MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
A Constituent Institution of Manipal University

Department of Computer Science and Engineering

Artificial Intelligence and Machine Learning

A mini project synopsis on

Title: Face Emotion Recognition for better education

By

Drumil Haresh Ved 210962006
Punith B Nayak 210962012

I. Abstract

We propose a model for online education that utilizes live input from students to provide real-time feedback to teachers, enabling dynamic adaptation of teaching methods based on student comprehension levels. Data preprocessing involves loading and normalizing images into numpy arrays and tensors, followed by encoding into category vectors. The neural network architecture comprises convolution layers for feature extraction and dense linear layers for classification, with a higher dropout rate to prevent overfitting. We aim to make a sophisticated model with more than 4 million parameters.

II. Literary Review

A novel method combines Principal Component Analysis (PCA) for dimensionality reduction in facial emotion recognition with deep learning for optimal facial feature extraction. Because significant variation is preserved, using PCA's compression function on deep learning features increases accuracy and computing efficiency. Empirical findings exhibit the method's potential to improve facial emotion detection systems, as it outperforms direct feature extraction with the VGG-Face model [6].
This paper highlights face detection and recognition, addressing the need for automated comprehension and assessment of image and video datasets, particularly in face identification, appearance recognition, and human-computer interaction applications. Emphasizing the significance of facial recognition in biometrics [1]. The research provides a noise reduction technique for iris recognition systems to target localized high-frequency information, such as eyelashes and eyelids in segmented iris areas utilizing radial suppression. This method performs better in obtaining lower equal error rates, improves segmentation accuracy, and includes an automated iris identification system prototype. It also uses a one-dimensional Log Gabor wavelet for iris feature extraction[2].

The paper compares five algorithms for real-time emotion recognition from facial images, including deep learning (AlexNet, Affdex CNN, FER-CNN) and conventional methods (SVM, MLP) utilizing Histogram of Oriented Gradients features. It evaluates these approaches on four basic emotions. It demonstrates the superiority of the proposed FER-CNN model, highlighting its potential for diverse applications such as medicine, e-learning, entertainment, and marketing[10]. This project aims to develop a FacE Emotion Recognition (FER) System for Mental Stress Detection in response to the growing concern about mental stress among college students. Using CV and ML techniques, the system analyzes facial expressions to recognize emotions, including happiness, sorrow, anger, and fear. It primarily focuses on stress detection. The study uses a multidisciplinary methodology that combines questionnaire methodology, historical research, and quantitative methodologies. Using the FER2013 dataset, machine learning algorithms are trained on facial expression recognition using a quantitative approach. The survey highlights the potential of facial emotion recognition systems for mental stress detection in terms of accurately assessing emotional states, assisting people in finding the proper care or support, and providing insightful information about the intricate connection between stress, subjective experiences, and facial expressions[9]. This article primarily targets mental stress detection in face recognition utilizing a deep transfer network (DTN) with 3D morphable models (3DMMs). By synthesizing faces with different poses using 3DMMs, the study tackles the problems of labeled face picture shortages and dataset

bias between synthesis and natural photos. In the suggested approach, real-face and synthetic face photos are used to train a deep neural network (DTN) designed to reduce distribution disparities between the two domains. The Labeled Faces in the Wild dataset experiment findings demonstrate how effective the DTN is at mitigating dataset bias[7].Using a pre-trained GoogLeNet architecture and Transfer Learning, the study presents a Facial Emotion Recognition (FER) system that achieves 63.39% accuracy on the FER-2013 dataset. It draws attention to the widespread use of FER across various fields and underscores the effectiveness of deep learning-based methods, especially regarding facial picture emotion classification[11].
To efficiently manage imbalanced FER datasets, this research presents a unique method for Facial Expression Recognition (FER) in educational contexts called HoE-CNN. This method makes use of ensemble convolutional neural networks. With an emphasis on online learning applications, it overcomes FER problems, especially in multi-class labeling and imbalanced data, and shows enhanced performance compared to single deep learning methods[12]. This work presents a Deep CNN-based system for Facial Expression Recognition (FER). It suggests a Transfer Learning strategy with facial expression data fine-tuning, exceeding earlier Deep TL methods to achieve FER accuracies of 98% and 83% on the CK+ and FER2013 datasets, respectively[13]. In this study, a Dual-Direction Attention Network (DDAN) head is combined with a Mixed Feature Network (MFN) backbone to present the Dual-Direction Attention Mixed Feature Network (DDAMFN) for facial expression recognition (FER). DDAMFN is a state-of-the-art model in the field, achieving improved performance on FER datasets such as AffectNet, RAF-DB, and FERPlus by utilizing mixed-size kernels and a new attention mechanism. Using a novel attention mechanism and mixed-size kernels within its Mixed Feature Network (MFN) backbone, the Dual-Direction Attention Mixed Feature Network (DDAMFN) offers a lightweight and robust solution for facial expression recognition (FER), outperforming existing models on a variety of FER datasets, including AffectNet, RAF-DB, and FERPlus[14].

III. Dataset:

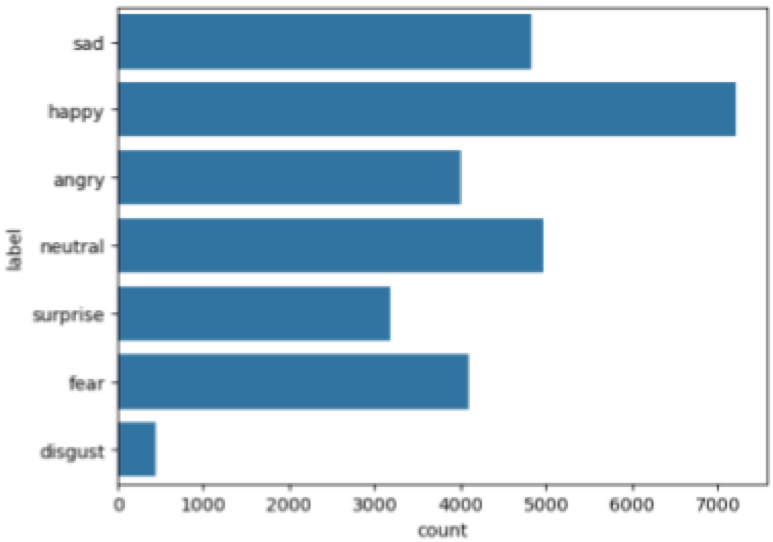
The dataset is called Facial Expression Dataset (FED-2013). The dataset contains 7 classes of emotions, Fear, Neutral, Angry, Happy, Surprise, Disgust, Sad.

A sample subset showing all the classes.

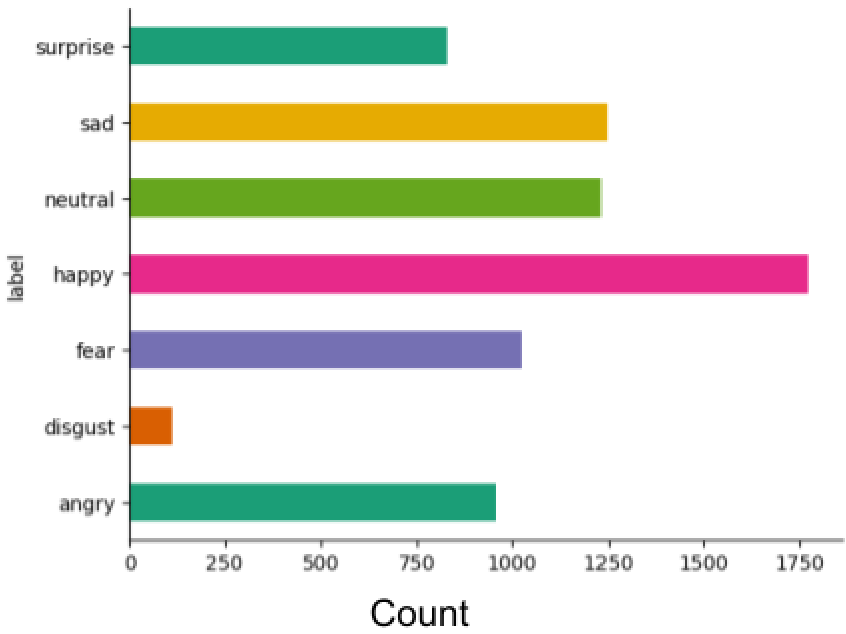
Each image is 48 x 48 pixels and greyscale.

A 48 x 48 grayscale image from dataset.

The dataset had about 35900 images of all the classes combined.



This graph shows the distribution of various classes among train directory.



