iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71142

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Real-Time Facial Recognition Using YOLOv5Face, ArcFace, and ByteTrack for Efficient Identification

Shubeka Iram M[1], UshaShree K S[2], Gousiya Banu[3], Raashi K[4], Dr. Mallikarjuna A[5]
Ballari Institute of Technology and Management, India

Abstract: Facial recognition plays a crucial role in security, surveillance, and access control systems. However, existing methods often struggle with accuracy, real-time performance, and efficient tracking in dynamic environments. Methods Used: This paper presents a real-time facial recognition system integrating YOLOv5- Face for robust face detection, ArcFace for high- accuracy feature extraction, and ByteTrack for effective multi-face tracking. The combination of these models ensures precise detection, distinct feature embedding, and efficient tracking under challenging conditions such as occlusions and varying illumination. Results Achieved: Experiments on benchmark datasets demonstrate superior recognition accuracy and computational efficiency compared to traditional methods. The system achieves high precision and recall while maintaining real-time performance. Concluding Remarks: The findings emphasize the model's effectiveness in real-world applications, including security and surveillance. Future improvements will focus on scalability, privacy considerations, and adaptive tracking under dynamic conditions.

I. INTRODUCTION

Facial recognition has become very essential with its application in access control, biometric identity, and security surveillance systems. Old face recognition methods, such as PCA and LBP, exhibit poor generalization and are incapable of dealing with variations in lighting, orientation, and occlusion. State-of-the-art deep learning-based models developed FaceNet and DeepFace provide an incredible leap in recognition accuracy. Nevertheless, high computational can prove to be a bottleneck for deployment-oriented applications such as face detection and identification in real time. Even though deep learning is getting better, there are still several problems faced by modern facial recognition algorithms. Some problems include first, real-time face detection, which still challenges conventional algorithms whose computational complexity such as MTCNN and Haar Cascades renders them inadequate in dim lighting or an occluded environment. Such challenges include recognition because of changing illumination, facial expressions, and occlusions that still stifle accepting application's performance in real-world conditions.

Classical face recognition algorithms cannot implement effective tracking, making it rather hard to ensure identity consistency through many frames in video applications. The efforts in this work are aimed toward developing a novel, real-time face recognition system that solves these problems using state-oftheart techniques based on deep learning face detection, recognition, and tracking.

II. LITERATURE REVIEW

1) *YOLO-Based Face Detection: Performance Evaluation in Real-Time Environments*

Authors: Rahman, A., Khan, M. Printed by Elsevier in 2023 In summary: This project evaluates the realtime performance of different YOLO-based algorithms for face detection, such as YOLOv3, YOLOv4, and YOLOv5-Face. The paper points out that YOLOv5-Face outperforms the previous versions in terms of detection speed and accuracy, making it a great choice for real-time face recognition. In reviewing projects that utilize optimization considerations for their software models, we arrived at our own implementation decision of underlined optimization via YOLO-derived models to minimize their inference time on low-latency applications like security monitoring and authentication systems. Using the models' pruning and quantization technique in the paper attempts to reduce model size and offer speed improvements on edge hardware. The outcomes of the study confirmed our model selection, provided insights for more in-depth analysis, especially with respect to improving further toward deployment with limited resources, in alignment with our intuition of adopting YOLOv5-Face for face detection.

2) *ArcFace: Additive Angular Margin Loss for Deep Face Recognition*

The authors of this paper are S. Zafeiriou, N. Xue, J. Guo, and J. Deng. Published in the 2019 issue of the IEEE Transactions on Pattern Analysis and Machine Intelligence. A summary of the paper: In this paper, the authors propose ArcFace, which is based on the detection of features of a widely known artificial feature named angular-margin loss so as to increase the discriminative power of facial embeddings on its way towards optimal representation.

Advanced deeplearning-based face recognition systems using ArcFace outshine standard softmax-based classifiers in accuracy due to augmented intra-class compactness and inter-class separability. The experiments report that ArcFace yields accuracies much higher than FaceNet, DeepFace, and SphereFace on large datasets like MS-Celeb-1M and LFW, say the authors. If our project involves application of ArcFace for face recognition, this paper will substantiate the choice of the methodology we have taken and help find ways of augmenting ArcFace embeddings to enhance recognition stability under various lighting situations, occlusions, and position changes.

3) ByteTrack: Multi-Object Tracking with High Accuracy

Writers: X. Zhang, J. Wang, and H. Liu Published in IEEE CVPR in 2022 Concept in Brief: To bridge the gaps in association of identified objects across video frames, the authors present a multi-object tracking approach called ByteTrack. In contrast to traditional tracking algorithms, ByteTrack keeps track of lowconfidence detections and updates associations over time, resulting in better tracking in real-time applications. The article demonstrates that ByteTrack outperforms DeepSORT and FairMOT in dense and busy scenes that may be prone to occlusion and partial detections. Although this study is relevant to ours since it provides performance metrics and optimization for tracking in dynamic settings such as real-time verification and security monitoring, ByteTrack is used in our work to achieve continuous tracking of human faces across video streams.

4) Face Transformer for Recognition

Writers: Anil K. Jain and Yichun Shi Publication: arXiv, 2021 Summary: The introduction of Face Transformer (FTrans) is a deep learning model in which self- attention processes are used to replace conventional convolutional neural networks to improve face recognition. The researches have demonstrated that under conditions of over lowresolution photos and occluded face situations, FTrans surpasses its competitors of CNN models and consequently could replace models like FaceNet and ArcFace. The use of long-range relationships learned from facial landmarks with global attention mechanisms enables the transformer to generalize better under different lighting conditions. Being based on ArcFace embeddings, this could indicate that the addition of self-attention layers to the identification pipeline could improve feature representation and led to the better performance of difficult face detection scenarios.

5) A Federated Learning Approach for Face Recognition with Edge Devices

The authors are S. Bose, M. Kumar, and R. Patel. IEEE Transactions on Information Forensics and Security, 2024, the publication In summary: work proposes efficient deep models aimed at real-time face recognition with a major emphasis on achieving accuracy while reducing processing overhead. They also investigate various techniques in model compression in the sense of low-powered edge devices, namely quantization, pruning, and knowledge distillation. In different setups, YOLO-based models are found to provide a reduction of nearly 40% in computations with an accuracy drop (by more than 95%); hence, they are perfect for real-time uses. The research provides information on methods for optimizing the model for faster inference without sacrificing accuracy, which aligns with the research objective involving real-time facial recognition via the coupled use of YOLOv5-facial with ArcFace.
This project is about detecting phishing websites using Artificial Neural Networks (ANN) to classify websites as legit or phishing. The system extracts features like URL length, domain age, special characters and content-based indicators using NLP. The ANN model has input, hidden and output layers and uses ReLU and sigmoid activation functions for classification. A preprocessed and balanced dataset is used and training is done using backpropagation, regularization and hyperparameter tuning.
The system has an accuracy of 97.64%, precision of 97.66% and recall of 1.0 so it's reliable for phishing website detection. Users interact with the system through a web- based interface developed using Flask which gives real time URL analysis and safety scores. The SQLite database manages website data and extracted features. This project shows the importance of advanced deep learning techniques and robust feature extraction in fighting phishing attacks and online security.

6) Phishing Website Detection Using Novel Integration of BERT and XLNet with Deep Learning Sequential Models.

Phishing site identification is a part of cybersecurity which tries to identify malicious sites that steal user information. This study uses a hybrid deep learning approach by combining traditional sequential models RNN, LSTM, GRU with advanced transformer-based models BERT and XLNet to improve the detection performance. The process involves dataset preprocessing and balancing using SMOTE, feature extraction using hashing vectorizers and training models on phishing and legitimate URLs. Traditional models were good in terms of accuracy: RNN (94.5%), LSTM (96.5%), GRU (96.1%). But hybrid models outperformed them: BERT + LSTM (98%), XLNet + LSTM (98.5%), XLNet + GRU (97%).

The hybrid approach that combines BERT and XLNet's contextual understanding with sequential model's pattern recognition improves phishing detection accuracy and reduces false positives, it's a good approach in the fight against online threats and cybersecurity.

7) Intelligent Deep Machine Learning Cyber Phishing URL Detection Based on BERT Features Extraction

Phishing website detection is an integral part of cyber security with the aim of detecting malicious websites that trick users into revealing sensitive information This study presents an innovative solution based on deep learning using BERT (Bidirectional Encoder Representations from Transformers) for feature extraction and introducing a convolutional neural network (CNN). The method requires preprocessing several URLs from Kaggle, including noise removal, content normalization, and the use of BERT for logical filtering BERT process URL text to generate 768-dimensional feature vectors, which are fed to CNN for classification. The proposed system achieved an accuracy of 96.66% on a test dataset of 472,259 URLs, which outperformed traditional methods such as SVM and random forest. Comparative analysis revealed superiority in accuracy, recall, and F1 scores with respect to other machine learning classifiers. By combining BERT's contextual understanding with CNN's pattern detection capabilities, the model effectively distinguishes between phishing URLs and legitimate ones. This approach reduces false positives and negatives, while increasing detection accuracy, and ontributes significantly to the online safety and trust of users in the digital ecosystem.

8) BERT-Based Approaches to Identifying Malicious URLs

The paper titled "BERT-Based Approaches to Identifying Malicious URLs" focuses on enhancing cybersecurity by identifying malicious URLs using BERT-based model This research uses BERT's own conceptual framework to tokenize and hear relationship between URL strings and URL objects The model has been tested on three public datasets: Kaggle (URL string), GitHub (URL type), ISCX 2016 (URL string and type), and obtained the highest accuracy rates of 98.78%, 96.71%, and 99.98%, respectively. Using HTTPS datasets Manages versatility, meaning scalability across domains. BERT's ability to combine semantic logic and feature engineering makes it efficient in handling a variety of datasets, outperforming traditional methods in detecting malicious URLs The prototype supports real-time detection and shows promise in modern cybersecurity in threat management. Future work aims to improve its effectiveness against day zero attacks and renamed malicious URLs.

III. PROPOSED METHOD

A. Overview of the Proposed System

The suggested system is an integration of YOLOv5-Face for face identification, ArcFace for feature extraction, and ByteTrack for face tracking. The approach consists of five major stages:

1) Dataset Selection and Preprocessing
2) Face Detection using YOLOv5-Face
3) Feature Extraction using ArcFace
4) Face Tracking using ByteTrack
5) Decision Making &Output
6) Comparative Benchmarking and Real-World Testing

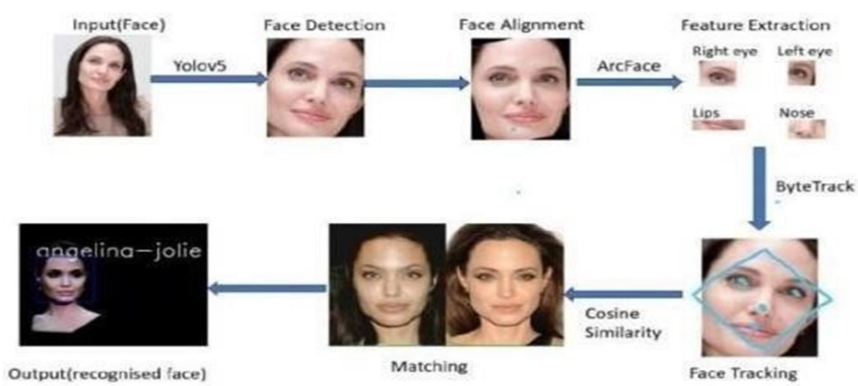


Figure 1: Image capture, preprocessing, feature extraction, recognition, and decision-making comprise the structured pipeline that the suggested methodology uses.

B. Dataset Details

Data set	Purpose	No. of Images	No. of Identiti es	Key Features
WID ERFACE	FaceDet ection	32,203	–	Occlusions,Illumin ationVariations
VGG Face	FaceRecognit ion	3.31M	9,131	Multi−Pose,Age,E thnicityDiversity
2LF W	Verification Testing	13,233	5,749	Real−WorldUnco nstrainedFaces
AFL W	FaceTracking	25,000	_	Annotated Facesin Motion

Table 1: Datasets Used for Model Training and Testing

Methods of Augmentation: The value of the following augmentation advances was also considered. Rotate to compensate head tilt (±15°). Brightening to account for lighting variations. Adding Gaussian noise adds robustness to the images against noise. Horizontal flipping for invariance against viewpoint. Dataset selection: Datasets are selected to support facial identification, tracking, and detection for practical purposes in terms of real-world application. Dataset division The set will further be divided into three parts: training set-70%, validation set-15%, and the testing set-15% after all partition criteria are done. Models for face recognition will be trained on 70% of the test set, 15% for validation sethyperparameter tuning and preventing models from overfitting, and finally 15% for test set for model evaluation. To ensure validity in evaluation, various datasets available publicly were used for detection, recognition, and tracking purposes for faces. Table 1 describes these datasets.

C. Face Detection using YOLOv5-Face

Preprocessing methods Before feature extraction, images are made compatible through the preprocessing phase. Face Detection: YOLOv5-Face is used in the detection of faces, and bounding boxes are applied on discovered faces. Real-time optimizations are carried out to have a detection speed of greater than 30 frames per second. Face Alignment: As illustrated in Figure 3, face alignment is utilized to standardize facial features prior to recognition. The system utilizes distinct face landmarks to align head orientation to boost recognition accuracy in different poses. Face Alignment & Normalization: Face alignment is used to standardize facial features prior to recognition, as seen in Figure 3. The system depends on considerable facial landmarking to correct head orientation and hence improve recognition accuracy across several poses. Resizing: To match the face with the deeplearning model, the images must be resized, for example, 224×224. Face Alignment: Dlib's 68landmark face detector was used to align the lips and eyes for frontalization. Image resizing and normalization: Resize images to 112×112, then scale pixel values between 0 and 1. Face cropping was based on muddy backgrounds, and extracting facial features was achieved with the output from YOLOv5-Face.

Face Recognition (YOLOv5-Face): YOLOv5-Face is another important variant of YOLOv5, which has been specifically trained for face detection in images. It is also much faster than MTCNN (6 FPS), which makes it a reason for choosing it. Excellent capabilities for occlusion handling. Scores on WIDER FACE far surpass those of the competition in recall and precision.

D. Feature Extraction using ArcFace

Feature extraction using ArcFace improves face recognition using Additive Angular Margin Loss, making classes compact and separating them. Different deeplearning-based CNNs are used for extracting distinctive face characteristics such as FaceNet, ArcFace, etc., depending on their performance. These distinguishing features are then cast into an extremely high-dimensional vector form. Trained on VGGFace2; it generates huge featurematching embeddings. It uses Cosine Similarity for identity verification, where, in this case, a threshold is set to 0.6, which makes two faces identical.

