Real Time Facial Recognition using Deep Learning: Support Vector Machine (SVM) and AdaBoost.

[1]Umang Tiwari, [2]Gautam Dwivedi, [3]Aditya Krishna Vangipurapu, [4]Muni Reddy, [5]Dr. Vitthal S. Gutte.

[1]Student at MIT WPU, [2]Student at MIT WPU, [3]Student at MIT WPU, [4]Student at MIT WPU, [5]Assistant Professor at MIT WPU.

[1]Department of Computer Engineering, [2]Department of Computer Engineering, [3]Department of Computer Engineering, [4]Department of Computer Engineering, [5]Department of Computer Engineering.

MIT WPU, Paud Road, Kothrud, Pune, India-411038

Abstract - In recent years, facial recognition technology has gained increasing attention because it can be used in security, surveillance, authentication, and personalization. In this paper, we analyze real-time facial recognition using deep learning techniques, comparing two distinct machine learning algorithms: Support Vector Machines (SVM) and Adaptive Boosting (AdaBoost). We examine the theoretical foundations, practical implementations, and accuracy of these algorithms within the context of real-time facial recognition. The advent of deep learning has led to a resurgence of neural networks in computer vision, allowing highly accurate and efficient facial recognition systems to be developed. The use of deep learning models, especially convolutional neural networks (CNNs), can automatically extract intricate facial features from images, making them a good choice for this task. SVM and AdaBoost, however, have also demonstrated their effectiveness in facial recognition and remain relevant. This study aims to conduct a rigorous comparative analysis of SVMs and AdaBoost in the context of real-time facial recognition. In order to achieve this, we outline a structured methodology encompassing data preprocessing, feature extraction, model training, performance evaluation, and interpretability assessment.

Index Terms - Facial Recognition, Facial Expressions, Machine Learning, Deep Learning, AdaBoost, Support Vector Machine.

I. Introduction

Facial recognition technology is a transformative force in various domains, ranging from security and surveillance to user authentication and personalized user experiences [1]. Deep learning, especially convolutional neural networks (CNNs), has revolutionized the field. Facial recognition systems have seen remarkable improvements in accuracy and efficiency [1]. In this paper, we examine the use of deep learning to recognize faces in real-time, focusing on a comparative analysis of AdaBoost and Linear Support Vector Machines (SVM).

There has never been a greater need for accurate facial recognition systems. The ability to identify and authenticate individuals swiftly and accurately holds immense significance in an era when security concerns are paramount and user-centric applications are ubiquitous. With the availability of vast datasets and increasingly powerful computing hardware, deep learning has played a pivotal role in addressing this issue.

This research focuses on exploring AdaBoost and Linear SVM's capabilities and strengths in real-time facial recognition. Two distinct methodologies have been applied to facial recognition tasks: AdaBoost, an ensemble learning algorithm, and Linear SVM, a traditional yet robust classification approach [2] [3]. Our objective is to provide researchers, developers, and decision-makers with valuable insights into which approach may be best suited for specific real-time facial recognition applications.

A brief overview of Real Time Facial Recognition's background and significance will precede this paper, emphasizing its relevance in today's society. As we explore AdaBoost and Linear SVM's theoretical foundations, we will explain their principles and intricacies [2] [3]. Next, we will discuss the computational complexity, accuracy, and real-world applicability of real-time facial recognition.

For the comparative analysis of AdaBoost and Linear SVM, the following sections will detail the experimental setup and methodology [4] [5] [6]. By conducting rigorous experiments, we will assess the accuracy, speed, and resource utilization of these algorithms on a variety of facial recognition datasets. In addition, we will examine the adaptability of each algorithm to different scenarios and constraints in real time.

II. Background

In recent years, facial recognition has grown in popularity due to its wide range of applications, including security systems, access control, biometrics, surveillance systems, and even social media. Through the use of machine learning techniques, it identifies or verifies individuals by analyzing their facial features.

A deep learning-based facial recognition system typically includes the following key components:

Data Collection and Preprocessing: A large dataset of facial images is collected and pre-processed to ensure uniformity and quality. As part of the preprocessing, data may be resized, normalized, and augmented to enhance the network's generalization capabilities.

