

# Real Time Face Recognition Using YOLO Based Deep Learning Techniques

Jyoti  
Research Scholar

Teerthanker Mahaveer University  
Moradabad  
[jyotisuryavanshi360@gmail.com](mailto:jyotisuryavanshi360@gmail.com)

Nupa Ram Chauhan  
Associate Professor

Teerthanker Mahaveer University  
Moradabad [nrcua80@gmail.com](mailto:nrcua80@gmail.com)

**Abstract**— Facial recognition is an important technology used in many real-time security systems. Just now, many innovative deep neural network algorithms have been developed to achieve impressive results in this area. This project came about because we need a more effective and affordable way to process facial recognition. We submitted a real-time facial recognition system that combines deep learning techniques, like Face Net, with traditional classifiers such as Random Forest, Neural Network and Logistic Regression. This system is designed to work well on medium hardware, making it suitable for less managed settings. A typical facial recognition system has two key components: detecting faces and recognizing them. For the face detection part, we use a technique called YOLO-Face, which is a quick and efficient real-time detector based on YoloV5. When it comes to recognizing the faces, we combine Face Net with a supervised learning method known as Random Forest (RF) to help classify them. Results from the experiments showed the Face Net + RF model reached an impressive accuracy rate of 99.8% when tested on the FDDB dataset. For same dataset, Face Net paired with Neural Network and Logical Regression achieved accuracies of 99.5% and 98.41%, respectively, while Face Net on its own attained an accuracy of 99.6%. In the end, the proposed system shows a recognition accuracy of 99.97% and processes data in just 49 milliseconds when both the face detection and classification phases are carried out at the same time.

**Keywords**— Face Recognition, security-systems, real-time security systems; machine learning; computer graphics; image process.

## I. INTRODUCTION

Facial recognition technology works correctly only when small parts come together to perform a larger task. It involves two main functions - detecting the face and recognizing it. This technology should be able to identify all types of faces in photos or videos, regardless of how they appear. Once a face is detected, the system processes the image to extract facial features using a feature extractor. These extracted features are then compared to those of enrolled faces to determine a match. Hence, implementing rapid algorithms can significantly enhance system effectiveness and boost recognition rates [1].

Humans have incredible skills to quickly identify a wide variety of objects, thanks to the brain's remarkable capacity to carry out complex operations for effective pattern recognition. Researchers have created a number of sophisticated algorithms in recent decades that allow computers to simulate some of these recognition tasks, whether in controlled environments or more unpredictable ones like video sequences. The groundwork for understanding how neurons and the brain work was established by McCulloch and Pitts' pioneering research, which marked the beginning of this quest to understand how machines can replicate human-like recognition [1].

Although face recognition has become a hot research topic in recent years, the accuracy of this technology varies greatly depending on the environment in which it is used. Performance may be hampered by a number of issues, specifically with non-restricted face recognition. Lower recognition rates can result from a number of factors, including complex lighting conditions, partial obstructions, scale variations, and pose changes. Deep learning techniques have gained popularity, and many branches of artificial intelligence have started using them for facial recognition. Strong learning capabilities, adaptability, and the capacity to generalize across various scenarios are just a few benefits of deep neural networks. Even so, there are still restrictions on real-time applications, particularly in uncertain settings where it is essential to achieve both high accuracy and computational efficiency. Face recognition is therefore still a major problem in real-time systems and is being actively researched in the fields of computer vision and deep learning. [3].

The ability of different, simpler subsystems to cooperate in order to accomplish more complicated tasks is what makes facial recognition technologies effective. At the core of facial recognition, we can identify two key operations: face detection and face recognition. A face recognition technology must operate under various conditions, detecting faces in images or videos, Irrespective of how they appear. Once a face is detected, the system processes the image to extract facial features using a feature extractor. These extracted features are then compared to those of enrolled faces to determine a match. Hence, implementing rapid algorithms can significantly enhance system effectiveness and boost recognition rates [4].

In recent times, to accelerate the face identification process, multiple deep learning techniques have adopted one-phase identifiers like YOLO (You Only Look Once) and SSD (Single Shot MultiBox Detector). Within these, the YOLO-Face technique stands out as an efficient solution for addressing the challenges of detecting faces at different scales. Built on the YOLOV5 framework, the YOLO-Face algorithm utilizes different stages of feature maps, similar to the Feature Pyramid Network (FPN), to enhance its detection capabilities [5].

Over the past ten years, the use of deep learning for face recognition has grown significantly. Nowadays, most methods use Convolutional Neural Networks (CNNs), which can range from simple to complex architectures. With techniques like DeepFace and DeepID, we have nearly achieved human-level accuracy in face recognition. In present day these successes, researchers have developed deep learning methods that can recognize faces even in challenging conditions, achieving over 99% accuracy on the dataset [6].

## II. LITERATURE REVIEW

This paper presents a face recognition technology designed to function effectively in unconstrained environments. It employs a modified YOLO-Face method in conjunction featuring a classification model aimed at enhancing detection accuracy while minimizing computation time. Additionally, we introduce an architecture that integrates a Convolutional Neural Network (CNN) with a supervised learning approach for classification purposes. While many deep learning models typically utilize the SoftMax loss function (and its variations) as the classifier in the fully connected layer, recent research indicates that alternative classifiers, such as support vector machines (SVMs), could potentially boost face recognition performance compared to the traditional SoftMax approach [7].

The author presented the latest techniques in real-time face detection and recognition. He introduce a modular automatic facial recognition system that is built on a new model for matching faces. This model utilizes features derived from Convolutional Neural Networks (CNNs) and incorporates a custom supervised learning classifier, like Support Vector Machines (SVM). Importantly, the modular design of our system allows for the easy replacement of specific algorithms, ensuring optimal performance on both GPU-enabled devices and machines that operate solely on CPUs.

The proposed face recognition technology is designed to function well in real-world environments, where circumstances can change at any time. The two primary phases of its operation are face detection and recognition. Even in difficult circumstances, it's critical to carefully determine the appropriate stages and any necessary adjustments to guarantee the best outcomes. We'll examine several face detection and recognition techniques in the following sections, along with the adjustments required to improve their effectiveness.

In past years, both traditional handcrafted methods and neural network-based learning techniques have grown in importance in the field of face recognition in recent years. Deep learning techniques adopt a different strategy than handcrafted techniques, which rely on manually chosen features to extract facial characteristics. These models use deep neural networks to automatically identify facial features by examining patterns and relationships in the input data.

The two main steps of traditional facial recognition techniques are feature extraction and classification. When it comes to feature extraction, the optimal algorithm should highlight individual differences while minimizing variations within the same person. This method streamlines the classifier's function in facial recognition. Consistent facial features such as the mouth, nose, eyes, eyebrows, and general face shape are usually the focus of feature extraction techniques. Local Binary Patterns, Gabor filters, Scale-Invariant Feature Transform, and HOG (Histogram of Oriented Gradients) are frequently employed descriptors for this purpose. Algorithms measure the separation between extracted facial features in order to recognize and differentiate people during the classification stage. Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF) are well-known classifiers in this field because of their excellent multiclass classification capabilities[10].

Over the past decade, research has shifted significantly in favour of deep learning for facial recognition. Nowadays, a lot of the common techniques use Convolutional Neural Network (CNN) architectures, which can be as basic as CNNs or as complex as more intricate variants. We are getting closer to attaining human-level performance in this area thanks to notable developments like Deep Face and Deep ID. Based on these achievements, scientists have created deep learning techniques that address the difficulties of face recognition in unrestricted settings, obtaining remarkable accuracy rates of more than 99% on the LFW dataset.

Schroff and colleagues' Face Net algorithm is one of the most notable contributions to face recognition. This approach uses a single deep neural network for both face collection and recognition. With more than 140 million settings to efficiently learn how to map face visuals, Face Net is made up of three layers that connect everything together to make sense of the data and eleven layers that assist in identifying patterns in images. This method uses triplet loss during training, which is a significant innovation. As seen in Figure 1, this method aims to increase the distance from an image of a different person while decreasing the distance between an anchor image and a matching image of the same person. An extensive dataset of 100–200 million 2D facial images, representing roughly 8 million distinct identities, was used to train the system. Consequently, on the LFW benchmark dataset, it obtained an outstanding accuracy of 99.6%.

