International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Convolutional Neural Network Based Image Classification

Archana Kumari*

M.Tech [Industrial Electronics], Mohan Garden, Uttam Nagar , New Delhi - 110059, India
Email: archanak.ie20@nsut.ac.in

Abstract

This paper presents the study of Convolutional Neural Network based Image Classification. In this study, different convolutional neural network model is implemented for image classification. There are two application which is performed in this namely, emotion detection from facial emotion recognition system and covid-19 detection from lung CT-scan. In this research work, AlexNet and ResNet-50 convolutional neural network models are used for comparison and evaluation is done based on its training accuracy and testing accuracy, confusion matrix and Area under the curve (AUC) of ROC graph. All the experiments are performed on MATLAB software. In the First application of CNN model is implemented to detect emotions from facial expressions and SVM classifier is used for classification of each emotion among eight facial expressions namely, surprise, contempt, happiness, sadness, fear, anger, disgust, and neutral. The following proposed work is carried out on CK+ datasets to determine the recognition. After performing on both CNN models accuracy of both the models are compared. The result shows that ResNet-50 model achieved the best accuracy of 97.32% which is better than 90.55% which got on AlexNet model. In Second application of CNN model is to detect covid-19 from lung CT-scans. The highest performing model, the ResNet-50 on a SARS-CoV-2 CT-scan dataset achieved an accuracy of 95.72% which is maximum than 85.50% which achieved on AlexNet Model. Gradient-Weighted Class Activation Mapping (Grad-CAM) is also used to display infected area in the lungs. After performing this experiment, the final results shows that the ResNet-50 model, performs much better as compared to AlexNet model.

Keywords: CK+; FER, Emotion Recognition; CNN models; AlexNet; Res-Net-50; SARS-CoV-2 dataset; Grad-CAM; SVM; AUC; ROC.

----

* Corresponding author.

1. Introduction

Emotions plays a significant part in our day-to-say life and will continue to do. It is an effective way of communication and the basis of mutual understanding, unity and cooperation between people. The use of PC frameworks, programming and organizations are developing quickly and generally. These frameworks have a significant impact in our day-to-day existence and make human existence a lot simpler. FER system assumes lots of importance during this time. The traditional technique has less accuracy as compared to deep learning methods. Deep learning methods give more accurate results as compared to traditional method. This project is aimed to implement different CNN model and compare their results to design Facial Emotion Recognition (FER) system to detect emotions from facial expression and to detect COVID-19 from facial expressions. In the first application, CNN is implemented for correct classification of eight facial expressions namely, surprise, contempt, happiness, sadness, fear, anger, disgust, and neutral. Whereas in second application, it is implemented to detect COVID-19 from CT-scans of lungs and Gradient-Weighted Class Activation Mapping (Grad-CAM) is used to display infected area in the lungs. The CNN module was built in MATLAB to design FER system and covid-19 detection method, but training module is performed on the Graphic Processing Unit (GPU) computational power to accelerate the process.

The COVID-19 epidemic has affected life of millions of people world widely. To date, a few million people have gone to the valley of death and many more are affected by it every day. Hundreds of studies have not been found to contain COVID-19 at a young age due to a lack of testing equipment. When these early undiagnosed studies with no symptoms come into contact with healthy people, they transmit the virus to healthy people and the series continues. The current method of testing COVID 19 is to use a reverse transcription-polymerase chain reaction (RT-PCR). In this method, sample of a patient is collected through a person's nose or throat and sent to a laboratory for diagnostic purposes. This is time consuming process and because of its high demand, there is a lack of these kits. Globally, researchers are trying to find other ways to diagnose corona-virus in infected patients. X-ray and CT-scan emerged as possible options to the diagnosis of COVID-19. In the 1st application, the following work consists of a multiple stage of image processing to extract feature representations and it is done in three phases: face detection, feature extraction and feature classification steps for predicting the facial expression label. The design of FER system is carried out on CK+ datasets to determine the emotions recognition. In this, FER system uses sequence of technique for detection of face, extraction of feature and classification are discussed. Initially, the dataset is acquired and image pre-processing is performed on dataset. All the images in datasets are of different size so, image batch processing is performed on entire dataset and all images are resized by 48x48. Next, face detection step is performed and it is done by using viola jones algorithm. Viola jones algorithm is one of the most popular techniques for object detection. After pre-processing dataset is split into training dataset and testing dataset and 80% of the data are used for training and left 20% for the testing set and splitting is done randomly to avoid biasing. Following this, image data is trained by using pre-trained CNN model. This CNN models are used for feature extraction. Next, Feature classification step is carried out by using Support Vector Machine classifier is used for classification of features namely, anger, contempt, disgust, fear, happy, neutral, sadness and surprise.

In 2nd application same initially, dataset is acquired and all images were resized to 256-by-256.Then, Dataset is split into three set i.e., Training dataset, validation dataset and testing dataset in the ratio of 60:20:20 respectively. After that, Data Augmentation is done to create dummy data from the original dataset. In this method, Image segmentation and feature extraction are automatically done by convolutional bases of CNN pre-trained models, while classification step is carried out by using a fully connected neural network. In this, activation function is ReLU layer and a Softmax output layer. For both the application Architecture of AlexNet and ResNet-50 are used to train the network with learning rate of 0.0001 and sgdm optimizer for AlexNet and adam optimizer for ResNet-50 model. This decrease in loss produces a small number of misclassifications. AlexNet model is trained over 50 epochs, and ResNet-50 over 100 epochs and learning rate is 0.0001 consider. Cross-entropy loss function is used for training progress and this Loss is ideally reduced by updating the weights. and then sgdm and adam optimizers are used for AlexNet and ResNet-50 model respectively. The motive of this study is to understand different pre-trained convolutional neural network models. Study the hyper parameters, properties, and methods for training CNN. Learn about Grad-CAM means how it can be used to extract features from input images. The objective of the project is to investigate and understand different pre-trained convolutional neural network models and tuning of model hyper parameters, their properties, and techniques for training CNN are done. There are two applications which are used in this studies :

- To design, implement, analyze FER system that classify correctly emotions from facial expressions by using a pre-trained CNN model.
- To design, implement, test, and analyse models to detect COVID-19 and classifies correctly between COVID-19 and NORMAL CT-scans by using a pre-trained CNN and Grad-CAM is implemented to show infected area in the lungs.

2. Background and Related Works

There are many studies for pattern recognition technique to recognize and classify human emotion. Pranav E. and his colleagues [1] presented a Facial Emotion Recognition that classifies 5 different human facial emotions Using a Deep Neural Network. For this model, Adam optimizer is use to minimize the loss function and it is tested to possess an accuracy of 78.04%. Sai Yeshwanth Chaganti and his colleagues [2] proposed an Image Classification using SVM and CNN and it is concluded that when SVM is implemented on a very small dataset it achieves an accuracy of 93% whereas SVM is a very strong technique to achieve maximum accuracy. So, by using data augmentation, the dataset size is tripled and on performing again SVM classifier, accuracy of 82% is achieved, a significant decrease. Not satisfied with the results, they implement CNN, and got an accuracy of a 93.57% on the same dataset. Balaji Balasubramanian and his colleagues [3] Analysis of Facial Emotion Recognition on different datasets and on different techniques that are used for the task of FER, and the conclusion is that CNN performs better as compared to Support Vector Machine (SVM). The structure is to perform face detection and train CNN to its output to performs well. G. Cao and his colleagues [4] used a CNN model to recognize and classify human emotion and brain signals respectively from ECG dataset. The system gives approx. 83% accuracy on testing. Recently, many researches were conducted on detecting COVID-19 using Machine learning, artificial intelligence techniques. In [5] D. F. Eljamassi proposes COVID-19 Detection using Machine Learning to detect the infected condition. Dataset contains chest x-ray images of COVID-19 patients, people with pneumonia and normal people. For feature extraction Histogram of oriented gradients

(HOG) and for classification different classifier namely, SVM, KNN and random forests are used. The classification accuracy obtained as 98.14%, 88.89% and 96.29% are achieved on SVM, KNN and random forests respectively. A. P. Hartono and his colleagues [6] propose and evaluate the performance CNN and Transfer Learning model to detect COVID-19 infections from lungs. Transfer-learning using pre-trained model resulted in a detection accuracy of 89% while proposed CNN show the classification accuracy of 97%. A. Q. A. Ameer and his colleagues [7] proposes covid-19 detection using CT scan based on gray level Co-Occurrence matrix (GLCM). Initially, image pre-processing step is done in which increase contrast in images, then lung is removed by labelling the most contrast connected pixels and lastly subtract labelling pixels from original image. Then removal of noise is performed by applying filters namely, mean, median and Gaussian, and then GLCM was applying in four directions (0^o,45^o,90^o and 135^o), then extract features from GLCM. They got an accuracy of 94% for detect the location of infection.

3. Materials and Methods

In this section, different CNN model is studied and further analysed based on its training and testing accuracy, performance metrices, and AUC of ROC graph.

Then there are two applications which are studied in this thesis:

1. To design, implement, analyze FER system that classify correctly emotions from facial expressions by using a pre-trained CNN model.
2. To design, implement, test, and analyse models to detect COVID-19 and classifies correctly between COVID-19 and NORMAL CT-scans by using a pre-trained CNN and Grad-CAM is implemented to show infected area in the lungs.

3.1. Dataset

For FER system CK+ dataset [9] and for covid-19 detection, SARS-CoV-2 CT scan dataset [10] are used.

CK+ Dataset: The CK+ dataset [9] contains 1534 images of 210 adults within the age of 18 to 50. In this dataset, images are of 640x490, 8 bit Gray-scale. The images in dataset are labelled with classes of emotions such as Anger, Contempt, Disgust, Neutral, Sadness, Happy, Fear, and Surprise. Ck+ Dataset image sample is illustrated in Fig.1. Table 1 represents No. of images in each class in CK+ dataset.

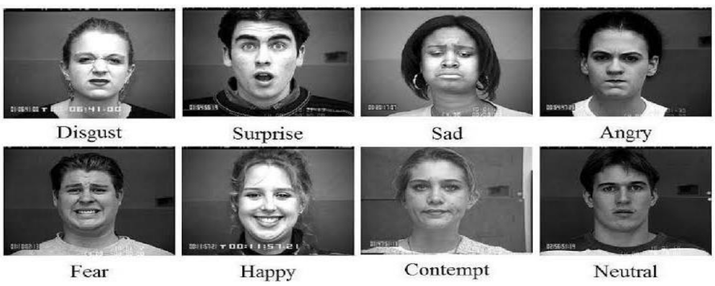


Figure 1: CK+ dataset image sample.

Table 1: List of No. of images in each class in CK+ dataset.