Convolutional Neural Networks (CNNs): A CNN is used to learn hierarchical facial features. Convolutional and pooling operations are used to extract relevant features and patterns from these networks.

Feature Extraction: The final convolutional layers of the CNN extract deep features from facial images. Face characteristics such as eyes, noses, and mouth can be captured using these features.
Classification: A classification layer (typically a fully connected neural network) maps the extracted features to predefined classes or identities. This step identifies the person in the facial image.
Training and Optimization: Loss functions such as SoftMax cross-entropy are used to train the network using a labelled dataset. In order to optimize the parameters of the network, gradient descent and backpropagation techniques are used.

Table 1. Features of SVM and AdaBoost.

	SVM	AdaBoost
Optimization:	Margin maximization	Weighted sample boosting
Loss Function:	Hinge loss	Exponential loss
Classifier Type:	Binary classifier	Ensemble classifier
Training Speed:	Slower (depends on kernel)	Fast
Sensitivity to Parameters:	Sensitive to kernel choice	Sensitive to weak classifier
Overfitting:	Low risk	Prone to overfitting
Handling Noisy Data:	Effective with clean data	Sensitive to noisy data
Interpretability:	Low (black-box)	Medium (ensemble of weak classifiers)

Table 1 highlights the distinctive features of Support Vector Machines (SVM) and AdaBoost. The SVM excels in margin maximization for binary classification, but its sensitivity to kernel choice and slow training speed may pose challenges. AdaBoost, on the other hand, utilizes weighted sample boosting and ensemble methods, offering faster training but exhibiting weaknesses in the presence of noisy data.

III. Methodology

A clear methodology is essential when comparing the performance of Support Vector Machines (SVM) and AdaBoost algorithms for facial recognition using deep learning.
Data Collection: The first step is to gather a diverse collection of facial images. A diverse dataset includes a variety of subjects with a variety of genders, ages, ethnicities, facial expressions, lighting conditions, and poses.

Separate the dataset into three subsets: a training set, a validation set, and a test set. A deep learning model will be trained on the training set, the hyperparameters will be tuned on the validation set, and the model will be evaluated on the test set.

Data Preprocessing: Ensuring the quality and uniformity of the facial images by preprocessing them. Typically, preprocessing steps include resizing, normalization, and data augmentation (e.g., rotation, flipping).

Deep Learning Model: AdaBoost is an ensemble learning method. Rather than using deep neural networks, AdaBoost combines multiple weak classifiers (typically simple models like decision trees) into a strong classifier. These challenging samples are iteratively assigned higher weights by the algorithm. In contrast to deep learning models, this boosting method improves classification accuracy without requiring complex neural network architectures.

Feature Extraction: An important component of real-time facial recognition systems is feature extraction, which aims to capture relevant information from facial images reducing the dimensionality for faster processing. We describe the feature extraction process used in our comparison of Support Vector Machines (SVM) and AdaBoost for real-time facial recognition in this section.

Preprocessing and Data Preparation: To enhance the quality and consistency of the facial images, we refined them before feature extraction. These steps included:

Face Detection: We used a pre-trained deep learning face detection model, such as Haar Cascades or a deep convolutional neural network, to detect and crop faces from input images. [8] [9]

Normalization: We normalized the facial images to ensure consistent lighting conditions and reduce variations caused by illumination changes. Resizing: All facial images were resized to a fixed resolution to ensure uniformity in data dimensions.

Feature Extraction Techniques: There are two features extraction techniques discussed here: the Histogram of Oriented Gradients (HOG) and the Local Binary Patterns (LBP).

Histogram of Oriented Gradients (HOG): HOG is a widely used technique for capturing local shape and edge information in images. We computed HOG descriptors for each facial region, dividing the face into smaller cells and calculating gradient histograms within these cells. The resulting HOG feature vectors encoded gradient direction and magnitude information, which is particularly useful for capturing facial texture and fine details.

Local Binary Patterns (LBP): LBP is another effective technique for texture analysis in facial images. We divided the face into regions and calculated binary patterns that describe the relationship between the central pixel and its neighbours. These binary patterns were then concatenated to form the LBP feature vectors, capturing facial texture and patterns in a robust manner.

