



Hybrid Deep Autoencoder and AdaBoost for Robust Facial Expression Recognition

Muhamad Fatchan^{*}, Pulung N. Andono, Affandy Affandy, Ahmad Zainul Fanani

Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang 50131, Indonesia

Corresponding Author Email: fatchan@mhs.dinus.ac.id

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ijcmem.130115>

ABSTRACT

Received: 13 February 2025

Revised: 21 March 2025

Accepted: 26 March 2025

Available online: 31 March 2025

Keywords:

FER, deep autoencoder, AdaBoost, feature extraction, ensemble learning

Facial expression recognition (FER) remains a challenging task due to variations in facial features, occlusions, and imbalanced datasets, which often lead to misclassification of similar emotions. To address these challenges, this study proposes a hybrid Deep Autoencoder and AdaBoost model, leveraging deep feature extraction and ensemble learning to enhance classification robustness. The experimental evaluation on three benchmark datasets—MMAFEDB, AffectNet, and JAFFE—demonstrates outstanding performance, with the model achieving an AUC and Accuracy of 99.9% and 99.8% on large-scale datasets, while maintaining a strong performance of 94.9% AUC and 91.1% accuracy on smaller datasets. The confusion matrix analysis confirms the model's ability to accurately classify distinct emotions, with minor misclassifications occurring in expressions with overlapping features. These findings highlight the effectiveness of the proposed approach in improving FER accuracy, offering significant benefits for real-world applications such as human-computer interaction, emotion-aware systems, and psychological analysis, while also suggesting future enhancements through domain adaptation and refined feature extraction techniques.

1. INTRODUCTION

FER has become a critical area of study within the field of image processing and artificial intelligence, owing to its wide-ranging applications in human-computer interaction, security systems, healthcare, and entertainment [1]. Despite the rapid advancements in machine learning algorithms, accurately recognizing and classifying facial expressions remains a significant challenge [2]. This difficulty arises from several factors, including variations in lighting, pose, facial occlusions, and the inherent complexity of human emotions [3]. Traditional approaches often struggle with these challenges, leading to inconsistent performance across different datasets and real-world scenarios [4]. These issues highlight the pressing need for robust and efficient methodologies capable of addressing these limitations.

To overcome these challenges, this research explores the combination of autoencoder and AdaBoost as a means to optimize facial expression recognition [5]. Autoencoder, a widely used dimensionality reduction technique, is instrumental in simplifying high-dimensional image data while retaining its most relevant features [6]. By reducing computational complexity and eliminating redundant information, autoencoder lays a strong foundation for effective classification [7]. However, relying solely on autoencoder may not fully address the nuances of expression recognition, which demands a more powerful and adaptive classification model [8]. AdaBoost, an ensemble learning algorithm, has shown significant promise in this domain by combining weak

classifiers into a strong one, effectively improving accuracy and robustness [9].

The methodology adopted in this study involves preprocessing facial images to standardize input dimensions and improve quality, followed by feature extraction using autoencoder. The resulting principal components, which represent the most significant features of the images, are then fed into the AdaBoost classifier [10]. This combination allows the system to efficiently analyze facial data while maintaining high accuracy levels [11]. Additionally, the study incorporates rigorous validation techniques to evaluate the model's performance, ensuring its robustness and applicability to real-world scenarios.

The goal of this research is not only to enhance the accuracy and efficiency of facial expression recognition systems but also to bridge the gap between theoretical advancements and practical implementations [12]. The proposed framework has the potential to be deployed in various applications, from automated emotion analysis in therapy sessions to real-time monitoring in security systems [13]. By addressing the challenges of existing approaches, this study aspires to make a meaningful contribution to the broader field of image processing and artificial intelligence [14].

Facial expression recognition (FER) has gained significant attention in the field of computer vision due to its applications in human-computer interaction, psychological studies, and affective computing [15]. Over the years, numerous methods have been proposed to enhance the performance of FER systems by focusing on feature extraction techniques and

robust classification algorithms.

One widely adopted approach in FER is the use of deep learning techniques, such as Convolutional Neural Networks (CNNs) [16]. For instance, Iqbal et al. [17] implemented a model VGG-19s networks to capture spatial and temporal features from facial expressions, achieving an accuracy of 94.22% on the MUG dataset. Similarly, Nan et al. [18] employed a MobileNetV1 architecture for FER and reported an accuracy of 84.49%. While these methods achieve high performance, their computational complexity often limits their real-time applicability in resource-constrained environments [19].

