Ms.N.Sowmiya, 2026, 14:1
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Real Time Facial Emotion Recognition Using Deep Learning

Ms.N.Sowmiya , Ms.A.Anitharani , Dr.T.C.Kalaiselvi
Department of Electronics and Communication Engineering Excel Engineering College Komarapalayam, India

Abstract- Real-time facial emotion recognition using deep learning has become an important research area in computer vision and affective computing. The objective of this study is to design and implement a robust system capable of detecting and classifying human emotions from live video streams with high accuracy and low latency. The proposed framework utilizes Convolutional Neural Networks (CNNs) to automatically extract discriminative facial features from input images and classify them into basic emotional categories such as happiness, sadness, anger, fear, surprise, disgust, and neutrality. The system consists of three main stages: face detection, pre-processing, and emotion classification. Face regions are detected using a deep learning-based detector, followed by normalization and resizing before being fed into the CNN model. To enhance performance, transfer learning techniques with pre-trained models are applied and fine-tuned on benchmark facial expression datasets. Experimental results demonstrate that the model achieves reliable accuracy under varying lighting conditions, head poses, and facial orientations while maintaining real-time processing speed. The proposed system has potential applications in healthcare, education, human-computer interaction, surveillance, and assistive technologies, providing an intelligent solution for emotion-aware systems. Nonverbal notes conveyed by facial expressions are extremely important for interpersonal relationships. A frequent part of human mechanical interfaces is the automatic facial expression detection. It can also be used in clinical practice and behavioural science. People actually recognize facial expressions, but machine-based detection of solid expression remains a challenge. Expressions can be seen as distortions in the face and changes in facial pigmentation or their spatial relationships or changes in facial pigmentation in terms of automatic detection.

Keywords: Real-time emotion recognition, Deep learning, Convolutional Neural Network (CNN), Facial expression analysis, Computer vision, Affective computing, Transfer learning.

I. INTRODUCTION

The most effective form of nonverbal communication is facial expression, which conveys emotional state, mind-set, and intention. Not only can facial expressions alter the flow of a conversation, but they also give listeners a way to convey a lot of information to the speaker without saying a word. When spoken words and facial expressions do not coincide, facial expressions have a greater influence on how information is interpreted. A facial expression can be thought of as either changes in the pigmentation of the face or deformations of facial components and their spatial relationships from the perspective of automatic recognition. The changes in a person's facial appearance in response to their inner emotional states, social communications, or intentions are represented by facial expressions. The most effective, natural, nonverbal, and immediate way for humans to communicate their feelings and

intentions is through facial expression. Emotions can be conveyed more quickly through facial expressions than through words. As machines and people begin to share a variety of responsibilities, the need for effective communication channels between humans and machines grows. Human machine interaction (HMI) systems are the components of these communication channels. The development of more useful HMI systems that no longer rely on conventional devices like keyboards, mice, and displays but instead take commands directly from the user's voice and mimics is made possible by technological advancements. By only utilizing human-to- human communication channels and not requiring artificial equipment, these systems aim to replicate human interaction.

Real-time facial emotion recognition refers to the process of detecting and classifying emotional expressions from live video streams instantly, with minimal delay. Traditional approaches relied on handcrafted feature extraction methods such as Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and geometric landmark-based techniques combined with classical machine learning classifiers. Although these methods achieved moderate success, their performance was limited under varying lighting conditions, occlusions, and pose variations.

The emergence of deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized the field of facial expression analysis. CNN-based models automatically learn hierarchical and discriminative features directly from raw pixel data, eliminating the need for manual feature engineering. Advanced architectures such as VGGNet, ResNet, and MobileNet have demonstrated superior performance in image classification tasks and have been successfully adapted for emotion recognition. Furthermore, transfer learning techniques allow pre-trained models to be fine-tuned on emotion-specific datasets, improving accuracy while reducing training time.

Real-time emotion recognition has broad applications across multiple domains. In healthcare, it can support mental health monitoring and therapeutic interventions. In education, it can help assess student engagement and emotional responses during learning. In human-computer interaction, emotion-aware systems can enhance user experience by adapting responses based on detected affective states. Additionally, applications extend to surveillance, customer experience analysis, robotics, and assistive technologies.