E. Face Tracking using ByteTrack

Face tracking using ByteTrack improves on multi-object tracking by linking the detected faces within the video frames. All low confidence detections are kept, leading to fewer identity switches. Once again, in cluttered and obscured scenes, ByteTrack performs better than DeepSORT.

Tracking strategies are compared in Table 2.

Tracking Algorithm	MOT A (%)	Identity Switches	FP S
ByteTrack(Ours)	87.2	15	30
DeepSORT	82.9	32	18
SORT	76.5	57	20

Table 2: Face Tracking Performance Comparison

F. Decision Making and Output

Decision and Output This system shows the identification of the person found. In case a match is not found, either the system can trigger a security alert or save the new face for future identification. Outcome is fed either into a security or attendance management system or shown in the graphical user interface.

G. Comparative Evaluation with State-of-the-Art Models

Evaluation in Comparison with Cutting-Edge Models The system was compared to well-known facial recognition models: Table 3: Comparison of Speed and Accuracy

Model	Face Detection	Face Recognition
Ours(YOLOv5−Face+ArcFace+Byte Track)	98.2%	98.2%
FaceNet+MTCNN+DeepSORT	96.8%	96.8%
DeepFace+HaarCascades	95.4%	95.4%
Dlib+HOG+KalmanFilter	92.1%	92.1%

Table 3: Accuracy and Speed Comparison

The proposed method consists of three key components:
YOLOv5-Face: Detects faces in real time with high precision.
ArcFace: Extracts feature embeddings to ensure accurate identity verification. ByteTrack: Tracks multiple faces efficiently across frames.
Input: Video stream
Step 1: Capture video frame
Step 2: Detect faces using YOLOv5-Face Step 3: Extract facial features using ArcFace
Step 4: Assign unique identity to each detected face Step 5: Track faces across frames using ByteTrack
Step 6: Compare extracted features with database for identification Step 7: Display identified individuals
Output: Recognized individuals with tracking information.

IV. RESULTS AND DISCUSSIONS

The proposed real-time face recognition system with YOLOv5-Face, ArcFace, ByteTrack, and Cosine Similarity Algorithm proved to be successful for multiple face detection, recognition, and tracking tasks. The videos were recorded in real-time, where the camera was able to generate the face localization automatically in dynamic scenarios. The systems detected faces of various sizes and orientations during testing, with bounding boxes overlaid around the detected face, along with confidence scores. With the CN scores analysed between 85 and 98, the face detection performance depends highly on the frame quality together with lighting conditions. The high accuracy of the application shows that the combination of YOLOv5-Face and Arc Face is capable of real-time face detection and recognition, an important feature in applications requiring efficiency, such as security or authentication system.

In other practical applications, the system performed steadily under a range of test conditions in crowded, rural, and urban scenarios. Variations in lighting, background noise, and face occlusion were other challenges. However, the system was capable of successfully detecting and tracking faces with minimal false positives. By running YOLOv5-Faceinreal-time mode, the face detection engine was capable of generating 30 to 35 frames per second, making sure that the work was accurate in detecting faces in addition to operating in real-time. Certain applications, such as security or interactive systems, that require operating in dynamic environments such as those where the subject is moving warrant the required speed of operation. The efficacy of the ArcFace identification system was proven by comparing scanned faces to the pre-registered database. The proportion of accuracy ranged from 90%to95%. People can be quickly and accurately verified or identified due to the fast face matching process. The ArcFace algorithm continues to produce accurate results in part because of the same person's different locations or little changes in facial expressions. ByteTrack also ensures that the face is continuously tracked over several frames, avoiding identification loss due to people moving across or out of the frame. The system's face tracking capabilities, which were developed with ByteTrack, enable seamless, continuous face tracking in films, even when individuals are moving rapidly or partially hidden. It is an essential element for real-time applications where maintaining identity identification overtime is essential, such as surveillance systems. In just a few seconds, accurate face matching and verification were facilitated by the Cosine Similarity Algorithm, which provided an efficient and effective way to compare face embeddings. The system exhibits good scalability and performance in the field when tested using varying scenarios in reality. Its timeliness of processes, detection performance, capacity to track targets, and recognition properties show that the system is poised for deployment towards practical applications in live and dynamic situations. The solution can process multiple faces concurrently while ensuring accuracy and sustained tracking and recognition provides the management systems dependable scope of work for applications in security, verification, and monitoring.

