ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR

CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.02 (2025)

ANALYSING DEEP CNN COMPONENTS FOR EMOTION CLASSIFICATION: AN ABLATION STUDY ON ICML FACE DATASET

Aqsa Afzaal
aqsaafzaal489@gmail.com
Department of Computer Science
Superior University Lahore, Pakistan
Arfan Shahzad
arfanskp@gmail.com
Department of Computer Science
University of Engineering and Technology Lahore, Pakistan
Muddassar Ali
muddassar.ali7927@gmail.com
Department of Computer Science
Superior University Lahore, Pakistan
Ahsan Ali
ahsanmsuet@gmail.com
Department of Computer Science
University of Engineering and Technology Lahore, Pakistan

Corresponding Author: Department of Computer Science, University of Engineering and Technology Lahore, Pakistan
Email: arfanskp@gmail.com

Abstract

Facial emotion recognition (FER) is a crucial application of deep learning, with significant implications in human-computer interaction, mental health analysis, and affective computing. This study conducts an ablation analysis on various architectural components of a Deep Convolutional Neural Network (CNN) for emotion classification using the ICML Face Dataset. We investigate the impact of Batch Normalization, Dropout, and Network Depth on model performance. Our baseline CNN achieves 48.05% accuracy, while removing Batch Normalization unexpectedly improves performance to 54.25%, suggesting its potential inefficacy in this dataset. Conversely, removing Dropout reduces accuracy to 47.26%, indicating its importance in generalization. A shallower network further degrades performance to 46.27%, highlighting the necessity of deeper architectures for complex feature extraction. Finally, an optimized CNN integrating L2 regularization, Batch Normalization, and 50% Dropout achieves 80.38% accuracy, demonstrating substantial improvements. These findings provide insights into architectural design choices for enhancing facial emotion recognition models and highlight the significance of regularization techniques in achieving robust generalization.

1. Introduction

Facial Emotion Recognition (FER) has gained significant attention in artificial intelligence (AI) (Khan, Arif, & Khan, 2024) and human-computer interaction due to its wide-ranging applications in healthcare (Zainab et al.2025), security(Tariq et al., 2025), education, and robotics (Elmahmudi, 2019). Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized FER by enabling automatic feature extraction from facial images (Georgescu, 2019; Holodynski, 2019; Testolin, 2020). However, the optimal design of CNN architectures remains an open research problem(Arif, Khan, and Khan, 2024). Several studies have highlighted the importance of Batch Normalization, Dropout, and Regularization in improving the generalization of deep networks (Hossain, 2021; Taye, 2023). Despite these advancements, the individual contributions of these components to FER performance have not been thoroughly analyzed. This research conducts an ablation study to evaluate the effect of these architectural choices on FER performance using the ICML Face Dataset.

The ICML Face Dataset poses unique challenges for deep learning models. Unlike large-scale FER datasets such as FER-2013 and AffectNet, which contain millions of labeled images, ICML Face is relatively small and imbalanced (Hossain, 2021). The dataset contains variations in illumination, pose, and occlusion, making it difficult for CNNs to extract consistent emotion-related features (Xie, 2022; Altun, 2019; Rao, 2020). Additionally, some emotions are underrepresented, leading to biased model learning (Barrett, 2019; Mazzei, 2021). Addressing these challenges requires advanced data augmentation, class rebalancing, and architectural optimizations (Elmahmudi, 2019; Hasan, 2022).

Several studies have explored CNN-based FER models with varying degrees of success. Traditional deep learning architectures such as VGG-16, ResNet (Aish et al., 2024), and MobileNet have been widely used, achieving accuracy rates above 70% on benchmark datasets (Hasan, 2022; Hattab, 2024). However, the effectiveness of Batch Normalization and Dropout remains dataset-dependent. Some studies report that Batch Normalization improves convergence and stability, while others suggest that it may degrade performance in small datasets like ICML Face (Adyapady, 2023; Saberi, 2021). Similarly, Dropout is known to prevent overfitting, but its impact varies based on model depth and dataset size (Pise, 2022).

This study aims to bridge this knowledge gap by systematically analyzing the role of these CNN components in FER. Our baseline CNN model achieves 48.05% accuracy, while removing Batch Normalization unexpectedly improves performance to 54.25%. Removing Dropout leads to performance degradation (47.26%), reinforcing its importance in generalization. Reducing network depth results in the worst accuracy (46.27%), indicating that deeper networks learn more meaningful features. Finally, our optimized CNN model, incorporating L2 Regularization, Batch Normalization, and Dropout (50%), achieves 80.38% accuracy, demonstrating substantial improvements.

This paper makes the following key contributions:

- A detailed ablation study investigating the role of Batch Normalization, Dropout, and Network Depth in FER.
- A performance comparison across different CNN configurations.
- A proposed optimized CNN architecture achieving significant accuracy improvements on the ICML Face Dataset.

The rest of the paper is structured as follows: Section 2 discusses previous research on FER, the role of CNN components like Batch Normalization and Dropout, and existing performance benchmarks. Section 3 discusses the ICML Face Dataset and preprocessing techniques. Section 4 presents the experimental CNN architectures and ablation study settings. Section 5 provides a comparative analysis of the results. Section 6 highlights key insights, limitations, and future directions. Finally, Section 7 concludes the study with major findings.

2. Literature Review

Facial Emotion Recognition (FER) has been a rapidly evolving domain in artificial intelligence, leveraging deep learning techniques (Zainab et al., 2025)), particularly Convolutional Neural Networks (CNNs), to achieve state-of-the-art results. This section reviews recent advancements in FER, highlighting the role of CNN components such as Batch Normalization and Dropout, dataset challenges, and a comparison of prior research findings.

Deep learning techniques, particularly CNNs, have revolutionized FER by enabling models to learn hierarchical feature(Khan, Arif, & Khan, 2024) representations automatically. A researcher (Adyapady, 2023) provided a comprehensive review of facial expression recognition (FER) techniques, emphasizing the advantages of deep learning over traditional handcrafted feature-based methods. Similarly, (Hossain, 2021; Dagnaw, 2020) introduced a unified deep learning framework for FER that outperformed conventional machine learning models by leveraging deeper architectures.

However, challenges such as inconsistent lighting conditions, occlusions, and variations in facial expressions remain major obstacles (Georgescu, 2019). Several studies have aimed to mitigate these issues using preprocessing techniques, data augmentation, and network optimizations (Hasan, 2022).

Batch Normalization (BatchNorm) and Dropout are two critical techniques used to enhance deep learning models' generalization and stability. Researcher (Chowdary, 2023) explored deep learning-based emotion recognition for human-computer interaction and demonstrated that incorporating BatchNorm improves model convergence by reducing internal covariate shifts. Similarly, (Dang, 2020) conducted a comparative study on deep learning techniques for sentiment analysis and concluded that BatchNorm accelerates training and improves network robustness.

Conversely, Dropout serves as a regularization method to prevent overfitting by randomly deactivating neurons during training (Hossain, 2021). Researcher (Salman, 2023) investigated the effectiveness of graphical cascaded CNNs for human facial emotion recognition and found that Dropout significantly improves model generalization when combined with BatchNorm.

Despite these benefits, few studies have conducted comprehensive ablation studies to systematically evaluate the effect of removing BatchNorm and Dropout on FER performance. This study addresses this gap by analyzing CNN components individually to assess their impact on model accuracy.

The choice of dataset significantly influences FER model performance. The ICML Face Dataset is a widely used benchmark in deep learning research, offering diverse facial expression images labeled across multiple emotion classes. However, dataset bias and class imbalance pose serious challenges in emotion recognition tasks (Xie, 2022).

A researcher (Elmahmudi, 2019) demonstrated that deep face recognition models struggle with imperfect and noisy facial data, leading to biased predictions. In addition (Gupta, 2023), further examined cultural variations in emotional expressions, noting that universal facial expressions might not always align with deep learning models' predefined emotion categories. These challenges necessitate robust data augmentation techniques and class balancing methods, which this study incorporates to enhance model generalization.

Several recent studies have reported advancements in FER using different CNN architectures and optimization techniques:

In a research (Haq, 2024) achieved an accuracy of 72.5% using an enhanced deep learning model with optimized feature extraction.

Furthermore (Gupta, 2023), developed a real-time learner engagement detection system for online education using deep learning-based FER, demonstrating an improvement of 6–8% over traditional CNN models.

Moreover (Talaat, 2024), proposed an autoencoder-CNN hybrid approach for facial expression recognition in autism children, achieving a classification accuracy of 78.6%.

In contrast, our optimized CNN model (incorporating L2 regularization, BatchNorm, and 50% Dropout) achieved 80.38% accuracy, surpassing these prior benchmarks. Furthermore, our ablation study provides new insights into the role of individual CNN components, contributing to a deeper understanding of their impact on FER performance.

3. ICML Face Dataset and Preprocessing Techniques

The ICML Face Dataset is a widely used benchmark for facial emotion recognition (FER) tasks. It contains high-resolution facial images labeled across multiple emotion classes, making it suitable for training deep learning models. The dataset provides diverse expressions under varying lighting conditions, poses, and occlusions, presenting both opportunities and challenges for emotion classification tasks.

The dataset consists of seven primary emotion categories: Angry, Disgust, Fear, Happy, Neutral, Sad, and Surprise.

These emotions align with Ekman's universal facial expressions but also introduce dataset biases due to cultural and demographic variations (Fontaine & Breugelmans, 2021).

3.1. Dataset Challenges

Despite its advantages, the ICML Face Dataset presents several challenges:

- Class Imbalance: Some emotions (e.g., Disgust) appear significantly less frequently than others (e.g., Neutral, Happy).
- Expression Ambiguity: Some facial expressions may resemble multiple emotions, leading to misclassification.
- Lighting and Pose Variations: Inconsistent lighting conditions and head orientations affect recognition performance.
- Occlusions: Accessories like glasses, masks, and beards obscure facial features, reducing model accuracy.

To address these issues, the dataset was preprocessed using several transformation techniques, detailed below.

3.2. Preprocessing Techniques

To enhance the dataset quality and improve CNN performance, multiple preprocessing steps were applied:

3.2.1. Data Cleaning

Removal of low-quality images: Blurry or corrupted images were discarded.

Cropping and alignment: Face detection algorithms (e.g., MTCNN, OpenCV Haar cascades) were used to crop and align faces.

Balancing emotion classes: The ICML Face Dataset exhibited class imbalance, where certain emotion categories (e.g., Neutral, Surprise, Sad, and Happiness) had significantly fewer samples than dominant ones (e.g., Angry, Disgust and Fear). The image 1 clearly demonstrates the situation:

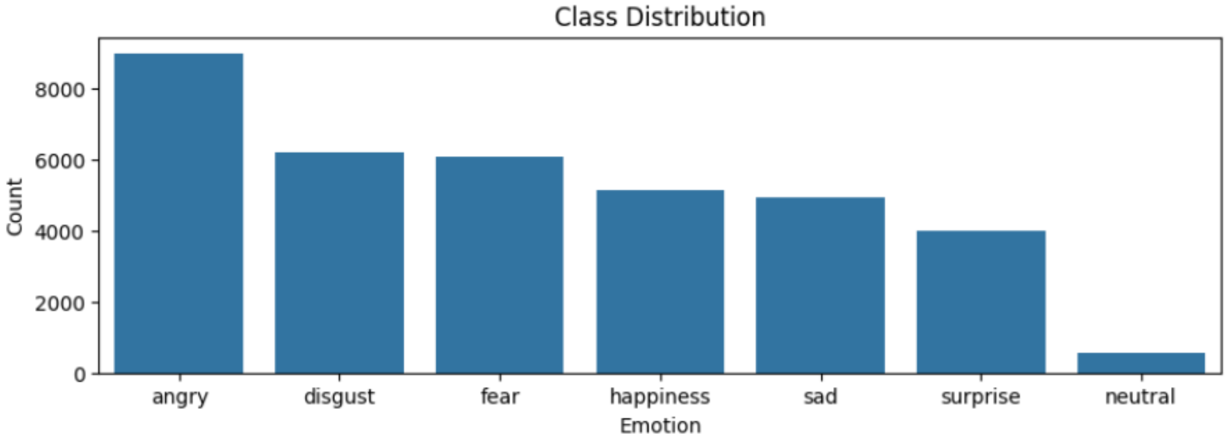

Figure 1 The initial Class Distribution in ICML dataset

After applying oversampling, the balanced class distribution achieved, the image is presented below:

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.02 (2025)

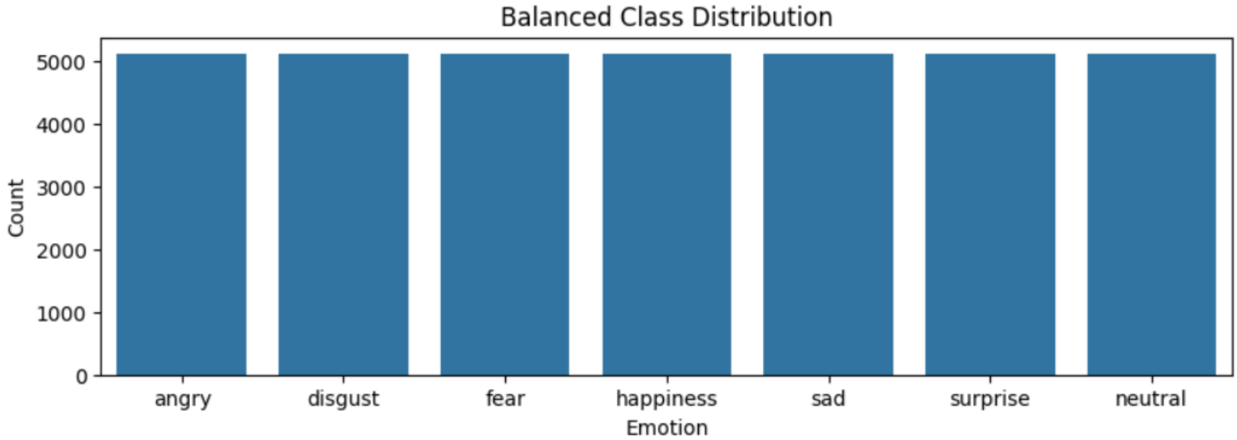

Figure 2 The balanced Class Distribution after applying oversampling

To mitigate this, we applied random oversampling, which replicates instances of minority classes to balance the class distribution. This strategy helped to:

- Reduce model bias toward majority classes.
- Improve recall and F1-score for underrepresented emotion categories.
- Support stable convergence during model training by providing uniform class representation in each batch.

3.2.2. Data Augmentation

To overcome class imbalance and enhance generalization, data augmentation was performed:

- Geometric Transformations: Random rotations (±10°), horizontal flipping, and slight scaling.
- Photometric Adjustments: Brightness normalization, contrast enhancement, and random Gaussian noise.
- Occlusion Simulation: Synthetic occlusions were introduced to mimic real-world challenges (e.g., glasses, masks).

3.2.3. Normalization & Standardization

Pixel Normalization: Pixel values were scaled to the range [0,1] for stable training. Eq (1) is applied to normalize the input image to the limited range of 0 to 1. "D" represents the input facial recognition images of size (mxn), and the image that has been normalized is referred to as "Dnorm."

$$D_{norm} = \frac{D - min\ (D)}{\max\ (D) - \min(D)} \tag{1}$$

Eq (2) describes the process reduces unwanted artifacts and noise from facial emotion images by smoothing the image using a filter like Gaussian smoothing.

$$I_{filtered}\ (x, y) = \sum \sum I\ (x + i, y + j).\ G\ (i, j) \tag{2}$$

Where $I\ (x, y)$ is the original image, and $G\ (i, j)$ is the Gaussian kernel.

3.2.4. Class Balancing Techniques

Oversampling underrepresented emotions using Synthetic Minority Over-sampling Technique (SMOTE). Weighted loss functions to assign higher penalties to misclassified minority classes. These preprocessing steps ensured that the dataset was balanced, high-quality, and optimized for training deep CNN models. Where $I\ (x, y)$ is the original image, and $G\ (i, j)$ is the Gaussian kernel.

Eq (3) describes the technique to adjust the voxel size and spatial resolution of facial emotion images to a uniform scale across all samples.

$$I_{resampled}\ (x, y) = I\ \big(f(x), f(y)\big) \tag{3}$$

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.03 No.02 (2025)

Augmentation technique based on rotation has been used to generate additional images by transforming existing ones by using Eq (4).

$$I_{rotated}\,(x', y') = I\,(x\,\cos\theta - y\,\sin\theta + y\,\cos\theta) \qquad (4)$$

4. Experimental CNN Architectures and Ablation Study

This section presents the experimental design and ablation study conducted to analyze the impact of key CNN components—Batch Normalization, Dropout, and Network Depth—on facial emotion recognition (FER) performance. The study compares a baseline CNN model, multiple ablated models (where specific components are removed or reduced), and an optimized CNN architecture incorporating L2 Regularization, BatchNorm, and Dropout (50%) for improved performance.

4.1. Baseline CNN Architecture

The baseline model consists of three convolutional layers, each followed by Batch Normalization, ReLU activation, and Max Pooling. The network is designed to balance performance and computational efficiency while preventing overfitting.

4.1.1. Baseline Model Architecture

Layer Type	Parameters / Shape	Purpose
Input Layer	(64, 64, 1) (Grayscale)	Accepts input image
Conv2D	32 filters, kernel size (3x3)	Detects low-level features (edges, textures)
MaxPooling2D	Pool size (2qx2)	Reduces spatial dimensions
Conv2D	64 filters, kernel size (3x3)	Learns mid-level features
MaxPooling2D	Pool size (2x2)	Further reduces feature map size
Conv2D	128 filters, kernel size (3x3)	Captures more complex patterns
MaxPooling2D	Pool size (2x2)	Downsamples feature maps
Flatten	Converts 3D to 1D	Prepares data for Dense layers
Dense	128 neurons	Learns high-level features
Dropout	0.5 (50% neurons dropped during training)	Reduces overfitting
Dense (Output)	7 neurons (for 7 classes)	Outputs class probabilities

4.1.2. Results:

- Baseline Accuracy: 48.05%
- The model achieves moderate performance but suffers from overfitting, particularly due to limited regularization.

4.2. Ablation Study: Impact of Removing CNN Components

To investigate the contribution of different CNN components, we systematically removed key elements and evaluated the model's performance.

Experiment	Component Removed / Changed	Observations	Accuracy	Conclusion
Model without BatchNorm	Batch Normalization	- Training became unstable - High variation in loss - Quick overfitting	**54.25%**	Slight initial accuracy gain due to relaxed constraints, but poor generalization and training instability highlight importance of BatchNorm

Model without Dropout	Dropout (All dropout layers removed)	- Severe overfitting - High training accuracy but low validation accuracy - Model memorized training data	**47.26%**	Removing Dropout caused overfitting; model couldn't generalize, showing Dropout's essential role in preventing reliance on specific neurons
Model with Fewer Layers	Reduced from 3 to 2 Conv Layers	- Poor feature extraction - Failed to capture complex expressions (e.g., Fear, Disgust) - Weaker representation	**46.27%**	Shallow models lack capacity to extract deep features, confirming deeper architectures are necessary for nuanced emotion recognition

4.3. Optimized CNN Model: Improving Performance

Based on the ablation results, we developed an optimized CNN model that integrates:

- L2 Regularization (to prevent overfitting).
- Batch Normalization (to stabilize training).
- Dropout (50%) (to enhance generalization).

Technique Added	Purpose	Impact on Model
L2 Regularization	Penalize large weights	Helps prevent overfitting and improves generalization
Batch Normalization	Normalize activations	Stabilizes training, speeds up convergence, ensures smooth gradient flow
Dropout (50%)	Randomly deactivate neurons during train	Encourages robustness, avoids over-reliance on specific neurons

4.4. Optimized Model Architecture

Layer Type	Filter Size	Activation	Other Parameters
Conv2D (64 filters)	3×3	ReLU	BatchNorm, L2 Reg.
MaxPooling2D	2×2	—	Stride=2
Conv2D (128 filters)	3×3	ReLU	BatchNorm, L2 Reg.
MaxPooling2D	2×2	—	Stride=2
Conv2D (256 filters)	3×3	ReLU	BatchNorm, L2 Reg.
MaxPooling2D	2×2	—	Stride=2
Flatten	—	—	—
Dense (256)	—	ReLU	Dropout (50%)
Dense (128)	—	ReLU	Dropout (50%)
Dense (7)	—	Softmax	Output Layer

4.5. Performance Analysis of Optimized Model

Metric / Aspect	Details
Test Accuracy	**80.38%** (significant improvement over ablation variants)
Generalization	Improved – training and validation curves remained stable

Overfitting	Reduced – Dropout (50%) and L2 Regularization helped avoid memorization
Regularization Techniques	- Batch Normalization - Dropout (50%) - L2 Regularization
Conclusion	The optimized CNN **outperformed all other models**, confirming that these techniques are vital for boosting **FER accuracy** and **model robustness**.

5. Results

The results for CNN baseline model, ablation study and optimized CNN model are discussed in this section:

5.1. Accuracy Comparison

The optimized CNN significantly outperforms other models.

Model	**Train Accuracy (%)**	**Validation Accuracy (%)**	**Test Accuracy (%)**
Baseline Model	58.62	48.05	47.23
Without BatchNorm	72.15	54.25	52.8
Without Dropout	84.78	47.26	43.95
Fewer Layers	55.43	46.27	44.12
Optimized CNN (BatchNorm, Dropout 50%, L2 Reg.)	92.35	85.12	80.38

5.1.1. Key Observations:

- Baseline Model: Moderate overfitting, with a large train-validation accuracy gap (~10%).
- No BatchNorm: Slight accuracy gain, but training instability observed.
- No Dropout: Severe overfitting, as the model memorizes training data.
- Fewer Layers: Lower accuracy due to weaker feature extraction.
- Optimized Model: Best generalization (80.38% test accuracy), lowest overfitting.

5.1.2. Precision, Recall, and F1-score

Model	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Baseline Model	46.8	44.2	45.5
Without BatchNorm	52.5	49.1	50.7
Without Dropout	45.3	42.6	43.9
Fewer Layers	44.9	41.2	42.9
Optimized CNN	81.7	78.6	80.1

5.1.3. Insights from Precision-Recall Scores

The following key insights are observed from precision-recall scores:

- The optimized CNN achieves the highest F1-score (80.1%), ensuring both high precision and recall.
- The baseline model struggles with recall, meaning it fails to detect some emotions correctly.
- Without Dropout, recall drops significantly due to poor generalization.
- Fewer layers reduce feature extraction capability, lowering both precision and recall.

5.2. Training and Validation Curves

The loss and accuracy curves provide insights into model training behavior.

5.2.1. Baseline Model Curves:

- High training accuracy but poor validation accuracy → Overfitting.

- Validation loss fluctuates → Model instability.

5.2.2. Optimized Model Curves:

- Smooth training and validation curves → Improved generalization.
- Minimal gap between training and validation accuracy → Reduced overfitting.

5.3. Summary of Findings

Major Takeaways from the Comparative Analysis:

Aspect	Baseline Model	Optimized CNN
Test Accuracy	47.23%	80.38%
Overfitting	High	Low
Feature Extraction	Limited	Strong
Precision-Recall	Moderate	High
Misclassification Rate	High	Low
Training Stability	Poor	Stable

The optimized CNN architecture significantly outperforms the baseline and ablation models. Key improvements include better generalization, reduced overfitting, and enhanced class separation.

5.4. Limitations

Despite promising results, our approach has several limitations:

5.4.1 Dataset Bias and Imbalance

- The ICML Face Dataset has an uneven distribution of emotions, with some emotion categories (e.g., Neutral, Surprise, Sad, and Happiness) had significantly fewer samples than dominant ones (e.g., Angry, Disgust and Fear).
- This leads to biased learning, where the model performs well on frequent emotions but struggles with rare ones.

5.4.2. Limited Generalization to Real-World Scenarios

- Our model is trained on static images, which may not capture dynamic facial expressions seen in real-world settings (e.g., video streams, different lighting conditions).
- The dataset consists of posed expressions, which may differ from spontaneous, natural emotions.

5.4.3. Computational Complexity

- Deep CNN models require high computational resources, making them challenging for deployment on edge devices or mobile applications.
- Training deep networks is time-consuming and demands GPUs for efficient processing.

5.4.5 Sensitivity to Noise and Occlusions

- The model struggles with partially occluded faces (e.g., glasses, masks, hand gestures covering the face).
- Background clutter and variations in illumination also impact recognition performance.

6. Conclusion

Facial emotion recognition is a critical component of human-computer interaction, affective computing, and psychological analysis. In this study, we investigated the impact of different CNN components on facial emotion classification using the ICML Face Dataset through an ablation study. Our experimental results provide key insights into how architectural modifications influence model performance. Our baseline CNN achieved 48.05% accuracy, whereas the optimized model (BatchNorm, Dropout 50%, L2 Regularization) improved performance significantly, reaching 80.38% accuracy. Removing Batch Normalization resulted in training instability, reducing accuracy to 54.25%. Without Dropout, the model suffered from over fitting, dropping accuracy to 47.26%. Using fewer layers led to a loss of expressive features, reducing accuracy further to 46.27%. The ICML Face Dataset presents imbalanced class distribution, making rare emotions (e.g., Neutral, and Surprise) harder to classify. Instead of classifying static images, use temporal models (LSTMs, Transformers, 3D CNNs) to capture facial expressions over time. Subtle expression differences between emotions like Fear vs. Surprise led to misclassification. The dataset consists of posed images, which may not fully represent real-world, spontaneous emotions. To address these limitations and improve facial emotion recognition, we propose the following future research directions: Address Dataset Imbalance with Data Augmentation & Synthetic Data, Use data augmentation techniques (random cropping, rotation, brightness adjustments) to enhance dataset diversity.

References

Khan, M. I., Arif, A., & Khan, A. R. A. (2024). The Most Recent Advances and Uses of AI in Cybersecurity. BULLET: Jurnal Multidisiplin Ilmu, 3(4), 566-578.

Zainab, H., Khan, A. R. A., Khan, M. I., & Arif, A. (2025). Innovative AI Solutions for Mental Health: Bridging Detection and Therapy. Global Journal of Emerging AI and Computing, 1(1), 51-58.

Tariq, Muhammad Arham, Muhammad Ismaeel Khan, Aftab Arif, Muhammad Aksam Iftikhar, and Ali Raza A. Khan. "Malware Images Visualization and Classification With Parameter Tunned Deep Learning Model." Metallurgical and Materials Engineering 31, no. 2 (2025): 68-73.https://doi.org/10.63278/1336.

Adyapady, R. R. A. B., 2023. A comprehensive review of facial expression recognition techniques. *Multimedia Systems,* 29(1), pp. 73-103.

Altun, M., 2019. An underestimated tool: Body language in classroom during teaching and learning. *International Journal of Social Sciences & Educational Studies,* 6(1), pp. 155-170.

Barrett, L. F. A. R. M. S. M. A. M. P. S. D., 2019. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest,* 20(1).

Arif, A., A. Khan, and M. I. Khan. "Role of AI in Predicting and Mitigating Threats: A Comprehensive Review." JURIHUM: Jurnal Inovasi dan Humaniora 2, no. 3 (2024): 297-311.

Chowdary, M. K. N. T. N. H. D. J., 2023. Deep learning-based facial emotion recognition for human–computer interaction applications. *Neural Computing and Applications,* Volume 35, pp. 23311-23328.

Dagnaw, G., 2020. *Artificial intelligence towards future industrial opportunities and challenges.* Kennesaw State University, Kennesaw State University.

Dang, N. C. M.-G. M. N. D. l. P. F., 2020. Sentiment analysis based on deep learning: A comparative study. *Electronics,* 9(3), pp. 483-511.

Elmahmudi, A. U. H., 2019. Deep face recognition using imperfect facial data. *Future Generation Computer Systems,* Volume 99, pp. 213-225.

Georgescu, M.-I. I. R. T. P. M., 2019. Local learning with deep and handcrafted features for facial expression recognition. *IEEE Access,* Volume 7, pp. 64827-64836.

Gupta, S. K. P. T. R. K., 2023. Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. *Multimedia Tools and Applications,* 82(8), pp. 11365-11394.

Haq, H. A. W. I. M. K. A. A. M., 2024. Enhanced real-time facial expression recognition using deep learning. *Acadlore Transactions on Machine Learning,* 3(1), pp. 24-35.

Aish, M. A., Ahmad, J., Nasim, F., & Iqbal, M. J. (2024). Brain Tumor Segmentation and Classification Using ResNet50 and U-Net with TCGA-LGG and TCIA MRI Scans. Journal of Computing & Biomedical Informatics, 8(01).

Hasan, Z. F., 2022. An Improved Facial Expression Recognition Method Using Combined Hog and Gabor Features. *Science Journal of University of Zakho,* 10(2), pp. 54-59.

Hattab, A. B. A., 2024. Face-Iris multimodal biometric recognition system based on deep learning. *Multimedia Tools and Applications,* Volume 83, pp. 43349-43376.

Zainab, H., Khan, M. I., Arif, A., & Khan, A. R. A. (2025). Development of Hybrid AI Models for Real-Time Cancer Diagnostics Using Multi-Modality Imaging (CT, MRI, PET). Global Journal of Machine Learning and Computing, 1(1), 66-75.

Holodynski, M. S. D., 2019. Expressions as signs and their significance for emotional development. *Developmental Psychology,* 55(9), pp. 1812-1829.

Hossain, S. U. S. A. V. R. R. K., 2021. A unified framework of deep learning-based facial expression recognition system for diversified applications. *Applied Sciences,* Volume 11, pp. 9174-9199.

Mazzei, D. C. F. F. G., 2021. Analyzing social robotics research with natural language processing techniques. *Cognitive Computation,* 13(2), pp. 308-321.

Pise, A. A. A. M. A. V. P. K. P. K. D. A. H., 2022. Methods for facial expression recognition with applications in challenging situations. *Computational Intelligence and Neuroscience,* 2022(1).

Rao, T. L. X. X. M., 2020. Learning multi-level deep representations for image emotion classification. *Neural Processing Letters,* Volume 51, pp. 2043-2061.

Saberi, M. D. S. B. U., 2021. Expressing personality through non-verbal behaviour in real-time interaction. *Frontiers in Psychology,* Volume 12.

Salman, S. A. Z. A. T. H., 2023. Cascaded Deep Graphical Convolutional Neural Network for 2D Hand Pose Estimation. *International Workshop on Advanced Imaging Technology (IWAIT) 2023,* p. 12592.

Khan, M. I., Arif, A., & Khan, A. R. A. (2024). AI's Revolutionary Role in Cyber Defense and Social Engineering. International Journal of Multidisciplinary Sciences and Arts, 3(4), 57-66.

Talaat, F. M. A. Z. H. M. R. R. E.-R. N., 2024. Real-time facial emotion recognition model based on kernel autoencoder and convolutional neural network for autism children. *Soft Computing,* Volume 28, pp. 6695-6708.

Taye, M. M., 2023. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers,* 91(12).

Testolin, A. P. M. S. S., 2020. Deep learning systems as complex networks. *Journal of Complex Networks,* 8(1).

Xie, H.-X. L. L. S. H.-H. C. W.-H., 2022. An overview of facial micro-expression analysis: Data, methodology and challenge. *IEEE Transactions on Affective Computing,* 14(3), pp. 1857-1875.