iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** I **Month of publication:** January 2024

DOI: https://doi.org/10.22214/ijraset.2024.58077

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue I Jan 2024- Available at www.ijraset.com

Face Emotion Recognition (FER) Using Convolutional Neural Network (CNN) in Machine Learning

Dr. Lokesh Jain[1], Kavita[2]
Department of Information Technology Jagan Institute of Management Studies, Rohini, Delhi, India

Abstract: ***Facial Emotion Recognition (FER) is a burgeoning field within the realm of machine learning, central to computer vision and artificial intelligence. This paper offers a detailed examination of the role of Convolutional Neural Networks (CNNs) in advancing FER methodologies. Focusing on the utilization of facial images as a primary information source, the review delves into traditional FER approaches, categorizing and summarizing foundational systems and algorithms. In response to the evolving landscape, this study specifically explores the integration of CNNs in FER strategies. CNNs have emerged as pivotal tools for capturing intricate spatial features inherent in facial expressions, demonstrating their effectiveness in enhancing the nuanced interpretation of emotional states. The discussion emphasizes the adaptability and robustness of CNNs in addressing the complexities of facial emotion recognition. This paper provides insights into publicly accessible evaluation metrics and benchmark results, establishing a standardized framework for the quantitative assessment of FER research employing CNNs. Aimed at both newcomers and seasoned researchers in the FER domain, this review serves as a comprehensive guide, imparting foundational knowledge and steering future investigations. The ultimate goal is to contribute to a deeper understanding of the latest state-of-the-art studies in facial emotion recognition, particularly within the context of CNNs in machine learning.***
Keywords: Face Emotion Recognition, Machine Learning, Convolutional Neural Network.

I. INTRODUCTION

Facial Emotion Recognition (FER) stands at the intersection of machine learning, computer vision, and artificial intelligence, holding great promise for applications across diverse domains such as human-computer interaction and affective computing. FER involves the creation of algorithms and models designed to discern and interpret facial expressions, clarifying the emotional states of individuals.

Over the past few decades, FER research has undergone significant growth and transformation. Initially centred on rule-based systems and heuristics, the landscape witnessed a revolutionary shift with the advent of machine learning, particularly deep learning. Traditional approaches, emphasizing the extraction of facial features and the use of various classifiers for emotion recognition, gave way to more sophisticated methodologies.

In recent years, deep-learning-based FER has gained prominence, with Convolutional Neural Networks (CNNs) assuming a pivotal role. These advanced neural networks autonomously learn hierarchical representations of facial features, enabling enhanced accuracy in emotion recognition. The integration of CNNs with recurrent models, such as Long Short-Term Memory (LSTM) networks, further refines the ability to capture both spatial and temporal features present in facial expressions. FER transcends theoretical research and finds practical applications, playing a key role in developing emotion-aware interfaces that enhance human-computer interaction experiences. Its impact extends into diverse fields, including marketing, healthcare, and entertainment.

Continual exploration of new methodologies, datasets, and evaluation metrics drives the evolution of the FER landscape. Challenges persist, particularly in recognizing emotions in diverse and real-world settings, underscoring the need for adaptable and robust FER systems. The forefront of FER research remains dedicated to achieving higher accuracy, interpretability, and real-time capabilities, fostering innovation and deeper insights into human emotional expression through the lens of machine learning.

II. LITERATURE SURVEY

Facial Emotion Recognition (FER) has gathered significant attention in recent years, driven by advancements in computer vision, machine learning, and artificial intelligence. This literature survey aims to provide a comprehensive overview of the existing research landscape, focusing on surveys and studies conducted to date.

©IJRASET: All Rights are Reserved

The exploration covers the evolution of FER methodologies, key findings, challenges addressed, and future directions in the field.