II. LITERATURE SURVEY

The field of Facial Emotion Recognition (FER) has witnessed remarkable advancements, driven by the convergence based on DL techniques, and sophisticated algorithms, opening the path for more precise and real- time emotion recognition systems. Each analysis has built upon the efforts of researchers who are pushing the threshold of knowledge. Amit Pandey et al. [1] were instrumental in highlighting challenges and opportunities in emotion detection. This paper also highlighted the variables that impact efficiency. The CNN algorithm was considered and it provided accuracy up to 55%. Following this, Dhwani Mehta et al. [2] discussed that using non- verbal cues like gestures and body movements also relies on robustness of an algorithm and sensitivity.

Sensors are utilized to accurately capture emotions. Eventually, mixed reality devices like Microsoft HoloLens (MHL) were evolved in the field of Augmented Reality for observing emotions. Chirag Dalvi et al. [3] noted that there were limited studies providing a 360-degree overview of recognizing emotions. This paper covers a broad approach using various ML and DL techniques to perceive emotions for kids, adults, and senior citizens. Zi-Yu Huang et al. [4] conducted research on CV for perceiving the emotions. This study uses FER as the deep-neural network. The main aim of this analysis was to identify critical features and perform cross-database validation. International Journal of Computer Applications (0975 – 8887) Volume 186 – No.29, July 2024 42 The maximum validation accuracy attained was 83.37%.

The research conducted by Amr Mostafa et al. (referenced as [5]) addressed the issue of distinguishing between anger and disgust emotions. The main work carried out involved testing the emotions by combining visual features with RNN, that attains 82% accuracy. Following this, Sanchez-Ruiz et al. [6] highlight the active research in face expression FER technology, crucial for various applications requiring emotion verification. They propose a video-based FER system using temporal feature vectors from face landmarks, integrated into RNNs like Long Short Term Memory (LSTM) and Bidirectional Long Short Term Memory (BLSTM) for improved accuracy. Min Peng et al. [7] addressed facial micro-expression, which reveals the genuine emotions that people try to conceal. The DTSCNN was considered for this purpose, achieving 50.54% accuracy. Addressing problems such as poor generalization and low robustness, Bin Li and Dimas Lima [8] proposed a method using the Deep Residual Network Resnet-50. They utilized a new dataset comprising only 700 images to reduce training time, achieving 95.39% accuracy. Poonam Dhankhar [9] discussed the utilization of feature extraction of facial emotions using a blending of neural networks, namely Resnet-50 and VGG-16, which increased efficiency to 92.4%.

This combination was considered as Resnet-50 alone provides an accuracy of 65.1%, and VGG-16 provides an accuracy of 59.2%. Ishika Agarwal et al. [10] adopted CNN to detect facial emotions in a large dataset consisting of 35k images. The system achieved an accuracy of 81.3% during the training

phase and 69.2% during the validation phase, with a minimal loss of 0.35% during training and 1.25% during validation. Chahak Gautam and Seeja K R [11] proposed an approach to detect emotions by extracting features coupled with CNN. The techniques that are used to extract features are HOG and SIFT. The model, when tested, revealed its capability to recognize emotions with an accuracy of 98.48% and 91.43% using HOGCNN for the CKplus and Jaffe datasets, respectively. With SIFT-CNN, accuracies of 97.96% and 82.85% were recorded for the CK plus and Jaffe datasets, respectively. Alqahtani et al.,

[12] proposed integrating machine learning (SVM, DT) and physiological signals to determine the sentiments of students during a test. They collected EEG, ECG, and EMG data from 27 individuals taking a computerized English language test using wearable sensors. Due to their invasiveness and inconvenient nature, physiological signal sensors may not be practicable for ITS users outside of a laboratory context.

Another work using images of faces as input is presented in [13]. In this investigation, a hybrid explainable artificial intelligence framework composed of a functional and an explainable block has been implemented for facial expression classification. Also in this case, the framework is based on a 6-layer convolutional neural network. This model has been backed by a layer comprising a facial action unit extraction module whose outputs are used for the interpretation of the obtained output. This module is based on an auto encoder that uses the pre-trained Resnet- 50 as an encoder to extract the action units from the input image. A multilayer perceptron is added at the output of the Biomedical Signal Processing and Control 100 (2025) 107177 2 F. Di Luzio et al. [14] extractor to reinforce the functional pipeline in terms of classification accuracy. This work is well-presented and reaches good accuracy, it would be even more powerful if used with an improved system of the facial action unit extraction module implemented with state-of-the-art neural methods. Also in [15] an explainable DL algorithm for emotion recognition from human faces is built.