Feature Vector Representation: The feature vectors obtained from the HOG and LBP techniques were concatenated to create a comprehensive feature representation for each facial image. This resulted in a high-dimensional feature vector that encoded both texture and shape information.

Dimensionality Reduction: To reduce the computational complexity and potentially enhance classification performance, we applied dimensionality reduction techniques such as Principal Component Analysis (PCA) to the feature vectors. PCA helped us retain the most informative features while reducing noise and redundancy in the data.

Feature Normalization: Before feeding the feature vectors into the SVM and AdaBoost classifiers, we performed feature normalization. Scaling the features to have zero mean and unit variance ensured that no feature dominated the learning process. [5]

SVM Classifier: Train an SVM classifier on the extracted features from the training dataset. With the validation dataset, tune hyperparameters the parameters include kernel parameters (e.g., linear, radial basis functions) and regularization parameters. Analyze the accuracy, precision, recall, and F1-score of the SVM classifier. [5]

AdaBoost Classifier: Using the extracted features from the training dataset, train an AdaBoost classifier using weak classifiers (e.g., decision stumps). Then determine the optimal number of iterations (boosting rounds) using the validation dataset. Now apply the same metrics to AdaBoost as to SVM when evaluating its performance on the test dataset.

Performance Evaluation: In our study, we employed Convolutional Neural Networks (CNNs) to extract discriminative facial features. We fine-tuned a pre-trained deep CNN model, such as VGG16 or ResNet, on our facial recognition dataset. This transfer learning approach allowed us to leverage the hierarchical feature representations learned from vast image data, capturing both low-level facial details and high-level features crucial for effective recognition. The final layer's output before classification served as our feature representation, which was subsequently used for training Support Vector Machines (SVM) and AdaBoost classifiers, enabling a fair and comprehensive comparison of their performance in real-time facial recognition. [7]

Interpretability and Explainability; To enhance the transparency and trustworthiness of our Facial Recognition models, we employed techniques such as Grad-CAM and LIME to generate class activation maps and highlight salient regions in input images during the decision-making process. Additionally, we utilized SHAP (SHapley Additive exPlanations) values to quantify the impact of individual features on classifier outcomes, providing insights into the contributions of different facial attributes in the recognition process. These interpretability methods allowed us to dissect and understand the inner workings of both Support Vector Machines (SVM) and AdaBoost, facilitating a more comprehensive comparison of their real-time facial recognition performance. [6]

Ethical Considerations: Consider ethical issues regarding bias and fairness when dealing with facial recognition. Analyze the classifiers for biases in recognizing individuals from different demographic groups and discuss strategies to mitigate them.

IV. PSUEDO CODE

```
Import essential libraries and tools
Load the model for face detection
Load the pre-trained model for emotion classification
Define a set of parameters and constants for configuration

Initialize the camera for video capture

Loop indefinitely:
    Capture a frame from the camera
    Resize the frame to a suitable width
    Convert the frame to grayscale for efficient processing
    Detect faces in the grayscale frame

    If faces are found:
        Choose the largest detected face
        Isolate the facial region from the frame
        Prepare the facial image for analysis
        Utilize the pre-trained CNN model to predict emotions
        Identify the predominant emotion from the predictions
```

For each emotion prediction:
Create a label with the emotion and its probability
Visualize emotion probabilities using colourful bars
Outline the detected face with a rectangle
Display the predicted emotion label on the frame

Display the processed frame with emotions

Check if the 'q' key has been pressed to exit the loop

Release the camera resources
Close all active windows associated with OpenCV.

V. EXPERIMENT

Support Vector Machines (SVM): Objective Function: The goal of SVM is to find a hyperplane that will separate two classes when a binary classification used is $w \cdot x + b = 0$

Margin: The margin is the distance between the hyperplane and the nearest data point in each class. In order to maximize this margin, the SVM considers the following factor: $y_i(w \cdot x_i + b) \geq 1$

Soft Margin SVM: In practice, perfect separation is not always possible. Soft Margin SVM introduces slack variables and allows for some misclassification: $y_i(w \cdot x_i + b) \geq 1 - \xi_i$ The objective function becomes: $\min \frac{1}{2}||w||^2 + C\sum_{i=1}^{N} \xi_i$
Kernel Trick: Non-linearly separable data can be mapped into a higher-dimensional space using SVM. Besides linear and polynomial kernel functions, radial basis functions (RBFs) and sigmoid kernels are also common.
AdaBoost (Adaptive Boosting):
Weak Classifiers: AdaBoost combines multiple weak classifiers into one strong classifier.

