

Emotion Analysis Using Convolutional Neural Network

Renu D.S

Associate Professor/CSE

Mar Ephraem College of Engineering and Technology,
Marthandam, Tamil Nadu .

Tintu Vijayan

Computer Science and Engineering

Mar Ephraem College of Engineering and Technology,
Marthandam, Tamil Nadu.

Dr. D. Dhanya

Assistant Professor/CSE

Mar Ephraem College of Engineering and Technology,
Marthandam, Tamil Nadu

Abstract—Human facial expressions convey a great deal of data visually rather than articulately. Facial expression is that the best thanks to determine human feeling. Happy, sad, neutral, disgust, angry, surprise and fear are the different facial expression. Here I create a model by using Keras with Convolutional Neural Network (CNN) and trained the model with using FER 2013 dataset provided by Kaggle. The model is trained to classify mainly seven different facial expressions from any person's face. This model consists of basically two parts. The first part will remove all backgrounds from the picture and save as a new image. In second part the model will work on the new image, for emotion recognition. Emotion analysis through facial gestures may be a technology that aims to enhance product and services performance by observance client behavior to bound product or maintenance staff by their analysis.

Keywords—Convolutional Neural Network (CNN), Facial Emotion Recognition (FER)

I. INTRODUCTION

Facial expression is crucial to identify one's feelings. Facial emotions are the nonverbal way of communication. It is troublesome to search out feeling, those that face stress and in bed. So machine-driven facial emotion recognition techniques facilitate us plenty in such cases. We can classify the emotion recognition techniques into two:

- Vision-based technique
- Bio signal/Physiological based technique

In the Vision-Based technique, the camera can detect the face, and based on the facial features deep learning model can acknowledge the feeling. But, this method not invariably provides true emotion of the mind. Individuals may also pretend up the emotions on the face. To detect the real emotions biosignal/physiological based techniques are used.

Bio signal/Physiological based technique works on real body sensors, so nobody will devour the emotions. Biosensors can catch the reactions happening in the body. This technique can be performed by using the cardio signal, brain waves, and so on. By incorporating regression problems, the intensity of feeling replicate on the face may also discover.

The vision-based technique is a type of classification problem. This proposed model will detect faces from the webcam and according to the facial expression, it will classify

the expression as Happy, sad, neutral, disgust, angry, surprise, or fearful. FER 2013 is the dataset selected for doing this project. FER 2013 is provided by Kaggle. This proposed approach focused on images having a background, most of the CNN models are confused by the backgrounds, and this led to giving the false output. The first step of our approach will remove the background from the image and make it more realistic. In the second part, our model classifies facial expression into any of the seven categories. The main objective of this proposed work is not only to develop a machine-driven system but also to enhance the accuracy and performance of the model as compared to other models.

II. LITERATURE REVIEW

Nowadays the emotional aspects attract the attention of many research areas, not only in computer science, but also in psychology, healthcare, communication, etc.[26]. Earlier approaches of facial detection and recognition are using binary pattern from the features. By using grey scale and rotation invariant texture classification based on local binary patterns is very effective and robust and also very low computational effort [1].

Ahonen et.al proposed local binary pattern technique, which considers both texture and shape to consider image features. This method has so many advantage as like efficiency, fast feature extraction and so on. Main advantage of this system lies on the length of feature vector used for face representation [2].

Kim et.al [3] proposed a multi-class classifier based Adaboost algorithm for efficient classification of multi class data. Instead of series of binary classifier here uses multi class classifiers. So that it attain low training time, more stable and more accurate classification results. It is one of the popularly used practical boosting algorithms. Although Adaboost is more resistant to over fitting than many machine learning algorithms, it is repeatedly sensitive to noisy data and outliers[4].

Viola et.al proposed a real time application method for object detection, that is capable of processing pictures quickly and of high rates. This consists of three factors. New image representation called integral image, feature extraction is take place by using adaboost algorithm and last make image

clean by using classifiers. This method works approximately 15 times faster than previous models [5].

Yang et.al demonstrated a new method for image representation and detection by using principal component analysis and Fisher linear discriminant method. This method is used for low dimensional representation of images. Yang et.al compared this method of feature extraction with all other existing nonlinear feature extraction technique, they find out that this kernel method do not need nonlinear optimization only they need the solution of eigen value[6].

