Automatic Classification of Oral Pathologies Using Orthopantomogram Radiography Images Based on Convolutional Neural Network

Anuradha Laishram*, Khelchandra Thongam

Department of Computer Science and Engineering, National Institute of Technology Manipur (India)

Received 20 July 2020 | Accepted 31 July 2021 | Published 29 October 2021

Abstract

An attempt has been made to device a robust method to classify different oral pathologies using Orthopantomogram (OPG) images based on Convolutional Neural Network (CNN). This system will provide a novel approach for the classification of types of teeth (viz., incisors and molar teeth) and also some underlying oral anomalies such as fixed partial denture (cap) and impacted teeth. To this end, various image preprocessing techniques are performed. The input OPG images are resized, pixels are scaled and erroneous data are excluded. The proposed algorithm is implemented using CNN with Dropout and the fully connected layer has been trained using hybrid GA-BP learning. Using the Dropout regularization technique, over fitting has been avoided and thereby making the network to correctly classify the objects. The CNN has been implemented with different convolutional layers and the highest accuracy of 97.92% has been obtained with two convolutional layers.

Keywords

Classification, CNN, Dropout, Image Pre-processing, Orthopantomogram Radiography Images.

DOI: 10.9781/ijimai.2021.10.009

I. Introduction

In human anatomy, the mouth is one of the principal organs and being the first portion of gastrointestinal tract, it plays a significant role in communication, breathing and digestion. This cavity organ involves different organs, for example, teeth, tongue, hard sense of taste and delicate sense of taste. Great oral wellbeing can give numerous extraordinary advantages. It improves our visual appearance, the inspiration of our mentality just as improving the personal satisfaction. An unhealthy mouth, exceptionally gum malady and tooth rot are the principal two reasons for tooth loss and furthermore offer ascent to numerous clinical issues and causes inconvenience, draw out agony, distortion. So, it makes sense to give oral health the best possible care. There are various kinds of dental x-rays to check the status of our oral wellbeing. Dental x-rays are a helpful demonstrative instrument of any dental consideration treatment plan which utilized low degree of radiation to catch pictures of the inside of the mouth region. The utilization of Dental x-ray images had denoted an unprecedented milestone in clinical discovering in light of its brief openness and for the most part, lower radiation estimation. One such radiograph image utilized in our work is Orthopantomogram radiograph.

An Orthopantomogram (OPG) by and large known as panoramic radiography catches wide perspective of the mouth in a solitary x-ray image including the upper and lower jaw. It distinguishes the situation of completely developed just as affected teeth and furthermore recognize other bone anomalies An Orthopantomogram radiograph image is shown in Fig. 1.

* Corresponding author.

E-mail address: annu0286@gmail.com

Fig. 1. An Orthopantomogram X-ray.

Over the recent years Convolutional Neural Network has made a significant contribution in the field of medical science [1]-[5]. Many researchers started working on the classification of dental images and signals recently [6]-[12]. Some researchers have explained the use of Convolutional network in the field of dentistry.

Oprea et al. [13] proposed how image processing techniques can help to check the x-ray and examine the extent to which the caries lesion is present and then classify the type of caries present in the dental radiograph. Two different operations such as thresholding images, image differentiation are used in order to detect the dental carries and hence the efficiency of the algorithm is compromised. The database used for the experimentation is not of a good quality due to the lack of better instrumentation (x-ray machines) at that time.

Prajapati et al. [14] made an endeavor towards precise arrangement of three dental ailments, specifically Dental Caries, Periapical Infection and Periodontitis. Marked dataset comprising of 251 Radiovisiography (RVG) x-ray images is utilized. In this paper, CNN is used for determination of small labelled dental dataset. Likewise, Transfer Learning is utilized to improve the precision. The overall accuracy of 88.46% is accomplished. In Yu. [15], a method in which SIFT, HOG2Ö2, HOG3Ö3, and Color features are used as image descriptors, the Bag-of-Words (BoW) is applied and also a spatial pyramid pipeline. K-nearest neighbors algorithm (k-NN) and an error correcting output coding support vector machine (ECOC-SVM) classifiers are deployed. 4-layer CNN model and 16-layer CNN model methods were proposed. Sharpness and precision of the x-ray images were not determined due to which the system was not very reliable. A maximum accuracy of 90.36% was achieved but the dataset size was very small.

Srivastava et al. [16] developed a Computer Aided Diagnosis (CAD) system that enhances the performance of dentists in detecting wide range of dental caries. Annotations were made using loose polygons around the carries and classification is done using a Fully Convolutional Neural Network. The dataset was collected from 100+ different clinics and hence providing different image quality and also required some type of normalization. Since it was difficult to annotate the carries, it was done using a loose polygon around the carries. Also the amount of False Positive test results was also high. Veena et al. [17] presented two distinct image processing algorithms for detection of dental anomalies. The first approach for the detection of dental caries uses hybridized negative transformation. The second approach uses statistical texture analysis for the dental images containing cysts along with dental caries. Gray Level Co-occurrence Matrix (GLCM) is used for feature extraction from panoramic images. The system was experimented on several datasets and high percentage of false positive were output in the first method.

Zakirov et al. [18] proposed a system that achieves 96.3% accuracy in tooth localization and an average of 0.94 AUROC for 6 common tooth conditions using deep convolutional neural networks and algorithmic heuristics. Cone-Beam CT imaging is a time-consuming process that requires a physician to work with complicated software. Since it a new imaging technique in dentistry, the availability of the machines is questionable. Also, the very small amount of dataset cannot be considered to present the validity of the proposed system.

In our study a Computer Aided Diagnosis system is being proposed in order to assist dentists in detection and diagnose the dental pathology faster and more accurately and also provide better judgment in the treatment of any oral diseases. The study will provide a novel approach in classifying the type of teeth such as incisors and molar teeth and also some underlying oral anomalies such as fixed partial denture and impacted teeth present in an OPG using deep learning technique. The Convolutional Neural Network with Dropout is used to classify the images and is implemented using Keras, a python neural network API. The fully connected layer of the CNN has been trained using hybrid GA-BP learning and provides a classification accuracy of 97.92%. The remaining paper is organized as follows: Section II describes the proposed methodology. Section III gives the details of Experimental Results. Section IV gives the comparison details of our proposed method with other standard method. Finally, the conclusion is given in section V.