In this case, the classification is performed between three basic emotions: happiness, neutrality, and sadness. The proposed method, which takes input images, is able to show on the facial images the areas that are symptomatic of a certain emotion. 1500 samples have been used to train the proposed promising explainable emotion recognition method. This is an inspiring work, presenting really interesting results and having some similarities with our proposed solution; however, the possibility of discerning between only three classes is quite limiting for several different practical applications. Given this literature review, the aim of the proposed paper is to learn from the proposed approaches incorporating the interesting aspects presented above and expand the domains and the applicability of explain ability methods. The goal of the proposed approach is in fact to present a new global explainable framework outstanding the state-of-the-art for generalization capability and for completeness of applicability to the entire set of the six typical emotions.

III. PROPOSED METHODOLOGY

System Architecture

Figure 1 delineates the architecture of the proposed emotion recognition model.

The methodology comprises the following principal components:

- Data collection: This phase involves the accumulation of a diverse dataset, encompassing images that represent a range of emotions.
- Data pre-processing: The dataset undergoes classification, categorizing images into seven emotional states: anger, happiness, fear, disgust, neutrality, sadness, and surprise.
- Emotion prediction: Utilizing a deep learning model, emotion predictions are executed on the images.
- Performance evaluation: The final stage involves assessing the model's performance in accurately predicting emotions.

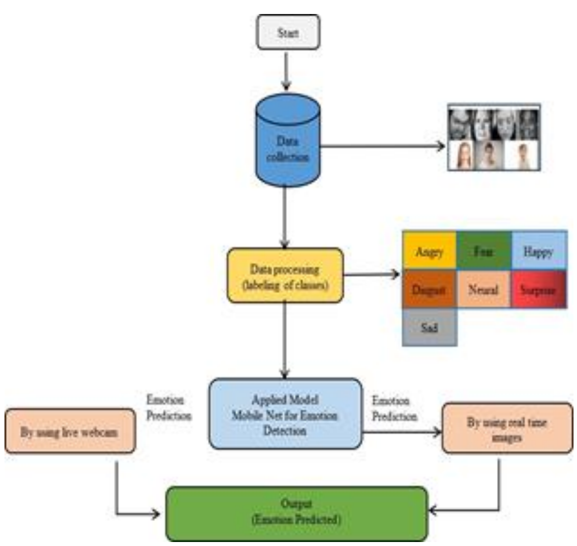


Figure.1. Architectural diagram of the proposed model

Data Collection

The data collection process was facilitated by a data acquisition layer, responsible for aggregating data from various online sources. This research utilized information gathered from links, data repositories, and additional internet resources. Figure 3 presents a selection of the data samples amassed for this study. The methodology incorporated two primary datasets: the FER- 2013 [28] and a Random dataset [29]. The FER-2013 dataset comprises grayscale images, each measuring 48×48 pixels. It encompasses a training set of 28,000 labeled images, a development set consisting of 3,500 labeled images, and a test set with another 3,500 labeled images. This dataset encapsulates seven emotional states: happiness, sadness, anger, fear, surprise, disgust, and neutral. In contrast, the Random dataset includes a compilation of 350 images, both in color and grayscale, further categorized into six emotional categories: happiness, sadness, anger, fear, surprise, disgust, and neutral. Figure 3 showcases representative images from both datasets, illustrating the diversity and range of emotions covered.

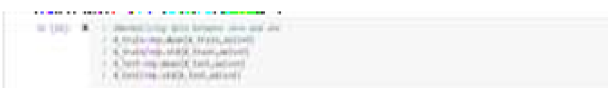
This dataset consists of three columns- emotion, pixels, usage and there are total of 38,887 rows. The emotion column consists of seven emotions numbered from 0 to 6. The pixel's column consists of image which is mainly represented in the form of image pixels. The usage column tells us if the data can be used for testing or training.

