

**Department of Computer Science and Engineering**

**Artificial Intelligence and Machine Learning**

*A mini project synopsis* *on*

**Title: A Step Toward Better Facial Emotion Recognition Models**

*By*

*Drumil Haresh Ved 210962006*

*Punith B Nayak 210962012*

1. **Introduction**

This idea represents a considerable advancement in online education by providing teachers with real-time student input. This transform raw student interactions into insightful knowledge by employing advanced data pretreatment techniques like picture encoding and standardization. Educators can adapt their lesson plans dynamically based on students' responses and comprehension levels with this strategy.

This model, which has over 4 million parameters including convolution layer, effectively uses convolutional layers to extract features and dense linear layers for precise classification. In order to close the gap between instructional refinement and student engagement, this ensures that the model is responsive to fluctuations in student feedback by including a higher dropout rate to minimize overfitting. This integration takes online learning to unprecedented efficacy by allowing educators to personalize their students' learning experiences.

The rest of the paper is divided as follows. Section II is the abstract, and section III is related to work. Section IV elaborates on the objectives. Section V describes the dataset in detail; section VI is the methodology followed for this paper. Section VII describes the results. Section VIII discusses the conclusion and future work; section IX is for citations.

1. **Abstract**

Images are loaded, normalized, and then encoded into category vectors using numpy arrays and tensors as part of the data preprocessing step. Convolution layers are used for feature extraction in the neural network design, while dense linear layers—which have a greater dropout rate to minimize overfitting—are used for classification. The destination is to create a complex model with about 4 million parameters.

1. **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. One disadvantage of this method is that, especially in complex emotional circumstances, PCA may only be able to capture some of the subtleties and intricacies of facial expressions, which could result in information loss during dimensionality reduction

[6].

This study focuses on face detection and recognition to meet the requirement for automated understanding and evaluation of picture and video datasets, especially in face identification, appearance recognition, and human-computer interaction applications. Highlighting the role facial recognition plays in biometricsThis paper's exclusive focus on biometrics-related facial recognition may be one of its limitations. It may have ignored other crucial facets of face detection and recognition, such as privacy issues, moral considerations, and possible biases in algorithmic decision-making[1]. The study offers iris identification systems a noise reduction method to focus on localized high-frequency data, such as eyelashes and eyelids, in segmented iris sections via radial suppression. In addition to improving segmentation accuracy and obtaining reduced equal error rates, this method incorporates an automated iris identification system prototype. A one-dimensional Log Gabor wavelet is employed to extract iris features. This study's dependence on particular environmental factors or picture capture configurations for best results may be one of its limitations. The accuracy of iris recognition and the effectiveness of the noise reduction technique may be impacted by changes in illumination, occlusions, or image quality[2]. The study uses a Histogram of Oriented Gradient characteristics to look at five algorithms that identify emotions in real time from facial pictures. These algorithms cover conventional methods (SVM, MLP) and deep learning techniques. It evaluates these techniques based on four core feelings. It demonstrates the usefulness of the proposed FER-CNN paradigm and emphasizes how many applications it may be used for, such as e-learning, marketing, entertainment, and healthcare. One of the study's possible shortcomings is that it only evaluated the algorithms on a few basic emotions. Since emotion detection is a complex process, restricting the evaluation to only four basic emotions might not accurately capture the nuances and complexity of human emotional expression. Because of this, applying the findings to situations in the real world when a more comprehensive range of emotions is present may be more difficult [10].

In terms of precisely evaluating emotional states, helping people find the proper care or support, and offering enlightening information about the complex relationship between stress, subjective experiences, and facial expressions, the survey emphasizes the potential of facial emotion recognition systems for mental stress detection[9]. This paper deals with identifying mental stress on face recognition utilizing a deep transfer network (DTN) and 3D morphable models (3DMMs). By utilizing 3DMMs to synthesize faces in different expressions, the study solves the problems of a need for annotated face images and dataset bias between artificial and natural photographs. The suggested technique uses real-faces and synthetically generated face images to train neural networks to minimize distribution differences between the two domains. The trial results of the Labeled Faces in the Wild dataset demonstrate how effectively the DTN minimizes dataset bias. One potential limitation on the study could be the use of synthetic facial images generated by 3D morphable models (3DMMs) to augment the training data. Even though synthetic data can help alleviate the shortage of labeled face photos, it may not fully capture the complexity and diversity inherent in real-world facial expressions and emotions. A performance gap could occur when the model is used in real-world scenarios where individuals exhibit different stress indicators and facial expressions that synthetic data might not sufficiently reflect.[7].