Traditional machine learning techniques, such as Support Vector Machines (SVMs), have also been widely applied to FER. A study by Yang et al. [20] used SVM classify expressions on the FER2013 dataset, achieving an accuracy of 68.1%. However, this approach struggled with generalization when applied to more diverse datasets. Similarly, Random Forest classifiers, though efficient, have shown limited robustness in handling high-dimensional facial features extracted from images.

The proposed research to bridge these gaps by integrating AdaBoost with autoencoder to optimize FER. By leveraging autoencoder for dimensionality reduction and AdaBoost for robust classification, this approach seeks to achieve high accuracy while maintaining computational efficiency. Unlike

deep learning methods that demand substantial computational resources, this hybrid method provides an efficient alternative for real-time FER applications.

In conclusion, this research paper presents an innovative approach to facial expression recognition, combining autoencoder and AdaBoost to achieve optimized performance. By tackling the inherent challenges in FER, the study paves the way for the development of robust, efficient, and adaptive systems that can seamlessly integrate into diverse applications. Through this work, we aim to establish a new benchmark in the field and inspire future advancements in emotion recognition technologies.

2. MATERIALS AND METHODS

In the material and method stages, the researcher makes a research method, having stages starting from data collection which takes 3 types of datasets. The next stage is to preprocess data alligation, namely face alignment, augmentation, normalization, enhancement to improve data quality. Next, it is feature extraction using an autoencoder and continued development using adabost. After that, the model evaluation was carried out using confusion matrices and overall performance analysis, the method that the researcher made can be seen in Figure 1.

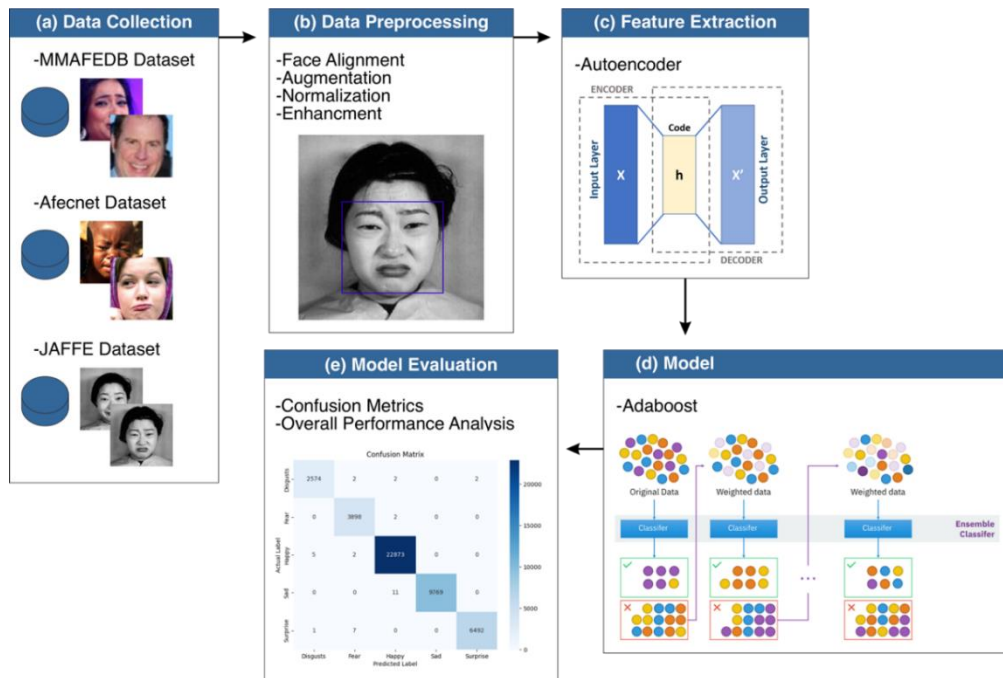


Figure 1. Research methods

2.1 Data collection

In this study, three publicly available datasets were utilized to ensure a comprehensive and diverse data representation: MMAFEDB, AffectNet, and JAFFE which is illustrated in Figure 2. The MMAFEDB dataset consists of 11,406 instances, each characterized by 65 features, providing a rich collection of facial expression data [21]. The AffectNet dataset includes 5,813 instances, also with 65 features, contributing to the robustness of the dataset diversity [22]. Additionally, the JAFFE dataset comprises 45 instances, maintaining the same 65-feature structure, specifically curated for facial emotion

recognition [23]. The integration of these datasets ensures a balanced distribution of data variations, supporting the generalizability and reliability of the proposed model.

