ISSN: 0974-5823 Vol. 7 No. 1 January, 2022

International Journal of Mechanical Engineering

Facial Emotion Threshold Based Data Dispatcher

Kurakula Tanu Sri
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu India.

Kurakula Rupa Sri
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu India.

Nimmagadda Padmaja
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu India.

Rajkumar Rajasekaran
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu India.

KamalaKannan.J
School of Information Technology and Engineering, Vellore Institute of Technology Vellore, Tamilnadu India.

Abstract - Predicting human emotions is not so easy without analyzing their expressions on the face. Face emotions are one of the most informative gestures. So, we make use of these gestures on the face for the emotional recognition of the person. For recognizing the emotions of a person, we make use of most informative gestures on the face. To continuously monitor the expressions of a person and predicting the emotions of that person is difficult. So, the automation process makes it completely flexible to analyze and dispatch the emotions of a person in the form of data. The proposed idea is a hybrid model of facial emotion recognition and data dispatcher that records the feelings of an individual in each frame of a video taken via webcam and dispatches the information. Dispatching of emotional data can be used for alerting the status of a person's emotion by means of the message (mail or message) at the instant when the particular emotion will reach the threshold. This model can predict different facial expressions such as happy, sad, disgust, anger, etc in real-time. The main objectives include classifying and predicting emotions in real-time, storing the data of a person indicating the status of the emotions, and dispatch data when needed. Neural networks have been used to get better results. We have taken the model that is trained over thousands of datasets to obtain higher accuracy. However, utilizing this model can be an effective tool in mental rehabilitation and Health agencies that help monitor the patient's condition and can be used for short-term or long-term goals.

Index Terms - Facial Detection, Emotion Recognition, CNN (Convolutional Neural Network), FER-2013 dataset, Data dispatching, SMTP

INTRODUCTION

It is important to know a person's emotions while they are doing something. It can be used to find their interest in doing it. All kinds of thinking (understanding, decision making) are done by our brain's neural networks. This model tries to develop these decision-making and classification skills by training the machine. The classification skills and decision making are developed by the model which is trained by the machine. It classifies the face and predicts different emotions of a person. To acquire higher accuracy, we take the model which is trained over the FER-2013 dataset. Charles Darwin, the first scientist, stated that recognizing the expressions on the face is one of the most powerful and important means for human beings to communicate their emotions, intentions, and opinions to each other. Facial expression can also provide information about the cognitive states, such as stress, confusion, boredom and interest. Human life is a complex social structure. Humans are difficult to understand without reading other people's emotions. They do it by identifying the faces. The mood of the person decides the state of response. This concept can be used in many areas such as posting ads or what kind of products they are more interested in, by continuously observing their expression in whatever they watch. If they are watching it with more happiness and surprise, then we can say that they are more interested in it.The model identifies the face by recognizing the facial expression of a person. But in the emotion recognition of a face, the model needs to identify the emotion expressed as well as the nodal points of the face. The recent study shows that deep learning applications have shown remarkable performance in the field of computer vision.
This Application can be used by psychiatrists, company recruiters, testers in product development, etc.

- In order to predict the mental state of the patient, psychiatrists can use this technology to keep track of a patient mental state and monitor them by continuously getting updated by their emotions each second
- In company placements, recruiters can monitor the state of the candidate and can judge the candidate's personality whether they are a good fit for the job.
- To monitor the criminals by keeping track of their emotions.
- During the development of a product like video games, there is a phase where it needs to be tested to get the feedback of the users. So, the creators of the game can monitor the emotions of the tester during the gameplay and product feedback can be taken.

LITERATURE SURVEY

Facial recognition has very interesting real-life applications such as biometric systems, iris recognition, etc. However, detecting the emotions of the person is a challenging task in facial recognition systems. Facial emotion recognition comprises three steps namely face detection, feature extraction, and expression classification. Many algorithms and pre-processing techniques are proposed such as segmentation, morphological operations, and masking to detect the face and identify the emotion using skin color segmentation. For the detection of emotions in the face, a novel algorithm is proposed. Firstly, the emotion recognition and face detection are done on the frames that are extracted from the video. The quality of the image can be found by using some image quality metrics. Region of the human face can be detected by segmenting the frames into the skin and non-skin regions by using the Viola Jones algorithm and extracting the boundary of the non-skin region by using some morphological operations. After the segmentation, the expression can be identified by calculating the area of the mouth region [1]. The neural network is a set of algorithms that mimics a human brain and can be extensively used in determining age, gender. In neural networks, mathematical operations are performed to know the relationship between the data. Traditional neural networks fail in solving complex problems such as image classification, pattern recognition, etc. Whereas CNN (convolutional neural network) is a neural network that has achieved great success in solving these problems. One of the CNN models for facial emotion recognition is the two-layer convolution network model. It classifies 5 different emotions from the image dataset. To reduce the loss function, an Adam optimizer can be used. This model can be used in real-life applications to detect emotions [2]. One of the novel techniques is facial emotion recognition using CNN(FERC). The FERC happens in two stages. The background is removed by the first part and the extraction of facial feature vectors is concentrated by the second part. The model is used to find the five different types of regular facial expressions [3]. Detecting face changes during communication is the most challenging task for researchers and are interested in exploring this field. Even though many algorithms have been introduced to detect the expression it is not easy to detect the emotion accurately due to the human head pose, gender, background, luminosity and age. One of the main problems that restricts the model accuracy to certain extent is the occlusion. A high precision is achieved by the researchers in facial expression recognition (FER) by applying Convolutional Neural Network with sequential data and spatial data. In CNN, the SoftMax function and Adam optimization algorithms are used widely. Several datasets are trained to this model and observed in increased accuracy [4]. By implementing various pre-processing techniques and feature extraction techniques to the input images, the accuracy of the model can be increased. To tackle the overfitting issues, data augmentation can be used. The accuracy of 73% is obtained by training the model with the FER2013 dataset which is used for training and testing purposes. The model can be useful for the various applications like lie detection, business promotions and in the areas that require high security [5].

One way of optimizing the convolutional neural network is to increase the accuracy of facial emotion recognition. Using Random Search Algorithm, the best parameters were identified by generating and training models and are applied to the search space defined by the hyperparameters. An approx. accuracy of 72.16% is achieved when the best model is trained and evaluated by the FER2013 dataset [6].The convolutional network-based architectures like LeNet can be used for facial expression recognition. A methodology is proposed to get the high accuracy in detecting the emotions by training the dataset which is formed by combining three other datasets namely KDEF, JAFFE and the custom dataset is trained to LeNet architecture for the classification of emotional state. The accuracy obtained is almost 97% for the classification of the 7 different emotions. In this way, many works are proposed by training the different datasets to the different CNN-based architecture to improve their accuracy. The hybrid algorithm combined with CNN can give more efficient results and also the training time [7].Deep neural networks can be trained on the image datasets with facial features and emotion labels. The videos which are captured in real-time can be divided into frames and are considered as the input images to the pre-trained model. The facial expressions can be predicted and can be displayed on the user interface. There are many factors other than training; the model is to be taken into consideration for attaining maximum efficiency. The robustness of the model needs to be considered apart from the training as some factors may affect the quality of an image and the shape of the objects. The brightness, rotation, and scaling of the facial images are such factors. To Fine-tune the model introducing, fine-tuning by adopting the two-stage training procedure may also be helpful [8].

DCNN (Deep convolutional neural network) is one of the convolutional neural networks that can be used for facial emotion recognition. The proposed algorithm is to detect video-based emotion detection with no manual design for features using DCNN [9].In Facial Emotion Recognition, Deep NNs (DNNs), Convolutional Neural Networks (CNNs) have drawn attention for their feature extraction mechanism from the input images [10].One of the real-life applications of facial emotion recognition is to teach young autistic children to recognize facial expressions with the help of computer vision and image processing. In this paper, the FER2013 dataset is used to train the deep convolutional neural network (DCNN) to detect emotions. The flow of the application given is the camera points to the face of a person. The next step is facial expression recognition with the help of deep convolutional networks and computer vision. The images in the datasets have been modified to different brightness in the images to examine the model performance in different environments. Many studies have been done about the problems that children with autism disorder face to recognize the expressions [11]. Automatic Extraction of emotions is done by using ML algorithms and CNN models that help in learning and analyzing complex emotions. The Framework used in the Proposed methodology are Gabor filters that are used for feature extraction and CNN is used for classifying the emotions. As Gabor Filter extracts the sub feature of the image and gives it to a neural network, CNN's learning speed has been increased profoundly [12].In non-verbal communication, facial expression plays an important role. While investigating a subject, facial expressions cannot be recorded by the investigator as they are diverted by the surroundings, voice, or argument. So, to keep track of such emotions, an automatic method to record facial expressions may facilitate the investigator to observe the subject in more detail. Such systems can be beneficial in many more fields like psychiatry, gaming, animations, criminal investigations, etc. Geometric features (shape of the face) and Appearance features (Texture) are two main features considered for facial expression recognition. FACS is a system designed to observe the changes in facial expressions in terms of observable facial muscle actions known as face action units. CNN can be extensively used in determining gender, age, etc. [13].Facial Action Coding System (FACS) could be a helpful

structure that classifies the human facial actions by using Action Units (AU). AU is one among forty-six minor components of visible facial motion or its connection type changes. They proposed a deep learning technique within the context of emotional recognition so as to classify emotion labels from the pictures. The fundamental commitment of this work is to show the Spatio-worldly development of outward appearances of a person inside the pictures utilizing a Recurrent Neural Network (RNN) that is embedded with a Convolutional Neural Network (CNN) in the form of CNN–RNN design [14].

The arousal-valence continuous emotion space model is used to find facial emotion expression. This method uses the concept of deep learning i.e. CNN pre-trained model which is a trained feature extractor and SVR (support vector regression) helps to predict the values of the current model. So, the predicted values they recorded helps to predict facial emotion which provides better recognition results when compared to other common methods. SVR's performance proves that it works better with the arousal-valence model that can avoid the risk of overfitting and it is reliable. The future work focuses on the Soft-max layer of CNN to be replaced with the projection layer to perform the regression well [15].Static facial Expression Recognition's objective is to classify emotions in static images extracted from the movies using the Transfer Learning approach. Supervised Fine-tuning is done on a pre-trained ImageNet dataset. Fine-tuning with combined datasets gave better results when compared to the single-stage fine-tuning method [16].Recent Methods to achieve Automatic facial expression analysis have reached near to human accuracy but some cases might not get the desired results. The change in the type of parameters and number of parameters have influenced to get better results. The final classification rule that assembles the result of the committee has outperformed the traditional methods. The proposed method involves supervised learning on the ensemble computation to improve the assembling of the committee. The current state-of-art algorithms that use ensembles of CNN's give better accuracy than normal CNN classifiers. The model has a 5% higher accuracy based on averaging classifiers [17].The two important techniques of facial emotion prediction are Machine Learning and Deep Learning. The Machine Learning model achieves more comparable results with Deep Learning models when they are analyzed in emotion prediction. ML techniques are considered as a benefit with AUs as it justifies the predicted emotions by AU contribution. Clinicians and other health related professionals identify the abnormalities and the causes by recognizing the contribution of each AU through the models. This method helps to get more details and insights and is used in many related applications [18].There are multiple perspectives identified on emotion detection. 'Viola-Jones Face detection is a rearranged technique using a rearranged technique and various classifiers are used for further emotion classification and localization of faces. 'DCT transform', 'Zernike moments' or 'Local Binary Patterns (LBP)' are the techniques adapted for extracting the features. In the initial stage of facial emotion recognition, it removes the background of the image and later, facial component vectors are extracted. Using various algorithms and techniques, computer vision has played an important role in detecting the emotions by inspecting the human faces [19].In this paper, the implemented CNN network has 4 Convolutional Layers, 4 Max Pooling layers along with 1 dropout and 2 Fully Connected Layers. Altogether, the model has 899,718 parameters. The data set used is iCV MEFED (Multi-Emotion Facial Expression Dataset). The image is passed to the model and the filter (Conv2D) is passed through the image and ReLu is applied. The dimensions of the image are decreased by the MaxPooling 2D that gets the maximum values from the convolutional layers. After this step, Flatten and Dropout layers are applied. Using the data augmentation along with the same model give better results. The model predicts the emotion by deciding the highest weight of the expression among all the expressions and finalizes it. This way many neural network models which are trained with different datasets can detect facial expressions accurately and can be used in many real-life applications [20].

PROPOSED METHODOLOGY

The objective is to detect the face of a person and predict his/her emotion. The basic human emotions are classified into 7 types and are recognized universally namely sad, fear, happiness, neutral, anger, surprise and disgust.

Dataset: A dataset is used to train the model to identify the emotions. 'FER2013' is a dataset which contains the images of people showing different expressions is used to train the model by applying it to convolutional neural networks.

Face Detection: To detect the face, HAAR course work is used. The input is an image of a person which is changed to gray scale and also is resized to match with the size of the images in the dataset. The HAAR like features are used in object detection that is to identify a face of a person in a particular image.

Recognizing the emotion: The emotion of the person in the video is recognized at each frame and the identified emotion is displayed on the screen along with the video. The HAAR cascade function returns the coordinates of the face and these coordinates help in recognizing the emotions of a person.

Parsing the data and emotions: The input is a real time video which is parsed frame by frame and the output will be in the form of a text file containing emotions frame by frame. A limit is set for certain emotions like fear for which it sends the data if the limit is crossed.

TOOLS AND TECHNOLOGIES

Computer vision: Computer Vision algorithms perform image processing to extract features and use them for classification. For example, we can think that they are handcrafted.

Deep learning: We use deep learning in the form of Convolutional Neural Networks to perform facial recognition. The deep learning model learns itself and extracts features automatically.

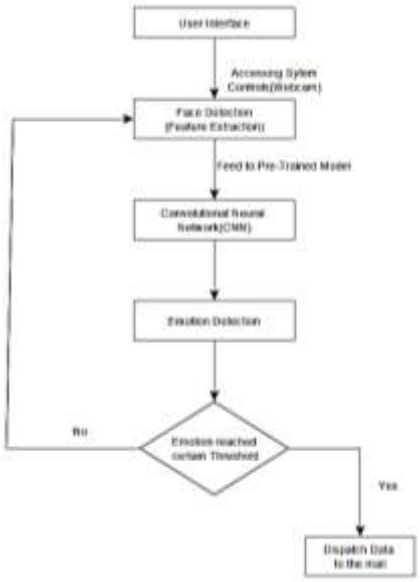
DCNN (Deep Convolutional Neural Network): In the DCNN model, the convolution and sampling layers are combined into a single layer. Based on the already trained network, the image recognition rate is greatly improved.

GPGPU (General-purpose graphic processing unit): Reduce the complete load on the CPU for the real-time use of the video capturing for the face recognition part.

SMTP (Simple Mail Transfer Protocol): It helps in handling the mailing server to mail from an email account.

Tkinter Python Module: The Tkinter module of python is to be used for the interface part of the application.

ARCHITECTURE



EXPERIMENTAL RESULTS AND ANALYSIS:

Description of the dataset used: The Dataset used here is FER-2013 that is a collection of various google images for a specific emotion.

FER2013 is an open-source dataset shared publicly in the Kaggle competition consisting of 35.887 grayscale, 48x48 sized face images with various emotions. Every image in the FER-2013 dataset is Labeled with its particular emotion. 'labelledEmotion' labels in the dataset are:

0: -4593 images- Angry, **1**: -547 images- Disgust, **2**: -5121 images- Fear
3: -8989 images- Happy, **4**: -6077 images- Sad, **5**: -4002 images- Surprise,
6: -6198 images- Neutral

Angry Disgust Fear Happy Sad Surprise Neutral

The Mini-Xception model is trained on the FER-2013 dataset as it gives higher accuracy when compared to different models.

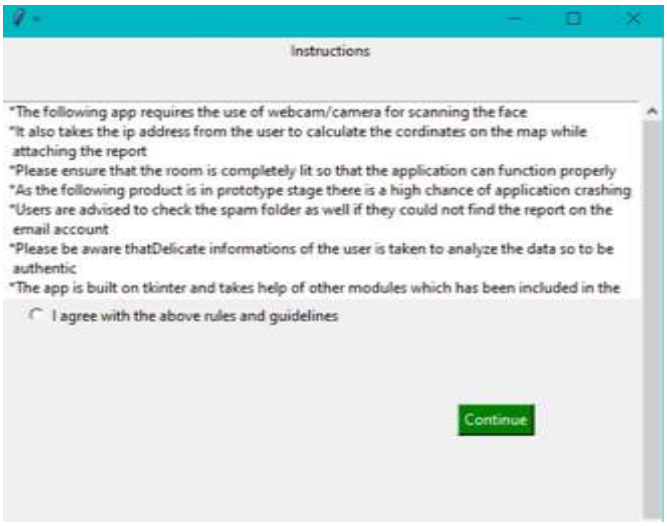
Accuracy is calculated using the Confusion Matrix. Performance metrics for experimental results are considered including Accuracy. True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are the values required to find the Confusion Matrix.

IMPLEMENTATION

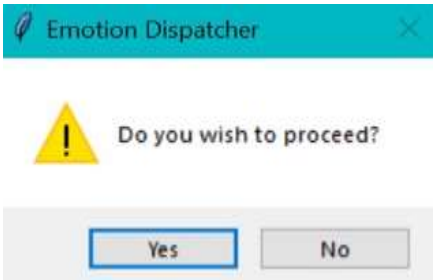
The use case considered in this paper is Psychiatry. In order to predict the mental state of the patient, psychiatrists can use this technology to keep track of a patient's state and monitor them by continuously getting updated by their emotions each second. In particular, fear recognition will be higher in residents with high anxiety and hostility scores, and enthusiastic qualities will be connected with the recognition of disgust. It is considered that accurate emotional recognition and analysis may have a significant impact on the patient-doctor relationship. It is valuable to know how the ability of facial emotion recognition would impact the quality of patient-clinician interaction.

The admin has to enter a username and password for authentication purposes to log in to the application.

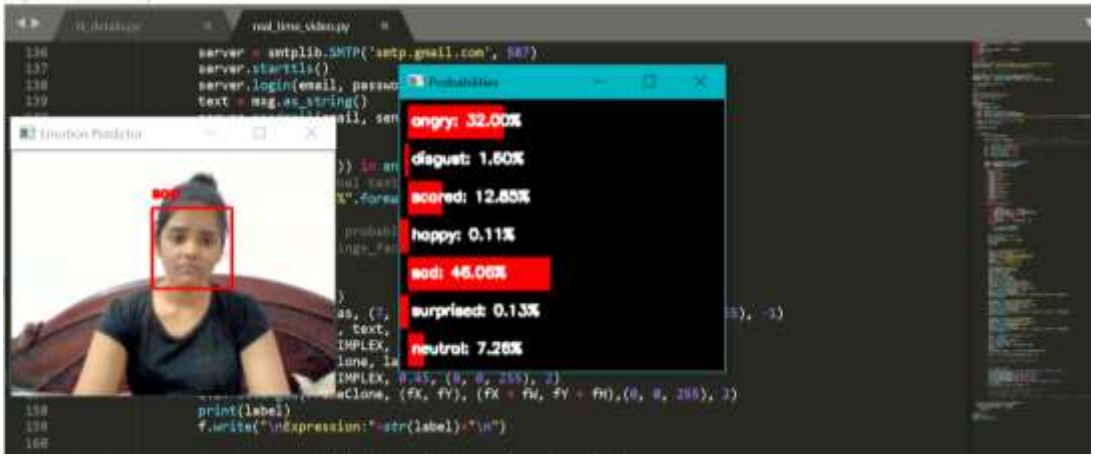
This is the registration form to fill in the details of the user and the details will be sent to the mail after detecting their emotions. This form helps the psychiatrists to recognize the identity of the patient when the data is sent to the email.



To detect the face, asking the user's permission to allow access to the webcam.



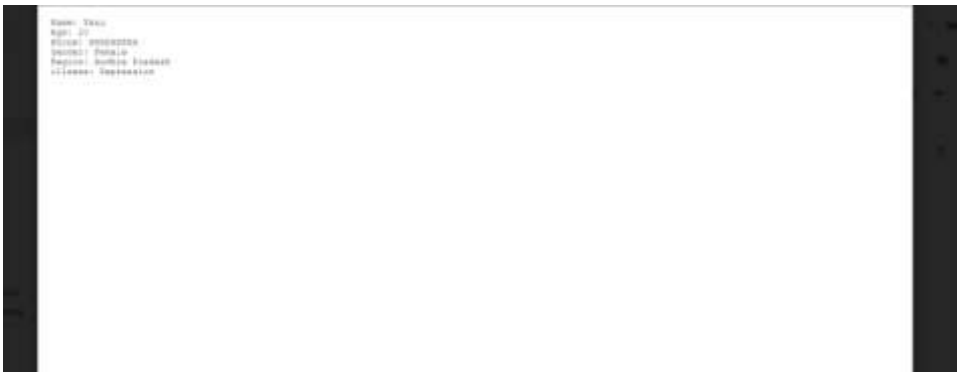
The program will create a window to display the scene captured by a web camera and a window representing the probabilities of detected emotions.



Statistics.txt: Emotions of the user at each frame.
The expressions at each frame are recorded and stored in a text file.

After reaching a threshold, detected emotions and the details of the user are dispatched through the mail. Three files will be dispatched at every detection.
Statistics.txt → file contains the emotions of the user at each frame ,Details.txt → file contains details of the user which is filled in the form. Emotions.txt → file contains the counter of each emotion.

details.txt: Details of the user



Emotions.txt: Counter of each emotion.

COMPARATIVE STUDY

FER-2013 dataset is the most commonly used dataset for Facial Emotion detection which is available in Kaggle. This dataset is trained on several models for facial emotion detection. Out of all the models that are trained over the FER-2013 dataset, the Mini-Xception model proves to attain higher accuracy of over 95%.

Comparison of Models

Model	Accuracy
Mini-Xception	95.05
Ensembler ResMaskingNet	76.82
Residual Masking Network	74.14
VGGNet	73.28
VGG	72.70
Res-Net	72.40
CNN Hyper Parameter Optimization	72.16
Inception	71.60
DeepEmotion	70.02

CONCLUSION

Facial emotion recognition is an emerging field now-a-days. Non-verbal communications like facial expressions are used in many applications in human-computer and interaction which is used to convey facial emotions. There is a lot of complexity and variability involved in recognizing facial emotions. This paper proposed a new method for facial recognition and mentioned its various applications in real-time scenarios. Predicting emotions gives us accurate information about how the user feels. So, this proposed method helps psychiatrists to analyze their patients effectively. We believe that this method has given promising results in detecting the emotions and dispatching the data via email to the person monitoring the user.

REFERENCES

1. Rani, A. M., & Durgadevi, R. (2017). Image Processing Techniques To Recognize Facial Emotions. *International Journal of Engineering and Advanced Technology (IJEAT)*, *6*(6).
2. Pranav, E., Kamal, S., Chandran, C. S., & Supriya, M. H. (2020, March). Facial emotion recognition using deep convolutional neural network. In 2020 6th International conference on advanced computing and communication Systems (ICACCS) (pp. 317-320). IEEE.
3. Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). *SN Applied Sciences*, *2*(3), 1-8.
4. Mellouk, W., & Handouzi, W. (2020). Facial emotion recognition using deep learning: review and insights. Procedia Computer Science, 175, 689-694.
5. Gowri, S. M., Rafeeq, A., & Devipriya, S. (2021, May). Detection of real-time Facial Emotions via Deep Convolution Neural Network. In *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1033-1037). IEEE.
6. Vulpe-Grigoraşi, A., & Grigore, O. (2021, March). Convolutional Neural Network Hyperparameters optimization for Facial Emotion Recognition. In *2021 12th International Symposium on Advanced Topics in Electrical Engineering (ATEE)* (pp. 1-5). IEEE.
7. Ozdemir, M. A., Elagoz, B., Alaybeyoglu, A., Sadighzadeh, R., & Akan, A. (2019, October). Real time emotion recognition from facial expressions using CNN architecture. In *2019 medical technologies congress (tiptekno)* (pp. 1-4). IEEE.
8. Wu, C. D., & Chen, L. H. (2019). Facial emotion recognition using deep learning. arXiv preprint arXiv:1910.11113.
9. Abdulsalam, W. H., Alhamdani, R. S., & Abdullah, M. N. (2019). Facial emotion recognition from videos using deep convolutional neural networks. *International Journal of Machine Learning and Computing*, *9*(1), 14-19.
10. Akhand, M. A. H., Roy, S., Siddique, N., Kamal, M. A. S., & Shimamura, T. (2021). Facial Emotion Recognition Using Transfer Learning in the Deep CNN. Electronics, 10(9), 1036.
11. Haque, M. I. U., & Valles, D. (2018, November). A facial expression recognition approach using DCNN for autistic children to identify emotions. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)* (pp. 546-551). IEEE.
12. Zadeh, M. M. T., Imani, M., & Majidi, B. (2019, February). Fast facial emotion recognition using convolutional neural networks and Gabor filters. In *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI)* (pp. 577-581). IEEE.
13. Mayya, V., Pai, R. M., & Pai, M. M. (2016). Automatic facial expression recognition using DCNN. Procedia Computer Science, 93, 453-461.
14. Jain, N., Kumar, S., Kumar, A., Shamsolmoali, P., & Zareapoor, M. (2018). Hybrid deep neural networks for face emotion recognition. Pattern Recognition Letters, 115, 101-106.
15. Yang, Y., & Sun, Y. (2017, December). Facial Expression Recognition Based on Arousal-Valence Emotion Model and Deep Learning Method. In *2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC)* (pp. 59-62). IEEE.
16. Ng, H. W., Nguyen, V. D., Vonikakis, V., & Winkler, S. (2015, November). Deep learning for emotion recognition on small datasets using transfer learning. In *Proceedings of the 2015 ACM on international conference on multimodal interaction* (pp. 443-449).
17. Pons, G., & Masip, D. (2017). Supervised committee of convolutional neural networks in automated facial expression analysis. *IEEE Transactions on Affective Computing*, *9*(3), 343-350.
18. Nadeeshani, M., Jayaweera, A., & Samarasinghe, P. (2020, December). Facial Emotion Prediction through Action Units and Deep Learning. In *2020 2nd International Conference on Advancements in Computing (ICAC)* (Vol. 1, pp. 293-298). IEEE.
19. Moolchandani, M., Dwivedi, S., Nigam, S., & Gupta, K. (2021, April). A survey on: Facial Emotion Recognition and Classification. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 1677-1686). IEEE.
20. Begaj, S., Topal, A. O., & Ali, M. (2020, December). Emotion Recognition Based on Facial Expressions Using Convolutional Neural Network (CNN). In *2020 International Conference on Computing, Networking, Telecommunications & Engineering Sciences Applications (CoNTESA)* (pp. 58-63). IEEE.