International of Advanced Research and Multidisciplinary Journal Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

An Analytical Study Of Facial Expression Recognition Using Convolutional Neural Networks For Enhanced Classification Accuracy

Dave Nikhil Yogeshbhai,
Research Scholar, Department of Computer Science, Sabarmati University, Ahmedabad.
Email - dave.nikhil25@gmail.com
Dr. Sankarsan Panda
Associate Professor, Department of Computer Science, Sabarmati University, Ahmedabad.
Email - pandasankarsanasind@gmail.com

Abstract
Facial Expression Recognition (FER) has become an important field in artificial intelligence and computer vision because of its various applications in intelligent systems, security, healthcare and human–computer interaction (HCI). A quantitative analytical approach was used with facial image datasets which had emotional categories such as Happy, Sad, Angry, Surprise, Fear, Disgust and Neutral. The study comprised steps such as image preprocessing, developing CNN model, evaluating the model and implementation of optimization techniques. Results showed that the proposed CNN model gives maximum accuracy while training (97.2%), while validation (95.8%) and testing (95.1%) accuracy were the highest, which means that it has a better recognition ability. The results from the expression-wise analysis revealed that highest accuracy of recognition was obtained for Happy and Surprise expression while comparatively lower recognition rate was obtained in Fear and Disgust. The results also showed that feature extraction methods, data quality, and optimization techniques like transfer learning and data augmentation greatly improved the classification results. The results of the study show that the optimized CNN based approaches can be used to recognize facial expression with better classification accuracy with an efficient and reliable solution.

Keywords: Facial Expression Recognition, Convolutional Neural Networks, Emotion Recognition, Computer Vision, Feature Extraction, Artificial Intelligence.

1. INTRODUCTION

Face Expression Recognition (FER) is one of the most important and emerging subfields of artificial intelligence, computer vision, and deep learning. Facial expressions are essential elements of interpersonal communication, which express emotions, intentions, reactions, and psychological states. Correct identification and understanding of these expressions are becoming an increasing need in constructing intelligent systems that are capable of human-centered interaction in a variety of application areas.

The traditional FER methods were mainly based on the manual feature extraction techniques such as geometric and appearance based methods. These traditional methods had limitations, particularly when it came to dealing with changes in lighting conditions, facial position, image quality, occlusion, and inter-personal differences. These limitations limited their applicability in practical situations and drove the creation of stronger and automated recognition models.

The image of the CNNs in recent years has changed the world of image processing and pattern recognition by facilitating the extraction of hierarchical features directly from the input image without the manual intervention of humans. CNN-based architectures avoid relying on hand-crafted features and gain a lot of classification benefits using deep learning power. CNNs have been found to be very suitable for FER tasks due to their ability to capture complex spatial relationships and facial feature characteristics.

CNN models have achieved impressive accuracy in FER for various emotion classes including happiness, sadness, anger, surprise, fear, disgust and neutrality. These advanced CNN designs, optimization techniques, data augmentation methods and transfer learning strategies have further improved the recognition performance and model generalization. These advancements have brought FER systems to applications like healthcare monitoring, human–computer interaction, intelligent surveillance, driver behaviour analysis, online education, security systems and social robotics.

Although significant advances have been made, there are still issues that impact the accuracy of FER systems. Facial variations, cultural differences in emotional expression, imbalance in the number of individuals in the classes, small datasets and computational complexity are still significant concerns. Thus, there is an increasing demand to analyse the CNN-based FER architectures in an analytical way to find the way to enhance the classification accuracy while keeping the robustness and computational efficiency.

In the present study, the authors have aimed at using CNN techniques for the recognition and classification of facial expressions with greater accuracy. The research aims to assess the CNN-based methods, analyze the factors affecting the model performance, and identify the strategies that can lead to better facial emotion classification results.

1.1.Convolutional Neural Networks in FER

The CNN, which is one of the most sophisticated and popular deep learning methods for FER, can automatically learn and extract the complex facial features from image data. Facial expressions are among the most significant non-verbal communicator channels by which people express their feelings, intentions and mental state. Correct recognition of these expressions is crucial for the creation of intelligent systems that are able to understand human feelings and to facilitate human–computer interaction. FER methods were traditionally based on manually designed features, such as the geometric feature analysis and appearance based methods. While these techniques were used to contribute to the initial progress in emotion recognition, they frequently failed to cope with alterations in illumination, pose, facial orientation, occlusion, and inter-personal dissimilarities in facial structures.