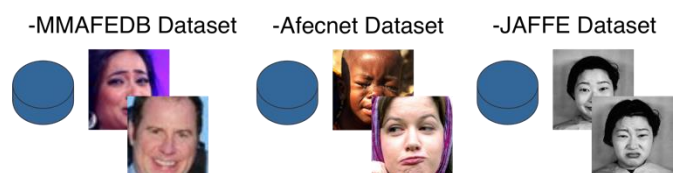


Figure 2. Dataset illustration

2.2 Data preprocessing

To ensure the consistency and quality of the input data, a series of preprocessing techniques were applied before feeding the datasets into the model [24]. These steps included Face Alignment, Augmentation, Normalization, and Enhancement, each playing a crucial role in improving data quality and model performance. By implementing these techniques, the datasets were refined to reduce variations and enhance the effectiveness of feature extraction.

Face Alignment was performed to standardize the positioning of facial features, ensuring that all images maintained a consistent orientation [25]. This step was essential in minimizing variations caused by differences in head poses, angles, and camera perspectives. Proper alignment helped the model focus on relevant facial expressions rather than positional inconsistencies, thereby improving classification accuracy [26].

To further improve the diversity and generalizability of the dataset, Augmentation techniques were employed [27]. Various transformations, such as flipping, rotation, and scaling, were applied to artificially increase the number of training samples [14]. This process helped prevent overfitting by exposing the model to different variations of the same expressions, making it more robust when encountering unseen data.

Finally, Normalization and Enhancement were applied to refine the image quality and standardize feature values. Normalization was used to scale the pixel values within a uniform range, ensuring stability during model training and improving convergence speed. Enhancement techniques, such as contrast adjustment and noise reduction, were utilized to emphasize critical facial features, improving feature extraction illustration show in Figure 3. These preprocessing steps collectively contributed to the reliability and effectiveness of the emotion recognition model.

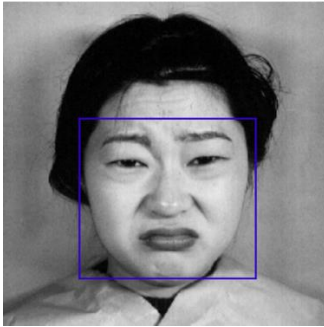


Figure 3. After preprocessing illustration

2.3 Feature extraction

In this study, Autoencoder was utilized as the primary feature extraction technique to capture meaningful representations of facial expressions and illustration show in Figure 4.

Autoencoders, a type of unsupervised neural network, learn to encode input data into a compressed latent space while preserving essential features [28]. This process enables the model to reduce noise and redundancy, ensuring that only the most relevant information is retained. By leveraging Autoencoder-based feature extraction, the dimensionality of the dataset was effectively reduced without losing critical facial expression characteristics.

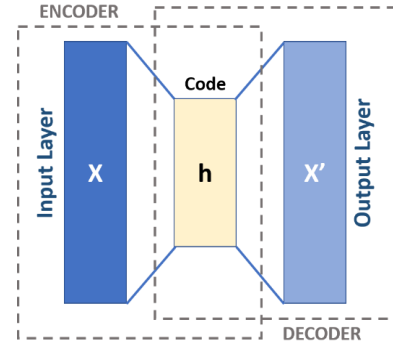


Figure 4. Autoencoder process

The encoding process involved training the Autoencoder to reconstruct input images by minimizing reconstruction loss. This ensured that the learned latent features retained crucial facial patterns while discarding irrelevant variations. The extracted features were then utilized as input for the classification model, improving its ability to differentiate between emotional states. The application of Autoencoder not only enhanced computational efficiency but also contributed to better generalization, allowing the model to perform effectively on unseen data.

