Second International Conference on Smart Computing Communication for Sustainable Convergence (ISCCSC-2025)

Facial Expression Recognition using Deep Learning Techniques

Grishma Babu Shetty[1] Kruthika V Shetty[2] Kshama shetty[3] Shazma Gulzar Ahamed[4] Kasilingam N[5]

[1,2,3,4]*Alvas Institute of Engineering and Technology, Department of Computer Science and Engineering (Data Science), Moodbidri, Manglore, Karnataka, India. [5] Alvas Institute of Engineering and Technology, Department of Artificial Intelligence and Machine Learning, Moodbidri, Manglore, Karnataka, India.,*

kasilingamn@gmail.com

Abstract

This chapter surveys recent advances in FER. Then, it focuses on the structure of automatic FER systems: face acquisition, feature extraction, representation, and classification. It discusses several key challenges such as variation in expression intensity, individual differences, head pose variation, image resolution, and unreliable ground truth. Overfitting, illumination or pose variation, and other expression-independent variations are reviewed with respect to how deep learning techniques address it. The most popular datasets, including JAFFE and CK+, and some deep neural networks for static and dynamic expressions are analyzed. Methods are compared by benchmark performance, underlining the strengths and limitations, and future challenges for real-world FER.

Keywords Facial Expression Recognition (FER), Automatic Facial Expression Analysis, Deep Learning, Over fitting, Head Pose, Pre-processing, Classification, Datasets, JAFFE, CK, Neural Net works, Static Images, Dynamic Sequences, Computer Vision, Machine Learning, Pattern Recognition.

1. INTRODUCTION

FER is an important area of research in computer vision and AI with equally wide-ranging applications in healthcare, education, security, marketing, and human-computer interaction. The first generation FER systems were based on manually designed features such as facial landmarks, LBP, Gabor filters, and PCA, which have proved to be weak in coping with real-world variations such as changes in illumination, pose, occlusion, and inter-person variability. But recent advancements in deep learning, especially in the area of CNNs, have allowed end-to-end learning and efficient hierarchical identification of image features in FER. However, micro-expression recognition, inter-person variability, less availability of diversely labeled datasets, and less generalization towards real-world settings have remained challenging. The current research in this area focuses on a multimodal approach, incorporating facial expressions with voice, non-verbal communications, and physiological signals to make it more accurate and robust. The efforts towards identification of compound emotions, analysis of micro-expressions in a better manner, and less dependency on large labeled datasets via transfer learning and unsupervised learning have a lot more to achieve in making FER systems used in a non-arbitrary, meaningful, and privacy-preserving manner.

2. PREVIOUS RESEARCH

State-of-the-art performance of Emotion Recognition Systems has been massively improved due to recent breakthroughs in Facial Expression Recognition using deep learning methods [1]. MobileNet V2, the combination of CNN & RNN, and other methods have advanced image and video analysis systems such as some recently developed systems, which include emotional change points identification and thermal imaging systems, impressively performing in this area [2,3]. Regarding their applications, these systems could be used in healthcare systems, educational institutions, autism identification systems, stress analysis systems, sign language identification systems, and music suggestion systems based on emotions [4,5]. Although different systems are emerging in the field of emotion recognition, there are a few aspects of culture variation, micro-expressions, and ethics, which need further exploration in online analyses [6,7].

3. METHODOLOGY

The improvement approach for FER is strategic enhancement at every stage, which includes data gathering, model construction, and deployment. The datasets contain faults that can be addressed more directly by focusing the development of the balanced and diversified datasets from wide-ranging expressions to diverse demographics, and real-world conditions. Other supplement techniques are crowd sourcing and synthetic data generation by GAN that may be able to supplement a dataset with otherwise rare and difficult to capture scenarios such as occlusions or extreme emotions. Models, multimodal architectures are highly important in their design. They integrate visual information with other inputs, such as audio, text, and physiological signals. For example, combining facial features with speech tone or textual sentiment improves the correctness of emotional recognition. Methods of 3D modeling, which include depth data and complex 3D CNNs, allow for a more detailed analysis of the facial geometry, which is not affected by variations in pose or occlusions. In the context of self-supervised learning, models are designed to predict patterns on unannotated data, thus cutting down huge annotations. Optimization for real-time applications is achievable with light architectures like MobileNet and techniques that allow knowledge distillation thereby ensuring FER's effective deployment on edge devices. At the methodological level, these will amplify performance, scalability, and adaptability of FER systems to real-world applications.

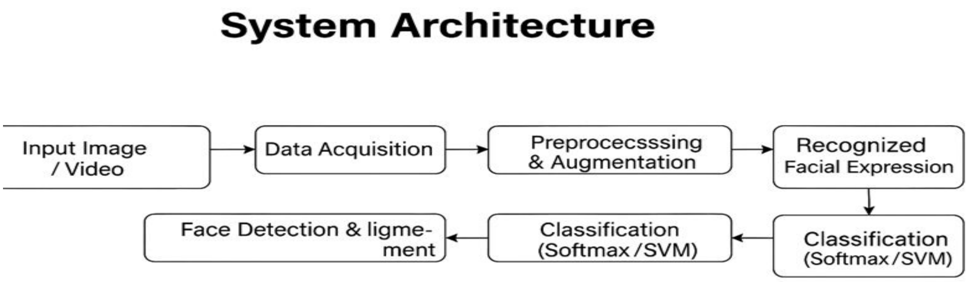

Figure 3. 1 System Architecture

4. RESULT ANALYSIS

Recent studies using FER have shown that deep learning-based methods have a great advantage over traditional methods based on handcrafted features in terms of both accuracy and robustness. Performance in standard evaluation datasets such as FER2013 and CK+, with a more efficient computation, can be raised based on deep learning methods, especially CNN models with lightweight variants such as MobileNetV2. Hybrid models using both CNN and RNN have achieved better results in processing video-based FER by assessing temporal patterns and emotional changeover. Multi-modal and 3D-oriented FER systems have increased robustness in processing light, pose, and occlusion variation. The performance in imbalanced datasets, micro-expression classification, and real-world FER is impaired due to reduced accuracy, complexity in micro-expression analysis, and poor model generalizability, respectively.

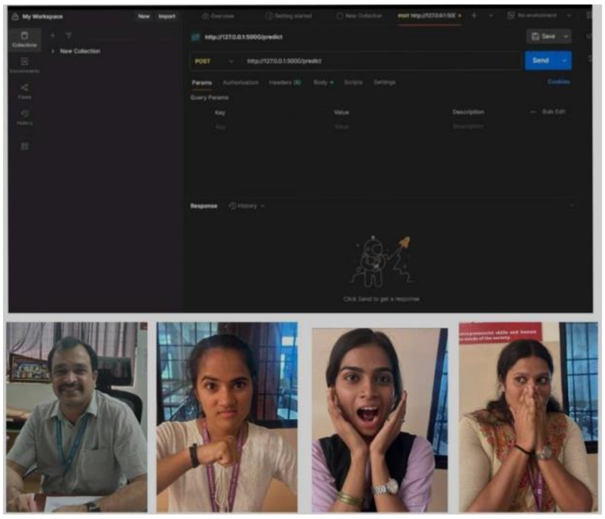

Figure 4. 1 Result

5. CONCLUSION

With recent advancements in Facial Expression Recognition based on deep learning approaches, the performance of Emotion Recognition Systems has substantially improved. MobileNet V2 and combination models involving CNN & RNN have enhanced image and video analysis systems, including some newly developed approaches such as emotional change point detection and thermal imaging, which have demonstrated impressive results in this domain. Facial Expression Recognition systems can be used in healthcare, education, autism identification, stress analysis, sign language identification, and music recommendations based on emotions. Even though different systems have evolved in this area, some aspects related to culture variation, micro-expressions, online analysis, and ethics have to be explored further.

6. REFERENCES

[1] S. Li and W. Deng, "Deep facial expression recognition: A survey," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1195 1215, 2022.

[2] I. M. Revina and W. S. Emmanuel, "A survey on human face expression recognition techniques," *Journal of King Saud University-Computer and Information Sciences*, vol. 33, no. 6, pp. 619–628, 2021.

[3] Y. Huang, F. Chen, S. Lv, and X. Wang, "Facial expression recognition: A survey," *Symmetry*, vol. 11, no. 10, p. 1189, 2019.

[4] F. J. Ming, S. Shabana Anhum, S. Islam, and K. H. Keoy, "Facial emotion recognition system for mental stress detection among university students," in 2023 *3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, 2023, pp. 1–6.

[5] N. Sarode and S. Bhatia, "Facial expression recognition," *International Journal on computer science and Engineering*, vol. 2, no. 5, pp. 1552 1557, 2010.

[6] O. S. Ekundayo and S. Viriri, "Facial expression recognition: A review of trends and techniques," *Ieee Access*, vol. 9, pp. 136944–136973, 2021.

[7] D. G. Lobo, J. S. Goutham, M. Aman, D. Vishma, H. Kunder, and K. M. KiranRaj, "Emolens: Research on enhancing recommendation systems through facial emotion recognition with cnn and mobilenetv2," in 2024 *15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2024, pp. 1–6.

Biographies

Grishma Babu Shetty is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects a strong interest in core and emerging domains such as Artificial Intelligence and Machine Learning, Natural Language Processing and Deep Learning.

Kruthika V Shetty is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Machine Learning, Natural Language Processing and Data Science.

Kshama Shetty is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Artificial Intelligence, Machine Learning and Web technologies.

Shazma Gulzar Ahamed is currently pursuing a Bachelor's degree in Computer Science and Engineering (Data Science) at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka. Her academic journey reflects interest in emerging domains such as Gen AI, Web technologies and Machine Learning.

Kasilingam N is a senior assistant professor in the Department of Artificial Intelligence and Machine Learning at Alva's Institute of Engineering and Technology (AIET), affiliated with Visvesvaraya Technological University (VTU) in Karnataka.