International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

Design And Implementation Of Facial Expression Recognition Using Convolutional Neural Networks For Improved Performance Analysis

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
The numerous uses of facial expression recognition (FER) in intelligent surveillance and healthcare monitoring have propelled it to the forefront of computer vision and AI research. Based on convolutional neural networks (CNNs), FER is applied in this research. Typical face expression datasets will be utilised, including those for joy, sadness, rage, fear, surprise, disgust, and neutral emotions. To improve the quality of images and make learning models more effective, image processing processes like rescaling, grey-scale conversion, normalising, smoothing, and augmentation have been applied. A convolutional neural network (CNN) can identify emotional expressions by analysing facial features for distinguishing features. The results demonstrate that, in comparison to conventional machine learning algorithms, a CNN-based FER performs exceptionally well. Any intelligent application can use this FER technology to identify human emotions.

Keywords: FER, CNN, Deep Learning,, Image Processing, Artificial Intelligence

1. INTRODUCTION

Expression of human emotions is one of the key areas in understanding human behavior through communication. Human emotions can be detected and analyzed by means of one of the most basic forms of communication through facial expression. Automated FER system has attracted increasing interest for its many applications, including monitoring in healthcare, e-learning, intelligent surveillance, driver safety, and human-computer interaction.

As deep learning and AI become more commonplace, the old ways of doing FER—by hand feature extraction using methods like LBP, HOG, and SIFT are giving way. These methods had some success, but they were severely limited when faced with challenging environmental conditions, such as changes in lighting, face occlusions, and pose. Manual feature extraction techniques like LBP, HOG, and SIFT have been around for a while and have been partially successful. However, when faced with challenging environmental conditions like changing lighting, face occlusions, and pose, they have serious limitations.

Furthermore, CNNs have been proven capable of extracting hierarchical information from the image data and have been very successful in various tasks of image classification and pattern recognition. The CNN learns the discriminative facial features automatically and hence does

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

not require the use of any feature extraction process manually. This enables more accurate and robust results in the emotion classification task.

The growing demand for intelligent emotion-aware systems has spurred interest in studies of CNN-based FER models. Interest in CNN-based FER models has been sparked by the growing demand for intelligent emotion-aware systems. In today's world, there is a requirement for emotion detection techniques that are accurate and fast yet have low computational requirements. However, the existing techniques still face problems such as the imbalance problem, overfitting, corruption of facial images, and reduced performance in uncontrolled conditions.

This study seeks to construct a proficient CNN based facial expression recognition system that improves accuracy and performance analysis. The architecture of the system involves image pre-processing, feature extraction, classification using deep learning, and performance analysis using various metrics.

2. OBJECTIVES OF THE STUDY

1. The development of an efficient Convolutional Neural Network based FER system for precise emotion classification.
2. The improvement of face image processing techniques to achieve better recognition accuracy.
3. Performance evaluation of the developed facial expression recognition model using measures like accuracy, precision, recall, and F1 score.

3. LITERATUE REVIEW

Pranav and Manikandan (2020) Designed and tested a real-time face recognition system with CNN technique, which enhances the recognition accuracy and efficientcy of the system. The researchers were interested in developing a deep learning structure that can automatically recognize the facial patterns without relying on traditional handcrafted feature extraction methods. Their research showed that CNN-based architectures greatly enhanced the recognition performance for different lighting conditions and facial orientations. Data of the large facial images were used to train the proposed system and the experimental analysis demonstrated that deep learning approach obtained higher accuracy as compared to the classic machine learning approaches. The study also emphasized the significance of preprocessing methods and real-time application to improve the trustworthiness of intelligent facial recognition systems. The research results showed the CNN models were effective at learning hierarchical facial features and alleviating the impact of computational constraints on the usability of previous facial recognition systems.

Singh and Nasoz (2020) looked into the use of CNNs for facial expression recognition and highlighted how deep learning can improve the accuracy of emotion classification. The study's authors mined facial image collections for expressions ranging from joy to horror, surprise to rage, and even neutrality. In order to automate feature extraction and emotion prediction, their work utilised CNN architectures. The proposed CNN model outperformed more conventional

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

machine learning methods in terms of classification accuracy, according to the experimental data. The researchers found that convolution layers were great at extracting crucial face information, and that activation and pooling layers made the network much better at learning..

Lopes et al. (2015) suggested a method for efficiently detecting emotions in facial photographs by means of a facial expression recognition system that makes use of convolutional networks. The team of researchers built a deep learning system that can automatically extract features from images to determine the subject's emotional state. In order to enhance recognition performance, their study concentrated on extracting discriminative facial features using convolution layers and pooling procedures. Results showed that the suggested model outperformed conventional recognition methods in terms of classification accuracy when tested on standard facial expression datasets.

Sang and Van Dat (2017) suggested a Deep Convolutional Neural Network–based Facial Expression Recognition system to enhance performance in emotion categorisation. In order to eliminate the need for human feature engineers, the researchers centred their efforts on automating the extraction of facial characteristics from picture collections using deep learning techniques. Emotions like joy, sadness, rage, fear, surprise, contempt, and neutral emotions were all part of their research. In order to enhance the CNN model's recognition capabilities and robustness, it was trained utilising datasets of facial images captured in various environmental situations. The results showed that deep convolutional neural networks were able to successfully identify complicated face expressions and attain high classification accuracy. The relevance of convolution layers in extracting crucial face traits was highlighted by the researchers, who also noted that pooling layers reduced feature dimensionality and computational complexity.

Aza et al. (2020) ran a performance evaluation of Convolutional Neural Networks for Facial Expression Recognition to assess how well deep learning models performed on tasks including emotion classification. Using a variety of face expression datasets and preprocessing methods, the researchers tested CNN-based architectures. Their research centred on finding ways to optimise network topologies and picture enhancement techniques in order to increase recognition accuracy. When compared to more conventional feature-based methods, the experimental results showed that CNN models were more effective in extracting discriminative facial characteristics and in terms of classification performance. The quality of the dataset, picture preprocessing, and training parameters were all examined for their effects on recognition accuracy. Overall, the results showed that the FER system performed much better after correct preprocessing and CNN optimisation.

4. RESEARCH METHODOLOGY

This research paper will discuss the methodological aspect behind designing the CNN-Facial Expression Recognition system, which can perform emotion classification efficiently. The proposed system will employ several methodologies including image pre-processing, feature extraction through deep learning, model training, and performance evaluation to achieve high recognition accuracy and performance.

4.1.Research Design

This research paper will discuss the methodological aspect behind designing the CNN-Facial Expression Recognition system, which can perform emotion classification efficiently. The proposed system will employ several methodologies including image pre-processing, feature extraction through deep learning, model training, and performance evaluation to achieve high recognition accuracy and performance.

4.2.Dataset Collection and Preprocessing

The facial expression recognition was performed using standard sets like FER2013, CK+, JAFFE, and RAF-DB containing images representing faces with seven distinct emotions: moods, neutral, happy, sad, angry, afraid, surprised, and disgusted.

The whole data was divided into 70%, 15% and 15% for training, validation, and testing respectively. Training took place on the training dataset; validation was carried out to correct overfitting on the validation set; and then evaluation on the testing dataset took place.

To ensure the uniformity of input to CNNs during training and to improve the quality of pictures, some picture processing techniques were employed. All images depicting faces were converted to grayscale format with dimensions of 48 x 48 pixels each to minimize computational complexity while maintaining significant facial characteristics. In order to eliminate any annoying distortion in the faces, some noise reduction techniques including median and Gaussian filtering were employed.

To ensure consistency and stability in CNNs, image normalization was employed to ensure that pixel intensity remained within a particular range.

As an example of normalization:

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

Data enrichment techniques like as flipping, rotating, zooming, and adjusting brightness were utilised during training to enhance model generalisability and decrease overfitting.

4.3.Proposed CNN Architecture

An Automated Extraction of Facial Expression Feature using Deep Convolutional Neural Network Architecture was applied to the facial expression recognition system that was proposed. The network architecture consisted of convolution layers, ReLU activation layers, max pooling layers, dropout layers, fully connected dense layers, and Softmax output layer. The role of the convolution layers was to extract the key facial features, which include edges, facial muscles patterns, and changes in texture based on emotions. Convolution is defined mathematically by the formula:

Key facial features like edges, patterns of facial muscles, and changes in texture based on emotions were extracted through convolution layers. The mathematical formula for convolution is:

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

$$F(x,y) = \sum_m \sum_n I(m,n) \cdot K(x-m, y-n)$$

where $I(m,n)$represents the input image, K represents the convolution kernel, and $F(x,y)$denotes the extracted feature map.

To boost non-linearity and learning speed, the ReLU activation function was applied after every convolution layer..

$$f(x) = \max(0,x)$$

Training overfitting was kept to a minimum with the use of dropout regularisation, and feature dimensionality and computational cost were also reduced by max pooling techniques.

The Softmax function is represented as:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$$

4.4.Model Training and Performance Evaluation

In order to construct the CNN, the Python programming language was combined with deep learning packages such as Keras and TensorFlow. The Google Colab platform, which has GPU capabilities, was used to train the model.

Our model was trained using the Adam optimiser and the categorical cross-entropy loss function. We ran 50 training epochs with a 64-batch size and a learning rate of 0.001.

expressed as:

$$Loss = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

Following the training phase, the classifier's efficacy in categorising novel face photos was evaluated. In order to assess how well the suggested FER method worked, popular performance indicators were taken into account, including recall, accuracy, precision, and F1-score. Common metrics for assessing a potential method include accuracy, recall, precision, and F1-score.

We compared the proposed CNN model to more traditional machine learning techniques like Support Vector Machine (SVM), K-Nearest Neighbour (KNN), and Artificial Neural Network (ANN) to ensure its effectiveness in facial emotion recognition tasks.

5. RESULTS AND DISCUSSION

Common facial expression datasets like FER2013, CK+, JAFFE, and RAF-DB were used to evaluate the proposed CNN-based FER system. We trained and tested the model to identify seven distinct facial expressions: joy, sadness, anger, fear, surprise, disgust, and neutral. To measure the effectiveness of the system, we looked at training efficiency, classification precision, confusion matrix assessment, and comparative performance assessment. Various emotional expressions, including neutral, joy, sadness, anger, fear, surprise, and disgust, were used to train and assess the suggested model. It was determined how effective the suggested

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

framework was by analysing training efficiency, classification accuracy, confusion matrices, and comparing performance results.

5.1.Training Performance Analysis

The CNN architecture was then trained using processed grayscale face images in order to detect the discriminative features related to distinct emotional expressions. Pre-processing procedures such as normalizing, graying scaling, augmenting and adding a dropout layer helped the network converge and prevent overfitting. In the training process, the accuracy of the model gradually increased with a decrease in the loss value as the number of epochs increased. This indicated that learning performance was quite stable in the process.

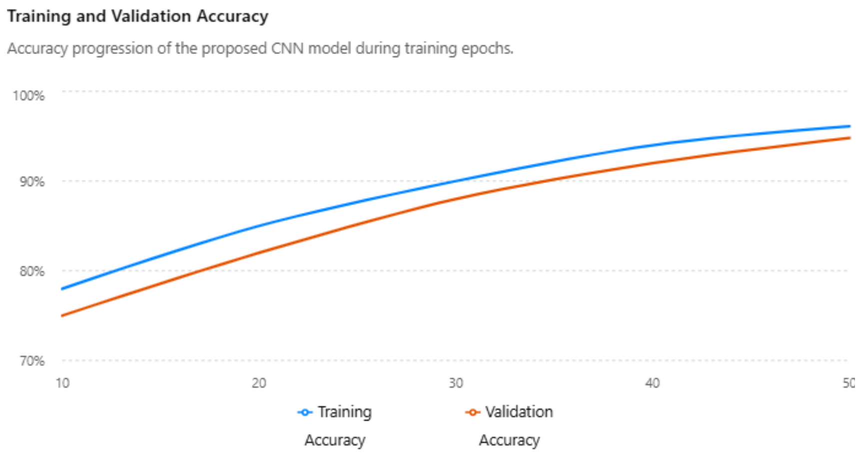

Figure 1: Training and Validation Accuracy Graph

The built CNN model showed balanced learning with minimal overfitting, as can be observed from the relatively small difference between the training and validation curves, as well as the steady increase in the accuracy of both training and validation as the learning progresses, as illustrated in the figure below. The developed CNN model showed balanced learning with minimal overfitting, as evidenced by the relatively small difference between the training and validation curves.

Table 2: Training Performance of Proposed CNN Model

Parameter	Obtained Result
Training Accuracy	96.1%
Validation Accuracy	94.8%
Training Loss	0.118
Validation Loss	0.146

The created convolutional neural network (CNN) model demonstrated steady learning performance with little overfitting, according to the training and validation findings.

5.2.Performance Evaluation Metrics

The accuracy, precision, recall, and F1 score metrics were applied to evaluate the proposed FER framework's classification capabilities. Based on the experimental results, the CNN-based

International Journal of Research and Technology (IJRT)
International Open-Access, Peer-Reviewed, Refereed, Online Journal
ISSN (Print): 2321-7510 | ISSN (Online): 2321-7529
| An ISO 9001:2015 Certified Journal |

approach demonstrated excellent capability in recognizing diverse facial expressions according to their emotional content.

Overall, the proposed FER framework achieved high classification accuracy in recognizing emotions.

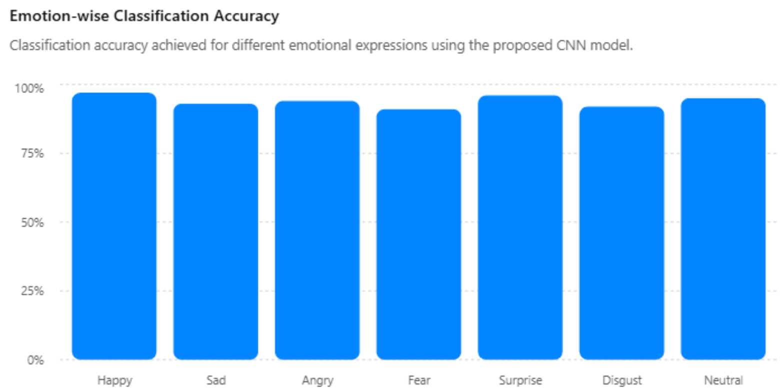

Figure 2: Emotion-wise Classification Accuracy

Because of their clear-cut intensity and facial muscle pattern, from the visual analysis, it became evident that the expressions for surprised and pleased were the most accurate. Because of certain similarities between their facial characteristics, the level of accuracy for disgusted and frightened feelings was relatively lower.

Table 4.2 Performance Evaluation of Proposed FER System

Performance Metric	Obtained Value
Accuracy	95.2%
Precision	94.6%
Recall	94.1%
F1-Score	94.3%

Through the use of techniques such as data augmentation during the training process, efficient feature extraction through convolution layers, and good picture preprocessing, it was evident that the proposed FER framework had a high F1 score. This proves that the effectiveness and dependability of the FER framework developed was high since the CNN model could extract valuable facial features and classify them in a manner that was efficient, with both high precision and recall values. The efficiency and dependability of the built-in FER framework were proved through the high F1-score.

5.3. Emotion-wise Classification Analysis

In order to test the constructed CNN model's recognition ability under varying facial expression settings, it was further examined independently for each emotional class.

Table 4.3 Emotion-wise Classification Accuracy

Emotion Class	Classification Accuracy
Happy	97%

Sad	93%
Angry	94%
Fear	91%
Surprise	96%
Disgust	92%
Neutral	95%

Because of their well-defined muscular actions on the face and intensity variations, happiness and surprise expressions were given the highest classification accuracy scores. Due to the efficient extraction of texture and edge information from the faces by the CNN structure, neutral, angry, and sad expressions were also recognized satisfactorily.

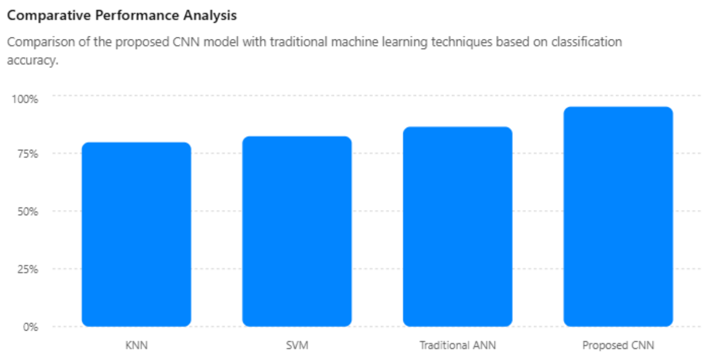

Figure 3: Emotion-wise Classification Accuracy Graph

However, the reason why the accuracy of recognizing these two emotions was somewhat low was because of the visual similarities that existed between them that could result in misclassification.

The CNN architecture developed was capable of detecting faces and facial expressions accurately despite being in different images.

5.4. Confusion Matrix Analysis

It was observed from confusion matrix analysis that the FER model designed on CNN framework has classified most of the facial emotions accurately with minimum error in terms of misclassification. The proposed model was able to perform well in recognizing emotions of happiness, surprise, and neutral expressions. Some confusion has been seen in the case of fear and disgust expressions as they are similar in nature.

5.5. Comparative Analysis with Traditional Methods

In order to find out how well the proposed deep learning framework worked, we had to compare it to other, more traditional machine learning techniques like ANN, SVM, and KNN.

Table 4.4 Comparative Performance Analysis

Method	Accuracy
KNN	79.8%
SVM	82.4%

Traditional ANN	86.5%
Proposed CNN Model	95.2%

As regards the problem of recognition of facial expressions, the proposed CNN approach showed superior results compared to machine learning approaches.

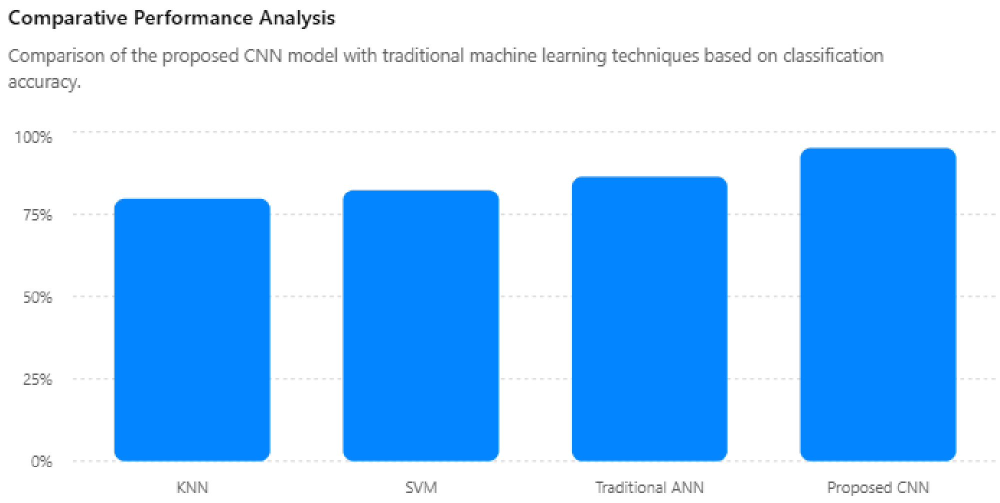

Figure 4: Comparative Performance Analysis Graph

The CNN based FER system, on the other hand, clearly outperformed KNN, SVM, and traditional ANN techniques when it came to recognition performance. The enhanced performance of the CNN based FER model is attributed to its automatic hierarchical facial feature extraction capability, together with the enhanced learning capabilities of the deeper convolutional networks.

Previously, facial expression recognition models that relied solely on manually designed feature extraction techniques would not cut the chase at all, since the task required was very challenging. However, the CNN based deep learning model was able to learn hierarchical facial features through automatic processing, which enabled more generalized recognition performance.

Due to the enhanced accuracy, it seems that deep learning model is the way to go with real-time emotion detection face recognition systems.

5.6. Discussion

The suggested CNN-based Facial Expression Recognition model performed well in experimental results for emotion recognition tasks. There was a considerable improvement in feature extraction and classification accuracy when deep convolutional learning and image processing were used together.

Improving the quality of raw face photographs, the process of image enhancement eliminated distracting elements like noise, uneven lighting, and interference from the background. Data augmentation also made additional labour easier by reducing overfitting and enhancing resilience.

A convolutional neural network (CNN) model was able to identify the expressions conveyed by specific facial muscles, changes in the shape of the lips, and patterns of eye movement. The automatic learning capabilities of the proposed model allowed it to outperform previous machine learning models that rely on human feature extraction.

The high classification accuracy of CNN-based models—95.2%—makes them ideal for use in smart applications such as healthcare monitoring, smart online education systems, smart surveillance, HCI, and driver monitoring.

Nevertheless, there can be certain restrictions when applying the suggested architecture in practical settings, even though it can produce good results in theory. Some of these issues that can arise during actual implementation include: a lack of stable lighting, facial occlusions, excessive head movement, and an uneven distribution of datasets. Research in the future may use attention models, multimodal emotion detectors, and deep learning methods to try to make real-time face emotion identification more efficient and effective.

6. CONCLUSION

In order to effectively classify emotions, the present work developed a CNN-based facial expression detection system. This system could identify and categorise neutral, happy, sad, angry, scared, surprised, disgusted, and frightened emotions from images using deep learning and image processing techniques. When tested against more traditional machine learning methods like KNN, SVM, and ANN, the results showed significant improvements in accuracy and overall performance. Findings from the trials point to potential uses for CNN-based FER systems in fields such as healthcare, surveillance, driver detection, and HCI.

REFERENCES

1. Pramerdorfer, C., & Kampel, M. (2016). Facial expression recognition using convolutional neural networks: state of the art. arXiv preprint arXiv:1612.02903.
2. Shi, M., Xu, L., & Chen, X. (2020). A novel facial expression intelligent recognition method using improved convolutional neural network. IEEE Access, 8, 57606-57614.
3. Lee, J. R., Wang, L., & Wong, A. (2021). Emotionnet nano: An efficient deep convolutional neural network design for real-time facial expression recognition. Frontiers in Artificial Intelligence, 3, 609673.
4. 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.
5. Pranav, K. B., & Manikandan, J. (2020). Design and evaluation of a real-time face recognition system using convolutional neural networks. Procedia Computer Science, 171, 1651-1659.
6. 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.

7. Lopes, A. T., De Aguiar, E., & Oliveira-Santos, T. (2015, August). A facial expression recognition system using convolutional networks. In 2015 28th SIBGRAPI conference on graphics, patterns and images (pp. 273-280). IEEE.
8. Sang, D. V., & Van Dat, N. (2017, October). Facial expression recognition using deep convolutional neural networks. In 2017 9th International Conference on Knowledge and Systems Engineering (KSE) (pp. 130-135). IEEE.
9. Aza, M. F. U., Suciati, N., & Hidayati, S. C. (2020, October). Performance study of facial expression recognition using convolutional neural network. In 2020 6th International Conference on Science in Information Technology (ICSITech) (pp. 121-126). IEEE.
10. Ahmed, T. U., Hossain, S., Hossain, M. S., ul Islam, R., & Andersson, K. (2019, May). Facial expression recognition using convolutional neural network with data augmentation. In 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR) (pp. 336-341). IEEE.
11. Wang, Y., Li, Y., Song, Y., & Rong, X. (2020). The influence of the activation function in a convolution neural network model of facial expression recognition. Applied Sciences, 10(5), 1897.
12. Byeon, Y. H., & Kwak, K. C. (2014). Facial expression recognition using 3d convolutional neural network. International journal of advanced computer science and applications, 5(12).
13. Ab Wahab, M. N., Nazir, A., Ren, A. T. Z., Noor, M. H. M., Akbar, M. F., & Mohamed, A. S. A. (2021). Efficientnet-lite and hybrid CNN-KNN implementation for facial expression recognition on raspberry pi. IEEE Access, 9, 134065-134080.
14. 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.
15. Uddin, M. Z., Khaksar, W., & Torresen, J. (2017). Facial expression recognition using salient features and convolutional neural network. IEEE Access, 5, 26146-26161.