Label	No. of Images
Anger	135
Contempt	54
Disgust	177
Fear	75
Happy	207
Neutral	553
sadness	84
surprise	249

Sars-Cov-2 CT Scan Dataset [10] used in this study. It consists of 702 images and it is split into two classes, 328 COVID-19 and 374 Normal images. The images in the dataset having different dimensions and there are no clinical labelling’s found in images which usually found in other datasets. Some of the Samples from CT scan dataset is shown in Fig. 2. Table 2 represent No. of images in each class in this dataset.

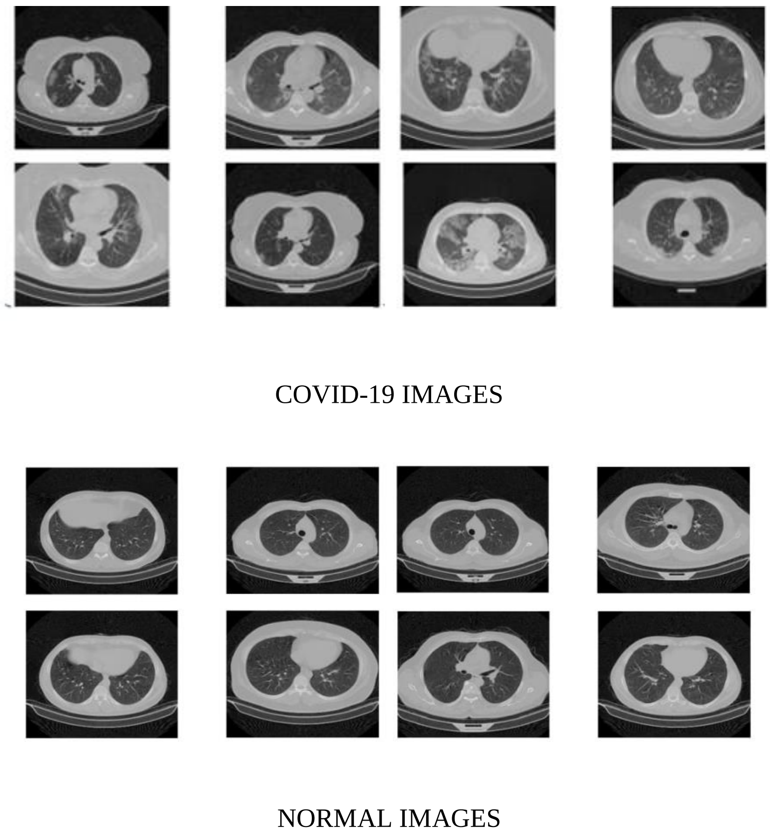


Figure 2: samples from SARS-CoV-2 CT scan dataset.

Table 3: List of No. of images in CT- scan dataset.

Label	No. of Images
COVID-19	328
NORMAL	374

3.2. Proposed Methodology

The proposed method of the FER system and covid detection is presented in Fig.3 and Fig.4 respectively.

In this section, it gives the proposed methodology which is used in this work is arranged in following three steps:

1. Face Detection step
2. Feature extraction step
3. Feature classification step

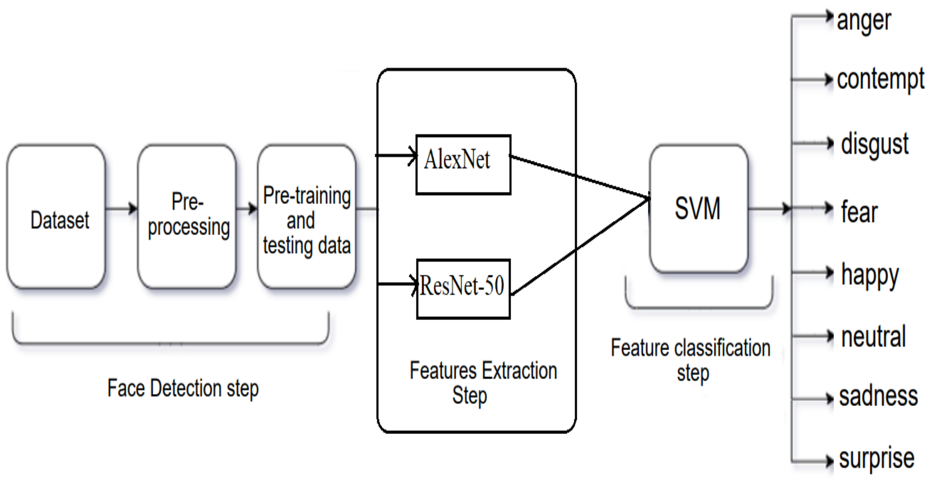


Figure 3: flowchart of the proposed methodology of FER system

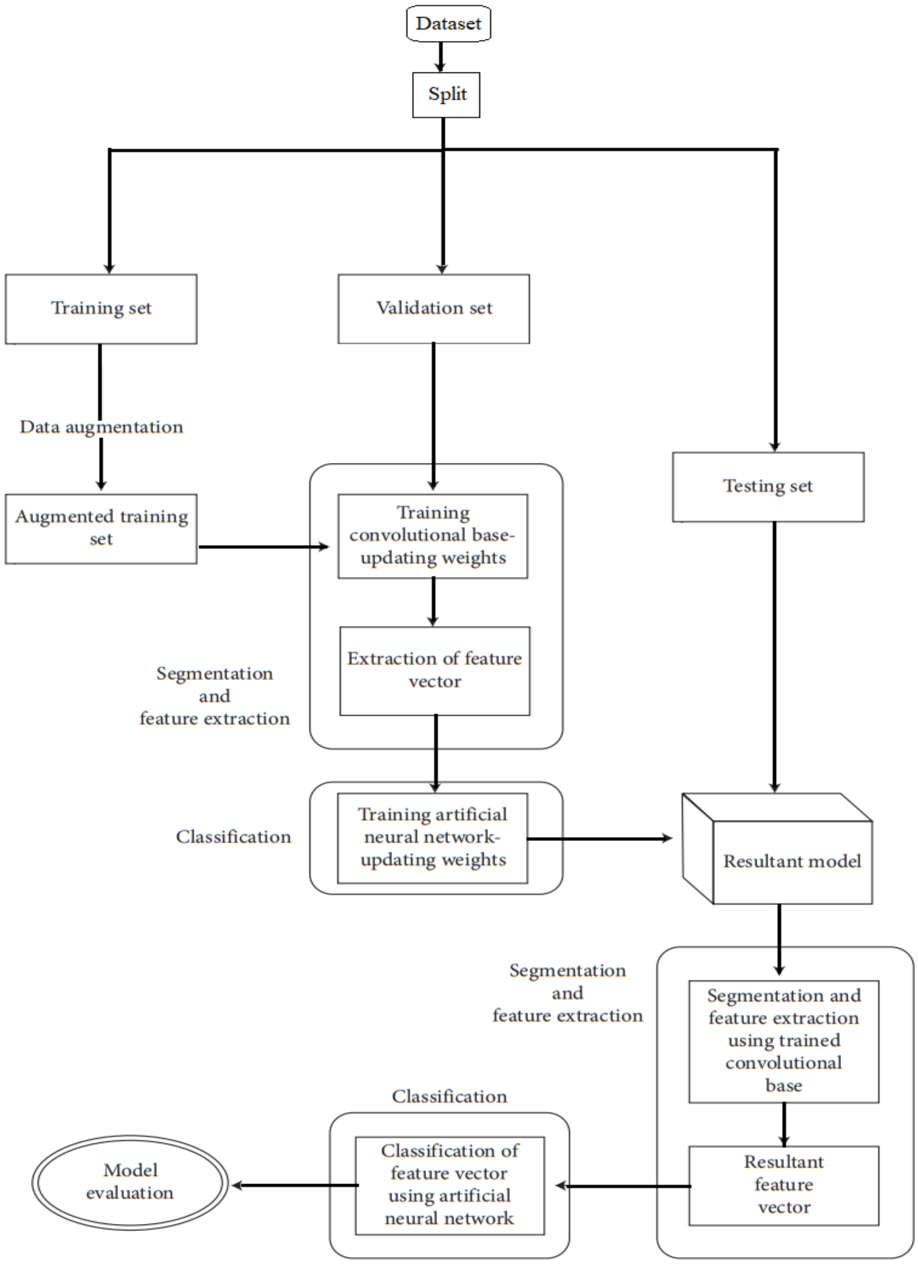


Figure 4: proposed methodology of COVID-19 detection.

3.3. Pre-processing

In this stage, face detection and data modification tasks are carried out for CK+ dataset and for SARS-CoV-2 CT scan dataset only modification is needed.

3.3.1. Face detection and data modification for CK+ dataset

In this step, Viola jones algorithm framework is used in face detection. It is most popular techniques for object detection. This technique uses cascade features to extract faces from the datasets. This approach used Haar like features of machine learning model to extract features from input image. The outcome of this is that it is reduced by "over one half the number of locations where the final detector must be evaluated". After this step, the images were resized to 48x48 by image batch processor in MATLAB and converted to grayscale images.

3.3.2. Data modification for lungs CT scan dataset

In this step, all the images in dataset were resized to 256x256 by image batch processor in MATLAB and converted to grayscale images because all the images in datasets are of different sizes.

3.3.3. Prepare Training Image Sets and Testing Image Sets

After performing the pre-processing stage, the CK+ dataset is splitting into training and testing Image set and it should be done random-y to avoid biasing. Datasets are split in the ratio of 80:20 for training and testing data. Table 3 represents Splitting of CK+ dataset.

Table 3: Splitting of CK+ dataset.

No. of class	8
No. of training images	1227
No. of testing images	306

SARS-CoV-2 CT scan dataset is split into 60:20:20 for training set, validation set and testing set respectively. Splitting of dataset is done randomly to avoid biasing. Table 4 represents Splitting of CT scan dataset.

Table 4: Splitting of CT scan dataset.

No. of class	2
No. of training images	422
No. of validation images	140
No. of testing images	140

3.3.4. Training and Testing stage

In this step, it focuses on training the models by changing the hyper parameters values, and then apply data augmentation techniques, and further studying the different pre-trained CNN models.

3.4. CNN Model

In this section, AlexNet and ResNet-50 CNN Models are implemented for feature extraction, and the hyper parameters is need to be adjusted to achieve most accurate performance.

AlexNet: AlexNet is a pre-trained deep learning model. It is the arrangement of 15 layers which consist of input image, 1 Convolution, 3 Rectified Linear Unit (ReLU), 1 Cross Channel Normalization, 2 Max Pooling, 2 Dropout, 3 Fully connected, 1 Softmax, 1 Classification Output layer. In this, ReLU layer comes after the Convolutional Layer. ReLU layer add non-linearity in the network and Max-Pooling Layer is add to decrease the number of variables and helps in managing overfitting. The fully connected layers are followed by dropout layers. Dropout is a technique used to decreases overfitting of neurons. Fig. 5 shows the architecture of AlexNet Model.