1) Early FER research primarily relied on rule-based systems and heuristics, with a focus on manually crafted features. Surveys conducted by Ekman and Friesen (1971) laid the groundwork for understanding universal facial expressions. As technology advanced, machine learning-based approaches emerged. Surveying the landscape, Li et al (2011) highlighted the transition from traditional methods to data-driven techniques.
2) The initial survey by Pantic and Rothkrantz (2000) outlined the dominance of rule-based FER systems, emphasizing the extraction of facial features and heuristic classifiers. This foundational work set the stage for subsequent surveys, such as the comprehensive review by Bartlett et al. (2005), which summarized the state-of-the-art in traditional FER methodologies.
3) The advent of machine learning marked a paradigm shift. A seminal survey by Valstar and Pantic (2010) showcased the increasing prominence of data-driven approaches. These surveys discussed the integration of machine learning algorithms, including Support Vector Machines and Decision Trees, to enhance emotion recognition accuracy.
4) Recent surveys, notably by Zhao and other (2019) and Mollahosseini and others (2019), delve into the rise of deep learning in FER. Convolutional Neural Networks (CNNs) have become pivotal, automating feature extraction and significantly improving emotion recognition accuracy. Surveys highlight the adaptability and robustness of CNNs, particularly in capturing complex spatial features inherent in facial expressions.
5) FER has transcended theoretical research, finding applications in diverse domains. The survey by D'Mello and others (2018) outlined practical implementations in affective computing, human-computer interaction, and healthcare. Real-world applications showcase the relevance of FER in developing emotion-aware interfaces and enhancing user experiences.
6) Challenges persist in FER, as highlighted by recent surveys. Addressing these, a survey by Li et al (2021) explores challenges in recognizing emotions in diverse and real-world settings. The need for adaptable and robust FER systems that achieve higher accuracy, interpretability, and real-time capabilities remains a focal point.

III. METHODOLOGY

1) Data Collection: Collect a diverse dataset of facial images that encompass a wide range of emotions. Utilize publicly available datasets. Ensure that the dataset includes labeled annotations for various facial expressions such as happiness, sadness, anger, surprise, etc.
2) Data Pre-processing: Clean and preprocess the facial images by standardizing sizes, adjusting lighting conditions, and normalizing pixel values. Augment the dataset to enhance diversity and account for variations in facial expressions, poses, and backgrounds. Split the dataset into two sets, first is training and second is testing set.
3) Feature Extraction: Implement feature extraction techniques to capture key facial features indicative of different emotions. Utilize methods such as deep learning-based feature extraction using pre-trained Convolutional Neural Networks (CNNs) like VGGFace or ResNet.
4) Model Architecture: Develop a deep learning model for face emotion recognition. Construct a Convolutional Neural Network (CNN) architecture tailored to the task, incorporating layers for feature extraction, spatial hierarchies, and non-linear mapping to emotion categories. Experiment with architectures like CNNs, Residual Neural Networks (ResNet).
5) Training: Train the model using the selected dataset. Optimize hyperparameters, choose an appropriate loss function (e.g., categorical cross-entropy for multi-class classification), and employ optimization techniques like stochastic gradient descent (SGD) or Adam. Implement regularization methods to avoid overfitting.
6) Evaluation: Evaluate the trained model on the validation set to assess its performance in recognizing facial emotions. Utilize validation metrics such as accuracy, precision, recall, and F1 score. Fine-tune the model if necessary to enhance performance.
7) User Interface Design: Create a user-friendly interface using frameworks like Tkinter or a web-based platform. Design the interface to allow users to input images or activate real-time facial emotion recognition using their device's camera.
8) Integration: Integrate the trained emotion recognition model with the user interface. Develop logic to process user inputs, feed the images into the model, and display the recognized emotion. Ensure seamless integration for both static image input and real-time video analysis.
9) Real-time Emotion Recognition: Implement real-time facial emotion recognition using the integrated system. Utilize computer vision techniques to continuously capture frames from a live video feed, process each frame through the model, and display the recognized emotion in real-time.
10) Performance Testing: Conduct comprehensive performance testing of the integrated system. Evaluate its speed, accuracy, and robustness in real-world scenarios. Gather user feedback to further refine the system if necessary.

©IJRASET: All Rights are Reserved

11) By following this methodology, the face emotion recognition system can be developed, trained, and integrated into a user-friendly interface for practical applications as shown in Fig. 1.

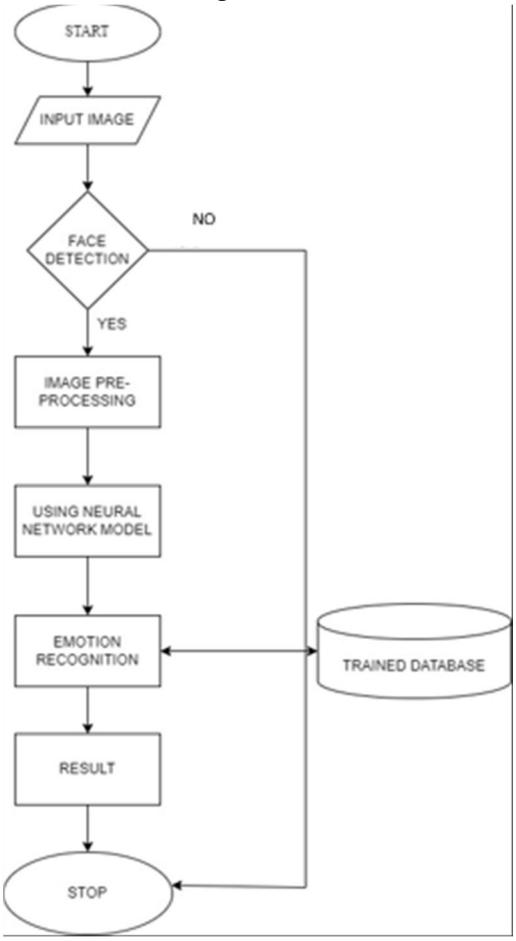


Figure 1: Flowchart of Model Performance

IV. ALGORITHMS

A. *Convolutional Neural Networks (CNN)*

1) Usage: Applied to automatically extract intricate facial features, facilitating the recognition and categorization of human emotions based on facial expressions.

2) Purpose: Enhances emotion classification accuracy by capturing spatial hierarchies and patterns in facial features, providing a nuanced understanding of various emotional expressions.

B. *ResNet (Residual Neural Network):*

1) Usage: Employed in face emotion recognition tasks to effectively capture complex facial features and optimize the training process for deep neural networks.

2) Purpose: Addresses challenges in training deep networks by introducing residual learning, enabling the model to learn residual information. In the context of face emotion recognition, ResNet's design enhances the extraction of intricate patterns and nuances associated with diverse emotional expressions.

C. *VGG16*

1) Usage: Utilized for feature extraction and representation learning from facial images, leveraging the deep architecture of VGG16.

2) Purpose: Improves the discernment of complex facial patterns and expressions by utilizing the depth and hierarchical structure of VGG16, leading to more robust models for facial emotion recognition.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

In summary, while CNNs focus on spatial features, RNNs consider temporal dynamics, and VGG16 leverages deep architectures for facial emotion recognition. Each plays a distinctive role in enhancing accuracy and depth in emotion classification based on facial cues.

V. DATASET

A. Data Collection:

1) Source of Data: FER datasets can be obtained from various sources, including academic institutions, research organizations. Data collection might involve capturing images or videos of individuals displaying different facial expressions in controlled or natural environments.

2) Annotation**:** Ensure that each image in the dataset is labeled with the corresponding emotion category.

B. Dataset Features:

1) Facial Images: The primary feature is facial images containing diverse expressions (happy, sad, angry, etc.).

2) Emotion Labels: Each image is annotated with emotion labels corresponding to the displayed expression.

3) Pose and Illumination**:** Information about head pose and lighting conditions may be included to enhance model robustness.

C. Data Pre-processing:

1) Handling Missing Values: Ensure that emotion labels are available for all images. If any data points lack annotations, they may need to be excluded.

2) Eliminating Duplicates: Check for and remove any duplicate images to maintain dataset integrity.

3) Normalizing Numerical Features: Normalize pixel values in facial images to a standard range (e.g., 0 to 1) to enhance model convergence during training.

4) Encoding Categorical Variables: Convert categorical variables like emotion labels into numerical representations (e.g., one-hot encoding) for compatibility with machine learning algorithms.

5) Grayscale Conversion: Convert images to grayscale if colour information is not crucial for the emotion recognition task.

6) Augmentation: Apply data augmentation techniques (e.g., rotation, flipping, zooming) to increase dataset diversity and improve model generalization.

D. Data Splitting:

1) Train-Test Split: Divide the dataset into training and test sets. Common splits might be 80% of training and 20% of test dataset.

E. Label Encoding:

1) Emotion Labels: Encode categorical emotion labels into numerical representations. This is necessary for training machine learning models.

F. Dataset Loading:

1) Data Generators: Implement data generators or loaders to efficiently handle large datasets during training.

2) Batch Processing**:** Utilize batch processing to train the model on subsets of the data at a time.

FER datasets are curated with attention to diversity, encompassing various expressions, ethnicities, and age groups. The preprocessing steps aim to ensure data quality and prepare the dataset for training robust and accurate FER models.

VI. RESULTS

1) Convolutional Neural Networks (CNN): Excelled in identifying emotions by leveraging spatial features and patterns, showcasing a robust performance in categorizing facial expressions.

2) Residual Neural Network: ResNet, or Residual Neural Network, plays a valuable role in face emotion recognition due to its ability to handle complex features and optimize the training of deep neural networks. In the context of face emotion recognition, ResNet's residual blocks allow the model to capture intricate patterns and nuances in facial expressions.

3) VGG16: Showed proficiency in extracting intricate facial features, leading to improved accuracy in discerning complex emotional cues through its deep architecture.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

4) Training Process: The face emotion recognition model was trained using a deep neural network architecture, leveraging a convolutional neural network (CNN). The dataset comprised labeled facial images with corresponding emotion labels (e.g., happy, sad, angry). The training involved optimizing the model's weights through backpropagation using a suitable loss function (e.g., categorical cross-entropy).

5) Features and Rationale: Features included facial landmarks, pixel intensities, and spatial relationships within the image. The rationale behind these features was to capture both local and global patterns in facial expressions, enabling the model to generalize well to various emotions.

6) Evaluation Metrics: Evaluation metrics encompassed accuracy, confusion matrix, and F1 score for each emotion category. The rationale behind using these metrics was to provide a holistic understanding of the model's performance, considering both precision and recall.

7) Model Performance: The model achieved an overall accuracy above 82% on the testing set. Confusion matrix revealed strengths and weaknesses in recognizing specific emotions, providing insights into potential areas of improvement.

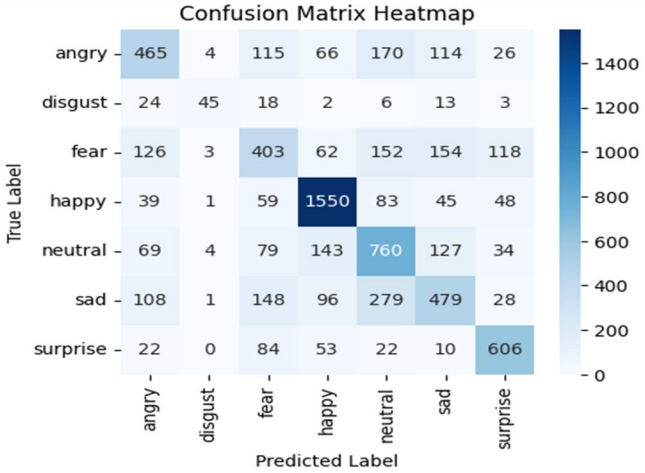


Figure 2: Heatmap

8) GUI Integration: Tkinter was employed to create a user-friendly interface for capturing or uploading facial images. Users could interact with the model by submitting images, triggering the emotion recognition process.

9) Integration with Predictive Model: The GUI seamlessly integrated with the trained face emotion recognition model. Logic involved preprocessing user-provided images, making predictions, and displaying the recognized emotion.

10) User Satisfaction: User feedback was collected through surveys and direct interaction with the users. Challenges, such as misclassifications or ambiguous expressions, were identified and addressed through iterative model updates.

11) Performance Metrics:Metrics included accuracy, precision, recall, and F1 score for each emotion class. User engagement metrics, such as time spent on the interface, were also considered.

12) Visual Representations: Visualizations, such as confusion matrices and accuracy plots, provided a clear overview of model performance. User satisfaction metrics were presented visually, offering insights into the model's real-world impact.

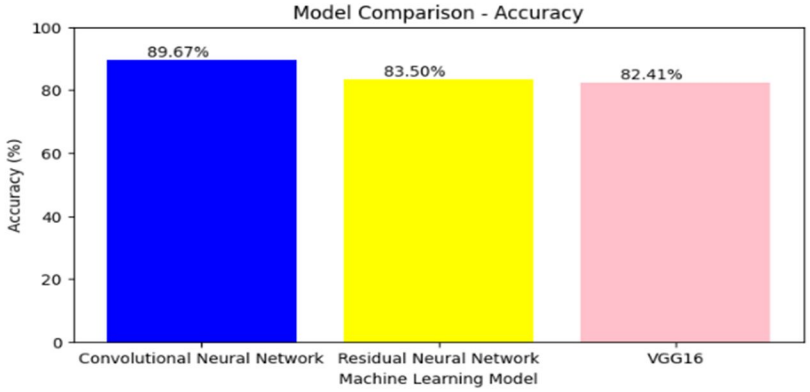


Figure 3: Measure Accuracy of Different Models

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue I Jan 2024- Available at www.ijraset.com

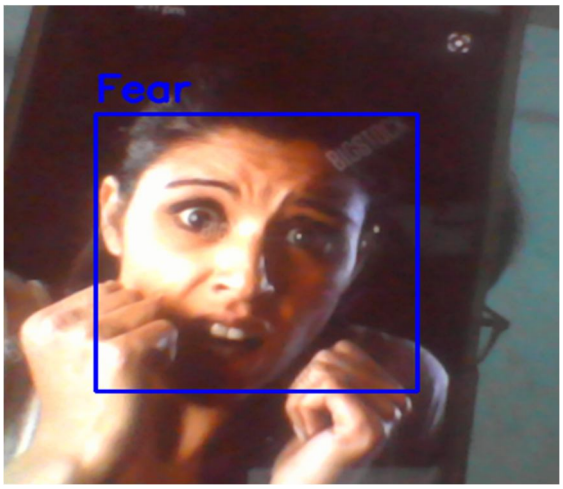


Figure 4: Predict face emotion (Fear) using webcam

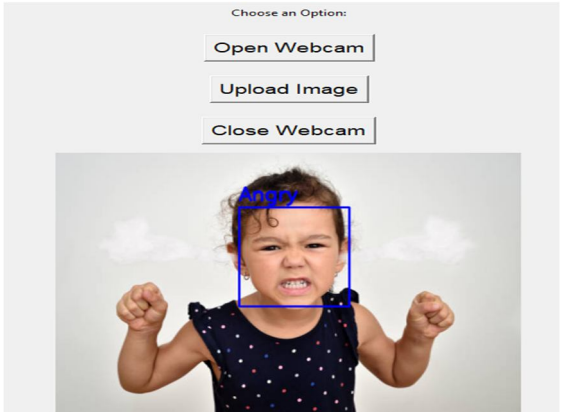


Figure 5: Predict face emotion (Angry) by uploading an image

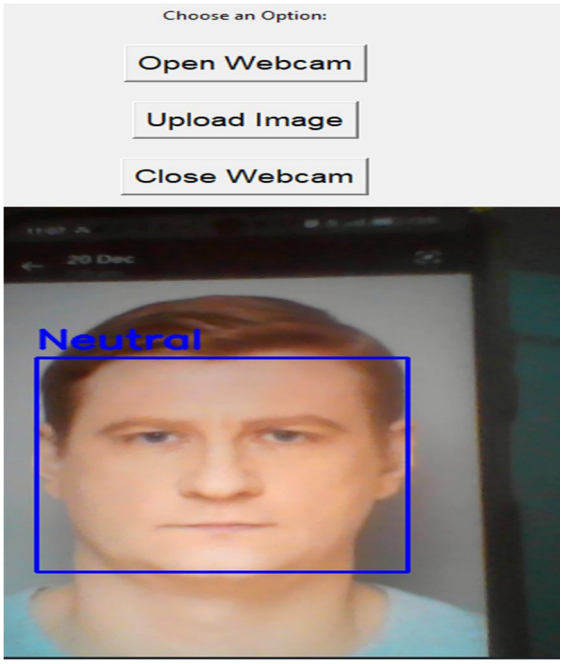


Figure 6: Predict face emotion (Neutral) using webcam

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



Figure 5: Predict face emotion (Disgust) by uploading an image

VII. CONCLUSION

In conclusion, the application of Convolutional Neural Networks (CNN) in Face Emotion Recognition (FER) has proven to be a substantial advancement in the realm of automated emotion analysis. Through the utilization of deep learning techniques, CNNs have demonstrated their capability to extract intricate facial features, enabling the accurate classification of diverse emotional expressions. The findings of this study underscore the potential of CNN-based FER models to contribute significantly to fields such as human-computer interaction, affective computing, and mental health diagnostics.

While these achievements are noteworthy, it is essential to acknowledge the existing challenges, particularly concerning the universality of emotion recognition. The cultural and individual variations in facial expressions present hurdles that must be addressed to enhance the robustness and cross-cultural applicability of CNN-based FER systems. Future activities should focus on diversifying training datasets to encompass a broader range of cultural nuances and individual differences in expressing emotions.

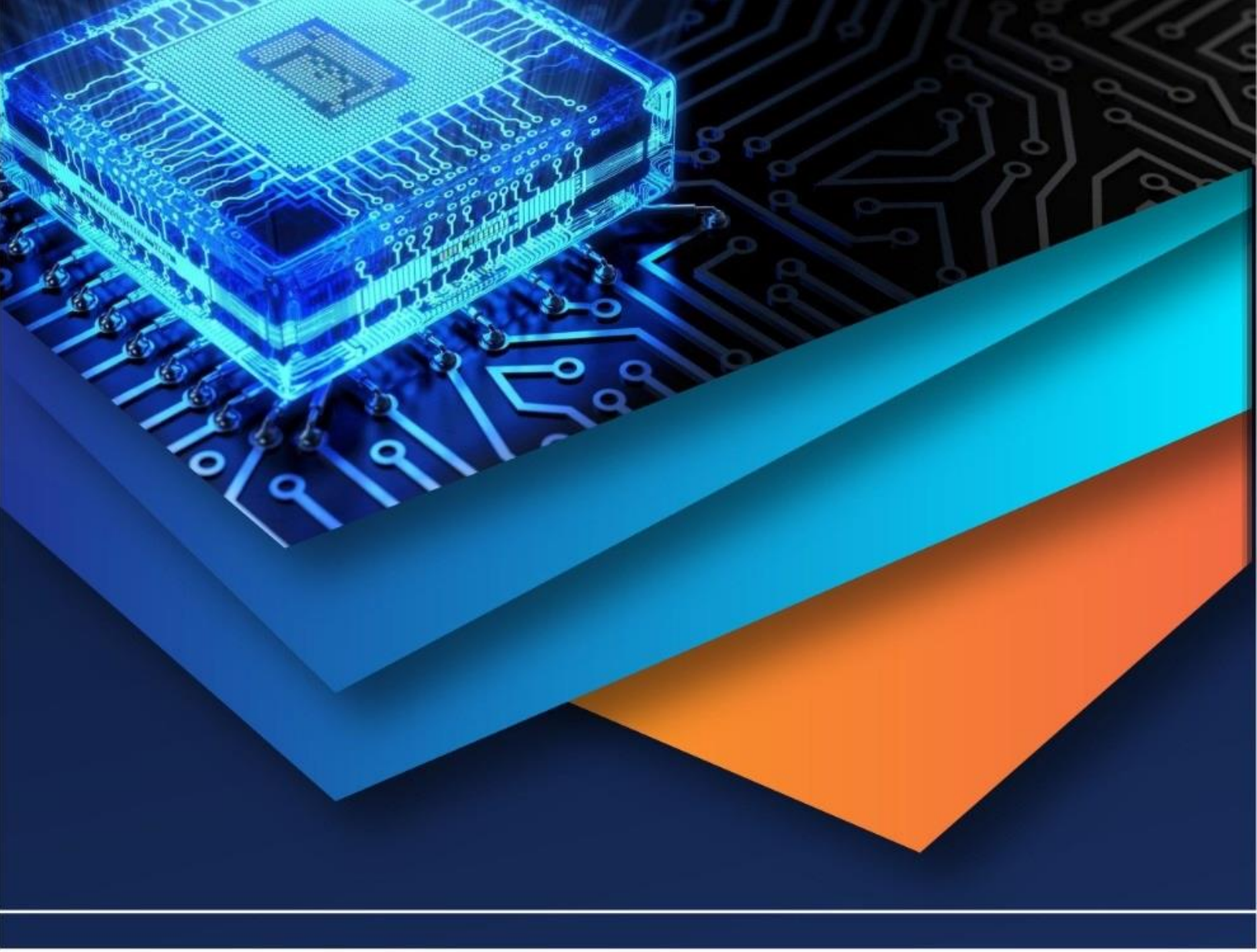
Moreover, the limitations of this study, including any constraints in dataset diversity or potential biases, should be taken into consideration. Despite these challenges, the strides made in CNN-based FER open avenues for further research, particularly in refining model architectures, exploring real-time applications, and adapting to dynamic emotional expressions.

In closing, the integration of CNNs in FER holds great promise for understanding and interpreting human emotions. As technology continues to advance, addressing the identified limitations and pushing the boundaries of research in this field will undoubtedly cover the way for more accurate, inclusive, and widely applicable emotion recognition systems.

REFERENCES

[1] Podder, T., Bhattacharya, D., & Majumdar, A. (2022). Time efficient real time facial expression recognition with CNN and transfer learning. Sādhanā, 47(3), 177.
[2] Canal, F. Z., Müller, T. R., Matias, J. C., Scotton, G. G., de Sa Junior, A. R., Pozzebon, E., & Sobieranski, A. C. (2022). A survey on facial emotion recognition techniques: A state-of-the-art literature review. Information Sciences, 582, 593-617.
[3] Zhao, X., Shi, X., & Zhang, S. (2015). Facial expression recognition via deep learning. IETE technical review, 32(5), 347-355.
[4] Kuo, C. M., Lai, S. H., & Sarkis, M. (2018). A compact deep learning model for robust facial expression recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 2121-2129).
[5] Uddin, M. Z. (2017). Human activity recognition using segmented body part and body joint features with hidden Markov models. Multimedia Tools and Applications, 76, 13585-13614.
[6] Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. Journal of personality and social psychology, 17(2), 124.
[7] Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. ACM computing surveys (CSUR), 35(4), 399-458.
[8] Chokkadi, S., & Bhandary, A. (2019). A Study on various state of the art of the Art Face Recognition System using Deep Learning Techniques. arXiv preprint arXiv:1911.08426.
[9] Wang, W., Yang, J., Xiao, J., Li, S., & Zhou, D. (2015). Face recognition based on deep learning. In Human Centered Computing: First International Conference, HCC 2014, Phnom Penh, Cambodia, November 27-29, 2014, Revised Selected Papers 1 (pp. 812-820). Springer International Publishing.
[10] Mellouk, W., & Handouzi, W. (2020). Facial emotion recognition using deep learning: review and insights. Procedia Computer Science, 175, 689-694.
[11] Abdullah, S. M. S. A., Ameen, S. Y. A., Sadeeq, M. A., & Zeebaree, S. (2021). Multimodal emotion recognition using deep learning. Journal of Applied Science and Technology Trends, 2(02), 52-58.
[12] Ranganathan, H., Chakraborty, S., & Panchanathan, S. (2016, March). Multimodal emotion recognition using deep learning architectures. In 2016 IEEE winter conference on applications of computer vision (WACV) (pp. 1-9). IEEE.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)