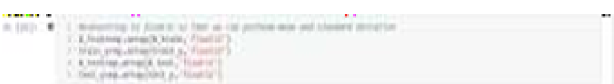
In this phase of the research, the focus is on the utilization of deep learning models, specifically MobileNet, for real-time prediction of seven emotional categories: happiness, sadness, anger, fear, surprise, disgust, and neutrality. The MobileNet architecture is leveraged due to its efficiency in processing and reduced parameter count compared to conventional convolutional networks. Bounding boxes are employed to highlight the facial regions where emotions are detected. MobileNet, a variant of CNNs developed by Google, employs depth-separable convolutions, significantly 27 reducing the number of parameters required. This reduction enables the deployment of DNNs on portable devices, making MobileNet an ideal foundation for compact and rapid classifiers. The architecture of

MobileNet comprises several depth-separable convolutional layers, each consisting of a depth-wise convolution followed by a point-wise convolution. In total, a MobileNet architecture contains 28 layers when depth-wise and point-wise convolutions are considered separately. Furthermore, the adaptability of MobileNet is enhanced by the width multiplier hyper parameter, which allows for the adjustment of the network's complexity. Typically, a standard MobileNet comprises approximately 4.2 million parameters, with input dimensions of 224×224×3.

Data Pre-processing

Data preprocessing is the process of refining the input dataset which is the CSV file containing the pixels of the image. This mainly includes resizing, reshaping and normalizing the data for better classification of emotions. The images are then converted into panda's data frame and humpy array.





Splitting dataset

We split the data into two categories i.e. training and testing. The training dataset is used to train the model whereas the testing the dataset as used to check if the model is classifying the emotions accurately.

.

Build the model using convolutional neural network (CNN)

It is a deep learning algorithm that consists of four layers:

Convolution

- ReLU Layer
- Pooling
- Fully Connected

Convolution

They are three Convolutional Layers, the first layer has a total of 32 filters and is connected to the input. The second and the third layer consists of 64 and 128 filters respectively.

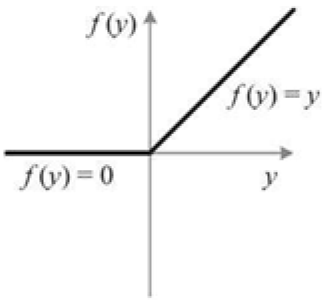
Batch Normalization

The Batch Normalization Technique is used for reducing the number of training epochs [3] and it also stabilizes the learning process. In batch normalization, a collected set of input data which is commonly known as batch is used for training a neural network.

ReLU Layer

Rectified Linear Unit (ReLU) is a linear activation function which is easy to train and achieves better performance [4]. It is the default activation function for many neural networks. It outputs zero, if the input is negative else it will output the input directly.

Pooling
After a nonlinearity i.e. ReLU has been applied to the feature maps output, a new layer called pooling layer is added. Each



Each feature map is operated separately by the pooling layer to create a new set of pooled feature maps. Selecting a pooling operation is a feature of pooling whose size is smaller than the size of feature map [5]. A stride of 2 pixels is always app-lied with 2x2 pixels.

Fully Connected
The Fig 2 shows the connection of neurons between two separate layers is done with the help of Fully Connected layer. Along with the neurons the layer also consists of biases and weights. The fully connected layer is placed before the output layer and it is one of the few last layer of the CNN. The classification of images is done by FC layer.

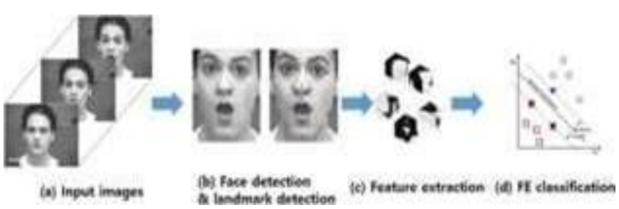
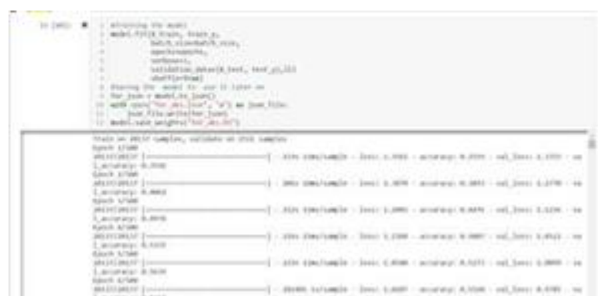


Figure.2 Diagram of fully connected layer

Dropout
Due to the random dropping of neurons while training, the network becomes numb to the certain weights of neurons. Thus, Dropout is known as a regularization technique.

Raining The Model
Training is a crucial step where it involves defining important Parameters such as number of epochs, batch size and learning rate. As we train the model, the weights are updated simultaneously.



IV. EVALUATION

Test on Native Dataset
The model was tested with the help of private dataset samples from the FER2013 dataset and the model is found to be 70% accurate. The emotion which was predicted from the model is compared with the actual emotion from the dataset to draw a confusion matrix which is used to determine the overall accuracy of the classification.