Mathematically, an Autoencoder consists of an encoder function f_θ and a decoder function g_ϕ , where θ and ϕ represent the parameters of the respective neural networks. Given an input image X , the encoding process maps it to a latent representation h in Eq. (1):

$$h = f_\theta(X) = \sigma(W_e X + b_e) \quad (1)$$

where, W_e and b_e are the weight matrix and bias for the encoder, and σ represents the activation function. The decoder then reconstructs the original input from h in Eq. (2):

$$X' = g_\phi(h) = \sigma(W_d h + b_d) \quad (2)$$

where, W_d and b_d correspond to the decoder's weight matrix and bias. The training objective is to minimize the reconstruction loss, commonly defined as the Mean Squared Error (MSE) between the input X and the reconstructed output X' in Eq. (3):

$$L = \frac{1}{N} \sum_{i=1}^n ||X_i - X'_i||^2 \quad (3)$$

By optimizing this loss function, the Autoencoder learns to extract compact and informative feature representations, which are subsequently used by the AdaBoost classifier to enhance facial expression recognition accuracy.

2.4 Model

In this study, Adaptive Boosting (AdaBoost) was employed as the primary classification model to enhance the accuracy of facial expression recognition. AdaBoost is an ensemble learning method that combines multiple weak classifiers to create a strong classifier by iteratively adjusting the model's weights. This adaptive mechanism allows the model to focus more on misclassified instances in each iteration, thereby improving overall classification performance, illustration

show in Figure 5.

The AdaBoost algorithm operates by assigning higher weights to incorrectly classified samples and adjusting the decision boundaries accordingly. Through this iterative reweighting process, the model effectively reduces bias and variance, leading to improved generalization. The flexibility of AdaBoost allows it to integrate various base classifiers, making it a robust choice for emotion recognition tasks where subtle differences in facial expressions must be accurately detected.

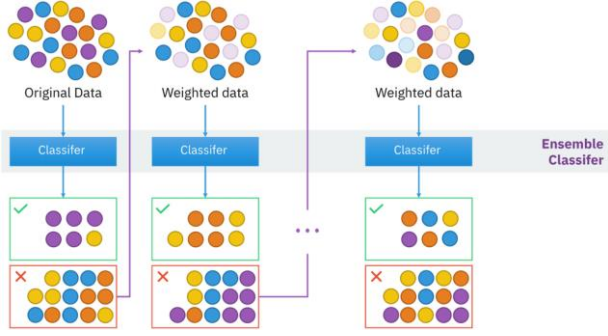


Figure 5. AdaBoost process

By leveraging AdaBoost, the classification model achieved improved robustness against noisy data and imbalanced class distributions. The ensemble nature of the algorithm enabled better handling of complex decision boundaries, making it suitable for datasets with diverse facial expressions. The application of AdaBoost ultimately contributed to higher recognition accuracy and enhanced the model's ability to generalize across different datasets.

AdaBoost operates by combining multiple weak classifiers to form a strong classifier through an iterative process. Given a training dataset $(X_1 Y_1), (X_2 Y_2), \dots, (X_N Y_N)$ where X_i represents feature vectors and Y_i denotes class labels, the algorithm assigns initial weights $w_i = \frac{1}{N}$ to each sample. At each iteration t , a weak classifier $h_t(X)$ is trained to minimize classification error in Eq. (4):

$$\epsilon_t = \sum_{i=1}^N w_i \mathbb{I}(h_t(X_i) \neq Y_i) \quad (4)$$

where, \mathbb{I} is the indicator function. The classifier's weight α_t is then computed as Eq. (5):

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right) \quad (5)$$

Indicating its importance in the final decision. The sample weights are updated to emphasize misclassified instances by Eq. (6):

$$w_i^{(t+1)} = w_i^{(t)} e^{\alpha_t \mathbb{I}(h_t(X_i) \neq Y_i)} \quad (6)$$

Ensuring that subsequent classifiers focus more on challenging examples. The final strong classifier is obtained as Eq. (7):

$$H(X) = \text{sign} \sum_{t=1}^T \alpha_t h_t(X) \quad (7)$$

Through this process, AdaBoost enhances the model's ability to differentiate subtle variations in facial expressions, leading to improved classification performance. Its adaptive weighting mechanism allows it to handle challenging datasets effectively, ensuring higher recognition accuracy across various test scenarios.

2.5 Model evaluation

The performance of the proposed model was evaluated using Confusion Matrix analysis and Overall Performance Metrics. The Confusion Matrix provided a detailed breakdown of correctly and incorrectly classified instances across different emotion categories, allowing for an in-depth assessment of the model's classification performance. Based on the Confusion Matrix, various evaluation metrics, including Classification Accuracy (CA), Precision (Prec), Recall, F1-Score (F1), and Area Under the Curve (AUC), were calculated to measure the effectiveness of the model.

CA represents the proportion of correctly classified instances to the total number of instances and is computed as Eq. (8):

$$CA = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

where TP (True Positive) and TN (True Negative) denote correctly predicted instances, while FP (False Positive) and FN (False Negative) indicate misclassified instances. Accuracy provides an overall measure of model performance but may not be suitable for imbalanced datasets.

Prec measures the proportion of correctly classified positive instances among all predicted positive instances and is given by Eq. (9):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

Precision is particularly important in scenarios where false positives must be minimized. Conversely, Recall (also known as Sensitivity) evaluates the model's ability to correctly identify all relevant instances, computed as Eq. (10):

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

A balance between Precision and Recall is captured by the F1, which is the harmonic mean of both metrics, ensuring a more reliable evaluation of model performance by Eq. (11):

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

AUC is used to assess the model's discriminatory power by analyzing the trade-off between True Positive Rate (TPR) and False Positive Rate (FPR). A higher AUC value indicates better model performance in distinguishing between different emotional states.

By leveraging these evaluation metrics, the effectiveness of the AdaBoost classifier in recognizing facial expressions was systematically analyzed. The combination of Confusion Matrix insights and statistical performance measures ensured a comprehensive assessment of the model's robustness and generalization capabilities.

3. RESULTS AND DISCUSSION

3.1 Confusion matrix

The evaluation of the classification model was conducted using confusion matrix for the three datasets: MMAFEDB, AffectNet, and JAFFE. These confusion matrix provide a comprehensive assessment of the model's predictive performance, illustrating the number of correctly and incorrectly classified instances across different emotion categories.

For the MMAFEDB dataset show in Figure 6, the model demonstrated high classification accuracy, particularly for the "Happy" and "Fear" classes, with 22,873 and 3,898 correctly classified samples, respectively. However, minor misclassifications were observed, such as 5 samples of "Happy" being classified as "Disgust" and 7 instances of "Surprise" being misclassified as "Fear." The overall distribution of predictions suggests that the model performed well, with minimal confusion among similar emotions.

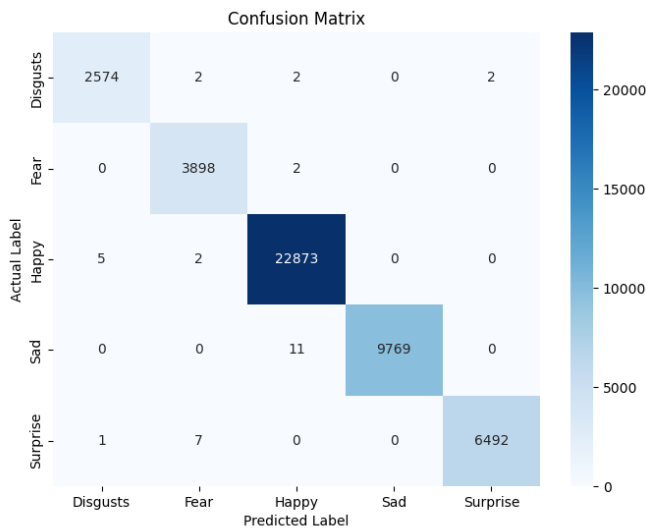


Figure 6. Confusion matrix MMAFEDB

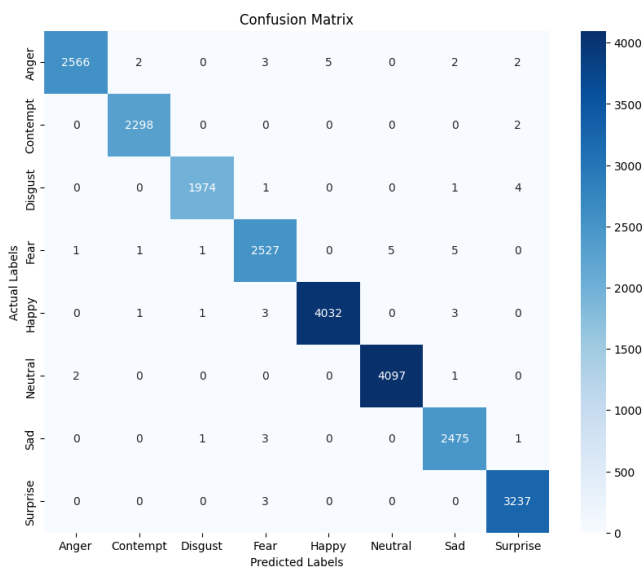


Figure 7. Confusion matrix AffectNet

In the AffectNet dataset show in Figure 7, the classification

model also exhibited strong performance. The majority of emotions, such as "Happy" (4,032 correctly classified), "Neutral" (4,097), and "Fear" (2,527), were accurately identified. However, some misclassification patterns emerged, particularly between "Fear" and "Sad," as well as between "Surprise" and "Anger." These misclassifications could be attributed to the subtle similarities between these emotions in facial expressions.

The JAFFE dataset was shown in Figure 8, which consists of a smaller number of samples, the confusion matrix indicates an overall robust classification performance. The model accurately classified most emotions, including 33 samples of "Happy" and 29 samples of "Sadness." However, slight misclassifications were noted, such as 2 instances of "Fear" being classified as "Anger" and 1 sample of "Surprise" being predicted as "Fear." These misclassifications suggest potential limitations in handling complex expressions due to the limited dataset size.

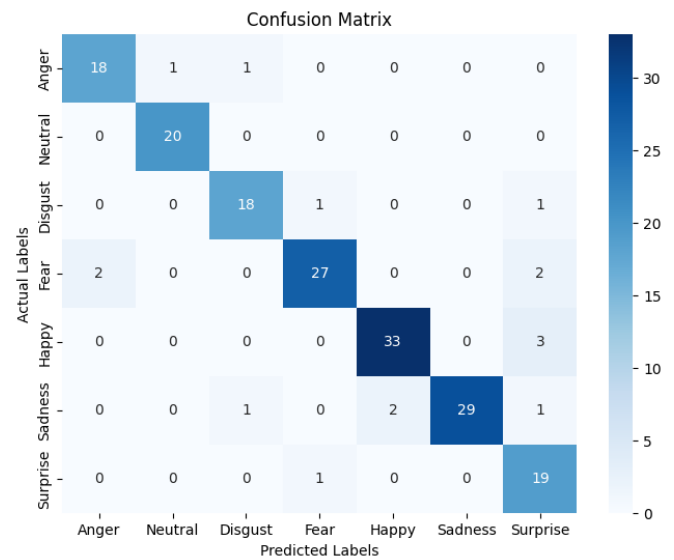


Figure 8. Confusion matrix JAFFE

The confusion matrix indicate that the classification model performs well across all datasets, with the highest accuracy for distinct emotions like "Happy" and "Neutral." However, minor misclassifications were observed for emotions that share similar facial features. These results suggest that additional fine-tuning, data augmentation, or advanced feature extraction techniques could further enhance model accuracy.

3.2 Overall performance analysis

The performance analysis presented in Table 1 highlights the effectiveness of the proposed model across three datasets: MMAFEDB, AffectNet, and JAFFE. The results indicate outstanding classification performance, particularly for MMAFEDB and AffectNet, where the model achieved an AUC, Accuracy, F1-Score, Precision, and Recall of 99.9% and 99.8%, respectively. These high scores suggest that the model demonstrates remarkable robustness and generalizability when dealing with large and diverse datasets.

For the JAFFE dataset, the model performance is slightly lower but still maintains a high level of accuracy. The AUC score of 94.9% and accuracy of 91.1% indicate that the model can effectively classify facial expressions in JAFFE, albeit with a small performance drop compared to the other datasets.

The F1-Score, Precision, and Recall values, which remain above 91%, further confirm the reliability of the model in handling smaller datasets with limited variations.

The highest results on MMAFEDB and AffectNet suggest that the model is well-suited for large-scale datasets with extensive feature variations. However, the lower scores on

JAFFE indicate that additional fine-tuning or domain adaptation techniques may be beneficial to further enhance the model's performance on smaller datasets. Overall, the high scores across all metrics demonstrate the robustness of the proposed method in facial expression recognition tasks.

Table 1. Performance analysis

Dataset	AUC	Accuracy	F1-Score	Precision	Recall
MMAFEDB	99,9%	99,9%	99,9%	99,9%	99,9%
AffectNet	99,9%	99,8%	99,8%	99,8%	99,8%
JAFFE	94,9%	91,1%	91,3%	91,9%	91,1%

4. CONCLUSIONS

This study introduced a hybrid Deep Autoencoder and AdaBoost model for robust facial expression recognition across MMAFEDB, AffectNet, and JAFFE datasets. The results demonstrated outstanding classification performance, particularly for large and diverse datasets, achieving an AUC and Accuracy of 99.9% and 99.8% on MMAFEDB and AffectNet, respectively. The confusion matrix analysis showed high accuracy in recognizing distinct emotions like "Happy" and "Fear," with minimal misclassifications among similar expressions. However, the model exhibited a slight performance drop on the JAFFE dataset, which has fewer samples, achieving an AUC of 94.9% and an accuracy of 91.1%. This indicates the need for further fine-tuning to enhance performance on smaller datasets.

