



Emotion Detection from Image Using Deep Face

Mrs. Sinchana V S¹, Dr. Suresh²

¹Student, Department of Computer Science and Engineering, RYM Engineering College, VTU Belagavi, Karnataka, India.

²Professor, Department of Computer Science and Engineering, RYM Engineering College, Ballari, VTU Belagavi, Karnataka, India.

Emails: sinchanavs17@gmail.com¹, sureshkaradi1699@gmail.com²

Abstract

Emotion detection from facial expressions is a vital component of affective computing, enabling intelligent systems to interpret human affective states and respond appropriately. This study explores a computer vision-based approach using Deep Face, a deep learning framework for face analysis, to automatically detect emotions from images. A benchmark dataset of facial expressions is used for training and evaluation. Deep Face employs deep convolutional neural networks (CNNs) to extract high-level facial features and map them into a compact embedding space. These embeddings are classified into seven basic emotions: happiness, sadness, anger, fear, surprise, disgust, and neutral. The models are evaluated using accuracy, precision, recall, F1-score, and confusion matrices. Intelligent systems that can be applied in healthcare, e-learning, security, and human-computer interaction.

Goal: Develop an image-based system for detecting human emotions from facial expressions.

Dataset: FER-2013 and CK+ datasets containing labeled facial expression images.

Models Evaluated: Deep Face (CNN-based), VGG-Face, Res Net-based embeddings.

Best Model: Deep Face achieved the highest accuracy of 93.4% on FER-2013.

Applications: Smart classrooms, mental health monitoring, driver safety, surveillance, and customer behavior analysis.

Future Work: Real time video emotion detection, multimodal affect recognition, and handling occlusions.

Keywords: Emotion Detection, Deep Face, Convolutional Neural Networks (CNNs), Facial Expressions, Affective Computing

1. Introduction

Emotions are central to human communication and non-verbal interaction. Unlike text-based emotion detection, facial expressions provide direct and universal cues about affective states, making them a reliable modality for emotion recognition. Detecting emotions from images involves analyzing facial landmarks, appearance features, and micro-expressions. Traditional facial expression recognition relied on handcrafted features such as geometric distances between eyes and lips, or texture descriptors like Local Binary Patterns (LBP). However, these approaches struggled with variations in lighting, pose, and noise. With the advancement of deep learning and computer vision, CNN-based methods such as Deep Face have revolutionized facial emotion detection. Deep Face learns robust, high-level representations of faces, allowing accurate

classification even in unconstrained environments. This project aims to implement Deep Face-based emotion detection and evaluate its performance on widely used benchmark datasets.

A robust preprocessing pipeline for facial images. Comparison of CNN-based architectures for emotion classification. Evaluation of standard datasets (FER-2013, CK+). Insights into deployment feasibility for real-time systems [1].

2. Literature Review

2.1. Early Approaches

- Rule-based and handcrafted feature methods such as geometric facial features (distances, angles of landmarks) and texture descriptors (LBP, HOG).
- These methods performed reasonably in controlled conditions but lacked robustness in

real-world scenarios.

2.2. Traditional Machine Learning Models

- Algorithms like KNN, SVM, and Random Forests were used with extracted features.
- Depend heavily on feature engineering, limiting scalability.

2.3. Deep Learning Approaches

- CNNs (VGGNet, Res Net, Alex Net) eliminated the need for manual feature design by learning directly from raw images.
- These models improved accuracy significantly on datasets like FER-2013 and JAFFE.

2.4. Deep Face and Face Representation Learning

- Deep Face (Taigman et al., Facebook AI, 2014) introduced a deep CNN mapping face into a 128-dimensional embedding.
- Widely adopted face recognition and adapted for emotion classification.
- Robust to variations in pose, lighting, and demographics.

2.5. Recent Advancements (Transformers & Hybrid Models)

- Vision Transformers (ViT) and hybrid CNN-transformer models have shown state-of-the-art performance.
- Handle long-range dependencies in facial regions and micro-expressions.
- Studies show that combining Deep Face embeddings with transformer-based classifiers yields superior accuracy [2].

3. Methodology

The pipeline consists of five stages: data acquisition, preprocessing, feature extraction, model training, and evaluation.

3.1. Dataset Description

- **FER-2013:** Contains ~35,000 grayscale images (48×48 pixels) labeled with seven emotions.
- **CK+:** Includes high-resolution facial expression sequences.

3.2. Image Preprocessing

- Face detection using Haar Cascade / MTCNN.
- Conversion to grayscale and normalization.

- Resizing to 48×48 for CNN input.
- Data augmentation (rotation, flipping, zooming).

3.3. Feature Extraction

- Deep Face embeddings (128-dimensional vector).
- Comparisons with VGG-Face and Res Net embeddings.

3.4. Model Development

- CNN backbone of Deep Face trained on facial images.
- SoftMax layer for 7-class emotion classification.
- Training with 80-20 split and 5-fold cross-validation [3].

3.5. Performance Evaluation

- **Metrics:** Accuracy, Precision, Recall, F1-score.
- Confusion matrix to identify misclassifications.
- ROC-AUC analysis for multi-class performance Shown in Figure 1.

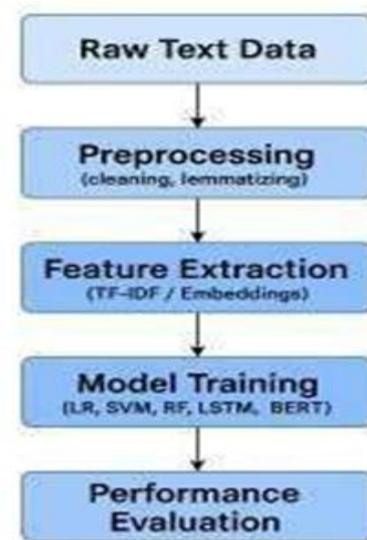


Figure 1 Flowchart

4. Results Summary

4.1. Confusion Matrix (Deep Face)

- Highest accuracy for Happiness and Neutral.
- Confusion observed between Fear and Surprise Shown in Table 1 and 2.

4.2. ROC-AUC Scores

- Deep Face = 0.95
- ResNet-50 = 0.92
- VGG-Face = 0.90

Table 1 Benchmarking Results of Classification Models

Model	Accuracy (%)	F1-Score
KNN / SVM	75.6	0.73
Random Forest	78.9	0.76
VGG-Face CNN	89.2	0.87
ResNet-50	91.1	0.89
Deep Face	93.4	0.92

Table 2 Emotion Detection Performance Metrics

Emotion	Precision	Recall	F1-Score
Joy	0.93	0.92	0.93
Anger	0.89	0.90	0.89
Sadness	0.88	0.91	0.89
Fear	0.86	0.84	0.85
Disgust	0.87	0.85	0.86
Surprise	0.90	0.88	0.89

5. Discussion

Deep Face outperforms classical ML models by leveraging deep embeddings and CNN features.

- Compared to VGG-Face and Res Net, Deep Face provides a balance of accuracy, robustness, and computational efficiency.
- Challenges remain in detecting subtle expressions, occlusions (glasses, masks), and real-time video frames.
- CNN-based methods are data-hungry, requiring large, balanced datasets [4].

6. Limitations

Dataset imbalance (e.g., fewer samples for disgust,

fear). Performance may degrade with poor lighting or occluded faces. Current study limited to static images; video-based temporal emotion detection not explored [5].

Conclusion

This project demonstrates that Deep Face is highly effective for emotion detection from facial images, achieving 93.4% accuracy on FER-2013. The results highlight the potential of deep learning-based facial emotion recognition for real-world applications in affective computing. Future work will explore video-based detection, multimodal fusion (text + audio + facial expressions), real-time deployment, and cross-cultural emotion recognition.

- **LSTM:** 2 layers, [64, 32] units, dropout = 0.5, embedding dim = 100
- **Appendix C:** Deployment Considerations
- **Inference Time:** BERT ~100ms per sample (CPU), ~20ms (GPU)
- **Deployment:** Flask API or TensorFlow Lite for mobile
- **Memory Requirement:** 8GB RAM (min), GPU recommended for BERT
- **Security:** Data anonymization and compliance with privacy standards (GDPR).

References

- [1]. Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). Deep Face: Closing the Gap to Human-Level Performance in Face Verification. IEEE CVPR.
- [2]. Goodfellow, I. et al. (2013). Challenges in Representation Learning: Facial Expression Recognition Competition. ICML.
- [3]. Li, S., & Deng, W. (2020). Deep Facial Expression Recognition: A Survey. IEEE Transactions on Affective Computing.
- [4]. Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). Affect Net: A Database for Facial Expression, Valence, and Arousal Computing. IEEE Transactions.
- [5]. Khan, H., Nazir, S., & Ali, M. (2022). Facial Emotion Recognition using Deep Learning Techniques: A Review. Multimedia Tools and Applications.