IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 26, 2022, accepted February 6, 2022, date of publication February 9, 2022, date of current version February 18, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3150358

Dental Caries Detection Using Score-Based Multi-Input Deep Convolutional Neural Network

ANDAC IMAK[1], ADALET CELEBI[2], KAMRAN SIDDIQUE[3], (Senior Member, IEEE), MUAMMER TURKOGLU[4], ABDULKADIR SENGUR[5], AND IFTEKHAR SALAM[3], (Member, IEEE)

[1]Department of Electrical and Electronic Engineering, Faculty of Engineering, Munzur University, 62000 Tunceli, Turkey
[2]Oral and Maxillofacial Surgery Department, Faculty of Dentistry, Bingol University, 12000 Bingol, Turkey
[3]Department of Information and Communication Technology, School of Electrical and Computer Engineering, Xiamen University Malaysia, Sepang 43900, Malaysia
[4]Department of Software Engineering, Faculty of Engineering, Samsun University, 55000 Samsun, Turkey
[5]Department of Electrical and Electronic Engineering, Faculty of Technology, Firat University, 23100 Elazig, Turkey

Corresponding authors: Kamran Siddique (kamran.siddique@xmu.edu.my) and Iftekhar Salam (iftekhar.salam@xmu.edu.my)

This work was supported by the Xiamen University Malaysia Research Fund under Grant XMUMRF/2019-C3/IECE/0005.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Firat University Noninterventional Clinical Research Ethics Committee under Approval No. 2020/1715.

ABSTRACT Panoramic and periapical radiograph tools help dentists in diagnosing the most common dental diseases, such as dental caries. Generally, dental caries is manually diagnosed by dentists based on panoramic and periapical images. For several reasons, such as carelessness caused by heavy workload and inexperience, manual diagnosis may cause unnoticeable dental caries. Thus, computer-based intelligent vision systems supported by machine learning and image processing techniques are needed to prevent these negativities. This study proposed a novel approach for the automatic diagnosis of dental caries based on periapical images. The proposed procedure used a multi-input deep convolutional neural network ensemble (MI-DCNNE) model. Specifically, a score-based ensemble scheme was employed to increase the achievement of the proposed MI-DCNNE method. The inputs to the proposed approach were both raw periapical images and an enhanced form of it. The score fusion was carried out in the Softmax layer of the proposed multi-input CNN architecture. In the experimental works, a periapical image dataset (340 images) covering both caries and non-caries images were used for the performance evaluation of the proposed method. According to the results, it was seen that the proposed model is quite successful in the diagnosis of dental caries. The reported accuracy score is 99.13%. This result shows that the proposed MI-DCNNE model can effectively contribute to the classification of dental caries.

INDEX TERMS Dental caries detection, score-based fusion, deep convolutional neural network, classification, periapical images.

I. INTRODUCTION

Dental caries is one of the most common dental diseases. Dental caries result from a complex interaction between acid-producing bacteria that adhere to the dental and fermentable carbohydrates. Acids in dental plaque can demineralize enamel and dentine in the cracks and smooth surfaces of the dental. The earliest visual sign of dental caries is the white spot lesion. If demineralization continues, white dot surfaces get pitted, causing small holes called cavities to form. It is particularly typical in children, teenagers, and older adults. However, anyone who has teeth, including babies, is likely to have dental caries. If cavities are left untreated, they can grow and affect the deeper layers of your teeth. They can cause severe toothache, infection, and dental loss [1], [2]. Periapical radiographs (peri means ''around'' and apical means ''end of the dental root'') record images of the outlines, location, and mesiodistal size of the teeth and surrounding tissues.

The associate editor coordinating the review of this manuscript and approving it for publication was Tyson Brooks.

The intraoral periapical examination obtains the view of the entire dental and surrounding structures [3]. Intraoral periapical radiography is a widely used intraoral imaging technique in dental radiology and provides considerable information about the teeth and surrounding bone. The film presents vital information to aid in the diagnosis of the most common dental diseases, especially dental caries, dental abscesses, periodontal bone loss, or gum disease, and shows the teeth and surrounding alveolar bone, all dental coatings, and dental roots. The application can supply an opportunity to identify additional significant findings, including the status of restorations, impacted teeth or broken dental fragments, and variations in dental and bone anatomy [3], [4].

Caries detection based on panoramic and periapical imaging techniques is performed manually by dentists. Mostly the caries dental diagnosis seen on the radiography cannot be made correctly because the existing dental caries can be overlooked due to the physician's inexperience or intense patient load. This situation can lead to the progression of caries, advanced dental infections, and dental loss. Automatic systems, which have been developed based on machine learning and image processing techniques to avoid these negativities and facilitate dentists' diagnoses, have gained considerable importance. Dental caries, periodontitis, interdental bone loss, gingivitis, impacted teeth, and cysts are some dental diseases. In the literature, there are few studies for detecting dental diseases. In these studies, the subjects such as caries detection (see 2nd literature review) and cysts detection [23]–[25] have been discussed. In most of these studies, traditional machine learning-based feature extraction methods were used to detect dental diseases [5], [7], [10], [11], [13], [15], [17]. In other studies [6], [8], [9], [12], [16], segmentation and classification processes were performed using deep learning models. Besides, there are few studies based on deep CNN architectures on dental x-ray images. Also, applications on dental caries detection using deep CNN architectures are even more limited in quantity and inadequate in this area.

This study presented a novel score-based multi-input deep convolutional neural network ensemble to detect dental caries using periapical teeth images. The proposed model used image processing techniques to sharpen raw caries images. Later, the multi-input deep convolutional neural network was designed using raw and processed images, and the training process was performed. Finally, the Softmax layers of the multi-input CNN architecture were combined with the score level-based algorithm. In experimental studies, 340 periapical images, including caries and non-caries, were used. These results have revealed that the model proposed for the classification of dental caries will provide remarkable contributions to dentists' performances.