II. Proposed Methodology

A flow diagram of the proposed methodology is given in Fig. 2. The proposed system, however, is an end-to-end solution, which classifies some of the dental pathology directly from the original radiographs with a bit of image pre-processing of the dental panoramic radiographs. This has been possible by using Convolutional Neural Networks with deep learning. A convolutional neural network (CNN or ConvNet) is one of the most popular algorithms for deep learning, a type of machine learning in which a model learns to perform classification tasks directly from images. The methodology consists of the following modules:

1. Image Acquisition
2. Image Pre-processing
3. Building of a CNN
4. Image Augmentation
5. Training, Validation and Testing

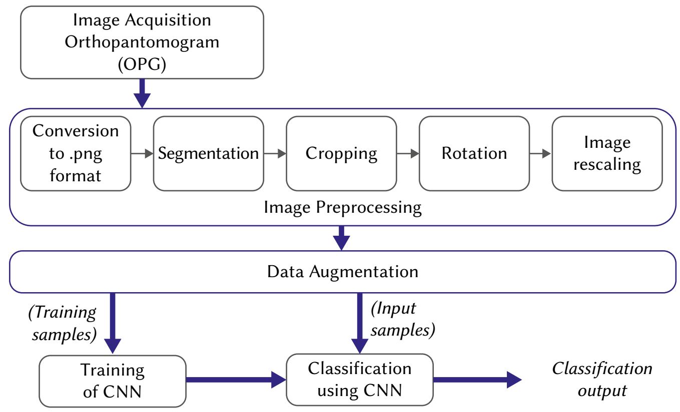


Fig. 2. Flow diagram of the proposed algorithm.

A. Image Acquisition

Image acquisition is the initial step of any vision framework. The images acquired are not a processed image. After the acquisition of images, further processing techniques has been applied to process the image. The process of image acquisition is done using Care stream (KODAK) Dental's CS 8100; an OPG machine that has machine has gained exceptional value in the market for its efficient performance and sturdiness.

B. Image Pre-processing

Image pre-processing is a technique to select and enhance the region of interest of acquired digitized images by applying various mathematical operations. The purpose of pre-processing was discussed in [19], [20]. In our study, we apply different image pre-processing techniques namely edge detection, cropping, rotation and image scaling. Before applying any preprocessing techniques, the raw data which is in .pano format is converted to .png format in order to be used for further processing.

Edge Detection: Edge detection is one of the vital steps of image preprocessing. Although cropping operation segments the teeth from the original image, it does not crop along the boundary of teeth. Edge detection technique are used to define the edge of the teeth to find the region of interest. In our study to perform the task of segmentation the Active Contour method [21] (also known as snake method) was used.

Cropping: Cropping is one of the most well-known image activity which is performed on the obtained images to expel the undesirable composed names and to complement the region of interest bringing about a rectangular shape comprising of just the region of interest which is required for the classification and furthermore lessen the potential outcomes of mistake.

Image Rotation: To get better visualization of the acquired images, the images are rotated about its center to a specified degree (θ=45, 90). This is performed to maintain uniformity in images representation.

Image scaling: In image pre-processing, image scaling alludes to the resizing of a picture. Image resizing is essential when we have to increment or diminishing the all-out number of pixels, though remapping can occur when you are correcting for lens distortion or rotating an image.

C. Building of CNN

Convolutional Neural Network consists of Convolution as the first layer to extract features from the image. The feature maps are passed on to activation function known as Rectified Linear Unit(ReLU) and then the input is passed on to pooling layers and finally given to the fully connected network, which gives the output [22]-[28]. Fig. 3 shows our CNN model which contains two convolutional layers with two pooled feature mapped layers. Our CNN architecture has been implemented using Dropout. Dropout is a regularization technique for neural network models to prevent the networks from over fitting by ignoring randomly selected neurons during training. Fig. 8 shows the proposed CNN architecture using two convolutional layers with dropouts.

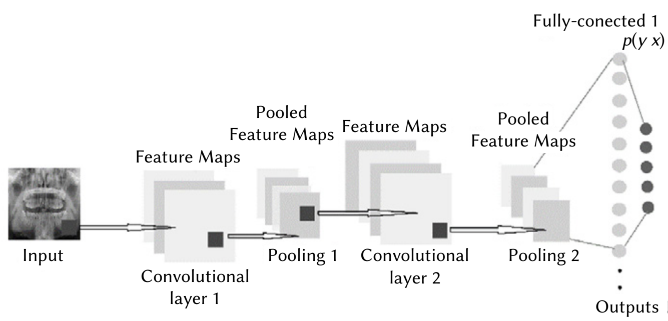


Fig. 3. Model of the Proposed CNN.

1. Convolution Layers

Convolutional Neural Network consists of Convolution as the first layer to extract features from the image also preserves the relationship between pixels by learning images features using small square of input data known as feature maps. The input image is convolved with a feature detector also called kernel or filter. To increase its values by the original pixel values is the task of this filter. The equation for this mathematical operation is given as:

$$s[t] = (x * w)[t] = \sum_{a=-\infty}^{a=\infty} x[a]w[a+t] \tag{1}$$

where $s[t]$ is the feature map, x is the input, w is the filter.

Fig. 4 shows convolutional layers with feature mapping function and activation function.

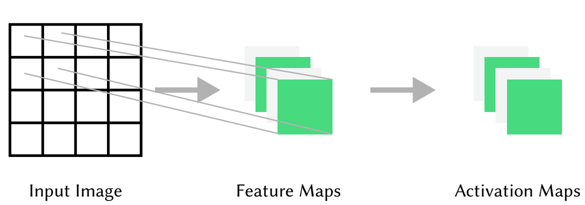


Fig. 4. Convolution Layers.