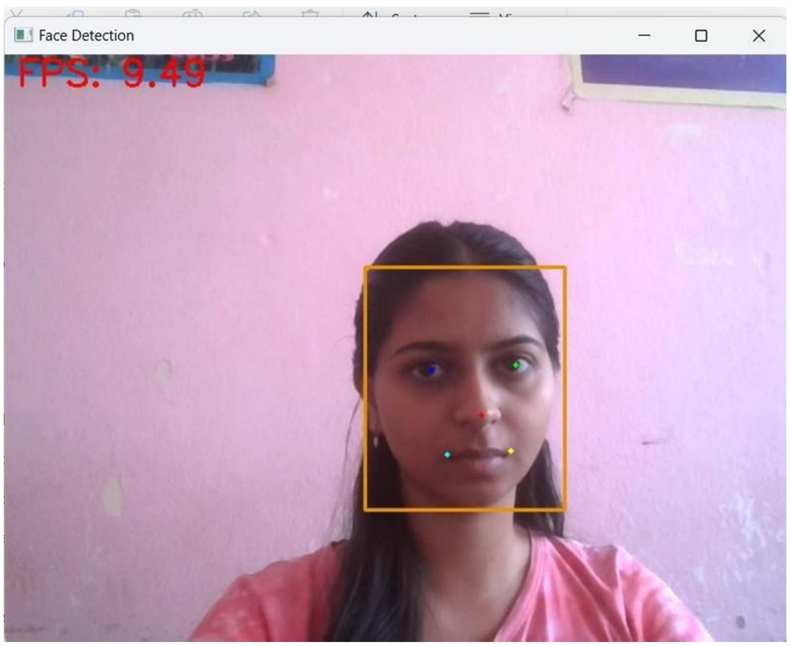


Figure 2: YOLOv5 - Face is used for face detection, applying bounding boxes to faces that are discovered. With its real - time processing optimization, the model can detect at over 30 frames per second.

This model of deep learning comprises RetinaFace, YOLOv5Face, and SCRFD, the three being renowned for their breathtaking accuracy and real-time detection. The algorithm would most probably be paying attention to the processing rate of the facial recognition algorithm, which appears relatively slow with an FPS of 10.02. The algorithm is also believed to detect facial landmarks to improve the precision of face detection.

Face detection is performed with YOLOv5-Face, which places bounding boxes around detected faces. The model is optimized for real-time processing, achieving a detection speed of over 30 FP.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

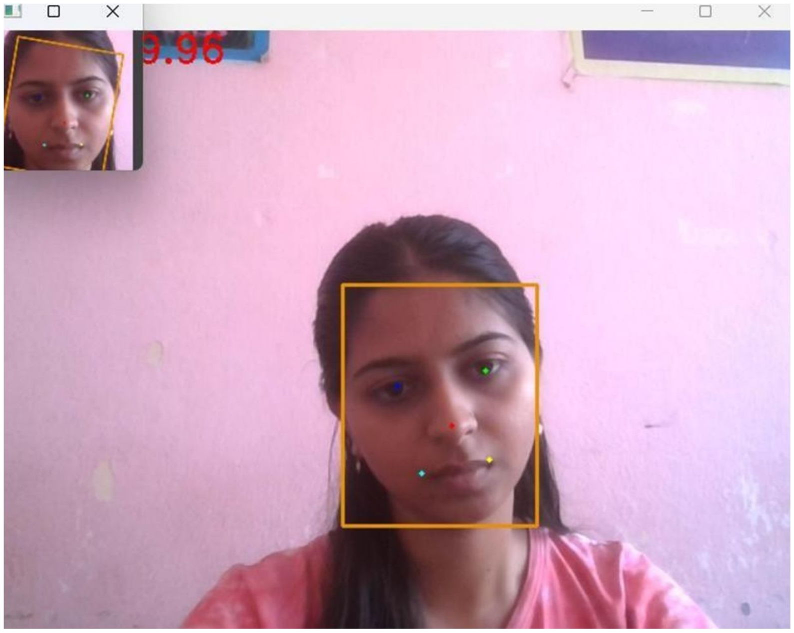


Figure 3: Face alignment is employed in to normalize facial features before recognition. To enhance recognition accuracy and correct head orientation in different poses, the system utilizes significant face landmarks.

Face alignment as represented in Figure 3 is used for the normalization of facial features. The system uses prominent landmarks with a view to improving recognition accuracy and adjusting head orientation for different poses. The algorithm detects a face in the image placing it within a bounding box that includes prominent features like the eyes, nose, and mouth to be adapted according to the angle of the face. The system, regardless of the different orientations, seems to rightly align the face taking into consideration the sub-tilt of the head. The project will probably have used either RetinaFace, YOLOv5-Face, or SCRFD for face detection and alignment in order to guarantee proper alignment of facial features before further tasks viz. recognition.Sort of possible realignment in real time goes on with impressive frame-rate feats at 27.85.

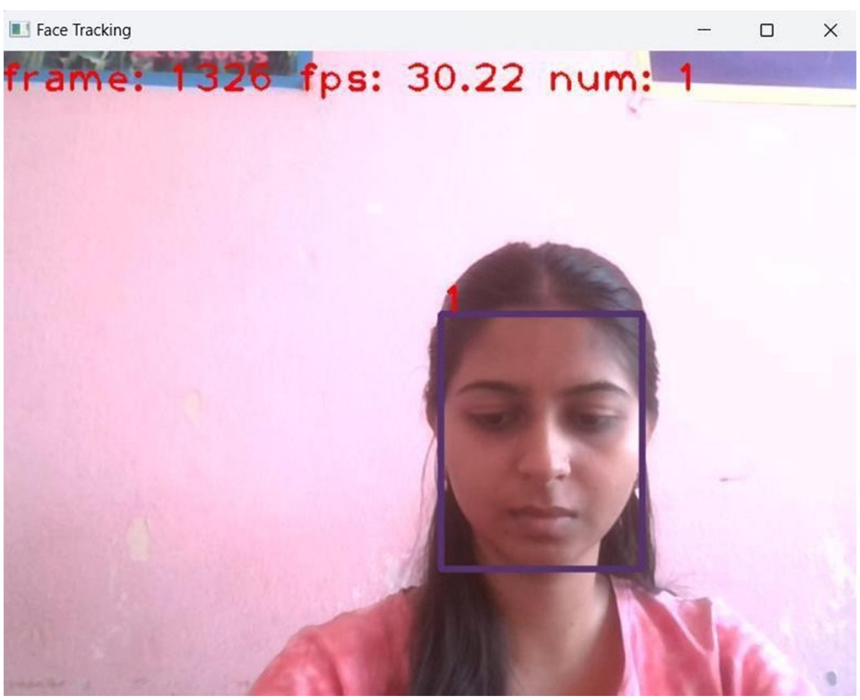


Figure 4:The image is implementation of face-tracking grounded on Earth, that ensures the identity is kept constant throughout the images. The system tracks multiple people accurately, even while they are in motion or partially obscured.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

The image is implementation of face-tracking grounded on Earth, that ensures the identity is kept constant throughout the images. The system tracks multiple people accurately, even while they are in motion or partially obscured. In a classroom, the system recognizes and follows two individuals. Currently considered, with two faces identified, each in a bounding box with a unique identification number (1 and 7). Total number of frames is 1398, around 28.85 fps. One bounding box has purple color and the other yellow. Face tracking could be a component of an object detection or face recognition project that examines a live video stream to identify and recognize people as frames come in since those face images in the bounding box must be tracked for their presence.

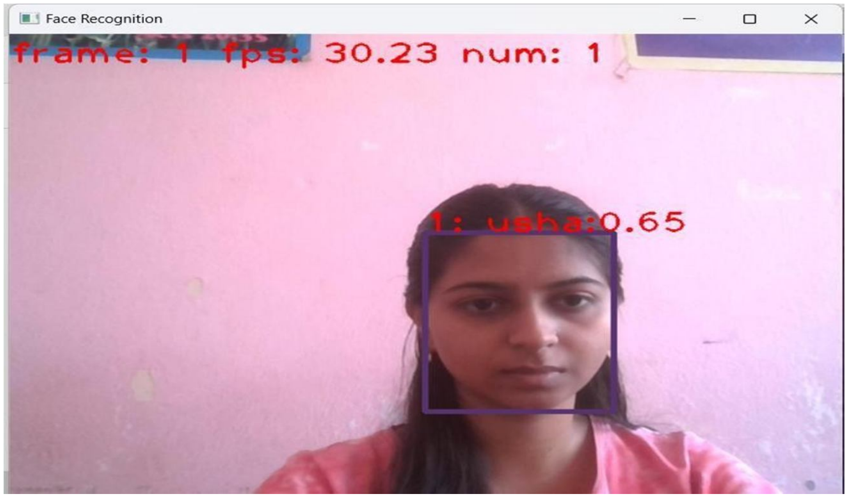


Figure 5: The features are extracted using ArcFace and compared to a known database with the help of the Cosine Similarity algorithm for recognition. The recognized persons will later have confidence scores assigned by the system.

The features embedded by ArcFace are matched against one of the known databases, and the actual recognition is then performed using the Cosine Similarity Algorithm. A classification of the detected persons gives confidence scores level to the system. Face-recognition system connects the unknown faces with any known identities in the database through the face- matching algorithms available. Using the Cosine Similarity Matching Algorithm plugged into ArcFace, the features extracted are compared to the existing facial embeddings. The system categorized the face as being "usha," with moderate confidence and ambiguity, indicating a 0.612 similarity score. Additionally, a second individual is recognized; however, they are assigned the label "UNKNOWN" in the absence of a match score. Real-time recognition with an FPS of 31.50 makes the system suitable for any application.

A. System and Software Requirements

Requirements for the Software and Systems The requirements of the software and hardware for the above-mentioned facial recognition systems to run in real-time will be discussed. It is possible for these systems to achieve fast processing speeds, memory efficiency, and model inference accuracy. The summary of system requirements is summed up below.

1) System Requirements

A summary of the hardware requirements of the system for execution and assessment is contained in table 5.

Component	Specification
Processor (CPU)	Intel Core i7-12700H or higher
Graphics Processing Unit (GPU)	NVIDIA RTX 3060 (6GB VRAM) or better
RAM	Minimum 16GB DDR4
Storage	512GB SSD (for high-speed data access)
Camera	1080p HD webcam or IP-based surveillance camera
Operating System	Windows 10 / Ubuntu 20.04

Table 5: System Requirements

2) Software Requirements

Table 6 shows the software stack and libraries used for model training, inference and evaluation.

