IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Convolution Neural Networking for Emotion Sensing through Facial Expression Recognition and its Application

[1]Krupali Trivedi, [2]Dhruv Patel, [3]Prince Patel, [4]Yogeshvari Makwana

[1]B.Tech Student, Department of Computer Science and Engineering, Indus Univeristy, Ahmedabad, India.

[2]B.Tech Student, Department of Computer Science and Engineering, Indus Univeristy, Ahmedabad, India.

[3]B.Tech Student, Department of Computer Science and Engineering, Indus Univeristy, Ahmedabad, India.

[4]Assistant Professor, Department of CSE, Indus University, Ahmedabad, Gujarat, India

Abstract:
This review paper explores Convolutional Neural Networking (CNN) for Emotion Sensing via Facial Expression Recognition (FER). The paper compares traditional FER methods with CNN-based FER methods and discusses the advantages of CNN in terms of accuracy and speed. Additionally, the paper highlights the various applications of FER in areas such as psychology, human-computer interaction, and healthcare. Overall, this paper provides a comprehensive overview of the current state-of-the-art in CNN-based FER and its potential impact on various industries.

Keywords:
CNN, Emotion Recognition, Facial Expression Recognition, FER

Introduction:

Understanding and interpreting human emotions has always been a crucial aspect of human civilization, as it enables us to connect with each other and build meaningful relationships. It is widely believed that women are better at detecting emotions, especially fear and disgust, than men. The ability to learn from others' emotions is essential in various fields, such as trading, religion, and fortune-telling. The eight basic human emotions are disgust, being neutral, contempt, happiness, joy, sadness, anger and getting surprised.

However, humans often hide their emotions, and in some cases, lie detection tests are necessary. With the advancement of technology, electronic sensing technology has made it possible to record and understand human behaviour and emotions in detail, changing the way we perceive emotions. In the future, we may not only be able to read a person's emotions but also predict their behaviour.

The emergence of emotion-sensing technology has led to a novel design approach that involves monitoring and analyzing behavior patterns, measuring actions, and taking note of facial expressions, voice intonation, and body language [1] [2] are combined to create a more comprehensive understanding of emotions. Smart devices can now assess the meaning behind certain emotions and respond accordingly, paving the way for a more personalized and intuitive user experience.

One area of research that has gained significant attention in recent years is Facial Expression Recognition (FER). FER involves using computer vision algorithms to analyse and interpret facial expressions, enabling

machines to recognize and respond to human emotions. This technology has a wide range of applications, including robotics, neuro-marketing, academia, and most notably, security [3].

As per survey by S. Li [4] [5] Facial Expression Recognition is possible through multiple deep learning techniques like Convolution Neural Network (CNN), Recurrent Neural Network (RNN) as well as Deep Belief Network (DBN).

The paper is divided in two sections. Section one of the paper focuses on use of Neural Networks, specifically Convolution Neural Network for FER. The paper surveys various problems existing and their solutions that have been through previous research. Along with that, in section two the paper also presents applications of FER in day to day life as we move ahead with digitization.

Facial Expression Recognition:

Facial expressions are one of the most important nonverbal cues that we use to convey our emotions and intentions to others [5]. They are innate, universal, and can be powerful indicators of our emotional states, ranging from joy and happiness to sadness, anger, and disgust. As technology continues to advance, we are seeing an increasing trend towards automation, with computers and machines taking over many of the tasks that were once performed by humans.

By accurately predicting facial expressions, FER technology can help enhance security measures, such as identifying potential threats at airports, train stations, or other public places. It can also be used in academic research to better understand how people express emotions, and how these expressions can be used to improve communication and collaboration.

In addition, FER can be used in the field of neuro-marketing, which involves using neuroscience to understand consumer behavior and preferences. By analyzing facial expressions, researchers can gain valuable insights into how people react to different products and advertising messages, helping companies design more effective marketing campaigns [6] [4] [3].

The emotions are divided on the base of time and each of these states are termed as emotional states of a human. On basis of these emotional states, it is possible to identify the behaviour of human which would help in bringing out various applications [1].

The feature recognition system follows particular steps [7]. Accurate classification of the face is the primary step in any facial processing system, followed by detecting the emotion from the facial expression. A standard emotion detection system involves four process flow stages: Face detection, preprocessing, feature extraction, and emotion recognition.

As seen in, the basic Neural Network Technique for FER can be seen in the Figure 1.

Facial Expression Recognition is a rapidly evolving field with enormous potential for improving human-machine interactions and enhancing our understanding of human emotions and behavior. As technology continues to advance, we can expect FER to become even more accurate and sophisticated, opening new possibilities for research, innovation, and improved security.

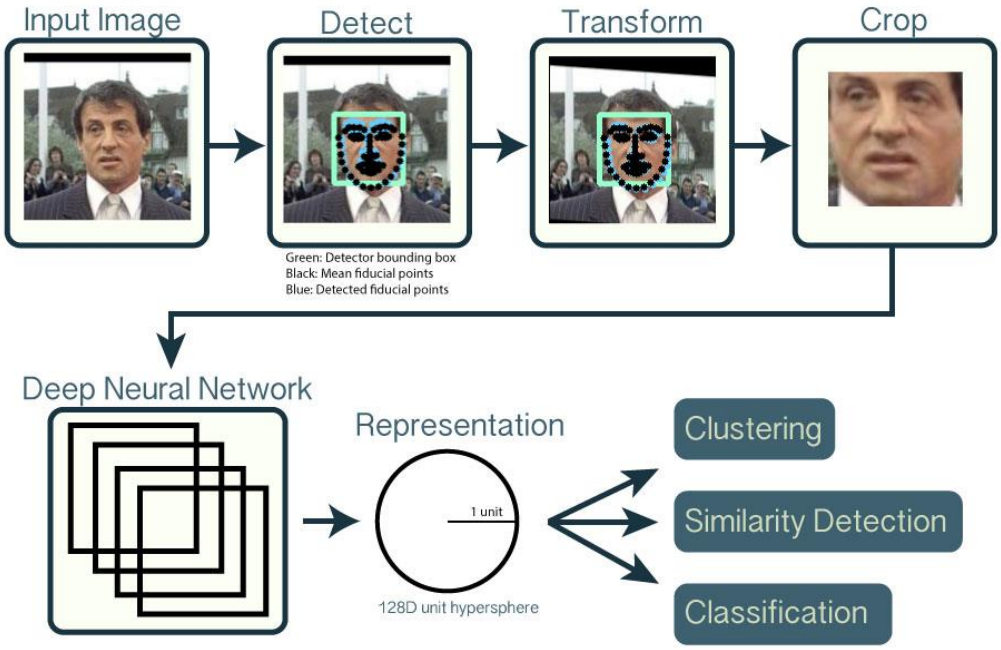

Figure 1: Basic Architecture of FER using Neural Networks [8]

Comparing the traditional techniques of FER to CNN:

As per survey by *Ankit V. et al* [5], traditionally, researchers in the field of facial expression recognition have proposed their methods using a variety of classifiers such as the Multi-layer Perceptron Model (MLP), Support Vector Machines (SVM), and k-Nearest Neighbours (k-NN) [9]. These classifiers typically use handcrafted features such as texture, face landmark features, Histogram of Oriented Gradients (HoG), gradient feature mapping, eigen vectors, and other similar features.

The process of feature extraction is crucial in facial expression recognition, as it helps to identify the most important visual cues in the facial expressions. Techniques such as Gabor filters, Local Binary Patterns (LBP), Eigen Faces, Linear Discriminant Analysis (LDA), and Principal Component Analysis (PCA) are commonly used to extract these features.

However, while these traditional methods have been successful in facial expression recognition, they often rely on handcrafted features that are limited in their ability to capture the full range of facial expressions. These problems usually arise as each picture passed through these techniques possess unique angles, position, and different facial features [7]. This has led to the development of deep learning approaches, such as Convolutional Neural Networks (CNNs), which have shown great promise in capturing complex features and achieving state-of-the-art performance in facial expression recognition tasks.