2. Non-linear Units

Since this proposed method is to classify dental pathology radiograph images and image classification is a nonlinear problem, non-linear unit increases non-linearity into the input images. As non-linearity is achieved through an activation function known as the ReLU (Rectified Linear Unit) function. Without this property a network would not be sufficiently intense and will not be able to model the response variable. Mathematically, ReLU converts all negative values to zero which decreases the computational complexity. Fig. 5 shows the graph of Rectified Linear Unit. The equation for ReLU activation function is as follows:

$$Equation: f(x) = \begin{cases} 0, x < 0 \\ x, x \geq 0 \end{cases} \tag{2}$$

$$Derivative: f'(x) = \begin{cases} 0, x < 0 \\ 1, x \geq 0 \end{cases} \tag{3}$$

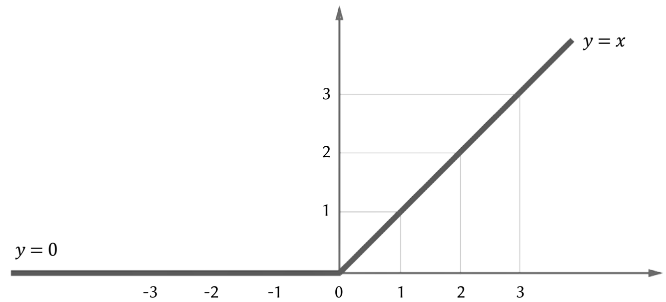


Fig. 5. Rectified Linear Unit (ReLU).

3. Pooling Layer

The next layer in a convolutional network comprises three steps which include pooling, down sampling and subsampling. This layer simplifies the output by performing nonlinear down sampling, reduces the spatial size of the representation to reduce the number of parameters and computation in the network. Out of many pooling techniques, Max pooling is used as it preserves the most matching feature in the activation map and also it helps to gain spatial invariance.

4. Fully Connected Layer

Fig. 6 shows a fully connected layer. Before the pooled layers are completely fed to the fully connected layer, the pooled feature map needs to be converted to a 1-D vector. This can be achieved by a process called the flattening process. The resultant 1-D vector contains the spatial structure information or pixel pattern and passed on to the fully connected network. The final layer of the CNN architecture which main purpose is to combine the feature into more attributes uses a classification layer such as softmax to provide the classification output. A fully connected layer works with high level features which strongly correlates to a particular class and particular weights that the products between the weights and the preceding layer has the accurate probabilities for the different classes.

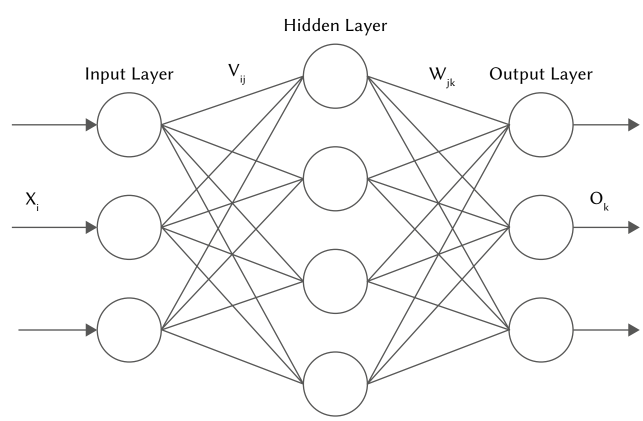


Fig. 6. Fully Connected Layer.

The proposed fully connected layer model contains three layers: input, hidden and output which is used to classify the images into different classes. The layers are connected by synaptic weights. The learning of the network is realized by a hybrid system of Genetic Algorithm (GA) and Back Propagation (BP). The GA is an evolutionary algorithm and is based on the principle of natural selection and bio-inspired operations.

GA is a global search algorithm good at determining the rough position of the global optimal solutions and BP is a local search algorithm good at fine tuning. It is possible to have both advantages by combining GA and BP. Use GA first to find the rough position of the global optimal solution and then use BP to refine.

Steps of the GA-BP hybrid learning:

1. Use GA learning first to find a good starting point
2. Use BP learning to find a local Solution
3. If the solution is worse than the current best result, return to Step 1; otherwise, continue
4. If the current best good enough, stop; otherwise, return to Step 1

Important parameters that are used during the learning process are as follows:

x_i: The i^{th} input

y_i: The output of the j^{th} hidden neuron

o_k: The output of the k^{th} output neuron

d_k: The desired output

v_{ij}: The weight from the i^{th} input to the j^{th} hidden neuron

w_{ij}: The weight from the j^{th} hidden neuron to the k^{th} output neuron

The steps of the GA learning are as follows:

i) Initialization of the weights

The weights of the connections of the neurons from one layer to another of the fully connected network will be represented by GA solution which is a binary string.

The number of total weights, TW is given by

$$TW = I * HN + HN * ON \tag{4}$$

where I is the size of each input sample, HN is the number of hidden neurons, ON is the number of output neurons.

In our work, I = 3136, HN = 64 and ON = 4, so TW comes out to be 200960.

The gene length, GL is given by the equation

$$GL = \big(NB * (I * HN + HN * ON)\big) \tag{5}$$

Where NB = Number of bits per weight

In our work, NB = 6, so GL comes out to be 1205760.

ii) Phenotype reconstruction using genotype

Consider,

$$y_i = \textstyle\sum_{k=1}^{NB} b_{ik}\, 2^{-k} \tag{6}$$

Where b_{ik} is the k^{th} bit for the i^{th} weight. Then,

$$w_i = y_i \times A + B \tag{7}$$

w_i is the i^{th} weight present in the string or solution, A is the scaling factor and B is the shifting factor.

In our application, we set A=20 and B=10 so that the weight will take value from [-10, 10].

iii) Output of the hidden layer and the output layer

We calculate the outputs of the hidden neurons using the relations:

$$s1 = \textstyle\sum_{i,j} v_{ji} \times x_{pi} \tag{8}$$

$$y_j = sigmoid(s1) \tag{9}$$

Where y_i is the output of the j^{th} hidden neurons.

We calculate the output of the output neurons with the equations:

$$s2 = \textstyle\sum_{j,k} w_{kj} \times y_j \tag{10}$$

$$o_k = sigmoid(s2) \tag{11}$$

Where o_k is the output of the k^{th} output neurons. The function sigmoid is a unipolar sigmoid function.

These two operations of finding the output are performed for all the input patterns. Then the Error is updated with the equation:

$$E = \tfrac{1}{2}\textstyle\sum_{k=0}^{k}(d_k - o_k)^2 \tag{12}$$

d_k is the desired output. This process is performed until all the training samples have been used.

iv) Calculate the fitness of string or solution.

The fitness definition to calculate the fitness of the string/solution is given by the equation

$$fitness = (1 - E)/N \tag{13}$$

Where, N is the number of patterns or training examples. The above processes are repeated from step ii) for all the strings or solutions of the population.

v) Selection

We find out the string with the highest fitness value. The weights representing this string with highest fitness value will be used for further operation.

vi) Reproduction

The population is modified using operators namely crossover and mutation. The above processes from step ii) are repeated for many generations.

