IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of
Information Technology & Computer Engineering
ISSN 2347–3657
Volume 12, Issue 2, 2024

FACIAL EXPRESSION RECOGNITION SYSTEM USING DEEP LEARNING MODELS BASED ON HUMAN EMOTIONS THROUGH CLASSIFICATION WITH CNN,RNN AND YOLO OBJECT DETECTION ALGORITHMS

1MR. VENKATESHWARLU,2N SANJU BHARGAVI,3THOUTAM VARUN KUMAR,4D RISHI KUMAR,5KATRAVATH SHIVA KUMAR,6PANDUGA JASHWANTH,7SAI KUMAR MUDHIRAJ

[1]Assistant Professor,department of information technology, malla reddy institute of engineering and technology(autonomous),dhulapally,secundrabad,chandrunit26@gmail.com

[2,3,4,5,6]UG students, department of information technology,malla reddy institute of engineering and technology(autonomous),Dhulapally,Secundrabad

ABSTRACT

Emotion is a key topic in a variety of professions, including biomedical engineering, psychology, neuroscience, and mental health. This component of emotion recognition is crucial, because it is commonly used in the diagnosis of human brain and psychiatric diseases. Deep learning has gotten a lot of users' interest in the field of picture categorization, according to a recent poll. These emotions are employed not just for brain diagnosis, but also as a recommendation system to help consumers select goods that meet their requirements and preferences. This inspired us to create a system that can accurately and efficiently discern emotions based on the user's facial expressions. In this proposal, we aim to create an application that can be used to anticipate expressions in both still and moving photographs. Then compare the results of the CNN with the recurrent neural network (RNN) model. Once the image is taken from the video sequences, the system uses HAAR cascade to detect faces, crops the image, resizes it to the necessary dimension, and sends it to the model for prediction. Seven probability values will be generated by the model, matching to seven expressions. We compare the two models to see which one provides better face expression detection accuracy for the image dataset.

I. INTRODUCTION

Project introduces an innovative approach to understanding and interpreting human emotions through facial expressions. Emotions play a significant role in human communication and interaction, and the ability to accurately recognize and interpret facial expressions is essential in various fields, including psychology, human-computer interaction, and affective computing. Traditional methods of facial expression recognition often rely on handcrafted features and shallow learning models, which may struggle to capture the complex and subtle variations in facial expressions. In response to this challenge, this project proposes the development of a comprehensive facial expression recognition system that leverages deep learning models, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and You Only Look Once (YOLO) object detection algorithms. By analyzing facial images and sequences, the project aims to classify facial expressions based on human emotions with high accuracy and efficiency. The integration of CNNs for feature extraction, RNNs for temporal modeling, and YOLO for object detection enables the system to capture spatial and temporal dynamics in facial expressions and accurately classify them into different emotion categories. Through the implementation of this advanced facial expression recognition system, the project seeks to contribute to the development of more intuitive and responsive human-computer interfaces, emotion-aware applications, and assistive technologies that can better understand and respond to human emotions.

II.LITERATURE REVIEW

Facial expression recognition (FER) has garnered significant attention in the field of computer vision and affective computing due to its wide-ranging applications in human-computer interaction, emotion-aware computing, and psychological research. Over the years, researchers have explored various approaches and techniques to improve the accuracy and robustness of FER systems. This literature review provides an overview of recent advancements in FER, with a focus on deep learning models and their application in recognizing human emotions from facial images.

1. Deep Learning Approaches for FER: Deep learning techniques, particularly Convolutional Neural Networks (CNNs), have revolutionized FER by enabling automatic feature learning directly from raw image data. Various CNN architectures, such as AlexNet, VGGNet, and ResNet, have been applied to FER tasks, achieving remarkable performance improvements. For instance, Zhang et al. (2018) proposed a multi-task deep CNN model for

simultaneous facial expression recognition and facial action unit detection, achieving state-of-the-art results on benchmark datasets.

2. Temporal Modeling with Recurrent Neural Networks (RNNs): Recognizing facial expressions often requires capturing temporal dynamics in facial movements over time. Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, have been employed to model temporal dependencies in sequential facial data. Liu et al. (2018) proposed a hybrid CNN-LSTM framework for FER, where CNNs extract spatial features from facial images, and LSTM networks capture temporal information from sequential facial data, leading to improved emotion recognition accuracy.

3. Object Detection for Facial Landmark Localization: Accurate localization of facial landmarks, such as eyes, nose, and mouth, is crucial for precise FER. Object detection algorithms, such as You Only Look Once (YOLO), have been utilized for facial landmark localization and feature extraction. Liu et al. (2020) proposed a YOLO-based facial landmark detection method combined with CNN features for robust facial expression recognition, achieving competitive performance on challenging datasets.

4. Transfer Learning and Data Augmentation Techniques: Transfer learning techniques, where pre-trained CNN models are fine-tuned on FER datasets, have been widely adopted to overcome limitations of limited training data and improve generalization performance. Additionally, data augmentation strategies, such as geometric transformations and facial expression synthesis, have been employed to augment training datasets and enhance model robustness (Lopes et al., 2017).

5. Challenges and Future Directions: Despite the significant progress in FER, several challenges remain, including handling occlusions, pose variations, and subtle facial expressions. Future research directions may focus on developing more robust and interpretable deep learning models, exploring multimodal fusion approaches combining facial images with other modalities (e.g., audio, text), and addressing ethical considerations related to privacy and bias in FER systems.

ISSN 2347–3657
Volume 12, Issue 2, 2024

III.EXISTING SYSTEM

try to discuss about face emotion classification and recognition in a real-time manner by using a deep learning model. The authors of this study attempt to extract the essential aspects of a face using deep learning, Haar cascade, and the VGG 16 model in order to develop a classification and identification system. The authors convincingly demonstrate that the network architecture created for this paper has more improvements than previous techniques based on the experimental results. When compared to various existing models utilized in the literature for facial expression recognition, the suggested deep learning models show a significant improvement.

IV.PROPOSED SYSTEM :

proposed an automated facial expression recognition system using neural network classifiers.They employed the Rough Contour Estimation Routine (RCER) technique to extract characteristics from a human face such as eyebrows, eyes, and mouth using the Point Contour Detection Method (PCDM) to increase and detect eye and mouth precision. The author of this proposed research tries to discover a revolutionary method called Action Units (AU) that allows us to see the basic movements of the face muscles.

V.CONCLUSION

In conclusion, the review of literature highlights the significant advancements made in the field of facial expression recognition (FER) using deep learning models and object detection algorithms. The integration of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and You Only Look Once (YOLO) object detection techniques has enabled researchers to develop robust and efficient FER systems capable of accurately recognizing human emotions from facial images. These deep learning-based approaches offer superior performance compared to traditional methods by automatically learning discriminative features directly from raw image data and capturing temporal dynamics in facial expressions.

Moreover, transfer learning techniques and data augmentation strategies have been instrumental in addressing challenges related to limited training data and improving model generalization performance. Despite the progress made, several challenges, such as handling occlusions, pose variations, and subtle facial expressions, remain to be addressed. Future research directions may involve exploring multimodal fusion approaches, incorporating

ISSN 2347–3657
Volume 12, Issue 2, 2024

additional modalities such as audio and text, and addressing ethical considerations related to privacy and bias in FER systems.

Overall, the reviewed literature underscores the potential of deep learning models and object detection algorithms in advancing FER capabilities, with implications for various applications, including human-computer interaction, affective computing, and psychological research.

VI.REFERENCES

1. Zhang, Z., Cao, Q., Shi, Z., et al. (2018). Joint face representation adaptation and identity preserving. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(11), 2702–2716.

2. Liu, W., Luo, W., Xie, X., et al. (2018). Facial expression recognition using a hybrid convolutional neural network. IEEE Transactions on Affective Computing, 9(1), 32–43.

3. Liu, Y., Du, S., Liu, J., et al. (2020). YOLO-based facial landmark detection for emotion recognition. Pattern Recognition Letters, 131, 128-135.

4. Lopes, A. T., de Lima Neto, F. B., & da Silva Torres, R. (2017). Data augmentation for facial expression recognition with convolutional neural networks. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 619-627).

5. Zhao, Z. Q., Zheng, P., Xu, S. T., et al. (2019). Object detection with deep learning: A review. IEEE Transactions on Neural Networks and Learning Systems, 30(11), 3212–3232.

6. Wang, X., & Gupta, A. (2018). Video object detection with an aligned spatial-temporal memory. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 6541-6549).

7. Song, S., Lan, C., Xing, J., et al. (2018). An end-to-end spatio-temporal attention model for human action recognition from skeleton data. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI) (Vol. 32, No. 1).

8. Li, J., Liang, W., Wei, Y., et al. (2019). Facial expression recognition using a deep learning approach based on convolutional neural networks. Journal of Visual Communication and Image Representation, 59, 114-119.

9. Zhu, X., Lei, Z., Liu, X., et al. (2017). Face alignment across large poses: A 3D solution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 146-155).

10. Zhang, X., Sugano, Y., Fritz, M., et al. (2018). S3FD: Single shot scale-invariant face detector. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 384-399).

11. Patacchiola, M., & Cangelosi, A. (2017). Head pose estimation in the wild using convolutional neural networks and adaptive gradient methods. Pattern Recognition Letters, 94, 220-226.

12. Wang, Y., Wang, L., & Zeng, W. (2019). Deep facial expression recognition: A survey. Neurocomputing, 324, 28-48.

13. Huang, G., Liu, Z., Van Der Maaten, L., et al. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4700-4708).

14. Zhou, H., Alvarez, J. M., & Xiao, J. (2019). Understanding image representations by measuring their equivariance and equivalence. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 10249-10258).

15. Liu, L., Wu, C., Tao, D., et al. (2018). Deep learning for generic object detection: A survey. International Journal of Computer Vision, 128(2), 261-318.

16. Oquab, M., Bottou, L., Laptev, I., et al. (2015). Is object localization for free? – Weakly-supervised learning with convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 685-694).

17. Li, Z., Zhang, Z., Huang, L., et al. (2020). A survey of deep learning-based face recognition. arXiv preprint arXiv:2001.08079.

18. Jiang, H., Learned-Miller, E., & Samaras, D. (2017). Face detection with the faster R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 2613-2621).

19. Ren, S., He, K., Girshick, R., et al. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6), 1137-1149.

20. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.