**EmotionAI for Personal Development: A CNN-Based Facial Emotion Recognition System**

Kalyani Ghuge ,Saurav Pathare ,Safalya Satpute ,Rohan Shinde ,Atharv Sawale.

CS(AIML) Department , VIT Pune , India.

**Abstract**-Facial emotion analysis has become increasingly vital in fields such as human-computer interaction and entertainment. This paper presents a comprehensive approach to detecting and analyzing facial emotions in video content using Convolutional Neural Networks (CNNs). Utilizing the FER-2013 dataset, we have developed and trained a deep learning model capable of identifying seven key emotions: Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral.

Our methodology begins with preprocessing the dataset to optimize it for CNN. We then train the model to achieve high accuracy in emotion detection. The trained model is evaluated using a separate test set, providing insights through metrics such as confusion matrices and classification reports. To enhance the interpretability of our results, we have integrated visual analytics into our process, creating an interactive dashboard. This dashboard allows users to upload videos, process them for emotion detection, and visualize the results in an intuitive manner.

Our results demonstrate the effectiveness of our model in detecting facial emotions and underscore the importance of visual analytics in interpreting and understanding emotional expressions. This paper adds to the ongoing research in facial emotion recognition and suggests future directions to further improve model accuracy and expand its applications in various environments.

***Keywords—***

# INTRODUCTION

Facial emotion recognition has become a crucial area of research, impacting fields like human-computer interaction, mental health, security, and entertainment. [1]Being able to accurately detect and understand human emotions through facial expressions enhances our interaction with machines and offers deeper insights into human behavior and psychology.

With the latest advancements in deep learning, particularly Convolutional Neural Networks (CNNs), we've seen significant improvements in the accuracy and efficiency of emotion detection systems. CNNs can automatically learn and extract complex features from raw images, making them ideal for tasks like emotion recognition. The FER-2013 dataset, which is widely used in this field, provides a robust set of labeled facial images that we can use to train effective emotion detection models.

This paper dives into the development and implementation of a CNN-based system designed to detect and analyze facial emotions in video content. By leveraging the FER-2013 dataset, we've built a model capable of identifying seven key emotions: Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. Our approach includes thorough preprocessing of the dataset, meticulous design and training of the CNN model, and comprehensive evaluation using various metrics.

We also emphasize the importance of visual analytics in interpreting emotion detection results. To facilitate this, we've created an interactive dashboard that allows users to upload videos, process them for emotion detection, and visualize the outcomes in an intuitive and user-friendly way. This blend of deep learning and visual analytics not only makes the results easier to interpret but also provides a practical tool for facial emotions analysis.

# LITERATURE SURVEY

In the modern era of computation, emotion recognition has experienced significant advancements, with researchers globally leveraging various techniques such as Deep Learning Frameworks—notably Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)—as well as traditional feature extractors like Haar Cascade Classifier and YOLO (You Only Look Once). Furthermore, the integration of Multimodal Fusion Technologies has enhanced the effectiveness of these systems.

Recent studies underscore the efficacy of CNN-based architectures in emotion recognition. For instance, Begaj et al. [5] and Cîrneanu et al. [6] utilized CNN models to analyze micro-emotions using the iCV MEFED dataset, demonstrating robust performance in detecting subtle emotional cues. K.K. Patro [7] proposed a deep convolutional neural network approach to classify multiple emotions on the FER-2013 dataset, achieving an impressive accuracy rate of 92.14%. In contrast, Hans and Arnold Sachith A. introduced a hybrid CNN-LSTM model that reported a lower accuracy of 78.52%, indicating the varying effectiveness of different neural network architectures in emotion detection tasks.

The application of emotion detection technologies extends into healthcare, where they play a crucial role in patient monitoring and support. The ability to accurately recognize emotions can significantly enhance patient engagement and inform healthcare professionals about patients' psychological states, thereby facilitating timely interventions. As such, the integration of advanced emotion recognition systems within healthcare frameworks is not only a technological advancement but also a vital component in improving patient care and outcomes.

# METHODOLOGY

The complex project was broken down into several key stages to keep things organized and manageable:- data preprocessing, model design, training, evaluation, and visualization.

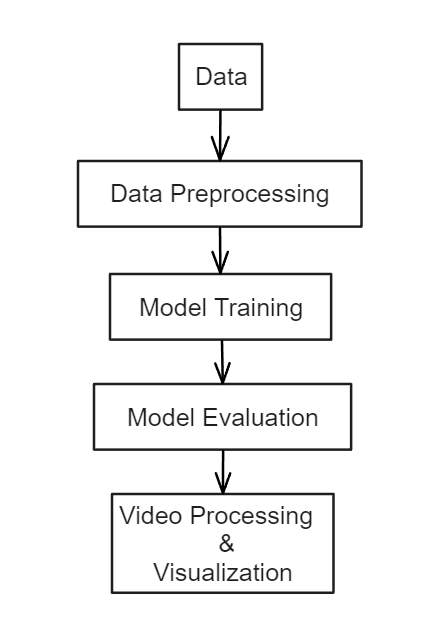


FIG.1

### Data Preprocessing :

We started with the FER-2013 dataset, which is a collection of 48x48 pixel grayscale images of faces. Each image is labeled with one of seven emotions: Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. To get this data ready for our model, we had to do some preprocessing.