The steps of the BP learning are as follows:

The size of each input sample is I. There are J hidden neurons and K output neurons.

1. Get a training input, calculate the output of hidden neurons and output neurons
2. Calculate the learning signals

The learning/error signal δ_k produced by the k^{th} output neuron is given as:

$$\delta_k = (d_k - o_k) * (1 - o_k) * o_k \tag{14}$$

Using δ_k, we can calculate δ_j as follows:

$$\delta_j' = \textstyle\sum_{k=0}^{k} \delta_k * w_{jk} \tag{15}$$

$$\delta_j = \big(1 - y_j\big) * y_j * \delta_j' \tag{16}$$

3. Update the weights

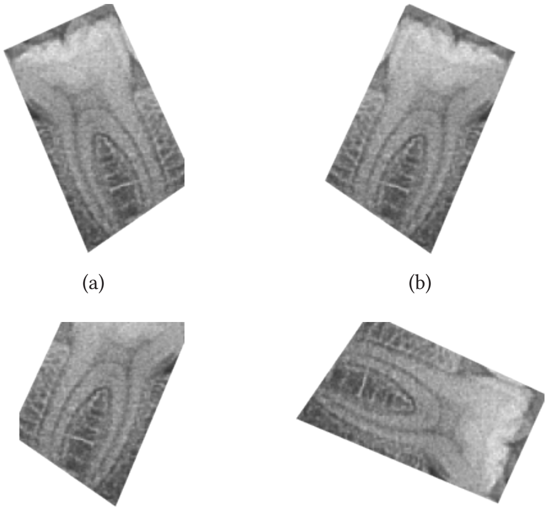
The weights of the output layer are updated using the equation:

$$w_{jk} = w_{jk} + \eta * \delta_k * y_j \tag{17}$$

Where η is the learning rate.

The weights of the hidden layer is updated using the equation

$$v_{ij}(n+1) = v_{ij}(n) + \eta * \delta_j(n) * x_i(n) \tag{18}$$



(a) (b) (c) (d)

Fig. 7. Image Augmentation: (a) Original image (b) Horizontal flip (c) Height shift (d) Rotation by 90°.

4. Update the Error

   Error is updated using the equation:

$$Error = \frac{1}{2}\sum_{k=0}^{k}(d_k - o_k)^2 \quad (19)$$

5. Check if all the training patterns have been used, if not return to Step 1.
6. If Error is smaller than desired error, then terminate otherwise repeat the whole process

D. Image Augmentation

Data augmentation is a way of creating new 'data' with different orientations [29], [30]. The benefits of this are twofold, the first being the ability to generate 'more data' from limited data and secondly it prevents over fitting making it less likely that the neural network recognizes unwanted characteristics in the data-set. Image data augmentation is applied to the training dataset, and not to the validation or test dataset as shown in Fig. 7. This is different from data preparation such as image resizing and pixel scaling; they must be performed consistently across all datasets that interact with the model.

E. Training, Validation and Testing

800 images belonging to 4 classes are used for training, 116 images belonging to 4 classes for validation. The training has been done with 16 numbers of batches, validation with 2 numbers of batches. Python Fit-generator method is employed for training, validation and testing of the dataset. The generator runs parallel to the model which in turn allows training the model on GPU and in parallel to perform real-time data augmentation on images on CPU.

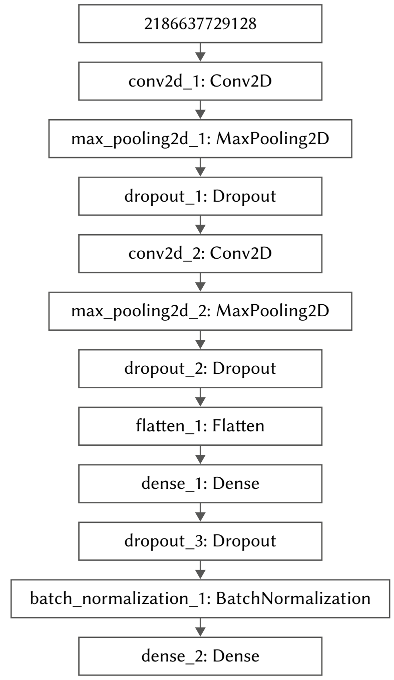


Fig. 8. Proposed architecture using two Convolutional Layer.

III. Experimental Results and Discussions

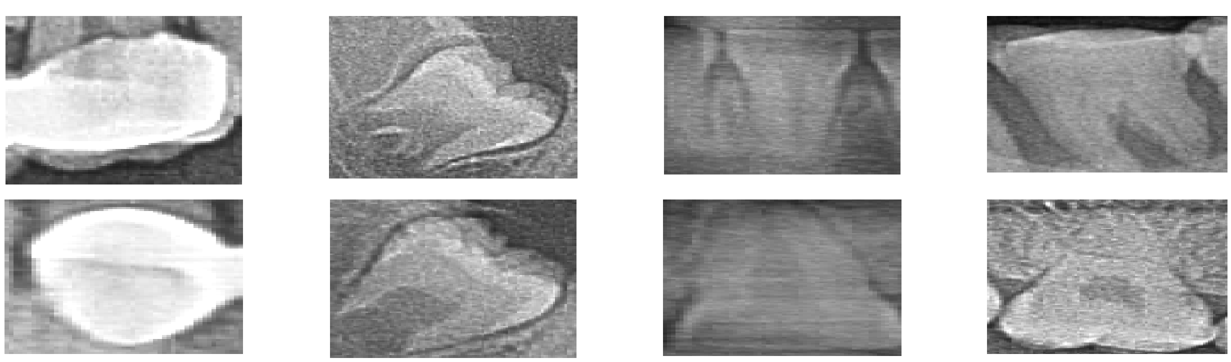
CNN is implemented for training, validation and testing using different convolutional layers. Using 2 convolutional layers, the best accuracy is obtained and the results are provided in the section.

A. Image Pre-processing

Different dental anomalies and different types of teeth are shown in Fig. 9 which is a result of image pre-processing. a) and b) are types of dental anomalies namely Cap(fixed partial denture) and Impacted teeth whereas c) and d) are types of teeth namely Incisor and Molar.

B. Dataset

As shown in Table I, 800 images belonging to 4 classes are used for training, 116 images belonging to 4 classes for validation. The training has been done with 16 numbers of batches, validation with 2 numbers of batches. Testing has been done with 120 random samples.



(a) Cap (Fixed partial denture) (b) Impacted (c) Incisor (d) Molar

Fig. 9. Input samples of Class A, B, C, D respectively.

TABLE I. Split of Samples in the Dataset

Dataset	Number of Images
Training Set	200 images for class A
	200 images for class B
	200 images for class C
	200 images for class D
Validation Set	29 images for class A
	29 images for class B
	29 images for class C
	29 images for class D
Test Set	120 random images for testing

C. GA Fitness and Accuracy and Loss for Training and Validation

GA is use at determining the rough position of the global optimal solutions and BP is use for fine tuning. We use GA learning first to find a good starting point.

The Number of bits per weight is 6. The GeneLength comes out to be 1205760. The population size is 60. The mutation rate is set as 0.001. The fitness of the solution/string obtained after some GA generations is given in Table II.

TABLE II. Highest Fitness of the Solution for Different GA Generations

Sl. No.	Generation	Training Loss	Fitness Value
1.	30	1.633	0.397
2.	60	0.792	0.484
3.	90	0.524	0.563
4.	120	0.501	0.621
5.	150	0.495	0.684

After GA learning generates global optimal weights, we use BP learning to refine the weights.

i) As shown in Fig. 10, using 2 convolutional layers and epochs upto 400, the training loss and accuracy comes out to be 0.19 and 97.57 respectively. The validation loss and accuracy comes out to be 0.18 and 97.94 respectively.

ii) As shown in Fig. 11, using 2 convolutional layers and epochs upto 1000, the training loss and accuracy comes out to be 0.23 and 96.32 respectively. The validation loss and accuracy comes out to be 0.69 and 83.85 respectively. We can see that validation loss rises more and validation accuracy drops down because of over fitting.

D. Model Graph

i) Fig. 12(a) shows Model Accuracy graph of Training and Validation between Epoch and Accuracy with epochs upto 400. We can observe that almost as the epoch increases, the accuracy increases with some drop down in between. The training accuracy and validation accuracy comes out to be 97.57 and 97.94 respectively.

ii) Fig. 12(b) shows Model Loss graph of Training and Validation between Epoch and Loss with epochs upto 400. Here, we can observe that almost as the epoch increases, the loss decreases with some rises in between. The training loss and validation loss comes out to be 0.19 and 0.18 respectively.

iii) Fig. 13(a) shows Model Accuracy graph of Training and Validation between Epoch and Accuracy with epochs up to 1000. The training accuracy and validation accuracy comes out to be 96.32 and 83.85 respectively. We can see that validation accuracy drops down because of over fitting.

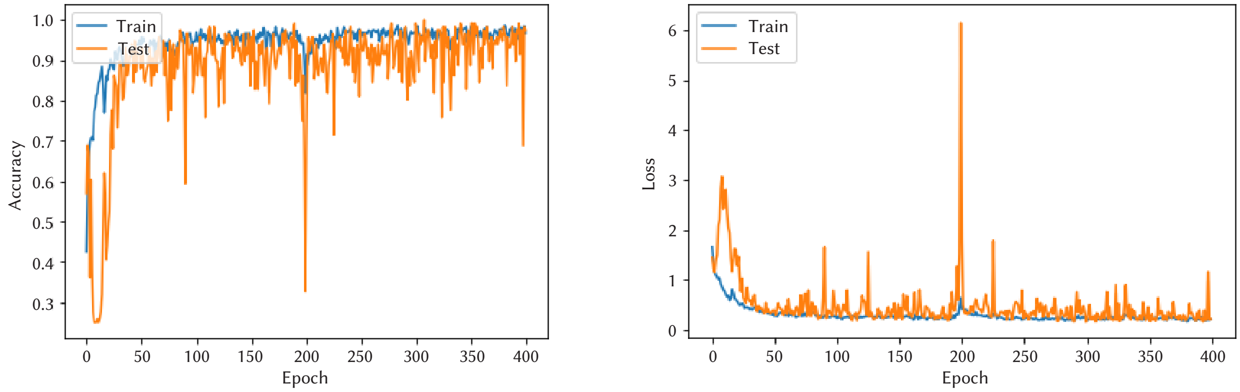
iv) Fig. 13(b) shows Model Loss graph of Training and Validation between Epoch and Loss with epochs up to 1000. The training loss and validation loss comes out to be 0.23 and 0.69 respectively. We can see that validation loss rises because of over fitting.

```
 16/16   [=========================] - 1s  80ms/step - loss:0.2243 - aacc:0.9782 - val_loss:0.1532 - val_acc:0.9932
Epoch 395/400
 16/16   [=========================] - 1s  78ms/step - loss:0.1410 - aacc:0.9852 - val_loss:0.2784 - val_acc:0.9247
Epoch 396/400
 16/16   [=========================] - 1s  83ms/step - loss:0.1932 - aacc:0.9798 - val_loss:0.1812 - val_acc:0.9805
Epoch 397/400
 16/16   [=========================] - 1s  80ms/step - loss:0.1713 - aacc:0.9883 - val_loss:0.1937 - val_acc:0.9818
Epoch 398/400
 16/16   [=========================] - 1s  79ms/step - loss:0.2184 - aacc:0.9648 - val_loss:0.6694 - val_acc:0.7612
Epoch 399/400
 16/16   [=========================] - 1s  81ms/step - loss:0.1937 - aacc:0.9875 - val_loss:0.1324 - val_acc:0.9878
Epoch 400/400
 16/16   [=========================] - 1s  79ms/step - loss:0.1927 - aacc:0.9757 - val_loss:0.1873 - val_acc:0.9794
```

Fig. 10. Accuracy and loss upto 400 epochs.

```
Epoch 995/1000
 16/16   [=========================] - 1s  74ms/step - loss:0.1944 - aacc:0.9732 - val_loss:0.3482 - val_acc:0.9098
Epoch 996/1000
 16/16   [=========================] - 1s  75ms/step - loss:0.1832 - aacc:0.9644 - val_loss:0.1892 - val_acc:0.9692
Epoch 997/1000
 16/16   [=========================] - 1s  77ms/step - loss:0.2145 - aacc:0.9631 - val_loss:0.2619 - val_acc:0.9472
Epoch 998/1000
 16/16   [=========================] - 1s  75ms/step - loss:0.1891 - aacc:0.9752 - val_loss:0.2591 - val_acc:0.9340
Epoch 999/1000
 16/16   [=========================] - 1s  75ms/step - loss:0.1653 - aacc:0.9792 - val_loss:0.1935 - val_acc:0.9711
Epoch 1000/1000
 16/16   [=========================] - 1s  74ms/step - loss:0.2311 - aacc:0.9632 - val_loss:0.6992 - val_acc:0.8385
```

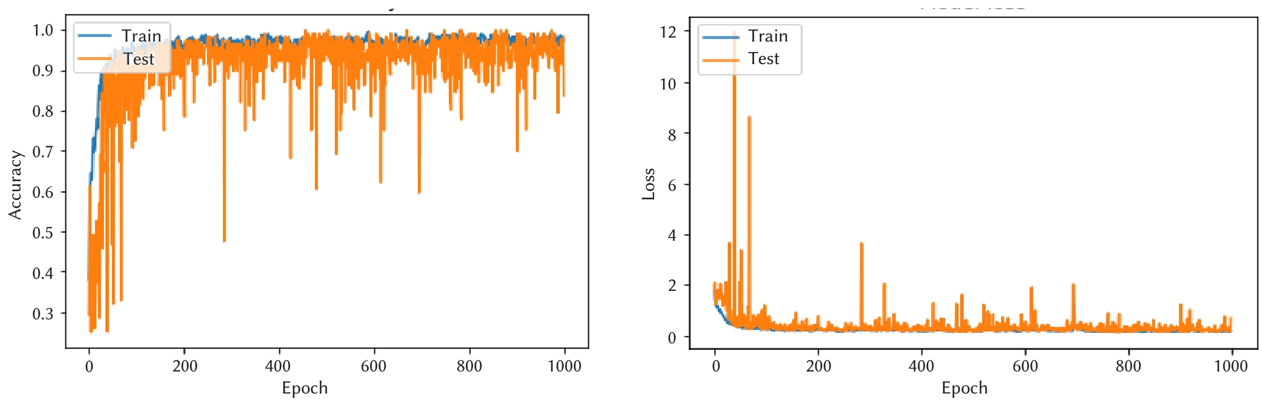
Fig. 11. Accuracy and loss upto 1000 epochs.



(a) Model accuracy

(b) Model loss

Fig. 12. Model Graph up to epochs 400.



(a) Model accuracy

(b) Model loss

Fig. 13. Model Graph up to epochs 1000.

E. Performance Analysis

For the performance analysis of the classification, Recall, Precision, F1 Score and Accuracy (Acc) are used.

$$Precision=\frac{TP}{TP+FP} \tag{20}$$

$$Recall=\frac{TP}{TP+FN} \tag{21}$$

$$F1Score=\frac{2*Recall*precision}{Recall+precision} \tag{22}$$

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN}\times 100 \tag{23}$$

where, TP, TN, FP and FN are the True Positive, True Negative, False Positive and False Negative respectively.

$$\begin{matrix} [[\,29 & 0 & 0 & 0\,] \\ [\,0 & 28 & 1 & 0\,] \\ [\,0 & 0 & 29 & 0\,] \\ [\,0 & 1 & 0 & 28\,]] \end{matrix}$$

Fig. 14. Confusion matrix (up to epochs 400).

Fig. 14 shows the Confusion Matrix of the classification done with our 29x4=116 samples and upto 400 epochs. The method has been implemented with 2 convolutional layers. From the matrix, the following information is obtained:

i) of the actual 29 Cap samples, our algorithm has predicted that 29 are Cap.

ii) of the actual 29 Impacted samples, our algorithm has predicted that 28 are Impacted and 1 is Incisor.

iii) of the actual 29 Incisor samples, our algorithm has predicted that 29 are Incisor.

iv) of the actual 29 Molar samples, our algorithm has predicted that 28 are Molar and 1 are Impacted.

The Precision, Recall, F1 Score and Accuracy of the four classes are provided in Table III. The Average Precision, Average Recall and Average F1 Score are found to be 98%, 98% and 98% respectively. The Average Accuracy for the classification is found to be **97.92%**.

TABLE III. Performance Analysis for Epochs 400

Category	Precision	Recall	F1	Acc
Cap	1.00	1.00	1.00	100.00
Impacted	0.97	0.97	0.97	96.45
Incisor	0.97	1.00	0.98	98.00
Molar	1.00	0.97	0.98	97.25
Average	0.98	0.98	0.98	**97.92**

Fig. 15 shows the Confusion Matrix of the classification done with our 29x4=116 samples and upto 1000 epochs. The method has been implemented with 2 convolutional layers.

The Precision, Recall, F1 Score and Accuracy of the four classes are provided in Table IV. The Average Precision, Average Recall and Average F1 Score are found to be 90%, 84% and 83% respectively. The Average Accuracy for the classification is found to be **83.64%**. The accuracy drops down as the epochs increases because of overfitting.

$$\begin{bmatrix} 27 & 0 & 1 & 1 \\ 0 & 27 & 1 & 1 \\ 0 & 0 & 29 & 0 \\ 0 & 1 & 14 & 14 \end{bmatrix}$$

Fig. 15. Confusion matrix (up to epochs 1000).

TABLE IV. Performance Analysis for Epochs 1000

Category	Precision	Recall	F1	Acc
Cap	1.00	0.93	0.96	95.93
Impacted	0.96	0.93	0.95	95.78
Incisor	0.70	1.00	0.82	81.52
Molar	0.88	0.48	0.61	61.27
Average	0.90	0.84	0.83	**83.64**

The algorithm has been implemented with 1 convolutional layer, 2 convolutional layers, 3 convolutional layers and 4 convolutional layers with CNN. As shown in Table V, the highest accuracy is obtained when the CNN is implemented with 2 Convolutional layers.

TABLE V. Loss and Accuracy of the Model With Different Convolutional Layer

No. of Convolutional Layer	Loss	Accuracy	Validation loss	Validation accuracy
1	0.6012	80.12	1.2054	73.78
2	0.1927	97.57	0.1873	97.94
3	0.2931	91.73	0.5774	79.59
4	0.2547	91.84	0.6711	77.87

IV. Comparison

Our method is compared with four standard algorithms and the comparison is shown in Table VI. The details of the four other methods are shown in first four rows and our proposed algorithm in last row.

The method proposed by [31] using CNN gives the classification accuracy of 85.3%.

The method proposed by [32] using Deep Convolutional Neural Networks (DCNNs) gives the Classification accuracy of 0.81(0.02). Mean (SD) sensitivity and specificity was 0.81 and 0.81 respectively. [33] Proposed a method using CNN which gives a classification accuracy rate of 83.0%.

[34] Proposed a method using Deep Convolutional Neural network (DCNN) which gives a classification accuracy of 95.6%.

Our proposed method uses CNN with Dropout and gives a classification accuracy of **97.92**%. Using the Dropout regularization technique, over fitting has been avoided and thereby the training and validation loss has been dropped down to significant level.

V. Conclusion

Over the past few years, the convolutional neural networks (CNNs) model has gain a huge potential in the field of medical imaging technology. In this paper, we have proposed a computer aided system which helps in classifying the type of teeth and also some underlying oral anomalies present in an OPG using CNN with deep learning technique to assist the dentists to have a better judgment in the treatment of any dental abnormalities. The fully connected layer of the CNN which does the classification has been trained using hybrid GA-BP learning. Dropout technique has been used to avoid over fitting thereby improving the performance of the network classification. The proposed algorithm is implemented with different convolutional layers and the model with two convolutional layers gives the highest accuracy of **97.92**%.

Acknowledgement

We are immensely thankful to NIT Manipur for providing a platform to do the research work and Cosmo Dental Clinic Manipur for providing the dataset.

References

[1] A. Qayyum, S.M. Anwar, M. Majid, M. Awais, and M.R. Alnowami, "Medical Image Analysis using Convolutional Neural Networks: A Review," *Journal of Medical Systems*, vol. 42, pp. 1-13, 2018, doi: 10.1007/s10916-018-1088-1.

[2] J. Prada, Y. Gala, and A. L. Sierra, "COVID-19 Mortality Risk Prediction Using X-Ray Images," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 6, pp.7-14, 2021, doi:10.9781/ijimai.2021.04.001.

[3] M. I. Khattak, M.'ath. Al-Hasan, A. Jan, N. Saleem, E. Verdú, and N. Khurshid, "Automated Detection of COVID-19 using Chest X-Ray Images and CT Scans through Multilayer- Spatial Convolutional Neural Networks," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 6, pp. 15-24, 2021, doi: 10.9781/ijimai.2021.04.002.

[4] S. Kaliyugarasan, A. Lundervold, and A. S. Lundervold, "Pulmonary Nodule Classification in Lung Cancer from 3D Thoracic CT Scans Using fastai and MONAI," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 7, pp. 83-89, 2021, doi:10.9781/ijimai.2021.05.002.

[5] B. S. Harish, M. S. Maheshan, and N. Nagadarshan, "A Convolution Neural Network Engine for Sclera Recognition," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 1, pp. 78-83, 2020, doi: 10.9781/ijimai.2019.03.006.

[6] J. J. Hwang, Y. H. Jung, B. H. Cho, and M. S. Heo, "An overview of deep learning in the field of dentistry," *Imaging science in dentistry*, vol. 49, no. 1, pp. 1–7, 2019, doi:10.5624/isd.2019.49.1.1.

[7] M. Mupparapu, C. W. Wu, and Y. C. Chen, "Artificial intelligence, machine learning, neural networks, and deep learning: Futuristic concepts for new dental diagnosis," *Quintessence international*, vol. 49, no. 9, pp. 687–688, 2018, doi:10.3290/j.qi.a41107.

TABLE VI. Comparison of Proposed Method With Other Standard Methods

Researchers	Year	Methods	Types of Radiography	Teeth Images class	Classification Accuracy
Kuo[31]	2017	CNN	Panoramic radiograph	Dental Anomalies	85.3%
Krois[32]	2019	Deep Convolutional Neural Network(DCNN)	Panoramic radiograph	Dental Anomalies	Classification accuracy was 0.81(0.02). Mean(SD) sensitivity and specificity was 0.81 and 0.81
Poedjiastoeti [33]	2018	CNN	Panoramic radiograph	Dental Anomalies	Sensitivity, specificity, accuracy were 81.8%, 83.3%, 83.0%.
Singh[34]	2020	CNN	Panoramic radiograph	Dental types	95.6%
Our proposed Method	2020	CNN with Dropout	Orthopantomogram Images (panoramic radiograph)	Dental types and Anomalies	**97.92**%

[8] S.B. Khanagar et al., "Developments, application, and performance of artificial intelligence in dentistry-A systematic review," *Journal of Dental Sciences*, vol. 16, pp. 508-522, 2021, doi: 10.1016/j.jds.2020.06.019.

[9] J. Yang, Y. Xie, L. Liu, B. Xia, Z. Cao, and C. Guo, "Automated Dental Image Analysis by Deep Learning on Small Dataset," *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC),* Tokyo, 2018, vol.1, pp. 492–497, doi:10.1109/COMPSAC.2018.00076.

[10] N. Karimian, H. S. Salehi, M. Mahdian, H. Alnajjar, and A. Tadinada, "Deep learning classifier with optical coherence tomography images for early dental caries detection," *Proc. SPIE 10473, Lasers in Dentistry XXIV,* 1047304, pp. 10-17, 2018, doi:10.1117/12.2291088.

[11] K.L. Devito, F.D. Barbosa, and W.N. Filho, "An artificial multilayer perceptron neural network for diagnosis of proximal dental caries," *Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics*, vol. 106, pp. 879–884, 2008, doi:10.1016/j.tripleo.2008.03.002

[12] R.B. Ali, R. Ejbali, and M. Zaied, "Detection and classification of dental caries in X-ray images using deep neural networks," *In: International Conference on Software Engineering Advances (ICSEA).* 2016, pp. 223-227.

[13] 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 *2008 31st International Spring Seminar on Electronics Technology*, 2008, pp.125-129, doi:10.1109/ISSE.2008.5276424.

[14] S. A. Prajapati, R. Nagaraj, and S. Mitra, "Classification of dental diseases using CNN and transfer learning," In *5th International Symposium on Computational and Business Intelligence (ISCBI),* Dubai, 2017, pp. 70-74, doi:10.1109/ISCBI.2017.8053547.

