

# A SURVEY ON FACIAL EXPRESSION RECOGNITION USING DEEP LEARNING APPROACH.

<sup>1</sup>Rachnaa Rajvanshi, <sup>2</sup>Kinjal Thakar,

<sup>1</sup>Student, <sup>2</sup>Assistant Professor

<sup>1</sup>Computer Department,

<sup>1</sup> Silver Oak College of Engineering & Technology, GTU, Ahmedabad, India

**Abstract**— In the course of the most recent couple of years, profound neural systems have gotten the most consideration in software engineering, particularly in design acknowledgment, machine vision and machine learning. One of its superb applications is in the feeling acknowledgment by means of Facial expression recognition region. Facial expression recognition investigation is valuable for some undertakings and the utilization of profound learning here is additionally growing quick. We survey some current research works in this space, present some new applications and demonstrate the general strides to actualizing every one of them.

**Keywords** - Deep Learning, facial expression recognition, emotion analysis. Image processing, Convolutional Neural Networks.

## I. INTRODUCTION

Human correspondence and social cooperation are one of the primary research regions which numerous researchers like therapists, sociologists and PC researchers are occupied with. As its software engineering viewpoint, the PC should help individuals for better association with both human and PC [1]. As of late, HCI (human-PC association) has been a dynamic research point in software engineering and it is about all the inquires about of the outline and utilization of PC innovation, which are centered around the interfaces amongst clients and PCs. Human feeling examination assumes a noteworthy part in giving appropriate human-PC cooperation and it talks about PC frameworks which endeavor to break down and perceive both facial component changes and facial movements from every visual recognition. Along these lines, facial component extraction is one of the primary parts of feeling investigation, which endeavors to locate the most proper data portrayal of facial pictures. Robotized Facial expression recognition acknowledgment has a vast assortment of uses, for example, information driven liveliness, intelligent recreations, stimulations, friendly mechanical autonomy, reconnaissance, swarm investigation, humanoid robots, intuitive TV, and numerous other human-PC cooperation frameworks [2][3].

In spite of the fact that there are numerous looks into in Facial expression recognition and acknowledgment, with a high exactness and execution in an online certifiable, despite everything it has a few challenges because of its intricacy and fluctuation.

The way that diverse individuals express their facial feelings may differ from each other besides, a picture can be fluctuated in shine, foundation, position and numerous other criteria which straightforwardly or by implication are influenced in the examination procedure. More exact and more elevated amount of learning and data is required for human feeling investigation[4].

Facial Expression additionally have some data about expectation, intellectual procedures, physical exertion, or other intra-or relational implications and understanding about these information can be finished and be more precise by setting, body signal, voice, singular contrasts, and social factors and additionally by facial setup and timing however the robotized Facial expression recognition investigation frameworks need to examine the facial activities and highlights paying little respect to setting, culture, sexual orientation, et cetera [5].

Automatic facial expression analysis can be done in three main steps:

- Face acquisition;
- Facial data extraction and representation, and
- Facial expression recognition.
- Face detection and
- Head pose estimation.
- 1) Geometric feature-based method
- 2) Appearance-based method, and
- 3) Hybrid-based method.

In the main strategy, it is imperative to quantify the shape and area of facial highlights. The Geometric estimations in light of the connections between these highlights are being utilized to develop an element vector for preparing reason. In static picture the undertaking will be done on the present picture, yet in powerful pictures, for example, video outlines, when we have a grouping of pictures, it should be possible by estimating geometrical removal of facial element focuses between the present casing and starting edge or another particular casing or even different edges. [6][7].

The second strategy separates the highlights by applying at least one channels to the feature pictures and here are some related methodologies, for example, PCA (main segment examination), LDA (straight segregate investigation), ICA (autonomous segment examination) and GW (Gabor wavelet). In appearance-based technique, changed facial locales contain distinctive data, for instance, eyes and mouth contain more data than the brow and cheek [8].

Also, the third technique is a mix of geometric and appearance based strategies, which gives better outcomes now and again. Both the first and second techniques have a few issues and blunders which can be secured by their combination, and subsequently, framework precision will be expanded [9].

The last advance in facial expression analysis is acknowledgment by grouping these highlights and numerous strategies can be utilized for it, for example, an ANN (manufactured neural system), SVM (bolster vector machine), BN (Bayesian system) and numerous different classifiers. Profound learning is a standout amongst the latest strategies that accomplishes outflanking comes about. It gives a successful answer for rough thinking, and productive avaricious calculation for applying in numerous applications, for example, Facial Expression Recognition. Various looks into have been done around there, yet when all is said in done they have comparative advances [10][11].

In Section 2, the execution ventures as a rule will be looked into. Area 3 quickly portrays the CNN technique, Section 4 depicts the execution ventures of the CNN and the last segment finishes up the article.

## II. Facial Expression Recognition System

As we mentioned before, in the general a FEs (facial expression systems) can be implemented in 3 major steps which we explain in detail here:

### A. Image Acquisition and Pre-processing

The Facial picture information can be grabbed from a database (static) or a live video stream (dynamic), in 2D or 3D mode. Here we additionally have a few stages as pre-preparing, for example, de-noising (which are identified with the gadgets that we send), et cetera which help to have better execution.

### B. Feature Extraction

Extricating the best highlights is a standout amongst the most imperative strides of any fruitful Facial expression recognition acknowledgment framework. The proficiency and viability of the facial picture portrayal could impact on the heartiness amid the acknowledgment procedure.

### C. Classification and Facial Expression Recognition

Numerous classifiers, for example, KNN, LDA, ANNs, HMMs, SVMs and CNN can be applied to the automatic expression recognition problem.

## III. CONVOLUTIONAL NEURAL NETWORK

For the first time, CNN was presented by Lecun et al.[12]. Having distinctive kinds of data portrayal, is the key point in CNN usefulness. Each layer can respond to the distinctive data, and when the layers stacked together, they can make an intricate portrayal. As of late it demonstrated great outcomes in feeling investigation, and Facial expression recognition, particularly in HRI and RRI. [13][14][15]. Fig. 1 demonstrates the general design for CNN.

## IV. FACIAL EXPRESSION WITH CNN

Here we describe four main ways of the facial expression recognition procedure by CNN:

### A. Normalization

The pictures in the database differ in numerous parameters which can influence specifically on acknowledgment precision and execution. These are a few challenges, for example, pivot, splendor and enlightenment changes notwithstanding for a similar individual's pictures. To address this issue, a standardization of the face picture, for example, identifying, de-noising and some other preprocessing, for example, amending the pivot is performed. The picture shine and difference varieties increment the multifaceted nature of the issue.

### B. Image Cropping

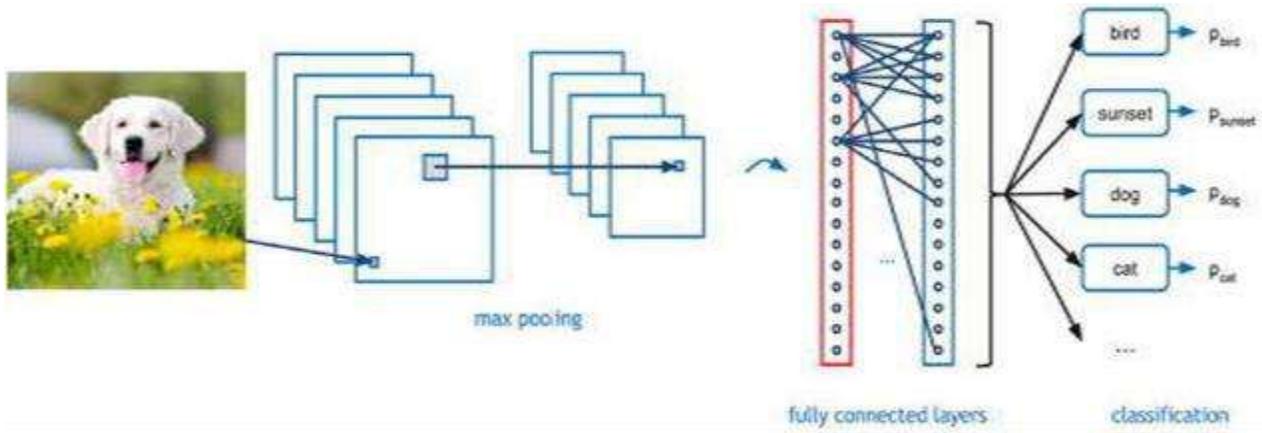
The original face images have background information that is not important and could make the output to be less accurate. The cropping region also tries to remove facial parts that do not contribute to the expression.

### C. Downsampling

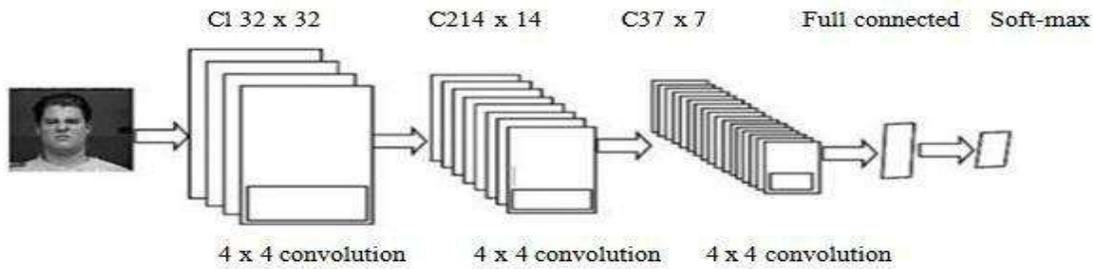
It is performing to guarantee a similar area of eyes, mouth, eyebrows, and other face segments each face picture. Down inspecting causes the CNN to realize which locales are identified with every particular articulation and furthermore empowers the CNN to be performed on the GPU all the more effectively.

### D. Convolutional Neural Network

The network receives an  $m$  image (can be specified in the down sampling step) as an input and then returns the confidence of each expression as an output. The first layer of the CNN is a convolution layer that applies a convolution kernel of  $m$  and outputs an image of  $m \times n \times 2$  pixels. This layer is followed by a subsampling layer that uses Max-pooling with kernel size  $k$  to reduce the size of the image by half. Subsequently, a new convolution is applied to the feature vector and is followed by another subsampling. The output is given to a fully connected layer that has  $L$  neurons. The network has six outputs.



**Fig.1 : The Convolutional Neural Network**



**Fig. 2 The ER architecture.**

Nodes that are fully connected to the previous layer. Each output node is corresponding to one of the expressions that outputs its confidence level [16][17]. Fig. 2 shows the expression recognition system architecture. Some operations such as Max pooling and convolution will be applied to the original image as filters to extract a different representation of images in each layer. The maximum pooling operation is used to reduce the dimensions of the extracted hidden features for training. Finally, in the last step, the Soft-max classifier is used to classify the facial expressions of the test samples from extracted features. Some issues in applying CNN are such as computation time for the many layers and data which we have here and the type of the filter which determine the new presentation of data as the core of CNN. There are some solutions for these issues such as using the GPU (graphics processing unit). The experimental results of researches, show the performance and the generalization ability of the CNN for creating a very accurate facial expression recognition.

## V. CONCLUSION

In this paper a concise depiction of Facial Expression Recognition using CNN has displayed. We finish up by saying that the innovation of Facial Expression Recognition has huge market potential and, sooner rather than later, it will improve most huma-computer interfaces. The new learning innovations especially CNN) help to get improved accuracy and performance.

## VI. ACKNOWLEDGMENT

No task can be accomplished without proper guidance, support and appraisal. I am deeply thankful to many people who helped me either directly or indirectly for this work and provided their invaluable cooperation to me to complete it. At Last, I heartily offer my regards and blessings to all of those who supported me in any respect during the completion of my work.

## REFERENCES

- [1] Ghayoumi, M., Tafar, M., and Arvind, K. B. 2016. "Towards Formal Multimodal Analysis of Emotions for Affective Computing." Presented at the 22<sup>nd</sup> International Conference on Distributed Multimedia Systems (DMS).
- [2] Li, S. Z., and Jain, A. K. 2011. "Handbook of Face Recognition." *Springer Science & Business Media*.
- [3] Michel, V., and Maja, P. 2010. "Induced Disgust, Happiness and Surprise: An Addition to the Facial Expression Database." In *Proc. 3rd Intern. Workshop on EMOTION*.
- [4] Ghayoumi, M. 2016. "Follower Robot with an Optimized Gesture Recognition System." Presented at the Robotics: Science and Systems (RSS), USA.
- [5] Susskind, J. M., Anderson, A. K., and Geoffrey, E. H. 2011. *The Toronto Face Database*. University of Toronto, Toronto, ON, Canada, Tech. Rep.
- [6] Tian, Y., Kanade, T., and Cohn, J. F. 2005. "Facial Expression Analysis." In *Handbook of Face Recognition*: 247-75.
- [7] Ghayoumi, M., and Bansal, A. 2015. "Unifying Geometric Features and Facial Action Units for Improved Performance of Facial Expression Analysis." *New Developments in Circuits, Systems, Signal Processing, Communications and Computers (CSSCC)*: 259-66.
- [8] Zee, T., and Ghayoumi, M. 2016. "Comparative Graph Model for Facial Recognition." Presented at the 2016 International Conference on Computational Science and Computational Intelligence (CSCI'16).

- [9] Ghayoumi, M., and Bansal, A. 2014. "An Integrated Approach for Efficient Analysis of Facial Expressions." Presented at the International Conference on Signal Processing and Multimedia Applications (SIGMAP).
- [10] Ghayoumi, M. 2015. "A Review of Multimodal Biometric Systems Fusion Methods and Its Applications." Presented at the 14th International Conference on Computer and Information Science (ICIS), USA.
- [11] Ghayoumi, M., Tafar, M., and Bansal, A. K. 2016. "A Formal Approach for Multimodal Integration to Derive Emotions." *Journal of Visual Languages and Sentient Systems*: 48-54.
- [12] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. 1998. "Gradient-Based Learning Applied to Document Recognition." In *Proc. IEEE*.
- [13] Ghayoumi, M., and Bansal, A. 2016. "Emotion in Robots Using Convolutional Neural Networks." Presented at the ICSR.
- [14] Ghayoumi, M., and Bansal, A. 2016. "Architecture of Emotion in Robots Using Convolutional Neural Networks." Presented at the RSS, USA.
- [15] Ghayoumi, M., and Bansal, A. 2016. "The Multimodal Architecture of Emotion in Robots Using Deep Learning." Presented at the Future Technologies Conference, San Francisco, United States.
- [16] Liu, P., Han, S., Meng, Z., and Tong, Y. 2014. "Facial Expression Recognition via a Boosted Deep Belief Network." Presented at the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [17] Ghayoumi, M., and Bansal, A. K. 2017. "Improved Human Emotion Recognition Using Symmetry of Facial Key Points with Dihedral Group." *International Journal of Advanced Studies in Computer Science and Engineering (IJASCSE)*.