This work presents a new dental caries detection model based on the multi-CNN ensemble. The major contributions of the proposed model are:

- The sharpening filter and intensity colormap procedures, which are used in this study, further highlight the decayed area in the unprocessed dental images.
- Although there are studies on dental caries in the literature, there is no publicly available data set. In this paper, a new dataset consisting of 340 images (caries and non-caries) was presented, and this dataset was made freely accessible for researchers.
- Pre-trained AlexNet architecture based on the transfer learning approach in dental caries detection was adapted. In this study, instead of starting the training process with random weights using a data set containing a limited number of images, the learned weights of pre-trained architecture were used. In the experimental results, it was observed that the transfer learning based AlexNet model used for the proposed model increased the success by approximately 10% compared to the Training from scratch approach.
- The proposed score-based multi-input CNN model has achieved notable success in detecting dental caries with an accuracy score of 99.13%. This result showed the usability of the proposed model in real-time applications.

The purpose of the present study was to determine the ability of deep convolutional neural networks to assist dentists in the automatic diagnosis of dental caries based on periapical images.

II. LITERATURE REVIEW

In the introduction part, general studies on dental disease detection were mentioned. Few studies on the detection of dental caries are available in the literature, and in these studies, commonly panoramic and periapical imaging systems have been used. In the research, pre-trained deep models have been used commonly, besides traditional feature extraction methods. Geetha *et al.* obtained dental X-ray images using the Laplacian filtering, window-based adaptive threshold, and back-propagation neural network. The resulting new views from a few statistical feature extractions were classified by the back-propagation neural network. A success of 97.1% was achieved in the caries detection using periapical 105 dental x-ray images [5]. Prajapati *et al.* used a transfer learning approach based on the pre-trained VGG16 model on 251 periapical dental X-rays. In the study performed for three distinct classes, 87.5% performance was obtained in the images of caries from the fully trained network architecture [6]. In their study, Singh and Sehgal proposed an automated caries detection system based on Radon Transformation (RT) and Discrete Cosine Transformation (DCT). The selected features were extracted using the PCA technique. Finally, the model was applied to the Random Forest classifier, and an accuracy of 86% was obtained [7]. Casalegno *et al.* have proposed a method for the detection of caries in dental images obtained with near-infrared transillumination (TI) imaging technology. With the proposed convolutional neural network architecture, ROC areas of 83.6% and 85.6% were obtained for occlusal and proximal caries lesions, respectively [8]. Cantu *et al.* conducted a study on the detection of caries lesions in 3686 bitewing radiographs. The U-Net

convolutional neural network architecture was used on dental caries. The proposed neural network showed an accuracy of 80% [9]. In the study by Datta *et al.*, special neutrosophic domain-based features were used. This proposed feature extraction method was obtained by combining certain features. Caries was detected from 120 periapical X-ray images using the active contour technique. As a result of the experimental studies, the performance was over 92% [10]. Tuan *et al.* proposed a hybrid approach consisting of segmentation, classification, and decision-making phases in their studies. A segmentation method based on semi-supervised fuzzy clustering is used. Besides, classification was carried out with a new graphic-based clustering technique. The method's performance recommended for five distinct carries classes has been observed as 90% [11]. Lakshimi and Chitra conducted a study to detect caries using 1700 panoramic dental x-rays. Performance results were compared with traditional segmentation techniques and deep CNN architecture. A success of 98.6% was achieved with the graph cut and deep CNN method [12]. The study by Naam *et al.* aimed to facilitate the identification of caries for experts in panoramic dental x-ray images and increase their accuracy. Multiple morphology gradient algorithms have been developed, besides well-known approaches in image processing techniques such as cropping, morphology dilation, and morphology erosion. The improved pre-treatment stage facilitated the caries detection by clarifying the dental x-ray image better [13]. Oprea *et al.* proposed a method to examine the size of the existing caries lesion and then classify the type of caries found in dental radiography. Segmentation of the dental was performed using the edge detection method in the x-ray. Later, they performed a study by determining the pixel number of caries (black area) affected on the dental and classifying them according to their size [14]. Obuchowicz *et al.* made comparisons in caries detection by using specific texture feature mapping methods. Original images were pre-processed by applying the Histogram Stretching (HSTR) algorithm. As a result of the comparison in the texture feature map methods, it was observed that the most useful information was provided by using the energy parameter. Besides, it was stated that the entropy parameter also performed well in visualizing the areas affected by caries [15]. Lee *et al.* used GoogLeNet Inception v3 CNN architecture to detect dental caries from periapical dental X-ray images. In the proposed study, a periapical data set consisting of 3000 images were used for 2400 training and 600 tests. In the results, the success rate of premolar, molar, and both premolar and molar caries models were 89%, 88%, and 82%, respectively [16]. Naam *et al.* aimed to improve the quality of X-ray images in determining proximal caries by experts. In the dataset they created, there are dental x-ray images containing 27 proximal dental caries. With the proposed Multiple Morphological Gradients method, the object edges have been clarified by producing clear images. The results suggest the model can be used to define appropriately proximal caries and progression levels [17]. Leo and Reddy [18] proposed a hybrid model consisting of Artificial Neural Network and Deep Neural Network techniques. In experimental studies, they used 300 training and 180 testing dental x-ray images. The proposed system yielded a 96% success for dental caries classification [18]. Huang and Lee performed a study on caries detection using optical coherence tomography (OCT) images as a different imaging technique. A total of 748 single tooth images were extracted using the cropping process from 63 OCT tooth images. In experimental studies carried out based on different CNN architectures, 599 images were used for training and the rest for testing. As a result, they achieved a 95.21% success rate with the ResNet-152 architecture [19]. Bui *et al.* proposed a hybrid model for the detection of caries based on support vector machine (SVM) classifier using deep and geometric features. In experimental studies, they used 229 carries, and 304 non-caries images. As a result, the accuracy score of the proposed system was obtained as 91.7% [20]. Majanga and Viriri developed a system based on morphological methods and deep learning. In the preprocessing stage, the boundaries of tooth images were determined by using thresholding methods with Gaussian blur filters. Later, the classification of dental images (caries/non-caries) was carried out by using a certain threshold value with features obtained based on deep learning. They used 11,114 tooth images obtained using data augmentation methods based on periapical dental X-ray images. As a result, the proposed model for caries detection was 97% successful [21]. Vinayahalingam *et al.* carried out a study based on the evaluation of the classification performance of deep learning architectures using panoramic x-ray images. 400 cropped panoramic x-ray images were used for training CNN architectures. In experimental studies, it was observed that they achieved 87% performance with the MobileNet V2 architecture [22].