Overall, the proposed model has proven its effectiveness in handling complex facial expressions with high accuracy and generalizability. The results suggest that the combination of Deep Autoencoder and AdaBoost is well-suited for large-scale datasets with extensive feature variations. However, for datasets with limited diversity, additional techniques such as domain adaptation or advanced feature extraction could further optimize classification performance. Future research could focus on refining the model's ability to distinguish subtle facial expressions and improve robustness across different datasets.

REFERENCES

- [1] Boughanem, H., Ghazouani, H., Barhoumi, W. (2023). Facial emotion recognition in-the-wild using deep neural networks: A comprehensive review. *SN Computer Science*, 5(1): 96. <https://doi.org/10.1007/s42979-023-02423-7>
- [2] Karnati, M., Seal, A., Bhattacharjee, D., Yazidi, A., Krejcar, O. (2023). Understanding deep learning techniques for recognition of human emotions using facial expressions: A comprehensive survey. *IEEE Transactions on Instrumentation and Measurement*, 72: 5006631. <https://doi.org/10.1109/TIM.2023.3243661>
- [3] Kuruvayil, S., Palaniswamy, S. (2022). Emotion recognition from facial images with simultaneous occlusion, pose and illumination variations using meta-learning. *Journal of King Saud University-Computer and Information Sciences*, 34(9): 7271-7282. <https://doi.org/10.1016/j.jksuci.2021.06.012>
- [4] Sarker, I.H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3): 160. <https://doi.org/10.1007/s42979-021-00592-x>
- [5] Pise, A.A., Alqahtani, M.A., Verma, P., K, P., Karras, D.A., Halifa, P.S.A. (2022). Methods for facial expression recognition with applications in challenging situations. *Computational Intelligence and Neuroscience*, 2022(1): 9261438. <https://doi.org/10.1155/2022/9261438>
- [6] Berahmand, K., Daneshfar, F., Salehi, E.S., Li, Y., Xu, Y. (2024). Autoencoders and their applications in machine learning: A survey. *Artificial Intelligence Review*, 57(2): 28. <https://doi.org/10.1007/s10462-023-10662-6>
- [7] Gong, X., Yu, L., Wang, J., Zhang, K., Bai, X., Pal, N.R. (2022). Unsupervised feature selection via adaptive autoencoder with redundancy control. *Neural Networks*, 150: 87-101. <https://doi.org/10.1016/j.neunet.2022.03.004>
- [8] Kopalidis, T., Solachidis, V., Vretos, N., Daras, P. (2024). Advances in facial expression recognition: A survey of methods, benchmarks, models, and datasets. *Information*, 15(3): 135. <https://doi.org/10.3390/info15030135>
- [9] Mienye, I.D., Sun, Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. *IEEE Access*, 10: 99129-99149. <https://doi.org/10.1109/ACCESS.2022.3207287>
- [10] Gamil, S., Zeng, F., Alrifay, M., Asim, M., Ahmad, N. (2024). An efficient AdaBoost algorithm for enhancing skin cancer detection and classification. *Algorithms*, 17(8): 353. <https://doi.org/10.3390/a17080353>
- [11] Ekundayo, O.S., Viriri, S. (2021). Facial expression recognition: A review of trends and techniques. *IEEE Access*, 9: 136944-136973. <https://doi.org/10.1109/ACCESS.2021.3113464>
- [12] Ali, W., Tian, W., Din, S.U., Iradukunda, D., Khan, A.A. (2021). Classical and modern face recognition approaches: A complete review. *Multimedia Tools and Applications*, 80: 4825-4880. <https://doi.org/10.1007/s11042-020-09850-1>
- [13] Rokhsaritalemi, S., Sadeghi-Niaraki, A., Choi, S.M. (2023). Exploring emotion analysis using artificial intelligence, geospatial information systems, and extended reality for urban services. *IEEE Access*, 11: 92478-92495. <https://doi.org/10.1109/ACCESS.2023.3307639>
- [14] Rifai, A.M., Raharjo, S., Utami, E., Ariatmanto, D. (2024). Analysis for diagnosis of pneumonia symptoms using chest X-ray based on MobileNetV2 models with image enhancement using white balance and contrast