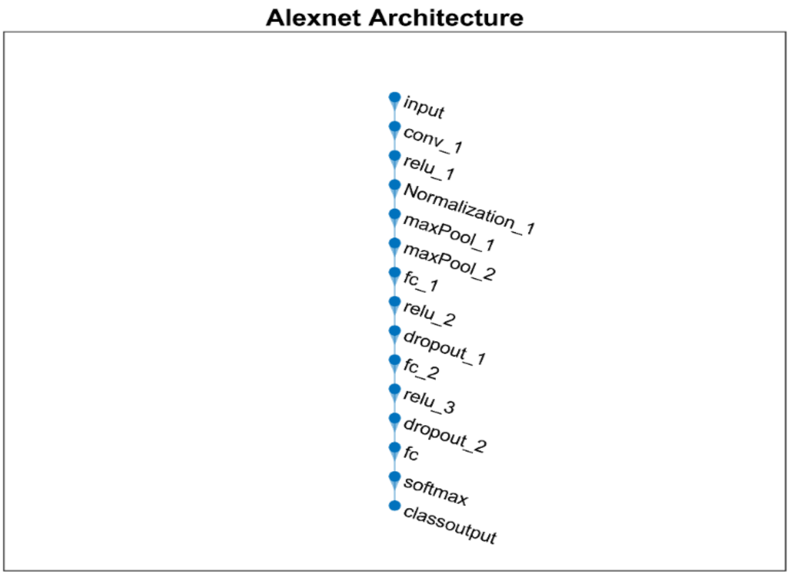


Figure 5: AlexNet Model Architecture.

ResNet-50: It is a pre-trained convolutional neural network model which is trained on millions of images in Image-Net database. The ResNet-50 architecture is the arrangement of 50 layers which consist of input image, 13 Convolution ,13 Batch-Normalization,6 Addition,13 Rectified Linear Unit (ReLU), 1 max- pooling,1 FC, 1 Softmax, 1 Classification Output layers. Our motive is to analyze the performance of the ResNet-50 model and compare it with AlexNet. Fig. 6 shows the ResNet-50 Model architecture.

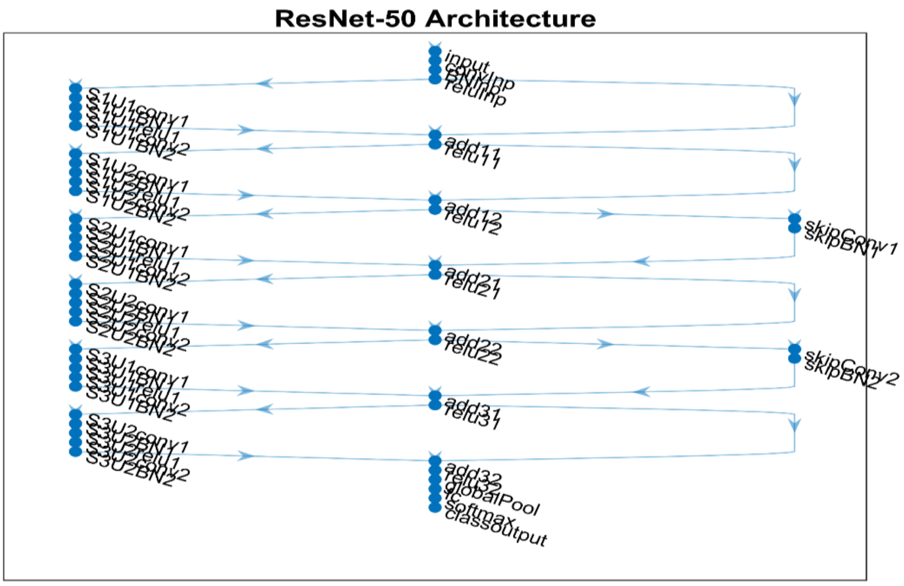


Figure 6: ResNet-50 Model Architecture

Table 5: Layers of AlexNet Architecture.

S.No.	INPUT LAYERS		PARAMETERS
1	'input'	Image input	Input size of images with 'zerocenter' normalization
2	'conv_1'	Convolution	96 11×11 convolutions with stride [4 4] and padding [0 0 0 0]
3	'relu_1'	ReLU	ReLU
4	'Normalization_1'	Cross Channel Normalization	Cross channel normalization with 5 channels per element
5	'maxPool_1'	Max Pooling	3×3 max pooling with stride [2 2] and padding [0 0 0 0]
6	'maxPool_2'	Max Pooling	3×3 max pooling with stride [2 2] and padding [0 0 0 0]
7	'fc_1'	Fully Connected	1024 fully connected layer
8	'relu_2'	ReLU	ReLU
9	'dropout_1'	Dropout	50% dropout
10	'fc_2'	Fully Connected	1024 fully connected layer
11	'relu_3'	ReLU	ReLU
12	'dropout_2'	Dropout	50% dropout
13	'fc'	Fully Connected	Number of fully connected layer
14	'softmax'	Softmax	Softmax
15	'classoutput'	Classification output	crossentropyex

Table 6: Layers of ResNet-50 Architecture.

S.No.	INPUT LAYERS		PARAMETERS
1.	'input'	Image Input	Input size of images with 'zerocenter' normalization
2.	'convInp'	Convolution	4 3×3 convolutions with stride [1 1] and padding 'same'
3.	'BNInp'	Batch Normalization	Batch Normalization
4.	'reluInp'	ReLU	ReLU
5.	'S1U1conv1'	Convolution	4 3×3 convolutions with stride [1 1] and padding 'same'
6.	'S1U1BN1'	Batch Normalization	Batch Normalization
7.	'S1U1relu1'	ReLU	ReLU
8.	'S1U1conv2'	Convolution	4 3×3 convolutions with stride [1 1] and padding 'same'
9.	'S1U1BN2'	Batch Normalization	Batch Normalization
10.	'add11'	Addition	Element-wise addition of 2 inputs
11.	'relu11'	ReLU	ReLU
12.	'S1U2conv1'	Convolution	8 3×3 convolutions with stride [2 2] and padding 'same'
13.	'S1U2BN1'	Batch Normalization	Batch Normalization
14.	'S1U2relu1'	ReLU	ReLU
15.	'S1U2conv2'	Convolution	8 3×3 convolutions with stride [2 2] and padding 'same'
16.	'S1U2BN2'	Batch Normalization	Batch Normalization
17.	'add12'	Addition	Element-wise addition of 2 inputs
18.	'relu12'	ReLU	ReLU
19.	'S2U1conv1'	Convolution	8 3×3 convolutions with stride [2 2] and padding 'same'
20.	'S2U1BN1'	Batch Normalization	Batch Normalization
21.	'S2U1relu1'	ReLU	ReLU
22.	'S2U1conv2'	Convolution	8 3×3 convolutions with stride [1 1] and padding 'same'
23.	'S2U1BN2'	Batch Normalization	Batch Normalization
24.	'add21'	Addition	Element-wise addition of 2 inputs
25.	'relu21'	ReLU	ReLU
26.	'S2U2conv1'	Convolution	8 3×3 convolutions with stride [1 1] and padding 'same'
27.	'S2U2BN1'	Batch Normalization	Batch Normalization
28.	'S2U2relu1'	ReLU	ReLU
29.	'S2U2conv2'	Convolution	8 3×3 convolutions with stride [1 1] and padding 'same'
30.	'S2U2BN2'	Batch Normalization	Batch Normalization
31.	'add22'	Addition	Element-wise addition of 2 inputs
32.	'relu22'	ReLU	ReLU
33.	'S3U1conv1'	Convolution	16 3×3 convolutions with stride [1 1] and padding 'same'
34.	'S3U1BN1'	Batch Normalization	Batch Normalization
35.	'S3U1relu1'	ReLU	ReLU
36.	'S3U1conv2'	Convolution	16 3×3 convolutions with stride [1 1] and padding 'same'
37.	'S3U1BN2'	Batch Normalization	Batch Normalization
38.	'add31'	Addition	Element-wise addition of 2 inputs
39.	'relu31'	ReLU	ReLU
40.	'S3U2conv1'	Convolution	16 3×3 convolutions with stride [1 1] and padding 'same'
41.	'S3U2BN1'	Batch Normalization	Batch Normalization
42.	'S3U2relu1'	ReLU	ReLU
43.	'S3U2conv2'	Convolution	16 3×3 convolutions with stride [1 1] and padding 'same'
44.	'S3U2BN2'	Batch Normalization	Batch normalization
45.	'add32'	Addition	Element-wise addition of 2 inputs
46.	'relu32'	ReLU	ReLU
47.	'globalPool'	Average Pooling	8×8 average pooling with stride [1 1] and padding [0 0 0 0]
48.	'fc'	Fully Connected	Number of fully connected layer
49.	'softmax'	Softmax	softmax
50.	'classoutput'	Classification Output	crossentropyex

3.5. Hyper Parameter Tuning

To Train the model hyper parameter tuning is an important process. In this process, model hyper parameters values are adjusting according to the requirement to attain good performance. The following model parameters are tuned:

- **Learning rate:** It control the step size and no. of iterations in the model. Its values range from 0.0001 to 1. However, a default value of learning rate is 0.01.
- **Optimizer:** Optimizer are techniques used to change the features of neural networks to reduce losses like weight and learning rate. sgdm and adam optimizers are used here.

  - **Sgdm:** sgdm optimizer update the parameters of the model. When loss is occurred during training, the model parameters are changed Eqn. (4.1) represents the mathematical expression of the sgdm optimizer.

$$\theta = \theta - \alpha \cdot \nabla J(\theta; x(i); y(i)) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.(1)$$

where, x(i) and y(i) are training examples.

  - **Adam:** adam stands for Adaptive Moment Estimation. This optimizer works with 1st and 2nd order momentums. It has an exponentially decaying average of previous gradients. To update the parameter eq. (4.2) is used.
    - $\widehat{m}_t = \frac{m_t}{1-\beta_1^t}$ ……………………………………………………………….(2)
    - $\hat{v}_t = \frac{v_t}{1-\beta_2^t}$………………………………………………………………........(3)
    - $\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t+\varepsilon}}.\widehat{m}_t$…………………………………….....………..(4)

- **epochs:** It controls the no. of iterations for the training progress. By the use of early stopping, number of epochs is not required to set.
- **Activation function:** ReLU activation functions are preferred to control the neurons.

3.6. SVM Classifier

SVM is a Support Vector Machine, supervised learning technique which conclude a relation from the training dataset. In this technique learns by analysing data and recognising patterns, and implemented for feature extraction and classification of pattern regression. And these features are being used for training and testing SVM classifier.

