Proceedings of the International Conference on Electronics and Renewable Systems (ICEARS 2022)
IEEE Xplore Part Number: CFP22AV8-ART; ISBN: 978-1-6654-8425-1

Deep Learning Framework for Facial Emotion Recognition using CNN Architectures

Mr. Rohan Appasaheb Borgalli
Research Scholar, Department of Electronics Engineering
Fr. Conceicao Rodrigues College of Engineering, Bandra, Mumbai University
Mumbai, India
rohanborgalli111@gmail.com

Dr. Sunil Surve
Professor, *Department of Computer Engineering*
Fr. Conceicao Rodrigues College of Engineering, Bandra, Mumbai University
Mumbai, India
surve@fragnel.edu.in

2022 International Conference on Electronics and Renewable Systems (ICEARS) | 978-1-6654-8425-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICEARS53579.2022.9751735

Abstract—**FER (facial expression recognition) is a significant study subject in the artificial intelligence and computer vision areas because of its widespread applicability in both academic and industrial sectors. Though FER can be carried out primarily utilizing multiple sensors, research shows that using facial images/videos for recognition of facial expression is better because visual expressions carry major information through which emotions can be conveyed.**

In the past, much research was conducted in the field of FER using different approaches such as the use of different sensors, machine learning, and deep learning framework with dynamic sequences and static images. The most recent state-of-the-art outcomes demonstrate In comparison to conventional FER techniques, deep learning Convolutional Neural Network (CNN) based systems are significantly more powerful. Deep learning-based FER methods utilizing deep networks enable extraction of features automatically instead of traditionally handcrafted feature extraction.

This paper focuses on implementing different custom and standard CNN architectures for training and testing them on facial expression static image datasets scenario KDEF, RAFD, RAF-DB, SFEW, and AMFED+, both lab-controlled and wild.

Keywords—Deep learning, Facial expression recognition, Convolution neural networks, Machine learning.

I. INTRODUCTION

FER (Facial Expression Recognition) is an essential research area since it may be used in a range of applications. A FER system model detects the human facial expression and identifies the corresponding induced emotion. Despite much study, accurately identifying facial emotions with high accuracy continues to be a tough task because of the complexity as well as variety of facial expressions.

According to Darwin &Prodger [1], human beings' facial expressions reflect their emotional states and intentions. Due to its importance in applications such as machine vision and machine learning, numerous attempts have been made in the past to implement a FER system practically. People may utilize this approach to communicate nonverbally and gain a lot of benefit from it. Detecting or extracting facial expressions from images is the most crucial part of the system. The system is growing attention due to being widely utilized in medical assessment, lie detection, human-machine interaction, driver safety, and solving other complex real-world problems with FER.

To implement the FER system effectively, it should work accurately on lab-controlled environment images and wild environments. Many researchers have developed datasets [2]-[6] to move the FER system to a wild scenario.

In this paper, we implemented different custom and standard CNN-based learning algorithms for facial expression recognition systems on different basic emotions datasets to achieve good performance on a particular dataset irrespective of whether it is lab-controlled or not.

II. LITERATURE SURVEY

In literature, The majority of facial expression recognition systems that have been suggested and deployed have been built on datasets that have been collected in a laboratory setting such as FER13 [7], CK+ [8], JAFFE [9], and many others, where images of faces are captured in frontal along with without any obstruction. Nevertheless, the FER systems that work well in the lab may not operate well in uncontrolled or wild situations. In order to fulfill this FER system gap in wild scenarios and lab-controlled datasets in the wild [2]-[6] are made available. Despite the fact that the use of data from the wild for facial expression identification has risen, it remains a difficult task owing to the presence of non-frontal and partially obscured faces.

Researchers Shan Li et al.[2] have developed a new real-world facial expression database, RAF-ML, to investigate and overcome the challenges of facial expression identification in the wild. DeepBi-Manifold (DBM)-based RAF-ML performs similarly on cross-database comparisons. Using CNN, it was discovered that RAF-ML can be used as a 'generic' database for facial expression analysis training data due to the wide variety of characteristics it has.
XIA et al. [10] used a successful facial recognition model based on the Inception-v3 model in TensorFlow.The Inceptionv3 model was retrained with facial data using a transfer learning strategy that reduced training time as much as feasible.

There are well-established methods with both handcrafted and automated features extraction through deep learning [11] for FER. Among them, the CNN-based [12] approach has proved suitable for image classification-based applications. The use of standard CNN architectures gives state-of-the-art outcomes in FER [11]. The fundamental advantage of a CNN is that it permits end-to-end learning directly from an input source and reduces or eliminates the need for physics-based models and/or other pre-processing approaches utilized for FER applications.

Past research on detecting seven basic emotions: fear, neutral, surprise, disgust, sad, anger, contempt, and happy is mainly focused on lab-controlled datasets. This lack of durability is demonstrated by the fact that most FER solutions are overly broad. Hence, there is a need for an efficient and robust real-life FER system that detects particular basic facial expressions regarding the environment.

We proposed a deep learning framework to correctly recognize basic human emotion on existing lab-controlled as well as wild datasets to build robust and efficient Facial Expression Recognition (FER) systems on KDEF [13], RAFD [14], RAF-DB [3], SFEW [4], and AMFED+ [5].

III. Dataset Overview

A. *Karolinska Directed Emotional Faces (KDEF) dataset [13]*

KDEF (Karolinska Directed Emotional Faces) is a dataset of human facial expressions with 4900 pictures of 70 individual subjects displaying seven different facial expressions, with each expression viewed from 5 different angles.

