International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 6 Volume No.6 October - November – 2022
DOI:10.46647/ijetms.2022.v06i06.097 ISSN: 2581-4621

# A Real-Time Visage Expression Detection Using Convolutional Neural Network (RTVED)

**Nandani Sharma[1], Deepali Verma [2], Prakriti Chaurasia[3]**
*[1,3] Assistant Professor, Department of IT, Shri Ramswaroop Memorial College of Engineering & Management, Lucknow, Uttar Pradesh – 226028, India.*
*[2] Research Scholar, Department of CSE, Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh, India.*

**ABSTRACT**
In this work, the purpose is to implement parts of seven Real-Time Visage Expression Detection using Convolutional Neural network (RTVED), Data Augmentation to get high accuracy, and focus on the Methodology of Real-Time Visage Expression Detection using CNN. In this paper, considered a convention Convolutional Neural Network model and utilized it to prepare and test diverse look pictures with the Keras, TensorFlow, and deep learning library. Real-Time human visage Expressions Detection has two sections, recognizer Validation, and data preparation, and a training model for data preparation and training. The recognizer contains a visage Expression detector and a visage Expression recognizer. The visage expression detector extricates facial pictures from the camera and the visage Expression recognizer recognizes the extracted pictures. The Data Training model utilizes the Convolutional Neural Network to prepare data and the recognizer likewise utilizes Convolutional Neural Network to identify the emotional condition through their visage Expressions. The framework perceives the six widespread emotions as anger, disgust, happiness, sadness, surprise neutral, and contempt.
**Keywords— CNN, Keras, TensorFlow, visage Expression detector and recognizer, CK+, FER2013.**

## 1. Introduction

For facial visage emotion recognition, the traditional approaches usually consider a visage image that is distinguished from an information picture, and facial visage fragments or achievements are perceived from the appearance locale. After that diverse spatial and common features are isolated from these facial visage fragments. At last dependent on the separated highlights a classifier, for example, Keras library, TensorFlow, sequential, is trained to produce recognition results [1]. Since the Convolutional Neural Network technique requires at least hours of training, the training phase must be separate from the front-end monitoring system. In this paper, design a proper Convolutional Neural Network for Real-time human Visage Expression detection. The other one is the facial visage identification and Recognition program that accesses a camera for input video and applies facial visage detection on video frames. Next, it utilizes the facial visage pictures from facial visage detection and the yield from the preparation program for expression recognition.

Significant research has been conducted in the past to increase efficiency and resolve the issues and problems, several methods have been proposed. The following (Table 1) are some of the most important and current works that address visage expression detection challenges.

**1.1. Artificial Neural Networks:** An ANN is a bunch of methodologies that execute label forecasts. On the off chance that the ANN is analyzed as a black box; the info would comprise labeled models, and the yield would be a vector containing a bunch of forecasts [1] [8].

**1.2. Affective Computing:** The fundamental establishments at the back full of Affective processing i.e. without feelings, people would not appropriately work as sane decision-making creatures. A few researchers show that there is nothing of the sort as "pure reason" [2] [10].

**1.3. Human Visage Expression Recognition:** This will give Computer graphical countenances that give a more normal communication. With regards to Recognition, computers have had the option to perceive some facial visage classes: happy, surprise, anger, and disgust [3] [9].