III. PROPOSED METHODOLOGY

This study presented a novel score-based multi-input CNN ensemble (MI-DCNNE) to detect dental caries. The proposed system consists of three phases: Pre-processing, Deep Convolutional Neural Network, and score-based fusion. In the preprocessing stage, filters are used to clarify the panoramic raw images. Later, a multi-input Convolutional Neural Network Model based on a pre-trained deep model is designed. In the last stage, the developed multi-input CNN model is combined at a score level. The general structure of the proposed system, including all these stages, is given in Figure 1. The stages composing the proposed system's general structure and dataset are detailed in subtitles.

A. DATASET

In this study, images obtained from two private oral and dental health clinics using periapical radiography devices (Planmeca Intra, Helsinki, Finland), (MyRay, Imola, Italy), as in retrospective studies, were evaluated. A total of 380 periapical images belonging to 310 patients were analyzed. 40 images including blurred images or images with fractures, cysts,

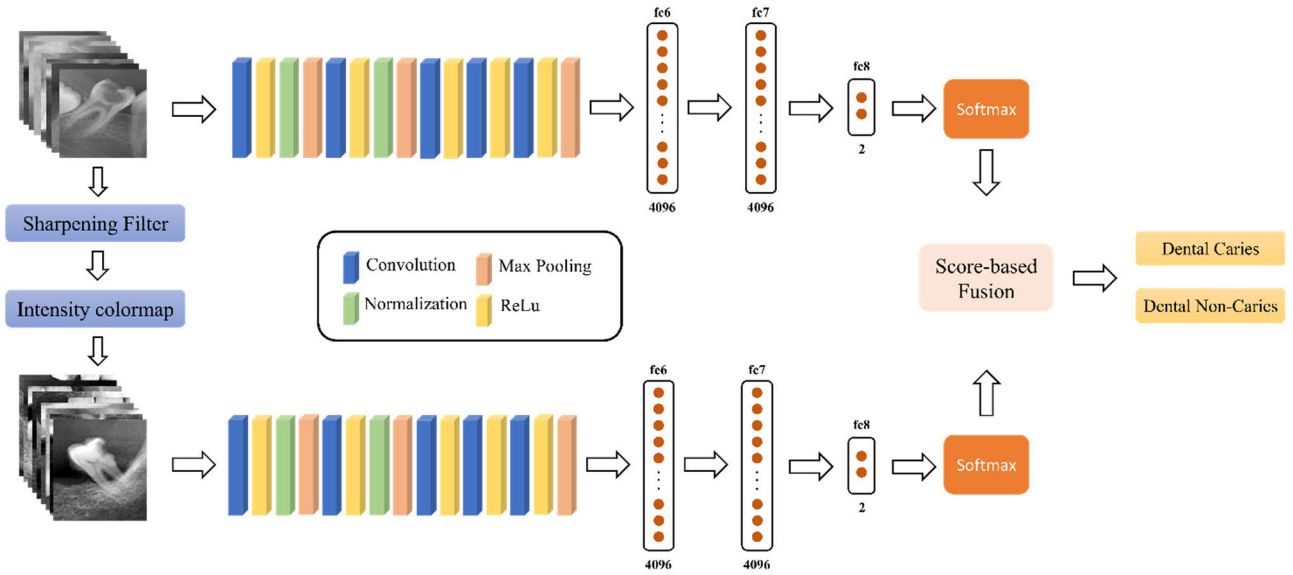


FIGURE 1. Structure of proposed architecture.

and infection were excluded from the study. 340 periapical images out of 380 images were included in the study. A data set containing periapical radiography images obtained from patients was created. These radiographs included images of 150 carious teeth and 190 non-carious teeth. This dataset was obtained under control by a dentist. These images were gray channels and variable sizes. Figure 2 shows some samples of dental periapical images from the dataset.

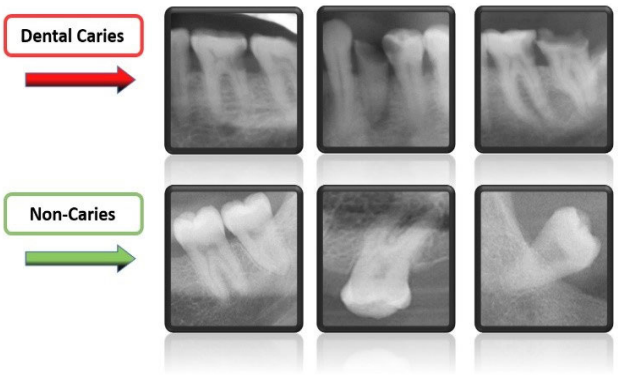


FIGURE 2. Dental periapical images used in this study.

B. PRE-PROCESSING

In the current study, panoramic images were used to classify dental caries. Several image processing techniques were applied to these raw images. This process has clarified the carious areas in the dental images and increased the classification performance. To this end, a sharpening filter and intensity colormap were applied for panoramic raw images, respectively. As a result, the same size new images were obtained. An example of this process is given in Figure 3. As seen in Figure 3, the sharpening filter was applied for raw gray images as $[-1\ -1\ -1;\ -1\ 9\ -1;\ -1\ -1\ -1]$. Then, the intensity values of the images were adjusted. In this way, the contrast of the new image was increased. Eventually, the caries zone was observed to become sharper in the recent pictures obtained as a result of filtering and density value adjustment processes.

C. DEEP CONVOLUTIONAL NEURAL NETWORK

A Convolutional Neural Network (CNN) has a multi-layer network structure trained in a specific system for the process of feature extraction and classification. It replaces time-consuming traditional methods of feature extraction and classification. The CNN architecture has a basic structure consisting of the input layer, multiple hidden layers, and output layers. Besides, the hidden layer consists of convolutional layers, pooling layers, and fully connected layers. High-level features obtained by filtering from input images are processed by convolutional and pooling layers. If we use a two-dimensional image as input, a two-dimensional kernel is used. Convolution is represented using the parameters shown in Equation (1).

$$X_j^l = f\left(\sum_{i \in M_j} X_i^{l-1}.k_{ij}^l + b_j^l\right) \tag{1}$$

where, the feature map of the previous layers is represented by the expression x_i^{l-1}. M_j, k_{ij}^l and b_j^l represent input map selection, convolution kernel and trainable additive bias terms, respectively. The Convolutional process is passed through an f activation function [26]. Filters are usually moved over the image with the specified stride (s). For an input image with width A, height B and channel C ($A_m \times B_m \times C_m$); with $r \times r$ size C_0 filters, the output volume will be

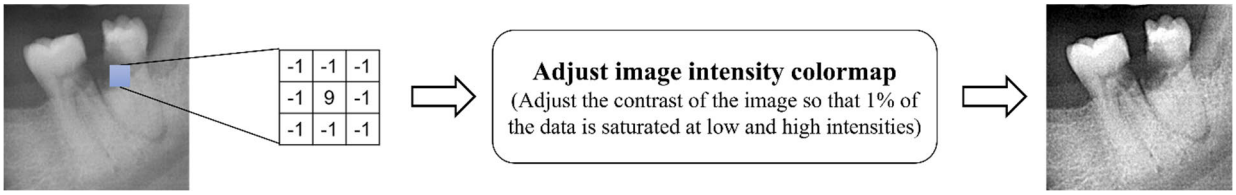


FIGURE 3. The proposed pre-processing phase.

$A_0 \times B_0 \times C_0$. As a result of the filtering, the new dimensions shown in Equation (2) are obtained;

$$A_0 = B_0 = ((m - r + 2p)/s) + 1 \tag{2}$$

where p represents the extra pixels to be added to the input matrix. The activation function enhances the nonlinearity in the network. Many studies in the literature use the Rectified Linear Unit (ReLU) function shown in Equation (3);

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \tag{3}$$

Pooling, also known as the sub-sampling layer, helps regulate over-compliance. The dimensions of output maps are reduced by using maximum, minimum, and average values. Finally, the full connect layer reduces the multidimensional property map to a one-dimensional vector [27].

In this study, pre-learned weights of the AlexNet architecture for each of the multi-input images and a transfer learning approach to adapt this architecture in dental caries detection have been used. In this process, new fully connected, Softmax, and classification layers have replaced the last three layers of pre-trained AlexNet architecture. Additionally, the proposed model has used a score-based fusion process in the Softmax output layers of multiple CNN architectures. The general structure and characteristic features of the proposed model are detailed in Table 1.

D. SCORE-BASED FUSION

In the proposed system, an MxN output matrix is obtained from the softmax layers in deep CNN architectures applied for each multi-input image [28]. Here, N and M are represented as the number of classes and the number of test sets. These CNN models were trained individually with transfer learning fashion.

Later, Equation 4 is used to calculate the new score value;

$$score = \sum_{i=1}^{M} \sum_{j=1}^{N} (A_1(i,j) + A_2(i,j))/2) \tag{4}$$

where A represents the score generated by the CNN-based models. A new score in MXN size was obtained using Equation 4. Later, using the new score value, the class label is determined according to the maximum value of each line (M). In the proposed system, two-CNN models based on two input images are used. In this context, the pseudocode of this approach is demonstrated in Algorithm Table 2.

TABLE 1. The proposed Multi-input AlexNet architecture.

	Name	Type	Filter Size/Stride	Output Size
1	Input_1/Input_2	Image	-	227x227x3
2	Conv1_1/Conv1_2	Convolution	11x11/4	55x55x96
3	Pool1_1/Pool1_2	Max Pooling	3x3/2	27x27x96
4	Conv2_1/Conv2_2	Convolution	5x5/1	55x55x256
5	Pool2_1/Pool2_2	Max Pooling	3x3/2	27x27x256
6	Conv3_1/Conv3_2	Convolution	3x3/2	13x13x512
7	Conv4_1/Conv4_2	Convolution	3x3/1	13x13x1024
8	Conv5_1/Conv5_2	Convolution	3x3/1	13x13x512
9	Pool3_1/Pool3_2	Max Pooling	3x3/2	6x6x256
10	Fc6_1 / Fc6_2	FC	-	4096
11	Fc7_1 / Fc7_2	FC	-	4096
12	Fc8_1 / Fc8_2	FC	-	2
13	Softmax_1/Softmax_2	Softmax	-	2
14	Score level	Fusion	-	2
15	Classoutput	Classification Output	-	-

TABLE 2. The pseudocode of the proposed score-based approach.

Input: score1, score2, YTest
Output: accuracy

```
for i=1 to M
    for j=1 to N
        new_score(i,j)=(score1(i,j)+score2(i,j))/2
    end for j
end for i
for k=1 to K
    if score(k,1)>=score(i,2)
        YPred(k)=1;
    else
        YPred(k)=2;
end for k
accuracy=mean (YPred==YTest);
```

IV. RESULTS

The experimental works were carried out on a computer equipped with the NVIDIA Quadro M4000 GPU with 8GB RAM. For all coding, MATLAB software was used. The dental periapical image dataset was collected in private dental clinics. All ethical issues were guaranteed.

In this study, a novel score-based multi-input CNN ensemble (MI-DCNNE) model was proposed to detect dental caries. The weights of pre-trained AlexNet architecture were used for multiple entry images in the proposed model. Later, the

softmax layer of the multi-input CNN model based on the score level-based algorithm was combined and the class tags were determined. The deep parameters used in the experimental studies for training the CNN architecture are given in Table 3.

TABLE 3. The deep network parameters used in the current study.

Mini-batch size	: 8-16
Maximum epoch number	: 5-20
Weight decay factor	: 0.01
Initial learning rate	: 0.0001
Optimization method	: SGDM (Stochastic Gradient Descent with Momentum)

As seen in Table 3, in experimental studies, the epoch number and batch size were investigated in values between 5-20 and 8-16, respectively. In these empirical studies, 70% was used for training and the rest for testing. This random separation was done only once, and all experimental studies used the same training/testing datasets. Also, the test procedure was repeated 100 times for the results of all experimental studies. The AlexNet architecture based on raw images, the AlexNet architecture based on filter approach, and the proposed architecture performance have been calculated separately by using these processes. These results are given in Table 4.

TABLE 4. Results (%) of experimental works-I (Transfer learning).

Methods	Minimum	Maximum	Mean	Standard Deviation
AlexNet	86.96	98.26	94.81	1.8309
Filter-AlexNet	87.83	98.26	94.33	2.0662
Proposed model	**93.04**	**99.13**	**96.19**	**1.5025**

The minimum, maximum, mean, and standard deviation values of the results obtained from the AlexNet, Filter-AlexNet, and proposed model through 100 times of running are given in Table 4. According to these results, the maximum accuracy score for the AlexNet model based on raw images and filtered images was 98.26% for both methods. As seen from Table 4, the proposed score-based multi-input CNN ensemble model obtained a 99.13% accuracy score. Additionally, the other performance measures such as sensitivity, specificity, precision, and F1_score values were 98%, 100%, 100%, and 98.99%, respectively. The confusion matrix and ROC diagram of this accuracy result is also given in Figure 4.

As shown from the confusion matrix, the proposed model achieved a 100% accuracy score for the non-caries class and 98% in the dental caries detection. Besides, the training accuracy and loss curves of the proposed method was given

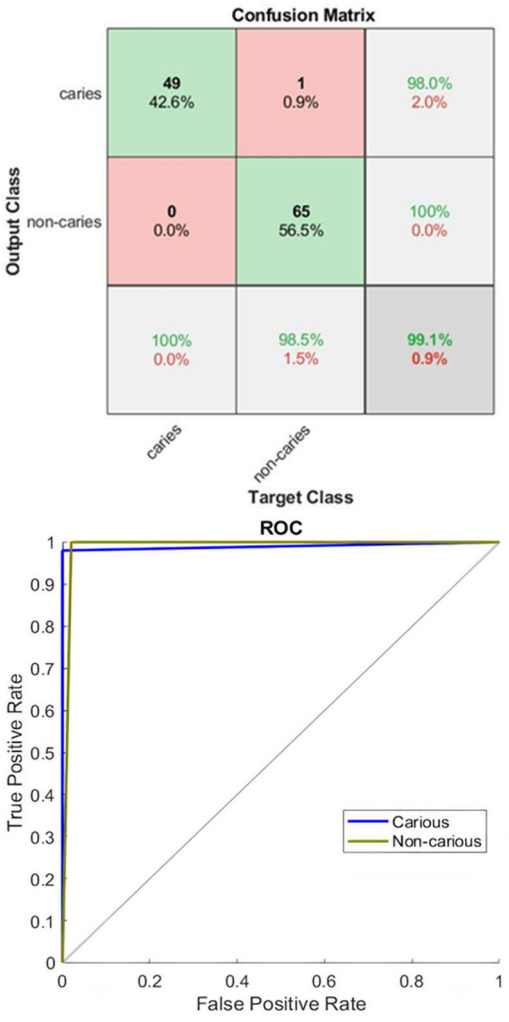


FIGURE 4. The confusion matrix and ROC diagram of the proposed model.

in Figure 5. As seen in Figure 5, the training accuracy was around 75% in the beginning and it dropped to almost 40% accuracy score in the first epoch. And, it suddenly increased through 90% accuracy score in the first epoch. The training accuracy curve increased to almost 100% accuracy score after fourth epoch. Similarly, the loss curve was around 1 in the beginning of the training procedure and it increased to around 3 in the first epoch. It gradually decreased through zero after fourth epoch.

In addition, to achieve interpretability of proposed model, we investigated network predictions using the gradient-weighted class activation mapping (Grad-CAM). The activation map for carious images is given in Figure 6. As seen in Figure 6, four test images and their corresponding output heat maps were given. From these heat maps, it was seen that the trained deep CNN model could learn the dental carries in the given panoramic dental images. The red regions in the heat maps showed the dental carries that the proposed model has learned.

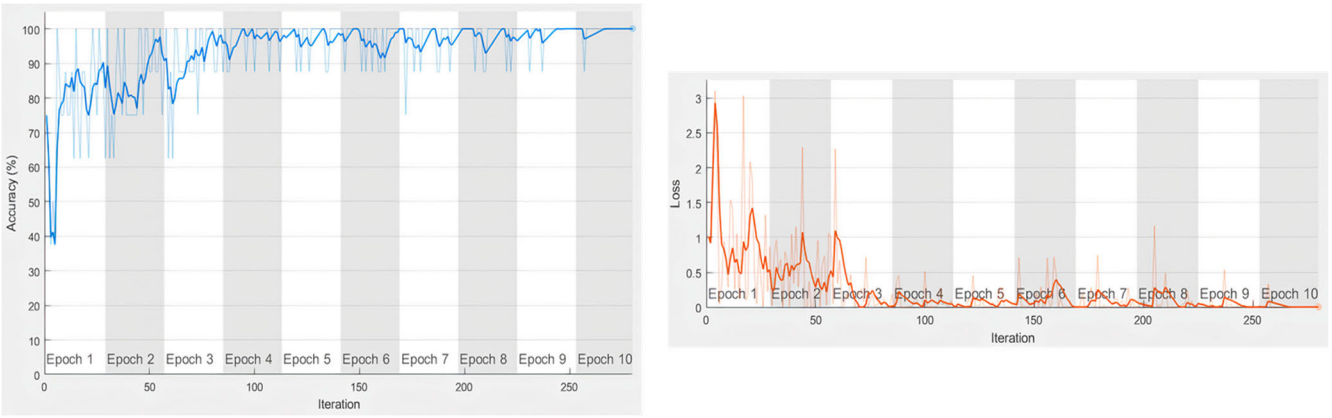


FIGURE 5. The graphs of training accuracy and loss against increasing iteration numbers of the proposed model.

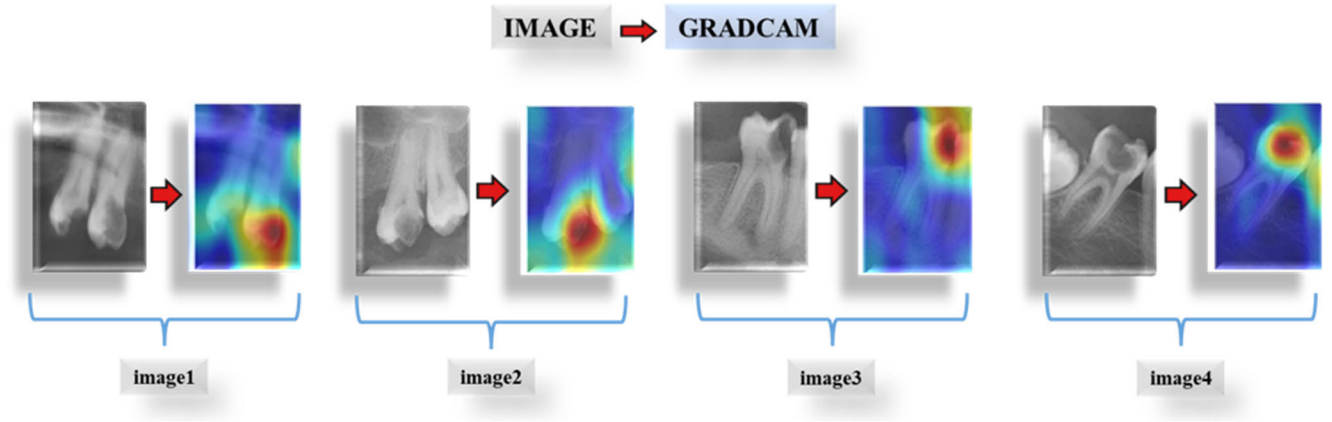


FIGURE 6. The examples based on activation maps of carious image.

A. EFFECT OF PRE-TRAINED PARAMETERS

Learned weights of a pre-trained AlexNet architecture were used in the proposed model. In this study, the pre-trained deep model has been used because it contains parameters trained by using an extensive multi-class data set [38]. In this study, a deep architecture based on the transfer learning approach was used and adapted to our work. The training from the scratch model was used for dental caries' classification to test the effect of this process. More precisely, the process started with random weights instead of trained parameters of the deep architecture. Accuracy scores obtained from these experimental studies are given in Table 5.

As seen from Table 5, the proposed model achieved the best accuracy score with 92.17%. Additionally, raw images based on AlexNet and filtered images based on AlexNet reached mean accuracy scores of 86.35% and 86.25%, respectively. According to all these results, the AlexNet model, which has been based on the transfer learning approach and used for the proposed model, provided approximately 10% more success than the Training from scratch approach. Besides, this performance rise was observed for other methods. Based on these results, it has been proven that using trained parameters based on pre-trained deep models for classification problems with a small data set provides a significant performance increase.

TABLE 5. Results (%) of experimental works-II (Training from scratch).

Methods	Accuracy score (%)			
	Minimum	Maximum	Mean	Standard Deviation
AlexNet	78.26	90.43	86.35	2.9083
Filter-AlexNet	74.78	90.43	86.25	2.8402
Proposed model	**81.73**	**92.17**	**87.57**	**2.0891**

B. COMPARISON OF THE STATE-OF-ART

There are many Deep convolutional Neural Network models developed for object recognition and classification in the literature. These architectures have been trained using a vast multi-class dataset. In this study, the proposed model was compared to seven pre-trained deep models: VGG16, VGG19, SqueezeNet, GoogleNet, ResNet18, ShuffleNet

TABLE 6. Results (%) of pre-trained deep models.

Deep Models	Minimum	Maximum	Mean	Standard deviation
VGG16 [29]	74.78	97.39	92.80	3.6473
VGG19 [29]	76.52	96.52	93.27	3.6556
SqueezeNet [30]	54.78	97.39	86.28	9.9827
GoogleNet [31]	88.69	96.52	92.62	1.7499
ResNet18 [32]	66.95	96.52	92.01	3.7495
Xception [33]	70.43	86.95	80.25	3.1831
DarkNet19 [34]	78.26	97.39	93.40	3.0454
MobileNetv2 [35]	90.43	93.91	92	1.0689
Inceptionv3 [36]	87.83	94.78	91.39	2.0269
ShuffleNet [37]	90.43	93.91	92.69	1.1738
Proposed MI-DCNNE model	**93.04**	**99.13**	**96.19**	**1.5025**

TABLE 7. The comparison of the previous studies with the proposed model.

References	Method	Dataset	Number of classes and images	Accuracy
Geeth et al. [5]	Statistical feature extraction and back-propagation neural network	Periapical image	2 class and 105 images	97.1%
Prajapati et al. [6]	VGG16 (fine- tuning)	Periapical image	3 class and 251 images	87.5%
Singh et al. [7]	RT - DCT- PCA and random forest classifier	Panoramic image	2 class and 93 images	86%
Tuan et al. [11]	Semi-supervised fuzzy clustering and graph-based clustering algorithm	Intraoral and Panoramic image	5 class and 87 images	92.74%
Lee et al. [16]	GoogleNet and Inception v3	Periapical image	2 class and 3000 images	89%
Leo and Reddy [18]	Artificial neural network and deep neural network	Periapical image	4 class and 480 images	96%
Bui et al. [20]	Deep, statistical feature and support vector machine	Panoramic image	2 class and 533 images	91.7%
Majanga and Viriri [21]	Morphological methods and deep features	Periapical image	2 class and 11,147 images	97%
Vinayahalingam et al. [22]	MobileNet V2	Periapical image	2 class and 500 images	87%
Proposed Model	**Score-based multi-input CNN ensemble**	**Periapical image**	**2 class and 340 images**	**99.13%**

Xception, MobileNetv2, DarkNet19, and Inceptionv3. The same deep parameters were used for these architectures and the proposed model (Table 3). Besides, the same training/testing datasets were used, and experimental studies were repeated 100 times for these test data. Minimum, Maximum, Mean, and Standard deviation values for these results are given in Table 6.

In Table 6, the performance results of the pre-trained deep model are given. According to these results, the highest mean accuracy score among deep models achieved was 93.40% with the DarkNet19 model, the worst mean accuracy score was 80.25% with the Xception model. The highest maximum accuracy score was 97.39% using VGG16, SqueezeNet, and DarkNet19 models. As understood from Table 6, the proposed model to detect dental caries has achieved higher performance than pre-trained deep models.

V. DISCUSSION

Many studies have been carried out to detect dental caries. Traditional machine learning and deep learning methods have been used in these existing studies. Besides, Periapical X-Ray images have been often used to test these methods used in previous studies. The details and accuracy scores of these

studies are given in Table 7. In the previous studies given in Table 7, different datasets were used.

As seen from Table 7, the proposed model produced a better accuracy score than the compared existing studies for dental caries detection. While Geetha *et al.* [5] among previous studies achieved the best accuracy level as 97.1%, Singh and Sehgal [7] produced the worst accuracy score with 86%. Furthermore, existing studies have produced an average accuracy score of 90%. As a result, the current study is one of the most comprehensive experimental studies conducted for dental caries detection, considering the previous research. Additionally, the proposed model having performance close to 100% can assist dentists in real-time computer-aided systems.

VI. CONCLUSION

With the developing technology, dental imaging tools improve day by day. These tools are highly efficient in supporting the professional skills of dentists. For the last decade, artificial intelligence techniques have been being used with dental imaging tools for developing decision support systems for dentists. This study proposes a novel decision support system for the automatic diagnosis of dental caries based on periapical images. The proposed approach is based on a multi-input deep convolutional neural network ensemble model (MI-DCNNE). A periapical image dataset covering 340 both caries and non-caries images has been used for performance evaluation of the proposed method. According to the obtained results, it is seen that the proposed model is quite successful in the diagnosis of dental caries, where the reported accuracy score is 99.13%.

In future works, the statistical analysis of the results from the various deep CNN models is investigated to determine how reliable the reported differences in mean accuracy scores.

REFERENCES

[1] Y. W. Tang, *Molecular Medical Microbiology*. New York, NY, USA: Academic, 2014.

[2] W. W. Johnson, "The history of prosthetic dentistry," *J. Prosthetic Dentistry*, vol. 9, no. 5, pp. 841–846, Sep./Oct. 1959, doi: 10.1016/0022-3913(59)90049-6.

[3] O. E. Langland, R. P. Langlais, and J. W. Preece, *Principles of Dental Imaging*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2002.

[4] S. C. White and M. J. Pharoah, *Oral Radiology-E-Book: Principles and Interpretation*. Amsterdam, The Netherlands: Elsevier, 2014.

[5] V. Geetha, K. S. Aprameya, and D. M. Hinduja, "Dental caries diagnosis in digital radiographs using back-propagation neural network," *Health Inf. Sci. Syst.*, vol. 8, no. 1, pp. 1–14, Jan. 2020, doi: 10.1007/s13755-019-0096-y.

[6] S. A. Prajapati, R. Nagaraj, and S. Mitra, "Classification of dental diseases using CNN and transfer learning," in *Proc. 5th Int. Symp. Comput. Bus. Intell. (ISCBI)*, Aug. 2017, pp. 70–74, doi: 10.1109/ISCBI.2017.8053547.

[7] P. Singh and P. Sehgal, "Automated caries detection based on radon transformation and DCT," in *Proc. 8th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2017, pp. 1–6, doi: 10.1109/ICCCNT.2017.8204030.

[8] F. Casalegno, T. Newton, R. Daher, M. Abdelaziz, A. Lodi-Rizzini, F. Schürmann, I. Krejci, and H. Markram, "Caries detection with near-infrared transillumination using deep learning," *J. Dental Res.*, vol. 98, no. 11, pp. 1227–1233, Aug. 2019, doi: 10.1177/0022034519871884.

[9] A. G. Cantu, S. Gehrung, J. Krois, A. Chaurasia, J. G. Rossi, R. Gaudin, K. Elhennawy, and F. Schwendicke, "Detecting caries lesions of different radiographic extension on bitewings using deep learning," *J. Dentistry*, vol. 100, Sep. 2020, Art. no. 103425, doi: 10.1016/j.jdent.2020.103425.

[10] S. Datta, N. Chaki, and B. Modak, "Neutrosophic set-based caries lesion detection method to avoid perception error," *Social Netw. Comput. Sci.*, vol. 1, no. 1, pp. 1–15, Jan. 2020, doi: 10.1007/s42979-020-0066-0.

[11] L. H. Son, H. Fujita, N. Dey, A. S. Ashour, V. T. N. Ngoc, and D. T. Chu, "Dental diagnosis from X-ray images: An expert system based on fuzzy computing," *Biomed. Signal Process. Control*, vol. 39C, pp. 64–73, Jan. 2018, doi: 10.1016/j.bspc.2017.07.005.

[12] M. M. Lakshmi and P. Chitra, "Tooth decay prediction and classification from X-ray images using deep CNN," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Jul. 2020, pp. 1349–1355, doi: 10.1109/ICCSP48568.2020.9182141.

[13] J. Naam, J. Harlan, S. Madenda, and E. P. Wibowo, "Identification of the proximal caries of dental X-ray image with multiple morphology gradient method," *Int. J. Adv. Sci., Eng. Inf. Technol.*, vol. 6, no. 3, pp. 343–346, 2016.

[14] S. Oprea, C. Marinescu, I. Lita, M. Jurianu, D. A. Visan, and I. B. Cioc, "Image processing techniques used for dental X-ray image analysis," in *Proc. 31st Int. Spring Seminar Electron. Technol.*, May 2008, pp. 125–129, doi: 10.1109/ISSE.2008.5276424.

[15] R. Obuchowicz, K. Nurzynska, B. Obuchowicz, A. Urbanik, and A. Piórkowski, "Caries detection enhancement using texture feature maps of intraoral radiographs," *Oral Radiol.*, vol. 36, no. 3, pp. 275–287, Jul. 2020, doi: 10.1007/s11282-018-0354-8.

[16] J.-H. Lee, D.-H. Kim, S.-N. Jeong, and S.-H. Choi, "Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm," *J. Dent*, vol. 77, pp. 106–111, Oct. 2018, doi: 10.1016/j.jdent.2018.07.015.

[17] J. Naam, J. Harlan, S. Madenda, and E. P. Wibowo, "Image processing of panoramic dental X-ray for identifying proximal caries," *Indonesian J. Elect. Eng. Comput. Sci. (Telkomnika)*, vol. 5, no. 2, pp. 702–708, Jun. 2017. [Online]. Available: https://pdfs.semanticscholar.org/7cd9/d2e1ff9afbe0f84a40dc32ef77c91eeff0be.pdf, doi: 10.12928/TELKOMNIKA.v15i2.4622.

[18] L. Megalan Leo and T. Kalapalatha Reddy, "Learning compact and discriminative hybrid neural network for dental caries classification," *Microprocessors Microsyst.*, vol. 82, Apr. 2021, Art. no. 103836, doi: 10.1016/j.micpro.2021.103836.

[19] Y. P. Huang and S. Y. Lee, "Deep learning for caries detection using optical coherence tomography," *medRxiv*, early access, doi: 10.1101/2021.05.04.21256502.

[20] T. H. Bui, K. Hamamoto, and M. P. Paing, "Deep fusion feature extraction for caries detection on dental panoramic radiographs," *Appl. Sci.*, vol. 11, no. 5, p. 2005, Feb. 2021, doi: 10.3390/app11052005.

[21] V. Majanga and S. Viriri, "Automatic blob detection for dental caries," *Appl. Sci.*, vol. 11, no. 19, p. 9232, Oct. 2021, doi: 10.3390/app11199232.

[22] S. Vinayahalingam, S. Kempers, L. Limon, D. Deibel, T. Maal, M. Hanisch, S. Bergé, and T. Xi, "Classification of caries in third molars on panoramic radiographs using deep learning," *Sci. Rep.*, vol. 11, no. 1, pp. 1–7, Jun. 2021, doi: 10.1038/s41598-021-92121-2.

[23] H. Yang, E. Jo, H. J. Kim, I.-H. Cha, Y.-S. Jung, W. Nam, J.-Y. Kim, J.-K. Kim, Y. H. Kim, T. G. Oh, S.-S. Han, H. Kim, and D. Kim, "Deep learning for automated detection of cyst and tumors of the jaw in panoramic radiographs," *J. Clin. Med.*, vol. 9, no. 6, p. 1839, Jun. 2020, doi: 10.3390/jcm9061839.

[24] H. Watanabe, Y. Ariji, M. Fukuda, C. Kuwada, Y. Kise, M. Nozawa, Y. Sugita, and E. Ariji, "Deep learning object detection of maxillary cyst-like lesions on panoramic radiographs: Preliminary study," *Oral Radiol.*, vol. 37, no. 3, pp. 487–493, Sep. 2020, doi: 10.1007/s11282-020-00485-4.

[25] O. Kwon, T. H. Yong, S. R. Kang, J. E. Kim, K. H. Huh, M. S. Heo, S. S. Lee, S. C. Choi, and W. J. Yi, "Automatic diagnosis for cysts and tumors of both jaws on panoramic radiographs using a deep convolution neural network," *Dentomaxillofacial Radiol.*, vol. 49, no. 8, Jul. 2020, Art. no. 20200185, doi: 10.1259/dmfr.20200185.

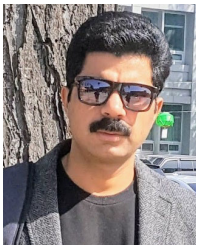
[26] M. Turkoglu, O. F. Alcin, M. Aslan, A. Al-Zebari, and A. Sengur, "Deep rhythm and long short term memory-based drowsiness detection," *Biomed. Signal Process. Control*, vol. 65, Mar. 2021, Art. no. 102364, doi: 10.1016/j.bspc.2020.102364.

[27] S. K. Khare and V. Bajaj, "Time–frequency representation and convolutional neural network-based emotion recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 2901–2909, Jul. 2021, doi: 10.1109/TNNLS.2020.3008938.

[28] S. Yavuzkilic, A. Sengur, Z. Akhtar, and K. Siddique, "Spotting deepfakes and face manipulations by fusing features from multi-stream CNNs models," *Symmetry*, vol. 13, no. 8, p. 1352, Jul. 2021, doi: 10.3390/sym13081352.
[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[30] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size," 2016, *arXiv:1602.07360*.
[31] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, Jun. 2015, pp. 1–9. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Szegedy_Going_Deeper_With_2015_CVPR_paper.html
[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html
[33] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1251–1258. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Chollet_Xception_Deep_Learning_CVPR_2017_paper.html
[34] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 7263–7271. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2017/html/Redmon_YOLO9000_Better_Faster_CVPR_2017_paper.html
[35] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/papers/Sandler_MobileNetV2_Inverted_Residuals_CVPR_2018_paper.html
[36] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html
[37] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6848–6856. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/papers/Zhang_ShuffleNet_An_Extremely_CVPR_2018_paper.html
[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105. [Online]. Available: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

ANDAC IMAK received the bachelor's and M.Sc. degrees in electrical and electronics engineering from Firat University, Turkey, in 2015 and 2019, respectively. He is currently pursuing the Ph.D. degree in image processing. He is a Research Assistant with the Department of Electrical and Electronics Engineering, Munzur University. His research interests include deep learning, biomedical signals, and image segmentation.

ADALET CELEBI received the B.Sc. degree from the Faculty of Dentistry, Istanbul University, Istanbul, in 2012, and the specialty degree from the Department of Maxillofacial Surgery, Faculty of Dentistry, Dicle University, in 2018. She is currently an Assistant Professor with the Faculty of Dentistry, Bingöl University.



KAMRAN SIDDIQUE (Senior Member, IEEE) received the Ph.D. degree in computer engineering from Dongguk University, South Korea. He is currently an Associate Professor with Xiamen University Malaysia. His research interests include big data, cybersecurity, machine learning, the IoT, and cloud computing.

MUAMMER TURKOGLU received the B.Sc. and M.Sc. degrees in electronics and computer education from Firat University, Turkey, in 2011 and 2013, respectively, and the Ph.D. degree in computer engineering from Inonu University, Turkey, in 2019. He was a Research Assistant with the Engineering Faculty, Bingol University, in April 2013. His research interests include signal processing, machine learning, and image processing.

ABDULKADIR SENGUR received the B.Sc. degree in electronics and computer education, the M.Sc. degree in electronics education, and the Ph.D. degree in electrical and electronics engineering from Firat University, Turkey, in 1999, 2003, and 2006, respectively. He was a Research Assistant with the Faculty of Technical Education, Firat University, in February 2001, where he is currently a Professor with the Faculty of Technical Education. His research interests include signal processing, image segmentation, pattern recognition, medical image processing, and computer vision.

IFTEKHAR SALAM (Member, IEEE) received the B.Eng. degree from Multimedia University, Malaysia, in 2008, the M.S. degree from Dongseo University, South Korea, in 2011, and the Ph.D. degree from the Queensland University of Technology, Australia, in 2018. From 2011 to 2018, he held sessional academic and research positions in Korean, Malaysian, and Australian universities, and research organizations. Since 2018, he has been an Assistant Professor with the School of Electrical and Computer Engineering, Xiamen University Malaysia. His research interests include cryptography, cryptanalysis, authenticated encryption, stream ciphers, and cryptographic protocols.