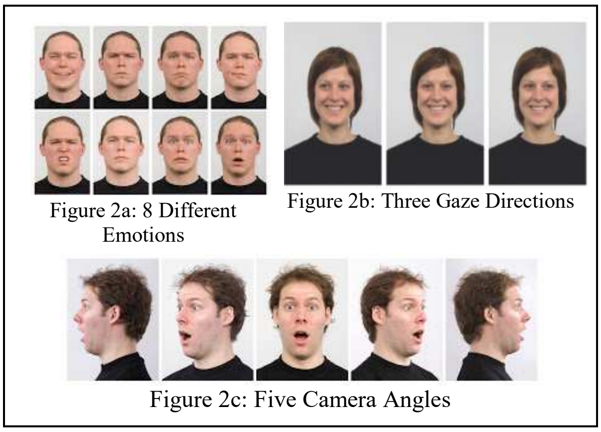
The population of subjects has 70 amateur actors of 35 females and 35 males. Subjects are selected with criteria of age between 20 and 30 years. During photo sessions, subjects are with no mustaches, beards, eyeglasses, or earrings, and preferably no visible makeup.

Fig. 1. KDEF Dataset Sample Images

B. *Radboud Faces Database (RaFD)[14]*

RaFD database contains two categories of subject's 19 female (39 Caucasian Dutch adults) along with 6 female children (10 Caucasian Dutch) with portrait images of 49 models. Eight different facial expressions and three different gaze orientations are illustrated in Figures 2a and 2b, which were performed by all of the models. Expressions were neutral, surprise, anger, disgust, fear, sad, contempt, and happy. At the same time, each emotion was exhibited in a variety of orientations, such as looking straight ahead, looking to the left, as well as to the right against a white background from 5 various camera angles. (see Figure 2c). Black t-shirts, no makeup, no hair on the face, along with no jewellery were all part of each model's 120 photographs in the collection.

Fig. 2. RAFD Dataset Sample Images



C. *Real-world Affective Faces Database (RAF-DB)[3]*

Around 30K well-diversified face images make up the RAF-DB(Real-world Affective Faces Database), a large-scale facial expression database collected from the Internet. Every image was labeled independently based on the crowdsourcing annotation by around 40 annotators. In this database, images are full of variability in subjects, their post-processing operations (for example, special effects and various filters), occlusions (for example, self-occlusion, facial hair, glasses), lighting conditions, head poses, ethnicity, gender, and age, etc. RAF-DB has large diversities rich annotated images in large quantity, including 29672 real-world images with seven basic expression distributions for each image.

Training, as well as test sets, were compared using phrases that were uniformly scattered across both sets in order to establish an objective measure of performance.

Fig. 3. RAF-DB Dataset Sample Images

D. *Afectiva-MIT Facial Expression Dataset (AM-FED+) [4]*

The AM-FED+ collection is comprised of 1,044 facial videos (.flv) obtained in the wild and captured in real-world situations. All the frames of each video have been manually annotated in FACS (Facial Action Coding System) coded. For all the videos, 34 automatically detected facial landmark locations are provided for every frame.

There are gender identifiers in every movie, and a significant portion (77 %) also includes information about age and country.

Fig. 4. AMFED+ Dataset Sample Images

E. *SFEW (Static Facial Expression in the Wild)[5]*

We have a facial expression dataset called SFEW "Static Facial Expressions in the Wild." Choosing static frames from the AFEW database of acted facial expressions resulted in this result. For the SReco sub-challenge in EmotiW 2015, we are employing the most regularly used version, which is SFEW 2.0. SFEW 2.0 was categorized into three main sets: Test (372 samples), validation (436 samples), along with Train (958 samples). Seven expression types are assigned to every image, i.e., disgust, neutral, surprise, anger, sad, contempt, happy and fear,

Fig. 5. SFEW Dataset Sample Images

IV. Proposed Methodology

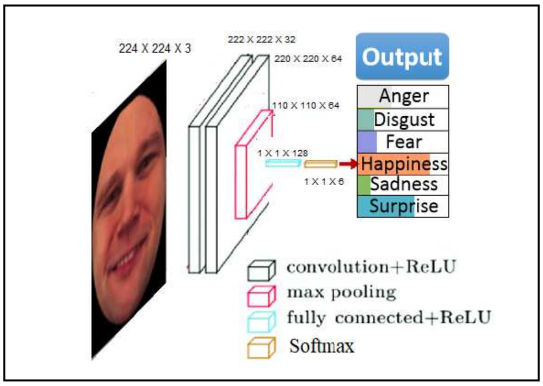
A. Custom CNN Architectures

The proposed methodology is based on a deep learning framework using Custom CNN Architectures. which has an input layer to take a facial image as an input which is then processed through a series of convolutional layers used for extracting features and max-pooling layers, which are used for reducing the size of extracted convolution feature and last layers some fully-connected layers and an output layer to learns task-specific features in our case its basic emotions such as neutral, surprise, anger, disgust, fear, sad, contempt, and happy

Custom CNN Architecture for Basic Emotion detection takes color RGB images of size (224,224) and generates softmax output for six classes indicating particular basic emotion in a given image. Fig. 6 shows Custom CNN architecture for basic emotion detection on the KDEF dataset.

Custom CNN Architecture for basic emotion detection on KDEF Dataset has five layers with 99.143 M Parameters

Fig. 6. Custom CNN Architecture for Basic Emotion detection on KDEF Dataset



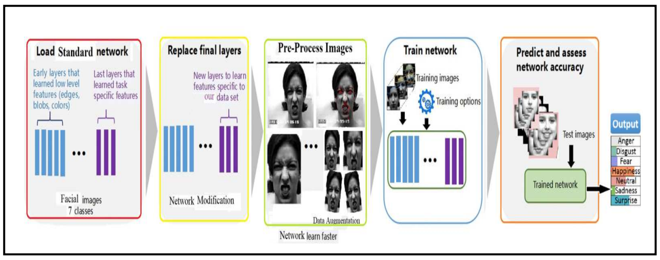
B. Standard CNN Architectures

The proposed methodology is based on using a standard Convolution Neural network-based deep learning framework. We use standard CNN architectures such as VGG and InceptionnetV2, which are well known for Image classification applications. These Standard Networks have starting layers that are utilized for to learning low-level features, for example, colors, blobs, edges. Along with the last layers, learn task-specific features, in our case, its Facial basic Emotions.