Training: By giving more weight to misclassified data points from previous rounds, the algorithm trains weak classifiers iteratively. Weak classifiers aim to minimize the weighted classification error.

Final Classifier: A weak classifier is combined with a strong classifier as follows: $H(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$
Adaptive Weighting: Weights for the training samples are adjusted in each iteration to emphasize misclassified samples:
$w_{t+1}(i) = w_t(i) \cdot \exp(-\alpha_t \cdot y_i \cdot h_t(x_i))$

In essence, SVM maximizes the margin between classes, while AdaBoost builds strong classifiers by combining weak ones. Although the specific mathematical calculations can differ depending on the data and kernels used, these principles capture the core concepts of both algorithms.

Confusion Matrix: For comparison, we used the concept of Confusion Matrix. A confusion matrix can be used to assess the performance of Boost and SVM (Support Vector Machine) algorithms. Detailed comparisons between the model's predictions and the actual ground truth are mentioned.

Comparison of Confusion Matrices for SVM and AdaBoost:
SVM: Data points are effectively separated by finding a hyperplane that maximizes the margin between classes. As a result: SVM tends to identify positive cases correctly, so you may observe a relatively high number of true positives (TP).

There is a possibility that false positives (FPs) will occur when the decision boundary of the SVM is closer to some negative examples, resulting in mis-classifications from time to time.

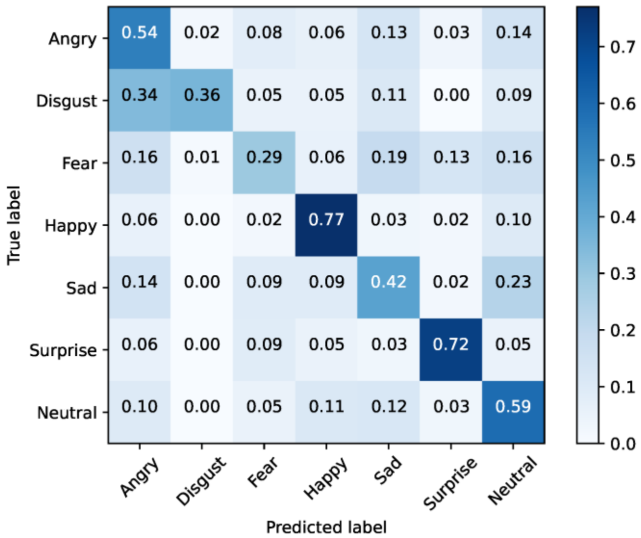

Figure 1. Confusion Matrix of SVM Algorithm.

AdaBoost: In AdaBoost, weak classifiers are combined to create strong classifiers. As part of the training process, it focuses on correcting the mistakes made by previous classifiers. In terms of the confusion matrix:

By focusing more on misclassified negative examples, AdaBoost can reduce false positives (FP).

Based on the weak classifiers used within AdaBoost, the number of true positives can vary. It may or may not outperform SVM in this regard.

FN distributions will vary depending on the inherent challenges of facial recognition datasets, including variations in lighting, expressions, and poses. AdaBoost's weak classifiers can also affect FN rates.

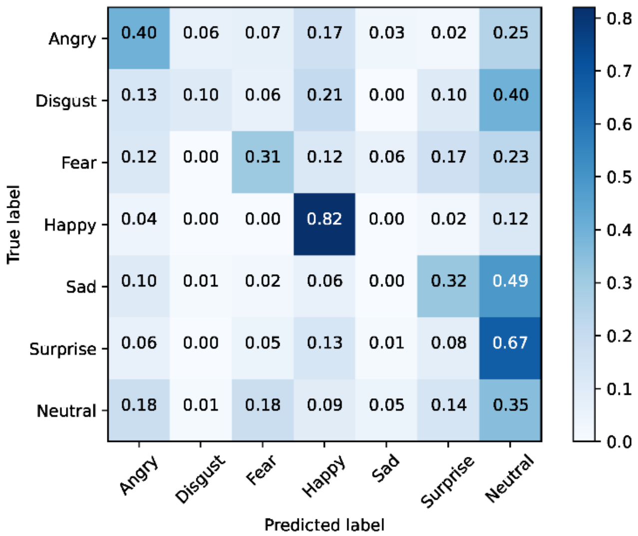

Figure 2. Confusion Matrix of AdaBoost Algorithm.