- limited adaptive histogram equalization (CLAHE). *Biomedical Signal Processing and Control*, 90: 105857. <https://doi.org/10.1016/j.bspc.2023.105857>
- [15] Ullah, S., Ou, J., Xie, Y., Tian, W. (2024). Facial expression recognition (FER) survey: A vision, architectural elements, and future directions. *PeerJ Computer Science*, 10: e2024. <https://doi.org/10.7717/peerj-cs.2024>
- [16] Rifa'i, A.M., Zy, A.T., Hadikristanto, W., Butsianto, S. (2024). Leveraging data augmentation and dropout layer in MobileNetV3 for accurate skin cancer detection ISIC dataset. In 2024 8th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, pp. 591-596. <https://doi.org/10.1109/ICITISEE63424.2024.10730105>
- [17] Iqbal, J.M., Kumar, M.S., Mishra, G., GR, A., AN, S., Karthik, A. N, B. (2023). Facial emotion recognition using geometrical features based deep learning techniques. *International Journal of Computers Communications & Control*, 18(4): 4644. <https://doi.org/10.15837/ijccc.2023.4.4644>
- [18] Nan, Y., Ju, J., Hua, Q., Zhang, H., Wang, B. (2022). A-MobileNet: An approach of facial expression recognition. *Alexandria Engineering Journal*, 61(6): 4435-4444. <https://doi.org/10.1016/j.aej.2021.09.066>
- [19] Rifa'i, A.M., Utami, E., Ariatmanto, D. (2022). Analysis for diagnosis of pneumonia symptoms using chest X-Ray based on Resnet-50 models with different epoch. In 2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, pp. 471-476. <https://doi.org/10.1109/ICITISEE57756.2022.10057805>
- [20] Yang, L., Zhang, H., Li, D., Xiao, F., Yang, S. (2021). Facial expression recognition based on transfer learning and SVM. In 3rd International Conference on Artificial Intelligence and Computer Science (AICS), Beijing, China, pp. 012015. <https://doi.org/10.1088/1742-6596/2025/1/012015>
- [21] Agrawal, I., Kumar, A., Swathi, D.G., Yashwanthi, V., Hegde, R. (2021). Emotion recognition from facial expression using CNN. In 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC), Bangalore, India, pp. 1-6. <https://doi.org/10.1109/R10-HTC53172.2021.9641578>
- [22] Uniyal, S., Agarwal, R. (2024). Analyzing facial emotion patterns in AffectNet with deep neural networks. In 2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N), Greater Noida, India, pp. 801-806. <https://doi.org/10.1109/ICAC2N63387.2024.10894781>
- [23] Owusu, E., Wiafe, I. (2021). An advance ensemble classification for object recognition. *Neural Computing and Applications*, 33(18): 11661-11672. <https://doi.org/10.1007/s00521-021-05881-3>
- [24] Maharana, K., Mondal, S., Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1): 91-99. <https://doi.org/10.1016/j.gltp.2022.04.020>
- [25] Poonia, R.C., Samanta, D., Prabu, P. (2023). Design and implementation of machine learning-based hybrid model for face recognition system. In *Information Systems for Intelligent Systems: Proceedings of ISBM 2022*, Bangkok, Thailand, pp. 59-68. https://doi.org/10.1007/978-981-19-7447-2_6
- [26] Kommineni, J., Mandala, S., Sunar, M.S., Chakravarthy, P.M. (2021). Accurate computing of facial expression recognition using a hybrid feature extraction technique. *The Journal of Supercomputing*, 77(5): 5019-5044. <https://doi.org/10.1007/s11227-020-03468-8>
- [27] Yang, S., Xiao, W., Zhang, M., Guo, S., Zhao, J., Shen, F. (2022). Image data augmentation for deep learning: A survey. *arXiv preprint arXiv:2204.08610*. <https://doi.org/10.48550/arXiv.2204.08610>
- [28] McClarren, R.G. (2021). Unsupervised learning with neural networks: Autoencoders. In *Machine Learning for Engineers: Using Data to Solve Problems for Physical Systems*, pp. 195-218. https://doi.org/10.1007/978-3-030-70388-2_8