ISSN: 3033-3555

Research Article

Journal of Stomatology & Dental Research

Enhancing Dental Caries Identification with Deep Learning: A Study of Convolutional Neural Networks and Transfer Learning Approaches

Fayqa Mannan*, Maymoona and Ayesha

Department of Computer Science, Government College Women University, Faisalabad, 38000, Pakistan

***Corresponding author**
Fayqa Mannan, Department of Computer Science, Government College Women University, Faisalabad, 38000, Pakistan.

Received: December 12, 2025; **Accepted:** December 18, 2025; **Published:** December 29, 2025

ABSTRACT
Background: Dental caries is a prevalent oral health issue, and early diagnosis using X-ray images can significantly improve treatment outcomes. Deep learning techniques have been increasingly employed for automated detection of dental caries in radiographic images.

Objectives: This study aims to evaluate the effectiveness of deep learning models, including Convolutional Neural Networks (CNNs) and transfer learning approaches, in identifying dental caries using periapical radiographs.

Methods: We utilized a traditional CNN model along with transfer learning models, including Visual Geometry Group (VGG16, VGG19), ResNet50, and Inception V3. The CNN model consisted of three sets of 2D convolutional layers followed by activation, max-pooling, flatten, dense layers, dropout, and final activation layers. For the transfer learning models, the top convolutional layers were frozen to prevent retraining, allowing only the last layers to be trained. Hyperparameters were optimized using a grid search approach, and model performance was validated using the Shuffle-Split-Cross (SSC) validation method.

Results: Ten images were generated for each original image, resulting in a total of 1,150 training dataset images. The accuracy achieved by the CNN, VGG16, VGG19, ResNet50, and Inception V3 models was 90%, 96%, 73%, 70%, and 73%, respectively. Among these, VGG16 exhibited the highest accuracy.

Conclusions: The findings demonstrate that transfer learning, particularly with VGG16, is highly effective in diagnosing dental caries from periapical radiographs. These results highlight the potential of deep learning models for improving automated dental diagnostics. Transfer learning, especially with VGG-16, achieved the highest accuracy (96%) in this study, outperforming both traditional CNN and related studies, highlighting its effectiveness for dental caries detection.

Keywords: Deep Learning, Dental Caries, Transfer Learning, X-Ray Images, Convolutional Neural Network

Introduction

Dental caries is caused by pathogenic bacteria. Pathogenic bacteria cause dental caries by attacking the dentin of the teeth and initiating their destruction. The primary pathogens responsible for dental caries include Villanelle, Aggregatibacter, Leptotrichia, Bacteroides, Granulicatella, Streptococcus, and Prevotella. Approximately 80% to 90% of the global population is affected by dental caries [1]. According to the World Health Organization (WHO), individuals residing in industrialized regions are more prone to dental caries due to unhealthy lifestyles, including increased sugar consumption and high fluoride exposure [2]. Radiography is considered one of the most effective methods for detecting all types of caries, particularly hidden ones [3].

Dental caries are curable if diagnosed at an early stage. In modern dentistry, early detection and recognition of caries initiation is a major concern [4]. Both traditional and convolutional approaches have been explored for the detection of dental caries

Citation: Fayqa Mannan, Maymoona, Ayesha. Enhancing Dental Caries Identification with Deep Learning: A Study of Convolutional Neural Networks and Transfer Learning Approaches . J Stoma Dent Res. 2025. 3(4): 1-12. DOI: doi.org/10.61440/JSDR.2025.v3.40

[5]. Among non-invasive techniques, the Quantitative Light-Induced Fluorescence-Digital (QLF-D) method is considered one of the most effective for detecting caries [6].

Despite ongoing advancements, detection and treatment methods have not shown significant improvements over the past few decades, primarily due to the complex morphological anatomy of teeth. While existing techniques are effective for detecting advanced or moderate caries, they often fail to identify caries at early stages [7]. These methods struggle with detecting deep fissures, tight interproximal contacts, and secondary lesions on teeth.

If not detected early, caries can spread beyond the affected tooth to neighboring teeth [8,9]. Evidence shows that individuals in underdeveloped countries, particularly those with low to middle incomes and limited education, are most affected by dental caries. The high cost of treatment makes dental care inaccessible for many, leading to a surge in oral health problems and potentially contributing to systemic diseases such as cardiovascular disorders, cancer, chronic respiratory diseases, and diabetes [10]. These challenges highlight the urgent need for an affordable, automated algorithmic system capable of detecting all types of caries at early stages, thereby minimizing dependence on extensive human expertise.

A human tooth primarily consists of three tissues: enamel, dentin, and pulp. Bacterial pathogens can attack any of these layers, resulting in enamel caries, dentinal caries, or pulpal caries, respectively [11]. Affected teeth often appear spongy, rough, decayed, broken, and may show discoloration. Various image processing techniques have been applied for dental caries detection, typically followed by machine learning or deep learning models [12]. These techniques generally involve steps such as preprocessing, localization, data splitting, model training, and classification [13]. However, accurate early-stage detection of dental caries remains a significant challenge. Recent studies have demonstrated that deep learning algorithms can achieve outstanding performance in the medical domain [14].

This study contributes by employing CNN and transfer learning models—specifically VGG-16 and VGG-19—for the detection of dental caries using periodic radiographs. To the best of our knowledge, these models have not been previously applied to this problem in the existing literature. The traditional CNN architecture used comprises three sets of 2D convolutional layers, activation layers, max-pooling layers, a flattening layer, dense layers, activation layers, dropout, and final output layers. In the VGG models, the top convolutional layers are frozen to preserve learned features, while only the final layers are retrained to generate the desired output.

Literature Review

Deep learning models are employed for preprocessing, segmentation, feature extraction, detection, and classification tasks. Lee and colleagues utilized deep convolutional neural network models to detect and diagnose dental caries using periapical radiography images. The GoogleNet Inception V3 (GNIV3) was used for data preprocessing [15]. Another study noted that dental images obtained through Computed Tomography (CT) have a limited range of properties such as sensitivity and geometric values. To address these challenges, the researchers found deep convolutional neural networks to be highly beneficial, particularly for resolution enhancement. They specifically highlighted two CNN models, U-Net and the Sub-Pixel network, in their work for enhancing image resolution. The proposed models exhibited significantly improved performance for CT images, aiding in the enhanced detection of medical attributes such as shape and size [16].

The study proposed an automated system that performs segmentation of dental images based on pixel values [17]. Furthermore, the researchers utilized a deep convolutional neural network for segmenting and labeling the affected areas of the teeth, such as those impacted by gingival issues. They employed a 3D tooth model as input, utilizing a CNN model with a two-level hierarchy. First, the teeth were labeled, and then the extracted features were fed into the neural network model. The approach by Cantu et al, achieved a remarkable 99% accuracy [18].

The literature indicates that deep learning models are widely utilized for the classification and detection of caries-infected teeth. Numerous researchers have proposed various models in different forms for classification purposes. Fully convolutional neural networks have been employed for estimating the probabilities of caries. One study proposed a model based on CNN called 'Automated Dental Red Autofluorescence Plaque Image Classification' for classifying QLF images. The study by A. Betul and Oktay demonstrated the use of a CNN model for the classification of tooth and non-tooth areas by employing tooth proposals [19]. Another article by Naam et al focused on the automated assessment of lower third molar development from panoramic radiographs using a pilot technique. Machine Learning (ML) algorithms were utilized to train the dataset, with deep convolutional neural networks employed for classification [20].

Moreover, a machine learning approach was employed for the classification of healthy gums versus unhealthy gums. The study utilized a database of 251 Radiovisiography (RVG) dental images, containing images related to three oral diseases: caries, periapical infection, and periodontitis. The research team utilized both a CNN model and VGG16 model for disease classification, achieving 73% accuracy with the CNN model and 88% accuracy with VGG16 [21]. Additionally, a Computer-Aided Diagnosis (CAD) system was proposed for the detection of dental caries, utilizing a bitewing radiographs database consisting of 3000 images. This system was based on a Fully Connected Convolutional Neural Network (FCNN) containing over 100 layers. The model effectively classified infected and non-infected teeth [22]. In another study, K. Moutselos et al, panoramic X-ray images were used and classified into incisors, molars, and premolars using a CNN model, specifically a modified version of AlexNet [23]. In situations such as major disasters like bomb blasts where identification becomes incredibly challenging due to the loss of many lives, dental forensic identification tests are crucial. However, labeling a large dataset during such disasters is considered highly complicated for general dentists. Deep CNN-based models have proven to be invaluable in these scenarios, as they are utilized for the detection and classification of teeth. The study employed AlexNet for the classification of identified teeth [23].

This model achieved an accuracy of 77.4%. Deep learning models prove to be very beneficial for comparative studies. For instance, CNN models are utilized to classify periapical teeth based on their condition after treatment into three classes: 'getting better,' 'getting worse,' or 'no change,' achieving a 74% F1 score [24]. In another study, the machine learning algorithm Support Vector Machine (SVM) was used to classify caries-infected and non-infected teeth using a panoramic image database for identifying proximal dental caries. This model achieved an accuracy of 81% [25]. A Neural Network was employed for the detection of dental caries using a database of periapical radiographs. They utilized an Adaptive Neural Network for this purpose, achieving an impressive 94% accuracy [25]. Machine learning models are divided into two categories: supervised learning models and unsupervised learning models. Unsupervised learning techniques are employed for the detection of dental caries through nonstandardized X-ray images [26]. CNN techniques are also used to differentiate between various treated teeth, such as cavity fillings, implants, and root canal-treated teeth [27]. In another study, the validity of CNN techniques was evaluated for diagnosing dental caries in different types of teeth, including proximal caries lesions on bitewing radiographs. [28].

Table 1: Performance in Term of Accuracy Comparison of the Detection of Dental Caries Techniques Mentioned in the Relevant Papers.

Reference	Accuracy	Reference	Accuracy
[2]	51%	[19]	94%
[3]	81%	[20]	81%
[7]	74%	[21]	71%
[8]	82%	[23]	88.46%
[9]	83%	[24]	84%
[13]	95%	[25]	61%
[14]	95%	[26]	90%
[15]	84%	[27]	74%
[17]	74%	[28]	81%
[18]	99%	[29]	96%

Table 1 presents a performance comparison of dental caries detection techniques along with their respective accuracies. CNN demonstrates exceptional performance in the fields of computer vision and image processing. CNN models offer solutions to the aforementioned challenges. They possess the learning ability to acquire features. There are several reasons for employing CNN models:

- CNN serves as an excellent feature extractor, autonomously learning features deeply from images.
- It has the capability to exploit spatial and temporal correlation data.
- CNN can learn feature representations specific to a dataset; for example, it can identify hidden features that may be overlooked by dentists, such as proximal caries.
- CNN models automatically discover all deep or hidden features from large datasets, eliminating the need for expert intervention in image processing.

Table 2: Techniques Mentioned by Relevant Papers for Detection of Dental Caries

Author	Technique(s)	Authors	Technique
[2]	F-CNN	[20]	CNN
[3]	Google Net Inception V3	[21]	CNN
[4]	U-net	[22]	ML
[5]	Adaptive Neural Network (ADA)	[23]	VGG-16
[6]	VGG16 CNN	[24]	CNN
[7]	Alex Net	[25]	F-CNN
[8]	Mask-RCNN	[26]	Modified Alex Net
[12]	Squeeze Net, MobileNet-v2, Google Net, ResNet-18 and ResNet-50	[27]	CNN model
[15]	Google Net Inception V3	[28]	SVM
[17]	ML	[29]	VGG19-CNN

Table 2 presents a performance comparison of dental caries detection techniques, where CNN not only demonstrates good performance in the field of dental caries but also in other domains. In dentistry, caries is detected through periapical and panoramic radiographs, as well as bitewing radiographs. Ocular caries is easily detected, but deep learning models yield impressive results in diagnosing proximal caries [29]. Deep learning models such as ANN, DNN, RNN, and CNN are also utilized in various medical fields, including the diagnosis of brain tumors, lung cancer, eye infections, digital pathology, chest and abdominal issues, musculoskeletal disorders, and dermatological diseases, among others. Deep Neural Networks (DNN) represent a class of deep learning models, employed for detecting and classifying Alzheimer's disease, as well as for segmenting brain tissues. Additionally, various ophthalmic diseases are diagnosed using deep learning models [30]. A simple CNN model is employed for diagnosing Color Fundus Imaging (CFI), which finds application in numerous ophthalmic disease-related tasks such as anatomical structure segmentation, detection of retinal abnormalities, eye disease diagnosis, and assessment of image quality [31]. CNN is such type of neural network [32, 33] that is a powerful tool [34] implemented by many researchers for feature extractions [35].

Proposed Work

This study utilized convolutional neural network models and transfer learning models for diagnosing dental caries from periapical radiographs. The primary research steps included image acquisition, data augmentation, image preprocessing, and training and validating the models using the cross-validation method.

Image Acquisition

The process of gathering data is known as image acquisition. The database of dental periapical images was collected in 2013 and published in 2016. This dataset was obtained from the University Technology Malaysia (UTM) Health Center's dental clinic. The images were captured during regular checkups of university students at the clinic. The age group of the students ranged between 25 and 35 years old. Patients were informed about the data collection process. For this study, periapical dental X-ray images were utilized. These dental X-ray data were collected using a specialized X-ray machine. An intraoral X-ray machine, directly linked to a digital scanner and software known as “SIDEX XG,” was used to produce digital periodic dental radiographs. The details of the dataset have been mentioned in Table 3. The hardware and software utilized in the study are from a German company named Sirona. The dataset specifications include 120 images in JPEG format, with dimensions of 748 × 512 pixels. There are three main types of tooth radiographs: panoramic, bitewing, and periapical. The dataset used in this research consists of periapical radiographs. Periapical X-rays display all the teeth within one section, either the upper or lower jaw. These X-rays are among the most cost-effective and commonly used for diagnosing and treating dental conditions, offering clear views of the tooth roots and surrounding bone structures. This dataset is also publicly available.[9]

Table 3: Details of Dataset

Dataset Type	Original Images	Augmented Images	Total Images (After Augmentation)	Training Data (80%)	Testing Data (20%)
Dental Images	120	1030	1150	920	230

Data Cleaning

Data cleaning is a straightforward process involving three steps: excluding blurred images, cropping, and flipping. All radiographs include dimension values on their sides. The next step involved flipping images to standardize all mixed images of the maxilla and mandibular jaw into the mandibular jaw orientation using a vertical flip.

Data Augmentation

The augmentation process was executed on the Keras framework using the ImageDataGenerator (IDG) function. Data was randomly generated with the following specifications: rotation range of 40 degrees, width shift range of 0.2, height shift range of 0.2, shear range of 0.2, zoom range of 0.2, horizontal flip set to true, and fill mode set to ‘nearest.’ Ten images were generated for each original image, resulting in a total of 1,150 training dataset images. Figure. 1 illustrates and describes the image augmentation techniques employed, including rotation, width shift, height shift, shear, and zoom range.

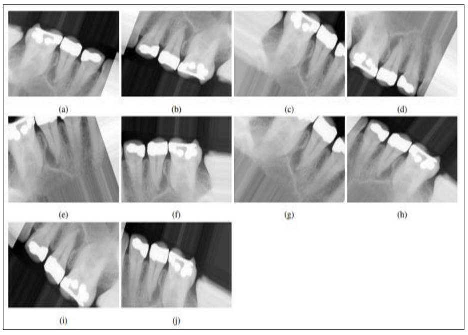


Figure 1: Image Augmentation Using Rotation-Range, Width Shift & Height Shift-Range, Shear & Zoom-Range

Data Preprocessing

After cleaning the data, image enhancement was conducted to improve the visibility of the images. In this study, images were enhanced by adjusting three factors: brightness, sharpness (as shown in Figure. 2b), and contrast (as shown in Figure. 2c). For this purpose, the ‘image enhance’ function from the PILLOW library was utilized. Figure. 2 illustrates and describes the sharpening and contrast adjustments made using the ‘image enhance’ function, as well as the enhancement of brightness in the images.

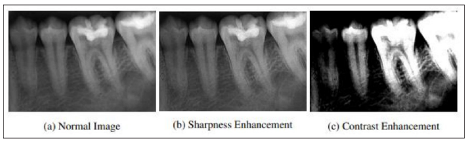


Figure 2: Image Enhancement

Image Segmentation

The images were segmented using a supervised segmentation method, employing three primary techniques: thresholding algorithms, the random walker method, and the active contour method. In thresholding algorithms, images are converted into pixels, and a range of pixel values is defined by the user in supervised data. This conversion transforms the images from grayscale to binary form, where grayscale images typically contain pixel values ranging from 0 to 255. The threshold method divides the image pixels into two categories: black and white, with pixels categorized below 150 and above 150 separately using the ‘threshold()’ function from the OpenCV library, as depicted in Figure. 3b. Its invariance is also shown in Figure. 3c. In the active contour method, a circle or line is drawn around the Region of Interest (ROI), and this contour expands or contracts based on light and edge information. In our dataset, where multiple caries may be present in an image, we required a method that could segment all pixels from areas of small caries using predefined labels. The random walker method proved to be the most suitable for this task. The random walker method is particularly effective for segmenting a small number of pixels from images. This approach utilizes predefined pixels and ‘walks’ randomly to all unlabeled pixels, assigning appropriate labels. It is especially useful for high-quality image segmentation.

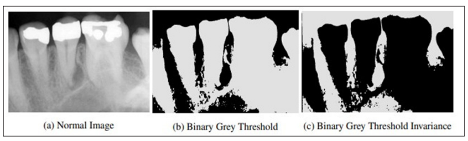


Figure 3: Image Segmentation Shows the Normal Image, Binary Grey Threshold and Binary Grey Threshold Invariance in Image Segmentations.

Feature Extraction

There are various methods for extracting features from images, such as using the Principal Component Analysis (PCA) method. PCA is utilized for the analysis of multivariate data and is particularly effective for extracting linear features and facilitating data comparisons. This method reveals only a few features by compressing the original image data [reference for claiming of compressing image]. However, deep learning models including CNN Sequential Model, VGG16, VGG19, ResNet50, and Inception V3, are also capable of automatically extracting features from images. In this study, these models are employed for detecting dental caries automatically extracted features. In previous studies, the said models were used on fewer images [limitation].

Training Models

This study utilized CNN Sequential Model, VGG16, VGG19, ResNet50, and Inception V3 models. It was discovered that these models had not been previously applied to the given dataset. Details on the utilization of these models are provided below.

CNN Model

As shown in Figure. 4, the input images were scaled down to 250 × 250 × 1 to reduce computational expenses. Processing high-dimensional images in their original dimensions would require significant computation. However, reducing the dimensions of images can sometimes result in the loss of information. The proposed network comprises three sets of convolutional layers, ReLU activation layers, and max-pooling layers. They are followed by a sequence of layers: flatten, dense, activation, dense, dense, and activation layer. This sequential model takes input images of size 250 × 250 × 1. Here is a summary of the inputs for all these layers:

- Input image size: 250 × 250 × 1
- Conv2D input layer size 250 × 250 × 1, Activation layer 248 × 248 × 32, Max pooling 248 × 248 × 32
- Conv2D input layer size 124 × 124 × 32, Activation layer 122 × 122 × 32, Max pooling 122 × 122 × 32
- Conv2D input layer size 61 × 61 × 32, Activation layer 59 × 59 × 64, Max pooling 61 × 61 × 32
- Flatten layer input size: 29 × 29 × 64
- Dense layer with 53824 nodes, followed by an activation with size 64, a dropout layer of 64 nodes, another dense layer of 64 nodes, and an activation of size 1.

The kernel size is 3 × 3, all activation functions are ReLU, and the final activation function is Sigmoid. Sigmoid activation is applied to the last dense layer. The dropout layer has a size of 0.5 to mitigate the problem of overfitting.

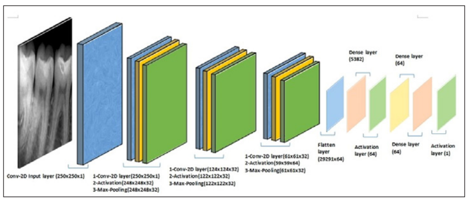


Figure 4: Architecture of CNN Model

Transfer Learning

As shown in Figure. 5, VGG architectures are typically trained on ImageNet. The workings of both VGG16 and VGG19 models are the same; the only difference lies in their number of layers, with VGG19 having more layers in its structure. Firstly, in the implementation setup, the relevant packages must be imported. For VGG16 and VGG19, input images are resized, scaling them down to 224 × 224. The weights for the first three convolutional layers of VGG16 are frozen, while the subsequent layers are used for training. The concept of transfer learning is applied for fine-tuning, enabling the model to learn basic low-level features from the images in the dataset. These models are pre-loaded with ImageNet weights.

The top convolutional layers are disabled. The weights of these models are set to "ImageNet," indicating that they are pre-trained based on ImageNet. The model type is sequential, with one flattening layer included without any parameters. Three dense layers are added with units=256 and ReLU activation functions in the first and second dense layers. The last dense layer has units=3 with a "softmax" activation function. All other optional layers are excluded to make the model more adaptive to the training data. The final dense layer with 'softmax' functions as a binary classifier, distinguishing between carious and non-carious teeth.

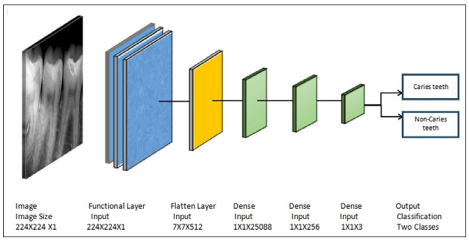


Figure 5: Architecture of Transfer Learning Models

Inception V3

Inception V3 architectures are typically pre-loaded with ImageNet weights. Firstly, in the implementation setup, the relevant packages must be imported. For Inception V3, input images are resized, scaling them down to 224 × 224 as shown in Fig. 6 [17]. The weights for the first three convolutional layers of Inception V3 are frozen, while the subsequent layers are used for training. The concept of transfer learning is employed for fine-tuning, enabling the model to learn basic low-level features from the dataset images. These models are pre-loaded with ImageNet weights. The top convolutional layers are disabled. The weights of these models are set to 'ImageNet,' indicating that they are pre-trained based on ImageNet. The model type is sequential, with one flattening layer included without any parameters. Three dense layers are added with units=256 and ReLU activation functions in the first and second dense layers. The last dense layer has units=3 with a 'softmax' activation function. All other optional layers are excluded to make the model more adaptive to the training data. The final dense layer with 'softmax' functions as a binary classifier, distinguishing between carious and non-carious teeth.

ResNet 50

ResNet 50 architectures are typically pre-loaded with ImageNet weights. Firstly, in the implementation setup, the relevant packages must be imported. For ResNet 50, input images are resized, scaling them down to 224 × 224 as shown in Figure. 7 [17]. The weights for the first 3 convolutional layers of ResNet 50 are frozen, while the subsequent layers are used for training. The concept of transfer learning is employed for fine-tuning, allowing the model to learn some basic low-level features from the dataset images. These models are pre-loaded with ImageNet weights. The top convolutional layers are disabled. The weights of these models are set to 'ImageNet,' indicating that they are pre-trained based on ImageNet.

```
Model: "sequential_1"
Layer (type)                  Output Shape            Param #
=============================================================
inception_v3 (Functional)     (None, 5, 5, 2048)      21802784
flatten_1 (Flatten)           (None, 51200)           0
dense_3 (Dense)               (None, 256)             13107456
dense_4 (Dense)               (None, 256)             65792
dense_5 (Dense)               (None, 3)               771
=============================================================
Total params: 34,976,803
Trainable params: 13,174,019
Non-trainable params: 21,802,784
```

Figure 6: Inceptionv3 model used for classification of dental caries and non-infected teeth.

In Figure. 7, it is shown that the model type is sequential, with one flatten layer included without any parameters. Additionally, three dense layers are added with units=256 and ReLU activation functions in the first and second dense layers. The last dense layer has units=3 with a 'softmax' activation function. No other optional layers are included. This is done to make the model more adaptive to the training data. The final dense layer with 'softmax' functions as a binary classifier, distinguishing between carious and noncarious teeth.

```
Model: "sequential"
Layer (type)                  Output Shape            Param #
=============================================================
resnet50 (Functional)         (None, 7, 7, 2048)      23587712
flatten (Flatten)             (None, 100352)          0
dense (Dense)                 (None, 256)             25690368
dense_1 (Dense)               (None, 256)             65792
dense_2 (Dense)               (None, 2)               514
=============================================================
Total params: 49,344,386
Trainable params: 25,756,674
Non-trainable params: 23,587,712
```

Figure 7: Resnet 50 Model Used for Classification of Dental Caries and Non-Infected Teeth

The adopted methodology and flow of this study is shown by figure. 8

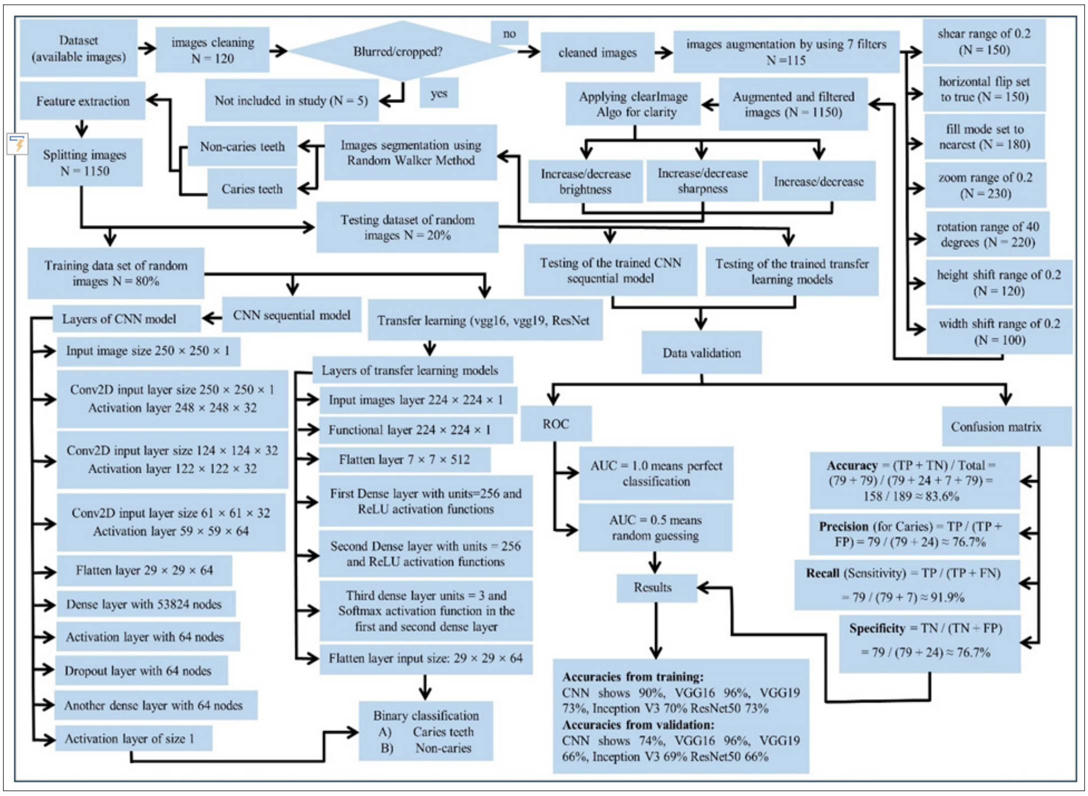


Figure 8: Methodology And Flow of This Study

Results

All employed models efficiently classify infected teeth from non-infected ones. However, the transfer learning models show better performance than the CNN sequential model. The evaluation of these models is measured by evaluation metrics named accuracy. A summary of the models' results is provided in Table 4. In the training process of the CNN sequential model, a total of 20 epochs are needed to converge with a trained model, achieving 90% accuracy. Table 3 illustrates the training process of VGG16, VGG19, ResNet 50, and Inception V3 models, each requiring a total of 5 epochs to converge with a trained model, achieving accuracies of 96%, 73%, 70%, and 73%, respectively. The VGG16 model demonstrates 6% higher accuracy than the CNN model. Accuracy is measured as TP+TN/TP+TN+FP+FN. True Positive refers to teeth that are truly infected by dental caries and are classified as caries-infected teeth. True Negative refers to samples classified as non-infected, which are truly healthy teeth. False Positive samples are healthy but erroneously classified as caries-infected teeth. False Negative indicates teeth classified as healthy by the model, but are actually infected by caries. These terms are denoted as TP=True Positive, TN=True Negative, FP=False Positive, and FN=False Negative.

Table 4: Performance Evaluation in the Term of Training and Validation Accuracy of Used Models

Model	Training	Validation
CNN	90%	70%
VGG16	96%	96%
VG19	73%	66%
Inception V3	70%	69%
ResNet50	73%	66%

TP+TN/TP+TN+FP+FN. True Positive denotes teeth truly infected by dental caries, also referred to as caries-infected teeth. True Negative refers to samples classified as non-infected, which are genuinely healthy teeth. False Positive represents samples classified as healthy but incorrectly labeled as cariesinfected teeth by the model. False Negative indicates teeth classified as healthy by the model, but in reality, they are infected by caries. These terms are denoted as: TP=True Positive, TN=True Negative, FP=False Positive, FN=False Negative.

However, these flexibilities are valuable for experimenting with large datasets. After validating the models, improvements in their performance are necessary. This is achieved by tuning parameters, which is both important and quite tricky. Finding the optimal parameter values for models is crucial, and the process can be complex. For parameter tuning, grid search is one of the best methods because it explores all possible combinations of parameters. Grid search is an approach that automatically selects the most accurate combination of parameters to assess performance validity and compare with other models. To implement grid search, GridSearchCV is imported from the sklearn.model_selection library. Hyperparameters are parameters provided to machine learning algorithms. In CNN algorithms, 'batch size' and 'epochs' are used as hyperparameters. The best combination is determined through the grid search method, as shown in Table 5.

Table 5: Performance Evaluation in Terms of Training, Validation Accuracy and Loss of Used Models

Models	Input Size	Training Accuracy	Testing Accuracy	Training Loss	Validation Loss	Number of Epochs	Batch Size
CNN sequential model	250 × 250 × 1	73%	53%	0.58	0.74	20	10
	200 × 200 × 1	68%	58%	0.67	0.67	10	5
	150 × 150 × 1	80%	58%	0.54	0.75	20	10
	50 × 50 × 1	77%	55%	0.55	0.86	20	10
Transfer Learning models	250 × 250 × 1	68%	50%	0.97	0.57	10	3
	250 × 250 × 1	70%	51%	0.55	0.47	14	2
	224 × 224 × 1	79%	49%	0.19	0.76	15	3
	224 × 224 × 1	69%	53%	0.63	0.09	10	3

The graphs below depict the accuracy and loss of all these models. Each model was trained using varying numbers of epochs and image sizes. Included are graphs for the CNN models and VGG models. Additionally, two other models, namely ResNet 50 and Inception V3, utilized a different package but share the same architecture as the VGG models. Accuracy metrics include training and validation accuracy, as well as training and validation loss.

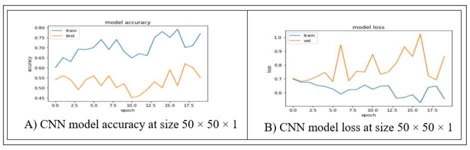


Figure 9: Cnn Model Graphs at the Size 50 × 50 × 1

Figure 9 shows the training and validation accuracy and shows training and validation loss in CNN model graph at the size 50×50 × 1.

Figure 10 shows the training and validation accuracy and also shows training and validation loss in CNN model graph at the size 150 ×150 × 1.

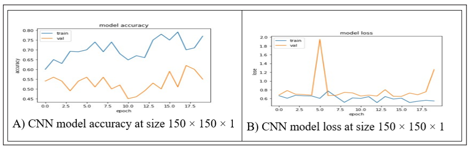


Figure 10: Cnn Model Graphs at the Size 150 × 150 × 1

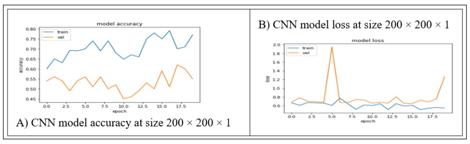


Figure 11: Cnn Model Graphs at the Size 200 × 200 × 1

Figure 11 shows the training and validation accuracy and also shows training and validation loss in CNN model graph at the size 200 × 200 × 1.

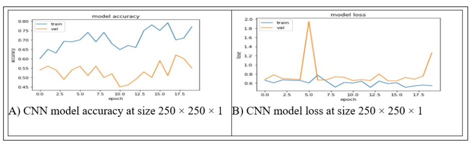
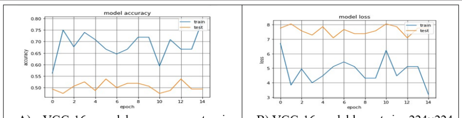


Figure 12: Cnn Model Graphs at the Size 250 × 250 × 1

Figure 12 shows the training and validation accuracy and also shows training and validation loss in CNN model graph at the size 250 × 250 × 1.

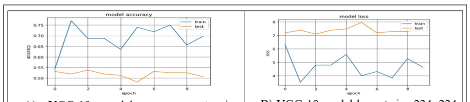
The graphs for all CNN sequential models are displayed in Fig. 10 through 15, showcasing the accuracy and loss of each model. These models were trained using varying numbers of epochs and different image sizes. The graphs illustrate the training and validation accuracy, as well as the training and validation loss, with image sizes set at 50 × 50 × 1, 150 × 150 × 1, 200 × 200 × 1, and 250 × 250 × 1 pixels. Various numbers of epochs and batch sizes are depicted, revealing the corresponding training accuracies and validation accuracies as detailed in Table 5.



A) VGG-16 model accuracy at size 224×224 ×1

B) VGG-16 model loss at size 224×224 × 1

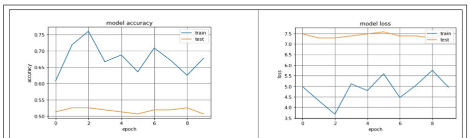
Figure 13: Vgg-16 Model Graphs at the Size 224 × 224 × 1



A) VGG-19 model accuracy at size 224×224 × 1

B) VGG-19 model loss at size 224×224 × 1

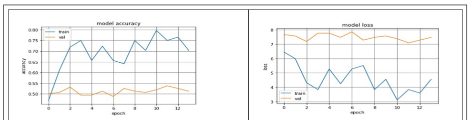
Figure 14: VGG-19 Model Graphs at the Size 224 × 224 × 1



A) ResNet-50 model accuracy at size 224×224 × 1

B) ResNet-50 model loss at size 224×224 × 1

Figure 15: ResNet-50 Model Graphs at the Size 224 × 224 × 1



A) Inception V3 model accuracy at size 224×224 × 1

B) Inception V3 model loss at size 224×224 × 1

Figure 16: Inception V3 Model Graphs at the Size 224 × 224 × 1

From figure 13-16 show the training and testing accuracies and training, validation by using VGG-16 , VGG-19- inception V3 and ResNet50. These graphs depict the training and testing accuracy, as well as the training and validation loss, with the input size of images set at 224 × 224 × 1 pixels. They illustrate various numbers of epochs and batch sizes, showcasing the varying training accuracies and testing accuracies detailed in the accompanying table.

Validation of Models

The models are validated using the cross-validation method. Among different types of crossvalidation methods, the Shuffle-split-cross (SSC) validation method is particularly flexible. In this research, this method is utilized to validate the models, wherein data is split into multiple training and test sets. For this purpose, the shuffle-split function is imported from the sklearn model-selection library. This type of cross-validation offers several advantages, including the ability to control data as needed. The number of iterations can be controlled independently of training and test sizes, allowing the same part of the data to be used in each iteration by adjusting the train-size and test-size settings, which do not necessarily add up to one. However, these flexibilities prove useful when experimenting with large datasets. After validating the models, improvements in performance are necessary. This is achieved through parameter tuning, a crucial yet challenging task. To find the optimal parameter values for the models, grid search is one of the best methods. Grid search systematically explores all possible combinations of parameters, automatically selecting the most accurate combination to assess performance validity and comparison with other models. For this purpose, GridSearchCV is imported from the sklearn model-selection library. Hyperparameters, which are parameters provided to machine learning programs, play a crucial role. In CNN algorithms, "batch-size" and "epochs" are used as hyperparameters. The best combination is determined through the grid search method.

Confusion Matrix

The confusion matrix is a graphical representation commonly used in the field of neural networks for validating models. It provides a clear statistical analysis of the data, particularly in binary classification tasks, such as detecting caries or non-caries data, which is our focus. This matrix serves as a visual tool to understand the differences between true positive, true negative, false positive, and false negative outcomes. True Positive indicates teeth truly infected by dental caries, while True Negative represents samples correctly classified as non-infected, or healthy teeth. False Positive refers to samples incorrectly labeled as caries-infected when they are actually healthy, while False Negative occurs when the model misclassifies healthy teeth as infected. In our case, the confusion matrix for our test data is visually represented in Figure. 17, providing a comprehensive overview of the model's performance.

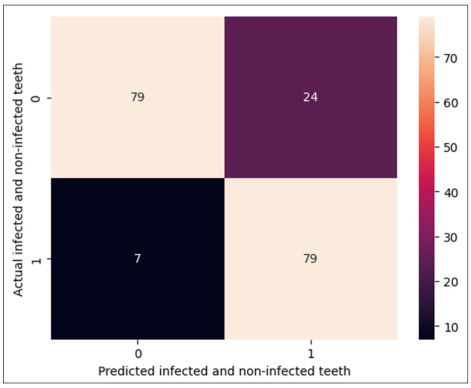


Figure 17: Confusion Matrix

ROC - Curve

The Receiver Operating Characteristic curve (ROC) is essentially a graphical representation that illustrates the performance of a model. Typically, in neural network models tasked with classifying images into binary or multiple classes, the results yield both real values and predicted values. These values serve as the basis for determining thresholds for the images. Researchers can utilize these thresholds to identify the maximum threshold value, aiming to achieve the highest possible accuracy.

This curve always utilizes two parameters: the actual label and the predicted label, also known as the True Positive Rate and False Positive Rate. Therefore, the ROC curve graph for the dental caries data is provided below in Figure 18.

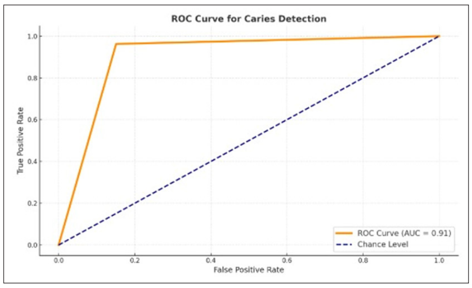


Figure 18: ROC Curve for Caries Detection

Discussion

The CNN model and transfer learning models employed for detecting dental caries demonstrate significant accuracy. Transfer learning exhibits superior accuracy compared to the CNN sequential model. The results of these models are validated through Shuffle Split-Cross validation, and hyper-parameter tuning is conducted via grid research. The studies conducted by Lee et al. [3] and Shreyansh et al. [17] are closely related to our work. Therefore, the contributions of our study are compared in Table 6 that highlights the novelty of our research.

Table 6: Comparison Of the Contributions of this and Related Studies

Features	Contribution by this study	Contribution by J.-H. Lee et. al. [3]	Contributions by Shreyansh et. al. [17]
Title	Enhancing Dental Caries Identification with Deep Learning: A Study of Convolution Neural Networks and Transfer Learning Approaches Applying Image Processing	Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm	Classification of Dental Diseases Using CNN and Transfer Learning
Data set preprocessing	image preprocessing through algorithms.	image preprocessing manually (Heading 2.1 and 2.2 of [3])	Not done
Steps on images preprocessing	1. Cleaning 2. Augmentation of the whole data 3. Editing, 4. Feature extraction 5. Segmentation 6. Splitting 7. Training 8. Testing 9. Validation	1. Cleaning 2. Augmentation only on training data, 3. Dataset splitting, 4. Model training and testing 5. Model evaluation	Data labelling only (not mentioned how)
Data Augmentation	Data augmentation performed • on whole data • using 7 filters	Data augmentation performed • on training data only • using 6 filters	Not done
Image editing	Algorithms were applied for image editing to enhance the quality of image • Increase/decrease brightness • Increase/decrease sharpness • Setting contrasts	Not done	Not done
Feature Extraction	The models employed for detecting dental caries automatically extracted features.	Learning features through network layers (not clearly mentioned)	Not done
Image segmentation and labelling	Images segmentation was performed by Random Walker Method that provided labeling: • Caries teeth • Non caries teeth	Not performed	Not performed
Model training and testing	Five models were trained and tested to obtain the results: • Sequential CNN model, • VGG-16, • VGG-19, • GoogleLeNet Inception V3 • GoogleLeNet ResNet50	Single model was used (Deep CNN based GoogleLeNet Inception V3)	• CNN • VGG-16

Results validation by comparison	To validate the accuracy of the obtained results, the five models were compared.	No comparison was done	Not done
Data validation	Two approaches were used: • Receiver Operating Characteristics Curve (ROC) • Confusion matrix	Single approach was used: • ROC	Not done
Data evaluation criteria	The evaluation was performed on the following criteria: • Accuracy • Sensitivity • Specificity • F1	The evaluation was performed on the following criteria: • Accuracy • Sensitivity • Specificity • F1 • MCC	Not done
Results	The five models showed accuracy: • CNN: 90% • VGG-16: 96% • VGG-19: 73% • Inception V3: 70% • ResNet50: 73%	82 % for premolar and molar	88%
Conclusion	Transfer learning, especially with VGG-16, achieved the highest accuracy (96%) in this study, outperforming both traditional CNN and related studies, highlighting its effectiveness for dental caries detection.	More improved deep-learning algorithms and high-quality and quantity datasets may be useful for dental caries detection and diagnosis in clinical dental practice.	Transfer learning with VGG16 pretrained model is used to achieve better accuracy.

Limitations

The study identifies several key limitations that need improvement for better and more accurate results. Firstly, the dataset size is notably small, which may not suffice for the convolutional neural network's requirements. Secondly, downscaled images are used as inputs to mitigate increased computational costs, training time, and storage space. Thirdly, while deep learning-based CNN methods demonstrate high accuracy and discriminatory power with high-resolution, large-scale images, this study's utilization of downscaled images may limit its potential. Fourthly, dental radiography presents challenges; grayscale images contain both light and dark regions, where distinguishing between various regions and shadow areas can be difficult due to improper camera placement. Lastly, the type of X-ray image poses challenges, particularly periapical X-rays, which capture images from crown to root and often exhibit ambiguous boundaries between bone and teeth. Additionally, teeth in periapical images may be diversely rotated, rendering image processing more challenging.

Experimental Setup

This section discusses the experimental setup used for implementing the proposed models. It covers details about the tools, hardware, and environment utilized. The tool employed for executing both the CNN model and transfer learning models is Python, with the environment being Google Colab. Python version 3.0 is utilized, with the Keras library used on the TensorFlow backend. TensorFlow is a free and open-source software library for dataflow, primarily used for machine learning applications like neural networks. Its import is essential as it serves as the backend for the implementation setup. Keras, on the other hand, is an open-source neural network library written in Python. It provides a straightforward workflow for training and evaluating models, comprising stages such as creation, configuration, training, and evaluation of the model. For implementation, the sequential model is imported from 'keras.models()', while layers such as Conv2D, MaxPooling2D, Sequential, Activation, Dropout, Flatten, Dense, Input, Lambda, Flatten, etc., are imported from 'keras.layers()'. Additionally, for creating the model, the Image Data Generator is imported from 'keras.preprocessing.image()'.

Other libraries like NumPy, which adds support for large, multi-dimensional arrays and matrices, along with a vast collection of high-level mathematical functions to operate on these arrays, and Pandas, a software library for data manipulation and analysis in the Python programming language, are also imported. Additionally, Matplotlib, a plotting library for Python, and its numerical mathematics extension NumPy, as well as Scikit-learn (Sklearn), a machine learning library written in Python, are imported. For the implementation of transfer learning models, additional packages are required for execution. One such package, VGG16, is imported from 'keras.application. vgg16()'. For image preprocessing, the 'preprocess_input' function is also imported from 'keras.application. vgg16()'. Similarly, VGG19(), ResNet(), and InceptionV3() are imported for VGG19, ResNet50, and InceptionV3 respectively. Regarding hardware specifications needed for the implementation, a laptop with the following configuration is required: an operating system (32-bit, x64-based processor), Intel® Core™ CPU (Intel® Core™ i3-3340 processor), CPU clocked at 2.70 GHz, Intel® integrated graphics, and 4.00GB of RAM.

Conclusion

Dental caries is among the most prevalent oral diseases across all age groups, underscoring the importance of early and accurate detection. The primary objective of this study is to diagnose early and proximal caries from periapical radiographs, employing CNN and transfer learning algorithms. The research is structured into several key stages: data augmentation, image preprocessing, segmentation, feature extraction, model training, and validation. Both CNN and transfer learning models are utilized for the detection of dental caries from periapical radiographs, leveraging a publicly available dataset comprising only 120 x-ray images. Transfer learning with VGG16 achieves the highest accuracy among all the models explored, reaching 96%. The experimental results of these models are discussed in detail in the results section, with validation conducted through shuffle-split-cross validation.

The results of this study demonstrate that transfer learning, particularly with the VGG-16 model, achieved the highest accuracy (96%) for dental caries detection, outperforming both the standard CNN (90%) and other pretrained models such as VGG-19, Inception V3, and ResNet50. Compared to related studies, which reported accuracies of 82% [3] and 88% [17], the VGG-16 model in this study shows a notable improvement. These findings suggest that well-selected transfer learning architectures, especially VGG-16, can significantly enhance performance in dental caries detection. Future work may further benefit from incorporating larger and more diverse datasets, as highlighted in previous studies

Acknowledgement
The authors would like to thank China's National Key R & D Program for providing the experimental facilities used to perform these experiments. The authors would also like to thank Artificial Intelligence and Data Analytics Lab CCIS Prince Sultan University for their support.

Funding Statement
This study is partly supported by the National Key R&D Program of China with project no. 2023YFB2704601.

Conflicts of Interest
The authors confirm that there are no conflicts of interest to disclose concerning the current study.

References
1. Jeong H, Kim J, Lee J, Kim D. Tooth-Related Disease Detection System Based on Panoramic Images and Optimization Through Automation: Development Study. JMIR Med Inform. 2022. 10: e38640.
2. Choi J, Eun H, Kim C. Boosting proximal dental caries detection via combination of variational methods and convolutional neural network. J Signal Process Syst. 2018. 90: 87-97.
3. Lee HJ, Kim HD, Jeong NS, Cho HS. Detection and diagnosis of dental caries using a deep learning based convolutional neural network algorithm. J Dent. 2018. 77: 106-111.
4. Litjens G, Kooi T, Bejnordi EB, Setio AAA, Ciompi F, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017. 42: 60-88.
5. Almazrooa S, Almotairy A, Alzahrani A. A Survey of Dental Caries Segmentation and Detection Techniques. J Dent. 2021. 100: 103425.
6. Ko YH, Kang MS, Kim EH, Kwon KH, Kim IB. Validation of quantitative light-induced fluorescence digital (QLF-D) for the detection of approximal caries in vitro. J Dent. 2015. 43: 568-575.
7. Casalegno F, Newton T, Daher R, Abdelaziz M, Lodi-Rizzini A, et al. Caries detection with near-infrared trans-illumination using deep learning. J Dent Res. 2019. 98: 1227-1233.
8. Ko YH, Kang MS, Kim EH, Kwon KH, Kim IB. Validation of quantitative light-induced fluorescence digital (QLF-D) for the detection of approximal caries in vitro. J Dent. 2015. 43: 568-575.
9. Mehta A, Bhalla S. Assessing consequences of untreated carious lesions using pufa index among 5-6 years old school children in Chandigarh city. Indian J Dent Res. 2014. 25: 150-153.
10. Patil S, Kulkarni V, Bhise A. Algorithmic analysis for dental caries detection using adaptive neural network architecture. Heliyon. 2019. 5: e01579.
11. Lee HJ, Kim HD, Jeong NS. Diagnosis of cystic lesions using panoramic and cone beam computed tomographic images based on deep learning neural network. Oral Dis. 2020. 26: 152-158.
12. Xu X, Liu C, Zheng Y. 3D tooth segmentation and labeling using deep convolutional neural networks. IEEE Trans Vis Comput Graph. 2018. 25: 2336-2348.
13. Lee HJ, Kim HD, Jeong NS, Choi HS. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. J Periodontal Implant Sci. 2018. 48: 114-123.
14. Rajpurkar P, Irvin J, Ball RL. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform. 2019. 20: 1024-1033.
15. Eun H, Kim C. Oriented tooth localization for periapical dental X-ray images via convolutional neural network. In: Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Jeju, South Korea. 2016. 1-7.
16. De Tobel J, Radesh P, Vandermeulen D, Thevissen WP. An automated technique to stage lower third molar development on panoramic radiographs for age estimation. A pilot study. J Forensic Odonto-Stomatol. 2017. 35: 42.
17. Prajapati AS, Nagaraj R, Mitra S. Classification of dental diseases using CNN and transfer learning. In: 2017 5th International Symposium on Computational and Business Intelligence (ISCBI), IEEE, Dubai, United Arab Emirates. 2017. 70-74.
18. Cantu AG, Gehrung S, Krois J, Chaurasia A, Rossi JG, et al. Detecting caries lesions of different radiographic extension on bitewings using deep learning. J Dent. 2020. 100: 103425.
19. Betul A, Oktay. Tooth detection with convolutional neural networks. In: 2017 Medical Technologies National Congress (TIPTEKNO), Trabzon, Turkey. 2017. 1-4.
20. Naam J, Harlan J, Madenda S, Wibowo EP. Identification of the proximal caries of dental X-ray image with multiple morphology gradient method. Int J Adv Sci Eng Technol. 2016. 6: 343-346.
21. Osterloh D, Viriri S. Caries detection in non-standardized periapical dental X-rays. In: Computer Aided Intervention and Diagnostics in Clinical and Medical Images, 1st Ed, Springer Nature Switzerland AG. 2019. 143-152.

22. Erickson JB, Korfiatis P, Akkus Z, Kline LT. Machine learning for medical imaging. RadioGraphics. 2017. 37: 505-515.
23. Moutselos K, Berdouses E, Oulis C, Maglogiannis I. Recognizing occlusal caries in dental intraoral images using deep learning. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Berlin, Germany. 2019. 1617-1620.
24. Tuzoff VD, Tuzova NL, Bornstein MM, Krasnov SA, Kharchenko AM, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. Dentomaxillofac Radiol. 2019. 48: 20180051.
25. Kim EJ, Nam EN, Shim SJ, Jung HY, Cho HB, et al. Transfer learning via deep neural networks for implant fixture system classification using periapical radiographs. J Clin Med. 2020. 9: 1117.
26. Zhang Y, Wang X, Li H, Ni C, Du Z, et al. Human oral microbiota and its modulation for oral health. Biomed Pharmacother. 2018. 99: 883-893.
27. AL-Malaisa AS. Detection of dental disease through X-ray images using neural search architecture network. Comput Intell Neurosci. 2023. 7395392.
28. Chen X, Guo J. Detection of proximal caries lesions on bitewing radiographs using deep learning method. Dentomaxillofac Radiol. 2023. 56: 455-463.
29. Khan A, Sohail A, Zahoora U, Qureshi SA. A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev. 2020. 53: 1-62.
30. Vinayahalinga S, Xi T, Berge S, Maal T, De Jong G. Automated detection of third molars and mandibular nerve by deep learning. Sci Rep. 2019. 9: 1-7.
31. Yun C, Zhiyan L, Chong Z, Jing L, Xin Z, et al. Illumina-based sequencing analysis of pathogenic microorganisms in dental caries patients of different Chinese ethnic groups. J Int Med Res. 2019. 47: 5037-5047.
32. Zhang J, Liu B, Wu J, Wang Z, Li J. DeepCAC: a deep learning approach on DNA transcription factors classification based on multi-head self-attention and concatenate convolutional neural network. BMC Bioinformatics. 2023. 24: 345.
33. Akhtar S, Adeel M, Iqbal M, Namoun A, Tufail A, et al. Deep learning methods utilization in electric power systems. Energy Rep. 2023. 10: 2138-2151.
34. Mahmood T, Li J, Pei Y, Akhtar F, Rehman MU, et al. Breast lesions classifications of mammographic images using a deep convolutional neural network-based approach. PLoS ONE. 2022. 17: e0265291.
35. Liu B, Li Q, Zhang W, Liu T. Deep learning-based facial emotion recognition: a survey. IEEE Access. 2022. 10: 124542-124563.