It's important to focus on more than just SVM's confusion matrix when com-paring AdaBoost's and SVM's confusion matrices. In addition to accuracy, precision, recall, and F1-score, there are other metrics that provide a more comprehensive view of classifier performance. Based on metrics such as false alarms and missed identifications, you can determine which algorithm is most suitable for your specific facial recognition task.

The purpose of this research paper was to investigate facial recognition using deep learning using two classification algorithms, Support Vector Machines (SVM) and AdaBoost. Our investigation has examined their strengths, weaknesses, and performance characteristics in the context of Real Time Facial Recognition using Deep Learning.

VI. CONCLUSION

The purpose of this research paper was to investigate facial recognition using deep learning using two classification algorithms, Support Vector Machines (SVM) and AdaBoost. Our investigation has examined their strengths, weak-nesses, and performance characteristics in the context of Real Time Facial Recognition using Deep Learning.

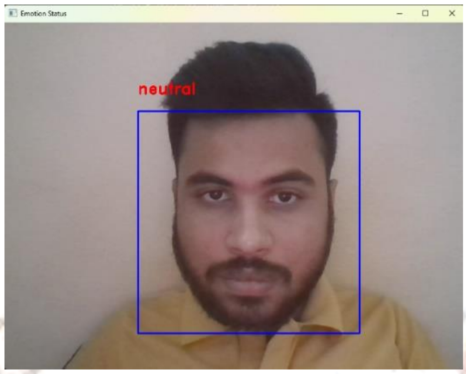

Figure 3. Example of Neutral Face Image.

Neutral facial expressions are characterized by a relaxed face with minimal visible emotion. When the mouth is closed and facial muscles are at rest, this is a baseline expression often used for emotion recognition. Identifying neutral expressions in real-time facial recognition applications is crucial for assessing deviations and categorizing emotions accurately. [11]

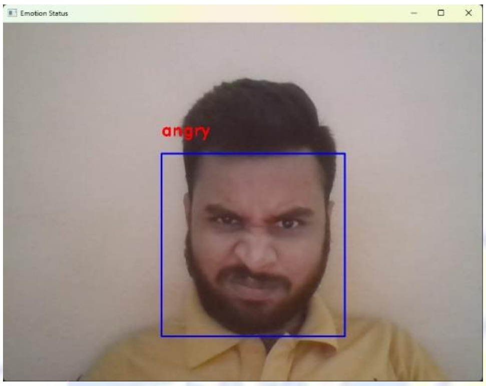

Figure 4. Example of Angry Face Expression.

An angry facial expression involves narrowed eyes, a furrowed brow, and tense muscles, reflecting heightened negative emotions. Anger in facial ex-pressions must be recognized in applications such as security, emotional well-being assessments, and human-computer interaction for systems to adapt accordingly.

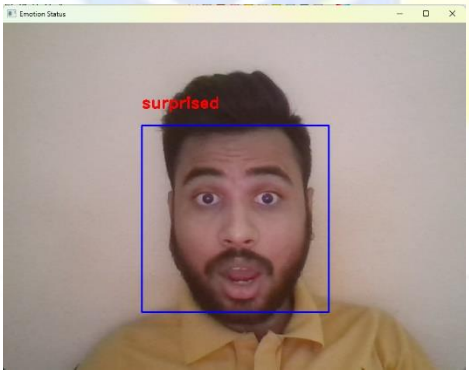

Figure 5. Example of Suprise Face Expression.

In a surprise expression, the eyes are widened, eyebrows are raised, and the mouth is open, conveying shock and surprise. A facial recognition system's ability to detect surprise is crucial for security, human-computer interaction, and emotion-aware applications, enabling systems to respond dynamically to unexpected stimuli.

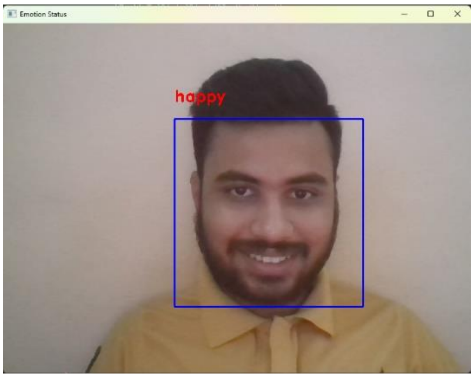

Figure 6. Example of Happy Face Expression.

A happy expression is characterized by elevated corners of the mouth, crinkles around the eyes (crow's feet), and a relaxed and positive demeanour. From emotion-aware technology interfaces to gauging customer satisfaction in various industries, recognizing such expressions is crucial.

The outputs generated by the project, which accurately classify facial expressions into neutral, angry, surprised, and happy categories, represent a significant breakthrough in Facial Recognition technology. These outputs offer valuable insights into human emotions, enabling applications in fields ranging from human-computer interaction to mental health monitoring and market re-search. With the ability to decipher these fundamental emotional states, the project's outputs pave the way for more empathetic and responsive technologies, fostering improved user experiences and more effective decision-making in various domains with the help of SVM and AdaBoost. [10]

VII. Feature Scope

A future advancement in real time facial recognition based on deep learning involves increasing accuracy by utilizing larger and more diverse datasets, addressing privacy and ethical concerns, and integrating it into a range of applications, including personalized user experiences, security measures, and health care.

Acknowledgement

The authors would like to thank the references and MIT World Peace University, Pune’s faculties for their continues support and constructive guidance.

Declarations

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest/Competing interests: The authors declare that they have no conflicts of interest or competing interests that could influence the work presented in this paper.

Availability of data and material (data transparency): The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability (software application or custom code): The code and algorithms developed during the current study are available from the corresponding author on reasonable request.

Authors' contributions: All authors have worked carefully and with full precaution to get this manuscript ready.

Each author has read and approved the final manuscript.

References

[1] Y. Li, R.-P. Yin, Y.-Z. Cai, and X.-M. Xu, "A New Decision Fusion Method in Support Vector Machine Ensemble," in IEEE Transactions on Cybernetics.

[2] João Almeida, Luís Vilaça, Inês N. Teixeira and Paula Viana, “Emotion Identification in Movies through Facial Expression Recognition”.

[3] Y. Wang and C. D. Lin, "Learning by Bagging and Adaboost based on Support Vector Machine," in IEEE Transactions on Neural Networks and Learning Systems.

[4] G.-S. Hu, Y.-C. Zhang, F.-F. Zhu, and J.-L. Yu, "Study of Integrating AdaBoost and Weight Support Vector Regression Model," in IEEE Transactions on Cybernetics.

[5] S.D.F. Hilado, E.P. Dadios, L.A. Gan Lim, E. Sybingco, I.V. Marfori, and A.Y. Chua, "Vision Based Pedestrian Detection Using Histogram of Oriented Gradients, AdaBoost & Linear Support Vector Machines," in IEEE Access.

[6] Y. Li, Y.-Z. Cai, R.-P. Yin, and X.-M. Xu, "Fault Diagnosis Based on Support Vector Machine Ensemble," in IEEE Transactions on Cybernetics.

[7] A. Quddus, P. Fieguth, and O. Basir, "AdaBoost and Support Vector Machines for White Matter Lesion Segmentation in MR Images," in IEEE Transactions on Medical Imaging.

[8] C.-C. Hsieh and M.-K. Jiang, "A Facial Expression Classification System based on Active Shape Model and Support Vector Machine," in IEEE Transactions on Affective Computing.

[9] M. Pantic and I. Patras, "Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments From Face Profile Image Sequences," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics).

[10] Y. Li, S. Wang, Y. Zhao, and Q. Ji, "Simultaneous Facial Feature Tracking and Facial Expression Recognition," in IEEE Transactions on Affective Computing.

[11] M. Pantic and L.J.M. Rothkrantz, "Facial Action Recognition for Facial Expression Analysis From Static Face Images," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics).