International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

CNN-Driven Enhancement in Facial Emotion Recognition Systems

V. Ajitha [1], T. S. Suganya *[2], S. Irin Sherly [3], P. Jose [4], C. Gnanaprakasam [5], M. Krishnaraj [6]

Submitted:11/03/2024 **Revised**: 26/04/2024 **Accepted**: 03/05/2024

Abstract: Facial Emotion Recognition is crucial for developing systems used in human-computer interaction, mental health diagnostics, and security applications. Accurately detecting and classifying human emotions from facial expressions is essential in these areas. This study aims to create a robust and high-accuracy model for recognizing and classifying human emotions from facial expressions using convolutional neural networks (CNNs). It emphasizes the use of transfer learning and data augmentation to improve the performance of the model on FER2013, which contains 35,887 black and white images representing seven emotions (surprise, happiness, sadness, disgust, anger, fear, and neutral). Each image is resized to 48x48 pixels to reduce computational complexity. The proposed model uses a pre-trained VGG16 CNN, fine-tuned with the FER2013 data. Preprocessing steps include resizing images to 224x224 pixels, normalization, and applying data augmentation techniques like rotation, scaling, and shifting to improve model robustness. The CNN analyzes features present in the input images and then runs them through fully connected layers to produce a distribution of probabilities across seven emotional categories using a softmax activation function. The algorithm is trained using the categorical cross-entropy loss function and the Adam optimizer, initializing with a learning rate of 0.001. This rate decreases tenfold every ten epochs. The batch size is 32, and training lasts for 50 epochs with early stopping. The use of transfer learning and data augmentation significantly improved the model's accuracy, achieving 83% on the test set. Performance measures like accuracy, completeness, and the F1-score were computed for each emotion category with the 'happy' class showing the highest precision (0.92) and recall (0.93). The confusion matrix revealed a high true positive rate and identified areas of misclassification, particularly between similar emotions like sadness and neutral. The training and validation loss curves indicated a good fit without overfitting, as shown by the close convergence of these values. Throughout the training period, the level of precision consistently rose, leveling off at approximately 83% for both the training and validation datasets, indicating a steady improvement in accuracy. The model shows potential for applications in mental health monitoring, customer feedback analysis, and personalized content delivery. Future work will focus on refining the model and applying it to real-world scenarios to enhance its robustness and accuracy in diverse environments.

Keywords: *Facial Emotion Recognition, Convolutional Neural Networks, Data augmentation, Transfer learning, FER2013 dataset.*

1. Introduction

Facial emotion recognition can be approached as a classification problem, aiming to label an input image according to the expressed emotion. To achieve this, we utilize a Convolutional Neural Network (CNN) architecture. This architecture is specifically structured to extract features from the input images and establish a connection between the input and output realms. It comprises various convolutional layers that capture spatial and spectral features from the images, alongside pooling layers that reduce spatial dimensions while preserving essential information. Subsequently, fully connected layers process the feature maps to generate a probability distribution covering seven emotions. Through training on a vast dataset of labelled images, the CNN can discern distinctive emotional features and apply this understanding effectively to new, unobserved images. By using a large dataset of labelled images, the CNN can learn to recognize features that are indicative of different emotions and generalize well to new, unseen images.

2. Related Works

In order to separate positive labels which, contain discriminative facial emotions like a smile from the other labels, Fujii et al., [1] first perform binary classification based on facial expression recognition. In certain cases, it takes human interaction to determine the subject's mental state. To automate this process, Ranjan et al., [2] investigate the use of the Fisher face method in conjunction with supervised classification techniques like artificial neural networks and K-nearest neighbor to recognize human emotions from still images. This approach results in an accurate, dependable, and automatic system. The effectiveness of the Fisher face model in

[1] Department of Artificial Intelligence and Machine Learning, Nalla Malla Reddy Engineering College, Hyderabad, India
ORCID ID: 0000-0001-5934-8712
[2] Department of Computer Science and Applications, Faculty of Science and Humanities, SRM Institute of Science and Technology, Ramapuram, Tamil Nadu, 600089, India.
ORCID ID: 0000-0003-1685-2032
[3] Department of CSE, SIMATS School of Engineering, Saveetha Institute of Medical and Technical Sciences, Thandalam, Tamil Nadu, 602105, India
ORCID ID: 0000-0002-4518-3213
[4] Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R &D Institute of Science and Technology, Avadi, Chennai, India
ORCID ID: 0000-0003-2340-6786
[5] Department of Artificial Intelligence and Data Science, Panimalar Engineering College, Chennai, Tamil Nadu, India
ORCID ID: 0000-0003-1401-9418
[6] Department of CSE, St.Joseph's Institute of Technology, OMR, Chennai, India
ORCID ID: 0000-0003-4065-2874
** Corresponding Author Email: tssuganya07@gmail.com*

identifying human mood based on information gleaned from facial photos has been investigated by Ranjan et al.,. A fuzzy-based method for extracting facial features for human emotion classification is presented by Sharma et al. [3] 136 photos overall from the JAFFE, NimStim, and MUG datasets for the joyful and neutral emotion classes were used to create the feature set. Using techniques from Naïve's Bayesian classification, support vector machines, random forests, deep learning, and linear discriminant analysis, classification models were built for happy and neutral emotion classes. The interpretation of the emotion landscape, or the distribution of emotion classes found in still photos or films with lots of people's faces visible, is developed for the second approach Tribedi et al., [4]. Though there is a clear distinction in the potential uses, this is comparable to group-level emotion classification. The distributed computer system's processors executed the threads, which classed face expressions based on their emotion identification. Using a distributed computer system, Korkmaz et al., [5] investigate the classification of human face expressions for emotion recognition. The categorization methodology made use of facial expression recognition algorithms and emotion analysis tools. In their work, Awale et al. [6] analyze face EMG data to classify emotions using the spectral kurtogram and CNN. The kurtogram is regarded as an input vector that CNN uses to classify emotions. Utilizing a two sensor wireless data gathering device, facial EMGs for five distinct facial moods were captured. The publicly accessible datasets used to train CNNs for facial emotion classification have issues with inter-class imbalance and noisy labeling. A Generative Adversarial Network by Kim et al., [7] to address issues with noisy labeling and inter-class imbalance. Local binary descriptors are often used as feature descriptors in face recognition software systems. Aryaet.al., [8] discuss the evaluation of various binary descriptors utilized in emotion classification, such as support vector machines (SVM), K-nearest neighbor (KNN), back propagation neural networks (BPNN), angular local directional patterns (ALDP), local binary patterns (LBP), local directional patterns (LDP), local directional number patterns (LDNP), and angular local directional patterns (LOOP). Abbassi et al., [9] provide a novel method based on the VGGNet-19 network. This method makes use of a dropout strategy and many convolution layers with tiny filters. Convolution layers are advised to be added to the selected model in order to improve image classification accuracy. Agastya et al., [10] suggest using a multi-label classification approach to translate Indonesian sentences with single and multiple structures into emotion classes. Multi-label text categorization with continuous values between 0 and 1 is the technique utilized for the emotion mapping procedure. The facial area is depicted as a network of nodes and edges, using the gSpan algorithm to identify common sub-structures in the emotional graph database. This novel approach to recognizing facial emotions through graph analysis aims to revolutionize how we represent the face region (Hassan et al.,). Six stages of binary classification are used to categorize the mood of the requested input facial image once the final sub-graphs have been encoded. Karbauskaite et al., [11] create a FER method that is completely distinct from this popular strategy. It has been confirmed that the kriging predictor is suitable for identifying facial expressions in small sets of images. Emotion is mostly expressed through facial expressions, and numerous research have focused on using facial expressions to categorize emotions. A neural network was employed to increase the classification accuracy, while the histogram of optical flow (HOF) was utilized to extract features Ragupathy et al., [12]. Convolutional neural networks, or CNNs, have been the subject of numerous studies lately on the identification of face emotions. A lightweight CNN architecture designed for embedded systems is presented by Kim et al., [13]. Fujii et al. [14] present a novel technique for identifying emotions at the group level. Through the identification of distinct areas or vertices accountable for the categorization of a given facial expression, Gund et al., [15] acquire enhanced comprehension of the facial dynamics during the manifestation of that sentiment. By doing this, Gund et al., offer a method that may be easily understood to comprehend the dynamics involved in convolutional-based prediction tasks on facial data sequences. Fnaiech et al., [16] present a sophisticated technique that, by quantitatively evaluating other negative emotions, may accurately identify fear-related facial emotions. In contrast to alternative cutting-edge and classification techniques, Fnaiech et al., documented the greatest accuracy of recognized dread emotion. The system's performance is negatively impacted by the restricted number of characteristics from facial photos, which also increases the degree of classification mistake. In their study conducted by Poulose et. al., [17] explored the integration of facial landmarks within the feature vector extraction process for identifying facial expressions. This approach resulted in an impressive 99.96% accuracy in classification and a model loss of 0.095 using the ResNet architecture for facial emotion recognition. In order to obtain an effective classification performance, in 2023, Wang introduce a sophisticated deep learning system for detecting emotions across various modalities by combining facial expressions and electroencephalogram (EEG) inputs. They utilize a pre-trained convolutional neural network (CNN) to extract facial characteristics from the facial expressions. Wang et al., [18] compared the six facial emotions to the most advanced classification techniques. According to empirical findings, the model performs more accurately than earlier techniques on two Facial Emotion Recognition (FER) benchmarks and the combined FER GiMeFive [19].