On the FER-2013 dataset, this article offers a Facial Emotion Recognition (FER) system that achieved 63.39% accuracy on a test set using Transfer Learning technique which was based on GoogLensNet architecture. It shows how widely FER is used in many different sectors and depicts how successful deep learning-based techniques are, particularly when classifying emotions in facial pictures[11].

To manage skewed FER datasets, this research paper presents a unique method for Facial Expression Recognition (FER) in educational contexts called HoE-CNN. This uses an ensemble of deep convolutional neural networks. Convolutional neural networks in an ensemble are used in this method. It addresses FER issues, particularly in multi-class labeling and imbalanced data, and demonstrates improved performance in comparison to individual deep learning techniques, with a focus on online learning applications.[12].

1. **Objective:**

This research paper aims to make a facial emotion recognition model that can help the teaching-learning process with low computing and the highest possible accuracy. The aim is to produce the best possible results with available datasets and known models to leverage and get the maximum possible outcome.

1. **Dataset:**

The dataset is called the Facial Expression Dataset (FED-2013). The dataset contains seven emotions: Fear, Neutral, Angry, Happy, Surprise, Disgust, and Sad.



Fig 1 shows each image is 48 x 48 pixels and greyscale.



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



The dataset was divided into two parts: train and test. The train: test split was

80: 20 % (28709: 7178).

1. **Methodology**

**Data Preprocessing:**

All the images were loaded to a numpy array and then to a tensor. Each image was normalized. The images are then encoded into category vectors.

**Model Architecture:**

The methodology to make the neural network is reasonably intuitive. The first few Convolution layers of the neural network are for feature extraction, and the following dense linear layers are dedicated to classification. In implementing this model, to curb overfitting and optimize training, a dropout of 40 % was chosen and have a relatively large neural network with over 4,232,199 trainable parameters. A convolution kernel size of 3 x 3 was chosen as it is a common practice to choose a smaller size and then try with a larger kernel size of 5 x 5. The standard Rectified Linear Unit (ReLU) was chosen to introduce nonlinearity. The convolution and max pooling layers kept decreasing the image from 48 x 48 to 1 x 1 while doubling the feature maps in every convolution layer and 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 decreases until the size equals the number of classes. Then softmax activation function at the final output layer to get a definitive answer.

The standard T4 GPU of Google Colab was used for this project.

**Architecture:**



1. **Results**

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



**Fig 6.) Showing the loss and accuracy matrices of the model**



**Fig 7.) Showing the exact model predictions**

**Comparison with other models:**

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

**Comparison with other datasets:**

| Sr No | Dataset | Accuracy |
| --- | --- | --- |
| 1. | JAFFE | 68.37% |
| 2. | FER3013 | 64.49% |

1. **Conclusion and Future Objective**

Even though this model for identifying emotions shows promise, a few issues must be resolved in further research. Two datasets, CK+ and FER2013, were used in this work; however, more dataset exploration is necessary for robustness, significantly increasing dataset size with larger subsets of CK+ and FER2013 to improve emotion identification results. Furthermore, future research could improve validation by creating task-specific CNN models for emotion recognition systems and combining various assessment approaches like k-fold cross-validation and statistical analysis. Many directions could be explored in the future. Research on ablation datasets like EMOTIC and EMO-DB could be done to increase generalizability. Moreover, continuous optimization efforts could enhance this model architecture by adding new blocks for increased precision and efficiency.

Moreover, adding speech and body language to the list of emotions that can be recognized provides creative ideas for valuable applications. Conclusion: Although this model shows encouraging results, it is essential to address the constraints that have been found and incorporate them into future developments. With continued study and cooperation, this model has great potential to advance the field of emotion recognition and eventually serve society by becoming more functional and integrating into real-world situations.

1. **References**
2. 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. -
3. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.
9. 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.
10. 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. -
11. 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.
12. Imane Bouslihim(B) and Walid Cherif, “Facial Emotion Recognition Using a GoogLeNet Architecture”
13. Rit Lawpanom, Wararat Songpan \* and Jakkrit Kaewyotha, “Advancing Facial Expression Recognition in Online Learning Education Using a Homogeneous Ensemble Convolutional Neural Network Approach”.
14. 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”
15. Saining Zhang, Yuhang Zhang, Ye Zhang , Yufei Wang and Zhigang Song 4, “A Dual-Direction Attention Mixed Feature Network for Facial Expression Recognition”