[15] Y. Yu, "Machine Learning for Dental Image Analysis," *ArXiv*, vol. abs/1611.09958, 2016, https://arxiv.org/abs/1611.09958.

[16] M. M. Srivastava, P. Kumar, L. Pradhan, and S. Varadarajan, "Detection of tooth caries in bitewing radiographs using deep learning," *ArXiv,* vol. abs/1711.07312, 2017, https://arxiv.org/abs/1711.07312 (accessed 16 May 2019).

[17] D. K. Veena, A. Jatti, R. Joshi, and K. S. Deepu, "Characterization of dental pathologies using digital panoramic X-ray images based on texture analysis," In *39th Annual International Conference of IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 592-595, doi:10.1109/EMBC.2017.8036894.

[18] A. Zakirov, M. Ezhov, M. Gusarev, V. Alexandrovsky, and E. Shumilov, "Dental pathology detection in 3D cone-beam CT," *ArXiv,* vol. abs/1810.10309, 2018.

[19] R.C. Gonzales, and R. E. Woods, "Digital Image Processing," *Prentice Hall, Inc, New Jersey*, 2008, ISBN: 9780131687288 013168728X 9780135052679 013505267X.

[20] Y.M. Abdallah, and T. Alqahtani, "Research in Medical Imaging Using Image Processing Techniques," *Medical Imaging - Principles and Applications,* IntechOpen, 2019, doi:10.5772/intechopen.84360.

[21] A. S. Unde, V. A. Premprakash, and P. Sankaran, "A novel edge detection approach on active contour for tumor segmentation," In *2012 Students Conference on Engineering and Systems,* Allahabad, Uttar Pradesh, 2012, pp. 1-6, doi:10.1109/SCES.2012.6199100.

[22] J. Wu, "Introduction to Convolutional Neural Networks," *National Key Lab for Novel Software Technology*, Nanjing University, China, 2017.

[23] T. Wiatowski, and H. Bolcskei, "A mathematical theory of deep convolutional neural networks for feature extraction," in *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1845-1866, March 2018, doi:10.1109/TIT.2017.2776228.

[24] A. Suruliandi, A. Kasthuri, and S. P. Raja, "Deep Feature Representation and Similarity Matrix based Noise Label Refinement Method for Efficient Face Annotation," *International Journal of Interactive Multimedia and Artificial Intelligence*, In Press (In Press), 1-12, http://doi.org/10.9781/ijimai.2021.05.001.

[25] Y. Li, J. Deng, Q. Wu, and Y. Wang, " Eye-Tracking Signals Based Affective Classification Employing Deep Gradient Convolutional Neural Networks," *International Journal of Interactive Multimedia And Artificial Intelligence*, In Press (In Press), pp. 1-10, doi: 10.9781/ijimai.2021.06.002.

[26] K. K. Verma, B. M. Singh, H. L. Mandoria, and P. Chauhan, "Two-Stage Human Activity Recognition Using 2D-ConvNet," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 2, pp. 125-135, 2020, doi: 10.9781/ijimai.2020.04.002.

[27] X. Jin, Y. Xiao, S. Li, and S. Wang, "Deep Learning-based Side Channel Attack on HMAC SM3," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 4, pp. 113-120, 2020, doi: 10.9781/ijimai.2020.11.007.

[28] A. Ghazvini, S. N. H. S. Abdullah, and M. Ayob "A Recent Trend in Individual Counting Approach Using Deep Network," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5, no. 5, pp. 7-14, 2019, doi:10.9781/ijimai.2019.04.003.

[29] M.D. Bloiceet, C. Stocker, and A. Holzinger, "Augmentor: An Image Augmentation Library for Machine Learning," *The Journal of Open Source Software*, vol. 2, no. 19, 2017, doi:10.21105/joss.00432.

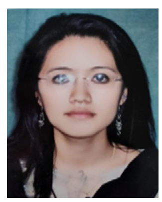
[30] J. Wang and L. Perez, "The effectiveness of data augmentation in image classification using deep learning," *Convolutional Neural Networks Vis. Recognit*, vol. 11, pp.1–8, 2017.

[31] Y.F. Kuo, S.Y. Lin, C.H. Wu, S.L. Chen, T. Lin, N.H. Lin, C.H. Mai, and J.F. Villaverde, "A Convolutional Neural Network Approach for Dental Panoramic Radiographs Classification," *Journal of Medical Imaging and Health Informatics*, vol. 7, pp. 1693-1704, 2017, doi:10.1166/JMIHI.2017.2257.

[32] J. Krois, T. Ekert, L. Meinhold, T. Golla, B. kharbot , A. Wittemeier, C. Doofer, and F. Schwendicke, "Deep Learning for the Radiographic Detection of Periodontal Bone Loss," *Scientific Reports*, vol. 9, 8495, 2019, doi: 10.1038/s41598-019-44839-3.

[33] W. Poedjiastoeti, and S. Suebnukarn, "Application of Convolutional Neural Network in the Diagnosis of Jaw Tumors," *Healthcare Informatics Research,* vol. 24, pp. 236-241, 2018, doi: 10.4258/hir.2018.24.3.236.

[34] P. Singh, and P. Sehgal, "Numbering and Classification of Panoramic Dental Images Using 6-Layer Convolutional Neural Network," *Pattern Recognition and Image Analysis*, vol.30, pp. 125–133, 2020, doi: 10.1134/S1054661820010149.



Anuradha Laishram

Anuradha Laishram received her B.Tech from Visvesvaraya Technological University, India in 2008 and M.E in Computer science and Engineering from pursuing her Ph.D degree from National Institute of Technology Manipur (NIT Manipur), India. She is presently working as Lecturer in NIT Manipur, India. Her main research interest includes Machine Learning, Deep Learning, Hybrid Intelligent System, and Medical Image Processing. Email: annu0286@gmail.com

Khelchandra Thonga

Khelchandra Thongam received his M.S degree in Computer science and Engineering from The University of Aizu, Japan in 2007 and his Ph.D also from The University of Aizu, Japan. He is currently working as Assistant Professor in National Institute of Technology Manipur (NIT Manipur), India. His main research interest includes Machine Learning, Soft Computing, Hybrid Intelligent System, Medical Image Processing, Speech Processing, and Mobile Robot Navigation. Email: thongam@gmail.com