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

Deep Learning Techniques for Facial Expression Recognition

Submitted by

Manju Sharma

1. Introduction

In communication, facial expressions reflect emotions of a human. Facial expression recognition (FER), as the main technology in emotional computing systems, is used not just in human-computer interaction, but in interactive gaming platforms, smart recommendation, auxiliary medical care, and safe driving too. It has a lot of potentials for use in a variety of fields. Face identification, feature extraction and expression classification are typically three steps in the FER. Among them, feature extraction is crucial in an expression recognition system because it affects the accuracy of recognition. Differentiation in facial expression with intensity is demonstra`ted in Figure 1 [Jiskoot et al., 2021].

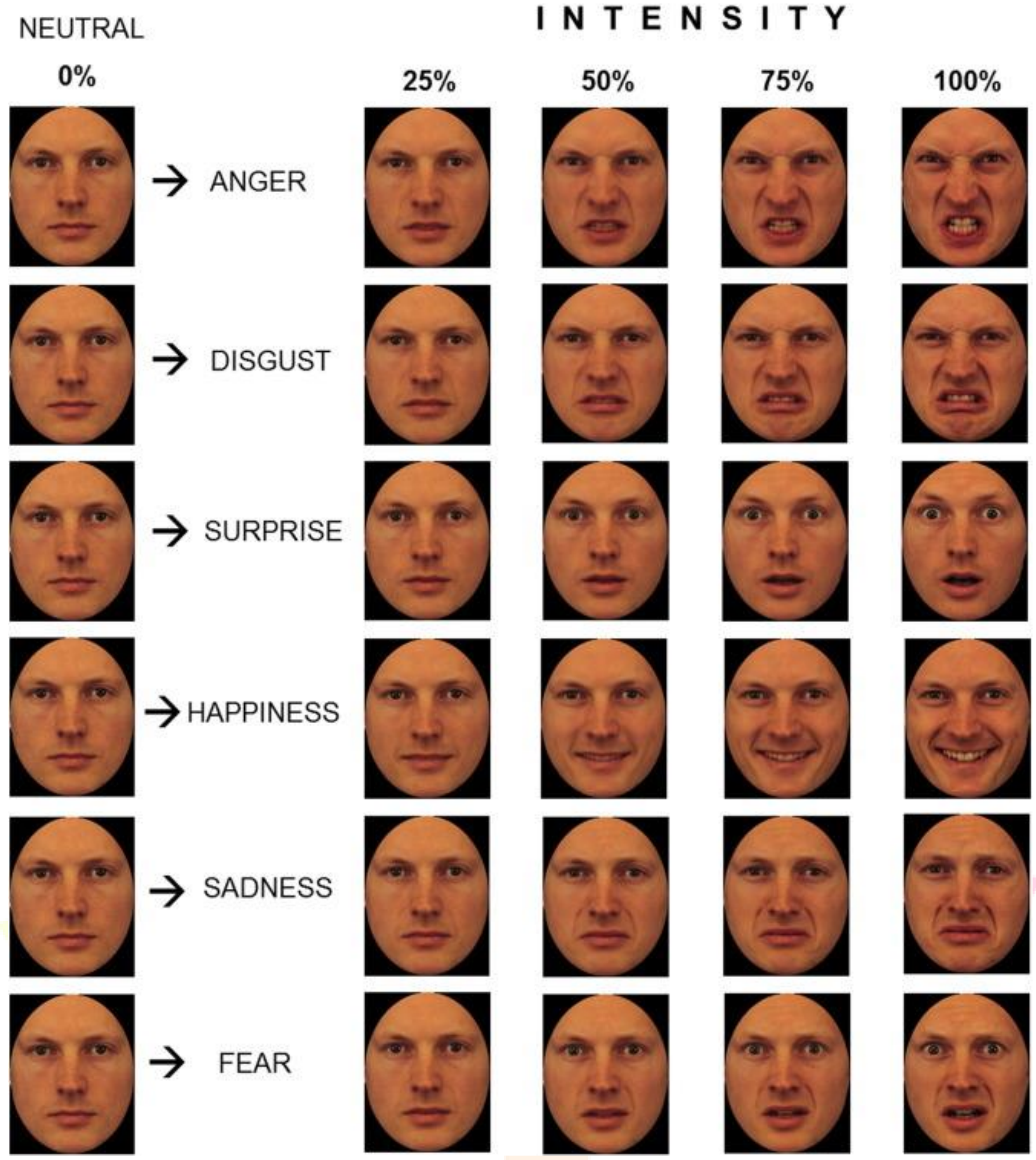

Figure 1: Differentiation in Facial Expression with Intensity

Newly, studies on FER have received lot of attention, and several approaches have been suggested, including Gabor Filter, Active Appearance Models, Active Shape Model, Local Binary Pattern, Histograms of Oriented Gradients and Principal Component Analysis. Different from traditional Machine Learning algorithms, Deep Learning can delete facial features without requiring intervention of human.