3.7. Data Augmentation

is a technique used to create “dummy data” from the original data. This study performs data augmentation to produce images in which a rotational change in the range of [−20° to 20°] to the original dataset images.

3.8. Performance Metrics

The performances are evaluated by calculating performance metrics of confusion matrix. The following parameters are calculated like, accuracy, precision, specificity, sensitivity, and AUC of ROC.

These are the following formula which is used to evaluate accuracy, precision, specificity, sensitivity.

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

$$Sensitivity = \frac{TP}{(TP+FN)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

$$Specificity = \frac{TN}{(FP+TN)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

$$Precision = \frac{TP}{(TP+FP)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8)$$

3.9. Receiver Operating Characteristics Curve

ROC curve is a graphical representation of showing the performance of the classification model. It is plot between True Positive Rate (TPR) and False Positive Rate (FPR). In ROC graph, TPR is shown against FPR. An Area under the curve value is calculated from the graph and it should be in the range of 0.5 and 1. Fig. 7 represents ROC curve.

True Positive Rate is defined as,

$$TPR = \frac{TP}{TP+FN} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.9)$$

False Positive Rate is defined as,

$$FPR = \frac{FP}{FP+TN} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.10)$$

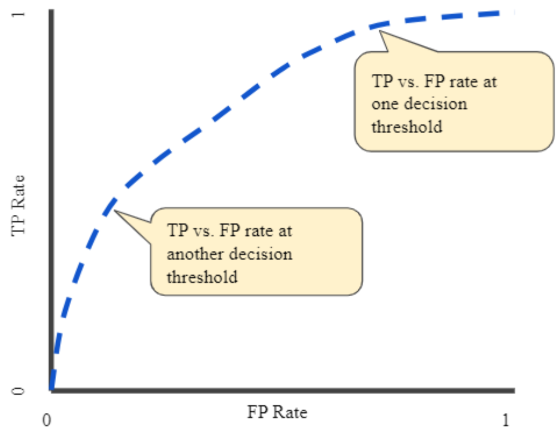


Figure 7: ROC Curve.

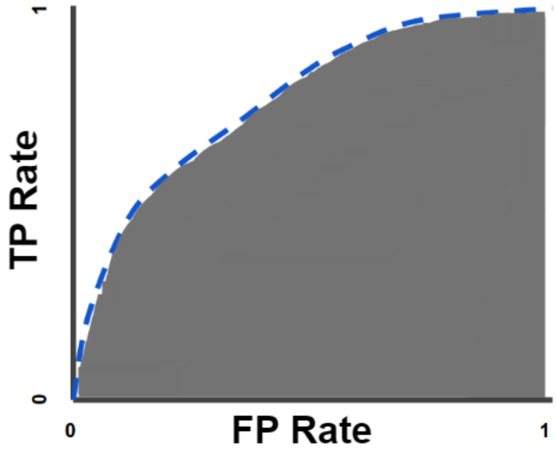


Figure 8: AUC.

3.10. Gradient-Weighted Class Activation Mapping

Grad-CAM is a classification decisions technique function. It uses the gradient of the classification result w.r.t to the convolutional features which is determined by the network to understand which region of the image is important for the purpose of classification. It can be applied for the purpose of non- classification. For example, regression or semantic segmentation. The places where this gradient is large are exactly the places where the final score depends most on the data. It computes the map by taking the derivative of the reduction layer results of the given class w.r.t to a convolutional feature map.

4. Experimental Results and Discussions

Here, the results of the implemented model is discussed. In this experiment, AlexNet and ResNet-50 CNN pre-trained models are used and their analysis are done by tuning their parameters.

In this experiment, there are two applications of CNN mode:

1. Emotion detection and their classifications from Facial Expressions.
2. COVID-19 detection from Lungs CT-scans.

Initially, dataset is acquired and all images were resized to 48-by-48 in CK+ datasets and 256-by-256 in CT scan dataset. Then Dataset is split into three set i.e., Training dataset, validation dataset and testing dataset in the ratio of 60:20:20 respectively. Next, Data Augmentation is done to create dummy data from the original dataset. Both the models are trained according to parameters which is tabulated in Table 7 and Table 10.

4.1. Emotion Detection And Their Classifications From Facial Expressions

In this section, firstly face detection is done by Viola-Jones algorithm.

Figure 9: Face Detection by Viola Jones algorithm.

After face detection, all the images are resized to 48-by-48 by batch image processor and converted into gray-scale in MATLAB. Following this, input images are trained on CNN models (AlexNet and ResNet-50) and the hyperparameter tuning of models are tabulate in Table 7.

Table 7: Model Parameters of CK+ dataset.

Model	Total images	Optimizer	Learning rate	Epoch	Iteration	Activation
AlexNet	1534	Sgdm	0.0001	300	600	RELU & Softmax
ResNet-50	1534	Adam	0.0001	173	346	RELU & Softmax

Fig. 12 and 13 shows the graph of training progress of AlexNet and ResNet-50 model. From the graph, it is observed that the model is not overfitting. The graph displays an increase in training accuracy and decrease in loss over time. From Fig. 12, it is observed that when datasets are trained over 300 epochs with learning rate 0.0001 and sgdm optimizer is used it attains training accuracy of 94.52% and Validation Accuracy of 88.24% and overall average accuracy of model is 90.55%.While from Fig. 13, it is observed that when same datasets are trained over 173 epochs with the learning rate 0.0001 and adam optimizer is used it attains training accuracy is 96.43% and validation accuracy is 98.53% and overall average accuracy of model is 97.32%.The results are tabulated in Table 8 and the classification results based on accuracy, precision, sensitivity, specificity is in Table 9. Fig.10 and Fig.11 represents the SVM scatter plot of original and prediction dataset of each class of CK+ datasets.

Experimental results which are performed in MATLAB.

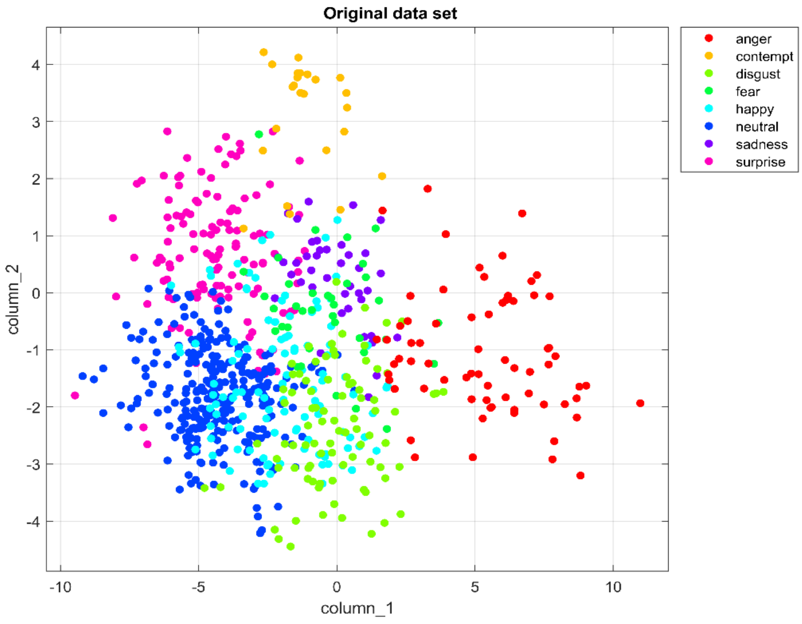


Figure 10: SVM Scatter Plot of Original Data Set of different emotion classes.

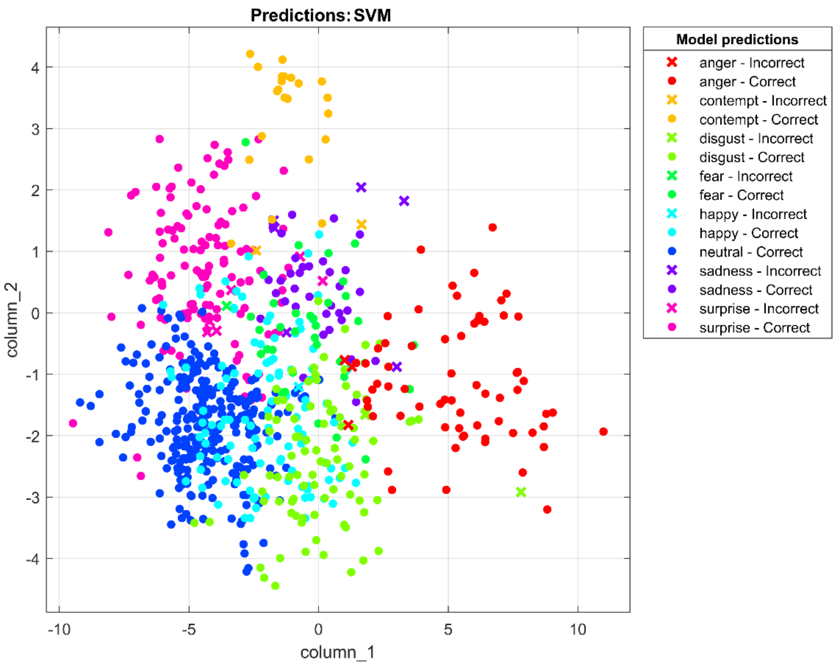


Figure 11: SVM Scatter Plot of Prediction Model of different emotion classes.