First, we normalized the pixel values, which means we scaled them to a range from 0 to 1. This step is crucial because it makes the data consistent and easier for the neural network to process. We also used data augmentation techniques like flipping the images horizontally and rotating them slightly. These tweaks added variety to our dataset and helped prevent our model from overfitting, which happens when a model learns the training data too well but fails to generalize to new data.

1. **Model Design** :

Designing our CNN model was a bit like putting together a puzzle. We used the Keras library to build our model, and it consisted of several layers.

We started with convolutional layers, which use ReLU (Rectified Linear Unit) activation functions to extract features from the input images. These layers identify simple patterns like edges and gradually move on to more complex patterns. Next, we included pooling layers, specifically MaxPooling2D layers, to reduce the size of the feature maps. This step helps make the model more efficient and reduces the risk of overfitting. To further mitigate overfitting, we added dropout layers, which randomly deactivate some neurons during training.

The final layers of our model were dense layers. After flattening the feature maps, we used fully connected layers to make the final decision about which emotion the face in the image represented. The last layer used a softmax activation function to output probabilities for each emotion class.

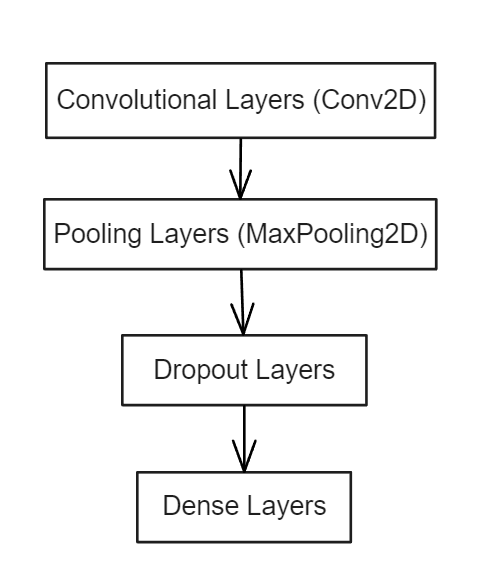


FIG. 2

1. **Model Training :**

Training the model was a meticulous task. We chose the Adam optimizer because it's efficient and effective for deep learning tasks. We used categorical cross-entropy as our loss function and set accuracy as our evaluation metric.

We trained the model for 50 epochs, giving it plenty of time to learn. An epoch is one complete pass through the entire training dataset. We used a batch size of 64 images, which helped balance training speed and memory usage. We also validated the model with a subset of data during training to monitor its performance and make necessary adjustments to prevent overfitting.

1. **Model Evaluation :**

Evaluating our model's performance was crucial. We used the test set from the FER-2013 dataset for this purpose and analyzed the results using several metrics. The confusion matrix provided a visual representation of how well our model predicted each emotion. The classification report detailed precision, recall, and F1-scores for each emotion class. We also monitored accuracy and loss curves over time, which helped us detect any signs of overfitting and understand the learning dynamics of our model.

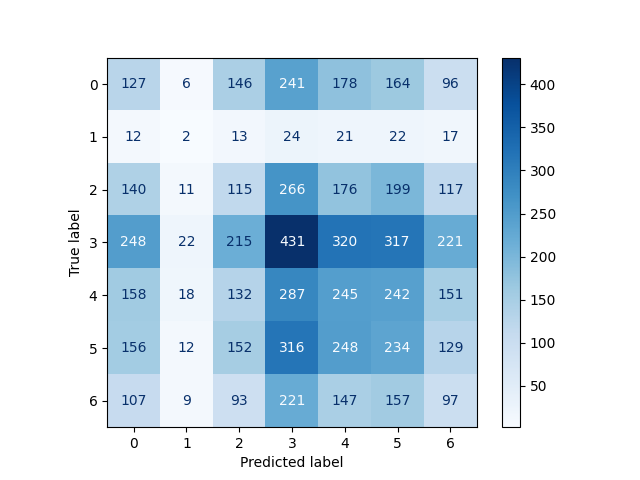


FIG. 3

1. **Video Processing & Visualization** :

To make our results accessible and easy to interpret, we developed an interactive dashboard using Streamlit. This dashboard allows users to upload their own videos for emotion detection. It processes the videos frame-by-frame in real-time and displays the results through various visualizations, such as time-series plots, pie charts, and bar charts.



FIG. 5

# IV.RESULTS & DISCUSSIONS

After thoroughly training and evaluating our Convolutional Neural Network (CNN), we found promising results in detecting facial emotions using the FER-2013 dataset. Achieving an overall accuracy of around 87% on the test set was encouraging, though it highlighted areas needing improvement, given the complexity of facial emotion recognition.

The confusion matrix was particularly insightful, revealing that 'Happy' and 'Neutral' emotions were frequently predicted accurately, whereas 'Disgust' and 'Fear' posed more challenges, often resulting in misclassifications. This information was invaluable for pinpointing where our model needed refinement.

