Classification of Dental Diseases Using CNN and Transfer Learning

Shreyansh A. Prajapati, R. Nagaraj and Suman Mitra
Dhirubhai Ambani Institute of Information and Communication Technology Gandhinagar, Gujarat, 382007 India
e-mail: shreyansh prajapati@daiict.ac.in, nagaraj ramrao@daiict.ac.in, suman mitra@daiict.ac.in

***Abstract*— Automated medical assistance system is in high demand with the advances in research in the machine learning area. In many such applications, availability of labeled medical dataset is a primary challenge and dataset of dental diseases is not an exception. An attempt towards accurate classification of dental diseases is addressed in this paper. Labeled dataset consisting of 251 Radio Visiography (RVG) x-ray images of 3 different classes is used for classification. Convolutional neural network (CNN) has become a most effective tool in machine learning which enables solving the problems like image recognition, segmentation, classification, etc., with high order of accuracy. It is found from literature that CNN performs well in natural image classification problems where large dataset is available. In this paper we experimented on the performance of CNN for diagnosis of small labeled dental dataset. In addition, transfer learning is used to improve the accuracy. Experimental results are presented for three different architectures of CNN. Overall accuracy achieved is very encouraging.**

***Keywords* — Convolutional Neural Network(CNN), Transfer Learning (TL), Machine Learning (ML), Computer Vision**

I. Introduction

There are various types of medical images like MRI (Magnetic Resonance Imaging), CT (Computed Tomography) scan, x-rays etc. which are useful for the diagnoses of the diseases. Deep learning algorithms have an incredible number of applications in the area of medical image processing. They are effectively used in problems like segmentation (brain tumor segmentation, liver tumor segmentation, anatomical brain segmentation, kidney segmentation [1]–[4]), detection (mitosis detection [5], glaucoma detection [6]) and many more. CNN performs exceptionally well in the classification problems (mass or normal breast tissue classification, lung nodule classification, lung pattern classification [7]–[10]). Hence analysis of medical images using deep learning algorithms have caught attention of many researchers. G. Litjens and T. Kooi, et al. [11], have summarized more than 300 applications of deep learning algorithms in medical imaging. But no literature was found wherein CNN was used for dental disease classification. In this research, CNN is used for dental disease classification.

Different types of x-ray detectors are used in dentistry [12]. Orthopantomogram (OPG) and Radiovisiography (RVG) x-ray images are the most widely used tools for the diagnoses of dental diseases. OPG image captures both the upper and lower teeth in one image. On the other hand, RVG x-ray images are used for the diagnosis of an individual tooth. Irregularities in RVG x-ray images (e.g, size, angle, shape, shadow etc.) make the analysis and classification difficult even for the the experienced radiologists. Such artifacts in the x-rays require domain knowledge, expertise and experience to reach to the conclusion. This process is time consuming and tiresome. Machines can do such tasks easily, rapidly and accurately. Dental diseases like dental caries, periapical infection and periodontitis can be identified by analyzing the RVG x-ray images [12]. The purpose of this research is to use CNN for training the machine so as to perform the classification of dental disease accurately. The input to the CNN is a set of RVG x-ray images. A library of labeled RVG images is used to train the CNN. Labeled images are those images in which the diseases type has already been classified.

Human tooth is mainly made up of two parts. One is the crown which is clinically visible and the other is the root which is not clinically visible but embedded in the jaw. Effect of disease on tooth, can be identified by analyzing the x-ray images. In particular, we have considered three diseases known as dental caries, periapical infection and periodontitis for the classification task. Healthy teeth can be considered as another class. It has not been taken into account in this task because of the data insufficiency.

Dental caries is one of the most common dental disease worldwide [13]. It is the medical terminology used for the common dental cavity or tooth decay. There can be different stages of dental caries, but the aim here is to classify disease and not the advancement in its stage. Figure-1 is an x-ray image of the patient suffering from dental caries. In the case of periapical infection, infection spreads around the root portion of the tooth. It results into dark sport around the root in the x-ray. Figure-2 is an x-ray image of the patient suffering from periapical infection. In the case of periodontitis disease, bacteria eats away the alveolar bone which provides support to the tooth. It results into the bone loss around the tooth. Figure-3 is an x-ray image of the patient suffering from periodontitis.

This paper discusses the CNN based classification of 3 major dental diseases. The paper is organized as follows. The background of CNN is discussed in next section followed by a description about the dataset. Section-IV discusses the architecture. Finally in Section-V, the experimental results are provided.

II. Background

A. Deep Learning

Neural Network (NN) is also known as *Artificial Neural Network (ANN)*. Neural Networks are inspired by the way

978-1-5386-1772-4/17/$31.00 ©2017 IEEE

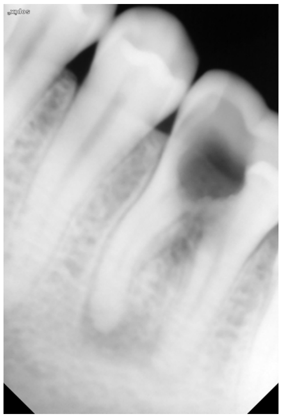

Fig. 1: RVG x-ray image of a patient suffering from dental caries disease.

Fig. 2: RVG x-ray image of a patient suffering from periapical infection.

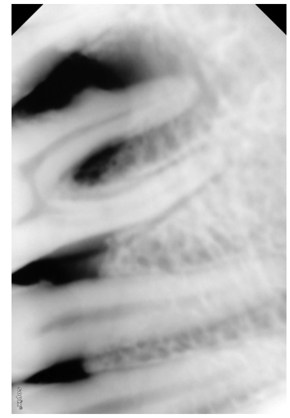

Fig. 3: RVG x-ray image of a patient suffering from periodontitis.

neurons work in human nervous system. ANN has layered architecture as described in Figure-4. Each layer has certain number of nodes with an activation function. Nodes are fully connected with the next layer. These connections have weights associated with them. Each neuron performs dot product of weight and input, and feed forwards the output to the next layer. ANN's architecture assumes fully connected (FC) layers.

These fully connected (FC) layers when used with the input being an image, increase the number of parameters drastically. Not only the rapid increase of parameters is a memory problem but may also cause overfitting. Overfitting refers to the state of the network when it perfectly learns the features on the train dataset but does not generalize pretty well on the test dataset. Therefore ANNs do not perform well with the images as an input.

The train dataset is divided into training and validation set. The network is trained on the training set and then evaluated on the validation set. The network tends to decrease the train error. The test error decreases first and then starts increasing after overfitting. The point where the test error is minimum is chosen to be the optimal setting for the network.

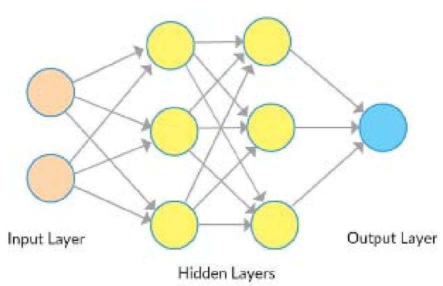

Figure 4.Artificial neural network with total 4 layers. First layer is an input layer, other two are hidden layers and last layer is an output layer.

Deep learning is state of the art for image classification problems. CNNs do not require features to be extracted beforehand. Designed CNN model can learn the features by itself. Convolutional neural network is one of the type of ANN which assumes that inputs are images. In ANN, fully connected (FC) layers increase the number of parameters in the network, but the CNNs consist of convolutional layers, FC, Dropout etc. As convolutional layers share the weights across the spatial dimension, it reduces the number of parameters and thus they can be easily scaled to image inputs where there are large number of input features.

B. Transfer Learning

1) Feature Extractor: A pre-trained model trained on a large dataset is used as a feature extractor. These features are then provided as inputs to the custom network. The pre-trained model learns the basic low level features from the large dataset.

2) Fine Tuning: The pre-trained model may learn high level features specific to the large dataset. To make the high level features compatible with the training data, fine tuning is performed. Weights of some layers are unfrozen and these layers are used for training. This is done in order for the pre-trained model to be more adaptive to the training data.

III. Labeling Dataset

Radiologists have large dataset of dental x-ray images, but x-ray images of individuals have privacy issues. Process of manually labeling the dataset is extremely tiresome and time-consuming. Large labeled dataset is required for training the CNN. There is no labeled dataset available online for the above mentioned dental diseases. So small dataset consisting of 251 images was labeled with the help of dentists and radiologists. Dataset can not be made public because of the privacy reasons. The dataset consists of total 251 grey images of dimension 1000 x 1496. Distribution of images is mentioned in Table-I.

IV. Network Architecture

We have used the dataset (mentioned in Table-I) for training all the models to compare the results. We have designed 2 different architectures. Architecture shown in Figure-5 is used for the first experiment. In architecture shown in Figure-6

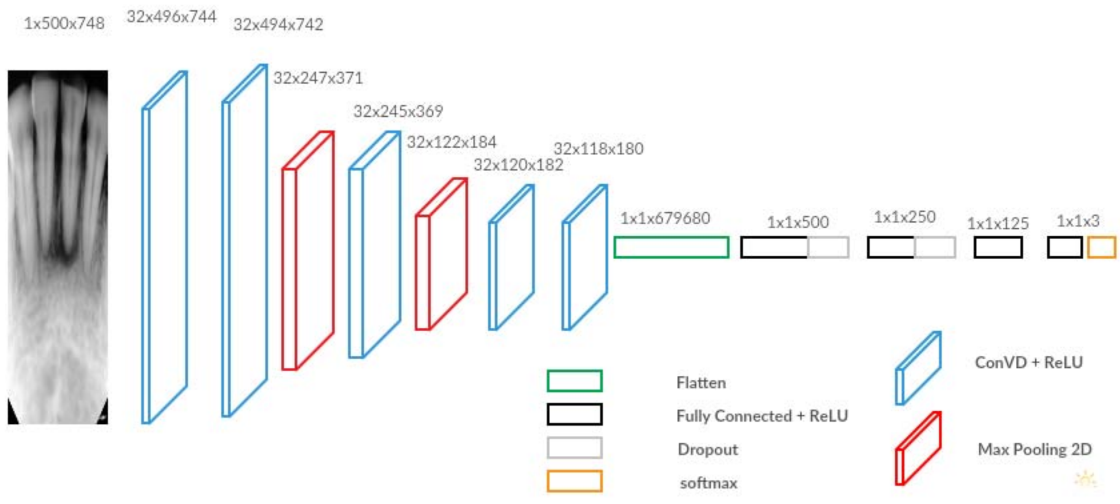

Figure 5. Designed CNN architecture. Images of dimension 1 x 500 x 748 are given as an input to the architecture. Different layers are represented by coloured blocks. The architecture has 5 Conv Layers, 2 MaxPooling Layers and 4 Fully Connected (FC) layers.

TABLE I. DATASET OF TOTAL 251 X-RAY IMAGES IS DISTRIBUTED IN THREE CLASSES

Disease Name	Number of Images
Dental Caries	80
Periapical Infection	110
Periodontitis	61
Total	**251**

VGG16 pre-trained network is used for transfer learning. Two different experiments were done on the second network.

A. CNN

The images are scaled down to 500 x 748 and fed as an input to the network. Reducing the image dimension results into information loss. On the other hand, processing high dimensional images is computationally expensive. Designed CNN architecture is show in Figure-5. The network consists of 5 convolution layers, 2 max pooling layers and 4 fully connected (FC) layers. All the convolution layers and the FC layers (except the last FC layer) have the activation function ReLU. Softmax activation function is applied to the last FC layer. Dropout of 0.4 and 0.2 is added between FC_1 and FC_2 and between FC_2 and FC_3 respectively [14]. Stride of (2,2) and kernel size of (2,2) is used in MaxPooling Layers to half the spatial dimension of the activation maps. Dropout is used to overcome the problem of overfitting.

B. Transfer Learning as a Feature Extractor

A pretrained VGG16 [15] is used as a feature detector. The VGG16 is trained on ILSVRC-2012 dataset with 1.3 million images and 1000 classes [16]. VGG16 model is trained for images of dimension 224 x 224. Therefore the images are scaled to dimension 224 x 224 and not to 500 x 748 as mentioned in previous architecture. Scaled down images of dimension 224 x 224 are fed as an input to the transfer learning architecture. Designed CNN and transfer learning architecture is show in Figure-6. Initial 9 layers (4 convolutional layers, 4 zero padding layers, 1 max pooling layer) of VGG16 are used as a feature extractor. The combined architecture has 8 Conv Layers, 4 Zero Padding Layers, 5 MaxPooling Layers and 8 FC Layers. ReLU activation function is applied at each Conv and FC layer (except the last FC Layer). Softmax is used as the activation function for the last FC Layer. Dropout of 0.2 added between all the fully connected layers except between the fully connected layer 7 and 8. Zero padding is used to keep the spatial dimension of the activation map constant. VGG16's pretrained weights were fixed and used as a feature extractor. The weights of the newly introduced layers were randomly initialized and used for training.

C. Transfer Learning with FineTuning

The images are scaled down to 224 x 224 and fed as an input to the network shown in Figure-6. The weights for the first 2 Conv Layers of VGG16 were frozen and all the subsequent layers were used for training. The idea of transfer learning to be used as fine tuning is to let the model learn some basic low level features of the image from a large dataset, but use the high level features to be more specific to the training data. This is done to make the model more adaptive to the training data.

V. EXPERIMENTAL RESULTS

All the 3 models discussed in the previous section are trained for 35 epochs. Out of total 251 x-ray images, 180 images are used for training, 45 images are used for validation and 26 images are used for testing purpose. Because of the

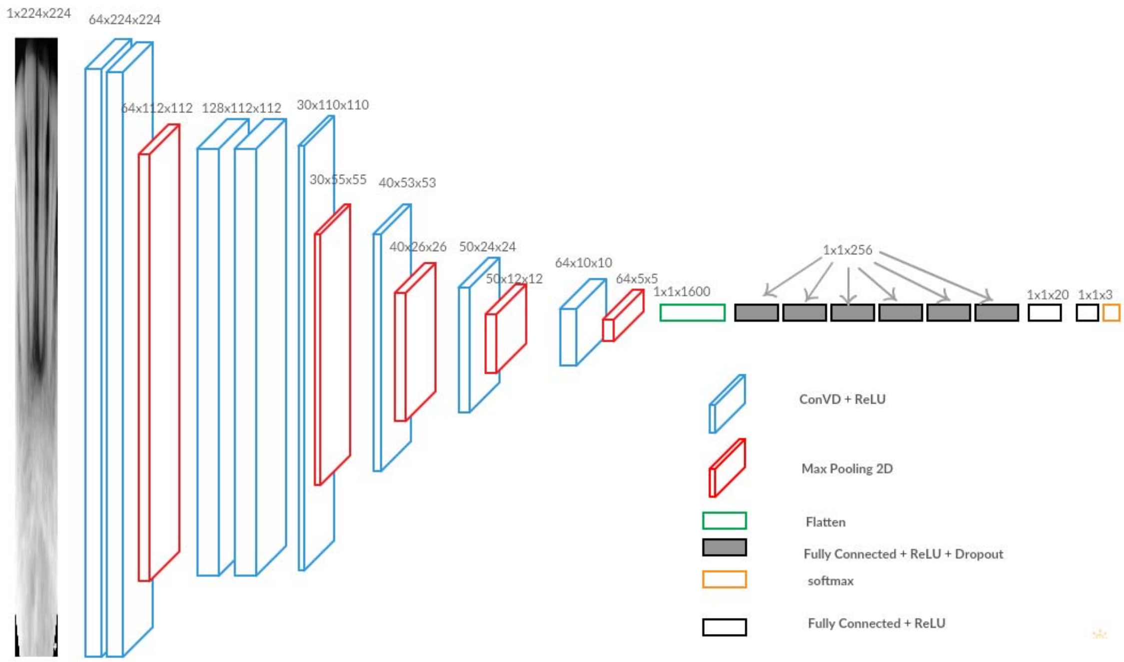

Fig. 6: Designed CNN and transfer learning architecture with VGG16 model. Images of dimension 1 x 224 x 224 are fed as an input to the architecture. Different layers are represented by coloured blocks. Initial 5 layers are the part of the pretrained VGG16 model. The architecture consists of total 8 Conv Layers, 5 MaxPooling Layers and 8 Fully Connected (FC) layers.

unavailability of the large dataset, CNN architecture mentioned in Figure-5 could not perform well in this classification task. Transfer learning can give better result even with the small dataset. Therefore we designed transfer learning architecture with VGG16 model mentioned in Figure-6. Accuracy is increased by 15.39% in transfer learning model as compared to CNN model. When transfer learning is used as a feature extractor, it may learn high level features which are not specific to the training data. Fine tuning is used to make those high level features suitable for our training set. Only initial 4 Conv Layers of VGG16 network are used as a feature extractor. These 4 layers must be learning very low level features (e.g, edges and curves) which are not specific to the training dataset. Hence fine tuning could not improve much in this task. Table-II shows the overall accuracy of 3 different models. As shown in Table-III, 23 out of 26 images are correctly classified and the highest accuracy of 88.46% is achieved. Confusion matrix is used to get the better understanding of the results. Table-III provides detailed analysis of the transfer learning model.

TABLE II. Results Summary

Model	**Accuracy**
CNN	0.7307
Transfer learning	0.8846
Transfer learning with fine tuning	0.8846

RVG x-ray images are used for the diagnosis of a particular tooth or portion. Angle, lightning, orientation etc. vary for different x-rays. OPG x-ray image covers all the upper and lower teeth in one image. All the OPG images have similar structure which means training model on dataset of OPG x-ray images may give better results.

TABLE III. Experimental Results for Transfer learning model

Disease Name	**Number of Samples**	**Correct results**	**Accuracy**
Dental Caries	8	7	0.875
Periapical Infection	10	9	0.90
Periodontitis	8	7	0.875
Total	**26**	**23**	**0.8846**

VI. Conclusion

Dental disease classification is done using convolutional neural network with transfer learning. Most common diseases known as dental caries, periapical infection and periodontitis are taken into account for the classification task. Three different models are designed on two different architectures. Transfer learning with VGG16 pretrained model is used to achieve better accuracy. Small dataset consisting of 251 RVG x-ray images is used for training and testing purposes. Experimental results for different models are discussed. The overall accuracy of 88.46% is achieved, which is very encouraging.

REFERENCES

[1] W. Li, F. Jia, and Q. Hu, "Automatic segmentation of liver tumor in ct images with deep convolutional neural networks," *Journal of Computer and Communications*, vol. 3, no. 11, p. 146, 2015.

[2] A. de Brebisson and G. Montana, "Deep neural networks for anatomical brain segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015, pp. 20–28.

[3] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in mri images," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.

[4] W. Thong, S. Kadoury, N. Piché, and C. J. Pal, "Convolutional networks for kidney segmentation in contrast-enhanced ct scans," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pp. 1–6, 2016.

[5] S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, and N. Navab, "Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1313–1321, 2016.

[6] X. Chen, Y. Xu, D. W. K. Wong, T. Y. Wong, and J. Liu, "Glaucoma detection based on deep convolutional neural network," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*. IEEE, 2015, pp. 715–718.

[7] B. Sahiner, H.-P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images," *IEEE transactions on Medical Imaging*, vol. 15, no. 5, pp. 598–610, 1996.

[8] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in *International Conference on Information Processing in Medical Imaging*. Springer, 2015, pp. 588–599.

[9] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, "Lung pattern classification for interstitial lung diseases using a deep convolutional neural network," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1207–1216, 2016.

[10] B. Q. Huynh, H. Li, and M. L. Giger, "Digital mammographic tumor classification using transfer learning from deep convolutional neural networks," *Journal of Medical Imaging*, vol. 3, no. 3, pp. 034 501–034 501, 2016.

[11] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *arXiv preprint arXiv:1702.05747*, 2017.

[12] A. G. Farman and T. T. Farman, "A comparison of 18 different x-ray detectors currently used in dentistry," *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, vol. 99, no. 4, pp. 485–489, 2005.

[13] R. H. Selwitz, A. I. Ismail, and N. B. Pitts, "Dental caries," *The Lancet*, vol. 369, no. 9555, pp. 51–59, 2007.

[14] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting." *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014. [Online]. Available: http://arxiv.org/abs/1409.1556

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.