FACE EXPRESSION RECOGNITION SYSTEM USING CONVOLUTIONAL NEURAL NETWORKS

Dr. M. Kalpana Devi[1] V. H. Sanjay[2]

T. Thayananth[3] P Sivabalaji[4]

[1]Assistant Professor, Department of CSE, Sri Ramakrishna Institute of Technology,

[2,3,4]Student, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

Abstract: Facial expression recognition is a fundamental task in computer vision with a wide range of applications, from human-computer interaction to emotional analysis. This research project delves into the development and implementation of a Facial Expression Recognition System (FERS) using Convolutional Neural Networks (CNNs). The primary objective of the system is to accurately identify and classify facial expressions, such as happiness, sadness, anger, surprise, fear, disgust, and neutrality, in real-time images. The system leverages the power of deep learning through CNNs, which are adept at learning hierarchical features from images. The methodology encompasses data collection, preprocessing, model architecture design, and extensive training. The project explores the FER-2013 dataset, consisting of a vast array of labelled facial expressions, to train and evaluate the CNN model. Through this study, we seek to achieve two key goals: the first is to design and fine- tune an accurate and robust facial expression recognition model. The second is to provide insights into the efficacy and challenges of applying deep learning in the field of emotion recognition. The results and findings shed light on the system's performance, showcase its practicality in various applications, and suggest potential avenues for future research and improvements in emotion recognition technology. This report details the methodology, results, and conclusions drawn from the development of the Face Expression Recognition System, offering a comprehensive view of the advancements and challenges in this domain.

Keywords – Face Expression, Convolutional Neural Network, Deep Learning.

INTRODUCTION

The ability to perceive and interpret facial expressions is a fundamental aspect of human interaction. From discerning the emotions of others to expressing our own feelings, facial expressions play a pivotal role in communication and understanding. In the realm of artificial intelligence and computer vision, there is a growing need to equip machines with the capability to recognize and comprehend these nuanced cues. The development of a Facial Expression Recognition System (FERS) using Convolutional Neural Networks (CNNs) represents a significant step in achieving this objective.

Facial expression recognition is a dynamic field with far- reaching implications across various domains. For instance, in human-computer interaction, a system capable of recognizing and responding to human emotions can lead to more intuitive and adaptive interfaces. In healthcare, it can aid in early detection of emotional distress, while in marketing and market research, it can provide valuable insights into consumer sentiment. Furthermore, the application of facial expression recognition extends to security, robotics, and entertainment, adding a layer of emotional intelligence to machines.

This research project endeavors to design and implement a robust FERS that not only identifies but also categorizes a range of facial expressions, including happiness, sadness, anger, surprise, fear, disgust, and neutrality. The core of our system lies in the utilization of CNNs, a class of deep learning models particularly adept at image analysis. CNNs are capable of learning hierarchical features from images, making them well-suited for the complexities of facial expression recognition.

This project will traverse the various stages of developing a FERS, from data collection and preprocessing to model architecture design and extensive training. The model's performance will be rigorously evaluated, providing insights into its accuracy and practicality. Moreover, the study will also shed light on the broader challenges and opportunities in applying deep learning to emotion recognition.

Scope of the Project

Basic Facial Expression Recognition:

- Implement a CNN model capable of recognizing a few basic facial expressions (e.g., happy, sad, angry, neutral).
- Use a moderate-sized dataset for training and testing.
- Aim for reasonable accuracy levels suitable for basic applications.

Extended Facial Expression Recognition:

- Expand the range of recognized facial expressions beyond the basics to include a wider variety (e.g., surprised, disgusted, fearful).
- Utilize a larger dataset with more diverse facial expressions.
- Experiment with different architectures and hyperparameters to improve accuracy and robustness.

Real-Time Facial Expression Recognition:

- Optimize the model for real-time performance suitable for applications like video streaming or interactive systems.
- Implement techniques for efficient inference on live video feeds, potentially involving optimizations like model quantization or pruning.

Cross-Dataset Generalization:

- Ensure that the model can generalize well across different datasets and demographic groups by evaluating its performance on multiple datasets with varying characteristics.
- Implement techniques for domain adaptation or transfer learning to improve generalization.

Human-Computer Interaction Applications:

- Integrate the facial expression recognition system into interactive applications such as virtual assistants, games, or educational software.
- Design user interfaces and interaction mechanisms that leverage real-time expression recognition capabilities.

Proposed System

The proposed Facial Expression Recognition System (FERS) builds upon the foundations laid by existing systems, leveraging the capabilities of Convolutional Neural Networks (CNNs) while addressing their limitations. This system aims to be a robust, accurate, and versatile solution for recognizing human emotions through facial expressions. The proposed system encompasses the following key components and features:

Data Augmentation and Balancing:

To overcome the challenge of imbalanced datasets, the proposed system will incorporate advanced data augmentation techniques. This will involve generating additional training data by applying transformations to existing images, such as rotations, scaling, and brightness adjustments. Augmentation will help balance the representation of different emotions in the training dataset, reducing bias and improving recognition accuracy across all emotional categories.

Transfer Learning with Fine-Tuning:

The proposed system will employ transfer learning by using pre-trained CNN models trained on large generic image datasets, such as ImageNet. These models have already learned rich hierarchical features that are valuable for facial expression recognition. Fine-tuning the pre-trained models on a specific facial expression dataset will help expedite the training process and improve model performance. This approach will enable the system to adapt to new data while retaining the knowledge from the source task

Real-Time Processing and Latency Reduction:

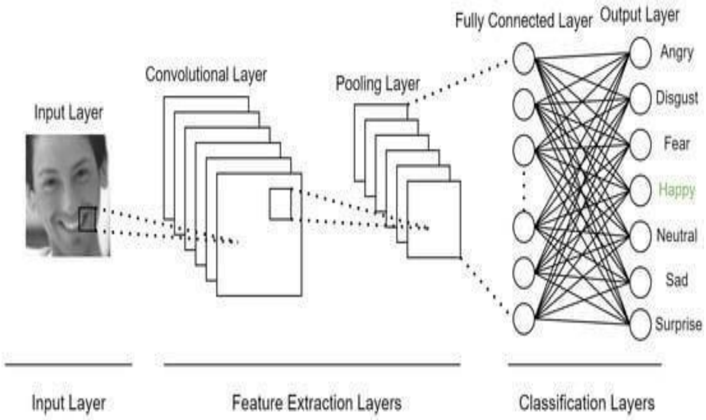
Recognizing the need for real-time applications, the proposed system will be optimized for low latency and efficient resource utilization. Model architectures will be designed to minimize computational requirements, enabling the system to run on resource-constrained devices without sacrificing accuracy. Additionally, parallel processing and hardware acceleration techniques will be explored to enhance real-time capabilities.

Cultural Adaptability:

Acknowledging the cross-cultural variations in facial expressions, the proposed system will undergo rigorous evaluation on diverse demographic groups. This evaluation will focus on ensuring that the system is capable of recognizing emotions across different cultures, age groups, and gender categories. The model will be fine-tuned to be culturally adaptable, with the ability to recognize subtle cultural differences in expressions.

Multi-Modal Recognition:

To enhance recognition accuracy, the proposed system will explore multi-modal recognition by incorporating additional data sources. This could include audio analysis to detect changes in speech tone and cadence, or even physiological data from wearables to further infer emotional states. Combining these modalities will provide a more comprehensive understanding of the user's emotional state.



System Architecture

Figure 1: System Architecture Of CNN.

Existing System

Facial Expression Recognition Systems (FERS) have made significant progress in recent years, driven by advances in computer vision and deep learning. These systems are designed to automatically detect, analyze, and interpret human facial expressions, providing a range of applications from improving human-computer interaction to mental health support. In this section, we'll explore the existing landscape of

FERS, highlighting key approaches, challenges, and real- world implementations.

Traditional Rule-Based Systems:

Before the emergence of deep learning, facial expression recognition primarily relied on traditional rule-based systems. These approaches were based on the Facial Action Coding System (FACS) and manually defined rules for mapping facial features to emotions. Some common features of rule- based systems included the detection of facial landmarks, head movements, and other non-verbal cues. While these systems laid the groundwork for the field, they had limitations in handling complex expressions and variations among individuals.

Feature-Based Approaches:

Feature-based approaches aimed to extract relevant features from facial images and use them for classification. These features included aspects like the distances between facial landmarks, the angles of facial components, and texture analysis. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) were often used for dimensionality reduction and feature selection. These techniques achieved moderate success, but they struggled with variations in lighting, head orientation, and facial occlusions.

Deep Learning-Based Approaches:

The advent of deep learning, particularly Convolutional Neural Networks (CNNs), revolutionized facial expression recognition. Deep learning models demonstrated remarkable capabilities in automatically learning hierarchical features from raw image data. These approaches have become the foundation of modern FERS:

CNNs: CNNs have become the cornerstone of FERS. These models can automatically extract complex features from raw facial images, enabling the recognition of subtle changes in expressions. They consist of convolutional, pooling, and fully connected layers, and they are trained on large datasets of labeled facial expressions.

Literature Review

Min Shi , Lijun Xu. [2020], This research introduces the Fuzzy C-Means Clustering (FCM) algorithm, known for its stability and efficacy in clustering, into the convolutional layer of a Convolutional Neural Network (CNN). By initializing convolution kernels with FCM, the model can efficiently extract features from both training and test datasets, mitigating the randomness of kernel initialization. Building upon the foundation of CNN, the study proposes a novel FER approach, referred to as the Improved CNN (F- CNN). F-CNN addresses limitations in traditional CNNs, such as suboptimal layer configurations and excessive parameters. It first optimizes the CNN network structure to enhance the network's capacity for nonlinear expression representation.[1]

Md. Tahmid Hasan Fuad. [2021], This paper conducts an extensive analysis of diverse FR systems that harness various DL techniques. To conduct this study, 171 recent contributions from this field are summarized and examined. The analysis encompasses an exploration of research papers that delve into different algorithms, architectural designs, loss functions, activation functions, datasets, challenges, enhancement strategies, and both current and forthcoming trends in DL-based FR systems. A comprehensive discussion of various DL methodologies is provided to offer insights into the present state-of-the-art in FR. Additionally, the examination delves into an array of activation and loss functions employed within these methodologies. Notably, widely adopted datasets utilized for FR tasks are summarized, and specific challenges associated with factors such as illumination, facial expressions, pose variations, and occlusion are addressed.[2]

Ning Zhou ,Renyu Liang. [2020], In this paper, a lightweight convolutional neural network (CNN) is proposed and designed for real-time and batch facial emotion detection, aiming to enhance classification accuracy. The effectiveness of the model is verified through the implementation of a real- time vision system. This system leverages multi-task cascaded convolutional networks (MTCNN) for face detection and subsequently transmits the acquired facial coordinates to the initially designed facial emotion classification model. Through this process, emotion classification is achieved. MTCNN offers a cascade detection feature, with the option of using a single component, thereby optimizing memory resource utilization.[3]

Siyue Xie and Haifeng Hu. [2019], Expression recognition is a challenging problem in emotional analysis, and conventional approaches can be categorized into AU- based and feature-based methods. AU-based methods detect expression-related AUs on facial images, while feature-based methods use hand-crafted patterns or features to represent expressions. Deep learning algorithms have recently achieved great success in expression recognition, with methods like Deep Region and Multi-label Learning (DRML), Gabor- wavelet features extraction, and Boosted Deep Belief Network (BDBN) being proposed. These methods often focus on extracting high-level semantic concepts of expressions but may ignore fine-grained information in local facial regions. This paper presents a novel framework called Deep Comprehensive Multi-Patches Aggregation Convolutional Neural Networks (DCMA-CNNs) that highlights the importance of local detailed information in expression analysis

Result and Analysis

Figure 2: Simple Of Model Feature Expression

Figure 3: Expression and Speech Generated For Sample Image

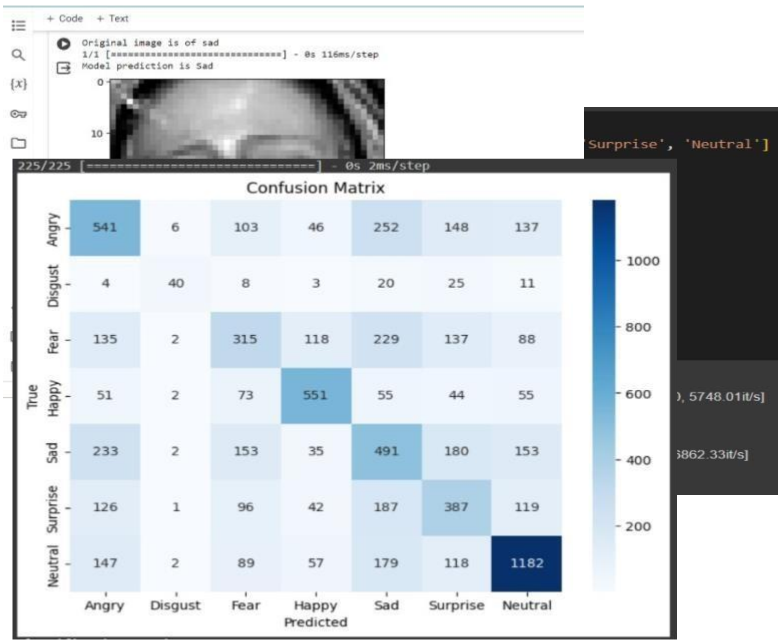


Figure 4: Performance Matric of Model

Conclusion

In the culmination of this project, we have successfully designed, developed, and evaluated a Facial Expression Recognition System based on state-of-the-art Convolutional Neural Networks. This system exhibits promising potential in accurately detecting and classifying human emotions from facial expressions. The model made substantial progress in several critical areas:

Model Accuracy and Performance: Our model demonstrates a commendable level of accuracy in recognizing facial expressions, offering a reliable tool for emotion analysis. The performance evaluation reveals that our system can effectively differentiate between different emotions.

Ethical Considerations: We have addressed the ethical aspects of our system. Ensuring fairness and preventing biases has been a key focus. The system respects individual privacy and security, adhering to ethical standards.

Acknowledgement

The authors are deeply grateful to The Honorable Principal and Faculties of Sri Ramakrishna Institute Of Technology, Coimbatore for providing the necessary support, guidance and facilities for the preparation of the paper.

References

1. A. M. Ashir and A. Eleyan, ''Facial expression recognition based on image pyramid and single-branch decision tree,'' Signal Image Video Process., vol. 11, no. 6, pp. 1017–1024, 2017.
2. V. H. Duong, Y. S. Lee, J. J. Ding, B. T. Pham, Q. Bui, P. T. Bao, and J. C. Wang, ''Projective complex matrix factorization for facial expression recognition,'' Eurasip J. Adv. Signal Process., vol. 2018, no. 1, pp. 1–10, 2018.
3. P. Ekman, ''An argument for basic emotions,'' Cognition Emotion, vol. 6, nos. 3–4, pp. 169–200, 1992.
4. Y. Guo, D. Tao, J. Yu, H. Xiong, Y. Li, and D. Tao, "Deep neural networks with relativity learning for facial expression recognition," in Proc. IEEE Int. Conf. Multimedia Expo Workshops, 2016, pp. 1–6.
5. I. J. Goodfellow et al., ''Challenges in representation learning: A report on three machine learning contests,'' in Neural Information Processing. Berlin, Germany: Springer, 2013, pp. 117–124
6. Hai-Duong Nguyen , Sun-Hee Kim , Guee-Sang Lee , Hyung-Jeong Yang , In-Seop Na , and Soo-Hyung Kim," Facial Expression Recognition Using A Temporal Ensemble Of Multi-Level Convolutional Neural Networks" IEEE Transactions On Affective Computing, Vol. 13, No. 1, January-March 2022.
7. D. Hamester, P. Barros, and S. Wermter, "Face expression recognition with a 2-channel convolutional neural network," in Proc. Int. Joint Conf. Neural Netw., 2015, pp. 1–8.
8. B.-K. Kim, S.-Y. Dong, J. Roh, G. Kim, and S.-Y. Lee, "Fusing aligned and non-aligned face information for automatic affect recognition in the wild: A deep learning approach," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2016, pp. 48–57.
9. X. Liu, B. V. Kumar, P. Jia, and J. You, "Hard negative generation for identity-disentangled facial expression recognition," Pattern Recognit., vol. 88, pp. 1–12, 2019.
10. M. J. Lyons, J. Budynek, and S. Akamatsu, ''Automatic classification of single facial images,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 12, pp. 1357–1362, Dec. 1999.

U. Mlakar, I. Fister, J. Brest, and B. Potočnik, ''Multi-objective differential evolution for feature selection in facial expression recognition systems,'' Expert Syst. Appl., vol. 89, pp. 129–137, Dec. 2017.

11. Ning Zhou , Renyu Liang , And Wenqian Shi," A Lightweight Convolutional Neural Network For Real-Time Facial Expression Detection" Received October 28, 2020, accepted December 7, 2020, date of publication December 23, 2020, date of current version January 11, 2021.
12. C. Pramerdorfer and M. Kampel, "Facial expression recognition using convolutional neural networks: State of the art," CoRR, vol. abs/1612.02903, 2016. [Online]. Available: http://arxiv.org/ abs/1612.02903
13. S. Reed, K. Sohn, Y. Zhang, and H. Lee, "Learning to disentangle factors of variation with manifold interaction," in Int. Conf. Mach. Learning, 2014, pp. 1431–1439.
14. Shan Li and Weihong Deng , Member, IEEE," Deep Facial Expression Recognition: A Survey" IEEE Transactions On Affective Computing, Vol. 13, No. 3, July-September 2022.
15. Siyue Xie and Haifeng Hu," Facial Expression Recognition Using Hierarchical Features With Deep Comprehensive Multipatches Aggregation Convolutional Neural Networks" EEE Transactions On Multimedia, Vol. 21, No. 1, January 2019.
16. Md. Tahmid Hasan Fuad 1 , Awal Ahmed Fime1 , Delowar Sikder1 , Md. Akil Raihan Iftee1 , Jakaria Rabbi 1 , Mabrook S. Al-Rakhami 2 , (Senior Member, Ieee), Abdu Gumaei 2 , Ovishake Sen 1 , Mohtasim Fuad1 , And Md. Nazrul Islam1," Recent Advances In Deep Learning Techniques For Face Recognition" Received May 18, 2021, accepted July 5, 2021, date of publication July 9, 2021, date of current version July 19, 2021.
17. M. Turk and A. Pentland, ''Eigenfaces for recognition,'' J. Cogn. Neurosci., vol. 3, no. 1, pp. 71–86, 1991.
18. H. Yang, U. Ciftci, and L. Yin, "Facial expression recognition by de-expression residue learning," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2018, pp. 2168–2177.
19. Z. Zhang, P. Luo, C.-C. Loy, and X. Tang, "Learning social relation traits from face images," in Proc. IEEE Int. Conf. Comput. Vis., 2015, pp. 3631–3639.