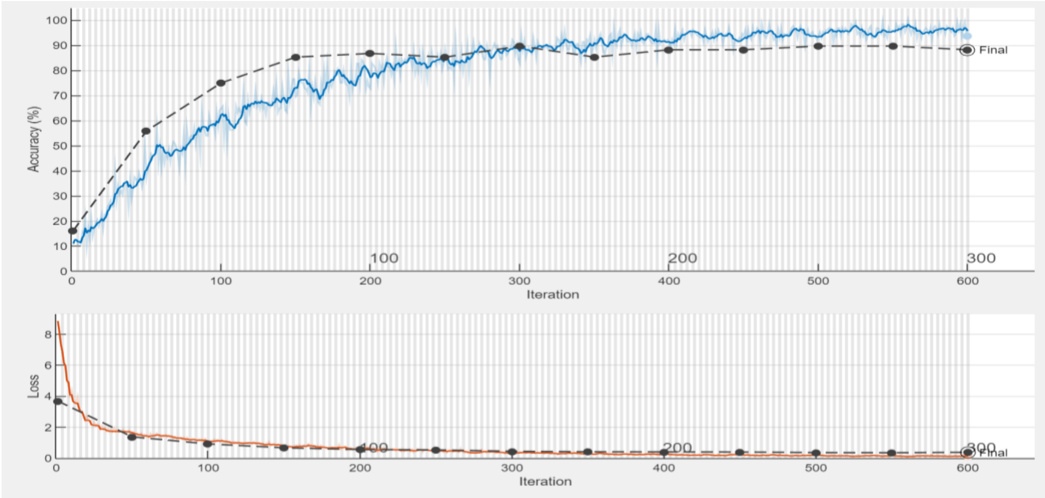


Figure 12: Accuracy (%) and loss curve v/s iteration with AlexNet Model of CK+ Dataset

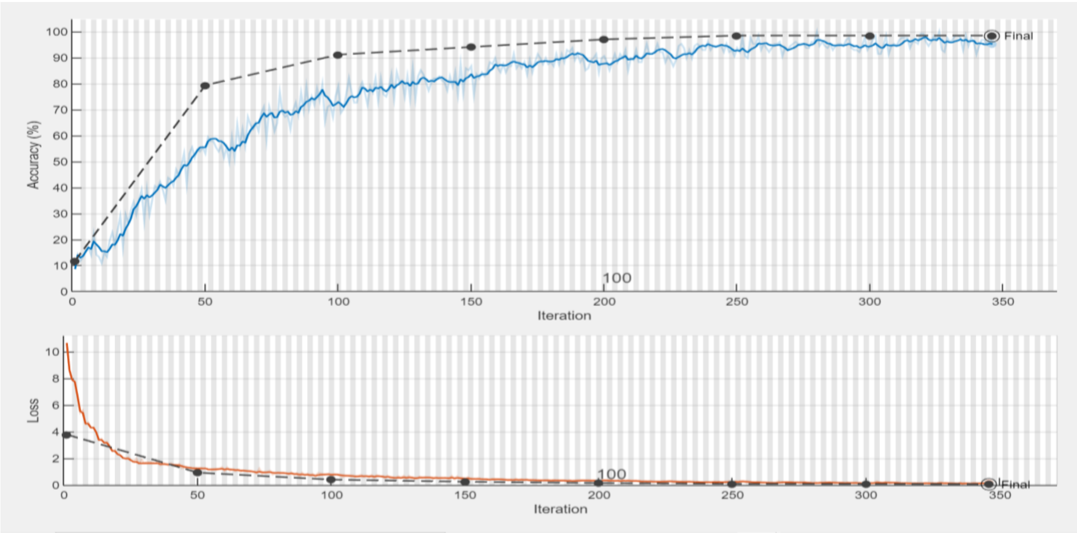


Figure 13: Accuracy (%) and loss curve v/s iteration with ResNet-50 Model of CK+ Dataset.

Table 8: Training and Validation Accuracy of AlexNet and ResNet-50 Models of CK+ dataset.

Model	Training accuracy	Validation accuracy	Average accuracy
AlexNet	94.52%	88.24%	90.55%
ResNet-50	96.43%	98.53%	97.32%

Table 9: Classification Results of emotion classes of AlexNet (A) and ResNet-50 (R) Model.

Classes	Accuracy		Precision		Sensitivity		Specificity		Accuracy (%)	
	A	R	A	R	A	R	A	R	A	R
Anger	0.98	1	0.91	1	0.91	1	0.99	1	85.3	91.2
Contempt	0.98	1	0.82	1	0.82	1	0.99	1	88.9	74.1
Disgust	0.98	1	0.94	1	0.94	1	0.99	1	88.8	95.5
Fear	0.99	1	0.93	1	0.93	1	0.99	1	15.8	84.2
Happy	0.98	1	0.96	1	0.96	1	0.99	1	91.3	95.2
Neutral	0.99	1	0.99	1	0.99	1	0.99	1	98.6	100
Sadness	0.99	1	0.92	1	0.92	1	0.99	1	61.9	81.0
Surprise	0.97	1	0.92	1	0.92	1	0.98	1	88.8	93.6

From Table 9, it is observed that when datasets is trained over AlexNet and ResNet-50 models Neutral emotions achieved maximum accurcacy of 98.6% and 100% respectively and Fear class achieved mimimum accuracy of 15.8% on AlexNet and contempt class achieved mimimum accuracy of 74.1% on ResNet-50.

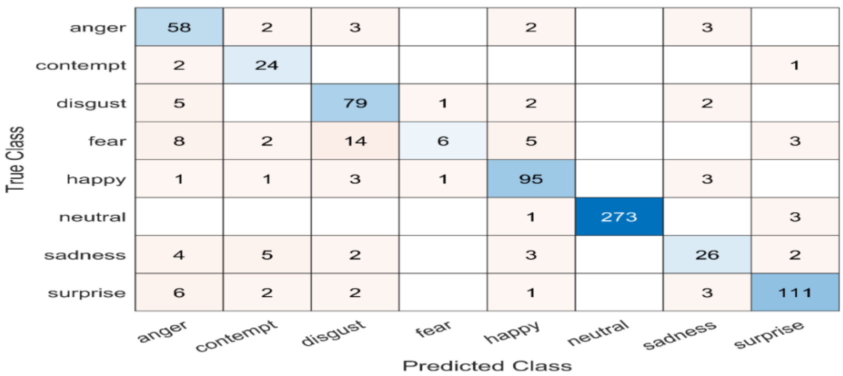


Figure 14: Confusion matrix of AlexNet Proposed Method of emotion classes.

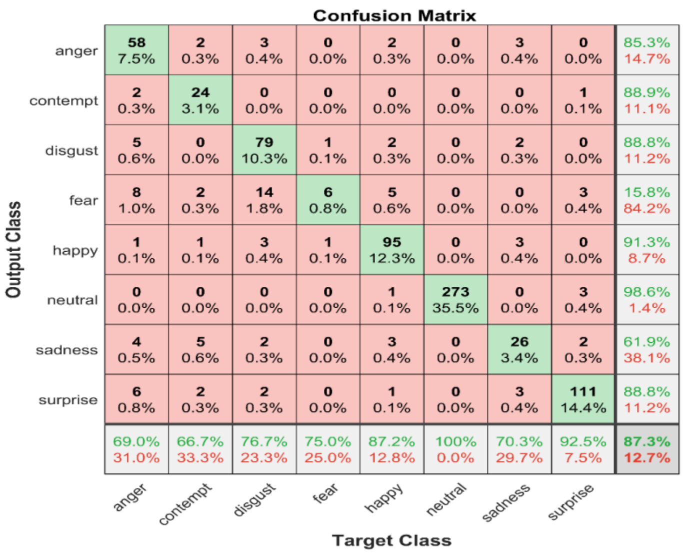


Figure 15: Classification report of AlexNet Proposed Method of emotion classes.

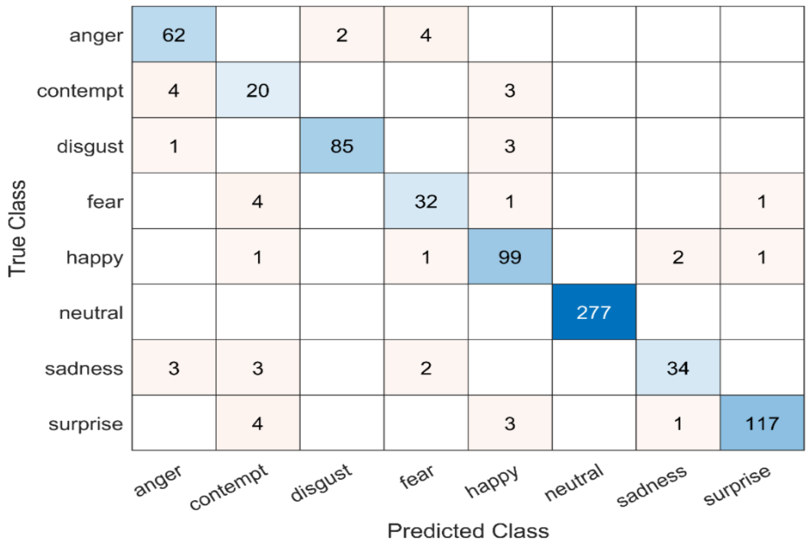


Figure 16: Confusion matrix of ResNet-50 Proposed Method of emotion classes.

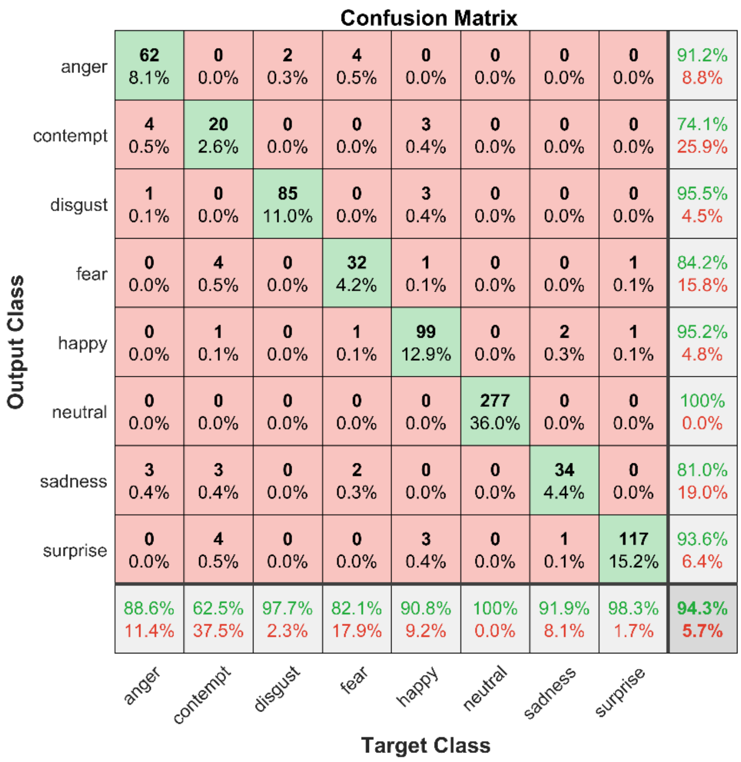


Figure 17: Classification report of ResNet-50 Proposed Method of emotion classes.

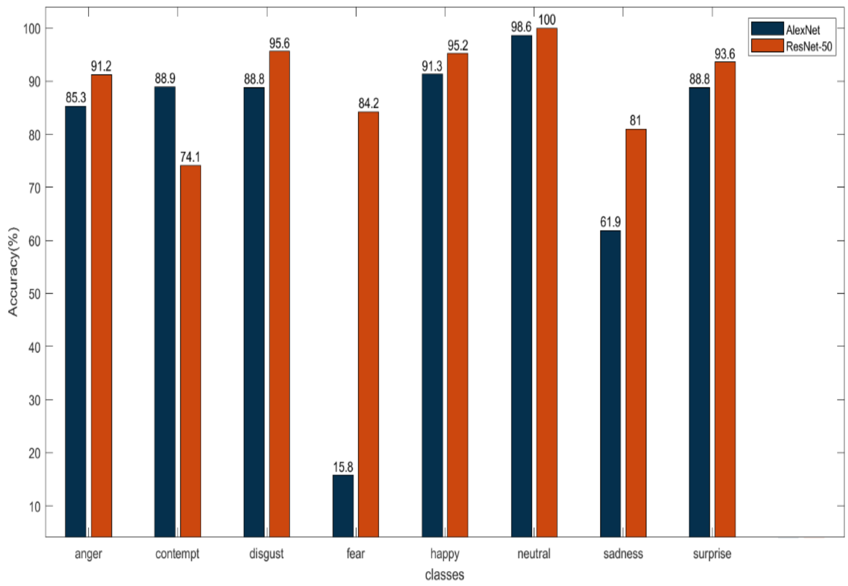


Figure 18: Graphical representation of comparison of both models.

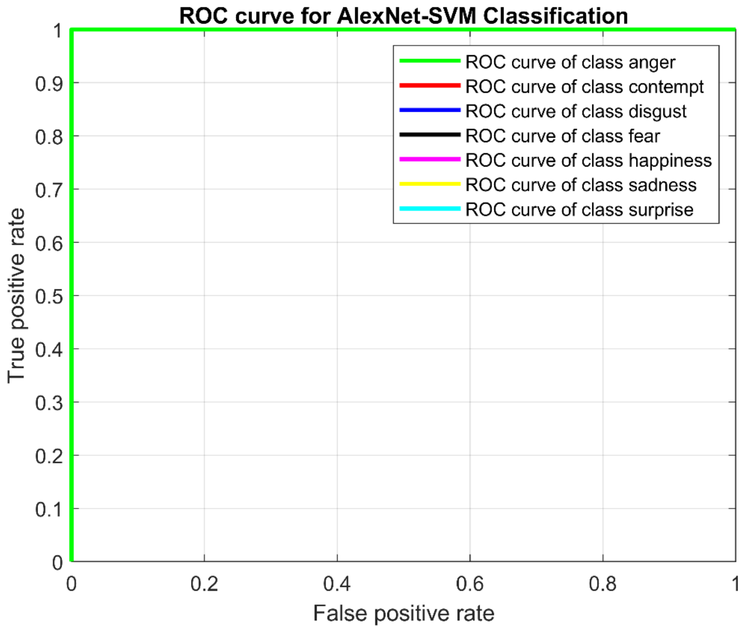


Figure 19: ROC curve of AlexNet SVM classifier of emotion classes.

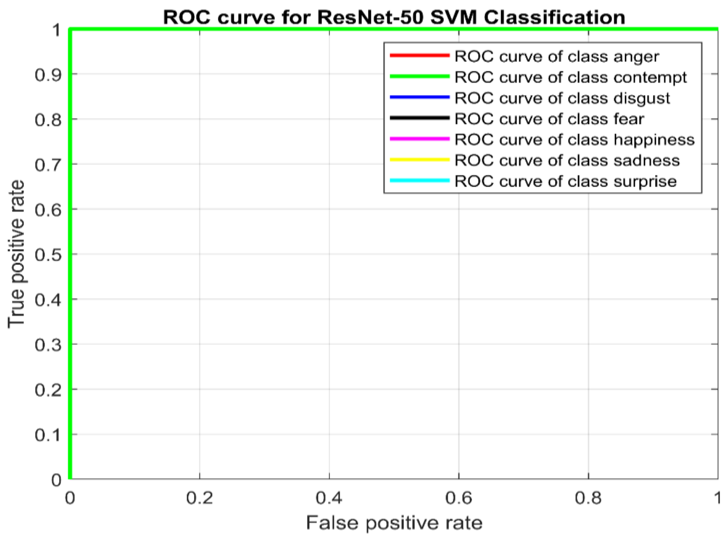


Figure 20: ROC curve of ResNet-50 with SVM classifier of emotion classes.

From Fig.15, it is observed that overall, 87.3% predictions are correct and 12.7% are incorrect in AlexNet while from Fig.17 it is clearly observed that 94.3% predictions are correct and 5.7% are incorrect in ResNet-50 model. Fig.19 and Fig.20 represents ROC curve of each model with SVM classifier and AUC is 1 for all emotions classes on both models. It means it is able to discriminate between classes.

4.2. Classification And Detection of Covid-19 From CT- Scan Dataset

Initially, all the images are resized to 256-by256 by batch image processor and converted into gray-scale in MATLAB. Following this, input images are trained on CNN models (AlexNet and ResNet-50) and the hyperparameter tuning of models are tabulate in Table 10.

Table 10: Hyper parameters of model of CT scan dataset

Model	Number of images	Optimizer	Learning rate	Epoch	Iteration	Iterations per epoch
AlexNet	702	Sgdm	1×10^{-4}	50	150	3
ResNet-50	702	Adam	1×10^{-4}	89	355	4

Fig. 21 and 22 shows the graph of training accuracy and loss curve of AlexNet and ResNet-50 model. The graph displays an increase in training accuracy and decrease in loss over time. From 21, it is observed that when datasets are trained over 50 epochs with learning rate 0.0001 and sgdm optimizer is used it attains Mini-batch accuracy of 88.72% and Mini-batch loss of 0.2861and overall average accuracy of model is 85.50%.While from Fig. 22, it is observed that when same datasets are trained over 89 epochs with learning rate 0.0001 and adam optimizer is used it attains Mini-batch accuracy of 98.44% and Mini-batch loss of 0.0974 and overall average accuracy of model is 95.72%.The results are tabulated in Table 11 and the classification analysis of each class is tabulated in Table 12 based on accuracy, precision, sensitivity, specificity.

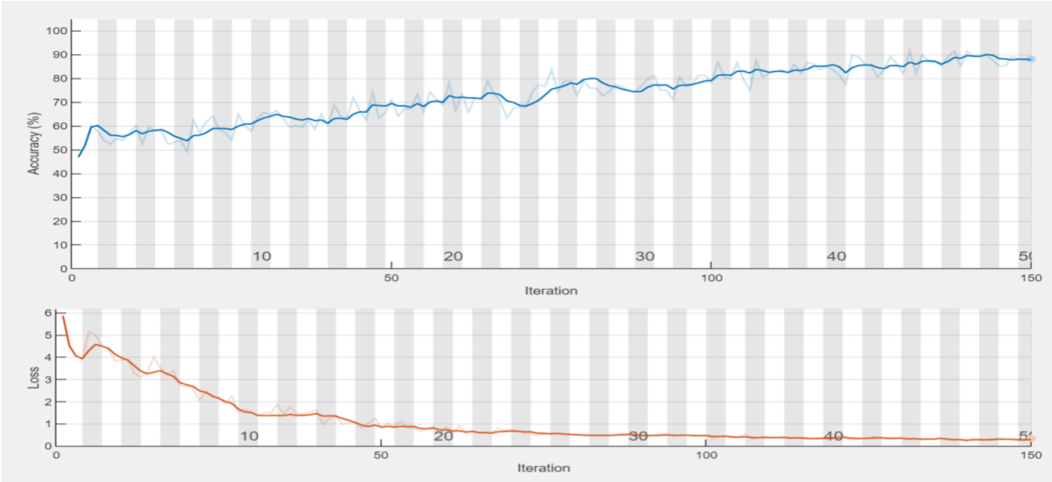


Figure 21: Accuracy (%) and loss curve v/s iteration of AlexNet Model of CT scan dataset.

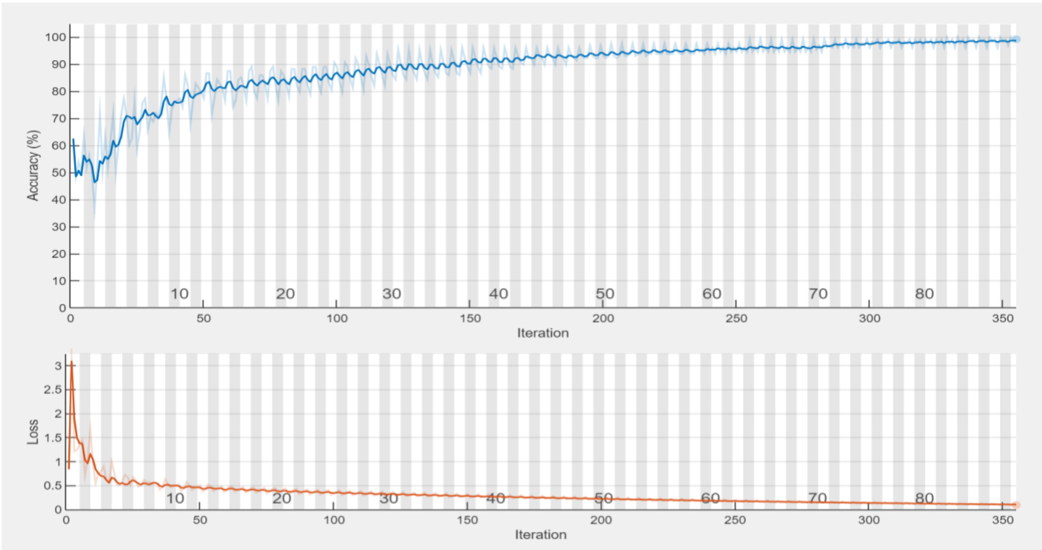


Figure 22: Accuracy (%) and loss curve v/s iteration of ResNet-50 Model of CT scan dataset.

Table 11: Results of Transfer learning of CT scan dataset.

Model	Mini batch accuracy (%)	Mini batch loss(%)	Average accuracy (%)	Training time
AlexNet	88.72	0.2861	85.50	10 min 50 sec
ResNet-50	98.44	0.0974	95.72	106 min 30 sec

Table 12: Comparison analysis of AlexNet (A) and ResNet-50(R) of CT scan dataset.

Classes	Accuracy		Precision		Specificity		Sensitivity		Accuracy(%)	
	(A)	(R)	(A)	(R)	(A)	(R)	(A)	(R)	(A)	(R)
COVID-19	90.14	96.55	89.24	96.27	90.90	96.79	89.24	96.27	92.1	94.5
NORMAL	87.11	94.91	88.08	95.26	85.97	94.51	88.08	95.26	79.7	96.8

In Fig. 24, the first two diagonal cells shows the number and percentge of classification by trained the network. Means, 151 are correctly classified as COVID-19 and it is corrsponds to 43.0% and 149 as NORMAL and it is corrsponds to 42.5%. similarly,13 and 38 are wrongly classified as COVID-19 and NORMAL and it is corresponds to 3.7% and 10.8% respectively. Out of 164 COVID-19 predictions, 92.1% are correct and 7.9% are incorrect. Out of 187 NORMAL predictions 79.7% are corect and 20.3% are incorrect. Out of 189 COVID-19 cases,79.9% are correctly classified as COVID-19 and 20.1% are predicted as NORMAL. Out of 162 NORMAL cases,92.0% are classified as NORMAL and 8.0% are classified as COVID-19. Overall, 85.5% of the predictions are correct and 14.5% are wrong.