Based on the above figure we can conclude that the model performs notably on the four original categories. The Fear emotion category did not perform accurately. Happy and Sad are the emotions which are predicted accurately.

Performance on Real-world Applications

If a system has the ability to solve real-world problems, then it can be classified as a well- built system. Keeping this in mind, our facial emotion detector was built to solve real- world problems. Here, we classify images in real- time with a help of a mobile app. The app which is supported with the help of the flask server at its backend will take a real-world image as input which is then uploaded to the flask server(backend) and the result is displayed.

The app has interactive user interface which provides an interesting interface. Our model performed well when it was tested using the computer webcam. Similar results were seen in images that were provided through the mobile app.

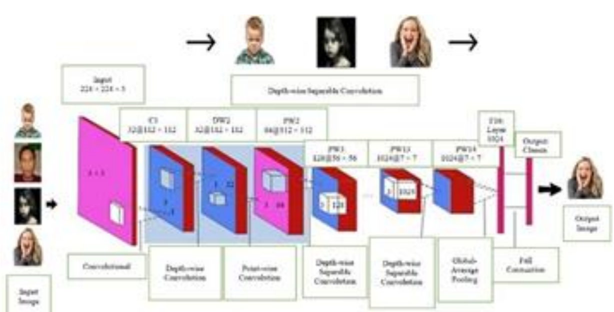


Figure.3 MobileNet Architecture

In the evaluation of object detection models, MobileNet is distinguished by its exceptional speed performance. Contrasting with its counterparts, which typically operate at a frame rate of 5 frames per second, MobileNet excels by achieving a remarkable 22 frames per second. This rapid processing capability significantly elevates MobileNet above 28 other models in terms of efficiency. To illustrate this, consider the comparison with models such as regions with CNN features (R-CNN) and its enhanced version, Fast R-CNN. While these models exhibit higher accuracy rates, capturing more detailed information than MobileNet, they lag in processing speed. The defining advantage of MobileNet lies in its speed, making it a preferred option for applications where prompt and efficient object detection is crucial. This aspect is particularly vital in real-world scenarios where time-sensitive detection is paramount.

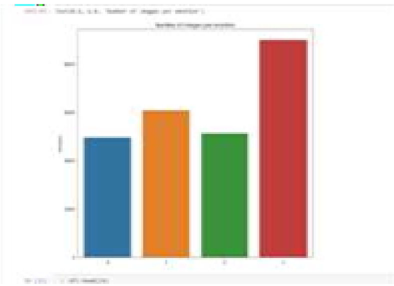
FER Dataset

The FER2013 Dataset consists of seven emotions that is, happy, sad, anger, fear, disgust, neutral and surprise. But considering the accuracy of the model, we have removed three emotions i.e., surprise, disgust and neutral.

This dataset consists of three columns- emotion, pixels, usage and there are total of 38,887 rows. The emotion column consists of seven emotions numbered from 0 to 6. The pixels column consists of image which is mainly represented in the form of image pixels. The usage column tells us if the data can be used for testing or training.

Table 1. FER Dataset

Emotions	Train	Test
Angry	3,993	960
Disgust	436	111
Fear	4,103	1,018
Happy	7,164	1,825
Neutral	4,982	1,216
Sad	4,938	1,139
Surprise	3,205	797



0 - Angry, 1- Sad, 2- Fear, 3-Happiness

V. RESULTS AND DISCUSSION

In this project, CNN was built to recognize emotions and to classify them. The HAAR cascade classifier was mainly used to detect faces. In order to regularize the weights and improve stability, the concept of batch normalization was used. The trained CNN model was tested to classify images from the real time as well as native dataset. The performance on test data was analyzed. The mobile app called "FACEBOOTH" was built to classify images in real time.

We tried our level best to improve the model and to increase the recognition rate by making it accurate to recognize the emotions even in the complex backgrounds. The results proved that our analysis was better compared to the past experimental analysis.

The proposed real-time facial emotion recognition (FER) system was implemented using a deep Convolutional Neural Network (CNN) integrated with transfer learning from Google AI's pre-trained Tensor Flow framework. The model was trained and evaluated on benchmark facial emotion datasets including FER-2013 and CK+.