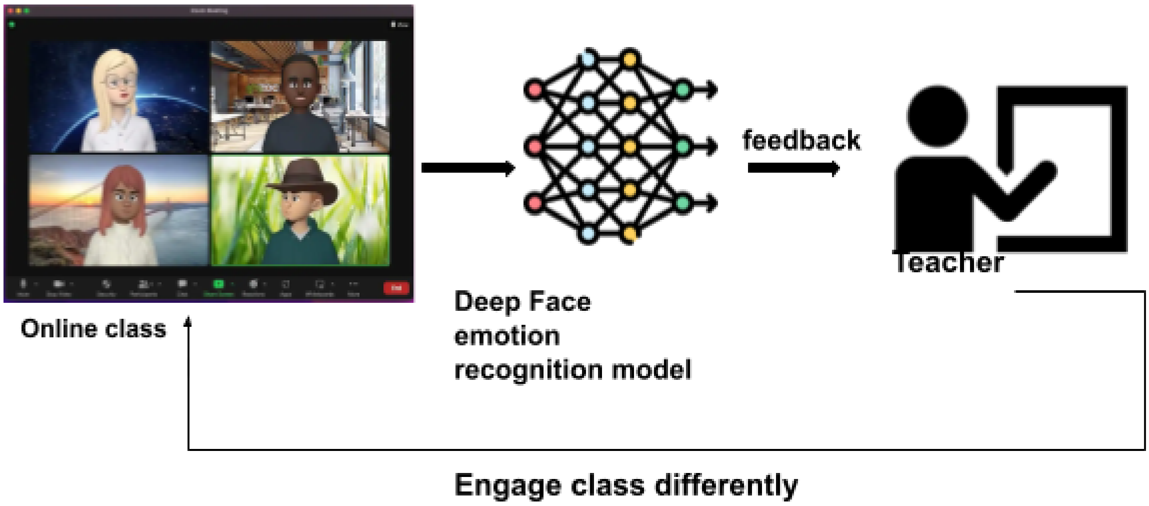
Test dataset distribution.

The dataset was divided into two parts: train and test. The train : test split was 80 : 20 %(28709 : 7178).

IV. Methodology

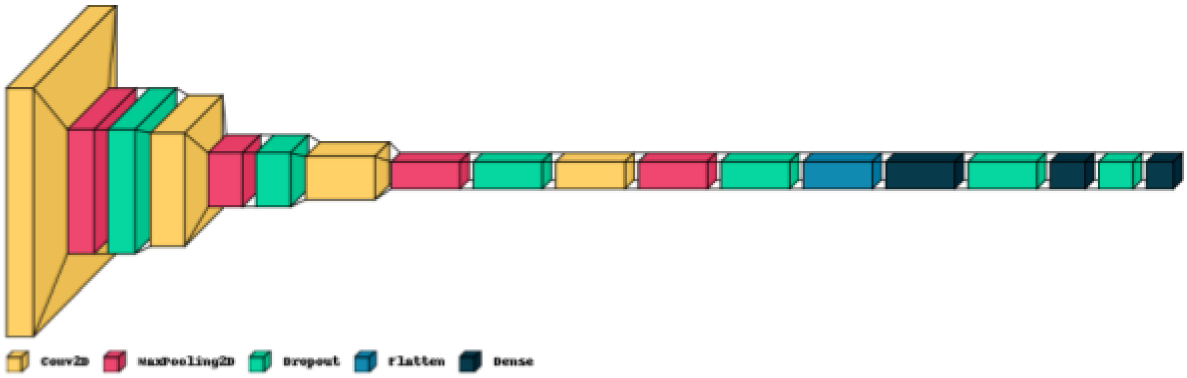
Proposed Model:

We propose a model which can take live inputs from the class and then give feedback to the teacher such that the teacher can dynamically engage the class based on the feedback from deep neural networks. This is more suitable for online classes where understanding the student requirements is vastly different from offline classes. If a tough concept causes the students to be confused then the teacher can immediately get feedback and immediately change the methodology of teaching.

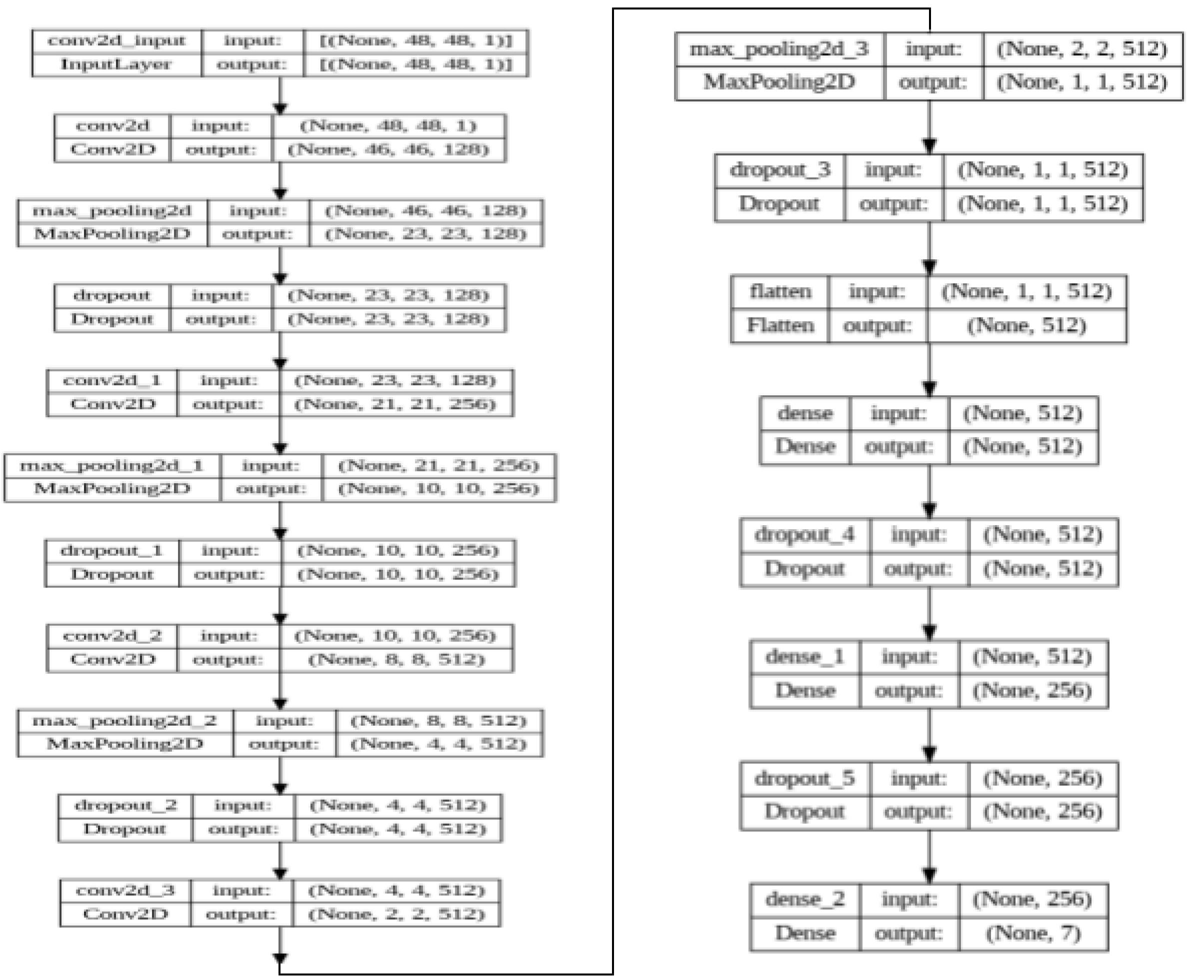


Data Preprocessing:
All the images were loaded to a numpy array then to tensor. Each image was normalized. We then encoded the images into category vectors.
Model Architecture:
The methodology to make the neural network is fairly intuitive. The first few Convolution layers of the neural network are for feature extraction and the next dense linear layers are dedicated to classification. In the implementation of this model, to curb overfitting and to optimize training, a dropout of 40 % was chosen as we have a fairly large neural network with over 4,232,199 trainable parameters. Convolution kernel size of 3 x 3 was chosen as it is a common practice to choose a smaller size and then try with larger kernel size of 5 x 5 or 11 x 11. To introduce non linearity the standard Rectified Linear Unit (ReLU) was opted for. The convolution layer and the max pooling layer kept decreasing the image from 48 x 48 to 1 x 1 while doubling the feature maps every convolution layer. Starting from 128 feature maps in the first convolution layer to 512 in the third convolution layer doubling every convolution layer. The later linear layers are for classification and the successive linear layer size keeps reducing in size till the size is equal to the number of classes. At the final output layer we use the softmax activation function to get a definitive answer.
The standard T4 gpu of Google Colab was used for this project.

Architecture:



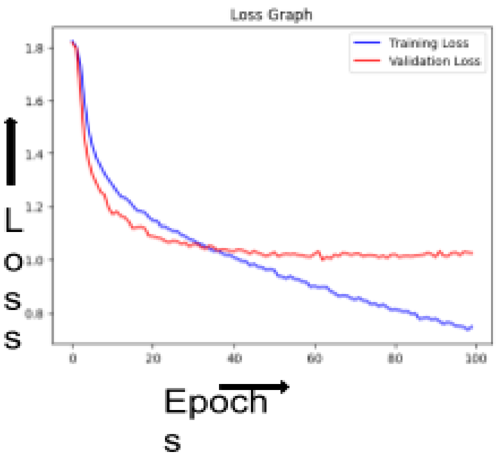
Visual representation of the model architecture.

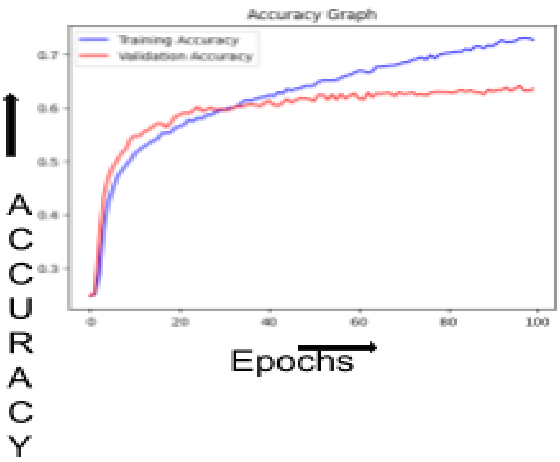


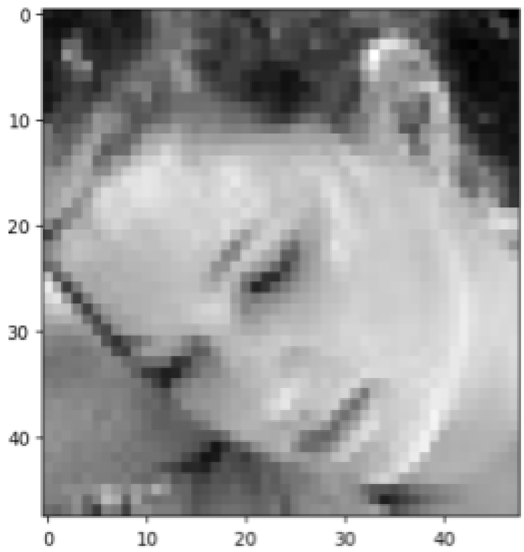
Model specifications

V. Results

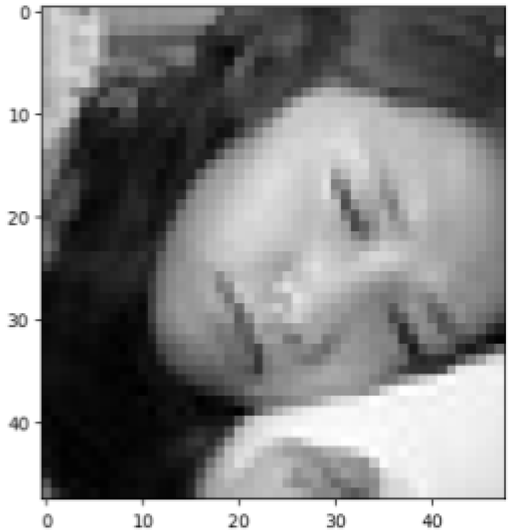
The model showed an accuracy of about 63.50 % while testing, the training was 100 epochs.







Original Output: sad
Predicted Output: neutral



Original Output: sad
Predicted Output: sad

Comparison:

Sr No	Dataset	Model	Accuracy.
1.	JAFFE	DenseNet161	96.51%
2.	FER2013	CNN	71.97%
3.	FER2103	Improved ResNet 18	83%
4.	FER3013	Out Model	64.49%

VI. Proposed Approach Limitations

While our model for emotion recognition shows promise, it has certain limitations to be addressed in future studies:

1. Our study utilized two datasets, CK+ and FER2013, for emotion detection, but further exploration of additional datasets is necessary for robustness. Expanding dataset size, particularly with larger subsets of CK+ and FER2013, could improve emotion recognition outcomes.
2. Future research could develop task-specific CNN models tailored for emotion recognition systems.
3. Incorporating alternative evaluation methodologies like k-fold cross-validation and statistical analysis could enhance the validation of our approach.

VII. Future Directions

Our model for emotion recognition offers numerous avenues for future exploration. Future studies could be conducted on ablation datasets such as EMOTIC and EMO-DB to improve generalizability. Furthermore, ongoing optimization efforts could refine our model architecture, potentially incorporating new blocks for improved efficiency and accuracy. Additionally, future research could expand emotion recognition to include speech and body gestures, exploring innovative approaches for practical applications.

VIII. Conclusion

In conclusion, while showing promising results, our model for emotion recognition also presents avenues for further enhancement and exploration. The limitations identified, such as the need for additional datasets, underscore the importance of ongoing research to ensure the robustness and generalizability of our approach. By addressing these limitations and incorporating them in future directions, such as optimization efforts and exploration of real-world applications, our model holds the potential to contribute significantly to the field of emotion recognition. Through continued collaboration and innovation, we can advance the capabilities of our model and pave the way for its integration into practical settings, ultimately benefiting society at large.

IX. References

1. A. Kumari Sirivarshitha, K. Sravani, K. S. Priya and V. Bhavani, "An approach for Face Detection and Face Recognition using OpenCV and Face Recognition Libraries in Python," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 1274-1278, doi: 10.1109/ICACCS57279.2023.10113066. -
2. P. Podder, T. Z. Khan, M. H. Khan, M. M. Rahman, R. Ahmed and M. S. Rahman, "An efficient iris segmentation model based on eyelids and eyelashes detection in iris recognition system," 2015 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2015, pp. 1-7, doi: 10.1109/ICCCI.2015.7218078.
3. C. -W. Tan and A. Kumar, "Integrating ocular and iris descriptors for fake iris image detection," 2nd International Workshop on Biometrics and Forensics, Valletta, Malta, 2014, pp. 1-4, doi: 10.1109/IWBF.2014.6914251.
4. P. C. Neto, J. R. Pinto, F. Boutros, N. Damer, A. F. Sequeira and J. S. Cardoso, "Beyond Masks: On the Generalization of Masked Face Recognition Models to Occluded Face Recognition," in IEEE Access, vol. 10, pp. 86222-86233, 2022, doi: 10.1109/ACCESS.2022.3199014.

5. S. Alfattama, P. Kanungo and S. K. Bisoy, "Face Recognition from Partial Face Data," 2021 International Conference in Advances in Power, Signal, and Information Technology (APSIT), Bhubaneswar, India, 2021, pp. 1-5, doi: 10.1109/APSIT52773.2021.9641286.
6. Y. Zhang, H. Wang, F. Xu and K. Jia, "A Deep Face Recognition Method Based on Model Fine-tuning and Principal Component Analysis," 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Honolulu, HI, USA, 2017, pp. 141-146, doi: 10.1109/CYBER.2017.8446331.
7. Z. An, W. Deng, T. Yuan and J. Hu, "Deep Transfer Network with 3D Morphable Models for Face Recognition," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 2018, pp. 416-422, doi: 10.1109/FG.2018.00067.
8. P. Dinkova, P. Georgieva, A. Manolova and M. Milanova, "Face recognition based on subject dependent Hidden Markov Models," 2016 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), Varna, Bulgaria, 2016, pp. 1-5, doi: 10.1109/BlackSeaCom.2016.7901570.
9. F. J. Ming, S. Shabana Anhum, S. Islam and K. H. Keoy, "Facial Emotion Recognition System for Mental Stress Detection among University Students," 2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), Tenerife, Canary Islands, Spain, 2023, pp. 1-6, doi: 10.1109/ICECCME57830.2023.10252617. -
10. A. Kartali, M. Roglić, M. Barjaktarović, M. Đurić-Jovičić and M. M. Janković, "Real-time Algorithms for Facial Emotion Recognition: A Comparison of Different Approaches," 2018 14th Symposium on Neural Networks and Applications (NEUREL), Belgrade, Serbia, 2018, pp. 1-4, doi: 10.1109/NEUREL.2018.8587011.
11. Imane Bouslihim(B) and Walid Cherif, “Facial Emotion Recognition Using a GoogLeNet Architecture”
12. Rit Lawpanom, Wararat Songpan * and Jakkrit Kaewyotha, “Advancing Facial Expression Recognition in Online Learning Education Using a Homogeneous Ensemble Convolutional Neural Network Approach”.
13. Rabie Helaly, Seifeddine Messaoud, Soulef Bouaafia, Mohamed Ali Hajjaji, Abdellatif Mtibaa, “DTL-I-ResNet18: facial emotion recognition based on deep transfer learning and improved ResNet18”
14. Saining Zhang, Yuhang Zhang, Ye Zhang , Yufei Wang and Zhigang Song 4, “A Dual-Direction Attention Mixed Feature Network for Facial Expression Recognition”