Figure 2 describes the basic difference of Traditional Methods in comparison to Deep Learning Techniques like CNNs. In traditional methods, the features have to manually selected from all the specific detected feature which in the end might bring errors in final result [10] [4] [7].

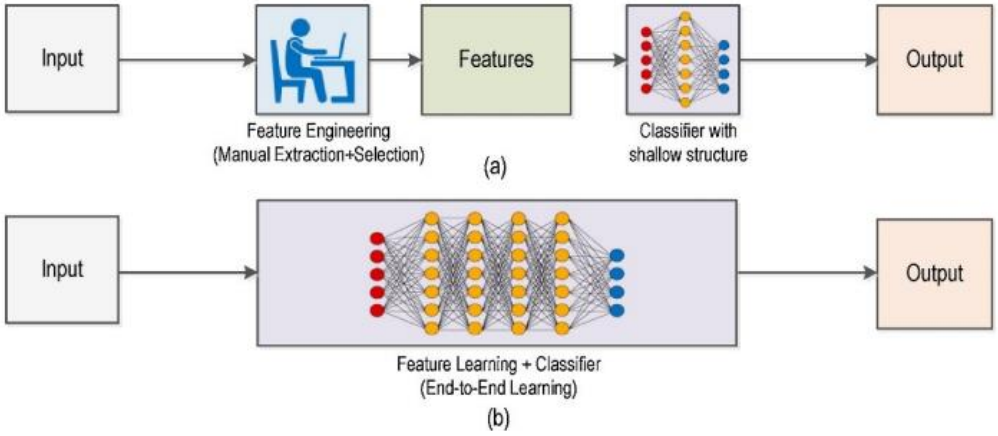

Figure 2: Traditional Feature Recognition vs Deep Learning Techniques [11]

Convolution Neural Network for Facial Expression Recognition

Convolutional Neural Networks comprises of neuron that optimise themselves through learning and they are similar to traditional ANNs [12]. In a Convolutional Neural Network (CNN), each layer is designed to detect specific patterns in the input image. The patterns detected in each layer are then combined and passed on to the next layer, which identifies larger and more complex patterns by aggregating the features detected in the previous layer. This process continues through the network until the final layer, which produces the output of the network. CNNs comprise of three layers which are fully connected to each other [13]. The three layers of CNN consists of the following:

1. The Input Layer that will hold the pixel values of the image
2. Then comes the convolutional layer or the kernal which converts the input into feature map in which the image dimensions is further filtered. The matrix strides till the whole image width is covered according to the dimensions. The objective of this layer to extract the features.
3. The rectified linear unit (ReLu function) that is performs the functioning piecewise and gives the result to see if the output will be negative, positive or zero.
4. The pooling layer performs down sampling of the input image. This helps in decreasing the computational power. The pooling layer comprises of Max Pooling and Average Pooling. This helps in extracting essential features from the images.
5. The fully-connected layer helps in learning the non-linear features that are of importance which have come out as output of convolutional layer.

The above layers are explained visually in form of image in Figure 3.

N. Lokeswari and K. Amaravathi [7] in their research about facial expression recognition make use of CNN and RNN to compare these techniques to the traditional techniques. As a result, it was found that using Neural Networks in emotion detection, it was easier to predict the detection even in conditions where the lighting was low or the image quality was of low resolution. In addition to that, the techniques bring state of art detection i.e no features have to be defined for each iteration.

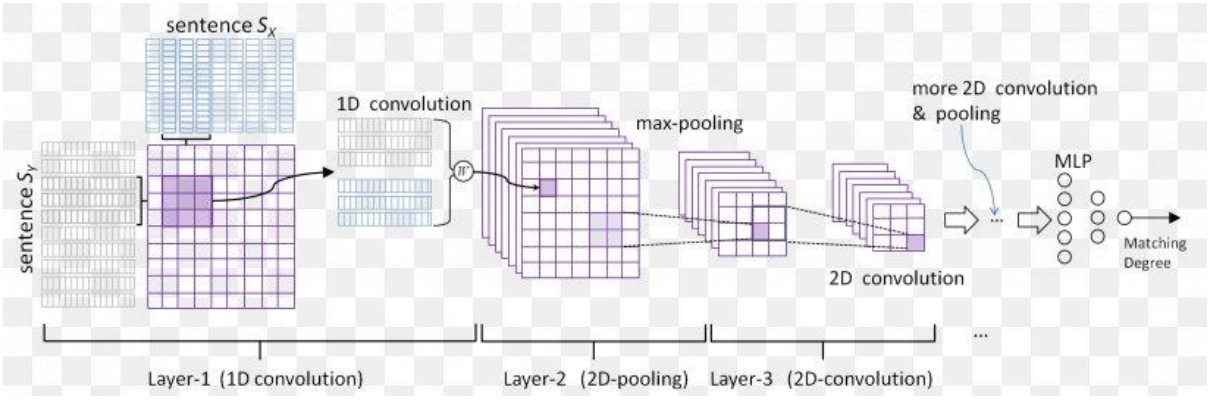

Figure 3:Convolutional Neural Network Architecture

Advancing upon the CNN techniques, Shao Jie et al [14] have proposed three variations of CNN in their research paper. Light CNN is a conventional CNN with six depthwise separable convolutional modules, followed by ReLu functions that helped in bringing out highly accurate edge features. The architecture has been adapted from ResNet and Xception.

The second one is dual-branched CNN [14] that presents global features and local texture features simultaneously. In order to achieve this, it uses three modules, two of them are conventional CNN and the other one is fusion model. The first branch gives global features, second will take pre-processed texture image as input and fusion model takes both outputs as input, thus resulting into maintained integrity of the facial expression while also maintaining the local textures.

The best result was achieved by deeper CNN in the model pertained CNN, which was constructed by exploiting ResNet101, this possesses five convolution layer, following average pooling layer and flattening layer.

Hamsa A. [15] has researched on integrating ANN with neuro-fuzzy fusion which in prospects of the future can also be implemented along with CNN to achieve the desired results.

Applications of Facial Expression Recognition:

The following section of the paper dives into the applications of Facial Expression Recognition technique that have been researched and introduced in the world till now.

Some possible applications according to A. Kolakowska [1], the FER can be used in order to do better software testing. A few of these applications include First Impression Test, Task-Based Utility Test, Free Interaction Test and Comparative Test. Further, the technology can be used productivity hacking, taking into measure the live detection of behaviour and finding out when the expression turns into boredom. FER can also be applied to compare the quality of code according the to the mood of the developer which needs continuous monitoring of expression changes.

Hugo Mitre-Hernández et al [16] explore in their research the application of CNN in a web-application developed by them which served the purpose of helping psychiatrists and psychologists in times of Covid-19 to sense the emotions conveyed by the patients. Optimized CNN models were applied and Residual mobile-based networks which also uses convolution were used to back the web-application. Five experts evaluated the usability (73.8 of 100) and utility (3.94 of 5) of this tool as good.

Joao Marcos Silva et al. [17] have worked upon solution to bring and research upon CNN based facial expression recognition to aid the visually impaired who cannot see so that the technique could sense how the person they are interacting with is reacting based on the emotion recognition.

P. Kasani et al. [18] have moved ahead with implementing the convolution neural network techniques in order to have video emotion sensing being made possible which can be used to sense the emotions real-time.

Jeong M et al. [19] researched upon an embedded system that helps in recognizing the features and emotions of drivers and thus the impact seen on their driving in order to achieve safety.

Conclusion:

In conclusion, Convolutional Neural Networks have shown significant promise in improving the accuracy and robustness of Facial Expression Recognition for emotion sensing. The advancements in deep learning techniques and availability of large datasets have made it possible to train deep CNNs to recognize subtle changes in facial expressions that were previously difficult for traditional FER methods. The performance of CNN-based FER can be further improved by optimizing the architecture of the network and data augmentation techniques. As demonstrated by the various applications of FER in different fields, it is evident that FER has enormous potential for improving human-machine interaction and can benefit several industries, such as healthcare, security, and entertainment.

References

[1] A. Kołakowska, A. Landowska, M. Szwoch, W. Szwoch and M. Wróbel, "Emotion Recognition and Its Applications," in *Human-Computer Systems Interaction: Backgrounds and Applications 3*, Springer, 2014, pp. 51-62.

[2] R. Picard, S. Papert, W. Bender, B. Blumberg, C. Breazeal, D. Cavallo and C. Strohecker, "Affective learning – a manifesto," *BT Technology Journal,* 2004.

[3] Y. Huang, F. Chen, L. Shaoh and W. Xiaodong, "Facial Expression Recognition: A Survey," *Symmetry,* 2019.

[4] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," *IEEE Transactions on Affective Computing,* pp. 1195-1215, 2022.

[5] A. S. Vyas, H. B. Prajapati and V. K. Dabhi, "Survey on Face Expression Recognition using CNN," in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 2019.

[6] P. Arnau-González, S. Katsigiannis, M. Arevalillo-Herráez and N. Ramzan, "Artificial intelligence for affective computing: an emotion recognition case study," *AI for Emerging Verticals,* 2020.

[7] N. Lokeswari and K. Amaravathi, "Automatic Facial Expression Recognition using CNN and RNN Algorithm's," *International Journal of Emerging Technologies and Innovative Research,* pp. 777-788, 2021.

[8] B. Amos, L. Bartosz and M. Satyanarayanan, "OpenFace: A general-purpose face recognition," 2016.

[9] C. Sawangwong, K. Puangsuwan, N. Boonnam, S. Kajornkasirat and W. Srisang, "Classification technique for real-time emotion detection using machine learning models," *IAES International Journal of Artificial Intelligence (IJ-AI),* pp. 1478-1486, 2022.

[10] C.-H. Hua, H.-T. Thien and H. Seo, "Convolutional Network with Densely Backward Attention for Facial Expression Recognition," *2020 14th International Conference on Ubiquitous Information Management and Communication (IMCOM),* pp. 1-6, 2020.

[11] S. Premkumar, V. Premkumar and R. Dhakshinamurthy, *Video Analytics on IoT devices,* 2021.

[12] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *CoRR,* 2015.

[13] I. Tabian, H. Fu and Z. S. Khodae, "A Convolutional Neural Network for Impact Detection and Characterization of Complex Composite Structures," *Internet of Things for Structural Health Monitoring,* 2019.

[14] J. . Shao and Y. . Qian, "Three convolutional neural network models for facial expression recognition in the wild," *Neurocomputing,* vol. 355, no. , pp. 82-92, 2019.

[15] H. Abdulkareem, "Neuro-fuzzy inference system based face recognition using feature extraction," *TELKOMNIKA (Telecommunication Computing Electronics and Control),* pp. 427-435, 2020.

[16] H. Mitre-Hernandez, R. Ferro-Perez and F. Gonzalez-Hernandez, "Convolutional Neural Network for emotion recognition to assist psychiatrists and psychologists during the COVID-19 pandemic: experts opinion," *arXiv preprint arXiv:2005.07649,* 2020.

[17] J. Silva, R. Silva, R. Veras, K. Aires and L. B. Neto, "Facial Expression Recognition to Aid Visually Impaired People," in *Anais do XVII Workshop de Visão Computacional*, 2021.

[18] P. Kasani, N. Jagini, S. Ramaraj and D. Jeyaraman, "Video-based emotion sensing and recognition using convolutional neural network based kinetic gas molecule optimization," *ACTA IMEKO,* 2022.

[19] M. Jeong and B. Ko, "Driver's Facial Expression Recognition in Real-Time for Safe Driving," *Sensors,* 2018.

International Research Journal
IJNRD
Research Through Innovation