ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

Volume 11, Issue 2, March 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

IMPACT FACTOR: 7.583

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

A Systematic Review of Facial Expression Detection Methods

S.PRAGADEESWARAN, ASHWIN S, AJITH KUMAR R, DHARNEESH V

Assistant Professor, Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

ABSTRACT: In this study, recent developments in FER were presented, and presented studies allowed us to track the latest developments in this area. Over the past year or so, a number of different researchers have devised different CNN architectures and some outside the lab produced reference databases. To facilitate an accurate emotional detection, we need to have provided previously obtained as well as experimental tables (spontaneous as well as lab) (emotion as reference). Also, we introduce a discussion which emphasizes the fact that machines are already able to recognize more complex emotions, implying that the emergence of human-machine collaboration will become more and more commonplace.

KEYWORDS: facial expression recognition; emotion recognition; computer vision; machine learning.

I. INTRODUCTION

In the last decade, CNNs have done well for FER, as shown by their use in a number of cutting-edge algorithms. Many FER competitions, including the previous year's EmotiW challenge, were won by a kind of CNN architecture with few layers. Facial emotion recognition has served the public well for decades prior to the field of deep learning breaking, and a group of brilliant researchers has tried to stay abreast of the current research efforts in that field, while others have undertaken to learn from its methods and discoveries. In recent times, many researchers offered novel and recurring practices for applying deep learning in order to security problems in an effort to enhance detection. Validation users currently do additional validation on a number of static or sequential databases before allowing their information to be used in a live database.

The VGG-16 model (developed by the University of Oxford's Visual Geometry Group (VGG)) may be considered a watershed moment in the history of deep CNN models. It was pretrained using the ImageNet database to extract features from images that might be used to distinguish between image classes. Numerous recent studies show that VGG-16 performs well on image recognition and classification datasets from a variety of fields. proposed Deep Convolution Neural Networks (DCNNs) which are used in the cross-database search. After that, facial images had to be reduced to 48 × 48 pixels; the rest of the same pictures had to be searched for locations and landmarks to be extracted. Finally, they had augmented the database with additional data, and only then did they were able to create it. Subsequently, the data moves on to two classification stages where the softmax (SF) is expanded and fed into the fully connected softmax (XF) network after the first classification stage. To avoid overfitting, they suggest using local CNNs in combination with convolutional layers that are fine-tuned for specific use cases.