Szegedy et.al proposed a new architecture by using deep convolutional network named as Inception for classification and detection of image. This architecture provides better result as compared to the existing narrower architectures.[7]

Deep neural network is kind of multi neural artificial neural network which consists of more than one hidden layer, Stuhlsatz et.al proposed a work on acoustic emotion recognition by using a new architecture called a Generalized Discriminant Analysis (GerDA) based on deep neural network.[8]

The work by Alice et.al developed a new algorithm for face recognition as human face processing models. To create a special representation of subject response pattern they used multidimensional scaling (MDS)[9].

Jian Yang et.al coined a new technique for image representation called two dimensional component analysis(2DPCA). Image feature extraction can done efficiently and accurately by using 2DPCA than PCA. But one disadvantage of this system is 2DPCA needs more coefficients for image extraction than PCA. By using this method image feature extraction can perform simpler and more straight forward[10].

Kwang et.al proposed a new method by combining by combining Active Appearance model(AAM) with Dynamic Bayesian Network (DBN) for facial emotion recognition. They used AAM for feature extraction and DBN for facial feature classification. The dataset used for this work is BioID, which shows an accuracy of more than 90% and also this model gives high recognition performance level compared with all other existing models[11].

Matre et.al has made a detail discussion on different methods used for facial emotion detection. Based on the detailed discussion they find out that performance metrics of these methods were calculated on basis of recognition rate, higher the recognition rate, higher the performance. They find out that tensor perceptual color framework has higher recognition rate and higher performance and they suggest higher studies on this topic need to be focused on gene matching to the geometric factors of facial expression[12].

Yelinkim et.al suggests a work which shows high order nonlinear relationships are much effective for facial emotion recognition. In this paper they tried to overcome the existing limitations as like feature selection model focus mainly on linear relationships of features. They overcome these limitation by proposing a deep belief network model focus on complex nonlinear features. This model shows a better performance gain in non-prototypical data[13].

Shojaeilangari et.al proposes Extreme sparse learning (ESL) approach. For the better classification of image with noisy signal and raw data they combines Extreme learning machine with the reconstruction property of sparse representation. This approach gives better performance rate on both acted and real time facial emotion. But still some disadvantages are there in this system, that is very high computational cost for both feature extraction and classification[14].

Because of the difficulty to train deep neural network, Kaiming et.al proposed residual learning framework to make this task easier. They used residual learning to train deep neural models consists of 152 layers, which is 8 times deeper than VGG net. Due to deep representations, they obtain relatively high performance in object detection as compared to COCO dataset. They used Faster CNN as object detection method. This work has made a remarkable change in VGG-16 with ResNet-101[15].

Neha et.al proposed a system that used semantics based approach to detect emotional activity by using temporal and spatial properties of the objects. Beyond the classification of emotional activity they also classified actions as sitting posture and standing posture. Classification is occurring based on the tracking of movements of arms, based on predefined set of rules and conditions behaviors are semantically detected. This model gives better adaptability and robustness and also eliminates the training required by machine learning models [16].

Yao et.al proposed a new well designed Convolutional Neural Network(CNN) for emotion recognition from videos. They named this new method as HOLONET, it has three important concerns in network design. First, to reduce redundant filters and enhance the non-saturated non-linearity in the lower convolutional layers, instead of Relu they used Concatenated Rectified Linear Unit(CRELU). Second, to construct middle layer they combine residual structure and CRELU and thus they attain accuracy and efficiency. Third, the topper layers are designed as a deviation of inception residual structure. This method has better accuracy and performance rate, but this method can only deploy on video based emotion [17].

Yadan et.al proposes a work mainly focus on facial expression recognition with the face components by face parsing. The model is trained by deep belief network and tuned by logistic regression. In this work they demonstrate that information for facial expression are not same for different faces, the idea of this work is to identify the components of face which are active in expression. The detectors first detect the active components like, face then nose, eyes, mouth hierarchically. The experiment is conducted on Japanese female facial expression database and Cohn-Kanade dataset. The results of this work shows that eyes and mouth can discriminate expressions clearly [18].

In this paper Kaiyue Li et.al propose a new light weight fully convolutional network (L-FCN), in order to overcome the drawbacks of existing methods as like higher hardware consumption and computational costs. They create this model by using traditional watershed algorithm and fully convolutional network. They uses ISBI 2012 dataset for evaluation and by the use of combination of CNN and watershed algorithm help them to eliminate useless non-edge

pixels in the raw image. So with the help of L-FCN, can avoid complex network structures and redundant parameters. Advantages of this model are avoid wastage of memory, low computing cost, efficiency etc. [19]

Tohidul et.al proposed a work on food image classification problem. It is a unique branch of classification problem. In this paper they proposed a system which will classify the food images into different food categories. Because of the nonlinear dataset of food images, this task is highly challenging. Here they used convolutional neural network (CNN) to classify images. The accuracy of the prescribed model is of 92.86% [20].

Sunitha et.al proposed an automated sign language recognition system. For sign language recognition they extract hand movements. In this paper they use a multimodal feature sharing mechanism with a four stream Convolutional neural network (CNNs) for RGB-D based sign language recognition. The proposed four stream CNN architecture with multi modal data sharing mechanism. RGB and depth data are used to train the CNN model with data sharing architecture. The proposed model is trained with all four input stream but it is tested with only two RGB special and temporal streams [21].

Zhong-Kiu et.al make a detailed review on deep learning based object detection frameworks with modification on RCNN, that handle different sub problems like clutter, occlusion, low resolution and so on. Their review begins with a study on deep learning and its tool, Convolutional neural network. Then they focus on generic object detection architectures. They also make a survey on several different tasks like face detection, object detection and pedestrian detection [22].

Santhosh kumar et.al proposed a deep learning approach for emotion detection from human body movements. They explain advantage of this system are, it can easily detect emotion of a person from any camera view and also it can easily recognize emotion if the person is far from camera. They used feed forward deep convolution neural network architecture for emotion recognition from body motion patterns. The proposed system is evaluated by emotion dataset with 15 types of emotions and GEMEP. The performance rate of this model is better than baseline models [23].

Kaviya et.al proposed a deep learning approach system for human facial sentiment recognition. They used Haar cascade filter to detect and extract face features. The system is developed by using convolutional neural network (CNN) and it is used to classify five different facial emotion namely happy, anger, sad, surprise and neutral. The proposed CNN model has an accuracy of 65% for facial expression recognition, by using FER2013 dataset and 60% accuracy for custom dataset [24].

In this paper Parvathi et.al proposed a model for emotion analysis using deep learning. They build the system by convolutional neural network (CNN) architecture. The proposed work has mainly four steps face detection, extraction, classification and recognition. Segmentation methods are used for splitting mouth region. In this paper they used a combination for supervised learning and reinforcement learning is being used [25].

III. METHODOLOGY

In this paper I proposed convolutional neural network architecture of 5 layers for facial emotion recognition. This proposed model is trained by using FER 2013 dataset which is provided by Kaggle. This dataset consists of gray scale images of different emotions of 48x48 pixel size. Also this dataset consists of 28,709 sample images, which are used for testing.



Fig 1. Sample FER 2013 dataset

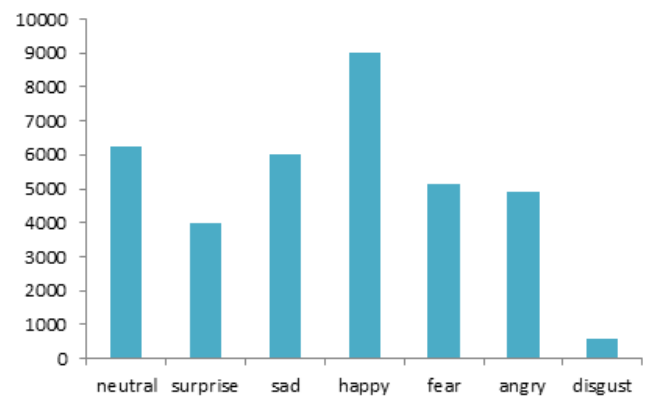


Fig 2. FER 2013 dataset

The model is created by using sequential function from Kera's library. The proposed architecture is shown below in fig 1. The first layer is input layer, this layer accepts input of size 48x48 and the images are of black and white. The first convolutional layer convolved with 5x5 kernel of size 100 for feature extraction and 64 filters. Conv2D layer is followed by a maxpooling layer of pool size 5x5 and strides 2x2. 5x5 Kernel quickly reduces the spatial dimension and learn larger features. Convolutional layer is followed by a max pooling layer to reduce spatial dimension in output volume. The second layer takes input from first layer and convolved with 3x3 kernel and filter 64 and it is followed by a average pooling2D layer of pool size 3x3 and strides of 2x2. The third layer takes input from the previous layer and convolved by using kernel 3x3 and filter 128. The filter number increases by reaching to output. It is followed by averagepooling2D of size 3x3 and strides 2x2 and then it is followed by flatten layer and finally followed by two drop out layer of 0.2 drop out and two dense layer of 1024 units. Final dense layer uses Softmax as activation function except that, remain all other layers uses Relu as activation function.

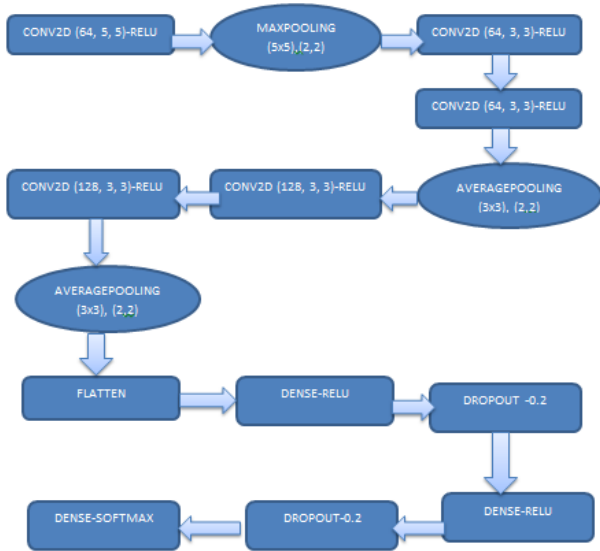


Fig 3. Proposed Architecture

Convolutional Layer

This model is built by using Keras’s CONV2D class. The first parameter is the number of Filter CONV2D has. Filters are always power of 2. The parameter filter determines the number of kernels used for convolve with the input image. The filter number increases with going to higher layers, which is reaching to output layer, filter number increases. Filters will extract the features of the input image. A convolution means multiplication of two matrices followed by a sum operation. Here the convolution layer takes performs the multiplication with input image matrix and the kernel and then it is followed by a sum. Activation map is the output of each convolution operation.

Max pooling

Maxpooling is an operation that is followed by individual convolutional layer. Output from the previous convolutional layer become input to the maxpooling layer, by reducing the number of pixels it will reduce the dimensionality of input images. Pooling downsamples the input representation. Pooling is of two types maxpooling and average pooling. In our model first convolutional layer is followed by maxpooling and then averagepooling is followed by all convolutional layers. Output size after first convolutional operation is (44,44,64), but after first maxpooling it is reduced to (20,20,64). The use of maxpooling is to reduce computational load and to reduce overfitting. Maxpooling reduces resolution of the image and it reduces the parameters, thus it reduces the computational load. Similarly maxpooling helps to extract most activated or high value pixels while discarding low value ones, thus it helps to reduce overfitting. Pooling layers are followed by flatten layer which flatten two dimensional input to 1 dimensional. Then it is followed by a dense layer and a drop out, drop out is used for regularisation of neurons. Drop out used in this model is 0.5. 3 dense layers are used in this model. Dense layer is also known as fully connected layer. The first two dense layers used Relu as activation function and the last dense layer uses Softmax as activation function and it divide the output into seven classes. For update the learning

rate I used Adam optimisation function and to calculate the loss categorical cross entropy function is used here.

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 44, 44, 64)	1664
max_pooling2d_1 (MaxPooling2D)	(None, 20, 20, 64)	0
conv2d_2 (Conv2D)	(None, 18, 18, 64)	36928
conv2d_3 (Conv2D)	(None, 16, 16, 64)	36928
average_pooling2d_1 (Average Pooling2D)	(None, 7, 7, 64)	0
conv2d_4 (Conv2D)	(None, 5, 5, 128)	73856
conv2d_5 (Conv2D)	(None, 3, 3, 128)	147584
average_pooling2d_2 (Average Pooling2D)	(None, 1, 1, 128)	0
flatten_1 (Flatten)	(None, 128)	0
dense_1 (Dense)	(None, 1024)	132096
dropout_1 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 1024)	1049600
dropout_2 (Dropout)	(None, 1024)	0
dense_3 (Dense)	(None, 7)	7175

Total params: 1,485,831
Trainable params: 1,485,831
Non-trainable params: 0

Fig 4. Model summary

IV. RESULTS AND DISCUSSION

FER2013 dataset is used to train the model, which consists of 28,709 samples of 48x48 pixel grey scale images of faces of different emotion. The proposed CNN model will extract the facial features and according to that it will classify the image to respective class, happy, sad, surprise, fear, neutral and disgust. The model gives the test accuracy of 57.481193% and test loss of 3.743610% in 25 epochs. And also the model gives a train loss of 0.429% and train accuracy of 98.766%. The model categorises the emotion of the image as bar chart by using bar() function of matplotlib API. Live image is captured from the webcam and by using Haar cascade classifier it will remove the unnecessary background and will focus on the face emotion and the model will detect the emotion.

To assess the performance of classification model Confusion matrix or error matrix is designed. This summarized model gives the count of emotion prediction and some insights into predictions.

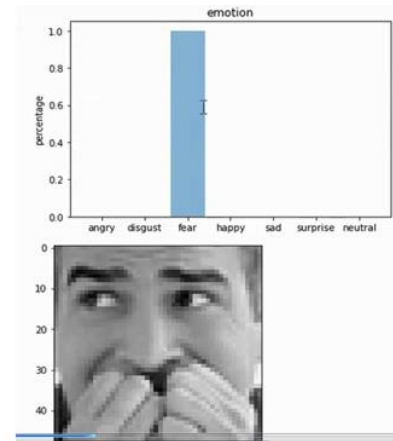
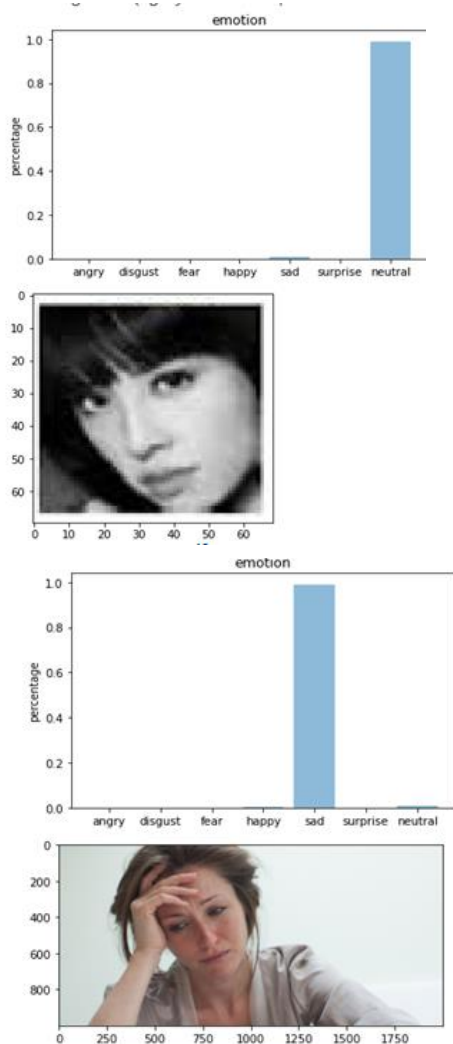
```
array([[234, 1, 62, 36, 57, 17, 60], angry
       [ 12, 26, 1, 2, 10, 1, 4], disgust
       [ 70, 3, 196, 26, 103, 37, 61], fear
       [ 42, 2, 25, 674, 54, 25, 73], happy
       [ 93, 3, 97, 43, 301, 16, 100], sad
       [ 11, 1, 26, 21, 17, 316, 23], surprise
       [ 53, 2, 51, 67, 108, 10, 316]] neutral
```

Fig 5. Confusion matrix

To compare the predicted value to the actual value Classification matrix is designed. Classification matrix is a very important tool to assess the performance of model. Here, in this classification matrix has five columns, which represents the emotion number, precision, recall, f1score and support. The proposed model works really well on positive emotions. The model gives a high precision score of 78% for happy and 75% for surprise. But the model seems to be work week on negative emotion as like fear and angry. The model gives a low precision score of 43% for fear and 45% for angry. But overall the model works well and gives overall accuracy of 98%.

	Precision	recall	f1-score	support	
0	0.45	0.50	0.48	467	angry
1	0.68	0.46	0.55	56	disgust
2	0.43	0.40	0.41	496	fear
3	0.78	0.75	0.76	895	happy
4	0.46	0.46	0.46	653	sad
5	0.75	0.76	0.76	415	surprise
6	0.50	0.52	0.51	607	neutral

Fig 6.Classification matrix



V. CONCLUSION AND FUTURE WORK

In this paper, a 5 layer CNN model is proposed which classify different emotion of face as angry, happy, sad, fear, disgust and neutral. FER2013 dataset is used to train the model, which consists of more than 28,708 images of 48x48 pixel greyscale images. The model gives a good result with an overall test accuracy of 57.481193% and test loss of 3.743610% in 25 epochs. And also the model gives a train loss of 0.429% and train accuracy of 98.766%. The model performs well on positive emotion with a precision score of 78% for happy and 75% for surprise. But the model seems to be work week on negative emotion as like fear and angry. The model gives a low precision score of 43% for fear and 45% for angry. On future this model can be improved by using different dataset and work which focus more on negative emotions.

REFERENCES

- [1] Timo Ojala, Matti PietikaE inen and Topi MaEenpaEaE, "Multiresoution gray-scale and rotation invariant texture classificstion with local binary patterns", in proceedings of the 2002
- [2] IEEE transaction on pattern analysis and machine intelligence, vol. 24
- [3] Timo Ahonen, Abdenour Hadid and Matti Pietik ainen, " Face recognition with local binary patterns", European conference on computer vision, Springer,2004
- [4] Tae-Hyun Kim, Dong-Chul Park, Dong-Min Woo, Taikyeong Jeong and Soo-Young Min, " Multi-class classifier- based Adaboost algorithm", International conference on intelligent sience and intelligent data engineering, Springer ,2011
- [5] Hou Y, Peng Q, " Face detection based on Adaboost and skin colour", In proceedings of international symposium on information science and engineering, 2009
- [6] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features ", in proceedings of the 2001 IEEE conference on computer vision and pattern recognition, Vol. 1
- [7] Misg-Hsuan Yang, "Eigenfaces vs. kernel fisherfaces: face recognition using kernel methods", In proceedings of the fifth IEEE conference on automatic face and gesture recognition, 2002
- [8] C. Szegedy, W. Liu, Y. Jia, P. Sermanent, S. Reed, D. Anguelov, D. Erhan, V.Vanhoucke and A.Rabinovich, "Going deeper with convolutions", IEEE, 2005
- [9] A. Stuhlsatz, C. Meyer, F. Eyben, T. Zielke, G. Meier and B. Schuller , "Deep neural networks for acoustic emotion recognition: raising the benchmarks", In proceedings of the IEEE conference on Acoustics, speech and signal processing, 2011
- [10] Alice J O'Toole, P. Jonathon Phillips, Yi Cheng, Brendan Ross, Heather A Wild, "Face recognition algorithms as models of human face processing", In proceedings of the IEEE transactions on pattern analysis and machine intelligence, vol 26 ,2004

- [11] Jian Yang, David Zhang, Alejandro F. Frangi, and Jing-yu Yang, "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition", IEEE, 2004
- [12] Kwang-Eun Ko, Heuksuk-Dong, Kwee-Bo Sim, "Development of a facial emotion recognition method based on combining AAM with DBN", In proceedings of the international conference on cyberworlds, IEEE, 2010
- [13] G N MATRE, S K Shah, "Facial expression detection", In proceedings of the international conference on computational intelligence and computing research, IEEE, 2013
- [14] Yelin Kim, Honglak Lee, Emily Mower Provost, "Deep learning for robust feature generation in audiovisual emotion recognition", In proceedings of the international conference on acoustics, speech and signal processing, IEEE, 2013
- [15] Sayedehsamaneh Shojaeilangari, Wei-Yun Yau, Karthik Nandakumar, Li Jun, Eam Khwang Teoh, "Robust representation and recognition of facial emotion using extreme sparse learning", In proceedings of the IEEE transaction on image processing, 2015
- [16] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", In proceedings of the IEEE conference on computer vision and pattern recognition, IEEE, 2016
- [17] Neha Shirbhate, Kiran Talele, "Human body language understanding for action detection using geometric features", In proceedings of the IEEE conference on contemporary computing and informatics, 2016
- [18] A Yao, D Cai, P Hu, S Wang, L Shan, Y Chen, "Holonet: towards robust emotion recognition in the wild", In proceedings of the ACM international conference on multimodal interaction, 2016
- [19] Yadan L V, Zhiyong Feng, Chao Xu, "Facial expression recognition via deep learning", IEEE, 2014
- [20] K. Li, G. Ding, and H. Wang, "L-fcn: A lightweight fully convolutional network for biomedical semantic segmentation", In proceedings of the IEEE conference on bioinformatics and biomedicine, 2018
- [21] Md Tohidul Islam, B M Nafiz Karim Siddique, Sagidur Rahman, Taskeed Jabid, "Image recognition with deep learning", In proceedings of the IEEE conference on intelligent informatics and biomedical sciences, 2018
- [22] Sunitha Ravi, M Suman, P V V Kishore, Kiran Kumar "Multi modal spatio temporal co-trained CNNs with single modal testing on RGB-D based sign language gesture recognition." Journal of Computer Languages, 2019
- [23] Zhong-Qiu Zhao, Shou-Tao Xu, Xindong Wu, "Object detection with deep learning: A review", In proceedings of the IEEE transactions on neural networks and learning system
- [24] R Santhoshkumara, M. Kalaiselvi Geethab, "Deep Learning Approach for Emotion Recognition from Human Body Movements with Feedforward Deep Convolution Neural Networks", In proceedings of the International Conference on Pervasive Computing Advances and Applications – PerCAA 2019
- [25] Kaviya P, Arumugaprakash T, "Group facial emotion analysis system using convolutional neural network Group facial emotion analysis system using convolutional neural network", In proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
- [26] D.S.L.Parvathi, N.Leelavathi, J.M.S.V.Ravikumar, B.Sujatha, "Emotion analysis using deep learning", In proceedings of the International Conference on Electronics and Sustainable Communication Systems, 2020
- [27] Çağlar Genç, Ashley Colley, Markus Löchtfeld, Jonna Häkkinen, "Face Mask Design to Mitigate Facial Expression Occlusion", In Proceedings of the 2020 International Symposium on Wearable Computers (pp. 40–44). Association for Computing Machinery, 2020
- [28] Amrit Kumar Bhadani, Anurag Sinha, "Facemask Detector Using Machine Learning And Image Processing Techniques", Elsevier, 2020
- [29] Paul Viola, Michael J Jones, "Robust Real-Time Face Detection", Springer, 2004
- [30] Zhong-Qiu Zhao, Peng Zheng, Shou-Tao Xu, and Xindong Wu, "Object Detection With Deep Learning: A Review", IEEE, 2019
- [31] Md Tohidul Islam, B.M. Nafiz Karim Siddique, Sagidur Rahman, Taskeed Jabid, "Image Recognition with Deep Learning", IEEE, 2018
- [32] R. Santhosh kumar, M. Kalaiselvi, "Deep Learning Approach for Emotion Recognition from Human Body Movements with Feedforward Deep Convolution Neural Networks", Elsevier, 2019
- [33] Neha Shirbhate, Kiran Talele, "Human Body Language Understanding for Action Detection using Geometric Features", IEEE, 2016
- [34] Dara, Suresh, Vaishnavai Pasari, Nageswara Rao Banoth, and JMSV Ravi Kumar, "Artificial Bee Colony Algorithm: A Survey and Recent Applications", International Journal of Pure and Applied Mathematics, 2018
- [35] C. Tomasi S.B. Goturk, J.-Y. Bouguet and B. Girod, "Model-based face tracking for view-independent facial expression recognition", IEEE, 2002
- [36] G. N. Matre, S. K. Shah, "Facial Expression Detection", IEEE, 2013
- [37] Jian Yang, David Zhang, Alejandro F. Frangi, and Jing-yu Yang. "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition", IEEE, 2004
- [38] D. S. L. Parvathi, N. Leelavathi, J. M. S. V. Ravikumar, B. Sujatha. "Emotion Analysis Using Deep Learning", IEEE, 2020
- [39] Shan Li and Weihong Deng. "Deep Facial Expression Recognition: A Survey" IEEE, 2020
- [40] Vaibhavkumar J. Mistry and Mahesh M. Goyani. "A literature survey on Facial Expression Recognition using Global Features", International Journal of Engineering and Advanced Technology (IJEAT), 2013
- [41] Y.-I. Tian, T. Kanade, and J. F. Cohn, "Recognizing action units for facial expression analysis," IEEE Transactions on pattern analysis and machine intelligence, 2001
- [42] Convolutional Neural Networks (CNN) With TensorFlow by Sourav from Edureka [https://www.youtube.com/watch?v=umGJ30-15_A]