By considering this fact, we are modifying the standard network by replacing the last layers so that features are specific to our dataset.

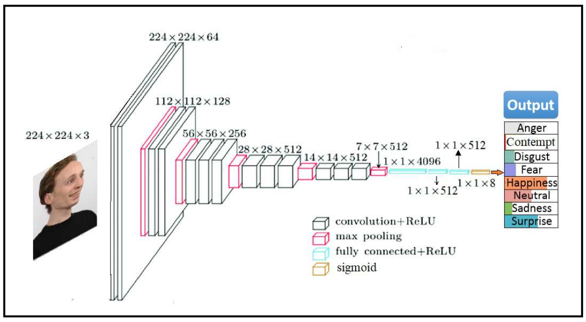
The database has a variety of images but to learn features faster. We are pre-processing Images, so only the required part of a face is captured using the Haar Cascade algorithm for detection of a frontal face. The proposed methodology for basic emotion detection using standard CNN architecture is shown in Fig. 7.

Fig. 7. Proposed Methodology for Basic Emotion detection using Standard CNN Architecture



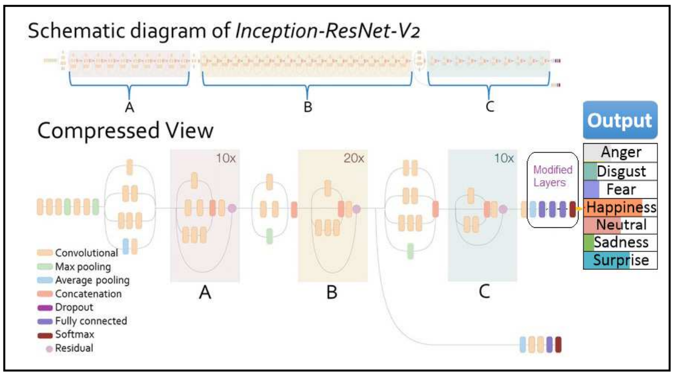
For the RaFD dataset to detect basic emotion VGG based architecture is used. It takes color RGB images of size (224,224) and generates softmax output for eight classes indicating particular basic emotion induced in the image. Fig. 8 indicates detailed VGG based CNN Architecture for basic emotion detection on the RaFD Dataset.

Fig. 8. VGG Based Architecture for Basic Emotion detection on RaFD Dataset



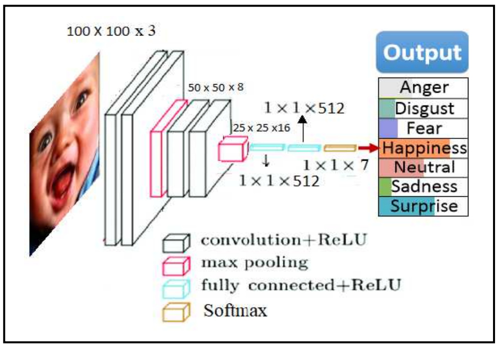
InceptionResNetV2 based CNN Architecture used for SFEW & AMFED+ dataset to detect basic emotion. It takes color RGB images of size (224,224) and generates softmax output for seven classes indicating particular basic emotion induced in the image. Fig. 9 indicates a schematic diagram of standard InceptionResNetV2 architecture and a compressed view of the same for basic emotion detection on SFEW and AMFED+ datasets.

Fig. 9. InceptionResNetV2 based CNN architecture emotion detection on SFEW & AMFED+ datasets



For the RAF-DB dataset, VGG based architecture is used for basic emotion detection. It takes color RGB images of size (100,100) and generates softmax output for 7 classes indicating particular basic emotion in a given image. Fig. 10 indicates detailed LIGHTVGG based CNN Architecture for basic emotion detection on RAF-DB Dataset. VGG based architecture for basic emotion detection takes color RGB images of size (100,100) and generates softmax output for 7 classes indicating particular basic emotion in a given image. Fig. 10 indicates detailed LIGHTVGG based CNN architecture for basic emotion detection on the RAF-DB dataset.

Fig. 10. LIGHTVGG Based Architecture for Basic Emotion detection on RAF-DB Dataset



A. MultiClass Confusion Matrix

Algorithm performance is shown via a table structure known as a confusion matrix or an error matrix. Each matrix row shows the predicted class, whereas each column shows the actual class (or vice versa). As the name implies, this indicates whether or not a system is confused between two different classes (such as usually mislabeling one as another). We utilized a confusion matrix for evaluating the accuracy between actual labels and predicted labels.

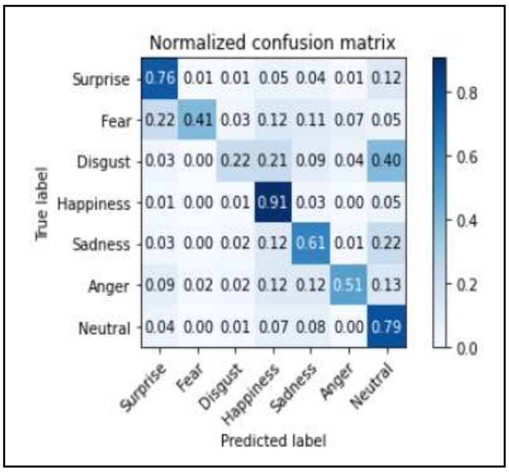
B. Accuracy Score

To evaluate model accuracy, a score is calculated, which has recall, precision, along with f1-score for every class on test samples, indicating the accuracy of each class predicted and overall accuracy.

1) Confusion Matrix of RAF-DB Database

RAF-DB basic emotion database has 7 Basic Facial emotions {neutral, surprise, anger, disgust, fear, sad, contempt, and happy}. Fig. 11 shows the normalized Confusion Matrix for RAF-DB Database indicates most of the classes are classified with good accuracy except disgust. Disgust is misclassified as neutral for the majority of cases. Accuracy score indicates recall, precision, along with an f1-score for every class on test samples Indicating overall accuracy of 75%.

Fig. 11. Confusion Matrix and accuracy report for RAF-DB Database

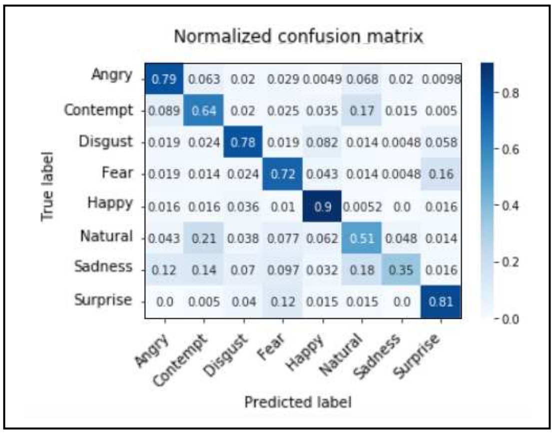


Accuracy Score :				
class	precision	recall	f1-score	support
0	0.74	0.76	0.75	329
1	0.81	0.41	0.54	74
2	0.51	0.22	0.31	160
3	0.85	0.91	0.88	1185
4	0.68	0.61	0.64	478
5	0.76	0.51	0.61	162
6	0.65	0.79	0.71	680
accuracy			0.75	3068
macro avg	0.71	0.60	0.63	3068
weighted avg	0.74	0.75	0.74	3068

2) Confusion Matrix of RAFD Database

RAFD Database has 8 Basic Facial emotions{neutral, surprise, anger, disgust, fear, sad, contempt, and happy }. Fig. 12 shows that the normalized Confusion Matrix for RAFD Database indicates that most classes are classified with good accuracy except sadness. Sadness is misclassified as Neutral, Contempt, and Angry for a majority of cases. Accuracy score indicates recall, precision, along with f1-score for every class on test samples, Indicating overall accuracy of 69%.

Fig. 12. Confusion Matrix and Accuracy report for RAFD Database

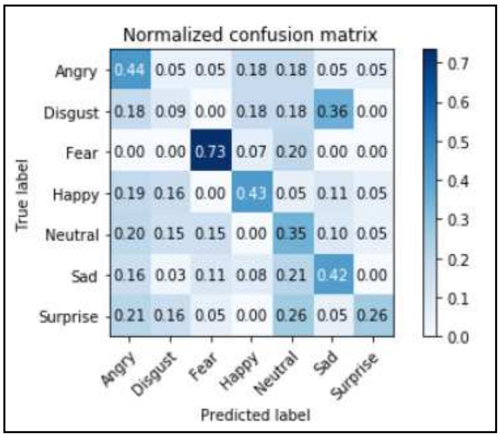


Accuracy Score :				
class	precision	recall	f1-score	support
0	0.73	0.79	0.76	205
1	0.58	0.64	0.61	202
2	0.77	0.78	0.77	208
3	0.67	0.72	0.69	207
4	0.76	0.90	0.82	193
5	0.53	0.51	0.52	208
6	0.77	0.35	0.48	185
7	0.74	0.81	0.77	200
accuracy			0.69	1608
macro avg	0.69	0.69	0.68	1608
weighted avg	0.69	0.69	0.68	1608

3) Confusion Matrix of SFEW Database

SFEW Database has 7 Basic Facial emotions {neutral, surprise, anger, disgust, fear, sad, contempt, and happy }. Fig. 13 shows that the normalized Confusion Matrix for SFEW Database indicates that most classes are classified with good accuracy except Disgust and Surprise. Disgust is misclassified as Neutral, Happy, and Angry for most cases. Also, surprise is misclassified as Neutral and Angry for most cases. Accuracy score shows recall, precision, along with f1-score for every class on test samples Indicating overall accuracy of 41%.

Fig. 13. Confusion Matrix and Accuracy report for SFEW Database



Accuracy Score :				
class	precision	recall	f1-score	support
0	0.42	0.44	0.43	39
1	0.06	0.09	0.07	11
2	0.52	0.73	0.61	15
3	0.55	0.43	0.48	37
4	0.21	0.35	0.26	20
5	0.55	0.42	0.48	38
6	0.50	0.26	0.34	19
accuracy			0.41	179
macro avg	0.40	0.39	0.38	179
weighted avg	0.45	0.41	0.42	179

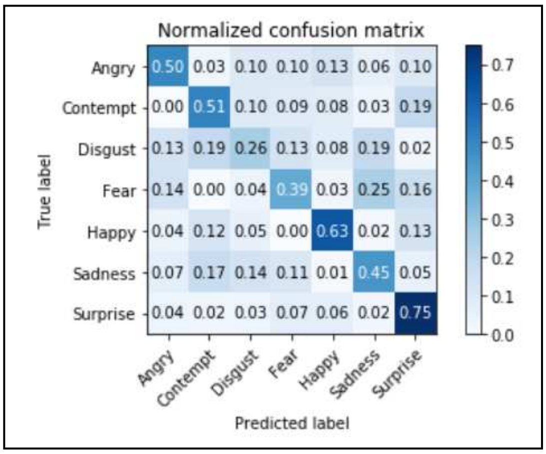
4) Confusion Matrix of AMFED+ Database

AMFED+ Database has 7 Basic Facial emotions { *neutral, surprise, anger, disgust, fear, sad, contempt, and happy* }. Fig. 14 shows the normalized Confusion Matrix for AMFED+ Database indicates most of the classes are classified with good accuracy except Disgust and Fear. Disgust is misclassified as Contempt, Fear, and sadness in a few cases, and fear is misclassified as sadness in a few cases. Accuracy score indicates recall, precision, along with f1-score for every class on test samples, Indicating overall accuracy of 54%.

5) Confusion Matrix of KDEF Database

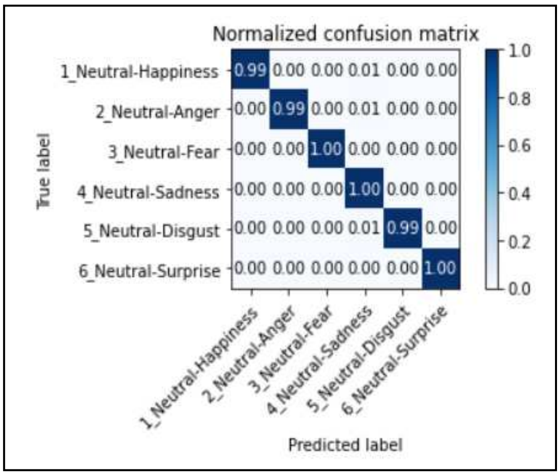
KDEF Database has 6 Basic Facial emotions {Surprise, Sadness, Fear, Anger, Happiness, and Disgust}. Fig. 15 shows the normalized Confusion Matrix for KDEF Database indicates most of the classes are classified with good accuracy. Accuracy score indicates recall, precision, along with f1-score for every class on test samples, Indicating overall accuracy of 100%.

Fig. 14. Confusion Matrix and Accuracy report for AMFED+ Database



Accuracy Score:				
class	precision	recall	f1-score	support
0	0.61	0.50	0.55	145
1	0.48	0.51	0.49	122
2	0.30	0.26	0.28	100
3	0.30	0.39	0.34	80
4	0.59	0.63	0.61	128
5	0.51	0.45	0.48	133
6	0.71	0.75	0.73	247
accuracy			0.54	955
macro avg	0.50	0.50	0.50	955
weighted avg	0.54	0.54	0.54	955

Fig. 15. Confusion Matrix and Accuracy report for KDEF Database

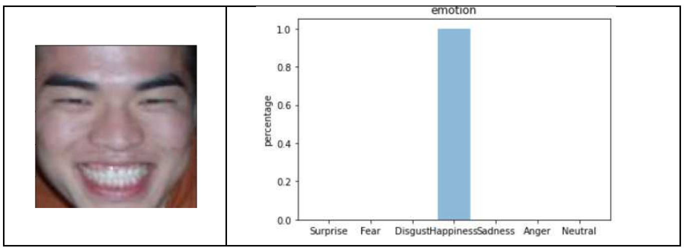


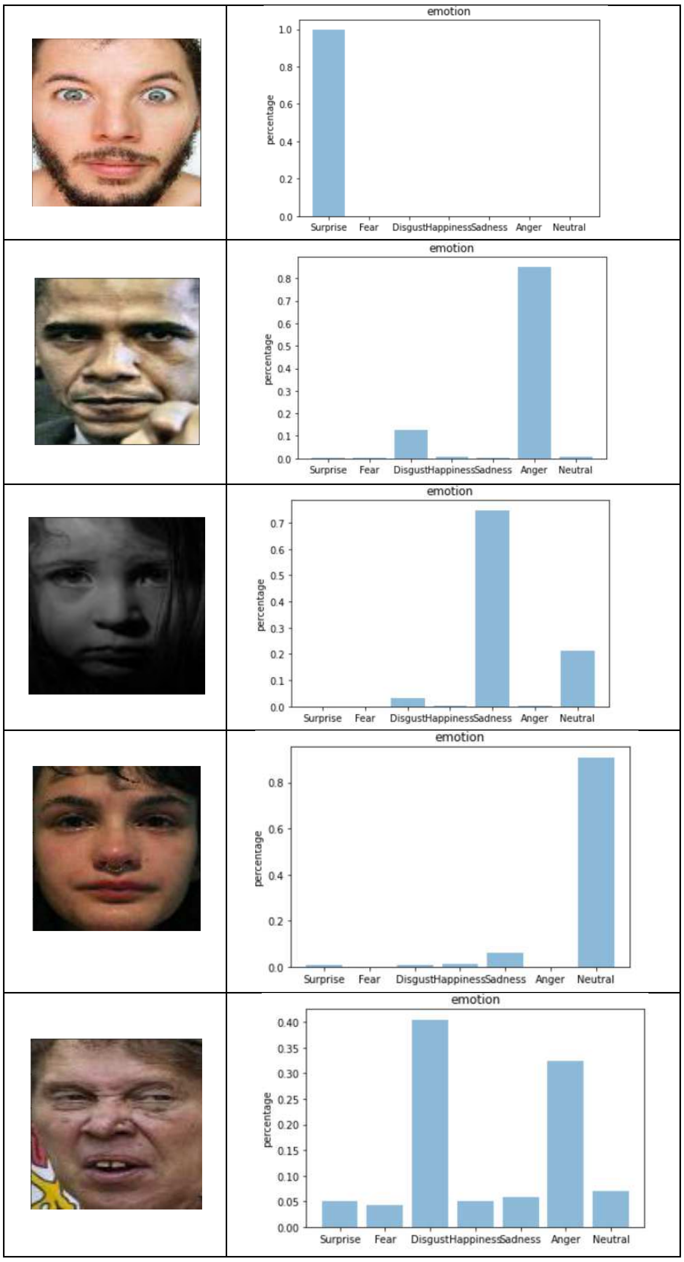
Accuracy Score :				
class	precision	recall	f1-score	support
0	1.00	0.99	1.00	154
1	1.00	0.99	1.00	154
2	1.00	1.00	1.00	158
3	0.98	1.00	0.99	142
4	1.00	0.99	1.00	140
5	1.00	1.00	1.00	52
accuracy			1.00	800
macro avg	1.00	1.00	1.00	800
weighted avg	1.00	1.00	1.00	800

C. *Emotion Analysis on Sample Test Image*

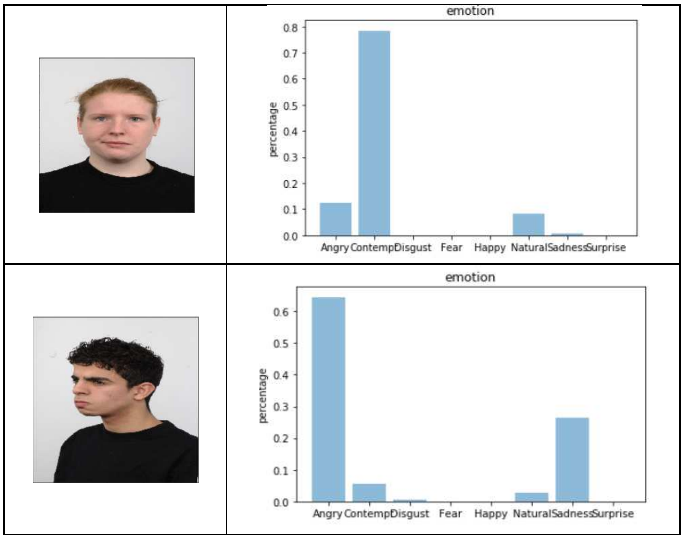
Emotion Analysis of Sample test Images is done at the end to check models' working from RAF-DB, RAFD, KDEF, SEFW, and AMFED+ Dataset. It shows for sample image softmax output indicating the probability of each emotion with a plotted graph the maximum will be considered as predicted emotion by the proposed model.

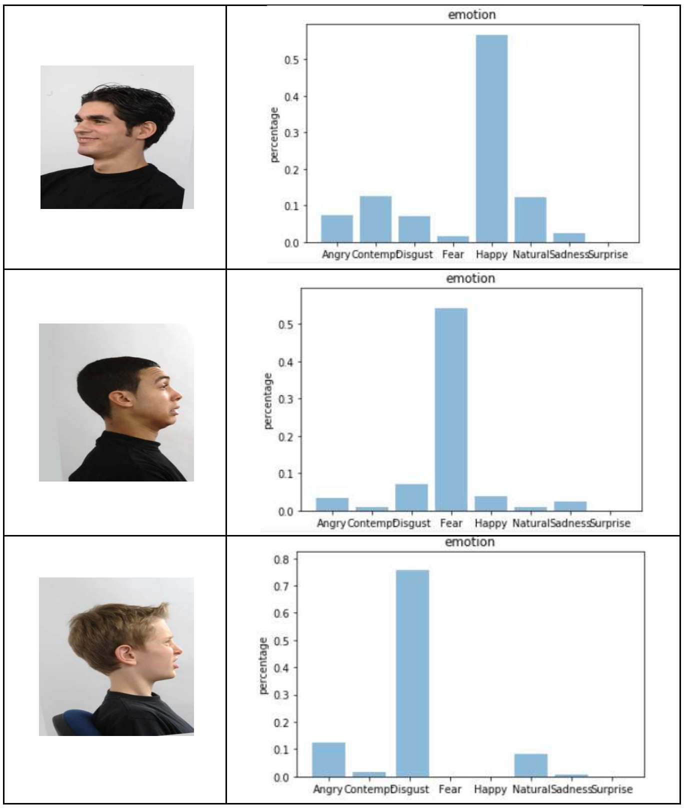
1) Emotion Analysis on RAF-DB Sample Test Image



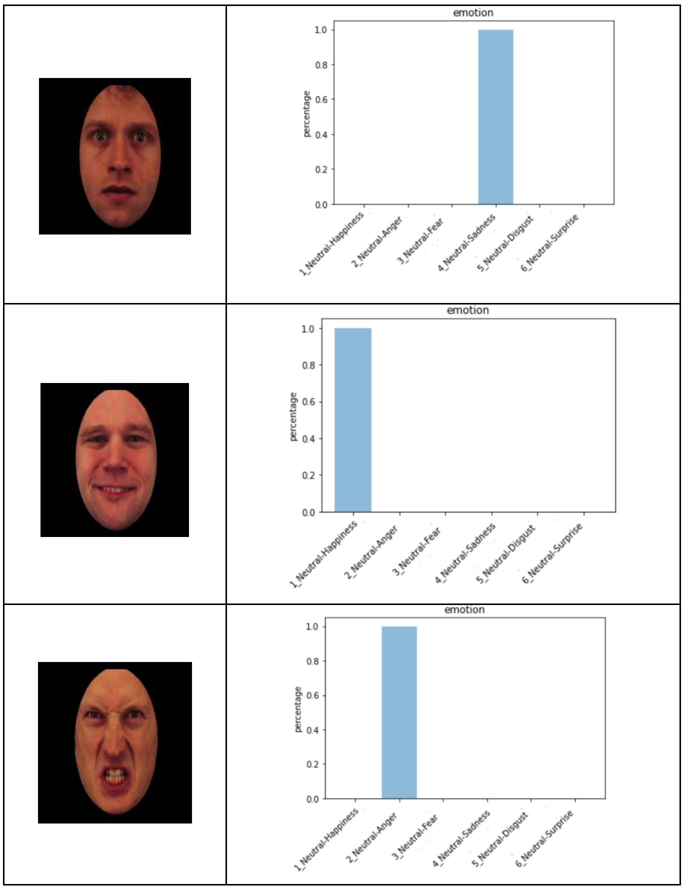


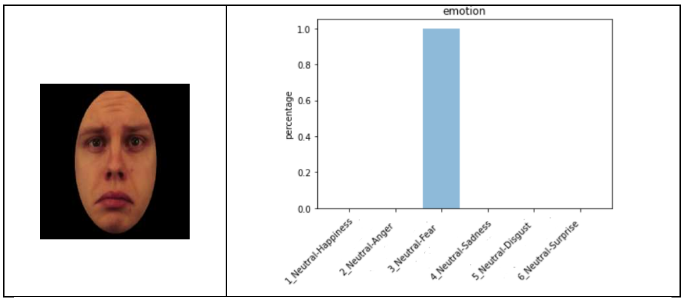
2) Emotion Analysis on RAFD Sample Test Image



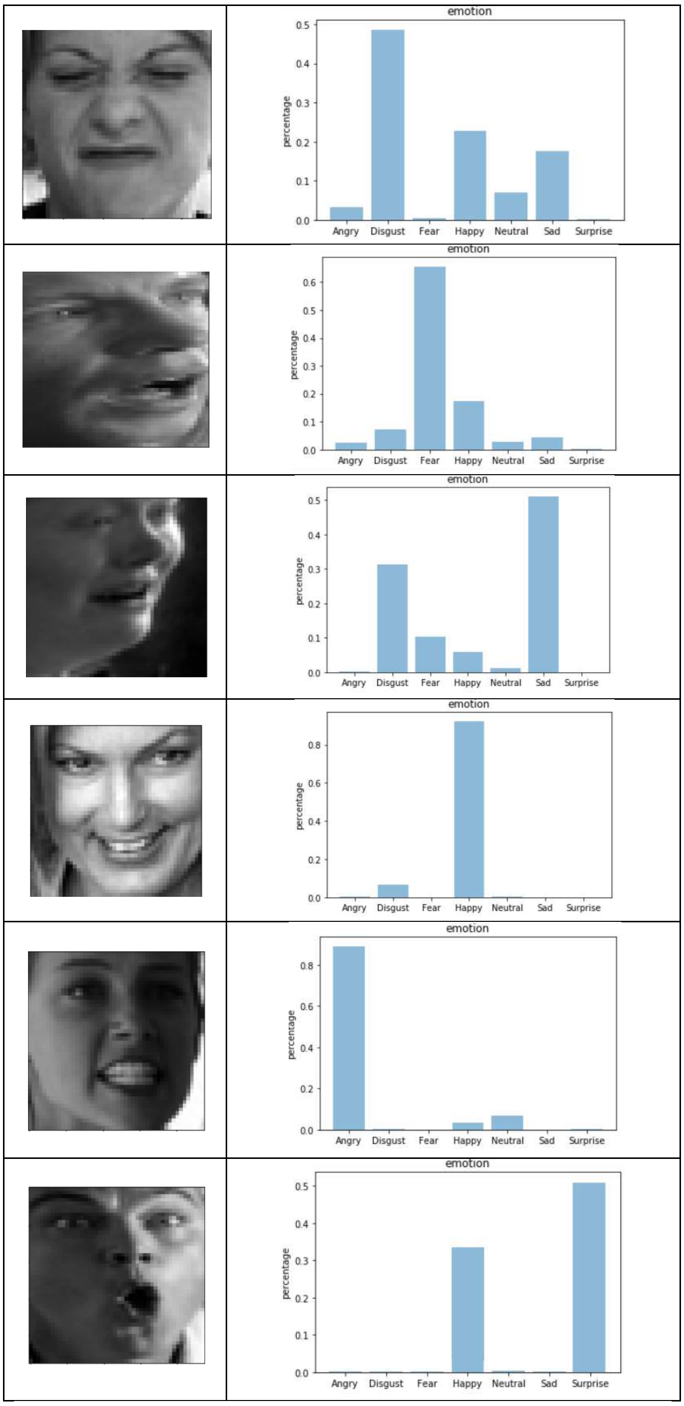


3) Emotion Analysis on KDEF Sample Test Image

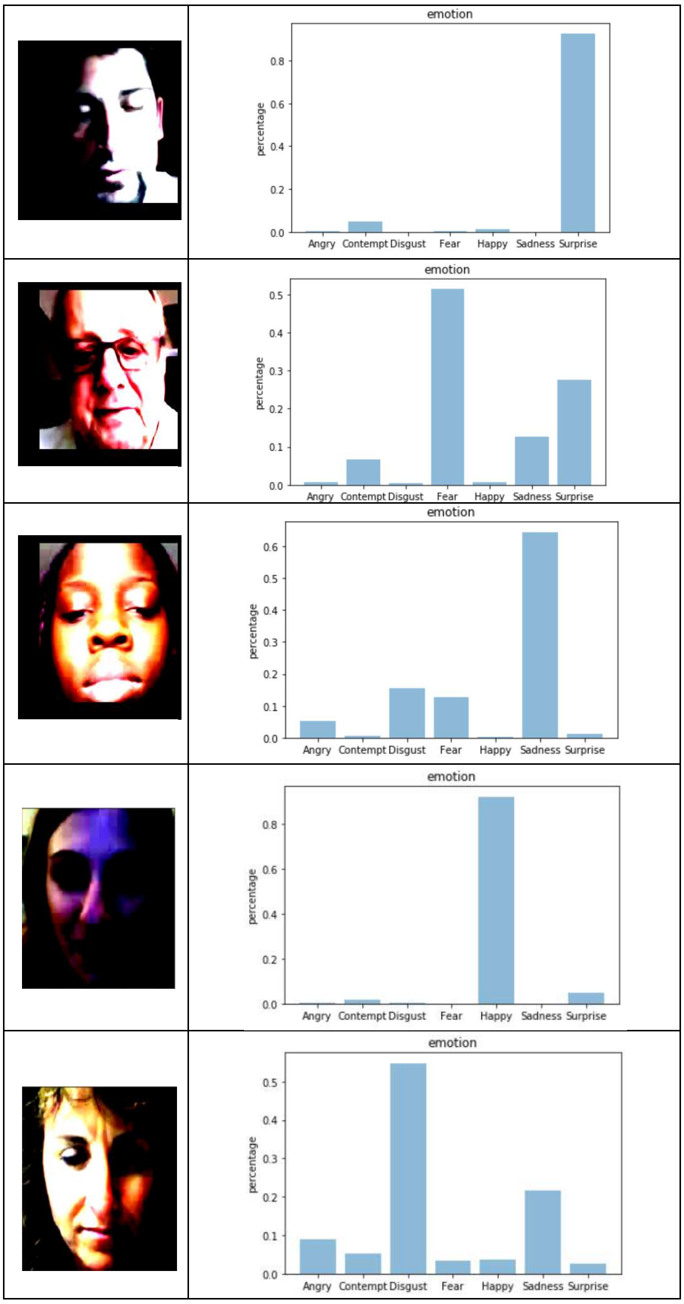




4) *Emotion Analysis on SEFW Sample Test Image*



5) *Emotion Analysis on AMFED+ Sample Test Image*



To summarize the results of implemented methodology, refer to Table 1, which indicates the result on basic emotion datasets such as RAFD, SFEW, AMFED+, RAF-DB, and KDEF. Few datasets are lab-controlled, and the remaining are real-life images taken from the web, which are not lab-controlled to get a wide variety of images.

A learning methodology based on the Custom and Standard CNN Model is adopted to get better results. Standard CNN architecture is based on InceptionResnetV2 and VGG architectures. To get an idea about the complexity of the network, no. of parameters are specified. For Pre-processing, the Viola-Jones algorithm for detection of frontal faces is used, which is based on Haar Cascade and to balance training images for unbiased results and generate more training, Image DataGenerator is used that performs a different operation on a given image, such as flipping, rotation, cropping, blurring, and adding noise to generate more images. There are six to eight fundamental emotion classes that the model is taught from begin. These datasets have facial images/videos with class labels indicating basic emotion class. Dataset with videos is converted into images by taking frames. To train models, these datasets are categorized into training, validation, along with test data groups.

To measure performance, the model is trained on training as well as validation dataset on each epoch, and accuracy and loss are calculated for each epoch. Once training is done, indicating no more improvement of accuracy on the validation dataset. The trained model is finally used to calculate accuracy performance (%) on the test dataset.

These results indicate that many factors affect the accuracy of detecting basic emotions from facial images. A few of them are the type of dataset, number of images, number of classes, the model architecture used, and training algorithm. Accuracy is good for the lab-controlled dataset as images are clearer, focused and all facial emotions are well distinguished. On the other hand, real-life images are non-lab-controlled. They are taken from the web. So, these images are not clear, focused and all facial emotions are not well distinguished.

KDEF shows outstanding results and almost 100% accuracy. The reason is images are lab-controlled, with only the facial part visible. The other part is masked and only has 6 Classes. Model architecture has more parameters.

TABLE 1: SUMMARY RESULTS ON VARIOUS FACIAL BASIC EMOTION DATASETS WITH CNN CUSTOM & STANDARD MODEL

Datasets	Lab Controlled	Model	Parameters	Accuracy (%)
RAFD	Yes	VGG16	15.244M	68.84%
SFEW	No	InceptionResNetV2	56.568M	40.78%
AMFED+	No	InceptionResNetV2	56.568M	54.13%
RAF-DB	No	LIGHT VGG	5.391M	75.26%
KDEF	Yes	Custom Model	5 Layers 99.143 M	99.63%

VI. CONCLUSION AND FUTURE WORK

Using unique and conventional CNN architecture, this study provides a deep learning framework for f ace emotion identification. The custom and standard CNN architecture are used to get better performance on a given database based on the types of a dataset, a number of images, and classes.

Experiment shows in lab-controlled dataset RAFD and KDEF got an accuracy of 68.84% and 99.63%, respectively. The custom model proposed for the KDEF dataset is outperformed with state-of-the-art results. On non lab-controlled dataset RAF-DB, SFEW, and AMFED+ got accuracy 75.26%, 40.78%, and 54.15% respectively. Comparatively, wild facial images in a non-lab controlled dataset accuracy are less compared to a lab-controlled dataset.

We may use an unsupervised pre-training technique from transfer learning in the future to increase the accuracy of non-lab-controlled datasets, which may reduce the recognition error. Pre-processing and feature extraction techniques can be applied before training models [11],[12]. Transfer learning [15] can also be tried on these datasets. FER based on Facial Action Units(AUs) showed promising results in a state of the art that can also be tested [16]. Some Advanced Models likeDBN (Deep Belief Network)along withGANs “Generative Adversarial Networks” can be applied to improve accuracy [17],[18].

References

[1] Darwin, C., & Prodger, P. (1998). The expression of the emotions in man and animals. Oxford University Press, USA.

[2] Li, S., Deng, W., & Du, J. (2017). Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2852-2861).

[3] Li, S., & Deng, W. (2018). Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition. IEEE Transactions on Image Processing, 28(1), 356-370.

[4] Dhall, A., Goecke, R., Lucey, S., & Gedeon, T. Static facial expressions in the wild: data and experiment protocol. CVHCI Google Scholar.

[5] McDuff, D., Kaliouby, R., Senechal, T., Amr, M., Cohn, J., & Picard, R. (2013). Affectiva-mit facial expression dataset (am-fed): Naturalistic and spontaneous facial expressions collected. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 881-888).

[6] Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, *10*(1), 18-31.

[7] Wolfram Research, "FER-2013" from the Wolfram Data Repository (2018)

[8] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010, June). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 ieee computer society conference on computer vision and pattern recognition-workshops (pp. 94-101). IEEE.

[9] Lyons, Michael, Kamachi, Miyuki, & Gyoba, Jiro. (1998). The Japanese Female Facial Expression (JAFFE) Dataset.

[10] Xia, X.L., Xu, C. and Nan, B., 2017. Facial expression recognition based on tensorflow platform. In ITM Web of Conferences (Vol. 12, p. 01005). EDP Sciences.

[11] S. Li, W. Deng, Deep facial expression recognition: a survey, arXiv preprint.(2018) arXiv:1804.08348.

[12] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25, 1097-1105.

[13] Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: a validation study. Cognition and emotion, 22(6), 1094-1118.

[14] Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. Cognition and emotion, 24(8), 1377-1388.

[15] Ramalingam, S. and Garzia, F., 2018, October. Facial Expression Recognition using Transfer Learning. In 2018 International Carnahan Conference on Security Technology (ICCST) (pp. 1-5). IEEE.

[16] Kim, S., & Kim, H. (2019, February). Deep explanation model for facial expression recognition through facial action coding unit. In 2019 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 1-4). IEEE.

[17] Caramihale, T., Popescu, D., & Ichim, L. (2018). Emotion classification using a tensorflow generative adversarial network implementation. Symmetry, 10(9), 414.

[18] Liu, P., Han, S., Meng, Z., & Tong, Y. (2014). Facial expression recognition via a boosted deep belief network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1805-1812).