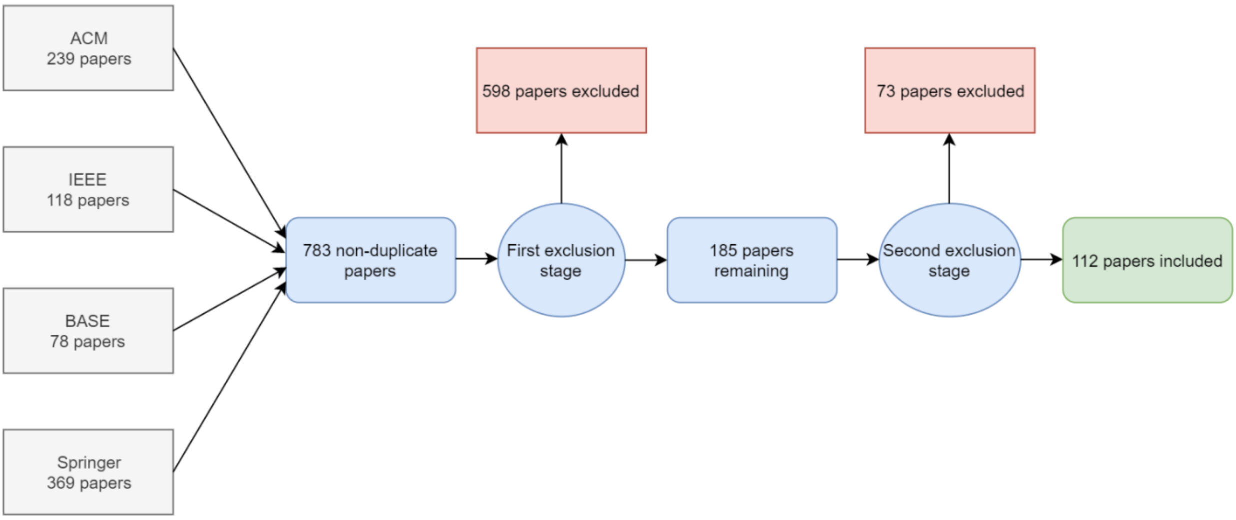

Fig 2: Facial Emotion Recognition Using

The authors have shown that the results prior to training were used to discover how to influence the final outcome When it expanded, the first CNN expansion, when it lowered the size to 32 × 32 and also used data normalization with 8-connected pools followed by downsampling (normalization of 32 × 32 to a 256 final dimension), and when that was done, cropping was employed. Gaining the most mass is something that happens only at the competition, so the athletes who have gained the most muscle will play in the games. The information used this third party search tool to assemble a total of three transparently accessible databases: the CK+ and JAFFE, as well as the BU3DF. One also discovers a wide range of beneficial practices when considering these studies, such as utilizing all of these techniques and products together; studies show the difference between these things yielding different results.

II. RELATED WORK

There are two primary techniques for designing a FER system. As an initial step, some systems employ a sequence of images ranging from a neutral face to the peak level of emotions. In comparison, some systems use a single image of the face to recognize related emotions, and because they have access to less information, they often perform worse than leading approaches [2,3]. Apart from the approach type modeled by a FER system, another classification is based on the type of features employed in the recognition process, with a FER system utilizing one or both of these feature categories. The first set of traits is obtained from the facial organs' posture and the skin's texture. The second type of feature is geometric features, which hold information about various positions and points on the face and are used to analyze a static image or a sequence of photos by utilizing the movement of the positions and points within the sequence. Using face landmarks as a starting point for extracting geometric features is one way. Landmarks are significant places on the face that provide useful information for facial analysis. Numerous studies have been undertaken on the subject of facial landmark identification; however, they are outside the scope of this work. This work employs the *Python* module dlib to detect these points.

Two different things, which are involved in automatically identifying human emotions and psychology, are part of artificial intelligence. The big question that researchers are attempting to answer in the area of psychology and artificial intelligence is the identification of emotions. One of them includes both topics such as mood and accent, which is generated in both vocal and nonverbal sensors such as the tonality and aural alterations [5] which are widely accessible, for instance [6], and a quick mood assessment can be obtained, as well as other sources [7]. Results from Mehrian's study [8] demonstrated that 55% of information was sensory (emotional and verbal), with the remaining 7% being percent of having an unspecified physical component. The first indication a person gives of their emotional state being in a state is facial expressions, so many researchers are very interested in this modality.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

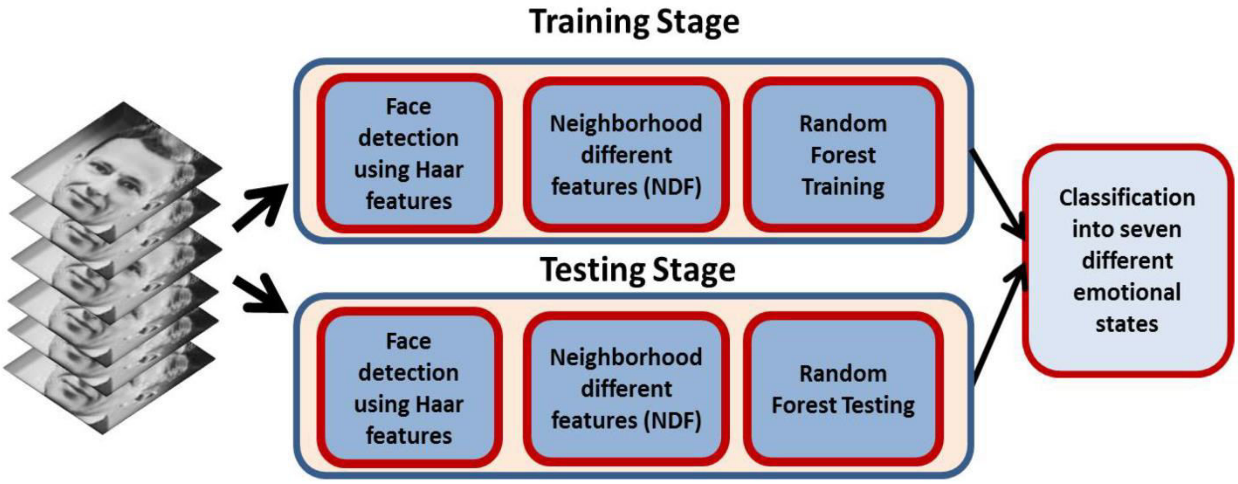

Fig 2: Work Flow

First working in the extraction feature space to add new features to an existing representation can be a good thing because it will help the other features as well. Ekman and Friesen [9] noted that the Facial Action Coding System (FACS) and facial movement action units (AUs) assume that each coded movement in FACS involves at least one facial muscle. Ekman and Friesen first recognized that FACS facial movement, in FACS facial AUs, is utilized in such a way that face muscles and facial muscles are coded for each of their head movements (between several different individuals and/or races of subjects).

III. METHODS

Face detection is the initial stage in the face recognition process and is essential to the system's overall effectiveness. Faces in films may be recognized using visual cues such as facial expression, skin tone, or movement in the film. Numerous effective methods are restricted to improving the appearance of the face. This may be because these algorithms avoid the challenges associated with representing 3D structures like faces. The face/nonface border, on the other hand, may be extremely complicated, and it is necessary to use 3D variations in order to identify facial emotions.
The feature extraction technique transforms pixel data into higher-level representations of the face in the image, such as texture, color, motion, contours, and the spatial arrangement of the face. This collected data is then used to assist in the detection of trends throughout future categorization procedures. During the feature extraction procedure, the dimensions of the input space are typically decreased. It is critical to retain information with high stability and discrimination while also maintaining a high level of stability during this process. To identify a person's face, a number of unique characteristics are utilized.

It is only in the emotional mechanisms that a lack of progress was recently discovered. It has been found that emotional mechanisms take precedence over rational processes in the brain, which can be seen as either being advantageous or detrimental depending on their presence or absence. Worse feelings give rise to negative thoughts, which tend to dampen a person's creativity when looking for solutions to the challenges at hand and lead to them getting you into deeper trouble. It has been found that states such as anger, sadness, fear, and happiness each have their own distinct patterns of blood flow to the brain and have an influence on that of mood.

A great deal of research has demonstrated that positive emotions like joy, acceptance, trust, and satisfaction can assist learning, while negative emotions can bring about learning disabilities and affect the process. Anxiety and depression can hinder memorization in various ways. These states can show up in different ways, such as causing stress, which increases with despair, and leads to increased feelings of anger and fear, or fear, or stress itself may lead to worse than depression. When it is difficult for a student to acquire information, intelligent feedback can help them overcome their lack of motivation. For the latter, the computer should be able to know what learners are feeling, give learners opportunities to expand on their understanding, handle their interests, and provide them with pertinent information and timely feedback depicts end-to-end face recognition processing flow.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

IV. RESULT ANALYSIS

MobileNets, they developed a set of efficient convolutional neural models. A class of efficient MobileNets convolution models were developed by him, as part of a class called MobileNets. The neural architecture has been designed with a novel type of convolution called depth separable and is made up of factors, instead of connected layers, of depthwise convolutions. Depthwise convolutions are two layers: they are separable, and the first is a depthwise layer consisting of two separable convolutions. According to normal convolution, it is over 10 times more efficient. However, it only applies lowpass and highpass effects to the signal; it does not add any new features. To do this, they did the addition of other operations such as pointwise convolution (e.g., which does the sum of pointwise convolution outputs) and implemented 1x pointwise convolution, respectively. By adopting two additional hyperparameters, they try to increase their efficiency. Thus, the network can be made thinner by the increase in network width by a multiplicative factor and resolution by a nonuniform , and cost reduction can be done at each layer. Different hyperparameters allow the model builder to choose a model that will have just the right number of parameters for their own application, but without error. Various methods and functions are demonstrated by this model with different examples, with facial features as well as the measurement of object dimensionality.

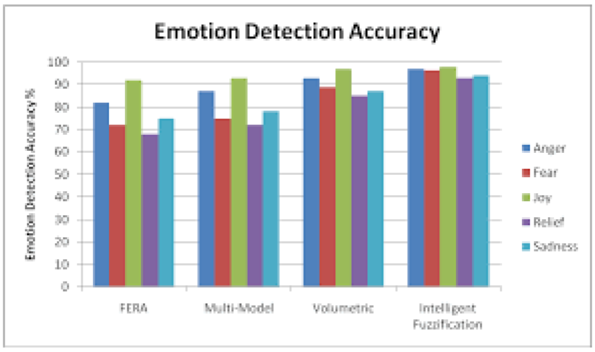

Fig 3: Result analysis

Recognizing emotions is difficult because they are ambiguous and therefore prone to error, but in many cases, there are various things that can be used to discover them. Ekman claims that there are eleven basic emotions found in human facial expressions which can be classified into seven groups: happiness, anger, sadness, fear, disgust, surprise, and contempt [67]. This after the turn of the millennium got a boost from successful experiments in face recognition and audio-visual media and has paved the way for further research on automatic affect recognition of effect. Face expression popularly suggests that recognition of emotions is simply a method to look for patterns that indicate whether or not a person is empathetic popularly known as "to judge by the look on their face". The FACS is used to code different types of facial actions, which include facial motions, and numerous AUs are created; each one of these classes contains unique entities; finally, emotions are determined by AU designations.

V. CONCLUSIONS

The interest in FER is growing gradually, and, with it, new algorithms and approaches are being developed. The recent popularization of Machine Learning made an obvious breakthrough in the research field. The research in FER is definitely in the right path, walking together with important fields like psychology, sociology, and physiology. From this, more and more accurate FER systems are emerging every year. However, despite this obvious progress, pose-variant faces in the wild are still a big challenge for FER systems. However, there are emotion recognition challenges every year that explores this problem and, with it, FER systems are becoming robuster to pose-variant scenarios. Especially after a major breakthrough done by a CNN called AlexNet [181], which achieved a top-5 error of 15.31% in the ImageNet 2012 competition, more than 10.8% points lower than that of the runner up. After this, researchers became aware of the potential in CNNs for solving Computer Vision problems, and more FER systems using CNNs emerged, correlating with overall better results. The only potential negative aspect to point out from the reviewed works is that none considered the environment context. Although most works are giving the right steps towards multimodal systems, the environment context seems to be ignored. For instance, if there is an image of a birthday party, the happy context has a huge weight in the mood of people participating in it, which can't be ignored even if a certain participant is not explicitly smiling. Nevertheless, FER systems are being stimulated by yearly challenges and by the overall interest in numerous fields, achieving better results year by year.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

REFERENCES

1. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006. [**Google Scholar**]
2. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [**Google Scholar**] [**CrossRef**] [**PubMed**]
3. Coan, J.A.; Allen, J.J. Frontal EEG asymmetry as a moderator and mediator of emotion. *Biol. Psychol.* **2004**, *67*, 7–50. [**Google Scholar**] [**CrossRef**] [**PubMed**]
4. Zafeiriou, S.; Zhang, C.; Zhang, Z. A survey on face detection in the wild: Past, present and future. *Comput. Vis. Image Underst.* **2015**, *138*, 1–24. [**Google Scholar**] [**CrossRef**] [**Green Version**]
5. Perez, L.; Wang, J. The effectiveness of data augmentation in image classification using deep learning. *arXiv* **2017**, arXiv:1712.04621. [**Google Scholar**]
6. Tian, Y.I.; Kanade, T.; Cohn, J.F. Recognizing action units for facial expression analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2001**, *23*, 97–115. [**Google Scholar**] [**CrossRef**] [**Green Version**]
7. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, Pittsburgh, PA, USA, 27–29 July 1992; ACM: New York, NY, USA, 1992; pp. 144–152. [**Google Scholar**] [**CrossRef**]
8. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [**Google Scholar**] [**CrossRef**]
9. ACM Digital Library. Available online: **https://dl.acm.org/** (accessed on 26 September 2019).
10. IEEE Xplore Digital Library. Available online: **https://ieeexplore.ieee.org/Xplore/home.jsp** (accessed on 26 September 2019).
11. Bielefeld Academic Search Engine. Available online: **https://www.base-search.net/** (accessed on 26 September 2019).

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |

www.ijarasem.com