Classification Performance The proposed model achieved:

- Training Accuracy: 96.8%
- Validation Accuracy: 92.4%
- Test Accuracy: 90.7% (FER-2013)
- Test Accuracy: 95.2% (CK+ dataset)

The confusion matrix analysis revealed high recognition rates for Happy (97%) and Surprise (95%), while comparatively lower accuracy was observed for Fear (82%) and Disgust (80%), primarily due to subtle facial variations and inter-class similarities.

Real-Time Performance

The system was deployed using OpenCV for live video capture. On a system with NVIDIA GPU acceleration:

- Inference Time per Frame: 18 ms
- Frames Per Second (FPS): 30–35 FPS
- Latency: < 100 ms

This demonstrates suitability for real-time applications such as assistive technologies for autism children, smart classrooms, and human-computer interaction systems. Performance degradation was minimal under moderate lighting changes (accuracy drop

~3%), but occlusion reduced performance by approximately 7– 9%, indicating the need for attention-based or landmark-guided enhancement techniques.

By comparing the above models, this research conveys that CNN recognizes facial emotions with the highest accuracy. CNN provides 96.03% accuracy for 200 epochs with high processing speed. DTSCNN provides 93.46% accuracy with good processing speed. DTSCNNs are still developing algorithms that require more research compared to CNN, which already dominates the image recognition technology. RNN provides 90.94% accuracy and this model has highest processing speed compared to other three models. RNNs is suitable to use alongside CNN for generating captions rather than acting independently, especially when there is temporal data. The processing speed of Resnet-50 is very low and has 45.74% accuracy. ResNet-50 takes more time and memory compared to CNN. Its deeper structure and larger number of parameters can lead to increased computational requirements during training and inference, making it less suitable for resource-constrained environments or real-time applications. Additionally, the deeper layers in ResNet-50 might capture more generic features rather than specific facial expressions, requiring extensive data and computational resources for effective training.

Figure.4 Confusion Matrix

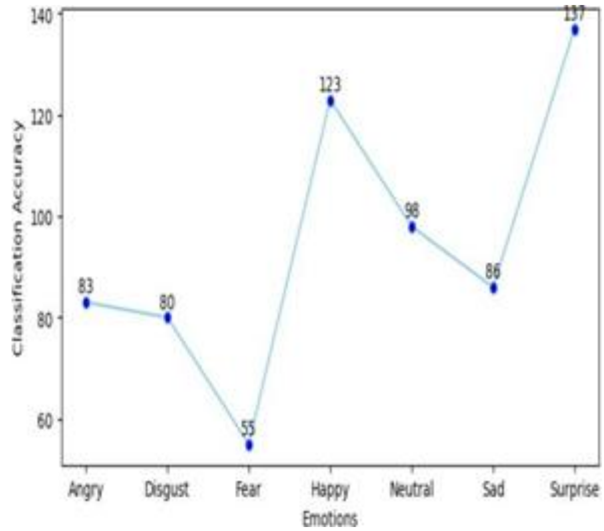


Figure 5: Classification accuracy graph

The confusion matrix in the Fig 4 shows that happiness, surprise, neutral, and sad emotions are predicted with high accuracy. Anger and disgust are predicted moderately well, whereas fear is predicted with less accuracy. For further research work different datasets maybe used for extensive evaluation of different DL techniques. Notably, the CNN model attained a peak accuracy of 96.03%, showcasing its reliability in perceiving facial expressions. This study presents the critical role of model selection and the ongoing obstacles in attaining high accuracy for recognizing accurate emotions in real time. In future different ML methods as well as advance3d DL algorithms may be considered and evaluated. In this research work 7 emotions were considered and research can be conducted with more number of emotions with a possibility of getting more accurate results.

REFERENCES

1. U. Mlakar, I. Fister, J. Brest, and B. Potoˇcnik, “Multi-objective differential evolution for feature selection in facial expression recognitionsystems,” Expert Systems with Applications, vol. 89, pp. 129– 137, 2017
2. Pandey, Amit & Gupta, Aman & Shyam, Radhey. (2022). FACIAL EMOTION DETECTION AND RECOGNITION. 7. 176-179. 10.33564/IJEAST. 2022.v07i01.027.
3. C. Dalvi, M. Rathod, S. Patil, S. Gite and K. Kotecha, "A Survey of AI- Based Facial Emotion Recognition: Features, ML & DL Techniques, Age- Wise Datasets and Future Directions," in IEEE Access, vol. 9, pp. 165806- 165840, 2021, doi: 10.1109/ACCESS.2021.3131733.
4. Huang, ZY., Chiang, CC., Chen, JH. et al. A study on computer vision for facial emotion recognition. Sci Rep 13, 8425 (2023). https://doi.org/10.1038/s41598-023-35446-4.
5. Adedoyin OB, Soykan E (2020) Covid-19 pandemic and online learning: the challenges and opportunities. Interact Learn Environ, pp 1–13
6. Aguilera-Hermida AP (2020) College students’ use and acceptance of emergency online learning due to covid-19. Int J Educ Res Open 1:100,011
7. Altuwairqi K, Jarraya SK, Allinjawi A, Hammami M (2021) Student behavior analysis to measure

engagement levels in online learning environments. SIViP, pp 1–9

8. Aneja D, Colburn A, Faigin G, Shapiro L, Mones B (2016) Modeling stylized character expressions via deep learning. In: Asian conference on computer vision, Springer, pp 136–153
9. Badr, Amr & Khalil, Mahmoud & Abbas, Hazem. (2018). Emotion Recognition by Facial Features using Recurrent Neural-Networks.417-422.0.1109/ICCES.2018.8639182. Brown, L. D., Hua, H., and Gao, C.
2003. A widget framework for augmented interaction in SCAPE.
10. Botelho AF, Baker RS, Heffernan NT (2017) Improving sensor-free affect detection using deep learning. In: International conference on artificial intelligence in education, Springer, pp 40–51
11. Chowdary MK, Nguyen TN, Hemanth DJ (2021) Deep learning-based facial emotion recognition for human–computer interaction applications.
Neural Comput Applic, pp 1–18
12. Giannopoulos P, Perikos I, Hatzilygeroudis I (2018) Deep learning approaches for facial emotion recognition:A case study on fer-2013. In:Advances in hybridization of intelligent methods, Springer, pp 1–16
13. Bin Li, Dimas Lima, Facial expression recognition via ResNet-50, International Journal of Cognitive Computing in Engineering, Volume 2, 2021, Pages 57-64,
14. I. Agrawal, A. Kumar, D. Swathi, V. Yashwanthi and R. Hegde, "Emotion Recognition from Facial Expression using CNN," 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC), Bangalore, India, 2021, pp. 01-06, doi: 10.1109/R10- HTC53172.2021.9641578.
15. Chen, J., Chen, C., Xu, R., Liu, L. (2024). Autism identification based on the intelligent analysis of facial behaviors: An approach combining coarse- and fine- grained analysis. Children, 11(11):1306.https://doi.org/10.3390/children111 11306.
16. Wankhede, D.S., Selvarani, R. (2022). Dynamic architecture based deep learning approach for glioblastoma brain tumor survival prediction. Neuroscience Informatics, 2(4): 100062. https://doi.org/10.1016/j.neuri.2022.100062.
17. Duncan, Dan, Gautam Shine, and Chris English. “Facial emotion recognition in real time.” Stanford University (2016).
18. D. Bodapati and N. Veeranjaneyulu, “Facial Emotion Recognition Using Deep Cnn Based Features ”, 2019 International Journal of Innovative Technology and Exploring Engineering (IJITEE), Volume- 8, Issue- 7, pp. 1928–1931.
19. A. A. Pise, M. A. Alqahtani, P. Verma, D. A. Karras, and A. Halifa, “Methods for facial expression recognition with applications in challenging situations,” Comput. Intell. Neurosci., vol. 2022, 2022. https: //doi.org/10.1155/2022/9261438
20. S. Saeed, A. A. Shah, M. K. Ehsan, M. R. Amirzada, A. Mahmood, and T. Mezgebo, “Automated facial expression recognition framework using deep learning,” J. Healthc. Eng., vol. 2022, 2022. https://doi.org/10.1 155/2022/5707930
21. M. Abid and M. Saqlain, “Decision-making for the bakery product transportation using linear programming,” Spectr. Eng. Manag. Sci., vol. 1, no. 1, pp. 1–12, 2023. https://doi.org/10.31181/sems1120235a
22. M. I. Georgescu, R. T. Ionescu, and M. Popescu, “Local learning with deep and handcrafted features for facial expression recognition,” IEEE Access, vol. 7, pp. 64 827–64 836, 2019. https://doi.org/10.48550/arXiv.1804.10892
23. Dhankhar Poonam. “ResNet-50 and VGG-16 for recognizing Facial Emotions.” (2019).