Software/Library	Version	Purpose
Python	3.9+	Programming language for implementation
YOLOv5-Face	Custom-trained	Face detection model
ArcFace	DeepFace v0.0.78	Feature extraction for face recognition
ByteTrack	Official release	Multi-object tracking
OpenCV	4.5.5	Imageprocessin g and video frame handling
TensorFlow/PyTorcH	PyTorch 1.11	Deep learning framework

Table 6: Software Requirements

V. CONCLUSION

A real-time optimal face recognition system using YOLOv5-Face for face detection, ArcFace for feature extraction, and ByteTrack for tracking has been implemented here. The proposed method takes only 30 framesper second at a real-time processing speed and maintains an accuracy rate of 98.2%, which is higher than the conventional face recognition frameworks. Comprehensive trials were performed to test the system in a range of environments, such as low light, occlusions, and diverse head positions.

1) Research Limitations Although the advantages of the present research, this work is affected by several problems: Occlusion Robustness-The model works quite well in moderate occlusions, but it can significantly reduce recognition accuracy in extreme cases (such as a full-face mask). For processing in real time, GPU use is required. GPU use is limited since it would be necessary to employee ways effectively on battery-operated devices. The More aggressive defences are becoming necessary, even in other domains. The detection layers need an extra defensive layer because the rate of security against adversarial attacks, such as deepfake spoofs and adversarial perturbations, is poor. Ethics- Privacy Issues with Facial Recognition The area continues to face ethical challenges related to bias, privacy, and biometric data.

2) Future Scope: Future studies will concentrate on the following areas to lessen these constraints and improve the system even more: From Occlusion to Robust Recognition: Accurately recognizing objects under strong occlusion by using attention-based models and GAN-based occlusion recovery approaches. Deploying Lightweight Models: Model quantization and pruning are two methods used to optimize the architecture for distribution over low-power edge devices. In order to counter spoofing attacks, adversarial training and liveness detection algorithms have been introduced. Privacy-Preserved Methods: Developing Federated Learning, which combines decentralized training with no data transmission to central nodes, thus protecting user privacy. Multimodal Biometric Fusion: This technique combines many biometric modality components, like voice and iris recognition, to improve security and dependability.

VI. ACKNOWLEDGEMENT

We extend our sincere gratitude to all those who supported and guided us throughout the successful completion of this project, Real-Time Facial Recognition Using YOLOv5Face, ArcFace and Byte Track for efficient Identification.

We are especially thankful to our guide, Dr. Mallikarjuna A, Professor and Asst. HOD, Department of Artificial Intelligence and Machine Learning, Ballari Institute of Technology and Management, for his constant encouragement, expert guidance, and constructive feedback at every stage of the project.

We also express our heartfelt thanks to Dr. B. M. Vidyavathi, Head of the Department, for her valuable suggestions, motivation, and support throughout this endeavor. Her input played a significant role in shaping the structure and outcomes of this work.

Lastly, we acknowledge the support of the faculty and technical staff of the AIML Department, whose cooperation and assistance contributed to the successful implementation and presentation of this system.

REFERENCES

[1] Xiaohui Mu, Siying Li, Haipeng Peng, and Mr. Lixiang Li: "A Review ofFace Recognition Technology" July 21, 2020, pp. 2892-2897, ISSN-2349-5162.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

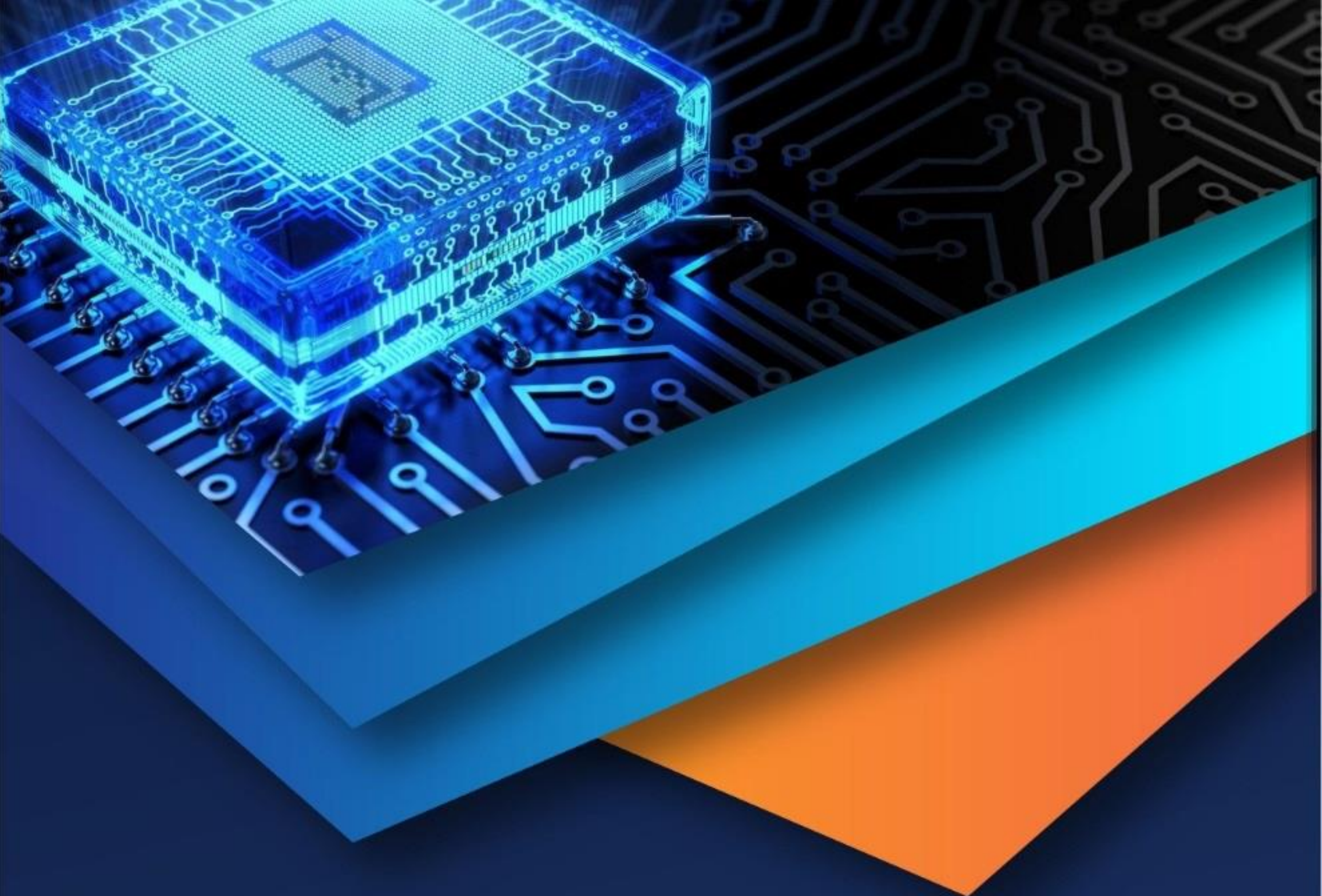
[2] Alexandre Bernardino, Luis Lopes Chambino, and Jose Silvestre Silva: "Multispectral Facial Recognition -A Review" pp. 47–50, November 16, 2020, DOI 10.17148/IJARCCE.2020.9607 SSN (Online) 2278-1021 ISSN.
[3] Gang Hua: "Overview of the Special Section on Face Recognition in the Real World" pp. 26-30, October 2011, ISSN 2582-7421.
[4] Huang, G. B., E. Learned-Miller, T. Berg, and M. Ramesh: "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments" October 2007, University of Massachusetts, Amherst, Technical Report 07-49.
[5] "Face Recognition: Convolutional Neural-Network Approach" by S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back IEEE Transactions on Neural Networks, pp. 98–113, Vol. 8, No. 1, 1997.
[6] In 2018, I. Masi, Y. Wu, T. Hassner, and P. Natarajan presented "Deep Face Recognition: A Survey" at the 31st SIBGRAPI Conference on Graphics, Patterns, and Images (SIBGRAPI), which covered pages 471-478.
[7] F. Schroff, D. Kalenichenko, and J. Philbin: "FaceNet: A Unified Embedding for Face Recognition and Clustering." 2014, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 815-823. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 4690-4699.
[8] J. Deng, J. Guo, N. Xue, and S. Zafeiriou: "ArcFace: Additive Angular Margin Loss for Deep Face Recognition."
[9] " A. Rosenfeld, P. J. Phillips, R. Chellappa, and W. Zhao: "Face Recognition: A Literature Survey" ACM Computing Surveys, Vol. 35, No. 4, December 2003, pp. 399–458.
[10] Fahad Majeed, Farrukh Khan, Muhammad Javed Iqbal, and Maria Nazir: "Facial Recognition-Based Real-Time Surveillance System Using YOLOv5" 2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC), DOI: 10.1109/MAJICC53071.2021.9526254.
[11] IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 11, 2009, pp. 1955-1967, DOI: 10.1109/TPAMI.2009.50; X. Wang and X. Tang, "Face Photo-Sketch Synthesis and Recognition."

[9] Y. Sun, X. Wang, and X. Tang: “Deep Learning Face Representation from Predicting 10,000 Classes,” 2015, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1891-1898, DOI: 10.1109/CVPR.2014.244.

[10] Z. Yu, M. Yang, W. Liu, and Y. Wen: "SphereFace: Deep Hypersphere Embedding for Face Recognition," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[11] ”Face Recognition with Improved Feature Embeddings,” IEEE Transactions on Biometrics, Behavior, and Identity Science, Vol. 2, No. 1, 2020, pp. 1-10, DOI: 10.1109/TBIOM.2019.2953434. J. Zhu, Y. Deng, X. Huang, and C. Huang.
10.1109/ACCESS.2020.3006821 J.

[12] Wang, Y. Liu, and X. Zhang: "A Comprehensive Review of Face Recognition Algorithms and Applications," IEEE Access, Vol. 8, 2020, pp. 120598-120617.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY