artificial intelligence (AI) model focused on gingival color assessment and avoided providing comprehensive periodontal diagnosis.
- The convolutional neural network demonstrated high diagnostic accuracy in identifying gingival color changes when compared with the reference diagnoses made by periodontists.
- The findings support further validation to improve AI performance before clinical implementation.

Introduction

Periodontal diseases are the 11th most prevalent group of diseases worldwide, affecting an estimated 20–50% of the global population.[1] Gingivitis is an inflammatory condition induced by bacterial plaque and/or bacterial toxins that is confined to the soft tissues of the periodontium. It is a mild and reversible form of periodontal disease that may progress, if left untreated, to periodontitis,[2,3] an inflammatory disease affecting both soft and hard tissues of the tooth-supporting apparatus and resulting in irreversible periodontal attachment loss.[4]

Clinical indices are standardized objective methods designed to evaluate dental and periodontal conditions.[5] Periodontal indices have evolved considerably over the years, from early forms such as Russel's periodontal index (1956) to more sophisticated and relatively recent indices, including the Genetic Susceptibility Index (GSI) introduced in 2007. The purpose of periodontal indices is to improve the diagnosis and assessment of periodontal diseases as well as to facilitate the development of appropriate preventive and treatment strategies.[6]

Some periodontal indices are designed to evaluate specific clinical conditions, such as the gingival index (GI),[7] whereas others assess oral hygiene status, such as the plaque index (PI).[5] More comprehensive indices have been developed to evaluate multiple periodontal conditions and estimate community treatment needs, such as the community periodontal index of treatment needs (CPITN), introduced by the World Health Organization (WHO) and assessed using the WHO periodontal probe.[8] The oral hygiene index (OHI) is used to examine overall oral hygiene status using numerical scores ranging from 0 to 3. It is simple, reproducible and reliable; however, it has some limitations. The index does not consider plaque and does not provide a substitution for missing teeth as it divides the oral cavity into sextants. Other indices focus on measuring periodontal destruction, including probing pocket depth (PPD) and clinical attachment loss (CAL). Although these indices provide valuable quantitative and qualitative information, they are more technique-sensitive, require greater examiner experience, and may show lower reproducibility than simpler clinical indices.

A good periodontal index should satisfy several criteria: (1) clarity, allowing easy interpretation and facilitating clinical decision-making; (2) simplicity; (3) objectivity, minimizing examiner-dependent variation; (4) validity, ensuring that the index measures the intended clinical condition; (5) reliability and reproducibility, providing consistent results when the clinical condition remains unchanged; (6) quantifiability, allowing statistical analysis; (7) sensitivity, enabling the detection of subtle clinical changes; and (8) acceptability, ensuring that the examination is well tolerated by patients. The strengths and limitations of periodontal indices are determined by the extent to which they satisfy these criteria.[9]

The main purpose of the GI is to assess the gingival condition, including the quality of the gingiva, and to identify the location of the lesion (buccal, mesial, distal, and lingual surfaces). The GI is non-invasive, atraumatic and easy to perform. However, it is inherently subjective and does not assess PPD, degree of bone loss or other quantitative indicators of the periodontium.[7]

The integration of digital technologies into dentistry has improved both clinical examination and treatment procedures.[10] Devices such as digital X-ray sensors, computed tomography (CT) scanners, microscopes, and loupes[11] have enhanced diagnostic precision.[12] More recently, smartphone applications have emerged as adjunctive tools in the process.[13]

Pain is one of the primary reasons for patients to seek dental care.[14,15] However, pain is generally absent during early stages of gingival disease, including gingivitis and peri-implant mucositis.[16] Consequently, particularly in countries without mandatory dental insurance systems[17] or among individuals lacking dental insurance coverage, periodontal disease may remain unnoticed until more advanced stages develop.[18]

Smartphone applications have transformed many aspects of everyday life, including advertising,[19] office applications, tools, and healthcare.[20] In dentistry, mobile applications have the potential to improve dental awareness and bridge the gap between patients and dentists.[21] Because many individuals postpone routine dental visits in the absence of pain or other obvious symptoms, user-friendly mobile applications capable of providing an initial assessment of gingival and periodontal status may encourage earlier professional evaluation.

However, digital diagnostic systems should be regarded as decision-support tools rather than replacements for professional clinical judgement.[22] The final diagnosis and treatment decision should always be made by qualified dental professionals.[23]

Online information and self-assessment applications may provide misleading interpretations of oral conditions,[24] potentially leading patients to overestimate disease severity or seek unnecessary treatment.[25]

Pain in the oral cavity or dental pain has multiple etiologies, with pulpitis and periodontal pain representing common causes.[26,27] However, patients' perception of pain, dissatisfaction and awareness is influenced by psychological factors.[28] The digital system developed in the present study was designed to provide users with a preliminary assessment of whether professional periodontal evaluation is advisable. At the same time, the system aims to reduce the ambiguity between GI scores of 0 and 1 without attempting to identify the exact location of inflammation or an opinion about the teeth's color. This approach preserves the essential role of the dentist in establishing the final diagnosis and treatment plan while reducing the likelihood of inappropriate self-diagnosis.

Artificial intelligence (AI) has become increasingly integrated into different fields,[29] especially medicine.[30] AI-based methods support decision-making and facilitate solving complex real-world problems.[31] Convolutional neural networks (CNNs) are artificial neural networks[32,33] designed for image analysis. Convolutional neural networks play an important role in computer-aided diagnosis (CAD)[34] by accurately detecting objects and classifying them by processing the patient's raw data.[35] In the present study, the CNN performed 2 principal tasks: object detection, which localizes the region of interest within an image; and image classification, which assigns the detected region to a predefined class using supervised learning.[36,37]

The aim of this study was to develop and evaluate a CNN capable of distinguishing between normal and abnormal gingival colors using standardized frontal photographs that had been previously annotated by experienced periodontists.

Material and methods

The study was conducted in accordance with the principles of the Declaration of Helsinki.[38] Written informed consent was obtained from all participants before enrollment. The study protocol was reviewed and approved by the Ethics Committee of the College of Dentistry, University of Al-Mashreq, Baghdad, Iraq (order No. 245, 2022).

Patient selection, grouping and exclusion criteria

Patients attending the Department of Periodontology at the College of Dentistry, primary healthcare centers, and specialized dental centers in Baghdad, Iraq, between September 2021 and October 2022 were screened for eligibility. Clinical examinations were performed by a team of 3 periodontists using direct visual inspection and dental mirrors to distinguish healthy gingiva from gingivitis. The final diagnosis was established after discussion and consensus among the 3 examiners. Demographic characteristics, including race and sex were not considered as selection criteria.

The exclusion criteria were as follows: mentally challenged patients; terminally ill patients; children aged <12 years; patients with traumatic gingival wounds or non-inflammatory gingival bleeding; and patients with drug-induced gingivitis.

To maximize the diversity of gingival color patterns at different stages of gingivitis, a large image dataset was collected. A larger and more diverse dataset enables the CNN to learn a wider range of normal and abnormal gingival color variations, thereby improving its ability to recognize subtle color changes in previously unseen images. Furthermore, including participants of different ages, sexes and ethnic backgrounds increases the reliability of the model while reducing the risk of overfitting and over-specialization.[39]

The image dataset was divided into training and test sets using an approximately 70:30 split (4,640 and 2,020 images, respectively), which is widely accepted for machine learning applications while allowing flexibility based on the complexity, size and type of the dataset.[40]

The training dataset consisted of 4,640 augmented images:
- group A – 1,950 images of clinically healthy gingiva;
- group B – 2,690 images of abnormally colored gingiva.

The test dataset consisted of 2,020 images:
- group A: 850 images of clinically healthy gingiva;
- group B: 1,170 images of abnormally colored gingiva.

After model training, all images from the test dataset were analyzed using the trained CNN models. The model predictions were compared with the reference diagnoses established by the periodontists to determine classification accuracy and reliability (Fig. 1).

Materials

The following equipment and software were used during the study:
- digital camera (Nikon D3200; Nikon Corporation, Tokyo, Japan);
- plastic lip retractor;
- dental mirror and WHO/CPITN periodontal probe (Hu-Friedy, Chicago, USA);

- cotton pellets and patient towels;
- class B autoclave sterilizer (Woson, Ningbo, China);
- 70% alcohol for cold sterilization;
- laptop computer (MSI GL63 8RD; MSI, New Taipei City, Taiwan) equipped with an Intel® Core™ i7-8750H processor, 16 GB RAM, and an NVIDIA GP107M (GeForce GTX 1050 Ti Mobile)/Mesa Intel® UHD Graphics 630 (CFL GT2) graphics card;
- Python3 programming language with TensorFlow 2.8.0.

Gingival index

The gingival index was used to examine and classify gingival condition. The present study focused primarily on the first two GI scores. The index was first introduced by Löe,[41] with the scores defined as follows:

- 0 – normal gingiva with no discoloration or redness and no bleeding on probing;
- 1 – reddened gingiva characterized by changes in color and/or texture without bleeding on probing;
- 2 – redness, edema, glazing, and slight bleeding on probing;
- 3 – marked redness and edema/hypertrophy, ulceration, amorphous/continuous bleeding on probing, or spontaneous bleeding.

Examiner calibration

To improve diagnostic consistency, the 3 periodontists underwent examiner calibration before the study. Clinical assessments were performed independently and in a blinded manner. The final diagnosis for each participant was accepted only when at least 2 of the 3 examiners reached agreement regarding the GI score.

Image acquisition and sample classification

Frontal intraoral photographs were taken using a digital camera (Nikon D3200; Nikon Corporation) fitted with an 18–55 mm lens. A plastic lip retractor was used to expose the gingiva, and the camera was focused on the contact point between the upper and lower central incisors. No chair light or camera flash was used while taking the photographs in order to preserve the natural appearance of gingival color. All patients were seated at the same location in the examination room. Illumination was provided exclusively by natural daylight entering through a wide window (2 m × 1 m) located 2 m in front of the participant. Photographs were acquired between 9 am and 11 am. For standardization, the camera was mounted on a tripod at a fixed distance of 30 cm from the patient's face, and no background plate was used.

Following image acquisition, all photographs were transferred to a computer and divided into 2 classes based on the periodontists' decision: normal gingival color (group A, GI score 0); or abnormal gingival color (group B, GI score ≥1).

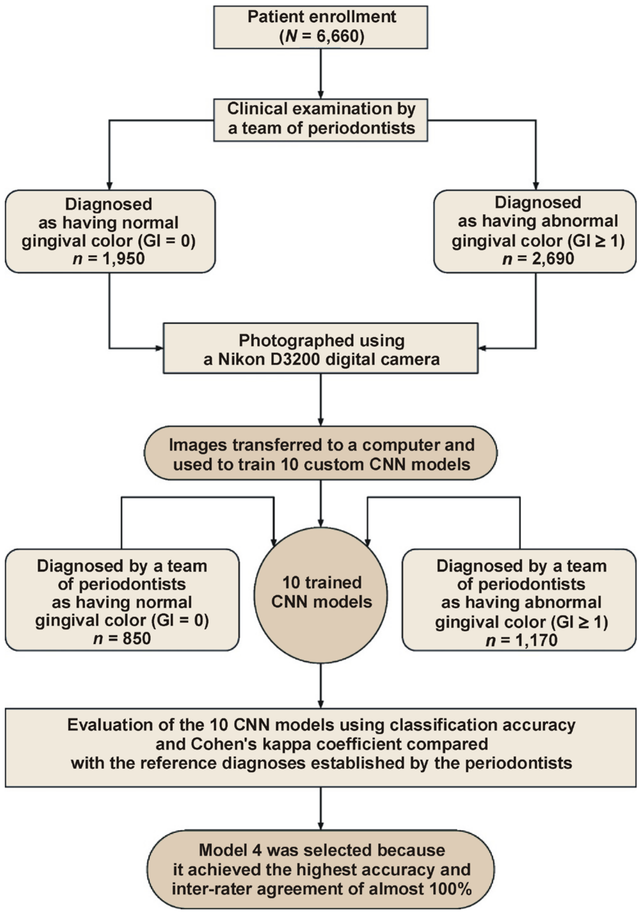

Fig. 1. Flowchart of the study design

CNN – convolutional neural network.

Image preprocessing

Raw images were resized to 1,024 × 768 pixels and stored in JPEG format using the RGB color space. Because the objective of the study was to analyze gingival color, the surrounding area (jaw, nose, lips, background) was excluded from further analysis. The gingival region was annotated using the LabelImg image annotation software.[42] Each image was manually labeled with a bounding box encompassing the gingival tissues. Subsequently, the gingival regions were detected and extracted using the Single Shot MultiBox Detector (SSD), a CNN-based model specifically developed for efficient object detection. The SSD model identified the gingival region and produced cropped images for subsequent classification.

Gingival detection using the Single Shot MultiBox Detector

The primary challenge of this research was the accurate extraction of the gingival region from images. Object detection techniques enable automatic localization of structures of interest within digital images. The SSD is a superior CNN-based method employed in object detection with lower computational demands. The Single Shot MultiBox

Detector processes an image in a single pass, using feature maps at various scales to predict object boundaries and categories. During training, the model requires only the input images and their corresponding ground-truth boxes for each object, simplifying the learning process while maintaining high detection accuracy. Consequently, SSD was selected for application in the present study (Fig. 2,3).

Image classification

Image classification involves assigning input images to predefined categories or classes using machine learning algorithms.[43] In the present study, the CNN was trained to classify gingival images into one of the 2 categories: normal gingival color or abnormal gingival color. When presented with new data, the trained network analyzed the learned image features and assigned each image the most probable class based on its similarity to the training data. This ability to automatically extract relevant image features is one of the principal advantages of CNNs and contributes to their superior performance.

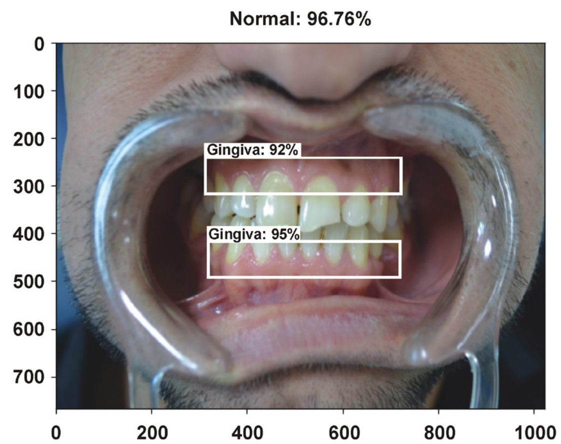

Fig. 2. Gingiva classified by the convolutional neural network (CNN) as having normal gingival color

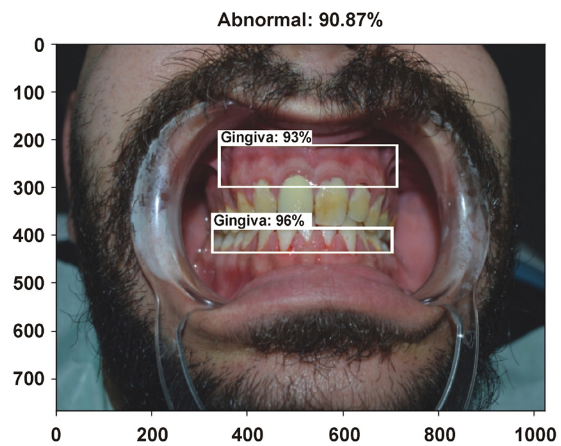

Fig. 3. Gingiva classified by the convolutional neural network (CNN) as having abnormal gingival color

Convolutional neural network

A convolutional neural network, also referred to as ConvNet, is a type of artificial neural network (ANN) specifically designed for image analysis. Unlike traditional fully-connected (FC) neural networks, CNNs employ a deep feed-forward structure and exhibit superior generalization performance. They can learn complex, abstract features of objects, particularly those that involve spatial data, and recognize them with high efficiency. A CNN consists of multiple layers, each serving a specific function (Fig. 4). The network begins with an input layer that receives the image data. In our study, the input images had dimensions of 224 px × 224 px × 3 px. The input layer was connected to a convolutional layer, which was responsible for extracting relevant image features, including color, texture and softness of the gingiva. Pooling layers were subsequently applied to reduce the number of samples. Furthermore, the FC layer was used to classify the features generated from the pooling layer. Finally, the output layer generated a prediction, classifying each image as either normal gingiva (0) or abnormal gingiva (1).

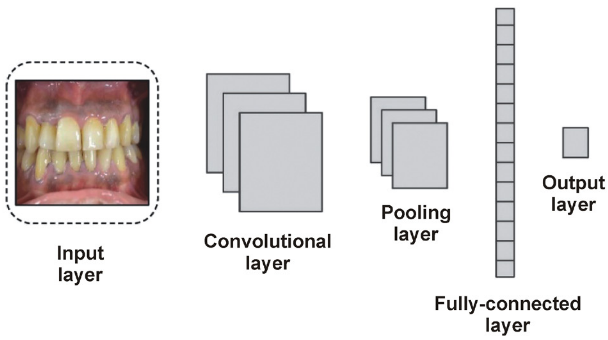

Fig. 4. Architecture of the convolutional neural network (CNN)

Development and training of the CNN models

Custom CNN models were developed specifically for the present study rather than employing standard deep learning models[44] such as Visual Geometry Group (VGG) or Residual Network (ResNet), which contain a large number of trainable parameters and require longer training time.[45] To optimize computational efficiency, 10 custom CNN models were designed and evaluated using the present dataset. These models provide maximum accuracy in the shortest amount of time and with the least number of parameters.

The dataset was divided into training (70%) and test (30%) subsets. The training process was iterated multiple times to facilitate learning by adjusting the weight and enabling the system to distinguish between normal and abnormal cases. The training was repeated several times until the model reached a state of stability.

Prevention of overfitting and underfitting

Several strategies were implemented to reduce the risk of overfitting during model development. First, the CNNs were trained using a relatively large dataset comprising approx. 6,660 images, which represented a diverse sample size. Second, data augmentation techniques were applied to increase the diversity of the training set and improve model generalizability. Finally, model performance was continuously monitored on both the training and test datasets. The alignment between the metrics indicated that the models showed no evidence of substantial overfitting or underfitting and maintained strong predictive capabilities of the test dataset.

Model evaluation

After completion of the training, the CNN models were tested using a dataset of 2,020 images. The predicted classifications were compared with the reference diagnoses established by periodontists at the Department of Periodontology, using the following formula:

$$accuracy = \frac{number\ \ of\ \ correct\ \ predictions}{total\ \ number\ \ of\ \ predictions}$$

Statistical analysis

Statistical analyses were performed using the IBM SPSS Statistics for Windows software, v. 26.0 (IBM Corp., Armonk, USA). Cohen's kappa coefficient was calculated to evaluate the agreement between the CNN classifications and the reference diagnoses (inter-rater reliability/agreement test).

According to Cohen, kappa values are interpreted as follows[46]:

- ≤0.00: no agreement;
- 0.01–0.20: poor to slight agreement;
- 0.21–0.40: fair agreement;
- 0.41–0.60: moderate agreement;
- 0.61–0.80: substantial agreement;
- 0.81–1.00: almost perfect agreement.

Results

Descriptive statistics of the training dataset

A total of 4,640 images were included in the training dataset and used to train the CNN models according to the reference diagnoses established by the 3 periodontists. Of these, 1,950 images were classified as healthy gingiva (GI score 0), whereas the remaining 2,690 images were classified as abnormal gingiva (GI score ≥1). The standard deviation of the test dataset was 0.493654. The middle quartile (Q2) value was 1.00, and the interquartile range (*IQR*) was 1.00 (Table 1).

Table 1. Descriptive statistics of the training dataset

Variable		Result
Valid samples, *n*		4,640
Missing samples, *n*		0
SD		0.49
Quartile	Q1	0.00
	Q2	1.00
	Q3	1.00

SD – standard deviation.

Descriptive statistics of the test dataset

The test dataset consisted of 2,020 images used to evaluate the trained CNN models. Based on the reference diagnoses established by the periodontists, 850 images were classified as normal gingiva and 1,170 images as abnormal gingiva (dark or abnormal color). The standard deviation of the test dataset was 0.493808. The middle quartile was 1.00, and the *IQR* was 1.00 (Table 2).

Table 2. Descriptive statistics of the test dataset

Variable		Result
Valid samples, *n*		2,020
SD		0.49
Quartile	Q1	0.00
	Q2	1.00
	Q3	1.00

Performance of the CNN models

The test dataset comprising 2,020 images was evaluated using 10 different CNN models. Then, the accuracy was determined, and the interrater reliability of every model was tested by calculating Cohen's kappa coefficient (Fig. 5).

The classification accuracy, Cohen's kappa coefficient and corresponding p-values for each model are presented in Table 3. Among the evaluated models, Model 4 achieved the highest accuracy (99.9%). The lowest accuracy was observed for Model 7 (81.1%), whereas Model 3 achieved an accuracy of 88.1%. All remaining models demonstrated classification accuracy exceeding 90%.

Model 4 exhibited perfect agreement with the reference diagnoses (κ = 1.00; p < 0.001). The remaining models

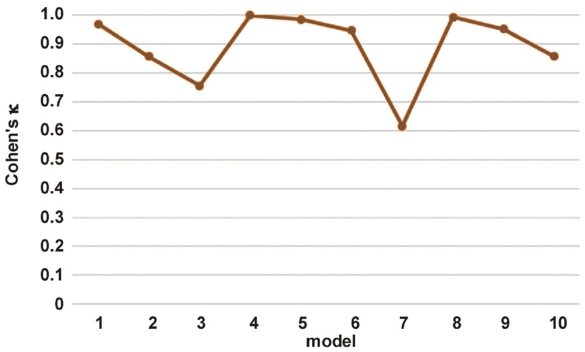

Fig. 5. Cohen's kappa coefficients for the agreement between the 10 convolutional neural network (CNN) models and the reference diagnoses established by the periodontists

achieved Cohen's kappa values ranging from 0.86 to 0.99, indicating almost perfect agreement with the periodontists' classifications. Model 7 was an exception (κ = 0.62), lying on the lower border of what is known as substantial agreement. Model 3 achieved a kappa value of 0.76, which also indicates substantial agreement. Despite their comparatively lower performance, both models remained statistically significant ($p < 0.001$).

Table 3. Classification accuracy, Cohen's kappa coefficients and *p*-values for the 10 convolutional neural network (CNN) models

Model	Accuracy	Cohen's κ	*p*-value
1	98.5%	0.97	<0.001*
2	93.1%	0.86	<0.001*
3	88.1%	0.76	<0.001*
4	99.99%	1.00	<0.001*
5	99.3%	0.99	<0.001*
6	97.5%	0.95	<0.001*
7	81.1%	0.62	<0.001*
8	99.7%	0.99	<0.001*
9	97.8%	0.95	<0.001*
10	93.1%	0.86	<0.001*

* statistically significant ($p < 0.05$).

Figures 6 and 7 show the comparison between the training curve pattern (expected pattern) and the test curve pattern during model development. Among the evaluated models, Model 4 demonstrated an almost perfect simulation of the expected curve pattern.

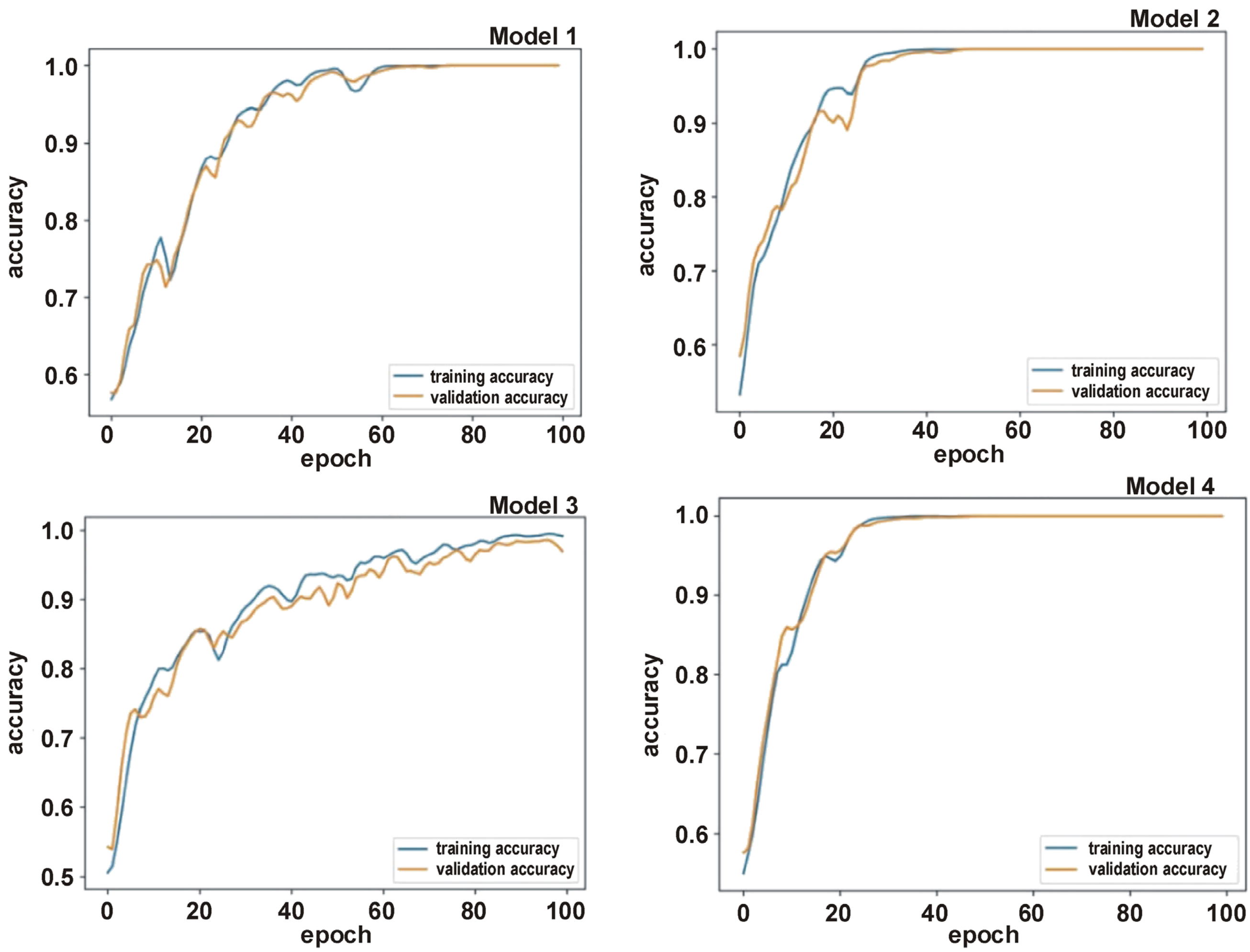

Fig. 6. Training and validation accuracy curves for Models 1–4

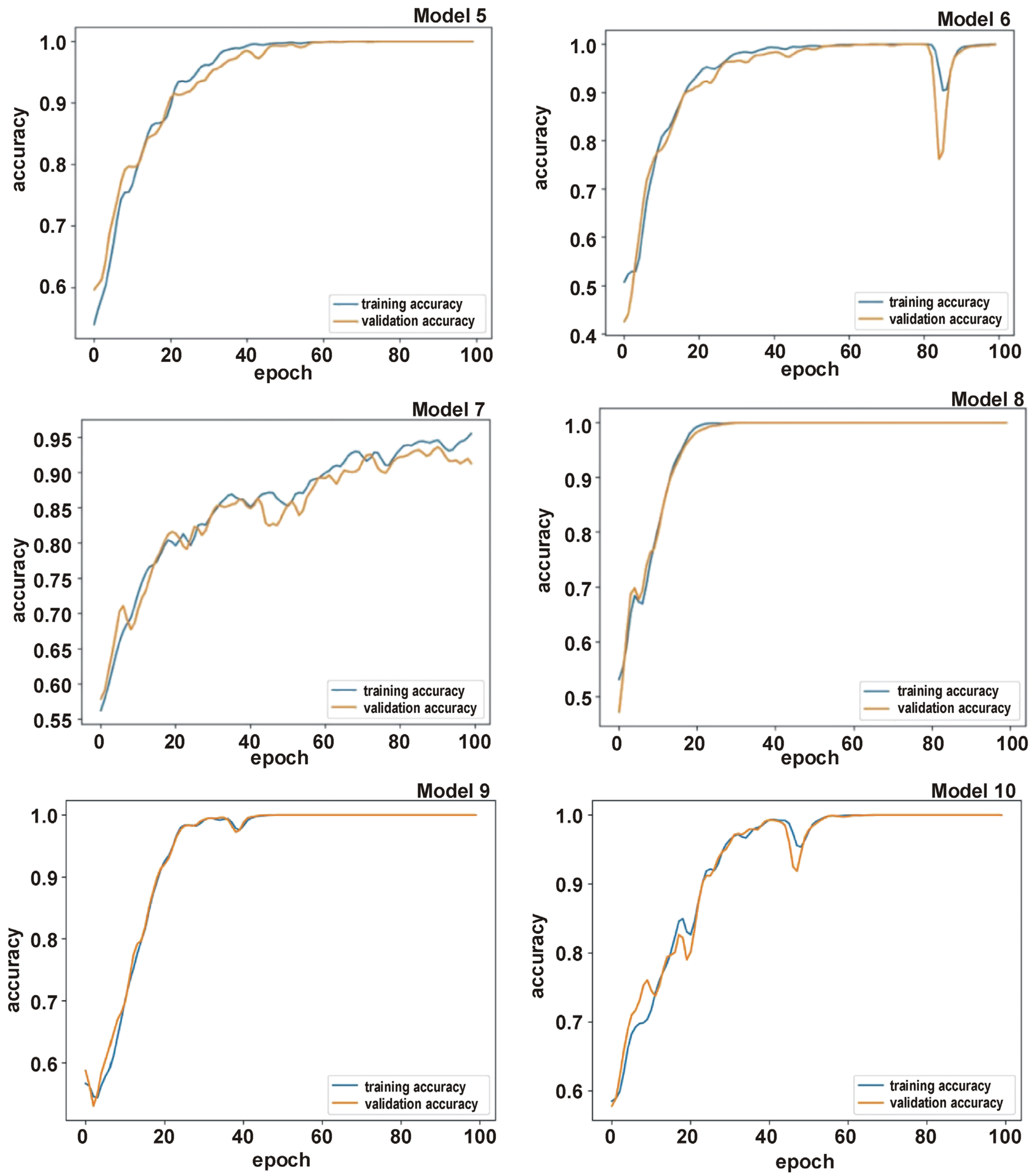

Fig. 7. Training and validation accuracy curves for Models 5–10

Discussion

The integration of digital technologies into dental diagnosis and treatment planning has expanded considerably over the past 2 decades. Applications including computer-assisted interpretation of diagnostic images, three-dimensional treatment planning, face-driven orthodontics, digital surgical planning, prosthetic design, and three-dimensional (3D) printing have transformed many aspects of clinical dentistry. These technologies aim to improve diagnostic precision and treatment outcomes by providing objective results that complement conventional clinical assessment.[47–49] Although experienced clinicians can interpret a wide range of clinical presentations, diagnostic performance may be influenced by stress and environmental factors.[50] Furthermore, many medical conditions present with overlapping clinical features,[51] making the distinction between normal and abnormal findings challenging.[52] As previously mentioned, factors such as long working hours, stress and fatigue, which are common in the medical field, may lead to a decline in concentration, resulting in what is known as human error in diagnosis.[53] Consequently, AI has become increasingly incorporated in this field for the detection of eye diseases, breast cancer, fetal anomalies, and numerous other medical conditions.[54]

Assessment of gingival color remains one of the most subjective components of the GI, particularly when differentiating between scores of 0 and 1, where only slight changes in color may be present.[41] The human eye is a subjective tool. Unless it is calibrated by other examiners, or accompanied by the use of adjunctive devices,[55,56] or, as we present in this article, digital methods, it is not objective.[48]

Convolutional neural networks are inspired by the mechanism of the visual cortex in animals. The manner in which animals perceive images is distinct from that of humans. In order to facilitate the retention and subsequent recognition of a particular image, it is necessary for animals to be exposed to multiple images and augmented photographs.[57]

Several CNN models were tested to detect gingival color alterations. The decisive factor in preferring one model over another is the trial-and-error process,[58] since there is no objective method to determine the best CNN model for different inputs and experiments.[59]

Raw images were imported into the system and then subjected to augmentation, including horizontal and vertical flipping, random cropping, and several variations in color or intensity. This led to the multiplication of the number of images.[59] The increased dataset obtained by applying such techniques improved the accuracy of the models. This method of image processing and augmentation was successfully adopted by Thaha et al. to detect brain tumors in magnetic resonance imaging (MRI) scans.[60]

Even though increasing the number of images in the model (up to thousands of images) can be useful, care must be taken to not overfit the model.[61]

Errors occur for several reasons, including a lack of data, the type of model used, or high correlations among the classes (normal and abnormal), which may be misdiagnosed by an expert.[62]

Ten CNN models were constructed based on the research problem. Model 4 (Fig. 6) was the simplest model, with the fewest parameters and layers, and achieved excellent accuracy.[63] The development of custom CNN models for gingival disease detection is just one example of how advanced machine-learning techniques can be applied in the medical field.[29,62]

In the present study, we followed a slightly different method of utilizing technology and software programming in dental practice. Unlike other researchers who have carried out extensive work to create software that can be used for self-diagnosis and even provide advice on self-treatment and suggestions for preventive actions (e.g., Liang et al.),[63,64] we tried to simplify the procedure and develop software that is limited to providing a primary idea or an initial indication for patients to have their oral cavity examined by a specialist. Additionally, the tool provides an objective method for periodontists and general dental practitioners to distinguish between GI scores of 0 and 1.

The reason for the development of this tool was to avoid giving the patient a preconceived mindset[61] that might cast a negative shadow on the examination procedure and the treatment plan to be established by the dentist. Otherwise, the patient may have already made a self-diagnosis and may wish to formulate a treatment plan independently.

An older research model was adopted by Seshan and Shwetha and tested on a small number of patients.[65] The researchers used Serif PhotoPlus 6 software (Serif Ltd, Nottingham, UK), which relied on counting pixels in cropped areas of red, inflamed or bleeding gingiva. The comparison was performed between the same areas in the same patient after proper plaque control, using a red disc as a reference.[65] The work model utilized in the present study differed, as we relied on a trained software model to digitally analyze images of the entire facial gingiva without using a foreign body as a reference. This method enabled a primary assessment of the gingival status.

Limitations

The present study has several limitations.

First, only frontal intraoral photographs were analyzed. Future studies should investigate whether incorporating additional views and angles improves classification performance.

Second, image acquisition was performed using a single camera system. Future investigations should evaluate the performance using smartphone cameras, to better reflect real-world applications.

Third, all photographs were obtained under standardized lighting conditions to increase reliability and generalizability. Future studies should assess model performance under a broader range of lighting conditions, including both indoor and outdoor environments.

Finally, the proposed model evaluated only 1 clinical feature, namely gingival color alteration. Future research should integrate additional variables used not only by patients but also by dentists themselves, such as vertical/horizontal radiological bone resorption level and the presence or absence of clinical attachment loss.

Conclusions

The results of the present study demonstrate that CNN software models can be trained to accurately detect slight changes in gingival color. These models have the potential to be incorporated into patient-oriented applications that promote oral health awareness and encourage timely dental visits before the progression of gingivitis.

Ethics approval and consent to participate

The study was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants before enrollment. The study protocol was reviewed and approved by the Ethics Committee of the College of Dentistry, University of Al-Mashreq, Baghdad, Iraq (order No. 245, 2022).

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Ali Raad AbdulAzeez https://orcid.org/0000-0003-0484-9370
Hashim Mueen Hussein https://orcid.org/0000-0002-1981-2768
Mahmood Ali Jumaah https://orcid.org/0000-0002-6232-3900
Athraa Ali Mahmood https://orcid.org/0000-0003-3095-7092

References

1. Nazir M, Al-Ansari A, Al-Khalifa K, Alhareky M, Gaffar B, Almas K. Global prevalence of periodontal disease and lack of its surveillance. *ScientificWorldJournal*. 2020;2020:2146160. doi:10.1155/2020/2146160
2. Hussein HM, Mahmood AA, Alberaqdar FA. The prevalence and relationship of root caries depth and gingival recession among different Iraqi groups. *MDJ*. 2015;12(1):144–155. doi:10.32828/mdj.v12i1.840
3. Abdulazeez AR, Kadhum AB, Ali BH. Impact of maternal periodontal health on fetus weight in Iraqi pregnant women: A clinical study. *Braz Dent Sci*. 2023;26(2):e3702. doi:10.4322/bds.2023.e3702
4. AbdulAzeez AR, Alkinani AA. The crucial role of plaque control in peri-implant mucositis initiation as opposed to the role of systemic health condition: A cross-sectional study. *Clin Cosmet Investig Dent*. 2021;13:257–268. doi:10.2147/ccide.s316838
5. Mahmood AA, AbdulAzeez AR, Hussein HM. The effect of smoking habit on apical status of adequate endodontically treated teeth with and without periodontal involvement. *Clin Cosmet Investig Dent*. 2019;11:419–428. doi:10.2147%2FCCIDE.S236747
6. Dhingra K, Vandana KL. Indices for measuring periodontitis: A literature review. *Int Dent J*. 2011;61(2):76–84. doi:10.1111/j.1875-595X.2011.00018.x
7. Gizem A, Burak D. Evaluation of plaque index, gingival index and oral health-related quality of life in obese patients. *Odovtos*. 2023;25(1):166–178. doi:10.15517/ijds.2022.52533
8. Singh A, Sinha RK, Richa S, Kumar R, Kishor A, Kumar A. Assessment of periodontal status and treatment needs among beedi factory workers. *J Pharm Bioallied Sci*. 2021;13(Suppl 1):S237–S240. doi:10.4103/jpbs.JPBS_699_20
9. Naveen Kumar PG, Peeran SW, Ramalingam K. 73. Gingival and periodontal indices. In: *Essentials of Periodontics and Oral Implantology*. 1st ed. Saranraj JPS Publication; 2021.
10. Marques S, Ribeiro P, Falcão C, et al. Digital impressions in implant dentistry: A literature review. *Int J Environ Res Public Health*. 2021;18(3):1020. doi:10.3390/ijerph18031020
11. Bud M, Jitaru S, Lucaciu O, et al. The advantages of the dental operative microscope in restorative dentistry. *Med Pharm Rep*. 2021;94(1):22–27. doi:10.15386/mpr-1662
12. Ou X, Chen X, Xu X, et al. Recent development in X-ray imaging technology: Future and challenges. *Research (Wash D C)*. 2021;2021:9892152. doi:10.34133/2021/9892152
13. Sharma H, Suprabha BS, Rao A. Teledentistry and its applications in paediatric dentistry: A literature review. *Pediatr Dent J*. 2021;31(3):203–215. doi:10.1016/j.pdj.2021.08.003
14. Zhou Y, Cuddy R, McNeil DW, et al. Oral health and related risk indicators in north-central Appalachia differ by rurality. *Community Dent Oral Epidemiol*. 2021;49(5):427–436. doi:10.1111/cdoe.12618
15. Saeed NA, Hussein HM, Mahmood AA. Prevalence of dental anxiety in relation to sociodemographic factors using two psychometric scales in Baghdad. *Mustansiria Dent J*. 2017;14(1):38–50. doi:10.32828/mdj.v14i1.753
16. Edens MH, Khaled Y, Napeñas JJ. Intraoral pain disorders. *Oral Maxillofac Surg Clin North Am*. 2016;28(3):275–288. doi:10.1016/j.coms.2016.03.008
17. Garla BK, Satish G, Divya KT. Dental insurance: A systematic review. *J Int Soc Prev Community Dent*. 2014;4(Suppl 2):S73–S77. doi:10.4103/2231-0762.146200
18. Steinvik LM, Svartdal F, Johnsen JAK. Delay of dental care: An exploratory study of procrastination, dental attendance, and self-reported oral health. *Dent J (Basel)*. 2023;11(2):56. doi:10.3390/dj11020056
19. Martins J, Costa C, Oliveira T, Gonçalves R, Branco F. How smartphone advertising influences consumers' purchase intention. *J Bus Res*. 2019;94(378–387). doi:10.1016/j.jbusres.2017.12.047
20. Ventola CL. Mobile devices and apps for health care professionals: Uses and benefits. *P T*. 2014;39(5):356–364. PMID:24883008.
21. Torous J, Bucci S, Bell IH, et al. The growing field of digital psychiatry: Current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry*. 2021;20(3):318–335. doi:10.1002/wps.20883
22. Chuchu N, Takwoingi Y, Dinnes J, et al.; Cochrane Skin Cancer Diagnostic Test Accuracy Group. Smartphone applications for triaging adults with skin lesions that are suspicious for melanoma. *Cochrane Database Syst Rev*. 2018;12(12):CD013192. doi:10.1002/14651858.CD013192
23. Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: Cross-sectional survey study. *J Med Internet Res*. 2020;22(1):e14679. doi:10.2196/14679
24. Aboueid S, Liu RH, Desta BN, Chaurasia A, Ebrahim S. The use of artificially intelligent self-diagnosing digital platforms by the general public: Scoping review. *JMIR Med Inform*. 2019;7(2):e13445. doi:10.2196/13445
25. Starcevic V. Keeping Dr. Google under control: How to prevent and manage cyberchondria. *World Psychiatry*. 2023;22(2):233–234. doi:10.1002/wps.21076
26. Seweryn P, Orzeszek SM, Waliszewska-Prosół M, et al. Relationship between pain severity, satisfaction with life and the quality of sleep in Polish adults with temporomandibular disorders. *Dent Med Probl*. 2023;60(4):609–617. doi:10.17219/dmp/171894
27. Mittal N, Thangamuthu T, Gupta S, Gupta S, Aggarwal H, Kharat S. Comparative evaluation of resin-based sealers and bioceramic sealers for postoperative pain after endodontic treatment: A systematic review. *Dent Med Probl*. 2024;61(2):293–300. doi:10.17219/dmp/155885
28. Lin CS, Lee CY, Chen LL, Wu LT, Yang SF, Wang TF. Magnification of fear and intention of avoidance in non-experienced versus experienced dental treatment in adults. *BMC Oral Health*. 2021;21(1):328. doi:10.1186/s12903-021-01682-1
29. Dey K, Nikhil V, Roy S. Machine learning approach with higher accuracy for simultaneous measurement of the dual parameter by MSM fiber structure. *Measurement*. 2023;221:113426. doi:10.1016/j.measurement.2023.113426

30. Amisha, Malik P, Pathania M, Rathaur VK. Overview of artificial intelligence in medicine. *J Family Med Prim Care*. 2019;8(7):2328–2331. doi:10.4103/jfmpc.jfmpc_440_19
31. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500–510. doi:10.1038/s41568-018-0016-5
32. Albawi S, Mohammed TA, Al-Zawi S. Understanding of a convolutional neural network. Paper presented at: 2017 International Conference on Engineering and Technology (ICET); August 21–23, 2017; Antalya, Turkey. doi:10.1109/ICEngTechnol.2017.8308186
33. Vangety N, Dey K, Roy S. Weight-location recognition in a plastic optical fiber using a convolutional neural network. *Opt Fiber Technol*. 2023;75:103166. doi:10.1016/j.yofte.2022.103166
34. Durán-López L, Dominguez-Morales JP, Conde-Martin AF, Vicente-Diaz S, Linares-Barranco A. PROMETEO: A CNN-based computer-aided diagnosis system for WSI prostate cancer detection. *IEEE Access*. 2020;8:128613–128628. doi:10.1109/ACCESS.2020.3008868
35. Al-Antari MA, Hua CH, Bang J, Lee S. Fast deep learning computer-aided diagnosis of COVID-19 based on digital chest X-ray images. *Appl Intell (Dordr)*. 2021;51(5):2890–2907. doi:10.1007/s10489-020-02076-6
36. Kisilev P, Sason E, Barkan E, Hashoul S. Medical image description using multi-task-loss CNN. In: Carneiro G et al., eds. *Deep Learning and Data Labeling for Medical Applications*. Springer International Publishing; 2016:121–129. doi:10.1007/978-3-319-46976-8_13
37. Jumaah MA, Shihab AI, Farhan AA. Epileptic seizures detection using DCT-II and KNN classifier in long-term EEG signals. *Iraqi J Sci*. 2020;61(10):2687–2694. doi:10.24996/ijs.2020.61.10.26
38. Carlson RV, Boyd KM, Webb DJ. The revision of the Declaration of Helsinki: Past, present and future. *Br J Clin Pharmacol*. 2004;57(6):695–713. doi:10.1111/j.1365-2125.2004.02103.x
39. Luo C, Li X, Wang L, He J, Li D, Zhou J. How does the data set affect CNN-based image classification performance? Paper presented at: 2018 5th International Conference on Systems and Informatics (ICSAI); November 10–12, 2018; Nanjing, China. doi:10.1109/ICSAI.2018.8599448
40. Nguyen QH, Ly HB, Ho LS, et al. Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Math Probl Eng*. 2021:4832864. doi:10.1155/2021/4832864
41. Löe H. The gingival index, the plaque index and the retention index system. *J Periodontol*. 1967;38(6P2):610–616. doi:10.1902/jop.1967.38.6_part2.610
42. Tzutalin. LabelImg. Git code (2015). https://github.com/HumanSignal/labelImg. Accessed June 1, 2024.
43. Sultana F, Sufian A, Dutta P. Advancements in image classification using convolutional neural network. Paper presented at: 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN); November 22–23, 2018; Kolkata, India. doi:10.1109/ICRCICN.2018.8718718
44. Mobeen-ur-Rehman, Khan SH, Abbas Z, Rizvi D. Classification of diabetic retinopathy images based on customised CNN architecture. Paper presented at: 2019 Amity International Conference on Artificial Intelligence (AICAI); February 4–6, 2019; Dubai, United Arab Emirates. doi:10.1109/AICAI.2019.8701231
45. Shreya S, Le DV. Completely automated CNN architecture design based on VGG blocks for fingerprinting localisation. Paper presented at: 2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN); November 29–December 2, 2021; Lloret de Mar, Spain. doi:10.1109/IPIN51156.2021.9662642
46. McHugh ML. Interrater reliability: The kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276–282. PMID:23092060.
47. Paradowska-Stolarz AM, Wieckiewicz M, Mikulewicz M, et al. Comparison of the tensile modulus of three 3D-printable materials used in dentistry. *Dent Med Probl*. 2023;60(3):505–511. doi:10.17219/dmp/166070
48. Paradowska-Stolarz A, Wieckiewicz M, Kozakiewicz M, Jurczyszyn K. Mechanical properties, fractal dimension, and texture analysis of selected 3D-printed resins used in dentistry that underwent the compression test. *Polymers (Basel)*. 2023;15(7):1772. doi:10.3390/polym15071772
49. Boyapati R, Peeta J, Dhulipalla R, Kolaparthy L, Adurty C, Cheruvu RNS. Comparative evaluation of the efficacy of probiotic, *Aloe vera*, povidine-iodine, and chlorhexidine mouthwashes in the treatment of gingival inflammation: A randomized controlled trial. *Dent Med Probl*. 2024;61(2):181–190. doi:10.17219/dmp/156425
50. Lourenco F, Casey BJ. Adjusting behavior to changing environmental demands with development. *Neurosci Biobehav Rev*. 2013;37(9 Pt B):2233–2242. doi:10.1016/j.neubiorev.2013.03.003
51. Zwaan L, Singh H. The challenges in defining and measuring diagnostic error. *Diagnosis (Berl)*. 2015;2(2):97–103. doi:10.1515/dx-2014-0069
52. Cotte F, Desvignes M, Ayasso H, Vignolle JM. A sparse dictionary representation approach for anti-scattering grid artifact removal in X-ray images. *Biomed Signal Process Control*. 2023;86:105247. doi:10.1016/j.bspc.2023.105247
53. Gates M, Wingert A, Featherstone R, Samuels C, Simon C, Dyson MP. Impact of fatigue and insufficient sleep on physician and patient outcomes: A systematic review. *BMJ Open*. 2018;8(9):e021967. doi:10.1136/bmjopen-2018-021967
54. Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: A survey. *Evol Intell*. 2022;15(1):1–22. doi:10.1007/s12065-020-00540-3
55. Sakiroff LM, Chennell P, Yessaad M, Pereira B, Bouattour Y, Sautou V. Evaluation of color changes during stability studies using spectrophotometric chromaticity measurements versus visual examination. *Sci Rep*. 2022;12(1):8959. doi:10.1038/s41598-022-13025-3
56. Śmielecka M, Dorocka-Bobkowska B. Comparison of two optical devices used for artificial tooth color selection. *Dent Med Probl*. 2022;59(2):249–253. doi:10.17219/dmp/141147
57. Gu P, Lan X, Li S. Object detection combining CNN and adaptive color prior features. *Sensors (Basel)*. 2021;21(8):2796. doi:10.3390/s21082796
58. Alzubaidi L, Zhang J, Humaidi AJ, et al. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J Big Data*. 2021;8(1):53. doi:10.1186/s40537-021-00444-8
59. Susilo AB, Sugiharti E. Accuracy enhancement in early detection of breast cancer on mammogram images with convolutional neural network (CNN) methods using data augmentation and transfer learning. *J Adv Inf Syst Technol*. 2021;3(1):9–16. doi:10.15294/jaist.v3i1.49012
60. Thaha MM, Kumar KPM, Murugan BS, Dhanasekeran S, Vijayakarthick P, Selvi AS. Brain tumor segmentation using convolutional neural networks in MRI images. *J Med Syst*. 2019;43(9):294. doi:10.1007/s10916-019-1416-0
61. Yang S, Chen LF, Yan T, Zhao YH, Fan YJ. An ensemble classification algorithm for convolutional neural networks based on AdaBoost. Paper presented at: 2017 IEEE/ACIS 16th International Conference on Computer and Information Science; May 24–26, 2017; Wuhan, China. doi:10.1109/ICIS.2017.7960026
62. Kim B, Yuvaraj N, Sri Preethaa KR, Arun Pandian R. Surface crack detection using deep learning with shallow CNN architecture for enhanced computation. *Neural Comput Appl*. 2021;33(15):9289–9305. doi:10.1007/s00521-021-05690-8
63. Liang Y, Fan HW, Fang Z, et al. OralCam: Enabling self-examination and awareness of oral health using a smartphone camera. Paper presented at: CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; April 25–30, 2020; Honolulu, HI. doi:10.1145/3313831.3376238
64. Ryan A, Wilson S. Internet healthcare: Do self-diagnosis sites do more harm than good? *Expert Opin Drug Saf*. 2008;7(3):227–229. doi:10.1517/14740338.7.3.227
65. Seshan H, Shwetha M. Gingival inflammation assessment: Image analysis. *J Indian Soc Periodontol*. 2012;16(2):231–234. doi:10.4103%2F0972-124X.99267