Kavitha et.al., [20] presented an intelligent flying robot for smart agriculture, using digital image processing and the VGG-SVM model to detect plant leaf diseases. Equipped with a high-resolution camera, the robot captures and analyzes real-time images, ensuring early disease detection and targeted interventions. This system optimized crop yield and resource use, achieving a high accuracy rate of 0.97, promoting sustainable agriculture. The study [21] explored using blockchain technology to enhance election security and reduce costs. It demonstrated that blockchain can provide durability, secrecy, and accountability, thereby improving system stability and security. The Artificial Intelligence Assisted Hybrid Learning Scheme (AIHLS) integrated Xception's feature extraction with the XGBoost Classifier for early autism detection. This approach harnessed fine-grained image features and accurate predictions, achieving a notable 97.11% accuracy [22]. The developed system assists visually impaired individuals by identifying obstacles and people, playing audio messages, and notifying caregivers. It uses OpenCV and face recognition for real-time image processing, requiring minimal memory and processing time [23].

S. Irin Sherly et. al., [24] highlighted deep learning methods for predicting cardiovascular disease, with convolutional neural networks (CNNs) achieving 94.2% accuracy in both training and testing. A series of thorough studies were performed using four databases, incorporating 14 clinical features into an ensemble learning model. The trials showed an accuracy rate of nearly 99% across all four datasets, surpassing previous machine learning methods. These findings indicate that the new ensemble learning technique outperforms other existing approaches [25]. The authors [26] presented a process for preprocessing noisy ECG signals, which includes muscle contraction and electrode touch noise, using the Delayed Normalized Least Mean Square (DNLMS) method. The QRS complex of the ECG signal is extracted using Discrete Cosine Transform (DCT) and Fast Fourier Transform (FFT). The FRCNN model's target detection box and anchor parameters are optimized with the HBA algorithm to reduce missed detections, overfitting, and computational costs. The ECG data used in the study comes from the BIDMC Congestive Heart Failure Database and the MIT-BIH Normal Sinus Rhythm Database. The author discusses the different types of noise that impact ECG signal quality and explores their causes. It highlights the challenge of filtering these noise components and reviews various filtering methods used to remove noise from ECG data [27].

3. Methodology

3.1. Dataset Information

The Facial Expression Recognition (FER) dataset, which contains of grayscale photos of various facial emotions, was employed in this investigation. Images representing seven distinct emotions are included in the dataset: surprise, happiness, sadness, disgust, rage, and neutral. To minimize computational complexity, every image is shrunk to 48x48 pixels. This is a labeled dataset in which a label for each emotion is attached to an image. Bounding boxes around the faces and labels indicating the emotions are used to annotate the photographs.The dataset is divided into three parts: 28,709 photos are allocated for training, 3,588 images are set aside for validation, and 3,601 images are designated for testing.