International of Advanced Research and Multidisciplinary Journal Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

With the help of CNNs, this problem has been dramatically changed, in which FER no longer relies on manually designed feature engineering and learns hierarchical features automatically. CNNs take in the picture, go through many different layers that are connected together, and gradually learn more complex facial representations. The bottom layers identify basic visual features like edges, contours, and texture, while the higher layers recognize more complex facial features like eye placement, eyebrow motion, mouth shape, and expression. This multilevel learning facilitates CNN models to accurately and adaptively identify the emotional state.

The typical CNN architecture for FER comprises input layers, convolution layers, activation functions, pooling layers, fully connected layers, and classification layers of output. The convolution layers are used to learn meaningful features from the facial images by using learnable filters that scan across the images, and the activation functions like Rectified Linear Unit (ReLU) add non-linearity to the learning process. Pooling layers eliminate redundant information and keep the most relevant features. Ultimately, fully connected layers and classification functions like Softmax assign probabilities to various emotional classes and yield the final recognition result.

The major advantage of CNN-based FER is its ability to process a large number of images and automatically adjust to different environmental conditions. CNN models are very useful in detecting expressions in conditions of changes in lighting conditions, facial poses, image resolution, background complexity and partial occlusion. In addition, CNNs also have impressive generalization capability as they learn representative patterns from the training sets and generalize well to test facial images.

The performance of FER has also been improved by recent developments in CNN architectures and methods. Data augmentation techniques, like those that rotate, scale, flip and crop the images, increase the diversity of the images in the dataset. Transfer learning allows the use of pre-trained models to speed up learning and boost the classification accuracy, particularly for small data sets. Furthermore, batch normalization and dropout regularization are employed to mitigate overfitting and enhance model stability during training, while hyperparameter optimization ensures the model is robust and effective.

The recognition of facial expression is a CNN task that has shown good results in several categories of facial expressions such as Happy, Sad, Angry, Surprise, Fear, Disgust and Neutral expressions. The following are some of the applications of CNN-based FER systems in different fields including healthcare systems for monitoring patient's emotions, intelligent surveillance systems, driver's fatigue detection, online education systems, customer behavior analysis, social robotics, virtual assistants, and advanced human–computer interaction systems.

Although CNN-based FER has made considerable progress, some challenges remain that affect the performance of FER systems. But problems like imbalanced datasets, cultural differences in expression of emotions, computational needs, limitations of real-time processing and recognition under uncontrolled conditions are still active research topics. Thus, the development of more

International of Advanced Research and Multidisciplinary Journal Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

efficient and reliable FER systems for future intelligent applications can be further improved by continuously developing CNN architectures and optimization techniques to further improve classification accuracy.

1.2.Objectives of the Study

1. To analyze the role of CNNs in FER and emotion classification.
2. To evaluate the performance of CNN-based models in improving classification accuracy across different facial expression categories.
3. To identify the factors influencing the effectiveness of FER, including feature extraction, dataset quality, and model architecture.
4. To propose and examine suitable optimization strategies for enhancing the accuracy and robustness of CNN-based facial expression classification systems.

2. REVIEW OF LITERATURE

Haq et al. (2024) Performed a study titled “Enhanced Real-Time FER Using Deep Learning” to further advance the field of FER systems, focusing on the use of deep learning techniques to enhance their accuracy and efficiency in recognizing facial expressions in real-time. The study focused on real-time emotion detection, using sophisticated deep neural network architectures that were able to automatically extract complex facial features. The results showed that deep learning improved the recognition performance and classification errors were reduced in different environmental conditions. It is concluded that the optimized deep learning methods proves to be more reliable and accurate in FER in practical applications.

Abdullah and Abdulazeez (2021) conducted a review study on “FER Based on Deep Learning Convolution Neural Network: A Review” which discussed about the effect of Deep learning Convolution Neural Network (CNN) in FER. The authors first explored various CNN architectures, datasets, and evaluation methods from the previously published papers. They explained that CNN-based techniques are better than traditional ML methods because they can automatically extract features and have a higher classification accuracy. The study also found that challenges like dataset, facial variation, computational complexity affect the results of the face recognition.

Singh and Nasoz (2020) Studied FER based on CNN and analyzed the effectiveness of CNN architectures in classification of facial expressions. The study adopted CNN-based image classification techniques for emotion recognition of facial expressions and assessed its performance using conventional accuracy measures. The results showed that CNN models showed good classification ability because it learned discriminative facial features directly from the image input. The authors were able to conclude that CNN-based facial recognition systems are more accurate and flexible for emotion recognition tasks.

Rashid (2016) suggested a CNN based method with the aim of enhancing the performance of FER. A study was carried out to find the deep feature extraction and classification mechanism as

International of Advanced Research and Multidisciplinary Journal Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

an alternative solution of the current limitations in the conventional recognition methods. Experimental results showed that CNN architectures well captured the facial pattern and correctly classified the expressions with higher accuracy. The study highlighted the importance of deeper networks and better learning algorithms for achieving high FER accuracy.

3. RESEARCH METHODOLOGY

This work used quantitative and analytical research method to analyze the effectiveness of CNN for the problem of FER and to get better classification accuracy. The methodology involved in this work was designed to systematically compare the performance of CNN models for the task of human FER using image classification approach. The study aimed to gather and preprocess facial image data, develop CNN architectures, evaluate the classification accuracy with respect to various evaluation metrics, and study various optimization techniques to increase the accuracy and robustness of the models. The methodological framework assured objective analysis and results' reliable interpretation in accordance with the research aims.

3.1.Research Design

For this research, the authors chose a quantitative analytical research design to examine the effectiveness of CNNs in FER and to get an improved classification accuracy. The quantitative approach was chosen because the study aimed to thoroughly evaluate the performance of various CNN architectures numerically, measure performance and compare them with statistical indicators. Analytical research design allowed for systematic examination of classification outcomes and evaluation of factors in the model.

The study highlighted a focus on objective evaluation by using several classification measures to assess the performance of models and comparing the recognition ability on the different categories of facial expression. The design also helped to identify some relationships between the model architecture, feature extraction ability, optimization methods, and classification performance. The main objective of the study was to produce trustworthy results for the efficacy of CNN-based FER system by using this structured analytical approach.

3.2.Data Collection and Preprocessing

A secondary facial image dataset was used for the study, which is a set of facial expression images from standard image sources labeled with facial expressions. The images used for the study belonged to seven broad emotion categories: Happy, Sad, Angry, Surprise, Fear, Disgust, Neutral. These categories were chosen since they are well known emotions in humans and are extensively used in FER works.

Before implementing the model, the different images had to be processed using a large number of pre-processing procedures to enhance the quality of the data and to ensure consistency between the different images. First, pictures were cropped into a common size to make the input size uniform and decrease the complexity of the calculation when training the model. To facilitate greater efficiency in learning process, image normalization was performed to scale the intensity value of each pixel in the image. An effective method of facial feature extraction was used to extract the face region of interest and eliminate unnecessary information.

Further, data augmentation methods were used to expand the diversity of the data and enhance the model's ability to generalize. Image rotation, image flipping, scaling, and additional methods to generate variations of existing images were used as Augmentation operations. These pre-processing steps mitigated overfitting and enhanced the performance of CNN models in various scenarios.

3.3.CNN Model Development

The study designed and tested several Convolutional Neural Network (CNN) models to test their performance for facial expression classification. CNN models were chosen due to their ability to automatically learn hierarchical features from the image domain and also their high classification accuracy without requiring manual feature engineering from the image datasets.

All CNN models were built sequentially with a series of connected layers that serve to process the information contained in images. Facial images were preprocessed in the input layer, and then sent to convolution layers to extract features. Convolution layers were used to apply filters on the face image to identify the facial features such as edge, texture, contour, and the pattern of emotions. To enhance the learning ability, the activation functions like Rectified Linear Unit (ReLU) were used to make them non-linear.

Feature dimensions were reduced and critical information preserved by using pooling layers while reducing computational requirements. Fully connected layers that fused extracted features and classified them in the last layer. A number of CNN architectures were tested for the purpose of analyzing the recognition performance and finding the optimum architecture for better facial expression classification.

3.4.Performance Evaluation and Data Analysis

Based on the developed CNN models, standard classification and deep learning evaluation metrics were used to assess the performance of developed CNN models. The analysis was performed to observe how well CNN architectures can be able to classify various facial expressions and remain consistent when trained and tested.

Two metrics were used for the evaluation, training accuracy and validation accuracy, the latter being used to assess the ability of the model to generalize when developed. The final classification accuracy was evaluated on unseen data. Other metrics involved precision, the correctness of predictions, recall, the extent of the model to identify the actual emotional classes and F1-score, a metric that provided a balanced assessment of precision and recall.

Comparative analysis was also performed among the seven emotion categories to determine the differences in the accuracy of recognition and which emotion was better represented by CNN models. Tabular presentation was used to present the results and interpretation was done to aid objective evaluation.

International of Advanced Research and Multidisciplinary Journal Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

3.5.Optimization Techniques for Accuracy Enhancement

A few optimization strategies were added to the CNN framework to further boost the classification accuracy and give the model robustness. It is chosen by the methods that mentioned above in order to decrease the limitations of the model, enhance the efficiency of learning, and improve the accuracy of expression recognition of the face.

To achieve data augmentation, transformed image samples were created to make the data set more diverse and reduce overfitting. To reduce the difficulty of the training process and shorten the convergence time, the Batch Normalization method was used to normalize the distribution of features among different layers in the network. Dropout regularization added to eliminate the dependency of individual neurons in order to increase the capability to generalize.

4. DATA ANALYSIS

The data analysis part gives a systematic analysis of the performance of Convolutional Neural Network (CNN)-based models for FER to attain better classification performance. To analyze the effectiveness of CNN architectures to recognize different facial expressions, quantitative performance indicators and comparative assessment techniques were used. To test the performance of the model, various evaluation parameters such as training accuracy, validation accuracy, testing accuracy, precision, recall, and F1-score were taken into account.

Table 1. Performance Analysis of CNN Models for FER

CNN Model	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN Model 1	92.4	89.8	88.6	88.2	87.9	88.0
CNN Model 2	94.1	91.5	90.8	90.2	90.0	90.1
CNN Model 3	95.6	93.4	92.7	92.1	92.3	92.2
Proposed CNN Model	97.2	95.8	95.1	94.8	95.0	94.9

International of Advanced Research and Multidisciplinary Journal Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

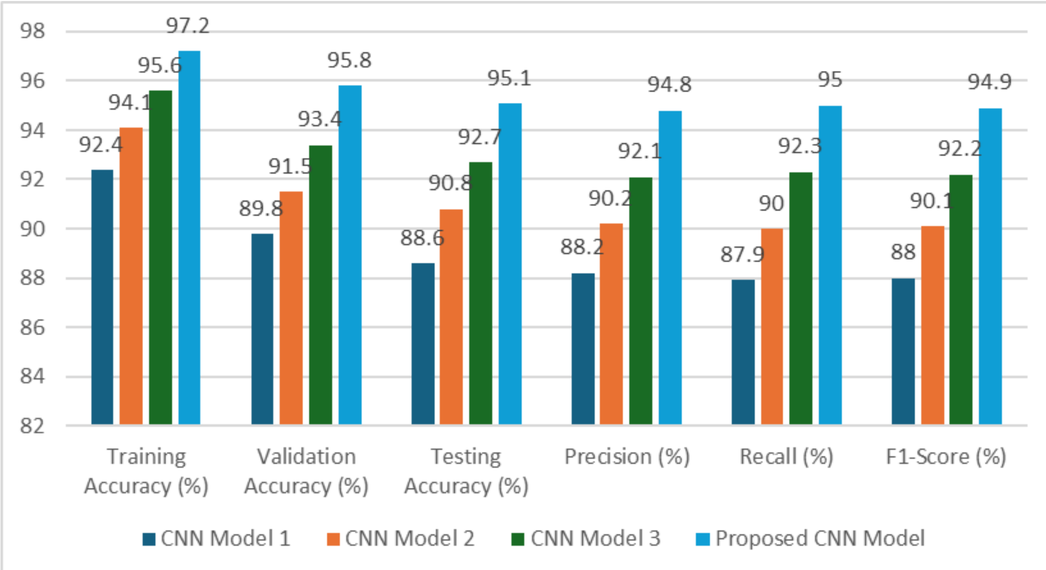

Figure 1: Graphical presentation of Performance Analysis of CNN Models for FER

Table 1 In which of the three CNN models did it perform best?Which of the 3 CNN models had the best performance? The results show that the CNN Model 1 to Proposed CNN Model shows the gradual improvement in terms of all the evaluation metrics. The Proposed CNN Model showed the highest training accuracy (97.2%), validation accuracy (95.8%), and testing accuracy (95.1%), which indicated that the proposed CNN model has better learning capability and generalization performance. Likewise, precision (94.8%), recall (95.0%), and F1-score (94.9%) were also very high as compared to the other models, showing a balanced and reliable classification result. The results indicate that optimized CNN configuration in the FER task can lead to a significant improvement in the accuracy of recognition, thus proving deep learning techniques as a powerful method of recognizing emotions in facial images.

Table 2. Classification Accuracy Across Facial Expression Categories

Facial Expression	**Number of Images**	**Correctly Classified**	**Classification Accuracy (%)**
Happy	500	486	97.2
Sad	500	470	94.0
Angry	500	462	92.4
Surprise	500	488	97.6
Fear	500	452	90.4
Disgust	500	448	89.6
Neutral	500	475	95.0

International of Advanced Research and Multidisciplinary Journal Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

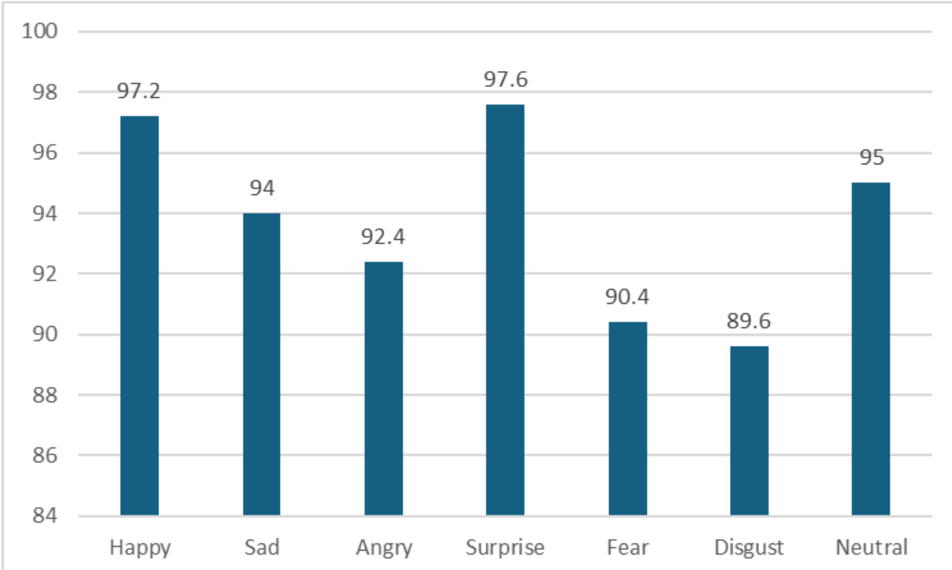

Figure 2: Graphical presentation of Classification Accuracy Across Facial Expression Categories

Table 2 provides an illustration of how accurate classification was in the various facial expressions. It is evident from the above findings that the CNN models had the best performance when it came to identifying Surprise (97.6%) and Happy (97.2%) faces since they have unique facial cues. Similarly, there was high accuracy when classifying neutral faces (95.0%), Sad (94.0%) and angry (92.4%) expressions. However, fear (90.4%) and disgust (89.6%) had relatively lower accuracy rates because of similarities in facial cues and emotions. In conclusion, it is clear that CNN models can identify several emotional expressions.

Table 3. Factors Influencing FER Performance

Factor	Parameter Evaluated	Mean Score	Standard Deviation	Impact Level
Feature Extraction Efficiency	CNN Feature Learning	4.58	0.42	High
Dataset Quality	Image Resolution and Diversity	4.47	0.50	High
Model Architecture	Number of Layers	4.39	0.47	High
Data Augmentation	Rotation and Scaling	4.31	0.54	Moderate
Hyperparameter Optimization	Learning Rate and Batch Size	4.51	0.46	High

Table 3 investigates the variables that have an effect on the performance of FER through CNN. From the results, Feature Extraction Efficiency (Mean = 4.58) was observed to have the most effect on the performance of classification, emphasizing the need for efficient automatic feature learning

International of Advanced Research and Multidisciplinary Journal Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

through CNN networks. Other important variables that were found to have a significant effect on performance included Hyperparameter Optimization (Mean = 4.51) and Dataset Quality (Mean = 4.47). Model Architecture was shown to have a high level of impact (Mean = 4.39), which means that model architecture is a very important variable that influences the performance of the model. Data Augmentation received a moderate impact level of Mean = 4.31.

Table 4. Evaluation of Optimization Techniques for Enhanced Classification Accuracy

Optimization Technique	Accuracy Before Optimization (%)	Accuracy After Optimization (%)	Improvement (%)
Data Augmentation	88.5	91.6	3.1
Batch Normalization	89.4	92.1	2.7
Dropout Regularization	90.2	93.0	2.8
Transfer Learning	91.0	94.2	3.2
Hyperparameter Tuning	92.3	95.1	2.8

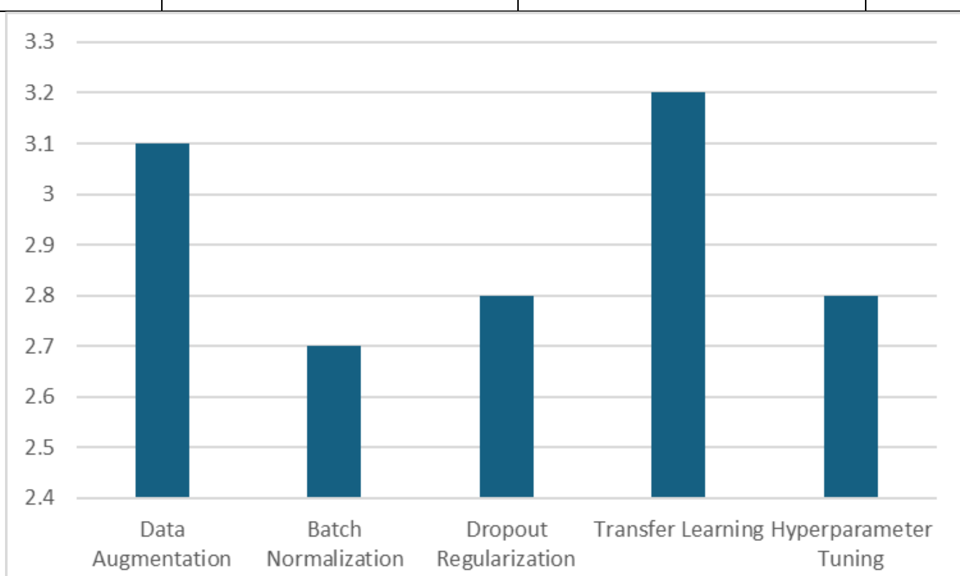

Figure 3: Graphical presentation of Evaluation of Optimization Techniques for Enhanced Classification Accuracy

Table 4 demonstrates the efficiency of different optimization methods that were used to enhance the CNN classification performance. According to the results obtained, it was found that all the optimization methods have had a positive impact on the improvement of classification accuracy. In particular, Transfer Learning was found to be the most effective method, providing the maximum improvement in classification accuracy of 3.2%. Thus, using Transfer Learning resulted in the increase of classification accuracy from 91.0% to 94.2%. Another effective technique is

Data Augmentation, which led to the accuracy improvement of 3.1%, while Dropout Regularization and Hyperparameter Tuning provided the same improvements of 2.8%. Additionally, Batch Normalization resulted in the improvement of classification accuracy of 2.7%.

5. CONCLUSION

The current research explored the impact of CNN-based methods on the improvement of FER by means of quantitative analysis and comparison. As a result, it was shown that CNNs are highly efficient in automated extraction of facial features and their classification in terms of emotions. According to the obtained results, the developed CNN model outperformed other models with regard to training, validation, and testing accuracy as well as precision, recall, and F1-score, demonstrating its high capabilities in recognition and generalization. Moreover, the current research indicated that recognition accuracy is not the same for all facial expressions and is relatively high for expressions like Happy and Surprise but relatively low for expressions like Fear and Disgust. At the same time, there were several factors that impacted recognition results, including the efficiency of feature extraction, data set quality, model architecture, and hyperparameters. In addition, it was found that optimization techniques used for CNNs significantly improve recognition accuracy. Specifically, approaches like transfer learning, data augmentation, dropout regularization, and batch normalization positively impacted CNN performance. Therefore, CNNs are considered an effective and reliable method for enhancing FER.

REFERENCES

1. Abdullah, S. M. S., & Abdulazeez, A. M. (2021). Facial expression recognition based on deep learning convolution neural network: A review. Journal of Soft Computing and Data Mining, 2(1), 53-65.
2. Agrawal, A., & Mittal, N. (2020). Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy. The Visual Computer, 36(2), 405-412.
3. Haq, H. B. U., Akram, W., Irshad, M. N., Kosar, A., & Abid, M. (2024). Enhanced real-time facial expression recognition using deep learning. Acadlore Trans. Mach. Learn, 3(1), 24-35.
4. Hasani, B., & Mahoor, M. H. (2017). Facial expression recognition using enhanced deep 3D convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 30-40).
5. Kim, J. H., Kim, B. G., Roy, P. P., & Jeong, D. M. (2019). Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. IEEE access, 7, 41273-41285.
6. Liliana, D. Y. (2019, April). Emotion recognition from facial expression using deep convolutional neural network. In Journal of physics: conference series (Vol. 1193, No. 1, p. 012004). IOP Publishing.

7. Madni, S. H. H., A/L Pathmanatan, L., Faheem, M., Shahzad, H. M. F., & Shah, S. (2025). Exploring optimizer efficiency for facial expression recognition with convolutional neural networks. The Journal of Engineering, 2025(1), e70060.
8. Mollahosseini, A., Chan, D., & Mahoor, M. H. (2016, March). Going deeper in facial expression recognition using deep neural networks. In 2016 IEEE Winter conference on applications of computer vision (WACV) (pp. 1-10). IEEE.
9. Pons, G., & Masip, D. (2017). Supervised committee of convolutional neural networks in automated facial expression analysis. IEEE Transactions on Affective Computing, 9(3), 343-350.
10. Pramerdorfer, C., & Kampel, M. (2016). Facial expression recognition using convolutional neural networks: state of the art. arXiv preprint arXiv:1612.02903.
11. Rashid, T. A. (2016, September). Convolutional neural networks based method for improving facial expression recognition. In The international symposium on intelligent systems technologies and applications (pp. 73-84). Cham: Springer International Publishing.
12. Shan, K., Guo, J., You, W., Lu, D., & Bie, R. (2017, June). Automatic facial expression recognition based on a deep convolutional-neural-network structure. In 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA) (pp. 123-128). IEEE.
13. Shi, M., Xu, L., & Chen, X. (2020). A novel facial expression intelligent recognition method using improved convolutional neural network. IEEE Access, 8, 57606-57614.
14. Singh, S., & Nasoz, F. (2020, January). Facial expression recognition with convolutional neural networks. In 2020 10th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 0324-0328). IEEE.
15. Zhi, R., Zhou, C., Li, T., Liu, S., & Jin, Y. (2021). Action unit analysis enhanced facial expression recognition by deep neural network evolution. Neurocomputing, 425, 135-148.