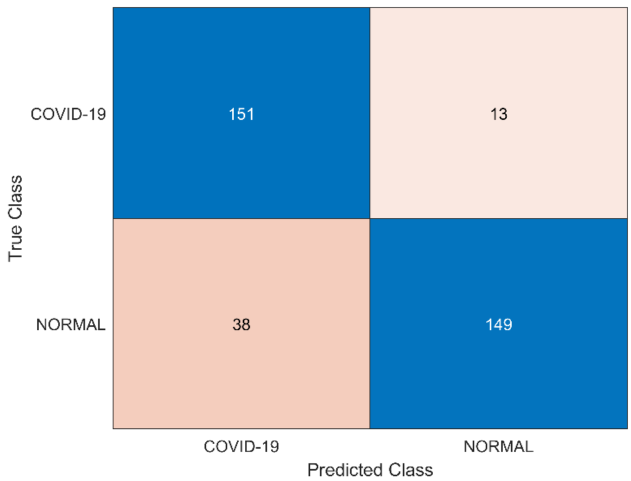


Figure 23: Confusion matrix of AlexNet Model of CT scan dataset.

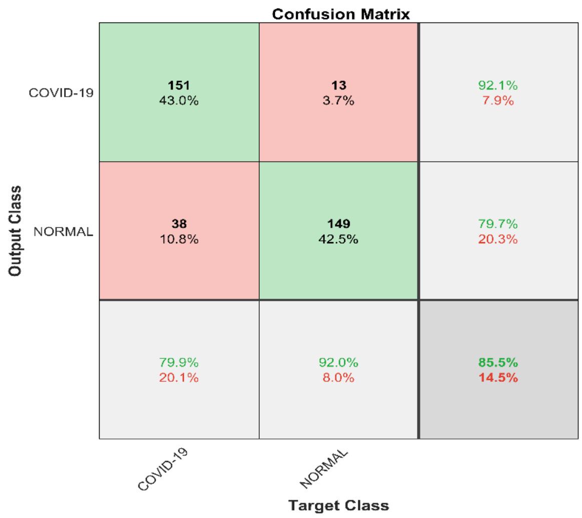


Figure24: Confusion report of AlexNet Model of CT scan dataset.

Similarly, in Fig.26, values across the diagonal cells means 155 are correctly classified as COVID-19 and it is corrsponds to 44.2% and 181 as NORMAL and it is corrsponds to 1.6%. similarly,9 and 6 arewrongly classified as COVID-19 and NORMAL and it is corresponds to 2.6% and 1.7% respectively. Out of 164 COVID-19 predictions, 94.5% are correct and 5.5% are incorrect. Out of 187 NORMAL predictions 96.8% are corect and

3.2% are incorrect. Out of 161 COVID-19 cases,96.3% are correctly classified as COVID-19 and 3.7% are predicted as NORMAL. Out of 190 NORMAL cases,95.3% are classified as NORMAL and 4.7% are classified as COVID-19. Overall, 95.7% of the predictions are correct and 4.3% are wrong.

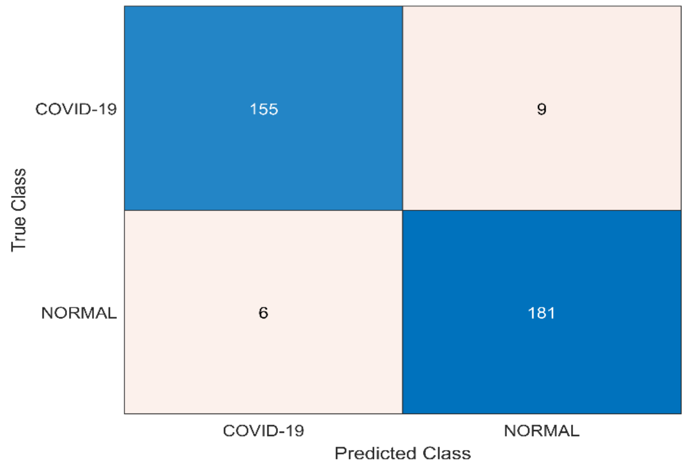


Figure 25: Confusion matrix of ResNet-50 Model of CT scan dataset.

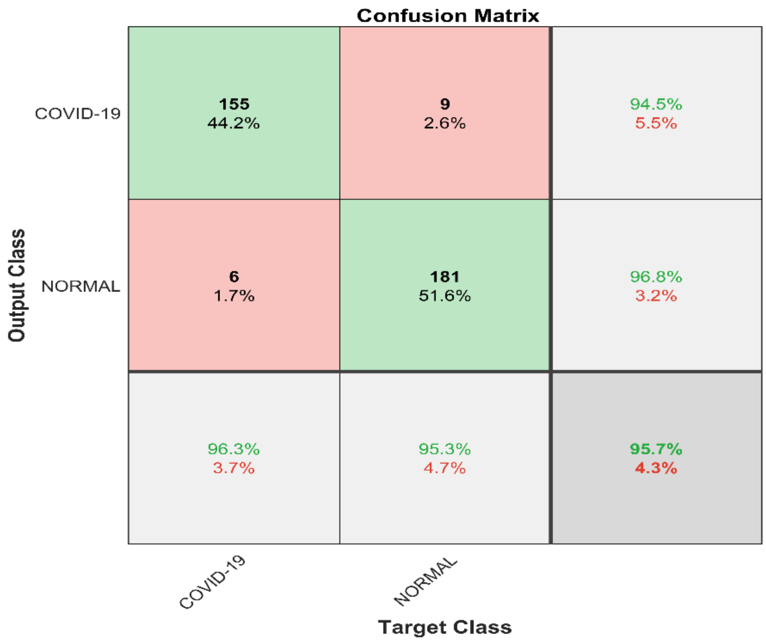


Figure 26: Confusion report of ResNet-50 Model of CT scan dataset.

In Fig. 27 it is observed that AUC for alexnet is 0.9587 for both class COVID-19 and NORMAL and from Fig.28 it is observed that, AUC is 0.9869 for both classes. The results are tabulated in Table 13.

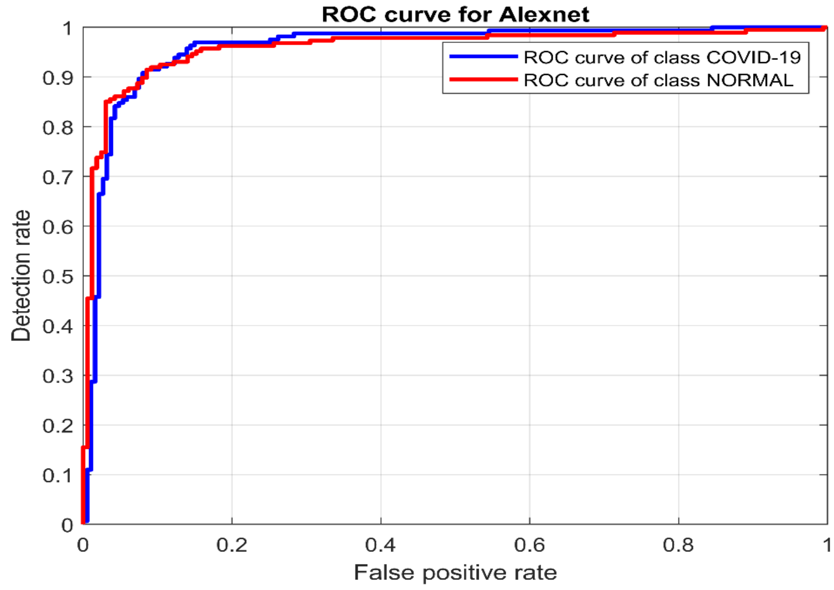


Figure 27: ROC curve of AlexNet Model of CT scan dataset.

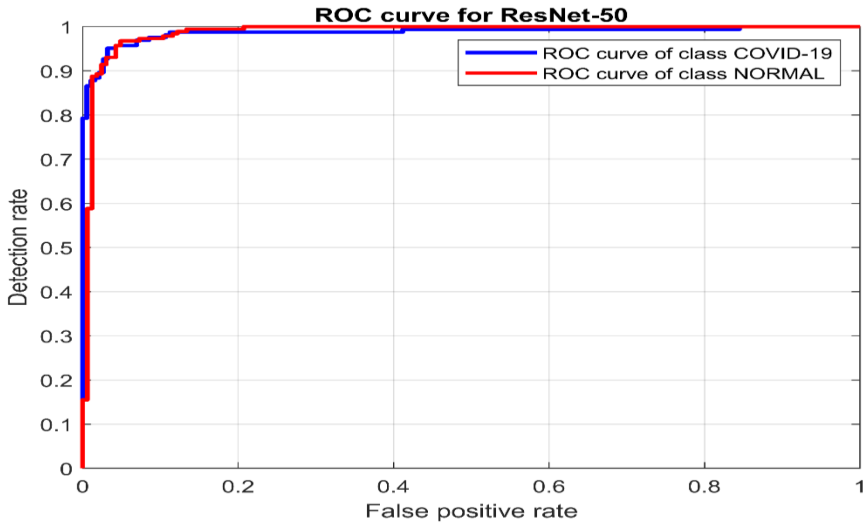
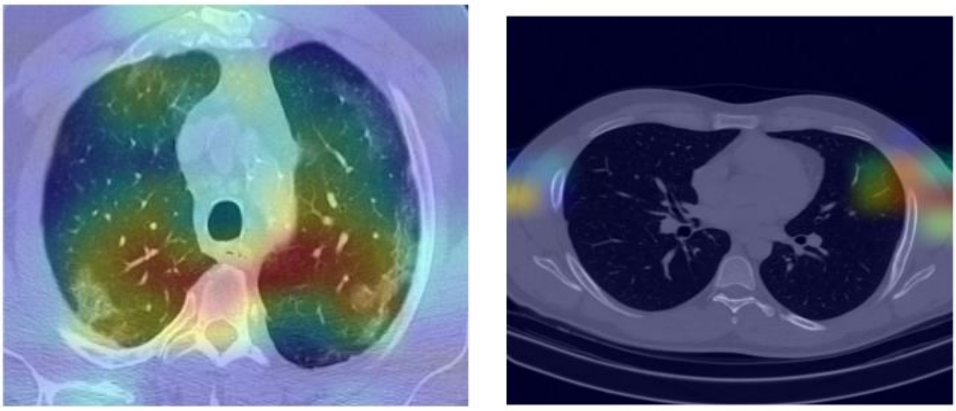


Figure 28: ROC curve of ResNet-50 Model of CT scan dataset.

Table 13: Area under the curve of ROC graph of COVID-19 and NORMAL classes

Classes	AUC for AlexNet	AUC for ResNet-50
COVID-19	0.9587	0.9869
NORMAL	0.9587	0.9869

In accordance with the testing of algorithm Grad-cam is implemented on a sample of COVID-19 and Normal CT-scans to see portions of the CT-scan which is utilized as a marker to show the beginning phases of the COVID-19 contamination. The CT-scan with a COVID-19 positive outcome has a 95.72% certainty level, while the COVID-19 outcome showed a 48.80% confirmation. Fig.5.20 shows Grad-cam result of infected and non-infeced lungs CT-scans.



(a) Heatmap showing covid-19 positive (b) Heatmap showing covid-19 negative

Figure 29: Grad-CAM result of ResNet-50 Model.

After performing this experiment, the final results shows that the ResNet-50 architecture, performs much better as compared to AlexNet architecture in both applications i.e., Emotion detection and their classifications from Facial Expressions and classification of COVID-19 from CT - scan dataset.

5. Conclusions

In this project, different types of convolutional neural network models are studied and implemented to detect emotions from facial expression and COVID from CT-scans. And SVM classifier is used to extracted features from convolution base of pre-trained models. In this work, AlexNet and ResNet-50 models are used and comparison of both the models are done on the basis of their training and testing accuracy, confusion matrices and ROC graph. In the 1st application FER system is designed and implemented on CK+ dataset to achieve the accuracy of 97.32% on ResNet-50 model which is better than 90.55% which got on AlexNet model. The Confusion Matrix of CK+ dataset clearly shows that the ResNet-50 model with adam optimizer and SVM classifier can classify emotions more efficiently and Neutral emotion show maximum accuracy i.e., 98.6% in AlexNet and 100% in ResNet-50 model. The area under the curve of ROC graph also shows that the model can distinguish all the emotions clearly. In the 2nd application, same CNN models are used to train the CT-scan datasets and to detect COVID-19. In this, the final results show that the ResNet-50 architecture, performs much better as compared to AlexNet architecture with an accuracy of 95.72%, and ROC-AUC of 98.69%. For COVID-19 testing different pre-trained CNN models like ResNet-50 can be used which is described in this paper, and it can become an alternative for covid-19 testing. Grad-CAM is implemented to detect the existence of COVID-19 infection in lungs CT scan. After comparison of both the model, it is concluded that ResNet-50 performs much better and give higher accuracy as compared to AlexNet model.

References

[1] Pranav E., Suraj Kamal, Satheesh Chandran C., Supriya M.H..: Facial Emotion Recognition Using Deep Convolutional Neural Network.6th International Conference on Advanced Computing &

Communication Systems (ICACCS), pp. 317-320.IEEE (2020).

[2] Sai Yeshwanth Chaganti, Ipseeta Nanda, Koteswara Rao Pandi, Tavva GNRSN Prudhvith, Niraj Kumar.: Image Classification using SVM and CNN, International Conference on Computer Science, Engineering and Applications (ICCSEA), pp. 1-5. IEEE (2020).

[3] Balaji Balasubramanian, Pranshu Diwan.: Analysis of Facial Emotion Recognition. 3rd International Conference on Trends in Electronics and Informatics, pp.945-949. IEEE (2019).

[4] Guolu Cao, Yuliang Ma*, Xiaofei Meng, Yunyuan Gao, Ming Meng," Emotion Recognition Based On CNN", 2019 Chinese Control Conference (CCC), Guangzhou, China, 2019, pp. 8627-8630.

[5] D. F. Eljamassi and A. Y. Maghari, "COVID-19 Detection from Chest X-ray Scans using Machine Learning," 2020 International Conference on Promising Electronic Technologies (ICPET), 2020, pp. 1-4.

[6] P. Hartono, C. R. Luhur, C. A. Indriyani, C. R. Wijaya, N. N. Qomariyah and A. A. Purwita,"Evaluating Deep Learning for CT Scan COVID-19 Automatic Detection," 2021 International Conference on ICT for Smart Society (ICISS), 2021, pp. 1-7.

[7] Aseel Qassim Abdul Ameer, Raghad Falih Mohammed, "Covid-19 detection using CT scan based on gray level Co-Occurrence matrix", Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.04.224.

[8] Saood A, Hatem I, "COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet", BMC Med Imaging, 2021 Feb, https://doi.org/10.1186/s12880-020-005295.

[9] https://github.com/spenceryee/CS229/tree/master/CK%2B.

[10] E. Soares, P. Angelov, S. Biaso, M. H. Froes, and D. K. Abe, "SARS-CoV-2 CT-scan Dataset: A Large Dataset of Real Patients CT scans for SARS-CoV-2 Identification", medRxiv, New Haven, CT, USA, 2020.

[11] Nahla Nour, Mohammed Elhebir and Serestina Viriri," Face Expression Recognition Using Convolution Neural Network (CNN) Models", International Journal of Grid Computing & Applications (IJGCA) Vol.11, No.1/2/3/4. December 2020.

[12] Y. Li, J. Zeng, S. Shan and X. Chen, "Occlusion Aware Facial Expression Recognition Using CNN With Attention Mechanism," in IEEE Transactions on Image Processing, vol. 28, no. 5, pp. 2439-2450, May 2019, doi: 10.1109/TIP.2018.2886767.

[13] Aayush Bhardwaj, Ankit Gupta, Pallav Jain, Asha Rani, Jyoti Yadav.: Classification of human emotions from EEG signals using SVM and LDA Classifiers. 2nd International Conference on Signal Processing and Integrated Networks (SPIN), pp. 180-185.IEEE(2015).

[14] G., Oztel, I., Kazan, S. et al.,"Facial expression recognition for monitoring neurological disorders based on convolutional neural network", Multimed Tools Appl 78, 31581–31603 (2019). https://doi.org/10.1007/s11042-019-07959-6

[15] Wafa Mellouk, Wahida Handouzi," Facial emotion recognition using deep learning: review and insights", Procedia Computer Science, Volume 175, 2020, Pages 689-694.

[16] C. Ruoxuan, L. Minyi, and L. Manhua," Facial Expression Recognition Based on Ensemble of Multiple CNNs", Lecture Notes in Computer Science (), vol 9967, pp. 511-578, Springer International Publishing AG (2016).

[17] Deepak Kumar Jain, Pourya Shamsolmoali, Paramjit Sehdev, “Extended deep neural network for facial emotion recognition”, Pattern Recognition Letters, Volume 120, 2019, Pages 69-74,

[18] D. H. Kim, W. J. Baddar, J. Jang and Y. M. Ro, "Multi-Objective Based Spatio-Temporal Feature Representation Learning Robust to Expression Intensity Variations for Facial Expression Recognition," in IEEE Transactions on Affective Computing, vol. 10, no. 2, pp. 223-236, 1 April-June 2019, doi: 10.1109/TAFFC.2017.2695999.

[19] André Teixeira Lopes, Edilson de Aguiar, Alberto F. De Souza, Thiago Oliveira-Santos, “Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order”, Pattern Recognition, Volume 61, 2017, Pages 610-628.

[20] Osama Shahid, Mohammad Nasajpour, Seyedamin Pouriyeh, Reza M. Parizi, Meng Han, Maria Valero, Fangyu Li, Mohammed Aledhari, Quan Z. Sheng, “Machine learning research Towards combating COVID-19: Virus detection, spread prevention, and medical assistance”, Journal of Biomedical Informatics, Volume 117, 2021,

[21] M. J. Horry, S. Chakraborty, M. Paul, A. Ulhaq, B. Pradhan, M. Saha, and N. Shukla, “Covid-19 detection through transfer learning using multimodal imaging data,” IEEE Access, vol. 8, pp. 149, 2020.

[22] H. X. Bai, R. Wang, Z. Xiong, B. Hsieh, K. Chang, K. Halsey, T. M. L. Tran, J. W. Choi, D.-C. Wang, L.- B. Shi, J. Mei, X.-L. Jiang, I. Pan, Q.- H. Zeng, P.-F. Hu, Y.-H. Li, F.-X. Fu, R. Y. Huang, R. Sebro, Q.-Z. Yu, M. K. Atalay, and W.-H. Liao, “Artificial Intelligence Augmentation of Radiologist Performance in Distinguishing COVID-19 from Pneumonia of Other Origin at Chest CT,” Radiology, vol. 296, no. 3, pp. E156– E165, 2020.

[23] S. Wang, Y. Zha, W. Li, Q. Wu, X. Li, M. Niu, M. Wang, X. Qiu, H. Li, H. Yu, W. Gong, Y. Bai, L. Li, Y. Zhu, L. Wang, and J. Tian, “A fully automatic deep learning system for covid-19 diagnostic and prognostic analysis,” European Respiratory Journal, vol. 56, no. 2, 2020.

[24] L. Sun, Z. Mo, F. Yan, L. Xia, F. Shan, Z. Ding, B. Song, W. Gao, W. Shao, F. Shi, H. Yuan, H. Jiang, D. Wu, Y. Wei, Y. Gao, H. Sui, D. Zhang, and D. Shen, “Adaptive feature selection guided deep forest for covid-19 classification with chest ct,” IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 10, pp. 2798–2805, 2020.

[25] S. Ahuja, B. K. Panigrahi, N. Dey, V. Rajinikanth, and T. K. Gandhi, “Deep transfer learning- based automated detection of COVID-19 from lung CT scan slices,” Applied Intelligence, vol. 51, no. 1, pp. 571–585, 2020.

[26] Garain, A. Basu, F. Giampaolo, J. D. Velasquez, and R. Sarkar, “Detection of COVID 19 from CT scan images: A spiking neural network-based approach,” Neural Computing and Applications, vol. 33,19 (2021).

[27] V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, and N. Mehendale, “Diagnosis of COVID-19 using CT scan images and deep learning techniques,” Emergency Radiology, vol. 28, no. 3, pp. 497–505, 2021.

[28] M. Maftouni, A. C. C. Law, B. Shen, Y. Zhou, N. Ayoobi Yazdi, and Z. Kong, “A robust ensemble-deep learning model for covid-19 diagnosis based on an integrated ct scan images database,” June 2021.

[29] Ahmad, A., Garhwal, S., Ray, S.K. et al," The Number of Confirmed Cases of Covid-19 by using Machine Learning: Methods and Challenges", Arch Computat Methods Eng 28, 2645–2653 (2021).

[30] Apostolopoulos, I.D., Mpesiana, T.A," Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks", Phys Eng Sci Med 43, 635–640 (2020)

[31] Selvaraju, R. R., M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization." In IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618–626. Available at Grad-CAM on the Computer Vision Foundation Open Access website.