The multilayered deep neural network (DNN) can classify the critical features independently by sample data. The efficiency of the DL system for extracting features is superior to conventional ML algorithms, as evidenced by numerous publications.

Numerous hidden layers in Deep neural networks, have outstanding competencies of learning, making them further conducive to classification. layer-by-layer initialization will effectively solve the challenge of DNN training. The deep structure model represents the multi-layer functionality. Its characterization capabilities are superior to those of the shallow structure model. Deep Belief Networks , Recursive NN , Convolutional NN , Stacked Auto-encoder , and other DL models are examples [Hinton et al., 2006].

1.1 Basic, 3D, and Micro expressions

Basic facial expressions or displays are ritualized to communicate. They conveniently express the emotions and define the emotional state of a person. Micro expressions are stifled expressions which can last for a very little time. These expressions are the result of suppression or repression of expressions [Reddy et al., 2019]. Recognition of micro expression is highly challenging and a complex task. It is difficult to understand these expressions since it requires a deeper understanding of the emotions. However, it is highly significant to understand these expressions since they assist in identifying the mental condition of an individual. Manual recognition of micro expression is erroneous and tedious. Various researchers have suggested the automation of micro facial expression recognition. But, this technique also involves lot of challenges. Especially automatic recognition through videos is difficult since they do not last for very long time. Previously, this problem was resolved by matching the similarities of hand-made micro expression analysts who manually extract the expressions from the images or videos. Recent times have seen significant advancements in this field, and researchers have suggested the application of deep learning techniques for recognizing micro expressions [Merghani et al., 2018]. On the other hand, micro expression recognition using 3D facial recognition has gained huge attention in recent times. The 3D techniques overcome the problem of conventional facial recognition techniques such as, performance degradation, inaccurate recognition due to large head rotation, variation in the postures, and changes in the light intensity etc. The 3D-based recognition techniques conduct an appropriate facial expression analysis which explores the surface information and extract facial expressions with high accuracy, which is beyond the ability of conventional techniques.

In this context, this research aims to propose a novel approach to separate human faces' images into different categories of emotions using CNN. The proposed approach incorporates pre-processing techniques along with feature extraction to improve the precision of the facial recognition.

2. Literature Review

The literature review of facial expression recognition is classified in three parts: Basic, 3-D and Micro FER.

Basic FER

There has been numerous researches and many studies over the last decade on developing methods and frameworks for recognizing the emotional content of facial expression recognition process. These researchers provide a comprehensive review of approaches.

Conventional histogram method based facial expression recognition models were used previously in various research works. A facial expression classification approach based on a feature vector histogram series is well-defined in [Aung et al., 2012]. This study identifies different facial emotions obtained from video streams and static photographs. The 4 key tasks are mouth segmentation, image pre-processing, feature extraction and classification, all of which are centered on histogram-based methods. Centered on the human mouth's geometrical features, the device can identify five human expressions: pleasure, indignation, sorrow, surprise, and neutral, with an average recognition accuracy of 81.6 %.

The advent of deep learning models has transformed the facial recognition process. The stacked neural network architecture can perform complex feature extraction tasks with high accuracy. CNN based techniques are used widely in facial recognition process to reduce difficulty of extracting artificial features from a face image [Xu et al., 2018]. This method preprocesses the facial expression images before extracting facial features with trainable convolution kernels, Then utilizes the largest pooling layer to decrease the number of measurements, and lastly uses the SoftMax layer to recognize seven different types of facial expression. The planned is tested using data from the Kaggle FER challenge [FER2013] dataset. The experimental results indicate that the procedure performs well in terms of recognition and generalization.

Transfer learning(TL) has the ability to learn the important features from the basic images. TL-based facial recognition models show better performance when applied for larger datasets. The implementation of TL-based method was used to extract relevant features from the input images in [Ramalingam and Garzia, 2018]. The data was used to train the model without deteriorating its generalization ability. Experimental evaluation were conducted using three different algorithms with TL, Results show that this approach achieves superior recognition accuracy and generalization ability.

Facial expression images are classified using transfer learning in [Orozco et al., 2018] wherein the images were categorized into eight different emotions. CNN was used for facial expression recognition wherein AlexNet, Virtual Force Field and ResNet models were trained using ImageNet dataset which is a larger dataset consisting of 1.2 million samples with different categories. The weights of all the network architectures were fine-tuned using TL which helped to achieve a superior accuracy of 90% during testing.

2.1. 3-D FER Techniques

The 3D facial analysis and 3D approaches for facial expression recognition have gained significant attention in recent years. Despite the availability of different techniques, there is a great scope of research in this field. A comprehensive analysis on 3D facial analysis and 3D approaches for facial expression recognition was proposed by [Nonis et al., 2019].

The study stated that despite the challenges such as high dimensionality, computational costs, and complexity in real-time applications, 3D techniques have achieved superior recognition accuracy as compared to 2D techniques. This provides a robust reason to employ a dataset which contains 3D facial images or 3D video sequences for better facial expression recognition. A deep feature fusion CNN for 3D recognition facial expressions is proposed in [Tian et al., 2019]. This proposed approach initially represents a 2D facial feature map which includes the depth, normal, and shape index values.

Further, these facial attributes were combined to learn the facial expressions using the deep feature fusion CNN approach. The method was trained to overcome the problem of overfitting. Experimental analysis was conducted on Bosphorus database and results show the efficacy of their proposed approach. Another similar approach for recognizing 3D facial expressions was presented by [Vo et al., 2019]. The proposed approach was based on multi view and prior knowledge fusion. The proposed approach has mainly learnt the expressions from the facial images and it also predicts the emotions of the images by analyzing the facial prior information. The model was tested using a BU-3DFE dataset which was obtained from a public 3D facial expression dataset. Results show that the proposed approach can learn more distinguishing features as compared to DL based approaches.

Deep learning based approaches have advantages such as better classification and prediction accuracy. Various researchers have used deep learning techniques for 3D facial analysis. These techniques can recognize different emotions from a single RGB image as shown in [Koujan et al., 2020]. This proposed work constructed an image dataset which consist of 3D pose variations. This dataset was employed to train a deep CNN model for accurately analyzing and predicting 3D expressions and their variations. Experimental analysis show that the proposed approach outperforms other 3D facial recognition techniques.

A point-based deep neural network for 3D facial expression recognition for identifying all prominent features of the facial images using different sets of key points is proposed in [Trimech et al., 2020]. This work also proposed to overcome the issue of overfitting of the neural network by effectively training the model on a larger dataset. Results of the experimental evaluation show that the proposed 3D facial recognition approach achieve a better performance in comparison to other techniques.

2.2.3 TL (Transfer Learning)

A part-centered Hierarchical Bidirectional Recurrent NN for analyzing different types of facial recognition in terms of temporal sequences was discussed by [Zhong et al., 2019]. Their proposed PHRNN model dynamical evolution and facial morphological variations of expressions has analyzed. The PHRNN approach allows the extraction of temporal features from consecutive frames using facial landmarks (geometry information). Meanwhile, a multi-signal CNN (MSCNN) is planned for extracting “spatial features” by still frames to supplement the still appearance information. They measure various loss functions using both verification and recognition signals as supervision, which helps in increasing the variations of diverse expressions while reducing the differences between similar expressions. This deep evolutional spatial-temporal network (STN) made up of MSCNN and PHRNN extracts geometry-appearance, dynamic-still, and partial-whole data to improve FER accuracy. Experiments validate that this technique outdoes State- of- the-art (SOTA) approaches through a wide margin.

2.2. Micro Facial Expression

Various approaches have been proposed for the recognition of micro facial expressions. A study conducted by [Takalkar et al., 2018] stated that as compared to basic-expressions, it is difficult to identify micro-expressions since their time frames are reduced to a fraction of a second and can only be identified using a larger classification scale. A similar approach based on CapsuleNet for recognizing facial expressions using apex frames was proposed by [Van Quang et al., 2019]. Their proposed approach reduced the computational complexity of the model and enhanced the generalization ability of the models while working on small-micro expression datasets. The performance of the proposed approach was validated using a cross-database micro-expression benchmark Experimental evaluation showed that the proposed approach achieved significantly higher accuracy compared to LBP-TOP approach and other CNN models.

A ML-based approach for recognizing micro facial features using an Extreme Learning Machine (ELM) approach was proposed by [Adegun and Vadapalli, 2020]. The ELM approach has the ability of fast learning compared to other ML algorithms. Support Vector Machine (SVM) was employed as a comparative model whose performance was tested with ELM model. The essential features were extracted on apex micro-expression frames using LBP. Results show that the proposed ELM model achieved better recognition accuracy compared to SVM. Also the learning time of ELM was faster than SVM.

A micro FER approach based on deep learning was discussed by [Takalkar et al., 2017]. The input image data for the proposed approach was collected from smaller datasets. The study generated an extensive training datasets which consists of synthetic images using data augmentation techniques. This study proposed a CNN based model for recognizing micro facial expressions. The performance of the proposed approach was validated using a thorough experimental analysis which proved the potential ability of their proposed model in recognizing micro facial expressions.

A three-dimensional flow based CNN model for video based micro-expression recognition was used for recognizing the micro facial expressions from the image database [Li et al., 2019]. Their proposed model is trained to extract relevant features which can easily characterize the fine motion flow of the images to analyze the facial movements. Their proposed model yields better results in terms of recognition accuracy.

A comprehensive analysis of different transfer learning based techniques for micro facial expression recognition is presented in [Peng et al., 2018]. The main aim of their proposed approach is to identify micro expression in two databases namely holdout-database recognition and composite database. Transfer learning was used to implement CNN for recognizing micro expressions. The results of the experimental analysis showed that the proposed approach achieved a better recall score for holdout-database as compared to other existing models.

2.3. Data Sets

A wide range of the datasets are used by the researchers. Some of the popular datasets are discussed here:

2.4.1 Basic FER Dataset

- **The Extended Cohn-Kanade (CK+) Dataset for micro facial expressions**

This dataset has 593 consecutive gray images of 123 titles, but only 309 labels have been followed. Each photo sequence begins with a neutral face and concludes with a facial expression. An images are 640 x 490 pixels in size. CK+ datasets images are shown in Figure 3.

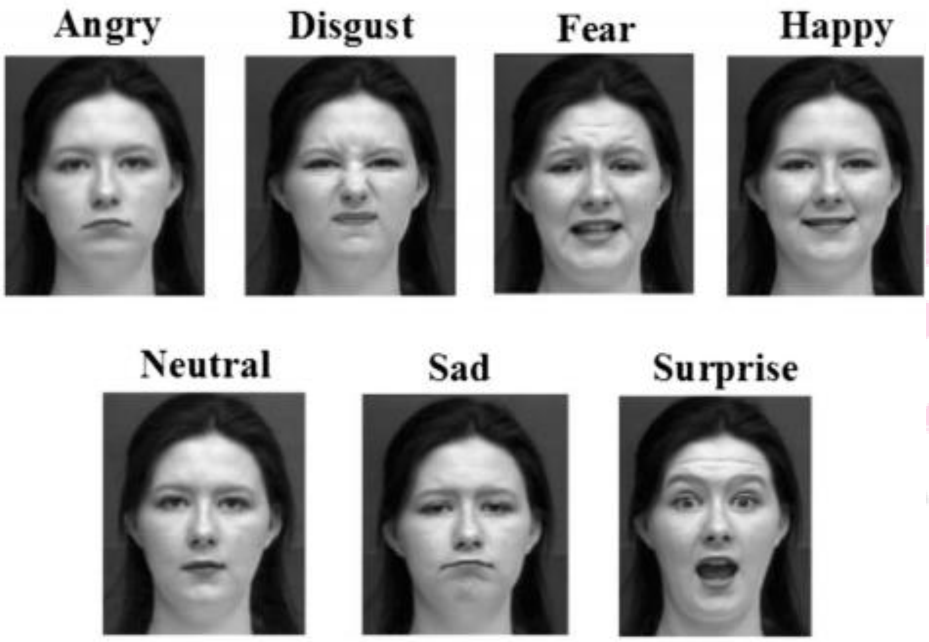

Figure 3: Sample Images of CK+ Datasets Individually

- **Japanese Female Facial Expression (JAFFE) Dataset**

This dataset contains 213 photos of 60 Japanese women. The gray images' size is 256 × 256 pixels .

Figure 4: Sample Images of JAFFE Datasets

- **Indian Spontaneous Expression Dataset (ISED)**

This database usages 428 videos found in 50 studies that seen videos which are emotional. By those videos, 425 photos released. Not like other databases, this one contains just 4 FEs (disgust, joy, sadness, and surprise).

The image size is 1,920 × 1,028, so they have a good resolution [Happy et al., 2015]. Indian Spontaneous Expression Dataset sample images shown in figure 4.

Figure 5: Sample Images of ISED Datasets

- **MMI (Man Machine Interface) Dataset**

This database contains 205 image sequences. About 500 images of 890 × 550 pixels were taken. Sample pictures of MMI dataset are shown in Figure 5 [Mokaya et al., 2020].

Figure 6: MMI Dataset's Sample Images

- **MUG (Multimedia Understanding Group) Dataset**

There are 658 image sequences in this set. All image sequences begin and end with a neutral voice, just like the MMI database; Facial Expression is in the mid of the sequence; 1,432 images of 896 x 896 pixels were made. The MUG Database is shown in Figure 6.

Figure 7: The MUG Face Database

2.4.2 Micro Expression Datasets

- **CAS(ME)b^2 Dataset**

The CAS(ME)b^2 dataset contains an overall 303 expressions out of which 250 are macro-expression samples and 53 are micro-expression samples. The expression samples were selected from more than 600 elicited facial movements and were coded with the onset, apex, and offset frames. The expression samples were recorded using a Logitech Pro C920 camera at 30 fps, and the resolution was set to 640 x 480 pixels.

The samples of the dataset with micro facial expressions are illustrated in Figure 8.

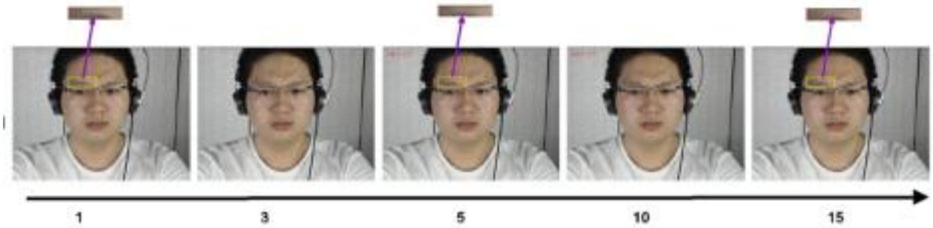

Figure 8. CAS(ME)b^2 dataset containing micro facial expressions

- **CASME II Dataset**

The CASME II dataset consists of 247 micro-expression samples collected from 26 participants. These samples are selected from nearly 3,000 elicited facial movements and are coded with the onset and offset frames, Action units (AUs)are marked and emotions are labelled and five main categories are provided such as happiness, disgust, surprise, repression, and other emotion-related facial movements. Micro expressions from a frame sequence using CASME II dataset are illustrated in Figure 9.

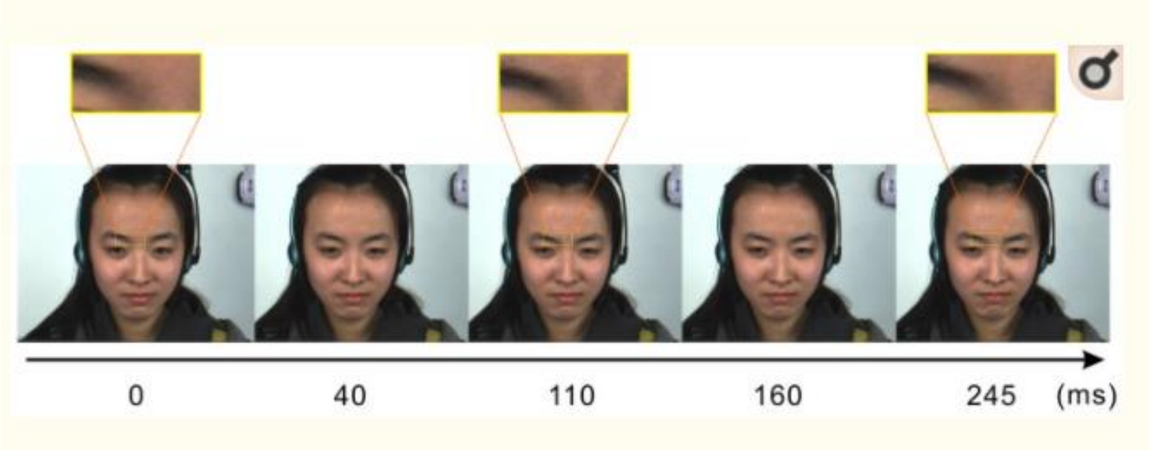

Figure 9. An illustration of frame sequence in a micro expression

- **SAMM Dataset**

The SAMM dataset contains micro-movements captured at 200 fps. The dataset consists of 159 spontaneous micro-facial movements . The micro-facial movements are obtained from 32 participants belonging to 13 different ethnicities with a mean age of 33.24 years, standard deviation of 11.32 with an age group between 19 and 57 where 16 were female participants and 17 were male participants .

2.4.3 3D FER Datasets

- **BU-3DFE Dataset**

The Binghamton University 3D Facial Expression (BU-3DFE) database consists of 100 subjects wherein 56% samples belongs to females and 44% samples belongs to male with an age group of 18 years to 70 years old, with different ethnic/racial ancestries, including White, Black, East-Asian, Middle-east Asian, Indian, and Hispanic Latino. Each user gave 7 expressions in front of the 3D face scanner. Excluding neutral expression, other expressions such as happiness, disgust, fear, angry, surprise and sadness contained four intensity levels. Considering this, there are 25 instant 3D expression models for each subject, resulting in a total of 2,500 3D facial expression models in the database. The database consists of 2,500 two-view's texture images and 2,500 geometric shape models.

The sample expressions of male and female subjects are illustrated in figure 10a and 10b respectively.

Figure 10a. Sample expressions of male subjects

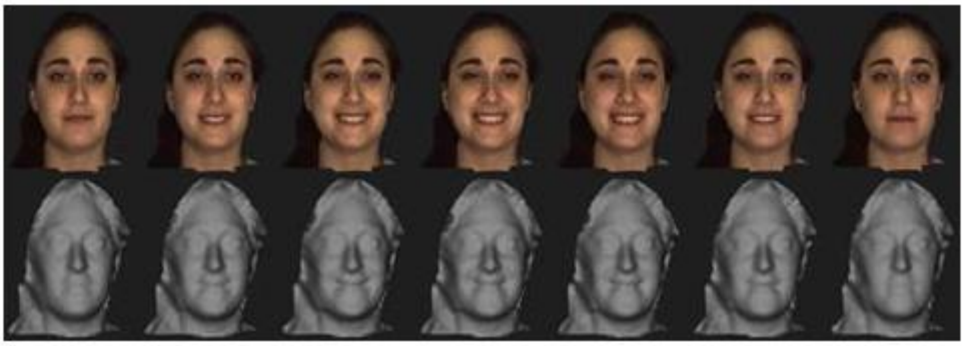

Figure 10b. Sample expressions of female subjects

- **Bosphorus Dataset**

  This dataset consists of 105 subjects in different poses, expressions and occlusion conditions. 18 subjects have beard/moustache and 15 subjects have short facial hair. Most of the subjects belong to an age group of 25 to 35 out of which 60 subjects are men and 45 subjects are women. In addition, 27 professional actors/actresses are included in the database. Up to 54 face scans are available per subject, where 34 subjects have 31 scans. Hence, the number of total face scans is 4652. Each scan is labelled manually labelled for 24 facial landmark points such as nose tip, inner eye corners, etc,

Figure.11 Sample Image for Bosphorus Dataset

3. Challenges and Research Gaps

One of the most difficult aspects of a FER task is coping with the high variability of data, for example, Facial Expressions can be influenced by expressiveness, race, or personality. When identifying Facial Expressions, the head posture and lighting conditions are also shown. Since it consumes time and also it is complex, numerous model functions points or specific areas are labelled by hand. Both face identification and changes are represented in the expression images.

As a result, derived characteristics are frequently a combination of identity characteristics and expression changes. After lots of studies shows how different directions and intensities of light on the face influence feature extraction, but few studies show how obscure main areas like the eyes, forehead, and mouth affect feature extraction and thus interfere with FER. However, there are certain challenges which needs to be addressed in order to achieve high facial recognition accuracy: Some of the prominent research challenges identified by this research are listed here:

- A major issue with DL is that it needs a huge amount of data to train good models.
- Despite the advances of DL in FER methods, there are two issues with the current databases: i.e., images are smaller in numbers and images are taken in extremely managed environments.
- It is highly challenging to perform feature extraction from 3D facial images due to the issue of high dimensionality in 3D image datasets.
- The complexity of facial recognition process increases with the changes in the intensity of lights and due to varying postures.

4. Objectives

1) To develop a novel design to classify the images of human face expression into different categories of emotions using CNN.
2) To develop pre-processing methods and methods of extraction for improving the precision of different databases.
3) To incorporate a deep learning based feature extraction and classification technique for identifying different types of images such as basic expressions, 3D, and micro facial expressions.
4) To achieve the highest levels of recognition accuracy through CNN formatting.

5. Methodology

In this section, CNN architecture and techniques to be used to get the accuracy on various datasets will be discussed. The workflow of the proposed approach is shown in Figure 12

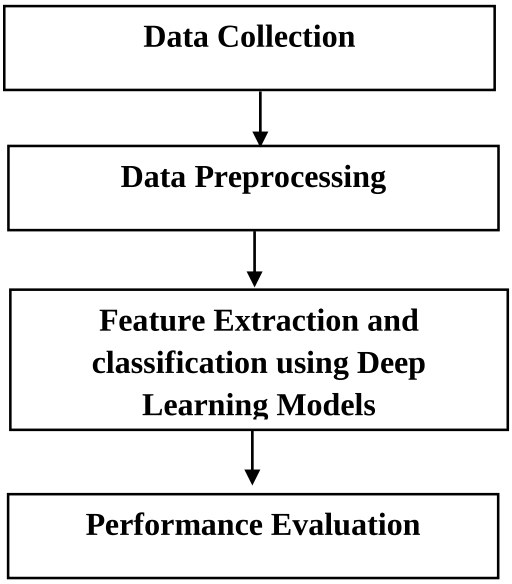

Figure 11. Workflow of the proposed approach

5.1 Pre-processing

Pre-processing can be utilized for improving the working of FER system and can be used to incorporate an extraction process. Pre-image processing involves a variety of processes, such as face alignment and detection, lighting adjustment, positioning, closing, and expansion of data.

When images are taken with different kinds of light, the display features are sensed incorrectly, therefore, the visual quality of the speech may be low and make the rendering of the feature very difficult.

5.2 Deep Learning Model

In this research, the CNN model is trained for recognizing different facial expressions. The CNN model is provided with facial expressions extracted from the input data. Feature extraction is an important step in the facial recognition process. In this step, the relevant and important features from the input data are extracted. The extracted features will be used by the CNN model for facial recognition.

CNNs contain artificial neurons' various layers. Artificial neurons are mathematical functions that calculate the activation value by weighing many inputs. They're intricate replicas of their biological counterparts. Every neuron's activity is determined through its weight. CNN artificial neurons choose a range of viewing characteristics when used with pixel values. NN is a set of algorithms that are designed for simulating the human brain and identify relationships amongst data to finding solutions.

Convolution is the mathematical function used to determine the relationship between data in CNN, which is a type of NN. When it comes to complex issues like image fragmentation, video fragmentation, pattern recognition, and so on, the Traditional NN (TNN) fails. CNN, on the other hand, has had a lot of success in these shows, offering outstanding accuracy.

In CNN, each layer creates multiple activation maps. Activation maps highlight the most important aspects of an image. Every neuron takes a pixel as input, multiplies it by its weight, shortens it, and then activates it in the activation feature.

Basic features such as horizontal, vertical, and broken edges are typically found in CNN's first (or lower) sheet. The installation of the next layer replaces the removal of the first layer, which eliminates complicated features like corners and composite combinations.

Figure 12 illustrates the CNN (Camacho Cezanne, 2018).

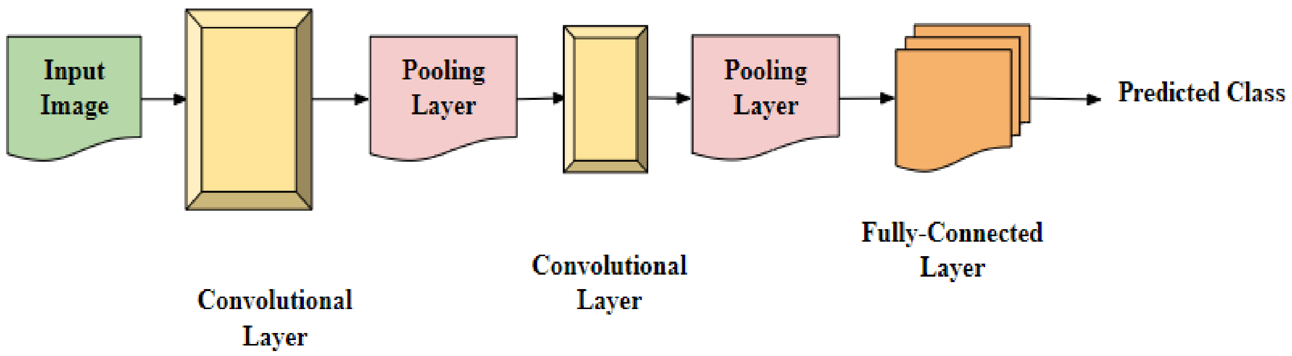

Figure 12: Convolutional Neural Network

5.4. Performance Evaluation

The performance of the proposed CNN-based facial expression recognition approach will be validated in terms of different performance metrics such as accuracy, precision, F1-score, and recall. In addition, the results of the proposed approach will be compared with other existing techniques in order to prove the effectiveness of the proposed approach.

6. Expected Outcomes

The aim of this research is to create an innovative architecture that uses CNNs to categorise images of human faces into distinct emotion groups, as well as pre-processing and feature extraction techniques for increasing precision on a different dataset. Our purpose is also to get the accuracy of seven FER by using CNN architecture. As deep networks necessity a large database for the training, various databases will be used to get improved results.

The major benefits of this research are as follows:

- A technique for recognizing the seven FEs centred on features extracted with CNNs will be proposed.
- Combining images from various databases aids in generalization will be showing.
- A framework for recognising MEs will be presented, as well as three techniques for creating a temporal-aggregated function vector.
- Lastly, there will be a prototype framework that can recognise expressions and MEs in various image datasets.

7. References

Adegun, I. P., Vadapalli, H. B. (2020). Facial micro-expression recognition: a machine learning approach. *Scientific African*, *8*, e00465.

Aung, Darli, Myint, Aye, Nyein,. “A Facial expression classification using histogram-based method.” *International Proceedings of Computer Science and Information Technology* 58 (2012): 1.

Camacho Cezanne, “Convolution Neural Networks”, June 3 2018, [online]. Available: https://cezannec.github.io/Convolutional_Neural_Networks/.

Cheng, Fei, Jiangsheng Yu, and Huilin Xiong. "Facial expression recognition in JAFFE dataset based on Gaussian process classification." *IEEE Transactions on Neural Networks* 21, no. 10 (2010): 1685-1690.

Davison, A. K., Lansley, C., Costen, N., Tan, K., & Yap, M. H. (2016). Samm: A spontaneous micro-facial movement dataset. *IEEE transactions on affective computing*, *9*(1), 116-129.

Happy, S. L., Priyadarshi Patnaik, Aurobinda Routray, and Rajlakshmi Guha. "The Indian spontaneous expression database for emotion recognition." *IEEE Transactions on Affective Computing* 8, no. 1 (2015): 131-142.

Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. “Reducing the dimensionality of data with neural networks.” *science* 313.5786 (2006): 504-507.

Jiskoot, Lize C., Jackie M. Poos, Manon E. Vollebergh, Sanne Franzen, Judy van Hemmen, Janne M. Papma, John C. van Swieten, Roy PC Kessels, and Esther van den Berg. "Emotion recognition of morphed facial expressions in presymptomatic and symptomatic frontotemporal dementia, and Alzheimer’s dementia." *Journal of neurology* 268, no. 1 (2021): 102-113.

Koujan, M. R., Alharbawee, L., Giannakakis, G., Pugeault, N., Roussos, A. (2020, November). Real-time Facial Expression Recognition “In The Wild” by Disentangling 3D Expression from Identity. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)* (pp. 24-31). IEEE.

Li, Q., Zhan, S., Xu, L., & Wu, C. (2019). Facial micro-expression recognition based on the fusion of deep learning and enhanced optical flow. *Multimedia Tools and Applications*, *78*(20), 29307-29322.

Lucey, Patrick, Jeffrey F. Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. "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, 2010.

Merghani, W., Davison, A. K., Yap, M. H. (2018). A review on facial micro-expressions analysis: datasets, features and metrics. *arXiv preprint arXiv:1805.02397*.

Mokaya, Victor, and Mukesh Kumar Gupta. "Performance evaluation of MMI dataset on six macro expressions using 2D-LDA and SVM kernels." *Performance evaluation* 7, no. 19: 2020.

Moschoglou, Stylianos, Athanasios Papaioannou, Christos Sagonas, Jiankang Deng, Irene Kotsia, and Stefanos Zafeiriou. "Agedb: the first manually collected, in-the-wild age database." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 51-59. 2017.

Nonis, F., Dagnes, N., Marcolin, F., Vezzetti, E. (2019). 3D Approaches and challenges in facial expression recognition algorithms—a literature review. *Applied Sciences*, *9*(18), 3904.

Orozco, D., Lee, C., Arabadzhi, Y., Gupta, D. (2018). Transfer learning for facial expression recognition. *Florida State Univ., Tallahassee, FL, USA, Tech. Rep*, *7*.

Peng, M., Wu, Z., Zhang, Z., Chen, T. (2018, May). From macro to micro expression recognition: Deep learning on small datasets using transfer learning. In *2018 13th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2018)* (pp. 657-661). IEEE.

Qu, F., Wang, S. J., Yan, W. J., Fu, X. (2016, July). CAS (ME) 2: A Database of Spontaneous Macro-expressions and Micro-expressions. In *International Conference on Human-Computer Interaction* (pp. 48-59). Springer, Cham.

Ramalingam, S., Garzia, F. Facial expression recognition using transfer learning. In *2018 International Carnahan Conference on Security Technology (ICCST)* (pp. 1-5). IEEE. 2018.

Reddy, S. P. T., Karri, S. T., Dubey, S. R., Mukherjee, S. Spontaneous facial micro-expression recognition using 3D spatiotemporal convolutional neural networks. In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE. 2019.

Savran, A., Alyüz, N., Dibeklioğlu, H., Çeliktutan, O., Gökberk, B., Sankur, B., Akarun, L. (2008, May). Bosphorus database for 3D face analysis. In *European workshop on biometrics and identity management* (pp. 47-56). Springer, Berlin, Heidelberg.

Sharma, Garima, Latika Singh, and Sumanlata Gautam. "Automatic Facial Expression Recognition Using Combined Geometric Features." *3D Research* 10, no. 2 (2019): 1-9.

Takalkar, M., Xu, M., Wu, Q., Chaczko, Z. (2018). A survey: facial micro-expression recognition. *Multimedia Tools and Applications*, *77*(15), 19301-19325.

Takalkar, M. A., Xu, M. (2017, November). Image based facial micro-expression recognition using deep learning on small datasets. In *2017 international conference on digital image computing: techniques and applications (DICTA)* (pp. 1-7). IEEE.

Tian, K., Zeng, L., McGrath, S., Yin, Q., & Wang, W. (2019, June). 3D facial expression recognition using deep feature fusion CNN. In *2019 30th Irish Signals and Systems Conference (ISSC)* (pp. 1-6). IEEE.

Trimech, I. H., Maalej, A., Amara, N. E. B. (2020, September). Point-based deep neural network for 3D facial expression recognition. In *2020 International Conference on Cyberworlds (CW)* (pp. 164-171). IEEE.

Van Quang, N., Chun, J., Tokuyama, T. (2019, May). CapsuleNet for micro-expression recognition. In *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)* (pp. 1-7). IEEE.

Vo, Q. N., Tran, K., Zhao, G. (2019, September). 3D Facial Expression Recognition Based on Multi-View and Prior Knowledge Fusion. In *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)* (pp. 1-6). IEEE.

Xu, Lei, Fei, Minrui, Zhou, Wenju, Yang, Aolei. "Face expression recognition based on convolutional neural network." *2018 Australian and New Zealand Control Conference (ANZCC)*. IEEE, 2018.

Yan, W. J., Li, X., Wang, S. J., Zhao, G., Liu, Y. J., Chen, Y. H., Fu, X. (2014). CASME II: An improved spontaneous micro-expression database and the baseline evaluation. *PloS one*, *9*(1), e86041.

Yin, Lijun, Xiaozhou Wei, Yi Sun, Jun Wang, and Matthew J. Rosato. "A 3D facial expression database for facial behavior research." In *7th international conference on automatic face and gesture recognition (FGR06)*, pp. 211-216. IEEE, 2006.

Zhong, L., Bai, C., Li, J., Chen, T., Li, S., Liu, Y. (2019, May). A graph-structured representation with brnn for static-based facial expression recognition. In *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)* (pp. 1-5). IEEE.