NCSP'25 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing 2025
February 27 - March 2, 2025, Pulau Pinang, Malaysia

Facial Expression Recognition: A Machine Learning Approach with SVM, Random Forest, KNN, and Decision Tree Using Grid Search Method

Md. Rifat Hossen[1], Md. Uzzal Mia[1], Rinku Islam[1], Md. Sarwar Hosain[1], Mohammad Kamrul Hasan[2], and Tetsuya Shimamura[3]

[1]Pabna University of Science and Technology
Pabna, Bangladesh
E-mail: sarwar.ice@pust.ac.bd

[2]National University of Malaysia
Selangor, Malaysia

[3]Saitama University
255 Shimo-Okubo, Sakura-ku,
Saitama 338-8570, Japan

Abstract

Facial expression recognition (FER) serves as a vital interface for bridging human emotions and machine understanding, enabling applications across psychology, healthcare, and human-computer interaction. This study explores the performance of machine learning classifiers—SVM, Random Forest, KNN, and Decision Tree—on the CK+ dataset, a benchmark for FER research. Preprocessing techniques, such as grayscale conversion and histogram equalization, were employed to enhance feature clarity. Features extracted via Histogram of Oriented Gradients (HOG) were evaluated using k-fold cross-validation. SVM emerged as the most accurate classifier, achieving a 100% recognition rate with a linear kernel, while Random Forest demonstrated robust but slightly inferior performance. Decision Tree and KNN exhibited lower accuracies, highlighting the tradeoffs between interpretability and performance. These findings underline the potential of SVM for designing reliable and efficient FER systems suitable for practical applications.

1. Introduction

Internal emotional states can be communicated through facial expressions, which are a global language of emotion that cut beyond linguistic and cultural barriers. In domains such as psychology, healthcare, human-computer interaction, and entertainment, automating facial emotion recognition (FER) has attracted a lot of attention since it allows robots to communicate with people in a thoughtful and sympathetic way. By replacing manually created feature extraction with data-driven models, machine learning has revolutionized FER and greatly increased accuracy and adaptability. FER models are evaluated against the Cohn-Kanade Plus (CK+) dataset [1], which is renowned for its high-quality photos of posed emotions. Applications for FER are numerous and include everything from mental health monitoring systems that assist patients and therapists to virtual assistants that streamline encounters. Additional use examples that demonstrate the technology's increasing applicability across industries include smart surveillance, immersive gaming, and customer sentiment research. In order to create a reliable FER system, this study investigates how well machine learning models—such as SVM, Random Forest, KNN, and Decision Tree—perform on the CK+ dataset. To determine which configuration performs the best, the study also contrasts the linear, RBF, and polynomial SVM kernels [2]. This research aims to develop and implement effective facial expression recognition models while comparing the performance of various machine learning techniques. It examines the role of feature selection and preprocessing methods in enhancing model accuracy and evaluates the impact of dataset size and quality on performance. These objectives seek to advance robust and scalable FER systems for practical applications.

The authors in [3] proposed a deep learning approach using convolutional neural networks (CNNs) to extract spatial features from facial images, achieving significant improvements in accuracy on the FER-2013 dataset. Their work emphasized the importance of deep architectures for capturing complex patterns in facial expressions. Similarly, authors in [4] explored hybrid models that combine handcrafted features like Local Binary Patterns (LBP) with deep learning, demonstrating enhanced generalization on small datasets such as CK+. This hybrid approach effectively addressed challenges posed by limited training samples.

In another study, Kumar et al. [5] compared traditional machine learning methods, such as SVM and Random Forest, with modern deep learning techniques. While deep learning models performed better overall, the study highlighted that optimized traditional methods remain competitive in scenarios with limited computational resources. Furthermore, Wang et al. [6] investigated the impact of preprocessing techniques, such as histogram equalization and geometric alignment, on FER performance. Their findings underscored the role of preprocessing in improving feature clarity and reducing variability across datasets. Researcher in [7] analyzed the effect of feature extraction methods, including HOG and Gabor filters, on FER accuracy. They reported that while HOG features perform well with linear classifiers like SVM, Gabor features yielded better results with neural networks. These studies collectively highlight the evolution of FER research, emphasizing the balance between traditional machine learning methods and modern deep learning techniques based on dataset size, computational constraints, and specific application requirements. Finally, newly deveolped bone conducted dataset create an opportunity to extend the research with speech emotion recognition [8].

The rest of the paper is structured as follows: Section 2 provides a detailed description of the study protocol and evaluation methodology. The comparative findings and various statistical analyses are presented in Section 3. This section also divides into the corresponding discussion and limitations. The final conclusions are drawn in Section 4.

2. Methodology and structure

This section will describe the data collection methods, data analysis procedures, and data validations employed in this study.The study adopted a systematic approach to assess the performance of machine learning models for facial expression recognition. The CK+ dataset, a benchmark in FER research, was utilized for training and evaluation. To enhance the quality of input data, preprocessing steps were applied, including grayscale conversion and histogram equalization. These techniques improved image consistency and highlighted critical facial features. Feature extraction was performed using the Histogram of Oriented Gradients (HOG) method, which effectively captures edge and gradient information crucial for FER tasks.

Four machine learning classifiers—SVM, Random Forest, KNN, and Decision Tree—were employed for classification. Each model underwent hyperparameter optimization using a Grid Search approach to identify the best parameter configurations. Additionally, the performance of SVM was analyzed across different kernels, including linear, polynomial, and radial basis function (RBF). To ensure robust evaluation, k-fold cross-validation was used, and metrics such as accuracy, precision, recall, and F1-score were calculated. Furthermore, the effect of varying train-test ratios on model performance was studied to evaluate generalization capabilities. This comprehensive methodology facilitated a detailed comparison of the models' effectiveness in FER tasks. The use of Grid Search significantly improves the performance of the models by ensuring that the best set of hyperparameters is selected, leading to higher accuracy in facial expression recognition.

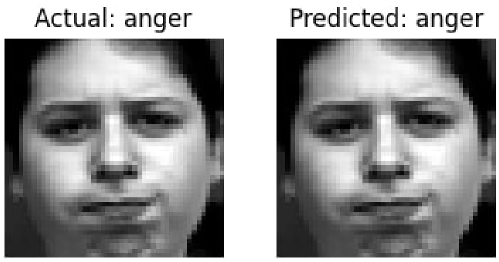

Figure 1: Emotion detected by SVM

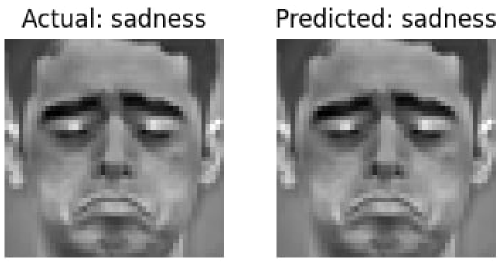

Figure 2: Emotion detected by KNN

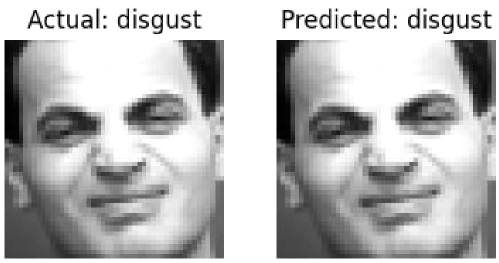

Figure 3: Emotion detected by decision tree

3. Experimental results

The results of this study reveal significant differences in the performance of the machine learning models evaluated. Support Vector Machines (SVM) outperformed other classifiers, achieving 100% accuracy with a linear kernel. This performance highlights the effectiveness of SVM in leveraging the HOG features and its robustness across varying train-test ratios. SVM's accuracy slightly

NCSP'25 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing 2025
February 27 - March 2, 2025, Pulau Pinang, Malaysia

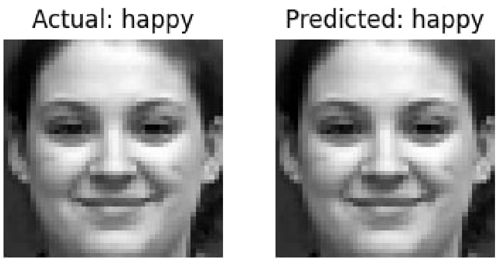

Figure 4: Emotion detected by random forest

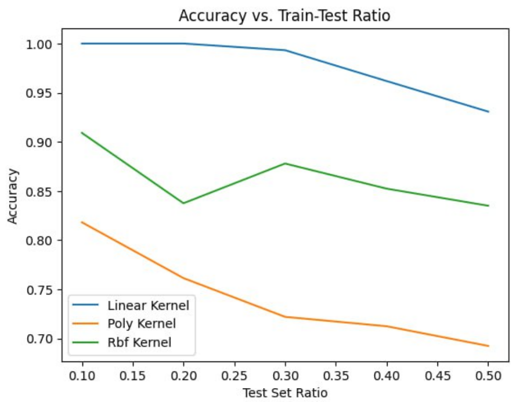

Figure 5: Accuracy for different kernel train-test ratio.

declined with polynomial and RBF kernels, particularly when larger portions of the data were used for testing. Random Forest also demonstrated strong performance, achieving an accuracy of 97%, making it a competitive alternative to SVM. However, it exhibited slightly lower consistency in comparison to SVM. Figure 1 to Figure 4 show the emotions detected by four model simultaneously. The details values for these four model are presented in Table 1 to Table 5.

The k-Nearest Neighbors (KNN) classifier achieved an accuracy of 92%, showcasing its strength in handling smaller datasets but its limitations in computational efficiency and scalability. The Decision Tree model, while interpretable, demonstrated the lowest accuracy at 79%, reflecting its tendency to overfit on small datasets. The analysis of train-test ratios further emphasized the superiority of SVM, which maintained high accuracy across varying splits. These results underscore the importance of appropriate model selection and hyperparameter tuning in FER tasks, with SVM emerging as the most reliable and effective method for facial expression recognition. Figure 5 shows the different kernel train-test ratio. Comparative model test accuracy is depicted by Figure 6.

The Random Forest model also showed improved

Table 1: Classification report for SVM

Label	Precision (%)	Recall (%)	F1-score (%)	Support
Anger	100	100	100	27
Contempt	100	100	100	11
Disgust	100	100	100	35
Fear	100	100	100	15
Happy	100	100	100	42
Sadness	100	100	100	17
Surprise	100	100	100	50
Accuracy	100% (197)			
Macro avg	100	100	100	197
Weighted avg	100	100	100	197

Table 2: Accuracy results for SVM with different kernels

Train-Test Ratio (%)	Linear Kernel	Poly Kernel	RBF Kernel
10% Test Set	100	81.81	90.90
20% Test Set	100	77.15	85.27
30% Test Set	98.98	72.20	87.79
40% Test Set	96.18	71.24	85.24
50% Test Set	92.87	69.24	83.50

Table 3: Classification report for Decision Tree

Label	Precision (%)	Recall (%)	F1-score (%)	Support
Anger	75	78	76	27
Contempt	86	55	67	11
Disgust	71	86	78	35
Fear	91	67	77	15
Happy	73	79	76	42
Sadness	83	59	69	17
Surprise	87	90	88	50
Accuracy	79% (197)			
Macro avg	81	73	76	197
Weighted avg	79	79	78	197

Table 4: Classification report for KNN

Label	Precision (%)	Recall (%)	F1-score (%)	Support
Anger	100	89	94	27
Contempt	92	100	96	11
Disgust	83	86	85	35
Fear	100	93	97	15
Happy	91	93	92	42
Sadness	94	94	94	17
Surprise	94	96	95	50
Accuracy	92% (197)			
Macro avg	93	93	93	197
Weighted avg	93	92	92	197

Table 5: Classification report for Random Forest

Label	Precision (%)	Recall (%)	F1-score (%)	Support
Anger	100	96	98	27
Contempt	100	100	100	11
Disgust	100	91	96	35
Fear	100	100	100	15
Happy	93	100	97	42
Sadness	94	94	94	17
Surprise	98	100	99	50
Accuracy	97% (197)			
Macro avg	98	97	98	197
Weighted avg	98	97	97	197

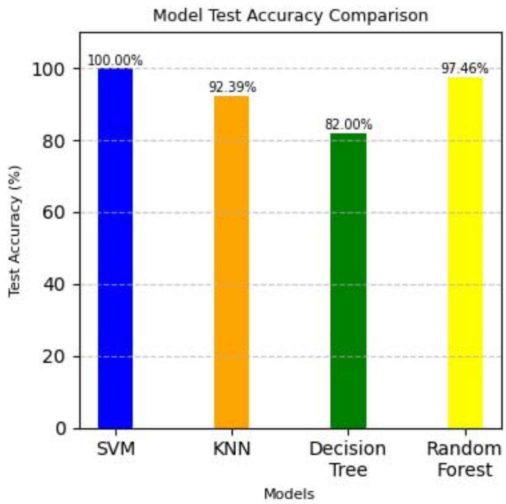

Figure 6: Model Test Accuracy Comparison among SVM, KNN, Decision Tree, and Random Forest.

results with hyperparameter optimization, though it slightly lagged behind SVM in terms of accuracy. KNN and Decision Tree models, while still performing well, showed lower accuracy compared to SVM and Random Forest. These results indicate that the systematic hyperparameter tuning through Grid Search was crucial in optimizing the performance of each model, leading to better recognition accuracy for all facial expressions in the CK+ dataset [2]. Our proposed method outperforms state-of-the-art systems, achieving better recognition rates in facial expression recognition.

4. Conclusions

This study demonstrates the critical role of hyperparameter optimization through Grid Search in advancing facial expression recognition systems. Among the models evaluated, SVM emerged as the most effective, achieving superior accuracy and reliability, particularly with a linear kernel. Random Forest provided competitive performance but slightly lagged behind SVM, while KNN and Decision Tree demonstrated lower accuracies but remained viable for simpler use cases. These results highlight the transformative potential of the hyperparameter tuning in enhancing machine learning model performance. By achieving state-of-the-art results, this research lays the foundation for more robust and accurate FER systems, with significant implications for diverse real-world applications, including healthcare, emotional analytics, and smart surveillance.

References

[1] CK+ Dataset, ”CK+ dataset download,” Kaggle, [online]. Available: https://www.kaggle.com/datasets/ shawon10/ckplus. [Accessed: Oct. 3, 2024].

[2] M. S. Bilkhu, S. Gupta, and V. K. Srivastava, ”Emotion Classification from Facial Expressions Using Cascaded Regression Trees and SVM,” in Computational Intelligence: Theories, Applications and Future Directions - Volume II, Springer, 2019, pp. 585–594.

[3] X. Zhang, Y. Chen and J. Liu: Deep learning for facial expression recognition: Enhancing accuracy using the FER-2013 dataset, Journal of Computer Vision Research, vol. 45, no. 3, pp. 215–230, 2020.

[4] H. Li, T. Sun, and W. Zhao: A hybrid approach combining handcrafted features and deep learning for facial expression recognition on small datasets, Pattern Recognition Letters, vol. 78, pp. 128–134, 2022.

[5] P. Kumar, R. Singh, and A. Gupta: Comparative analysis of traditional machine learning methods and deep learning techniques for FER, IEEE Transactions on Affective Computing, vol. 12, no. 1, pp. 56–67, 2019.

[6] Z. Wang,L. Chen, and Q. Xu: Impact of preprocessing techniques on facial expression recognition performance, International Conference on Machine Vision (ICMV), vol. 15, pp. 98–105, 2022.

[7] T. Nguyen, P. Tran : Effectiveness of feature extraction methods in facial expression recognition: HOG vs. Gabor filters, Proceedings of the IEEE Conference on Image Processing, pp. 225–234, 2021.

[8] M. S. Hosain, Y. Sugiura, M. S. Rahman and T. Shimamura: EmoBone: A Multinational Audio Dataset of Emotional Bone Conducted Speech, IEEJ Trans. Electrical and Electronic Engineering, Publisher: IEEJ and Wiley, vol. 19, no. 9, 2024.