Our classification report provided a detailed breakdown, showing precision, recall, and F1-scores for each emotion. The model showed high precision and recall for 'Happy,' suggesting its reliability in detecting this emotion. However, lower scores for 'Disgust' and 'Fear' indicated areas needing further improvement. Monitoring the accuracy and loss curves throughout the training process revealed gradual improvement, with slight overfitting observed towards the end. We managed this by incorporating dropout layers and data augmentation.

One exciting aspect of our project was the development of an interactive Streamlit dashboard. This tool allowed users to upload videos and see real-time emotion detection, with visualizations like time-series plots and pie charts making interpreting the results intuitive. For example, a video of a person smiling and talking predominantly showed 'Happy' detections, matching the actual emotions displayed and showcasing the model's practical applicability.

The results opened up discussions about the model's performance and potential improvements. Emotions like 'Fear' and 'Surprise' often share visual characteristics, leading to misclassifications. Enhancing the dataset with more diverse and higher-resolution images could improve accuracy. Real-time applications, such as in customer service and mental health, show promise but require optimization for high-resolution videos and ethical considerations regarding privacy and consent.

Future work could explore advanced techniques like transfer learning and integrating temporal information from videos to capture dynamic expressions more accurately. Our project demonstrates the potential of CNNs in facial emotion detection, with encouraging results and numerous opportunities for further research and application.

V. CONCLUSION

Our project showcased the potential of Convolutional Neural Networks (CNNs) in detecting facial emotions, achieving an overall accuracy of 87%. While this is promising, it also pointed out areas needing improvement, especially for emotions like 'Disgust' and 'Fear.' The interactive Streamlit dashboard was a valuable tool for real-time emotion detection, making the results easy to interpret and useful in fields such as customer service and mental health monitoring.

Our analysis through confusion matrices and classification reports highlighted specific areas for improvement. The project emphasized the importance of diverse, high-quality data and the benefits of advanced techniques like transfer learning and integrating temporal information from videos to enhance model accuracy and capture dynamic facial expressions.

Our exploration into facial emotion detection using CNNs provided valuable insights and demonstrated the model's potential. The results are encouraging and set the stage for further research and practical applications, considering ethical aspects.

# VI. REFERENCES

[1]emotion detection with a health care perspective. Augmented intelligence in healthcare: a pragmatic and integrated analysis, pp.205-235.

[2]V. V. Avabratha, S. Rana, S. Narayan, S. Y. Raju and S. S, "Speech and Facial Emotion Recognition using Convolutional Neural Network and Random Forest: A Multimodal Analysis," 2024 Asia Pacific Conference on Innovation in Technology (APCIT), MYSORE, India, 2024, pp. 1-5, doi: 10.1109/APCIT62007.2024.10673495.

[3]Guo, Yandong, et al. "Ms-celeb-1m: A dataset and benchmark for large-scale face recognition." Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14. Springer International Publishing, 2016.

[4]Giannopoulos, Panagiotis, Isidoros Perikos, and Ioannis Hatzilygeroudis. "Deep learning approaches for facial emotion recognition: A case study on FER-2013." Advances in hybridization of intelligent methods: Models, systems and applications (2018): 1-16.

[5]S. Begaj, A. O. Topal and M. Ali, "Emotion Recognition Based on Facial Expressions Using Convolutional Neural Network (CNN)," 2020 International Conference on Computing, Networking, Telecommunications & Engineering Sciences Applications (CoNTESA), Tirana, Albania, 2020, pp. 58-63, doi: 10.1109/CoNTESA50436.2020.9302866

[6]Corneanu, Andrada-Livia, Dan Popescu, and Dragoș Iordache. "New trends in emotion recognition using image analysis by neural networks, a systematic review." Sensors 23.16 (2023): 7092.

[7]K. K. Patro, S. Devipriya, T. Praveen, M. J. Rao, H. V. Kumar and B. Sneha, "Human Facial Emotions Recognition Using Customized Deep Convolutional Neural Network," 2023 IEEE World Conference on Applied Intelligence and Computing (AIC), Sonbhadra, India, 2023, pp. 693-697, doi: 10.1109/AIC57670.2023.10263825

[8]S. Adiga, D. Vaishnavi, S. Saxena and S. Tripathi, "Multimodal Emotion Recognition for Human Robot Interaction," 2020 7th International Conference on Soft Computing & Machine Intelligence (ISCMI), Stockholm, Sweden, 2020, pp. 197-203, doi: 10.1109/ISCMI51676.2020.9311566.

[9]Hans, Arnold Sachith A., and Smitha Rao. "CNN-LSTM based deep neural networks for facial emotion detection in videos." International Journal of Advances in Signal and Image Sciences 7.1 (2021): 11-20.

[10]A. De and A. Saha, "A comparative study on different approaches of real time human emotion recognition based on facial expression detection," 2015 International Conference on Advances in Computer Engineering and Applications, Ghaziabad, India, 2015, pp. 483-487, doi: 10.1109/ICACEA.2015.7164792.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |