ISSN: 2584-0495 Vol. 2, Issue 9, pp. 1187-1196

IJMIT International Journal of Microsystems and IoT

International Journal of Microsystems and IoT

ISSN: (Online) Journal homepage: https://www.ijmit.org

Emotion Detection using Deep Learning CNN Model

Swetha NG, Himasri Allu, KP Hari Chandana, Ujwal Kumar, Naga Sushwar and BL Swathi

Cite as: Swetha, N., Allu, H., Hari Chandana, K., Kumar, U., Sushwar, N., & Swathi, B. (2024). Emotion Detection using Deep Learning CNN Model. International Journal of Microsystems and IoT, 2(9), 1187-1196. https://doi.org/10.5281/zenodo.14099594

Published online: 23 Sept 2024

DOI: https://doi.org/10.5281/zenodo.14099594

International Journal of Microsystems and IoT
Vol. 2, Issue 9, pp.1187-1196; DOI: https://doi.org/10.5281/zenodo.14099594

Emotion Detection using Deep Learning CNN Model

Swetha NG, Himasri Allu, KP Hari Chandana, Ujwal Kumar, Naga Sushwar and BL Swathi

School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Vellore, Tamil Nadu, India

ABSTRACT

Facial Emotion Recognition (FER) is crucial in domains like human-computer interaction, mental health assessment, and marketing. This paper details the design and implementation of a FER model using Deep Convolutional Neural Networks (DCNNs) on the FER2013 dataset, which contains grayscale images labeled with seven emotions. Data augmentation and feature extraction are employed to enhance dataset diversity and reduce dimensionality. The DCNN architecture includes ReLU and Softmax activations for efficient non-linearity and multiclass classification, respectively, with Tanh and LeakyReLU showing promising results. The study explores the impact of pooling layers, identifying an optimal configuration of three layers. Hardware configurations significantly influence performance, with superior accuracy in System 2. Results highlight that balancing activation functions, pooling layers, and hardware specifications is key to optimizing CNN performance.

KEYWORDS

Programmable gain amplifier (PGA); Instrumentation Amplifier (INA); Common-mode-rejection-ratio (CMRR); Complementary Metal Oxide Semiconductor (CMOS); Phase Margin (PM)

1. INTRODUCTION

In recognising emotions, facial expression plays a vital role. Studies show that understanding facial expressions may significantly change one's understanding of spoken words. It also influences the direction of a conversation. The ability to recognise emotions is crucial for effective communication. Facial Emotion is analysed by their facial features such as their eyebrow, eye and mouth coordination. It infers emotions like happiness, sadness, anger, surprise, fear and disgust. Detecting emotions accurately by facial expression is the problem of facial emotion recognition. Creating the models which can effectively recognize these emotions across different individuals, ages, ethnicities, and environmental conditions is the main challenge. Some other factors were also taken into consideration such as facial occlusions, variations in lighting, and subtle nuances in the expression. Facial Emotion Recognition is used in many real-time applications such as human-computer interaction, surveillance, healthcare, and entertainment.

The necessity of addressing the problem of FER comes from its significant effects in a variety of fields, including marketing, security, healthcare, and human-computer interaction. This work is in line with the goal of developing Artificial intelligence (AI) to mimic human social and cognitive abilities. It leads to more meaningful and natural interaction between humans and AI systems. Facial emotion recognition represents a crucial area of study with profound implications across various practical domains. A common example involves the context of mental health services. Particularly those with illnesses like autism spectrum disorder (ASD)— frequently face difficulties in correctly reading facial expressions. Facial emotion identification has great potential to transform how teachers evaluate students' emotional health and engagement in the classroom. Through the analysis of facial expressions, teachers can identify small signs of student behaviours, such as boredom, confusion or anxiety, and modify they accordingly.

There is a scope for improvement in the business sphere in terms of improving customer experiences and market research.. Through the examination of people's facial expressions during interactions, entrepreneurs can use this technology to get insights into how customers react to different stimuli, such as products, advertisements, or services. In a nutshell, there is great promise For the field of facial expression recognition research to revolutionise a wide range of real-world uses, from market research to mental health treatment and education.

This paper consists of five sections namely introduction, related works, implementation, references, conclusion. The introduction consists of background information on the importance of understanding human emotions, and highlights the potential of deep learning techniques to address these challenges. In "Related Works" it consists of reviews and discusses previous studies, papers, and projects that are relevant to the topic of emotion detection using deep learning. In "Implementation" section of the research paper typically describes the technical details of how the proposed method or model was implemented. The "References" section lists all the sources cited within the paper. The "Conclusion" section provides a concise summary of the key findings of the study. It recaps the main results presented in the "Results" section and highlights their significance in relation to the research questions or objectives.

2. RELATED WORK

A two-part CNN for facial emotion recognition (FERC) is presented by the authors [1]. The first part of the CNN removes background, and the second part extracts facial features to create an Expressional Vector (EV) that can identify five different expressions. Tested on 750K images, this two-level architecture outperforms single-level CNNs, demonstrating its versatility.

However, changes in lighting and facial orientations may have an impact on the accuracy of recognition.

The authors [2] developed four emotion classification models: one using SVM with HOG descriptors and three using CNNs with different inputs. Model-3, which combines HOG and pixel data, outperforms others in accuracy and F1 score, highlighting the drawbacks of model-4's down sampling. Techniques like batch normalization, dropout, and L2 regularization are employed to reduce overfitting, with GPU acceleration enhancing training speed.

Using the Affect Net database, the authors [3] present a CNN-based Facial Emotion Recognition (FER) system, with performance measured against accurate recognition rates. The system can accurately identify emotions, but one major drawback is that it heavily relies on the AffectNet dataset, which might not adequately represent the variety of real-world facial expressions, especially in situations with distinct cultural contexts.

The authors [4] introduce Light-FER, a compressed version of the Xception model to address low accuracy and high resource demands in deep learning-based FER systems, especially on edge devices. Pruning and quantization reduce memory usage and optimize inference performance, making Light-FER an efficient solution for FER deployment on resource-limited devices. The dlib 68-facial landmark detector, however, has trouble picking up faces at sharp angles.

The authors [5] developed FLEPNet to address FER challenges like feature extraction and overfitting by using a texture-based feature-level ensemble network with multi-scale DCNNs to classify seven facial expressions. Despite its effectiveness, FLEPNet faces issues with high redundancy of low-level features due to occlusions and posture changes.

Through adaptive window and radial averaging, the authors [6] suggest an improved LBP method for better FER by reducing feature vector length and noise susceptibility. SVM evaluation of this method demonstrates enhanced identification rates in a variety of databases, even under noisy environments. On the other hand, difficult expressions or drastic lighting changes might pose problems for it.

The authors [7] present FER-net, an efficient CNN with a softmax classifier for facial expression differentiation, showing competitive performance on five benchmark datasets compared to 21 methods. However, issues include potential overfitting due to the model's complexity and reliance on large datasets, with performance variability across datasets and real-world scenarios, requiring further validation and optimization.

The authors [8] preprocess facial expressions using Wiener fPSNR values and high MSE filters before using CNN to detect them in seven important locations. CNN feature extraction manages high-dimensional data in a way that is insensitive to skewing, scaling, and distortions. However, when emotions are expressed through other modalities, such as speech or gestures, the method's reliance on facial expressions for emotion recognition may reduce accuracy.

The authors [9] explore advancements in multimodal emotion recognition using Brain-Computer Interfaces (BCIs), integrating EEG, fNIRS, and physiological signals for improved emotion decoding. They review cutting-edge methodologies in data fusion, feature extraction, and machine learning tailored for these tasks, while also addressing challenges in real-time emotion detection. Novel strategies are proposed to overcome obstacles in BCI-based emotion recognition.

The authors [10] present a synopsis of current techniques for detecting facial emotions while pointing out their shortcomings. They go into detail about the training procedure, hyperparameters, and evaluation metrics (F1-score, recall, accuracy, and precision). Test photos visualizations of the model's performance may provide more information about how effective it is.

The author [11] explores human emotions through Ekman's theory of six universal expressions. Electromyography (EMG) is complicated, the Facial Action Coding System (FACS) offers accurate but time-consuming analysis, and automatic systems, while effective, might not be as accurate. Difficulties include small sample sizes, methodological constraints, inconsistent theoretical interpretations, and scant investigation of dynamic aspects.

The authors [12] propose AFERS with face detection, feature extraction using AAM for feature extraction, and skin color detection and lighting compensation, and expression recognition via Euclidean distance. Enhancements with ANFIS achieve nearly 100% accuracy, but further validation is needed across diverse datasets and real-world scenarios.

The authors [13] describe a hybrid CNN method that combines Haar Cascade for face detection and CNN for emotion classification to automatically identify facial emotions. The research contributes to the field of Human-Computer Interaction (HCI) by showcasing high accuracy and minimal loss in emotion recognition. This could find practical uses in mental health screening and social media platforms.

The authors [14] introduce a weakly-supervised Deep Emotion Change Detection (DECD) framework for detecting significant facial expression changes in video sequences, using static facial photos to avoid costly temporal annotations. The framework includes a Multi-Task Emotion Recognizer (MTER) and a changepoint detector to identify key emotion shifts. The study compares DECD's performance with advanced temporal spotting algorithms and validates its effectiveness on three datasets (CASME II, MMI, and YoutubeECD), highlighting the importance of capturing emotional shifts.

The authors [15] utilize a Super-Resolution Generative Adversarial Network (SRGAN) to upscale images by a factor of four, minimizing mean square error (MSE). sing the MediaPipe framework and FACS principles, the model annotates 468 3D facial features, choosing 27 important landmarks to build an emotional mesh that monitors the movements of facial muscles. The next step involves using a K-Nearest Neighbours (KNN) classifier.

The authors [16] compare transfer learning models using MobileNet-V2 and Inception-V3 with a full learning model for facial emotion recognition. They achieve 96% accuracy with Inception-V3 on the Emognition dataset and demonstrate competitive results on JAFFE and KDEF datasets. The study highlights the effectiveness of transfer learning, particularly with Inception-V3, for facial emotion recognition across different datasets, offering insights into optimizing model architecture for improved accuracy and efficiency.

The authors [17] apply Soft Actor-Critic (SAC) to NB-IoT resource allocation, comparing it with DQN and PPO algorithms. Their experiments demonstrate SAC's superior performance in managing NB-IoT resources across various metrics. This study contributes to reinforcement learning applications in IoT network optimization, highlighting SAC's effectiveness for complex resource allocation challenges in NB-IoT environments.

The authors[18] explained significant power consumption, especially in the power-intensive RRC_Connected_ON state. To address this, researchers have explored energy-saving states like extended DRX (eDRX) and Power Saving Mode (PSM), which allow devices to remain in low-power states for extended durations. However, these states have limitations in efficiently handling small or bursty data packets.

The authors[19] used deep learning approaches, especially using CNNs, have showing improved performance in feature extraction and have explored hybrid models like CNN-RNN or CNN-SVM for better accuracy in HAR tasks. The combination of pre-trained CNNs (such as VGG16) with classifiers like SVM has proven effective in leveraging both deep feature learning and robust classification.

The authors[20] used hybrid models combining CNNs with RNNs or SVMs have been explored to improve accuracy. The use of pre-trained CNNs such as VGG16 with SVM classifiers has demonstrated effective integration of deep learning and robust classification techniques

3. MODEL DESIGN

3.1 Dataset Description

1. **Data Type**: Image data.
2.**Number of Samples**: The dataset contains a total of 28,709 images.
3. **Features/Columns:** Each image is represented as a pixel array, with the dimensions of the images depending on the resolution used for capturing.
4. **Target Variable**: The target variable will typically be facial expression depicted in each image. Emotions could include categories like happy, sad, angry, surprised, etc.
5. **Source:** The dataset's source is taken from KaggleFER2013. link:https://www.kaggle.com/datasets/msambare/fer2013

3.2 Cleaning and Pre-processing

Data Cleaning:

1. Removing duplicate Images: duplicate images are identified and removed from the dataset. Duplicate images can skew the analysis by introducing redundancy in the dataset. Removing duplicates helps in improving the efficiency and prevents bias in model training.
2. Handling corrupt Images: detection and removal of corrupted files that are corrupted and unavailable due to various reasons were removed. Because they can result in mistakes or inconsistencies.
3.Quality Control: images that are blurry, overly compressed, and flawed are identified and removed from the dataset which can make analysis and model training incorrect.

Data Pre-processing:

1.Resizing and Rescaling: This step involves standardizing image size and scale to ensure uniformity and compatibility with analysis or models. By improving image dimensions, this process helps the model quickly identify patterns and reduces computational complexity and memory usage.
2.Data Augmentation: This technique involves transforming images through rotation, flipping, grayscale conversion, cropping, and zooming. Augmenting the data enhances dataset diversity, reduces overfitting.
3. Feature Extraction: This process extracts relevant features from images to lower dimensionality and simplify analysis. Techniques like edge detection, Histogram of Oriented Gradients (HOG), and CNN-based methods are used.

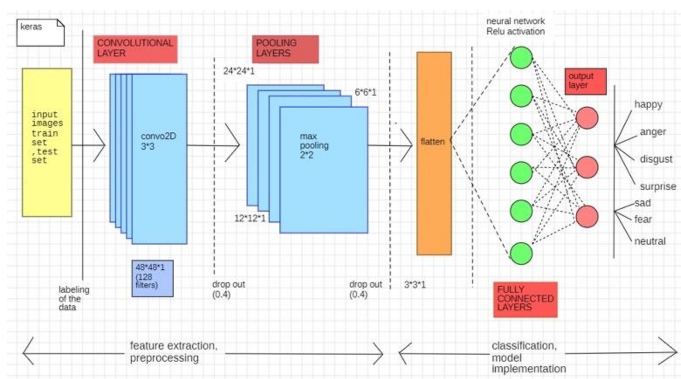


Fig 1. CNN Architecture Diagram

The FER2013 dataset on Kaggle includes 48x48 pixel grayscale images of facial expressions, labeled with seven emotions: anger, disgust, fear, happiness, sadness, surprise, and neutral. We used a Deep Convolutional Neural Network (DCNN) to classify these images. The network processes images through multiple convolutional layers, reducing dimensions from 48x48x1 to 24x24x1 by applying small kernels. It includes 10 pooling layers for further dimension reduction and flattens the output to 3x3x1 for fully connected layers. PReLu activation functions were used in these layers, and a Softmax output layer normalized the results across the seven emotion classes.

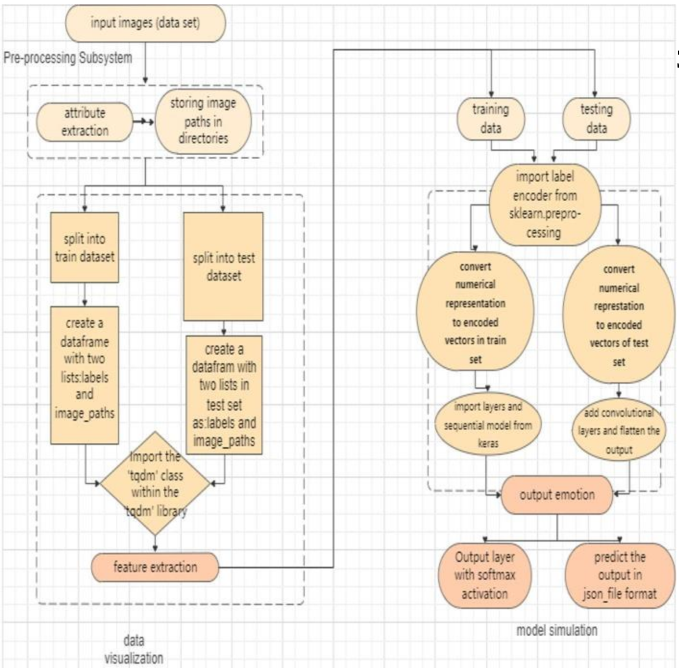


Fig 2. Flow Diagram of Action Recognition

The flow chart in facial emotion recognition model typically illustrates the flow of data and operations within a system

The images are taken as inputs from the data set and pre-processed by attribute extractions and stored into directories before splitting it into testing and training sets. After the split each testing set and the training set creates a data frame with two lists i.e.,Labels and images paths. In the next step we complete the feature extraction by importing tqdm class within the tqdm library and then one hot encoding is performed to convert numerical representation to encoded vectors before outputting the emotion.

3.3 Activation Function

Activation functions are vital in artificial neural networks, introducing non-linearity that enables the model to learn complex patterns and relationships. They determine a neuron's output and whether it should be activated. Key activation functions include:

Sigmoid: Outputs values between 0 and 1, useful for binary classification.

Tanh: Similar to sigmoid but ranges from -1 to 1, addressing the vanishing gradient issue.

ReLU (Rectified Linear Unit): Activates neurons for positive inputs, enhancing computational efficiency.

Leaky ReLU: Allows a small gradient for negative inputs, solving the dying ReLU problem.

Softmax: Produces probabilities for multiclass classification.

3.3.1 Hidden Layer Activation Functions

Choice: Rectified Linear Unit (ReLU) was chosen for the hidden layers.

Rationale: ReLU introduces non-linearity to the model efficiently by allowing positive values to pass through unchanged. This simplicity often leads to faster convergence during training, and the avoidance of vanishing gradient issues.

3.3.2 Output Layer Activation Functions

1. **Multiclass Classification (Softmax):**

   Choice: Softmax activation was selected for the output layer in the context of multiclass facial emotion recognition.

   Rationale: Softmax normalizes the outputs across multiple classes, producing a probability distribution over all classes. This is fundamental in multiclass scenarios

   Binary Classification (Sigmoid):

   Alternative: Sigmoid activation is considered when the task involves binary classification (e.g., distinguishing between positive and negative emotions).

   Rationale: Sigmoid squashes the output between 0 and 1, making it suitable for binary decisions.

2. **Addressing the "Dying ReLU" Problem:**

   Choice: Leaky ReLU or variants could be considered.

   Rationale: To mitigate the "dying ReLU" problem where neurons may become inactive during training, Leaky ReLU introduces a small slope for negative inputs, allowing a continuous flow of information through the network.

3. **Experimentation and Model Evaluation:**

   Consideration: The choice of activation functions involved experimentation and evaluation during the model development phase.

4. **Other Activation Functions:**

   Tanh: Hyperbolic Tangent can be considered, especially in hidden layers, as it is zero-centred and may help mitigate issues related to shifting the mean of the activations.

   Parametric ReLU (PReLU): PReLU introduces a learnable parameter to Leaky ReLU, allowing the network to adaptively learn the slope for negative inputs.

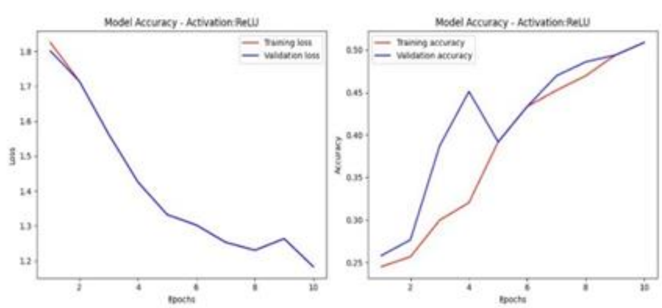


Fig 3. Loss and Accuracy with ReLU activation function

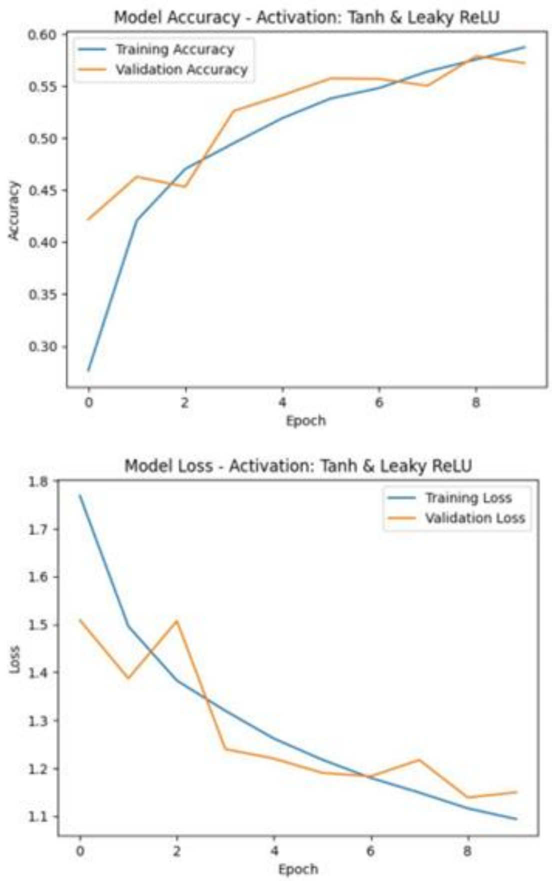


Fig 4. Loss and Accuracy of Tanh and Leaky ReLU

3.4 Summary and Choice

ReLU activation is more computationally efficient than Tanh and Leaky ReLU in our case, due to its simplicity and speed, which is advantageous for deep neural networks. ReLU's efficiency often outweighs its potential issues, such as the "dying ReLU" problem and sensitivity to outliers. While Tanh helps with mean-shifting issues by being zero-centered, and Leaky ReLU addresses the "dying ReLU" problem by allowing a small gradient for negative inputs, ReLU's performance remains satisfactory and efficient. The choice of activation function balances model simplicity, convergence speed, and the ability to capture complex patterns in facial expressions, guided by both theoretical considerations and practical performance.

3.5 Summary and Choice

3.5.1 Pooling Layers

Pooling layers are crucial in CNNs for performing feature extraction and down sampling. They operate on feature maps from convolutional layers, which contain scattered activations across the image. Pooling layers reduce the spatial dimensions of these feature maps while retaining significant information.

1. **Max Pooling:** This technique selects the maximum value from each local region of the feature map, preserving the most critical elements and discarding less important ones.

2. **Average Pooling:** This method averages the values in each local region of the feature map to determine the output, which can be useful for reducing noise.

Table 1. Pooling layers with accuracies

POOLING LAYERS	ACCURACY
2	0.5386
3	0.5770
4	0.5492
5	0.5331
6	0.5496
7	0.5348
8	0.5535
9	0.5368
10	0.5032

Layer 1 Pooling: This initial pooling layer, following the first convolutional layer, reduces the spatial dimensions of feature maps while preserving essential information. The choice of pooling operation (e.g., max or average pooling) and window size depends on the network architecture and task requirements.

Layer 2 Pooling: The second pooling layer further decreases the feature maps' dimensions after another set of convolutional layers. It helps retain crucial features for facial emotion recognition while discarding less important details.

Layer 3 Pooling: Continuing the down-sampling process, Layer 3 expands the network's receptive field, enabling it to capture more global information from the image.

Layer 4 Pooling: This layer further reduces the dimensions of the feature maps, summarizing hierarchical features.

Layer 5 Pooling: Located deeper in the network, Layer 5 continues dimensionality reduction, focusing on prominent features to improve resilience.

Layer 6 Pooling: It further down-samples feature maps while retaining significant features.

Layer 7 Pooling: Possibly the final pooling layer, Layer 7 ensures high-level features critical for facial emotion recognition are preserved before the data moves to fully connected layers.

Layer 8 Pooling: If an eighth pooling layer were present, it would continue the down-sampling process, further improving the network's ability to extract relevant information. However, such a deep network with many pooling layers is uncommon.

Layer 9 and 10 Pooling: Similarly, these pooling layers would continue reducing the spatial dimensions of feature maps, which by this point would be quite abstract. Excessive pooling could lead to significant detail loss, especially if precise facial feature localization is needed.

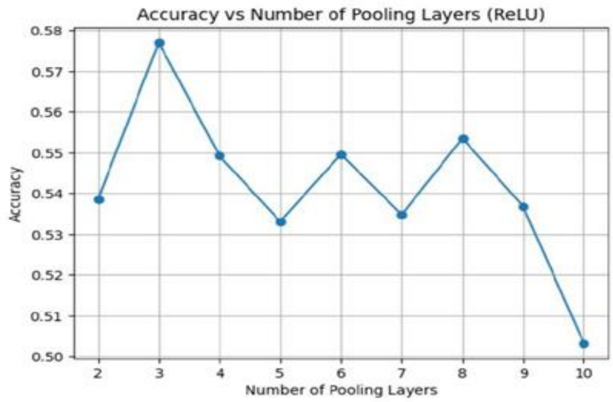


Fig 5. Accuracy of pooling layers

4. RESULT AND DISCUSSION

Parameters of Evaluation

1. Testing Accuracy

This parameter measures the accuracy of the model on a separate dataset, called the test dataset. It indicates how well the model generalises to new, unseen data. Test accuracy is important because it reflects the model's ability to make accurate predictions in real-world situations.

2. Training Accuracy

This parameter measures the accuracy of the model on the training dataset, which is the data used to train the model. It indicates how well the model fits the training data. While high training accuracy is desirable, it is important to ensure that the model does not overfit the training data and can generalise well to new data.

Activation Functions Used

ReLu, Softmax, LeakyReLu and Tanh were used.

Table 2 . Impact of Activation Functions (ReLu, Softmax, Tanh and LeakyRelu) on CNN Accuracy for Image Classification (10 Epochs)

Activation Methods	**Testing Accuracy**	**Training Accuracy**
ReLu and Softmax	52.08%	49.10%
Tanh, LeakyRelu and Softmax	57.18%	58.70%

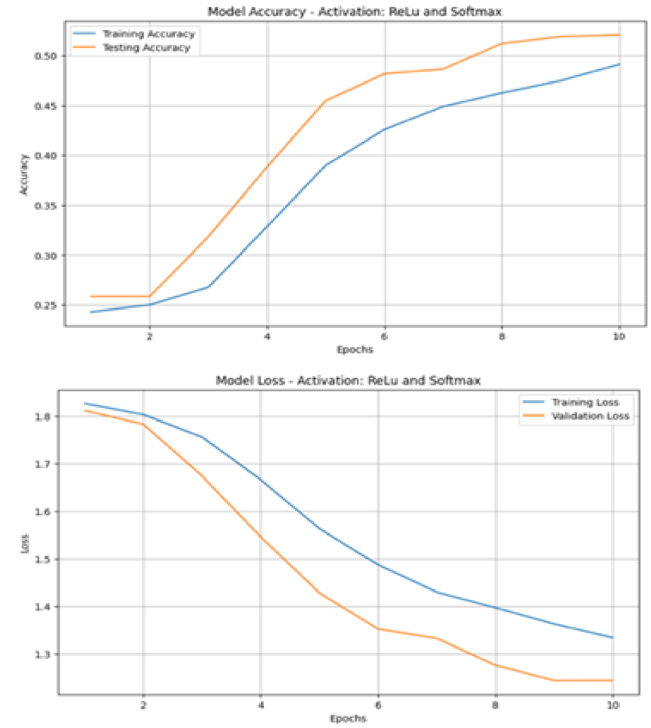


Fig 6. Training progress (10 epochs) correlates with improvement in both accuracy and loss, suggesting successful model training.

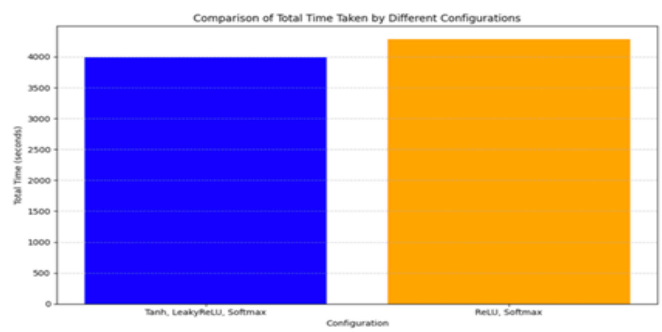


Fig 7. Comparison of total training time of each activation function

We can observe from graph 4 that the total training time taken by the model (Tanh, LeakyReLu and Softmax) gives us better time efficiency as it took 395 seconds less time to train for 10 epochs then the model (ReLu and softmax).

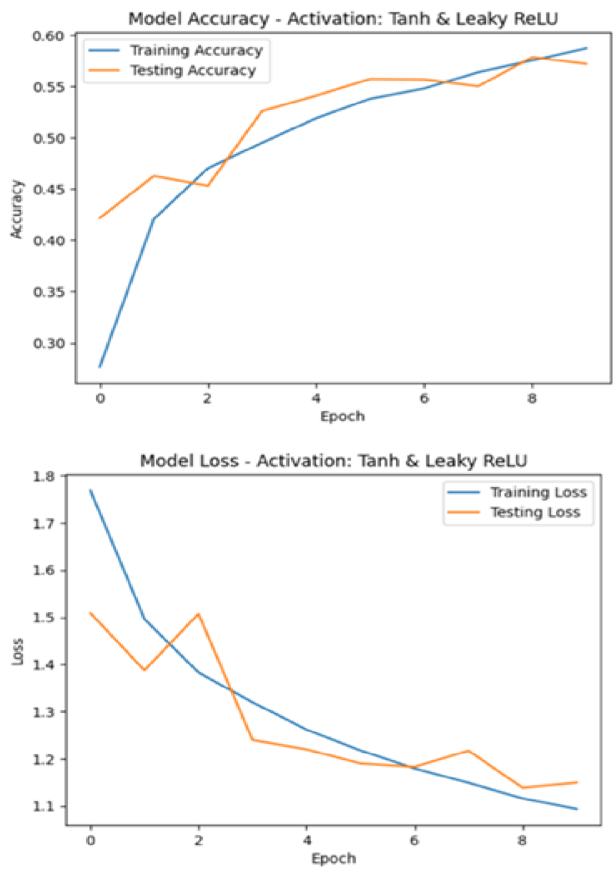


Fig 8. Accuracy (Tanh & Leaky ReLU) increases steadily over epochs, while Loss (Tanh & Leaky ReLU) decreases, suggesting successful model training

We can observe from Table 1 that the activation methods – Tanh, LeakyReLu and Softmax give us better results in terms of accuracy.

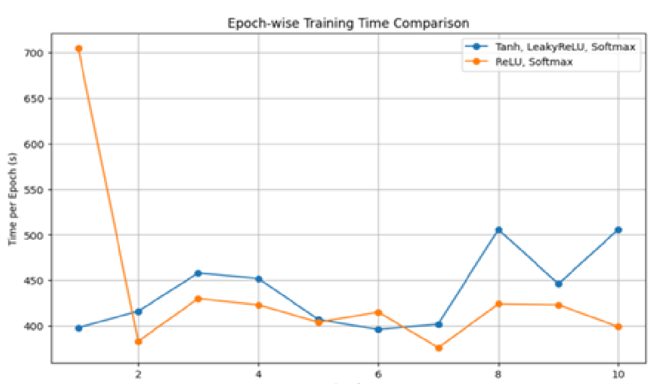


Fig 9. Epoch wise training time for each activation function

Variation of accuracy in correspondence to the no. of Pooling Layers

Table 3 . Accuracy Comparison of Pooling Layers (in CNN for Image Classification

Pooling Layers	Testing Accuracy	Training Accuracy
Pooling Layer 2	53.86%	53.11%
Pooling Layer 3	57.70%	55.41%
Pooling Layer 4	52.08%	49.10%
Pooling Layer 5	53.31%	49.22%
Pooling Layer 6	54.96%	50.62%
Pooling Layer 7	53.48%	49.30%
Pooling Layer 8	55.35%	50.53%
Pooling Layer 9	53.68%	48.68%
Pooling Layer 10	50.05%	54.78%

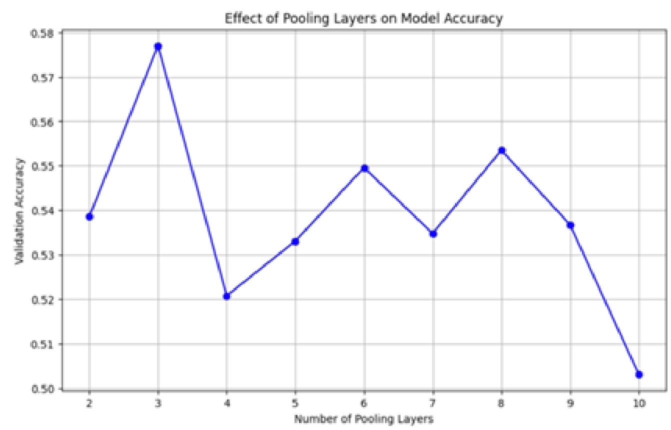


Fig 10. Effect of pooling layers on Model Accuracy

We Can observe from table 2 that pooling layer 3 gives us the best accuracy among all so it is ideal to run this model using 3 pooling layers.

Performance Variation Across Hardware and Software Configurations

Table 4. Accuracy of the CNN model across two different systems.

Systems	Testing Accuracy	Training Accuracy
System 1	52.08%	49.10%
System 2	54.37%	50.32%

Table 5 . System Specifications of System 1 and System 2.

	System 1	System 2
Processor	**1.4 GHz Quad-Core Intel Core i5**	**Apple M2 Chip**
Graphics	**Intel Iris Plus Graphics 645 1536 MB**	**10 core GPU**
Memory	**8 GB 2133 MHz**	**8 GB 1398 MHz**
GPU	**4 Core**	**8 Core**

Table 6 . Comparative analysis with other articles

Paper	Model	Accuracy
Measuring facial expression of emotion	DCNN	51%
Advances in Multimodal Emotion Recognition Based on Brain-Computer Interfaces	BCI/CNN	52%
A novel approach for facial expression recognition using local binary pattern with adaptive window	LBP/CNN	54%
A Lightweight Facial Emotion Recognition System on Edge Devices	CNN	40%
Proposed work	CNN	54.37%

Our study underscores the critical impact of activation functions and pooling layers on CNN accuracy for facial emotion recognition. Tanh, LeakyReLU, and Softmax activations yield higher accuracy and better time efficiency. Employing three pooling layers optimises testing accuracy, and hardware configurations affect model performance, with System 2 outperforming System 1. This emphasises the need to carefully balance activation functions, pooling layers, and hardware specifications for optimal CNN performance in emotion recognition.

5. CONCLUSION AND FUTURE WORK

In conclusion, our research highlights how important pooling layers and activation functions are to improving CNN accuracy in detecting face emotions. Tanh, LeakyReLU, and Softmax activations yielded results that were more accurate and efficient with less time. Three pooling layers improved testing accuracy, System 2 exceeded System 1 in terms of model performance, which was greatly impacted by hardware configurations. This emphasises the necessity of carefully balancing these elements in order to have CNN operate at its best when it comes to emotion recognition tasks.

Through hyperparameter adjustment, factors such as activation functions and pooling layers could be optimised in future work to improve model performance. Human-computer interaction, sentiment analysis in marketing, and healthcare could all benefit from the use of these models. Furthermore, investigating transfer learning techniques may improve model performance, particularly in situations where there is a deficiency of labelled data. To increase accuracy and resilience in facial emotion identification tasks, more studies should look into innovative designs or ensemble methods.

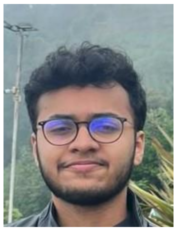
REFERENCES

1. Mehendale (2020). N. Facial emotion recognition using convolutional neural networks (FERC). SN Appl. Sci. 2, 446 https://doi.org/10.1007/s42452-020-2234-1
2. Singh V., Singh, R.K., Tiwary, U.S. (2023). Emotion Classification Through Facial Expressions Using Support Vector Machine (SVM) and Convolutional Neural Classifier. In: Zaynidinov, H., Singh, M., Tiwary, U.S., Singh, D. (eds) Intelligent Human Computer Interaction. IHCI 2022. Lecture Notes in Computer Science, vol 13741. Springer, Cham. https://doi.org/10.1007/978-3-031-27199-1_13
3. Do, H.N. et al. (2020). Automatic Facial Expression Recognition System Using Convolutional Neural Networks. In: Van Toi , V., Le, T., Ngo, H., Nguyen, TH. (eds) 7th International Conference on the Development of Biomedical Engineering in Vietnam (BME7). BME 2018. IFMBE Proceedings, vol 69. Springer, Singapore. https://doi.org/10.1007/978-981-13-5859-3_82
4. Pascual AM, Valverde EC, Kim J-i, Jeong J-W, Jung Y, Kim S-H, Lim W. (2022). Light-FER: A Lightweight Facial Emotion Recognition System on Edge Devices. Sensors.; 22(23):9524. https://doi.org/10.3390/s22239524
5. M. Karnati, A. Seal, A. Yazidi and O. Krejcar. (2022). FLEPNet: Feature Level Ensemble Parallel Network for Facial Expression Recognition, in IEEE Transactions on Affective Computing, vol. 13, no. 4, pp. 2058-2070 https://doi.org/10.1109/TAFFC.2022.3208309
6. Kola, D.G.R., Samayamantula, S.K. (2021). A novel approach for facial expression recognition using local binary pattern with adaptive window. Multimed Tools Appl 80, 2243–2262 https://doi.org/10.1007/s11042-020-09663-2

7. Mohan, K., Seal, A., Krejcar, O. et al. (2021). FER-net: facial expression recognition using deep neural net. Neural Comput & Applic 33, 9125–9136 https://doi.org/10.1007/s00521-020-05676-y
8. Monisha, G. S. et al. (2023). Enhanced Automatic Recognition of Human Emotions Using Machine Learning Techniques. Procedia Computer Science, 218, 1877-0509. https://doi.org/10.1016/j.procs.2023.01.020
9. He Z, Li Z, Yang F, Wang L, Li J, Zhou C, Pan J. (2020). Advances in Multimodal Emotion Recognition Based on Brain–Compu ter Interfaces. Brain Sciences., *10*, 687. https://doi.org/10.3390/brainsci10100687
10. Jaiswal, A., Krishnama Raju, A., & Deb, S. (2020). Facial Emotion Detection Using Deep Learning. *2020 International Conference for Emerging Technology (INCET)*, 1-5. https://doi.org/10.1109/INCET49848.2020.9154121
11. Wolf, Karsten. (2015). Measuring facial expression of emotion. Dialogues in Clinical Neuroscience, 17(4):457-62. https://doi.org/10.31887/DCNS.2015.17.4/kwolf
12. Vijayanand. G, Karthick. S, Hari. B, Jaikrishnan.V. (2020). Emotion Detection using Machine Learning, International Journal of Engineering Research & Technology (IJERT) NCICT https://doi.org/10.17577/IJERTCONV8IS08017
13. Sandhu, Nehmat & Malhotra, Aksh & Bedi, Mandeep. (2020). Human Emotions Detection Using Hybrid CNN Approach. International Journal of Computer Science and Mobile Computing. https://doi.org/10.47760/ijcsmc.2020.v09i10.001
14. ByungOk Han, Cheol-Hwan Yoo, Ho-Won Kim, Jang-Hee Yoo, Jinhyeok Jang. (2023). Deep emotion change detection via facial expression analysis, Neurocomputing, Volume 549, 0925-2312. https://doi.org/10.1016/j.neucom.2023.126439
15. Siam AI, Soliman NF, Algarni AD, Abd El-Samie FE, Sedik A. (2022). Deploying Machine Learning Techniques for Human Emotion Detection. Comput Intell Neurosci. https://doi.org/10.1155/2022/8032673
16. Agung, E. S., Rifai, A. P., & Wijayanto, T. (2024). Image-based facial emotion recognition using convolutional neural network on emognition dataset. Scientific reports, 14(1), 14429. https://doi.org/10.1038/s41598-024-65276-x
17. Anbazhagan, S., & Mugelan, R. K. (2024). Next-gen resource optimization in NB-IoT networks: Harnessing soft actor–critic reinforcement learning. Computer Networks, 252, 110670. https://doi.org/10.1016/j.comnet.2024.110670
18. Anbazhagan, S., & Mugelan, R. K. (2024). Energy efficiency optimization of NB-IoT using integrated Proxy & ERAI technique. Results in Engineering, 23, 102419. https://doi.org/10.1016/j.rineng.2024.102419
19. Athavale, V. A., Kumar, D., & Gupta, S. C. (2021). Human Action Recognition Using CNN-SVM Model. Advances in Science and Technology, 105, 282-290. https://doi.org/10.4028/www.scientific.net/AST.105.282
20. Tufek, N., Yalcin, M., Altintas, M., Kalaoglu, F., Li, Y., & Bahadir, S. K. (2020). Human Action Recognition Using Deep Learning Methods on Limited Sensory Data. IEEE Sensors Journal, 20(6), 3101-3112. https://doi.org/10.1109/JSEN.2019.2956901

AUTHORS

Dr. Swetha NG. is currently working as an Assistant Professor Senior in the department of Analytics, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore. She has completed her Masters from PSG College of Technology and her Ph.D from Anna University, Chennai. Her areas of research include Parallel Computing, Multi Criteria Decision Making, Quantum Computing and Machine Learning.
E-mail: swetha.ng@vit.ac.in



Ujwal Kumar is pursuing his Bachelor's of Technology in Computer Science and Engineering with a specialization in Blockchain from Vellore Institute of Technology, Vellore, India. He has a keen interest in the fundamentals and applications of blockchain technology, with a focus on building decentralized and tamper-proof systems. His research interests include the use of blockchain for creating secure digital identities, streamlining transactions in financial systems, and improving the efficiency of supply chain management. He is dedicated to exploring innovative blockchain solutions that address current challenges in data privacy, trust, and security across various industries.
E-mail: ujwal.kumar2021@vitstudent.ac.in

KP Hari Chandana is pursuing her Bachelor's of Technology in Computer Science and Engineering with specialization in Data Science from Vellore Institute of Technology, Vellore, India. She has a keen interest in deep learning, particularly in developing and optimizing neural networks to solve complex problems. Her focus lies in exploring cutting-edge deep learning architectures and algorithms, aiming to enhance model performance and adaptability across various applications.
E-mail: penchala.harichand2021@vitstudent.ac.in

BL Swathi is pursuing her Bachelor's of Technology in Computer Science and Engineering with specialization in Data Science from Vellore Institute of Technology, Vellore, India. She has a keen interest in machine learning, focusing on the development of intelligent systems that can learn from data and improve over time. Her research interests include exploring various machine learning algorithms, optimizing model performance, and applying these techniques to solve real-world problems across diverse domains.
E-mail: bl.swathi2021@vitstudent.ac.in

N Naga Sushwar is pursuing her Bachelor's of Technology in Computer Science and Engineering with a specialization in Blockchain from Vellore Institute of Technology, Vellore, India. He has a keen interest in blockchain technology, particularly in its applications for secure and decentralized systems. His focus lies in exploring the potential of blockchain for enhancing transparency, security, and efficiency in various domains, such as finance, supply chain management, and data integrity. He is passionate about developing innovative solutions that leverage blockchain's unique features to address complex challenges.
E-mail: nichenametla.sushwar2021@vitstudent.ac.in

Himasri Allu is pursuing her Bacherlor's of Technology in Computer Science and Engineering from Vellore Institute of Technology, Vellore, India. She has a keen interest in exploring new methodologies for data preprocessing and feature engineering, as well as developing robust models that can adapt to evolving data patterns.
E-mail: himasri.allu2021@vitstudent.ac.in