3.2. Mathematical Justification

During the process of feature extraction, characteristics are derived from the input pictures utilizing a pre-trained CNN model. Suppose we assign y as the label for the image x, and x represents the input image. The CNN model can be expressed as a function f(x), where z = f(x) is the feature vector that is produced from the input image. A fully connected neural network g(z) is then trained on the feature vector z to identify one of the seven emotions in the image. The function g(z), which takes the feature vector z and outputs a probability distribution over the seven emotions, can be used to represent the neural network.

$$g(f(x)) = p(y|x)$$

The neural network is trained using the categorical cross-entropy loss function, which assesses the difference between the actual and predicted probability distributions. The purpose is to train the learning process by adjusting these distributions.

The loss function L can be represented as:

$$L = -\sum i=1^7 yilog(pi)$$

where yi represents the genuine distribution of probabilities and pi indicates the estimated distribution of probabilities.

3.3. Proposed Method

The network comprises various layers, beginning with an initial layer, followed by a preprocessing layer as in figure 1. After the preprocessing step, there are several layers denoted as "Conv2D" and "MaxPool2D." These layers illustrate the structure of the convolutional neural network, commonly applied in tasks related to image processing and recognition. The CNN design aims to extract pertinent features from the input data by conducting convolution operations and downsizing via max pooling. These layers process the input data, extracting features and reducing spatial dimensions. The arrows between the layers indicate the flow of data through the network. The input data passes through each layer sequentially, with each layer performing specific operations on the data. Sequential processing in a neural network enables the system to progressively learn and capture more intricate characteristics from the input data. The ultimate layer of this network, designated as the output layer, furnishes the likelihoods for each category. Consequently, once the input information traverses all network layers, it generates a probability distribution signifying the chances of it falling into each category. The neural network's outcome consists of a range of class probabilities that are specified under "Class Probabilities Output," encompassing eight categories: "Angry," "Disgust," "Fear," "Happy," "Sad," "Surprise," "Neutral," and "Other." This neural network has been taught to categorize input images into these groups, with the output layer presenting the likelihoods for each class.

3.4. Steps in the proposed method

3.4.1. Preprocessing

a. Resizing: The images are resized to 224 x 224 pixels to fit the input requirement of the CNN model.

b. Normalization: Pixel values are standardized to have an average of 0 and a standard deviation of 1, enhancing model convergence.

3.4.2. Convolutional Neural Network (CNN)

a. Convolutional Layers: These layers contain various filters that execute convolution operations on the input image. The image undergoes multiple convolutional layers with filters of different dimensions and movements, detecting features like edges, corners, and textures. The outcome is a collection of feature maps.

b. Activation Function: Non-linear activation functions, like ReLU (Rectified Linear Unit), are applied element-wise to the feature maps, adding complexity for learning intricate relationships between the input image and identified features.

c. Max Pooling Layers: Post each convolutional layer, a max pooling layer reduces the dimensionality of the feature maps, averting overfitting. This process selects the maximum value in a pooling window, effectively down sampling the feature maps.

d. Flatten Layer: The final pooling layer's output becomes a one-dimensional vector for transition through fully connected layers.

e. Fully Connected Layers: The flattened vector goes through fully connected layers, gradually reducing the number of neurons to learn the relationships between features and emotions. Each neuron connects to all neurons in the prior layer.

f. Output Layer: The ultimate fully connected layer comprises neurons matching the emotion categories. Softmax activation function is employed to convert raw outcomes into probabilities denoting each emotion category likelihood.

3.4.3. Training

a. Loss Function: Categorical cross-entropy measures the deviation between predicted probabilities and true labels.

b. Optimizer: Utilizing the Adam optimizer updates the model weights during training.

c. Learning Rate: Initial learning rate at 0.001 decreases by a factor of 10 every 10 epochs to hinder overfitting.

d. Batch Size: Set at 32 to balance computation time versus model accuracy.

e. Number of Epochs: Training spans 50 epochs with early stopping mechanisms to curb overfitting.

3.4.4. Evaluation

Once trained, the model can classify the emotions of unseen facial images accurately. Evaluation includes accuracy, precision, recall, F1-score for each emotion group, and a weighted average F1-score for an overall assessment.

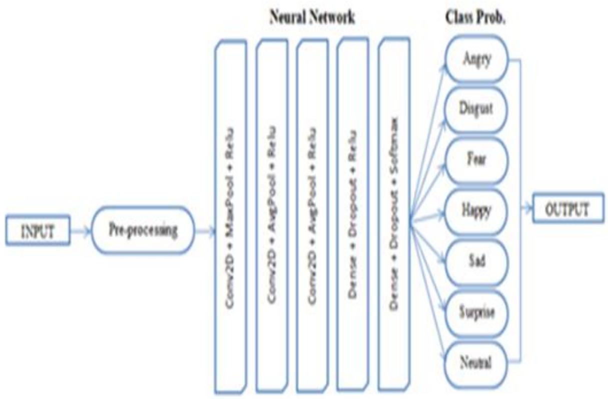

Fig.1. Architecture Diagram of prediction system

4. Results and Discussion

	precision	recall	f1-score	support
0	0.92	0.93	0.93	899
1	0.82	0.71	0.76	608
2	0.72	0.80	0.76	620
accuracy			0.83	2127
macro avg	0.82	0.81	0.82	2127
weighted avg	0.83	0.83	0.83	2127

Fig.2. Performance metrics for each emotion class

Emotions are categorized into three classes, possibly representing different emotions such as 'happy', 'sad', and 'angry'. Figure 2 shows the performance metrics for each emotion class. The accuracy of predicting class 0 stands notably high at 0.92, showing that when the model identifies an instance of this class, it is correct 92% of the time. The recall is also high at 0.93, suggesting that it successfully identifies 93% of all actual instances of this class. This class also has the highest f1-score at 0.93, which is a measure of the test's accuracy. Class 1 has a lower precision and recall, at 0.82 and 0.71 respectively, which means the model is less accurate in predicting this emotion and misses more actual instances of this emotion compared to class 0. The class 1 f1-score is 0.76, indicating the presence of lower precision and recall values. Class 2 has the lowest precision at 0.72 but a relatively high recall at 0.80, which could indicate that while the model is less precise in predicting this emotion, it is relatively good at identifying most instances of this emotion when they occur. The f1-score for class 2 is also 0.76, the same as for class 1, suggesting similar performance levels for these two classes despite the differences in precision and recall. The model demonstrates an overall accuracy rate of 0.83, meaning that across all classes, it correctly identifies the correct emotion 83% of the time. The macro average and weighted average scores are consistent with the individual class scores, indicating a balanced performance across all classes when considering both the precision and recall. The support values indicate the number of instances for each class that were used to compute these metrics, which is important for understanding the context of the model's performance. Class 0 has the highest number of instances at 899, followed by class 2 with 620 and class 1 with 608. This distribution can affect the weighted average metrics, as they take into account the number of instances in each class. In summary, the model performs best at recognizing the emotion represented by class 0 and has more challenges with the emotions represented by classes 1 and 2. The overall accuracy is quite high, but there is room for improvement, especially in increasing the precision for class 2 and the recall for class 1 to enhance the model's ability to correctly identify these emotions.

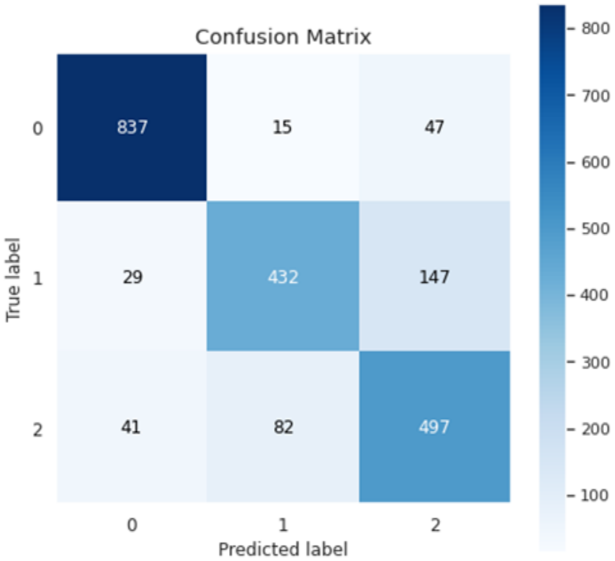

Fig.3. Confusion matrix on the test set

This matrix aids in demonstrating the model's accuracy by displaying the count of accurate and inaccurate predictions for each category, corresponding to different emotions. The main diagonal elements depict the correct positive predictions, while the elements off the diagonal represent the incorrect positive and negative predictions. The majority of the predictions for class 0 are correct, with 837 out of 899 instances correctly identified. There are relatively few instances where class 0 was confused with class 1 (15 instances) or class 2 (47 instances).

For class 1, the model correctly predicted 432 out of 608 instances. However, there is a notable number of instances where class 1 was misclassified as class 0 (29 instances) or class 2 (147 instances). Class 2 shows a good number of correct predictions at 497 out of 620 instances. Misclassifications include 41 instances where class 2 was mistaken for class 0 and 82 instances where it was mistaken for class 1.The confusion matrix complements the classification report from Fig 1 by showing that while the model has a high overall accuracy, there are specific areas where misclassification occurs, particularly between classes 1 and 2. This insight could be used to further refine the model, possibly by providing more training data for these classes or by improving the feature extraction process to better distinguish between the emotions that these classes represent.

The model converges quickly and achieves a low loss on both training and validation sets. The graph's x-axis displays the number of epochs, with the y-axis showing the loss value. Initially, both the training and validation losses decrease rapidly, indicating the model's quick learning and improvement.

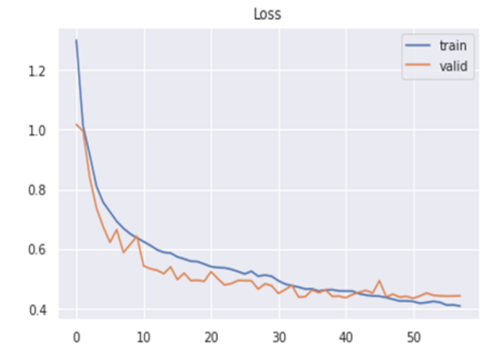

Fig. 4. Loss values of a facial emotion recognition model over a number of epochs during its training and validation phases

However, as the epochs progress, the rate of loss reduction slows down, suggesting that the model is starting to converge to a solution. The training loss continues to decrease steadily as the number of epochs grows, aligning with the expected trend as the model fits better to the training data. Similarly, the validation loss decreases alongside the training loss but begins to plateau as the epochs increase, which is a typical behavior indicating that the model is generalizing to new data that it hasn't seen before. The close similarity between the training and validation losses indicates that the model is not overfitting. This closeness signifies that the model's performance on the training set is indicative of its performance on the validation set, a positive outcome for facial emotion recognition tasks on new images.

During the initial epochs, both the training and validation accuracy of the model show rapid improvement, indicating effective learning in recognizing facial emotions. After the sharp initial increase, the accuracy of both training and validation phases begins to plateau, suggesting that the model is approaching its maximum capability given the current architecture and dataset. The training accuracy remains slightly higher than the validation accuracy throughout the training process, which is common occurrence as the model directly learning from the training data.

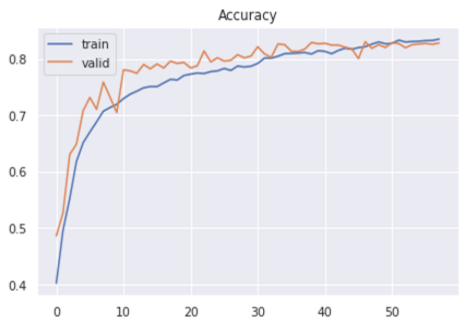

Fig.5. Training and validation accuracy of a facial emotion recognition model over a series of epochs.

The validation accuracy closely mirrors the training accuracy, indicating that the model is generalizing well rather than merely memorizing the training data. The epochs history demonstrates a gradual accuracy improvement reaching an 83% accuracy on both training and validation set.

Fig.6. Persons expressing different emotions

Fig.7. Facial images with labels indicating the true emotion and the predicted emotion for each face

We see examples of faces with various expressions, and each face has two labels. The first label indicates the true emotion, and the second label shows the emotion predicted by the model. For instance, a label might read "true: sad, pred: happy," which would indicate that the true emotion was sad, but the model incorrectly predicted the emotion as happy. This visual inspection can be particularly useful for identifying patterns in the types of errors the model is making. The 7th image in the first row appears more neutral than sad, which even our model accurately predicted as neutral. Whereas the last image in second row is very much sad. For example, if the model consistently mislabels 'sad' as 'happy', there might be an issue with how the model is interpreting certain facial features associated with sadness. Conversely, correct predictions like "true: neutral, pred: neutral" confirm that the model is capable of accurately recognizing certain expressions. Overall, the image captures the model's capacity to identify and differentiate between various facial emotions, which is crucial for refining the model and improving its accuracy.

5. Conclusion

This study concludes that a CNN-based approach for facial emotion recognition is effective, highlighting significant advancements in the accuracy and robustness of emotion classification systems. Leveraging the FER2013 dataset and a pre-trained VGG16 model, the research emphasizes the critical role of preprocessing steps, including resizing, normalization, and data augmentation techniques such as rotation, scaling, and shifting. These steps, combined with transfer learning, allowed the model to achieve an impressive 83% accuracy on the test set. The detailed performance metrics, particularly the high precision and recall for the 'happy' emotion, underline the model's proficiency, though it also identifies areas needing improvement, such as distinguishing between similar emotions like sadness and neutral. The training and validation curves reflect a well-fitted model with no overfitting, reinforcing its potential for real-world applications in mental health diagnostics, human-computer interaction, and security systems. Future work will aim to enhance the model's accuracy and robustness further, exploring more sophisticated data augmentation techniques and additional datasets to improve generalizability and applying the model to diverse environments to validate its practical efficacy.

Author contributions

V. Ajitha: Conceptualization, Methodology, Software,

T.S. Suganya: Field study, Data curation, Writing-Original draft preparation

S. Irin Sherly: Software, Validation,

P. Jose: Visualization, Investigation,

C. Gnanaprakasam: Writing,

M. Krishnaraj: Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] K. Fujii, D. Sugimura and T. Hamamoto, "Hierarchical Group-level Emotion Recognition in the Wild," 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), Lille, France, 2019, pp. 1-5, doi: 10.1109/FG.2019.8756573.

[2] Ranjan, Rakesh, and Bikash Chandra Sahana. "An efficient facial feature extraction method based supervised classification model for human facial emotion identification," In 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 1-6. IEEE, 2019.

[3] Sharma, Garima, Latika Singh, and Sumanlata Gautam. "Facial Feature Extraction for Emotion Classification Using Fuzzy C-Mean Clustering," Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science) 14, no. 7 (2021): 2210-2219.

[4] Tribedi, Sabyasachi, and Ranjit Kumar Barai. "Generating Context-Free Group-Level Emotion Landscapes Using Image Processing and Shallow Convolutional Neural Networks," In Progress in Computing, A Smart Agriculture Friendly Intelligent Flying Robot Design to Identify Plant Leaf Diseases by using Digital Image Processing with Learning ApproachAnalytics and Networking: Proceedings of ICCAN 2019, pp. 313-325. Springer Singapore, 2020.

[5] Korkmaz, Timuçin, and Hamza Erol. "Classification Of Human Facial Expressions For Emotion Recognition Using A Distributed Computer System," In 2020 5th International Conference on Computer Science and Engineering (UBMK), pp. 1-6. IEEE, 2020.

[6] .Kehri, Vikram, and R. N. Awale. "A facial EMG data analysis for emotion classification based on spectral kurtogram and CNN," International Journal of digital signals and smart systems 4, no. 1-3 (2020): 50-63.

[7] Kim, Jun-Hwa, and Chee Sun Won. "Emotion enhancement for facial images using GAN," In 2020 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), pp. 1-4. IEEE, 2020.

[8] Arya, R., and E. R. Vimina. "An evaluation of local binary descriptors for facial emotion classification," Innovations in Computer Science and Engineering: Proceedings of 7th ICICSE (2020): 195-205.

[9] Abbassi, Nessrine, Rabie Helaly, Mohamed Ali Hajjaji, and Abdellatif Mtibaa. "A deep learning facial emotion classification system: a VGGNet-19 based approach," In 2020 20th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), pp. 271-276. IEEE, 2020.

[10] Agastya, Wisnu, and Hanny Haryanto. "Sequential Model for Mapping Compound Emotions in Indonesian Sentences," Journal of Applied Intelligent System 5, no. 1 (2020): 32-46.

[11] Karbauskaitė, Rasa, Leonidas Sakalauskas, and Gintautas Dzemyda. "Kriging predictor for facial emotion recognition using numerical proximities of human emotions," Informatica 31, no. 2 (2020): 249-275.

[12] Ragupathy, P., and P. Vivekanandan,A modified fuzzy histogram of optical flow for emotion classification," Journal of Ambient Intelligence and Humanized Computing 12, no. 3 (2021): 3601-3608.

[13] Kim, Jaemyung, Jin-Ku Kang, and Yongwoo Kim. "A resource efficient integer-arithmetic-only FPGA-based CNN accelerator for real-time facial emotion

recognition," IEEE Access 9 (2021): 104367-104381.

[14] Fujii, Katsuya, Daisuke Sugimura, and Takayuki Hamamoto. "Hierarchical group-level emotion recognition," IEEE Transactions on Multimedia 23 (2020): 3892-3906.

[15] Gund, Manasi, Abhiram Ravi Bharadwaj, and Ifeoma Nwogu. "Interpretable emotion classification using temporal convolutional models," In 2020 25th international conference on pattern recognition (ICPR), pp. 6367-6374. IEEE, 2021.

[16] Fnaiech, Ahmed, Hanene Sahli, Mounir Sayadi, and Philippe Gorce. "Fear facial emotion recognition based on angular deviation," Electronics 10, no. 3 (2021): 358.

[17] Poulose, Alwin, Jung Hwan Kim, and Dong Seog Han. "Feature vector extraction technique for facial emotion recognition using facial landmarks," In 2021 International Conference on Information and Communication Technology Convergence (ICTC), pp. 1072-1076. IEEE, 2021.

[18] Wang, Shuai, Jingzi Qu, Yong Zhang, and Yidie Zhang. "Multimodal emotion recognition from EEG signals and facial expressions," IEEE Access 11 (2023): 33061-33068.

[19] Wang, Jiawen, and Leah Kawka. "GiMeFive: Towards Interpretable Facial Emotion Classification," arXiv preprint arXiv:2402.15662 2024.

[20] M.S.Kavitha, R. Harikrishnan, S. Vijay Shankar, S. Krishnakumari, S. Irin Sherly," A Smart Agriculture Friendly Intelligent Flying Robot Design to Identify Plant Leaf Diseases by using Digital Image Processing with Learning Approach," 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2023, IEEE, DOI: 10.1109/ICSES60034.2023.10465304.

[21] T.Judgi, R. Anitha, S.Padma, G.Nishanthi, S. Irin Sherly," A Secured Framework Model to Design Electronic Voting System using Block Chain Methodology," 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2023, IEEE, DOI: 10.1109/ICSES60034.2023.10465374.

[22] Geetha, M. Gomathi, A. Jayalakshmi, K. Neela, S. Irin Sherly," A Novel Methodology to Identify Autism Disorder in Earlier Stages using Artificial Intelligence Assisted Hybrid Learning Scheme," 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2023, IEEE, DOI: 10.1109/ICSES60034.2023.10465373.

[23] Irin Sherly, P. M. Kavitha, S. Jayachandran, M. Robinson Joel, R. Jeena," Empowering the Visually Impaired: Arduino Mega Enhanced Accessibility Device," International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 11 Issue: 8s, DOI: https://doi.org/10.17762/ijritcc.v11i8s.7223, 2023.

[24] Prabu Sankar N et. al., "Study of ECG Analysis based Cardiac Disease Prediction using Deep Learning Techniques," International Journal of Intelligent Systems and Applications in Engineering, 2023, 11(4), 431–438.

[25] S.Irin Sherly, G.Mathivanan," An Ensemble Basedheart Disease Predictionusing Gradient Boosting Decision Tree," Turkish Journal of Computer and Mathematics Education, Vol.12, No.10, 3648-3660, 2021.

[26] S.Irin Sherly, G.Mathivanan," An efficient honey badger based Faster region CNN for chronic heart failure prediction," Biomedical Signal Processing and Control, Vol. 79, 104165, 2023.

[27] S.Irin Sherly, G.Mathivanan," ECG Signal Noises Versus Filters for Signal Quality Improvement", 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), IEEE, DOI: 10.1109/ICAECT49130.2021.9392621, 2021.