DOI:10.46647/ijetms.2022.v06i06.097 ISSN: 2581-4621

**Table 1: Table representing the summary of the latest related works [7].**

| Issues/ Problems | Methods | Networks Types | Datasets | Expressions /Classes | Classifier | Performance % |
|---|---|---|---|---|---|---|
| Imbalance and over-fitting problems | [3] | ACNN | LEW<br>CelebA | 6 classes | kNN | 99.62%<br>88.78% |
| | [4] | DCNN | AffectNet | 8 classes | SVM | 60.70% |
| Intra- Class variation of FER and over-fitting problems | [4] | DCNN | AffectNet | 8 classes | SVM | 60.70% |
| | [5] | SD-CNN | CK+<br>Oulu-CASIA | 8 classes<br>6 classes | 6 DiffNets | 99.7%<br>91.3% |
| | [6] | CNN | RAF-DB<br>AffectNet | 7 Classes<br>8 classes | NN | 80.44%<br>65.20% |
| Illumination ,Contrast Variation and over-fitting problems | [8], [7] | CNN | FER2013, CK+,KDEF, JAFFE, SFEW2.0 ** | 7 & 8 classes | SVM | High and Good ** |
| occlusion problem | [9] | ACNN | RAF-DB<br>AffectNet | 7 Classes<br>8 classes | NN | 85.54%<br>54.84% |
| | [10] | ACNN | FERPlus, RAF-DB, AffectNet, SFEW | 7 & 8 classes | NN | ** |
| Eyeglasses | [11] | ERGAN | CelebA | 6 classes | NN | 96.35% |

**1.4. Machine Learning (ML):** The ML algorithm makes models based on the input data and information. These models produce a yield that is typically a bunch of expectations or choices. Then, at that point, when another necessity shows up, the model could deal with it or answers without the need adding new code [4].

**1.5. Deep Learning**: DNN is a piece of ANN with various mystery layers of units between the underlying and last layers. In their system, scale-invariant element change (SIFT) features contrasting with a lot of achievement centers are first removed from each facial appearance picture [5] [9].

**1.6. CNN (Convolutional Neural Network):** Convolutional Neural Network is a sort of feed-forward ANN. Where the organization plan between its neurons is pushed by the relationship of the animal's visual cortex. It is engaging for a few, Deep learning tasks like scene acknowledgment, NLP, and facial look appearance acknowledgment [1] [6].

## 2. Implementation and Framework

RTVD contains 2 parts:

**1. Recognizer for validation**

**2. Pre-train program for data training.**

The recognizer contains a facial visage detector and a visage expression recognizer. The pre-train program uses a CNN deep learning training model and all steps of the process are shown in Fig.1, 2, 3, and 4.

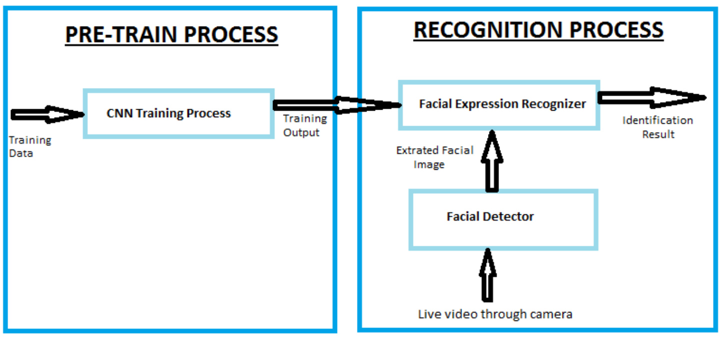



**Fig1: System Design**

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 6 Volume No.6 October - November – 2022
DOI:10.46647/ijetms.2022.v06i06.097 ISSN: 2581-4621

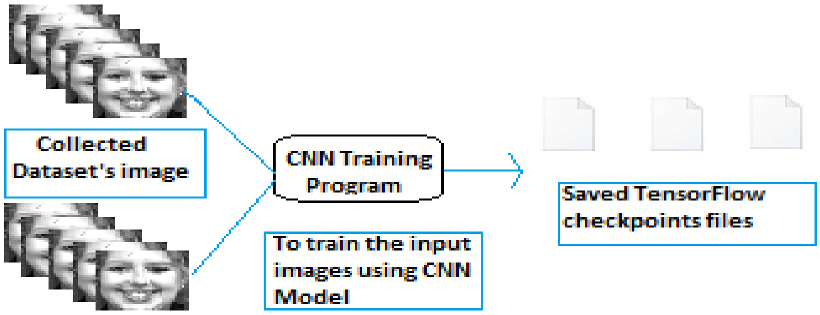



**Fig. 2. The facial visage expression recognizer requires the complete result from the pre-train program to validate the facial visage image.**

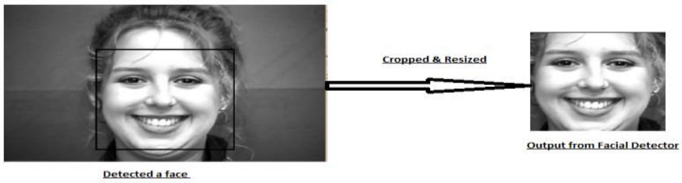



**Fig. 3. Facial visages Detector - detect and extract image from video.**

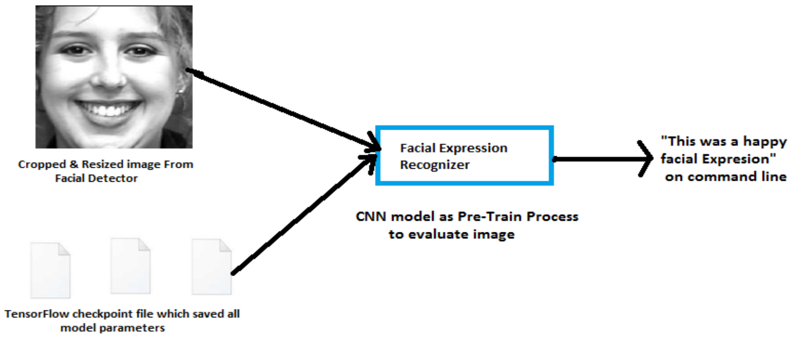



**Fig. 4. Facial visage Expression Recognizer design. It uses checkpoint files from the pre-train program to evaluate the images extracted from the facial visage detector.**

The facial visage expression recognizer uses the pre-train data for facial visage expression identification. The dataset is trained in the pre-train system using deep learning technology and the results are stored in checkpoint files. While the facial visage expression recognizer receives an image input, it transforms the input image into a tensor and uses the inference from the pre-train program to hypothesize the input image (Fig. 2). The output from the facial visage expression recognizer is a single expression label and it is the final system output. The facial visage detector was developed with OpenCV in python and the facial visage expression recognizer was developed with Tensorflow using the deep learning technique in python. Besides the six universal facial visage expressions: anger, disgust, fear, happiness, sadness, and surprise, the method also recognizes a neutral face. The proposed system has four options for the user. A first option is a train-only option with the command python system.py -t train file checkpoint name input size. In fig.4, the checkpoint name is the name of the checkpoint file without an extension and the name of a folder that is used to save the tensor board file. The second option is the validation-only option with the command python system.py –v validation file checkpoint name input size. This option requires a checkpoint file with completed

DOI:10.46647/ijetms.2022.v06i06.097 ISSN: 2581-4621

training data for validation purposes. The input can be either a tfrecords file or an image. The third option is the train-validate option with the command python system.py -b train file validation file checkpoint name input size. This option is a combination of options one and two. The last option is the live option with the command python system.py -l checkpoint name input size. This option requires a checkpoint file and an RGB camera for live video input.

**System Model**

The method used in the system is a custom CNN method to train and validate facial visage images. In the training phase, the learning rate is set to 1e-4, the batch size is set to 32, the quantity of epochs is set to 100 and 200 (mentioned in Table 2), and drop-out is set to 50%. All input data is divided into groups of 20 and each group needs to be trained 50 times. All images that use for training purposes need to be labeled manually and converted into a binary file, tfrecords which is recommended input format for TensorFlow, with labels as the input dataset. Test dataset/data can be a single tfrecords file or a single image file. An image converter can convert all labeled images into a tfrecords file and it provides an image re-size function and image channel transform function. The CNN train Process provides a re-size function and the input data can be re-sized to a fixed size after the input layer and before the first convolutional layer as shown in Figure 5.

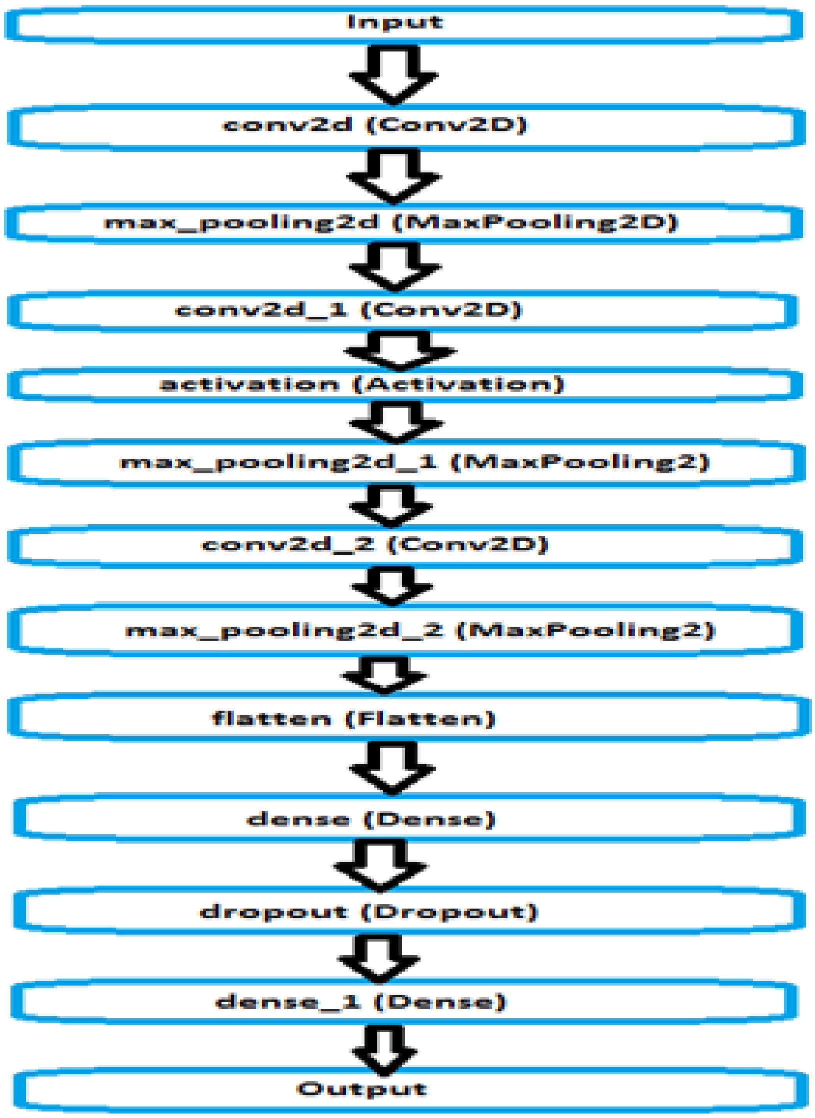



**Fig. 5. CNN Model**

DOI:10.46647/ijetms.2022.v06i06.097 ISSN: 2581-4621

## 3. Performance Evaluation

In this paper, applied CNN Model to our datasets and predict the outcomes. The outcomes are deciphered utilizing the confusion matrix (Fig. 7) and performance metric as accuracy metrics Fig. 6). The train: test split was 75:25. We additionally cross-validation the dataset to eliminate any predispositions. The worth of the split was picked as 5 because the resultant parts will have the same number of pictures as our 25% test set. The outcomes are Fig.6, 7, 8, and 9 as per the following:

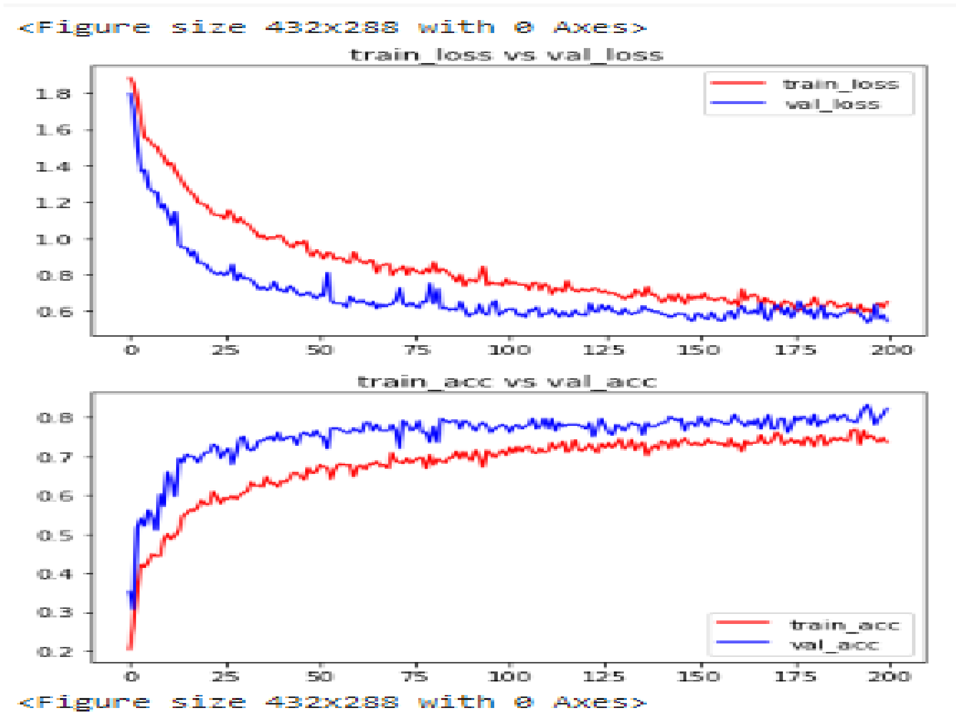



**Fig. 6. shows the Visualization of the predicted testable dataset loss and accuracy curve.**

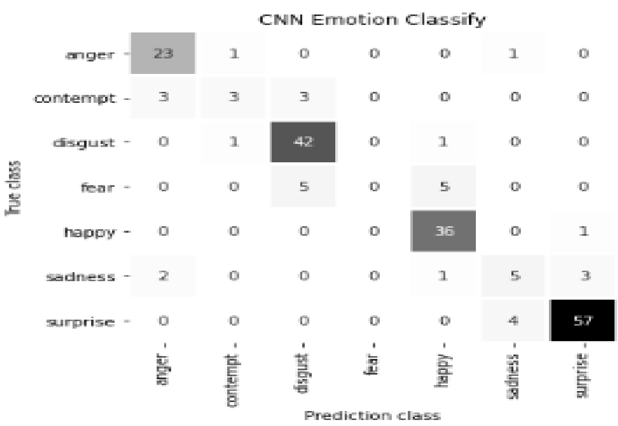



**Fig. 7. Shows the plotting of the confusion matrix.**

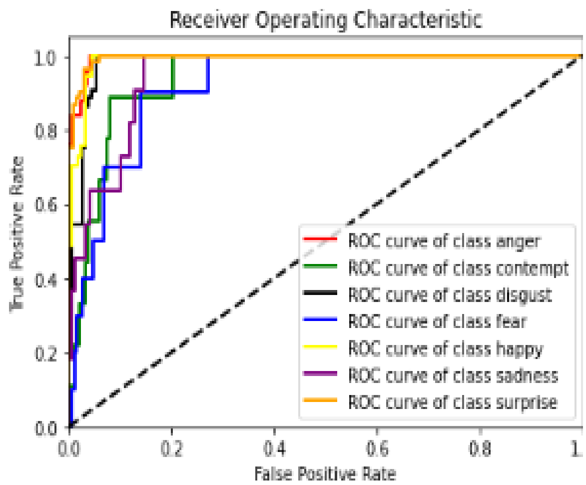
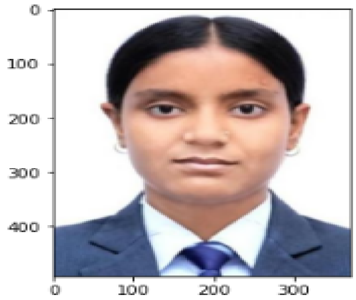



**Fig. 8. Shows the plotting of the Receiver Operating Characteristic curve (ROC).**

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 6 Volume No.6 October - November – 2022
DOI:10.46647/ijetms.2022.v06i06.097 ISSN: 2581-4621

```
img = cv2.imread('/content/drive/MyDrive/nandani/nandaniimage.jpg')
plt.imshow(img[:,:,::-1])
plt.show()
```




```
result = DeepFace.analyze(img, actions = ['emotion'])
```

**Fig.9. shows the result of the RBG image as an output.**

**Table 2: comparative summary results of ck+ and fer2013 datasets**

| **Factors/datasets** | **CK+** | | **FER2013** |
|---|---|---|---|
| | **100 epochs** | **200 epochs** | **100 epochs** |
| **Number of Facial visage Expressions** | 6 Universal emotions added with contempt | | 6 Universal emotions added with neutral |
| **Number of Experimental Images** | 981 | | 28812 |
| **Score Accuracy** | 79.08(H) | 83.23(H) | 69.34 (H) |
| **Score Loss** | 57.01(L) | 60.29(L) | 30.98 (L) |
| **Train Accuracy & Loss** | 69.68 & 30.31 | 73.50 & 26.49 | 62.43 & 30.32 |
| **Test Accuracy & Loss** | 78.06 & 21.93 | 82.14 & 17.85 | 70.0 & 34.09 |
| **Batch size** | 8 & 32 | 8 & 32 | 32 |
| **Value Index** | 4 | | 4 |
| **Total and Trainable Parameters** | 217983 | | 217983 |
| **Predicted Test Accuracy & Loss** | 82.23 & 48.21 | 84.26 & 40.39 | 70.56 & 30.43 |

The proposed scheme was compared with some other schemes table 1 to table 2 like DCNN [4], CNN [6] and ACNN [9]. It is found that the proposed scheme performs better in terms of in types of networks and layering with CK+ and FER2013 Datasets.

## CONCLUSION

This paper evaluation and categorize Human facial visage Expression over static facial visage pictures utilizing a Deep learning methodology. This is a difficult issue that has effectively been moved toward a few times with various Techniques and methodologies. While great outcomes have been accomplished utilizing visage learning techniques, this paper focused on including CNN, which is one of the Deep Learning guarantees with FER2013 and CK+ datasets.

In this paper, the efficiency of detection and recognition is increased and achieve this by planning a CNN model to recognize human facial visage emotions of the six widespread Expressions in addition

DOI:10.46647/ijetms.2022.v06i06.097 ISSN: 2581-4621

to Neutral and contempt utilizing TensorFlow and Keras.

**References**

**1.** Ali, M. F., Khatun, M., & Turzo, N. A. (2020). Facial emotion detection using neural network. the international journal of scientific and engineering research.

**2.** Picard, Rosalind W. "The promise of affective computing." The Oxford handbook of affective computing (2015): 11-20.

**3.** Spiers, D. L. (2016). Facial emotion detection using deep learning.

**4.** Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35(8), 1798-1828.

**5.** Bodapati, J. D., Srilakshmi, U., & Veeranjaneyulu, N. (2022). FERNet: a deep CNN architecture for facial expression recognition in the wild. Journal of The institution of engineers (India): Series B, 103(2), 439-448.

**6.** Li, Y., Zeng, J., Shan, S., & Chen, X. (2018). Occlusion-aware facial expression recognition using CNN with an attention mechanism. IEEE Transactions on Image Processing, 28(5), 2439-2450.

**7.** Sharma Nandani, Peeyush Kumar Pathak, Deepali Verma. (2021). A Survey Paper on Issues and Challenges in Facial Expression Recognition with the solution. International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN: 2349-5162, Vol.8, Issue 6, page no.f814-f824.

**8.** Sham, A. H., Aktas, K., Rizhinashvili, D., Kuklianov, D., Alisinanoglu, F., Ofodile, I., & Anbarjafari, G. (2022). Ethical AI in facial expression analysis: Racial bias. Signal, Image and Video Processing, 1-8.

**9.** Ge, H., Zhu, Z., Dai, Y., Wang, B., & Wu, X. (2022). Facial expression recognition based on deep learning. Computer Methods and Programs in Biomedicine, 215, 106621.

**10.** Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., & Zhang, W. (2022). A systematic review on affective computing: Emotion models, databases, and recent advances. Information Fusion.

**11.** Hu, Bingwen, Zhedong Zheng, Ping Liu, Wankou Yang, and Mingwu Ren. (2020). Unsupervised eyeglasses removal in the wild. IEEE Transactions on Cybernetics.