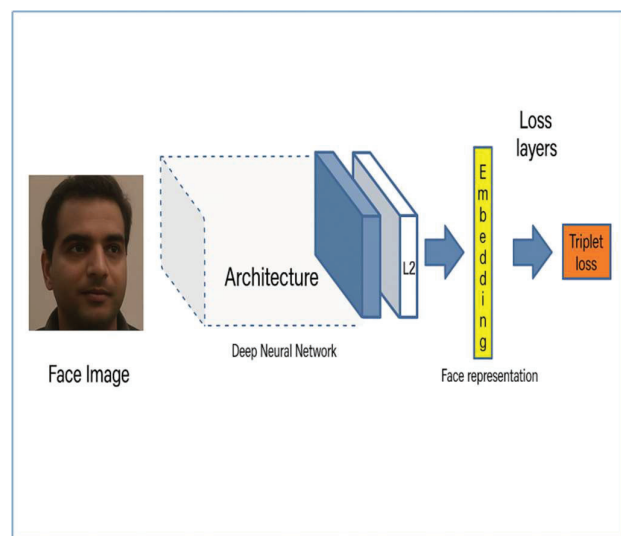


Fig.1: Framework for FaceNet

Chauhan and et al. [9] had presented a hybrid intrusion detection model that uses an Ensemble of Gradient Boosting (EGB) for classification and a 1D Residual Autoencoder (1D-RAE) for feature extraction. The approach tackles issues like as complicated attack patterns, skewed datasets, and high-dimensional data that arise in cloud systems. While EGB increases detection accuracy and lowers false positives/negatives, the 1D-RAE improves feature representation by identifying geographical and temporal patterns in traffic data. The model was effective for real-time, scalable cloud security when tested on benchmark datasets, outperforming conventional techniques in accuracy, precision, recall, and F1 score.

TABLE I. SUMMARY OF THE RELATED WORK

Author- Model / Algo. applied	Workflow	Implementation	Outcomes	Drawbacks	Future Directions
Paul Viola and Michael Jones (2014) Viola-Jones	Heer-like Feature and Ada boost classifier	Real-time face recognition	High recognition rate and speed	Prone to false positives in challenging Condition	Improve robustness to variations in lighting, pose, and occlusion
Navneet Dalal, Bill Triggs. (2005) Histogram of Oriented Gradients (HOG)	Feature extraction using edge orientation	Face detection and recognition	High accuracy in controlled environments	Computationally expensive	Explore faster and more efficient feature extraction techniques
Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton (2012) Deep Learning (CNN)	CNN Model	Face detection and recognition	Advanced performance	Requires a large dataset for training	Develop a smaller and more efficient CNN
Alexander M Rush Sumit Chopra Jason Weston (2015) Attention-Based Model	Attention mechanism to focus on relevant facial features	Face Recognition	Improved accuracy and robustness	Complex modules and training process	Investigate the attention mechanism for real-time face recognition
Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi (2016) YOLO (You Only Look Once)	Uses YOLO for real-time F. D. combined with feature extraction for recognition.	Implemented in Python with Tensor Flow or PyTorch frameworks.	Achieves fast and accurate face detection with recognition capabilities.	Struggles with very small faces or occlusions in the image.	Improving recognition accuracy under varying lighting conditions.
Yi Zhang, Jie Pang, Baicheng Li, Jianfeng Luo (2023) YOLOv5 + Siamese Network	YOLOv5 for detection, Siamese Network for face verification.	Integrated with GPU acceleration for speed.	Improved precision in identifying and verifying faces.	Requires high computational resources.	Optimization for edge devices to reduce resource dependency.

### III. PROPOSED METHODOLOGY

The methodology of this system is divided into four main phases: face detection, data collection, Data pre-processing and feature extraction. In the first stage - one or more faces are detected in the image or video frame. The second stage involves image processing techniques to pre-process the face image before passing it on to the machine learning model. In the third stage, would be the extraction of the face image features. We compare facial features with the previously saved facial features in the last stage.

#### 3.1. Face Detection Stage

The face identification step, which is in charge of identifying face pictures in the input frames in unconstrained contexts, is described in this section. This stage's face detector technique needs to be able to work with the high variability factors—like stance, scale, illumination, etc.—that are present in the photos. Following a thorough analysis of multiple face detection schemes, the chosen approach is intended for real-time applications such as YOLO-Face. The foundation of this technique is a one-stage deep neural network methodology. Unlike the quicker R-CNN, technique, these solutions estimate classes and their bounding boxes immediately without a second-stage proposal process. This method's

primary contribution is to address the speed issue that arises with two-stage approaches..

#### 3.2. YOLO-Face

The YOLO-Face system is a specialized face detection model built using the YoloV5 framework, which is based on Intelligent learning system techniques for object identification. This approach noticeably enhances the capability of face detection systems.

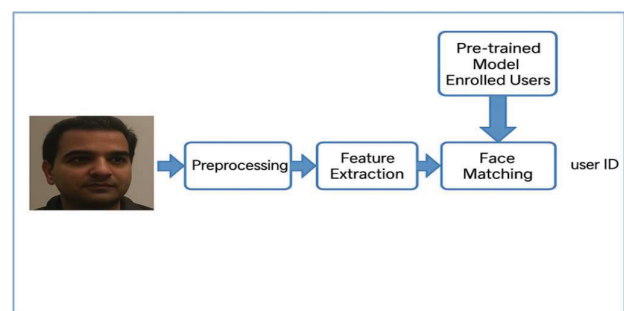


Fig.2: General real-time face recognition framework in video systems.

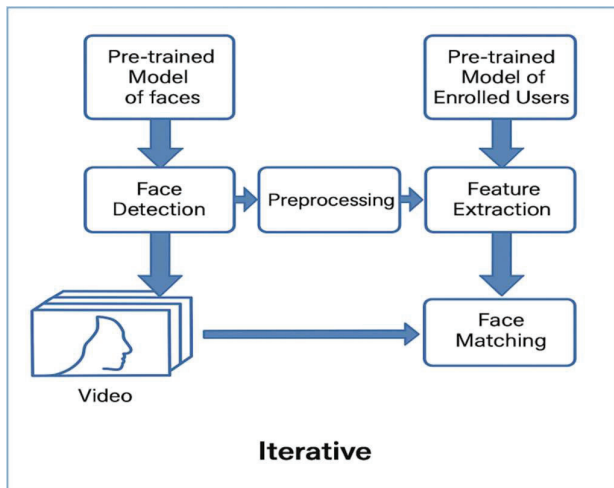


Fig.3: General real-time face recognition framework in video systems.

### Key Features of YOLO-Face

#### A. Class Prediction:

Unlike traditional multi-label classification models, YOLO-Face simplifies the task by treating it as a binary classification problem — identifying objects as either a "face" or "non-face." To improve accuracy, YOLO-Face uses a specialized loss function with four key components:

- Coordinate regression loss (to improve the accuracy of face positioning).
- Confidence loss for anchors with objects (to ensure detected faces are accurate).
- Confidence loss for anchors without objects (to reduce false positives).
- Classification loss (to improve the system's ability to distinguish faces from non-faces).

#### B. Network Architecture:

YOLO-Face is built on the darknet-53 network, which combines elements from dark net-19 (used in YOLOv2) and balanced networks. With 53 layers for pattern recognition, Darknet-53 offers improved performance compared to its predecessor, Darknet-19. YOLO-Face also outperforms ResNet-101 and achieves similar results to Res Net-152 but with better efficiency [9, 10].

#### C. Addressing Small Object Detection:

A challenge with YoloV5 is its reduced performance when detecting small-scale objects like distant or low-resolution faces. YOLO-Face addresses this by enhancing the original darknet-53 structure. Specifically, the number of layers in the first two balanced blocks is increased, from 1 and 2 layers to 4 and 8 layers, allowing the network to capture more detailed facial features early in the process. Additionally, YOLO-Face scales down feature maps more effectively using convolutional layers with a stride of 2, improving its ability to handle small faces.

#### D. Performance:

The YOLO-Face network includes 71 convolutional layers, referred to as the "deeper darknet." This deeper architecture

enhances the system's ability to extract fine details from face images. According to results shown in Figure 3, YOLO-Face achieves a detection speed comparable to the Face Attention Network (FAN), but with better overall efficiency [11, 12].

### 3.3. Datasets

This section gives a brief overview of the datasets used to evaluate the proposed system.

- **Fddb**: This dataset has 5,171 face images across 2,845 pictures. It's designed to test face detection in challenging conditions, like images with low resolution, major cover-ups, and difficult angles.
- **CelebA**: This dataset contains 202,599 celebrity face images. Each image includes 40 labeled attributes, offering a wide range of poses and background changes.
- **Easy, Medium, and Hard**: These terms likely refer to different difficulty levels in the dataset, though further details are not provided.
- **Honda/UCSD**: This video dataset is mostly used for evaluating face detection and tracking. It includes various facial expressions and poses, including rotations in both 2D and 3D planes.
- **LFW (Labeled Faces in the Wild)**: This popular dataset has over 13,000 web-sourced face visuals of 1,680 individuals, with at least two images per person. It's widely used for testing face recognition systems.
- **YTF (YouTube Faces)**: This dataset includes 3,425 video clips of 1,595 individuals. It's designed to test face recognition in real-world, unconstrained conditions.

### 3.4. The pre-analysis phase

The preprocessing stage plays a crucial role in improving the quality of features that are extracted in later stages of a face recognition system. Different preprocessing methods are applied depending on the type of data and the specific problem being addressed. In real-time facial recognition, detected face images often vary in size and lighting conditions.

**-Image Resampling:** Resampling is particularly important when using deep learning models like CNNs, which require input images of the same dimensions for effective processing. Bicubic interpolation, which smoothly resizes images to the desired size, was selected as the resampling technique in this investigation.

**-Super-Resolution Techniques:** When face images need to be enhanced to a higher resolution, super-resolution methods can improve image clarity and detail [13].

**-Normalization:** To improve system performance, the L2 norm method is used, a common normalization technique that minimizes variations in pixel values. This method reduces inconsistencies caused by differences in lighting conditions, enhancing recognition accuracy.

**-Colour Transformation:** The face detection process starts by converting colour images to grayscale, which simplifies the data and improves detection speed for methods like the Viola-Jones algorithm [14].

**-Image Rescaling:** This ensures face images are adjusted to the correct size before feature extraction. Without proper scaling, the extracted features may not align correctly with the trained model's data.

**-Image Normalization:** This step adjusts pixel value ranges to maintain consistency across images. Without normalization, the system may produce embeddings (feature representations) with excessive variation, reducing recognition accuracy [15, 16].

### 3.5. Feature Extraction Stage

Recent advancements in the process of extracting features have introduced deep convolutional neural networks (CNNs) as powerful tools for analyzing and recognizing face images. These CNN-based methods have become highly effective for learning facial features directly from image data. In the proposed system, CNN-based feature extraction (like FaceNet) is combined with traditional machine learning classifiers for improved performance. The system explores the following combinations:

- FaceNet + RF (Random Forest): Built using 100 decision trees and requiring at least 5 samples per leaf node.
- FaceNet + NN (Neural Network): Trained using a linear kernel and a regularization parameter of  $1 \times 10^3$ .
- FaceNet + KNN (K-Nearest Neighbors): Set up with 10 neighbors, equal weights, and a Murkowski distance of 2.
- FaceNet + LR (Linear Regression): It set up the categorical data from the dataset in the probabilistic function.

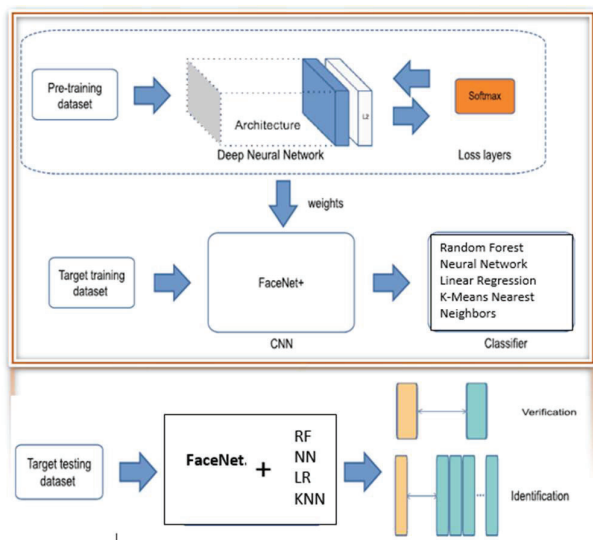


Fig. 4: A function of FaceNet + Classifier Technique

## IV. EXPERIMENTAL RESULT

This section provides a detailed evaluation of various face detection methods, including VJ (Viola-Jones), MTCNN (Multi-task Cascaded Convolutional Networks), FaceSDD, and YOLO Face. The goal is to identify the most suitable method for use in the proposed face recognition system. To assess their effectiveness, these face detection methods were tested on several well-known and challenging datasets:

- Face Detection Data Set and Benchmark
- CelebA dataset
- WIDER FACE dataset
- Honda/UCSD dataset

### 4.1. Face Recognition Evaluation

For the face recognition steps, the recommended models were evaluated using two widely recognized benchmarks:

- **Labeled Faces in the Wild (LFW)** — a dataset containing real-world face images.
- **YouTube Faces (YTF)** — dataset featuring video frames to test recognition accuracy [17, 18, 19].



Fig. 5: Face Detection of CelebA, Wider Face and FDDB

These evaluations ensure the system's effectiveness across diverse and challenging face detection and recognition scenarios

We also evaluated the same face detection models that are built using the CelebA dataset, which includes nearly 20,000 images of faces. The results, shown in Table 2, indicate that all models perform quite well. Notably, Face-SSD slightly edges out YOLO-Face with an accuracy of 99.6%, while YOLO-Face follows closely behind at 99.7%.

When we looked at the WIDER FACE validation dataset, YOLO-Face achieved scores of 99.8%, 98.2%, and 98.4% on the Easy, moderate, and Hard subsets, respectively, as highlighted in Table 3. Fig showcases some examples of face detection results from the WIDER dataset, and it's clear that YOLO-Face excels at detecting smaller faces.

Overall, these results highlight the effectiveness of YOLO-Face in various scenarios. As shown in the following Fig. 6: that it is more specific and clear image from the other images of the dataset.

In Table 1, we present the results from our experiments using the FDDB dataset. Our findings reveal that YOLO-Face achieves impressive scores of 99.8% in precision, 72.9% in recall, and 72.8% in accuracy. The only model that surpasses YOLO-Face is MTCNN, which boasts an accuracy of 81.5%.

**TABLE 1: Results of the different AI Techniques in accuracy percentage**

S.No.	Methods	Scenario A	Scenario B	Scenario C
1	FaceNet	99.41%	99.32%	99.26%
2	FaceNet+RF	99.87%	99.84%	99.81%
3	FaceNet+NN	99.58%	99.54%	99.48%
4	FaceNet+LR	98.61%	99.14%	99.21%
5	FaceNet+KNN	99.51%	99.36%	99.43%

As the implementation have don on the MTCNN and Face-SSD and YOLO-Face have compared with the following table, TABLE 2. .

**TABLE 2: Different method for the obtaining the result**

Videos	No. of frame	VJ [16]	MTCNN [25]	Face-SSD [19]	YOLO-Face [31]
1	377	75.3	93.9	97.1	96.9
2	323	58.7	92.0	98.9	99.5
3	383	58.6	83.2	94.0	96.9
4	482	51.6	81.9	96.3	97.1
5	371	46.2	83.8	96.0	95.1
6	345	44.6	93.0	93.8	97.1
7	289	46.2	83.8	96.9	98.2
8	311	43.5	81.3	89.1	94.4
9	659	81.3	98.2	98.5	99.5
10	421	61.7	91.9	95.1	94.9

## V. CONCLUSION AND FUTURE SCOPE

Real-time face recognition has become an essential part of many user authentication, security, and surveillance systems. The efficiency of face detection and recognition systems has been greatly improved by utilizing the YOLO (You Only Look Once) architecture, which is renowned for its remarkable speed and accuracy. The most recent iterations of YOLO, YOLOv8 and YOLO-NAS, have significantly improved model precision, lightweight design, and adaptability to complex scenarios, such as small-face and occluded detection.

Challenges like scale variation, low light conditions, and masked faces can be addressed by combining attention mechanisms, feature refinement, and lightweight modules, as shown by recent studies like YOLO5Face, YOLO-FaceV2, and GMC-YOLO. Face embedding and recognition accuracy have also been improved by developments in loss functions, such as triplet loss and anchor optimization (as demonstrated in Hambox and ADYOLOv5).

These models can be deployed on mobile and edge devices without sacrificing performance thanks to the combination of real-time detection and effective face recognition pipelines, as investigated in LSR-YOLO and AMLN. Furthermore, in-depth evaluations and comparative studies have produced useful standards that direct the choice of suitable YOLO-based models for particular uses.

Real-time face recognition is made reliable, scalable, and effective by combining deep learning methods with YOLO-based architectures. Future studies might concentrate on developing even more lightweight frameworks for broad implementation in real-world situations, improving privacy-aware recognition systems, and improving model generalization across a variety of demographics.



**Fig. 6: (a) Result detected by Face SSD, (b) Face Detected by MTCNN, (c) Face Detected by YOLO- Face**

## VI. REFERENCES

- [1] Khan, A., & Li, M. (2020). "A Novel Real-Time Facial Recognition System for Asian Adults Using FaceNet with SVM." In *Journal of Ambient Intelligence and Humanized Computing*, 11(2), 575-588.
- [2] Yu, Z., Huang, H., Chen, W., Su, Y., Liu, Y., & Wang, X. (2024). Yolo-facev2: A scale and occlusion aware face detector. *Pattern Recognition*, 155, 110714.
- [3] Qi, D., Tan, W., Yao, Q., & Liu, J. (2022, October). YOLO5Face: Why reinventing a face detector. In *European Conference on Computer Vision* (pp. 228-244). Cham: Springer Nature Switzerland.
- [4] Kumar, Y. J., & Valarmathi, P. (2023, March). YOLO Based Real Time Human Detection Using Deep Learning. In *Journal of Physics: Conference Series* (Vol. 2466, No. 1, p. 012034). IOP Publishing.
- [5] Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine learning and knowledge extraction*, 5(4), 1680-1716.
- [6] Omkar M. Parkhi, Andrea Vedaldi and Andrew Zisserman. "Deep Face Recognition". *Proceedings of the British Machine Vision Conference (BMVC)*, pages 41.1-41.12. BMVA Press, September 2015.
- [7] Yu, F., Zhang, G., Zhao, F., Wang, X., Liu, H., Lin, P., & Chen, Y. (2023). Improved YOLO-v5 model for boosting face mask recognition accuracy on heterogeneous IoT computing platforms. *Internet of Things*, 23, 100881.
- [8] Xu, B., Wang, C., Yu, B., Xu, W., & Du, B. (2023, September). A lightweight network face detection based on YOLOv5. In *Proceedings of the 2023 6th International Conference on Robot Systems and Applications* (pp. 157-162).
- [9] Liu, L., Wang, G., & Miao, Q. (2024). ADYOLOv5-Face: An Enhanced YOLO-Based Face Detector for Small Target Faces. *Electronics*, 13(21), 4184.
- [10] Fikry, M. (2024). Performance Analysis of Smart Technology With Face Detection Using YOLOv3 and InsightFace for Student Attendance Monitoring. *International Journal of Intelligent Systems and Applications in Engineering*, 12(4).
- [11] Zhang, X., Xuan, C., Xue, J., Chen, B., & Ma, Y. (2023). LSR-YOLO: a high-precision, lightweight model for sheep face recognition on the mobile end. *Animals*, 13(11), 1824.
- [12] Wang, Q., Zhang, P., Xiong, H., & Zhao, J. (2021). Face. evolve: A high-performance face recognition library. *arXiv preprint arXiv:2107.08621*.
- [13] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 7464-7475).
- [14] Liu, Y., Tang, X., Han, J., Liu, J., Rui, D., & Wu, X. (2020, June). Hambox: Delving into mining high-quality anchors on face detection. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 13043-13051). IEEE.
- [15] Yuan, X., Huang, H., Jiang, Z., & Xue, S. (2020, May). An Object Detection Algorithm Based on Attention Mechanism and Lightweight Network (AMLN). In *Proceedings of the 2020 the 4th International Conference on Innovation in Artificial Intelligence* (pp. 64-69).
- [16] Yuanxin, W., Deyu, Y., Meng, Y., & Meng, D. (2022, July). Anomaly Detection Model for Key Places Based on Improved YOLOv5. In *International Conference on Artificial Intelligence and Security* (pp. 51-61). Cham: Springer International Publishing.
- [17] S. Kumari, N. Kumari, and Nuparam, in *2022 11th Int. Conf. Syst. Model. Adv. Res. Trends* (2022), pp. 1466-1476.
- [18] N. R. Chauhan, R. K. Shukla, A. S. Sengar and A. Gupta, "Classification of Nutritional Deficiencies in Cabbage Leave Using Random Forest," *2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART)*, Moradabad, India, 2022, pp. 1314-1319
- [19] N. Ram and D. Kumar, "Effective Cyber Attack Detection in an IoMT-Smart System using Deep Convolutional Neural Networks and Machine Learning Algorithms," *2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE)*, Bangalore, India, 2022, pp. 1-6