Research Article

International Journal of Advanced Technology and Engineering Exploration, Vol 9(91)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2021.874862

Customized convolutional neural network to detect dental caries from radiovisiography(RVG) images

Dipmala Salunke[1*], Deepak Mane[2], Ram Joshi[1] and Prasadu Peddi[2]
Department of Information Technology, JSPM's Rajarshi Shahu College of Engineering Tathawade, Pune, Maharashtra, India[1]
Department of Computer Engineering, JSPM's Rajarshi Shahu College of Engineering Tathawade, Pune, Maharashtra, India[2]

Received: 25-September-2021; Revised: 12-June-2022; Accepted: 13-June-2022

Abstract
Dental caries is amongst the most common tooth disease that people of all age groups face around the world. Detection of dental caries from dental x-ray images or radiovisiography (RVG) images in its initial stages is a challenging task. Deep learning is being used in almost all medical fields to predict or detect different diseases. This growing trend of deep learning algorithms like artificial neural network, back propagation neural network, convolution neural network (CNN) and recurrent convolution neural network (RCNN) was also applied into dentistry but the size of dataset is small. However, no standard large dataset is available for dental RVG images. Therefore, a dataset comprising 1336 samples has been prepared for the proposed work, and the dataset size has been augmented by 13582 using the image augmentation technique. In this paper, a customized convolution neural network (CCNN) was proposed that learns features automatically to classify the tooth is having dental caries or not from dental x-ray images. We've committed to old approaches like input data, data preparation, segmentation, feature extraction, model building, model training, model testing, model consequences, and model interpretation. Out of total 1336 images, 1104 images were used for training, 111 images were used for validation testing and 121 images were used for testing purpose. With a learning rate of 0.0001 and 100 epochs, a CCNN with six convolution layers achieves an average precision of 94.59 %, recall of 95.89 %, specificity of 91.66 %, f1-score of 94 %, and average testing accuracy of 94.2 %.

Keywords
RVG images, CNN, Image processing, Deep learning, Image augmentation, Dental caries, Machine learning.

1.Introduction
Dental caries, also known as tooth decay, is one of the extensive chronic diseases worldwide caused by the breakdown of tooth enamel [1]. Tooth decay causing bacteria in the mouth make acid that attacks tooth enamel, leading to a small hole in the tooth. Dental caries develops as a result of multiple interactions between acid-producing bacteria and fermentable carbohydrates, teeth, and saliva [2]. The mouth's correct and efficient care affects a person's general health and beauty. Oral diseases have to be treated on time or else they lead to gum problems, dental caries, tooth loss, bone loss, and other dental issues [3]. Furthermore, oral disorders affect over 3.5 billion people worldwide; dental caries is the most common disease, with a rising prevalence in many low- and middle-income countries [4].

Medical imaging is a growing field of research in the healthcare domain. For early detection, diagnosis, and treatment of diseases, medical imaging plays an important role. Images of this type include x-rays, computed tomography, magnetic resonance, ultra-sound, and so on [5]. Deep learning techniques are successful due to the availability of powerful computing machines, large amounts of data, and the development of new algorithms. The requirement for feature extraction is the fundamental disadvantage of traditional machine learning methods. When employing deep learning, this step is skipped. In the medical arena, computer assisted diagnosis (CAD) software has been used to gain second views, but because deep learning techniques are difficult to implement, they have been integrated into CAD, with promising outcomes for a variety of medical applications [6]. In the last 25 years, digital radiography has been available in dentistry and its

*Author for correspondence

use of dental practitioners is growing day by day. Dental radiographs show cavities, roots of teeth, bone loss, and hidden dental structures that cannot be seen during physical examination [7]. Dental experts typically use dental radiographs to diagnose dental deterioration, but some of them are having difficulty determining the correct diagnosis. Deep learning has played a major role in the dentistry field for several years, with largely acknowledged results [8]. It is also successful due to the availability of high-performance computer equipment, enormous amounts of data, and the development of new algorithms. Therefore, improved deep learning neural networks can be customised to design the best CAD system for dental disease detection.

Deep learning models have been used to provide interesting results in the last few years for dental disease detection and diagnosis [9]. According to a literature survey, many studies are done to detect and diagnose dental caries, but the size of the dataset is small. No public or standard dataset is available for dental x-ray images. However, no significant efforts were made to create a large dental x-ray image dataset. A dental x-ray image dataset with 1336 samples were created. It is a challenging task to identify patterns from the dental x-ray images for dental caries detection as the caries size may be small or large. Convolution neural network (CNN) is a deep learning architecture that has been customised to detect dental caries from 2-dimension dental radiographs and provides better results. The challenging problem is the image segmentation and extraction of useful features defining the area of a tooth having caries from dental x-ray images. This paper compares the performance of six different deep learning algorithms, primarily visual geometry group (VGG16), ResNet50, Inception3, EfficientNetB0, EfficientNetB7, and AlexNet, with the proposed customized convolution neural network (CCNN) algorithm in the classification of dental disease in radiovisiography (RVG) images into caries or no caries.

The following is the outline of the paper: Section 2 provides a brief overview of the literature study and background information. Section 3 goes over the dataset and proposed architecture in detail. The section 4, summarizes the results. Section 5 summarises and interprets major findings. The paper is concluded in section 6.

2.Literature review

Deep learning has become successful due to the availability of graphical processing units, faster central processing units, big data sets, and improvement in artificial neural networks with more than two layers [10]. CNNs have been used successfully in medicine, notably in the field of cancer, for the automated detection of skin cancer, breast cancer and diabetic retinopathy in eye exams [11]. Because dental caries is the most frequent disease in the world and is becoming more widespread in many low and middle-income nations, it has sparked a lot of interest and research. Development of deep learning algorithms for dental caries detection is a challenging task due to variation in size of caries. Computer programmes can assist dentists in making decisions about prevention, diagnosis, and treatment planning, among other things. Casalegno et al. [12] employed convolution neural network (CNN) based methods on near-infrared transillumination images to detect early-stage carious legions. For proximal and occlusal lesions, the model has obtained an area under the receiver operating curve of 85.6% and 83.6%, respectively, in 2019. Ekert et al. [13] employed a deep learning-based CNN model to detect apical lesions on panoramic radiographs. The 2001 synthesised tooth images demonstrated satisfactory discriminatory capacity to detect apical legions in 2019 where hyper parameters were tuned using grid search. The area under the curve, sensitivity and specificity of the model were 85%, 65% and 87%, respectively. Employed three models: CNN, transfer learning, and transfer learning with fine tuning, trained on 180 images with 45 validation images and 26 test images, having an accuracy of 73.07%, 88.46%, and 88.46%, respectively. Lee et al. (2018) [14] evaluated the efficacy of the deep CNN algorithm to detect and diagnose dental caries on 3000 periapical radiographs with an average of 85% accuracy. Koch et al. 2019[15] employed a fully CNN based on the U-net architecture for the segmentation of 1500 panoramic radiographs with a dice score of 0.934 for a single network. Bouchahma et al. 2019[16] employed a deep CNN to predict dental decay treatment on 200 x-ray images and, based on classification results, different types of treatments were suggested: fluoride with 99% accuracy, filling with 98% accuracy, and root canal with 88% accuracy. Tuzoff et al. [17] designed the CNN-based architecture in 2019 for automatic detection and numbering of teeth on 1352 panoramic mages with a sensitivity of 0.9941 and a precision of 0.9945.

Dasanayaka et al. 2019[18] employed, U-Net deep CNN to assist doctors in locating dental foramen on 1000 dental panoramic tomography images with a mean dice similarity coefficient value of 95.4 for patients between 7 and 20 years of age. On 760 panoramic radiographs, a deep learning system performed diagnostically well in classifying the root morphology of mandibular first molars, with an accuracy of 86.9% [19]. Of a total of 88 intra-oral camera images, a deep learning model mask called recurrent convolution neural network (RCNN) was used to detect and categorise dental caries on a 7-class caries detection and assessment system scale with an average accuracy of 75% by Moutselos et al. 2019 [20]. Geetha et al. 2020[21] developed a back propagation neural network algorithm for the diagnosis of dental caries on a total of 105 digital radiographs with 97.1% accuracy. Lee et al. 2021[22] used a deep CNN model based on U-nets to detect caries in bitewing images. 304 bitewing images were used to train the model, and 50 images were used to test its performance. The model had a precision of 63.29 %, an F1-score of 64.14 %, and a recall value of 65.02 % on the test dataset. 23. Laishram and Thongam 2020 [23] used a faster RCNN algorithm to classify different tooth types like incisors, molars, premolars, and canine teeth for 146 images with an impacted teeth detection accuracy of 90%. With an accuracy of 92 %, F1 score of 90 %, recall of 87 %, and precision of 94 %, Kang et al. [24] proposed a random forest algorithm to identify dental caries. Deep CNN architecture was proposed by Chen et al. [25] to detect dental caries with severity levels from periapical images with precision and recall values between 0.5 and 0.6. To diagnose dental caries from dental x-rays, Imak et al. [26] proposed a multi input deep CNN ensemble model. The model has the potential to aid in the classification of dental caries. AL-ghamdi et al. [27] with 96 %, proposed a neural search architecture to classify x-ray pictures into cavity, filling, and the implant. Thanh et al. [28] used four deep learning models to detect cavities from intraoral photographs, including Faster region-based convolutional neural networks (Faster R-CNNs), You Only Look Once version 3 (YOLOv3), RetinaNet, and Single-shot multi-box detector (SSD), with sensitivity values of 87.4 % and 71.4 % for Faster RCNN and YOLOv3 respectively.

According to the results of the survey, there are certain challenges in the classification of dental diseases. These challenges are as under:

1. A complex tooth structures.
2. The size of caries.
3. Teeth are arranged in a random manner.
4. The dental image has insufficient contrast, and the uneven exposure causes the intensities of teeth, gums, and bones to resemble visual cues.

Because human inspection alone misses a large percentage of caries, visual inspection has a low sensitivity rate. Detecting cavities in their early stages is also impossible to notice with the naked eye and even more difficult to detect with the human vision system in x-ray images. The bulk of studies have focused on caries location, aetiology, and rate of development in order to develop an algorithm for the identification of caries and quantification of tooth damage. An observable trend in the majority of previous works is that they have focused on using smaller datasets.

3.Methodology

CNN is customized in such a way that it recognises dental caries images of 2-dimension shapes and gives better results than other deep learning models. CNN extracts feature very efficiently. Unlike vector algorithms for image processing, the spatial interaction of pixels is conserved in CNN [29]. Information related to adjacent pixels is not lost during downsizing the dimension of the original image.

3.1Dataset preparation and pre-processing

The standard dental x-ray dataset is not available online, so the authors have collected a dataset of around 1336 RVG x-ray images from the Tirupati Dental Clinic, Pune. Dental imaging sensors are used to capture the x-rays of a patient’s mouth, whereas Kodak’s RVG software [30], called CareStream 6100, was used to store RVG images. All RVG images are of the same size, with 748×512 dimensions. A licenced copy of RVG software is used to convert .rvg files into.jpg format, which is available in almost all dental clinics. Each tooth was cropped from the x-ray image separately. These cropped tooth samples were resized in a 224×224 size image. Each image is resized and saved into a folder labelled "caries or no caries." After building the dataset, we had a thorough review and cross-verification by the expert dentist to determine whether the images were classified properly or not. We used input from two dental experts to divide images into two classes and created a folder for images with caries and images without caries, which allowed us transmit the annotated dataset to the CCNN model.

The dataset size is increased by applying image augmentation techniques [31]. Image augmentation helped to increase the size of the dataset from 1336 images to 13582 images. Out of a total of 1336 images in the dataset, 647 images are of teeth having caries and 689 images are of normal teeth. Out of a total of 1336 images, 1104 images were used for training, 111 images were used for validation and 121 images were used for testing purposes. *Figure 1* shows the sample dental dataset showing teeth with caries and those without caries. The result of image augmentation on a sample dental image is shown in *Figure 2*.

(a) Caries (b) No caries

Figure 1 Sample RVG images

Different types of image augmentations result of dental x-rays are shown in *Figure 2*:

(a) Original image
(b) Rotate image by -50 degree to 30 degree
(c) Rotate image by 30 degree to 70 degree
(d) Adding noise to the data
(e) Cropping of image
(f) Shearing the image
(g) Flipping an image horizontally
(h) Flipping an image vertically
(i) Changing the brightness of the image
(j)Scaling the image below to 150% to 80% of the im age height/width
(k) Bounding box around the original image
(l)Displaying the bounding box on top of the original image.

Position augmentation techniques were applied to increase the size of the dataset which led to an increase in the performance of the model.

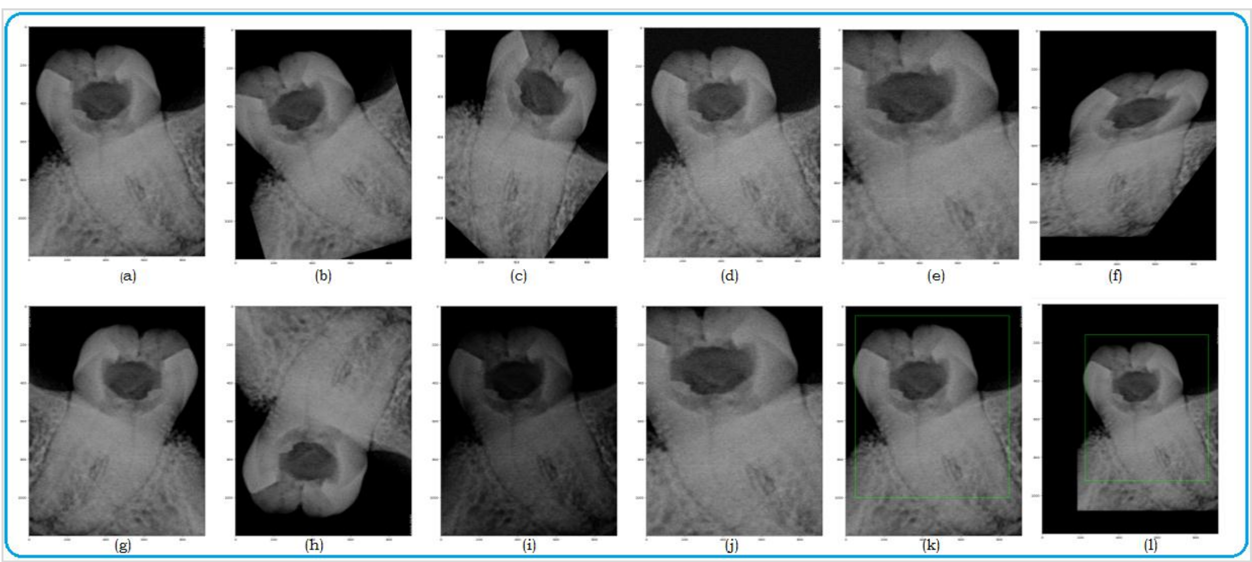

Figure 2 Image augmentation

3.2Proposed architecture of CCNN for dental caries detection

To get better accuracy, in the proposed CCNN model, six convolution layers are defined, and all images are resized to 224×224. CCNN is a CNN with 6 convolutional layers, 6 max-pooling layers, and 2 fully connected layers, giving a better model performance of the test data. The convolution layer is the first layer that extracts features from the input image to generate the feature map. Convolution is a mathematical procedure that takes place between an input image and a filter of the size of 3×3. The dot product between the filter and a part of an input image is calculated by sliding the filter across it [32]. The feature map provides details about the image, such as corners and edges. This feature map is then supplied to further layers, which learn a variety of other features from the input image. The pooling layer is applied to reduce the computational costs by reducing the size of the convolved feature map. Max pooling is used here to select the largest element from the feature map by decreasing connections between layers. A stride of (2, 2) is applied to the max-pooling layer. Weights and biases, as well as neurons, make up the fully connected layers, which are used to connect neurons from two separate layers. The input image from the previous layers is flattened and fed to the fully connected layer. When all features are integrated into fully connected layers, the model becomes overfitted. To avoid this, a dropout layer is employed, in which a few neurons are removed from the neural network during the training process, resulting in a reduced model. To add nonlinearity to the model, the first five convolutional layers utilise a rectified linear unit activation function, while the last fully connected layer uses a sigmoid activation function [33].

As shown in *Figure 3*, the fully connected layer receives the flattened image as an input, and the softmax layer determines the class of the input image.

Figure 4 depicts a block diagram for a dental caries detection system in which a dataset is acquired from dentists, pre-processed, and labelled with a medical practitioner's opinion. After then, dataset augmentation is used to increase the model's accuracy. To classify images into caries and no caries classes, the dataset was trained and tested using various models such as VGG16[34], ResNet50[35], Inception3[36], EfficientB0[37], EfficientNetB7[37], AlexNet[38], and the proposed CCNN model.

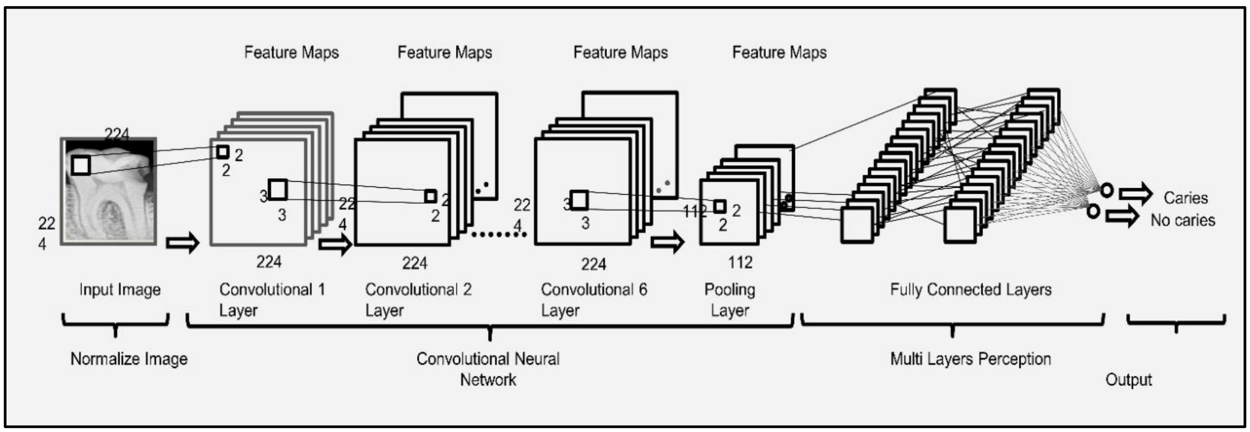

Figure 3 CCNN architecture for caries classification

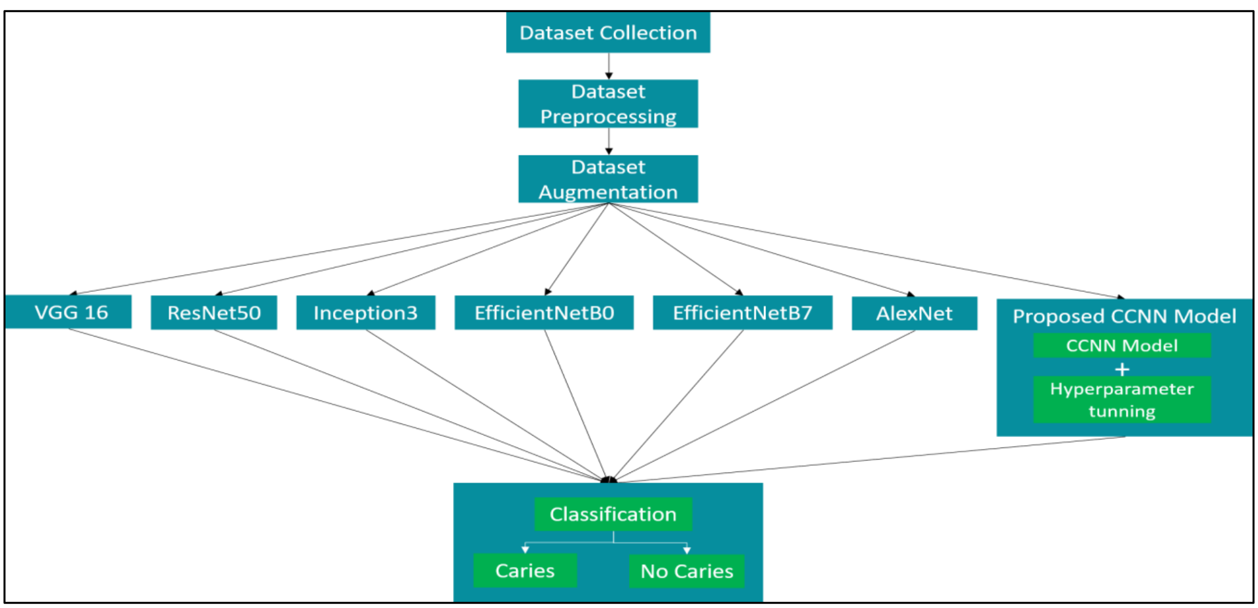

Figure 4 Block diagram

3.3Visualization of feature map

A feature map, also known as an activation map, shows the result of applying filters to an input image. An intermediate visualisation of the feature map is done on a sample input image to understand the kind of features CNN extracts after each layer. On a 224×224 input image, *Figure 5* shows the effect of applying the first convolution layer, which has 16 filters. *Figure 6* shows the effect of applying a second convolution layer having 32 filters on the input image size of 224×224. From the Figure, we can infer that the initial layer depicts low-level features like edges, textures, and colour of the input image and later layers depict high-level features like shape and object. Prominent features of the image are extracted to identify that the input image is of a tooth having caries. The feature map helped to understand the workings of the CNN model, and the model's inaccurate predictions were fine-tuned to improve its accuracy.

We have trained our system from scratch as pre-trained models are not available for dental image processing applications. *Table 1* shows the proposed CCNN model parameters, which is trained on input images until loss reaches 10-2 or more than that.

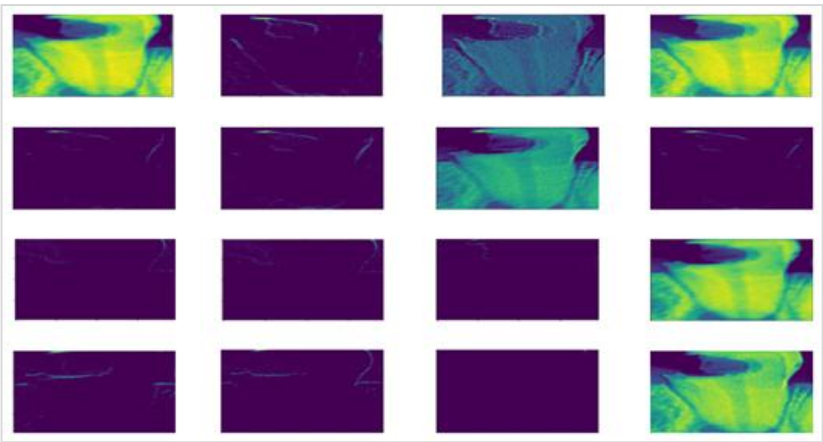

Figure 5 Feature map (16 Filters)

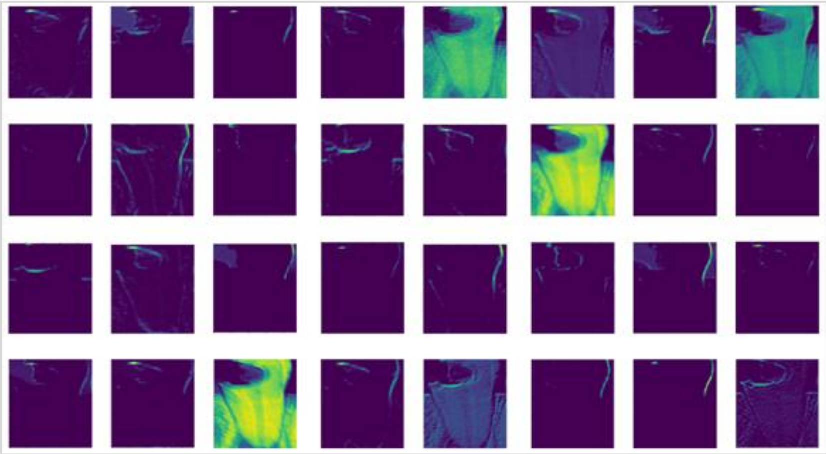

Figure 6 Feature map (32 Filters)

Table 1 Proposed CCNN model parameters

Layer (type)	Output shape param #	Output shap param #
conv2d (Conv2D)	(None, 222, 222, 16)	448
Max_pooling_2d (Maxpooling2D)	(None, 111, 111, 16)	0
Conv2d_1 (conv2D)	(None, 109, 109, 32)	4640
Max_pooling_2d_1(Maxpooling2D)	(None, 54, 54, 32)	0
Conv2d_2 (conv2D)	(None, 52, 52, 64)	18496

Layer (type)	Output shape param #	Output shap param #
Max_pooling_2d_2(Maxpooling2D)	(None, 26, 26, 64)	0
Conv2d_3 (conv2D)	(None, 24, 24, 64)	36928
Max_pooling_2d_3(Maxpooling2D)	(None, 12, 12, 64)	0
Conv2d_4 (conv2D)	(None, 10, 10, 64)	36928
Max_pooling_2d_4(Maxpooling2D)	(None, 5, 5, 64)	0
Conv2d_5 (conv2D)	(None, 3, 3, 64)	36928
Max_pooling_2d_5(Maxpooling2D)	(None, 1, 1, 64)	0
Flatten (Flatten)	(None,64)	0
dense (Dense)	(None,512)	33280
dense_1 (dense)	(None,1)	513
Total params:168,161 Trainable params: 168,161 Non-trainable params: 0		

Cross entropy is employed to calculate the model's loss value, as it is a binary classification problem and given by the Equation 1:

$$H(p,q) = -\frac{1}{N}\sum_{i=1}^{N} yi.log\big(p(yi)\big) + (1 - yi).log(1 - p(yi)) \quad (1)$$

Where, p=actual output
q = predicted output
N= all data points

The RMSprop optimizer is chosen because it averages each weight's squared gradients and divides the gradient by the square root of the mean square defined by Equation 2 and Equation 3:

$$E[g^2]_t = \beta E[g^2]_{t-1} + (1-\beta)(\frac{\delta c}{\delta w})^2 \quad (2)$$

$$w_t = w_{t-1} - \frac{n}{\sqrt{E[g^2]_t}}(\frac{\delta c}{\delta w}) \quad (3)$$

Where, E[g]= the squared gradients moving average
$(\frac{\delta c}{\delta w})$= the cost function's gradient with respect to the weight
n =learning rate
βE = Considering the default value of =0.9 for the moving average parameter

3.4 Steps

1. Start.
2.Open the user interface and upload the input image. The supported image formats for our system are jpeg, jpg, and png.
3. The uploaded image is pre-processed. The size of the image is normalised to a dimension of 224×224. This is an important step because it ensures that the data distribution in all images is the same. This leads to a faster convergence of the network towards a result.
4.After pre-processing, the image is fed to the model. To extract features from the image, we have Con2D layers.
5. The kernel size for the filter of our system is 3×3×3. The kernel slides over the image and thus multiplies values in the filter by the original pixel values. This leads to a transformation of the original image.
6. Multiplication is summed up to get one value for a particular sub matrix. This eventually generates a feature map of 16×16 for the first layer.
7. Then the image is processed further by sequential filters. Each filter is activated when an area of interest is found. At each layer of filter, a low-level feature map of the original image is made, while at the same time retaining all the information in the image.
8. After the generation of the final feature map, it is fed to the max pooling layers. For the given system, the filter of 2×2 with stride 2×2 slides over each patch of feature. The maximum value from each patch is saved. Thus, the feature map is down-sampled.
9. The output vector of max pooling is fed to fully connected layers. The fully connected layer predicts the given feature map and analyses it in which of the 2 classes our x-ray image strongly resembles.
10. The system is a binary classifier, so the output is either 0 or 1, and this is passed to the frontend for the user.
11. Stop.

4. Results

In total, we scanned 6,000 RVG images from patients who visited the Tirupati Dental and Implant Centre in Pune, Maharashtra, India. Out of those, 1336 images are finalised, and from those images, the individual teeth have been cropped to fit our model's needs. Out of a total of 1336 images, 1104 images were used for training, 111 images for validation testing, and 121 images for testing the performance of the model. The system is implemented using tensor flow and keras python libraries and trained with the help of Google

Colab [39], which provides free access to the GPU: 1xTesla K80. The dataset is trained using different deep learning models like VGG16, ResNet50, Inception3, EfficientNetB0, EfficientNetB7, AlexNet, and the proposed CCNN model.

The comparative analysis of deep learning models with proposed CCNN Model is shown in the *Table 2* and *Figure 7*. Out of 121 test images, our model correctly classified 114 images. Hence, we were able to achieve an accuracy of 94.2% which is acceptable.

Table 2 Comparative analysis of deep learning models with proposed CCNN model

Sr. No.	**Deep learning model**	**Number of epochs**	**Training dataset size**	**Testing dataset size**	**Testing accuracy**
1	VGG16	37	1104	121	75.3
2	ResNet50	100	1104	121	82.03
3	Inception3	79	1104	121	85.5
4	EfficientNetB0	97	1104	121	80.85
5	EfficientNetB7	81	1104	121	80.3
6	AlexNet	67	1104	121	53.98
7	Proposed CCNN model	100	1104	121	94.2%

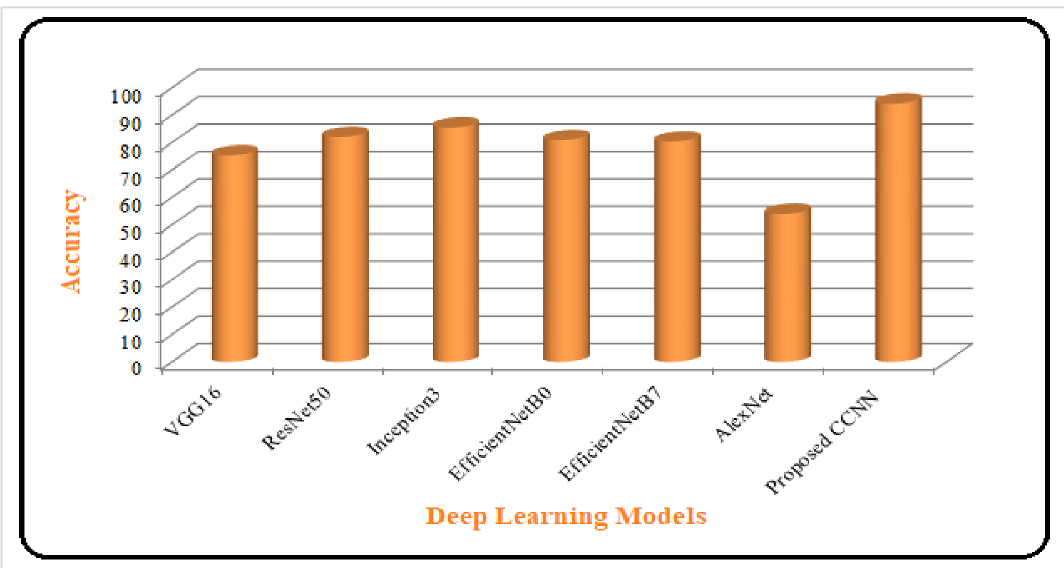

Figure 7 Test accuracy comparison of deep learning models with proposed CCNN model

As far as both precision and dataset size considered the proposed approach is ideal. *Figure 8* shows the confusion matrix [40] for the system where two classes are displayed as no-caries and caries. The model predicts 44 no-caries images and 04 caries images out of a total of 48 actual no-caries images.

The algorithm correctly predicted 70 images out of a total of 73 actual caries images, and three images were predicted as no-caries. The classification report shows the class-wise model performance evaluation parameters for class 0 (which represents caries) and class 1 (no caries). The efficiency of the proposed model is evaluated using different performance parameters like accuracy, precision, recall, f1-score, error rate and support values as shown in *Table 3* which can be calculated as below: (TP: True Positive, TN: True Negative; FP: False positive; FN: False Negative).

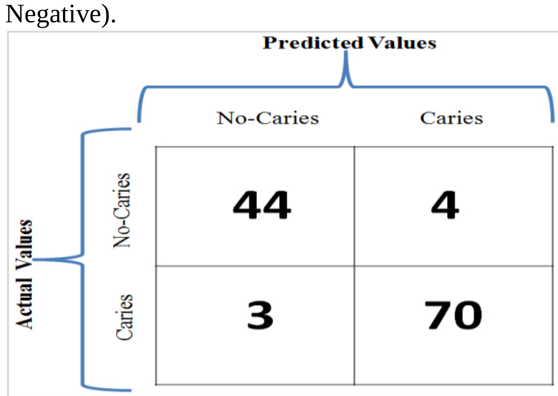

Figure 8 Confusion matrix

The performance measures used here are represented by Equation 4, 5, 6, 7 and 8.

Table 3 Classification report

Classification report:				
Class	**Precision**	**Recall**	**F1-score**	**Support**
1	0.94	0.92	0.93	48
0	0.95	0.96	0.95	73
accuracy	--	--	0.94	121
Macro average	0.94	0.94	0.94	121
Weighted average	0.94	0.94	0.94	121

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN}\times 100 = \frac{70+44}{70+44+4+3}\times 100 = 94.2\% \quad (4)$$

$$\text{Precision}=\frac{TP}{TP+FP}\times 100 = \frac{70}{70+4}\times 100 = 94.6\% \quad (5)$$

$$\text{Recall}= \frac{TP}{TP+FN}\times 100 = \frac{70}{70+3}\times 100 = 95.9\% \quad (6)$$

$$\text{F1-score}=\frac{2\times \text{Precision}\times \text{Recall}}{\text{Precsion}+\text{Recall}} = \frac{2\times 94.6\times 95.9}{94.6+95.9} = 95.24\% \quad (7)$$

$$\text{Error rate}=\frac{FP+FN}{TP+TN+FP+FN}\times 100 = \frac{4+3}{70+44+4+3}\times 100 = 5.785\% \quad (8)$$

Figure 9 shows the training and testing accuracy of the CCNN model for 100 epochs. The model has shown reliable dental caries detection performance as shown in the confusion matrix and classification report. In future, visual-tactile assessment can be used as the gold standard for confirming the existence of dental caries.

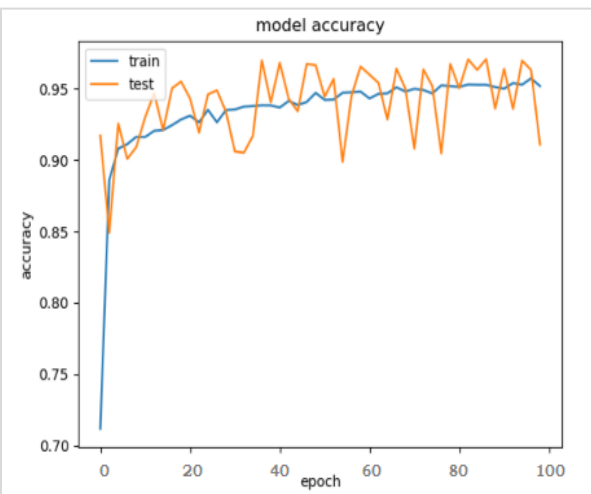

Figure 9 Train and test accuracy of CCNN model

5. Discussion

As demonstrated in *Figure 9*, the model has exhibited consistent dental caries detection ability. The proposed CCNN architecture outperforms other architectures with an accuracy of 94.2%, precision of 94.6%, the recall of 95.9%, and an F1-score of 95.24%. Since CNN prevents over-fitting [41] and incorporates local connections between the pixels, the suggested model scales well with the increasing data set. Furthermore, the larger data set enhances the filter's weights, which boosts performance even further. The number of parameters is reduced in the convolution layers of VGG16 and the conv kernels are of size 3×3 with a stride of 2 to improve training time. The accuracy of the VGG16 model is 75.35%. The ResNet50 model was tested with 82.03% accuracy where only two pooling layers were used, one at the start and the other at the end of the network. The Inception3 model's accuracy is 85.5%, which is focused on applying different size kernels for effective recognition of variable-sized features in x-ray. EfficientnetB0 and EfficientnetB7 networks are too wide and deep, having more parameters that are employed on the dataset, giving 80.85% accuracy for EfficientnetB0 and 80.3% accuracy for the EfficientnetB7 architecture. Five convolutional layers and three fully connected layers make up the AlexNet network. Data augmentation is used in Alexnet to reduce the overfitting. The AlexNet model was trained on a dataset, but the accuracy was 53.98%, which was the lowest accuracy of dental caries classification. After the first, second, and fifth convolution levels, the network employs an overlapped max-pooling layer to help to reduce top-1 error. The model assesses whether or not a tooth x-ray includes caries, however, it does not determine the depth or severity of caries. Only radiographs were used to assess the presence of dental caries, with no visual or clinical assessment data. Overall, it is observed that the proposed CCNN is the best classifier to classify dental caries based on RVG images, followed by Inception 3, ResNet50, EfficientNetB0, EfficientNetB7, VGG16, and AlexNet. A complete list of abbreviations is shown in *Appendix I.*

6. Conclusion and future work

This paper focuses on dental disease classification using a CCNN. The average testing accuracy of the model is 94.2%. The major merit of using this proposed model is that it increases the accuracy of detection of dental caries even in its initial stages, which might be difficult to detect by dentists, and makes the entire process faster. The automatic prediction of caries from x-ray images will support the patient's trust in the analysis so that he can acknowledge the prescribed treatment by the dentist. Our system doesn't work better for images with low resolution and some surprising patterns in images. This system not only helps dentists to get a second opinion about the condition of the patient's teeth, but also helps them identify some defects in the teeth that they might have missed.

With further expansion of our dataset, the system should be capable to predict most dental diseases and will be able to create a treatment plan for the patients. This model can be further trained to predict and detect a large variety of dental diseases, and with sufficient data, can also predict the severity of the diseases.

Acknowledgment

The authors would like to express their gratitude to Dr. Vikas Kashid (Tirupati Dental and Implant Centre, Pune, Maharashtra, India), who supported the dental RVG x-ray images and dental disease information for this study.

Conflicts of interest

The authors have no conflicts of interest to declare.

Author's contributions statement

Dipmala Salunke: Conceptualization, dataset creation, implementation, writing-original draft, reviewing and editing. **Deepak Mane:** Dataset augmentation, conceptualization, draft manuscript preparation and identification of challenges. **Ram Joshi:** Analysis and interpretation of results, reviewing and editing. **Prasadu Peddi:** Design and data analysis, manuscript preparation.

References

[1] Mutasa S, Sun S, Ha R. Understanding artificial intelligence based radiology studies: what is overfitting? Clinical Imaging. 2020; 65:96-9.

[2] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. Journal of Big Data. 2019; 6(1):1-48.

[3] You W, Hao A, Li S, Wang Y, Xia B. Deep learning-based dental plaque detection on primary teeth: a comparison with clinical assessments. BMC Oral Health. 2020; 20(1):1-7.

[4] Peres MA, Macpherson LM, Weyant RJ, Daly B, Venturelli R, Mathur MR, et al. Oral diseases: a global public health challenge. The Lancet. 2019; 394(10194):249-60.

[5] Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annual Review of Biomedical Engineering. 2017; 19:221-48.

[6] Rodríguez JH, Fraile FJ, Conde MJ, Llorente PL. Computer aided detection and diagnosis in medical imaging: a review of clinical and educational applications. In proceedings of the fourth international conference on technological ecosystems for enhancing multiculturality 2016 (pp. 517-24).

[7] Ramenzoni LL, Lehner MP, Kaufmann ME, Wiedemeier D, Attin T, Schmidlin PR. Oral diagnostic methods for the detection of periodontal disease. Diagnostics. 2021; 11(3):1-13.

[8] Wang CW, Huang CT, Lee JH, Li CH, Chang SW, Siao MJ, et al. A benchmark for comparison of dental radiography analysis algorithms. Medical Image Analysis. 2016; 31:63-76.

[9] Hwang JJ, Jung YH, Cho BH, Heo MS. An overview of deep learning in the field of dentistry. Imaging Science in Dentistry. 2019; 49(1):1-7.

[10] Park JH, Hwang HW, Moon JH, Yu Y, Kim H, Her SB, et al. Automated identification of cephalometric landmarks: part 1—comparisons between the latest deep-learning methods YOLOV3 and SSD. The Angle Orthodontist. 2019; 89(6):903-9.

[11] Prajapati SA, Nagaraj R, Mitra S. Classification of dental diseases using CNN and transfer learning. In international symposium on computational and business intelligence 2017 (pp. 70-4). IEEE.

[12] Casalegno F, Newton T, Daher R, Abdelaziz M, Lodi-rizzini A, Schürmann F et al. Caries detection with near-infrared transillumination using deep learning. Journal of Dental Research. 2019; 98(11):1227-33.

[13] Ekert T, Krois J, Meinhold L, Elhennawy K, Emara R, Golla T, et al. Deep learning for the radiographic detection of apical lesions. Journal of Endodontics. 2019; 45(7):917-22.

[14] Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. Journal of Dentistry. 2018; 77:106-11.

[15] Koch TL, Perslev M, Igel C, Brandt SS. Accurate segmentation of dental panoramic radiographs with U-Nets. In 16th international symposium on biomedical imaging 2019 (pp. 15-9). IEEE.

[16] Bouchahma M, Hammouda SB, Kouki S, Alshemaili M, Samara K. An automatic dental decay treatment prediction using a deep convolutional neural network on X-ray images. In IEEE/ACS 16th international conference on computer systems and applications 2019 (pp. 1-4). IEEE.

[17] Tuzoff DV, Tuzova LN, Bornstein MM, Krasnov AS, Kharchenko MA, Nikolenko SI, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks. Dentomaxillofacial Radiology. 2019; 48(4).

[18] Dasanayaka C, Dharmasena B, Bandara WR, Dissanayake MB, Jayasinghe R. Segmentation of mental foramen in dental panoramic tomography using deep learning. In conference on industrial and information systems 2019 (pp. 81-4). IEEE.

[19] Hiraiwa T, Ariji Y, Fukuda M, Kise Y, Nakata K, Katsumata A, et al. A deep-learning artificial intelligence system for assessment of root morphology of the mandibular first molar on panoramic radiography. Dentomaxillofacial Radiology. 2019; 48(3).

[20] Moutselos K, Berdouses E, Oulis C, Maglogiannis I. Recognizing occlusal caries in dental intraoral images using deep learning. In 41st annual international

conference of the IEEE engineering in medicine and biology society 2019 (pp. 1617-20). IEEE.
[21] Geetha V, Aprameya KS, Hinduja DM. Dental caries diagnosis in digital radiographs using back-propagation neural network. Health Information Science and Systems. 2020; 8(1):1-14.
[22] Lee S, Oh SI, Jo J, Kang S, Shin Y, Park JW. Deep learning for early dental caries detection in bitewing radiographs. Scientific Reports. 2021; 11(1):1-8.
[23] Laishram A, Thongam K. Detection and classification of dental pathologies using faster-RCNN in orthopantomogram radiography image. In 7th international conference on signal processing and integrated networks 2020 (pp. 423-8). IEEE.
[24] Kang IA, Ngnamsie NS, Lee KO, Kim JD. DCP: prediction of dental caries using machine learning in personalized medicine. Applied Sciences. 2022; 12(6):1-15.
[25] Chen H, Li H, Zhao Y, Zhao J, Wang Y. Dental disease detection on periapical radiographs based on deep convolutional neural networks. International Journal of Computer Assisted Radiology and Surgery. 2021; 16(4):649-61.
[26] Imak A, Celebi A, Siddique K, Turkoglu M, Sengur A, Salam I. Dental caries detection using score-based multi-input deep convolutional neural network. IEEE Access. 2022; 10:18320-9.
[27] AL-ghamdi AS, Ragab M, Alghamdi SA, Asseri AH, Mansour RF, Koundal D. Detection of dental diseases through X-ray images using neural search architecture network. Computational Intelligence and Neuroscience. 2022.
[28] Thanh MT, Van TN, Ngoc VT, Tra NT, Giap CN, Nguyen DM. Deep learning application in dental caries detection using intraoral photos taken by smartphones. Applied Sciences. 2022; 12(11):1-10.
[29] Yamashita R, Nishio M, Do RK, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. 2018; 9(4):611-29.
[30] Ilić D, Stojanović L. Application of radiovisiography (digital radiology) in dental clinical practice. Vojnosanitetski Pregled. 2012; 69(1):81-4.
[31] Nanni L, Paci M, Brahnam S, Lumini A. Comparison of different image data augmentation approaches. Journal of Imaging. 2021; 7(12):1-13.
[32] Albawi S, Mohammed TA, Al-zawi S. Understanding of a convolutional neural network. In international conference on engineering and technology 2017 (pp. 1-6). IEEE.
[33] Nwankpa C, Ijomah W, Gachagan A, Marshall S. Activation functions: comparison of trends in practice and research for deep learning. arXiv preprint arXiv:1811.03378. 2018.
[34] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014.
[35] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In proceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 770-8).
[36] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In proceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 2818-26).
[37] Tan M, Le Q. Efficientnet: rethinking model scaling for convolutional neural networks. In international conference on machine learning 2019 (pp. 6105-14). PMLR.
[38] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. 2012.
[39] Carneiro T, Da NRV, Nepomuceno T, Bian GB, De AVH, Reboucas FPP. Performance analysis of google colaboratory as a tool for accelerating deep learning applications. IEEE Access. 2018; 6:61677-85.
[40] Sammut C, Webb GI. Encyclopedia of machine learning and data mining. Springer Publishing Company, Incorporated; 2017.
[41] Salman S, Liu X. Overfitting mechanism and avoidance in deep neural networks. arXiv preprint arXiv:1901.06566. 2019.

Dipmala Salunke is working as Assistant Professor in JSPM's Rajarshi Shahu College of Engineering, Tathwawade, Pune, India. She is a Research Scholar at Department of Computer Engineering, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan, India. She has obtained her BE in Information Technology from TPCT, COE Osmanabad and ME in Computer Engineering from PICT, Pune, India. She has written books on Machine Learning, Artificial Intelligence and Software Architecture for Computer /IT students. Her research interests are Deep Learning, Image Processing, Artificial Intelligence.
Email: dipmala.salunke@gmail.com

Dr. Deepak Mane awarded a Ph.D. degree in Computer Science and Engineering from SGGS Institute of Engineering and Technology, Nanded, in 2019. He has written books on "Machine Learning", "Fundamental of Programming Languages" and "Principles of Programming languages" for Computer /IT students. He is a reviewer for many international conferences and journals. He has delivered many sessions on "Machine Learning, Deep Learning, High Performance Computing and Neural Networks" to students as well as faculty members at different Engineering institutes. His main research work focuses on Machine Learning, Pattern Classification, Supervised Clustering and Neural Network. He has 16.5 years of teaching experience. He is currently working as Associate

Professor in the Department of Computer Engineering, JSPM's Rajarshi Shahu College of Engineering, Pune. In 2020, he received I2OR national "Eminent Young Researcher Award" and in 2021 he received "BEST Researcher Award" in International Scientist Award 2021 on Engineering , Science & Medicine in Goa, India.
Email: dtmane@gmail.com

Ram Joshi is working as an academician with administrative and teaching experience of 27 years and 02 years with industrial experience as developer in the software industry. Currently serving as Professor and Dean Academics at JSPM's Rajarshi Shahu College of Engineering Tathawade, Pune, affiliated to SPPU Pune. He obtained his bachelor degree in Computer Engineering from SGGSCE&T Nanded and Masters in CSE from BVDUCOE, Pune. He guides the scholars in the field of Information Security, Cloud Security, Block Chain and Machine Learning.
Email: ramjoshi.com@gmail.com

Dr. Prasadu Peddi has over 13+ years of academic and industry experience. He is a Professor in the Department of Computer Engineering at Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan, India. He guides the scholars in the field of Machine Learning and Cloud Computing.
Email: Prasadupeddi27@gmail.com

Appendix I

S. No.	Abbreviation	Description
1	CAD	Computer Assisted Diagnosis
2	CCNN	Customized Convolution Neural Network
3	CNN	Convolution Neural Network
4	FP	False Positive
5	FN	False Negative
6	RCNN	Recurrent Convolution Neural Network
7	RVG	Radiovisiography
8	SSD	Single-Shot Multi-Box Detector
9	TP	True Positive
10	TN	True Negative
11	VGG16	Visual Geometry Group