Deep Learning Approaches for Criminal Face Detection and Recognition: A Comparative Study

B. Venkataratnam, PhD
Assistant Professor, Depart of Data Science, Malla Reddy University, Hyderabad, Telangana

Narsaiah Battu
Assistant Professor(C), Department of Computer Science and Applications
Dr.B.R. Ambedkar Open University, Hyderabad, Telangana

Enjamuri Avanthi Priya
Student B.Tech Final Year, Department of Computer science & engineering (CSE), Stanley College of engineering and technology for women, Hyderabad, Telangana

ABSTRACT
For contemporary law enforcement applications, especially in surveillance systems, criminal face detection and recognition are essential. Increasingly precise and effective deep learning models are required as automated systems are used increasingly frequently to identify suspects from photos or video feeds. Several cutting-edge deep learning methods for criminal face detection and recognition are examined and contrasted in this research. In particular, we examine Faster R-CNN, YOLO, SSD, and FaceNet using a proprietary criminal face database for recognition and the WIDER FACE dataset for face detection.

A thorough evaluation of face detection models based on recall, accuracy, precision, and real-time inference speed is part of our research. Faster R-CNN exhibits comparatively slower processing speeds but displays better accuracy in identifying faces in difficult situations like occlusions and changing stances. On the other hand, YOLOv4 performs quite well in real-time, which makes it perfect for applications that need to quickly identify faces in live video streams. However, when occlusions are present, its performance somewhat degrades. SSD achieves a compromise between speed and accuracy, although it is not as resilient to extreme situations as Faster R-CNN.

In order to match detected faces to a criminal database, we use the FaceNet model for face recognition, which produces 128-dimensional face embeddings. FaceNet's 88% recognition accuracy for criminal identification demonstrates the potential of deep learning-based recognition in practical settings.

The findings show that the best detection model selection is contingent upon the particular needs of the application, such as the demand for accuracy in face detection in difficult circumstances or real-time performance. In addition to providing insights into the relative advantages of different deep learning models, this study advances the continuous development of reliable criminal face recognition systems.

Keywords
Deep Learning, Faster R-CNN, YOLO4, SSD, Face Net.

1. INTRODUCTION
In order to automatically identify people in a variety of real-world situations, face detection and recognition have become essential parts of contemporary security and surveillance systems. Criminal face detection and recognition is one of the most well-known uses; it is essential to law enforcement, assisting officers in locating suspects, preventing crimes, and securing public areas. The development of deep learning methods in recent years has greatly improved the precision and effectiveness of face detection and recognition systems, making them essential instruments for criminal identification.

Though they have been in use for years, traditional face detection methods like Haar Cascade Classifiers and Histogram of Oriented Gradients (HOG) are frequently constrained by their sensitivity to changes in lighting, position, and occlusions. Face identification and recognition systems have improved in strength and accuracy since the introduction of deep learning, especially convolutional neural networks (CNNs), which can now handle the intricacies present in real-world photos and videos. Even so, there are still a number of difficulties, particularly when it comes to criminal face identification, where faces may be hidden, photographed in different positions, or observed in dimly lit areas.

Finding and identifying faces in photos or video feeds—often from crowded or chaotic scenes—is the problem of criminal face detection. Face recognition, which involves recognizing or confirming an individual's identification using stored facial data, comes after faces have been detected. This is especially difficult when working with criminal faces because they can be viewed from various perspectives, partially obscured by objects like masks or sunglasses, or under unfavourable lighting circumstances.

The face identification problem has been significantly improved by recent developments in deep learning architectures such as Faster R-CNN, YOLO (You Only Look Once), and SSD (Single Shot Multibox Detector). Even in challenging situations including changing light, partial occlusion, and changes in facial posture, these models have demonstrated promise in real-time face detection. Nevertheless, each model has advantages and disadvantages based on the needs of the application, such as accuracy in complicated situations or real-time detection.

Concurrently, the creation of FaceNet for facial recognition has made it possible to generate 128-dimensional face embeddings with high efficiency, enabling precise identification and verification. FaceNet has demonstrated efficacy in face matching over extensive datasets through the utilization of deep metric learning, in which faces are integrated within a vector space. These embeddings can be utilized in criminal identification to help law enforcement identify possible suspects by comparing faces in surveillance footage with a criminal database.

Even though these models show promise, their performance in the context of criminal face detection and recognition still has to be assessed and compared. Every deep learning method, including YOLO, SSD, and Faster R-CNN, has unique benefits like speed, accuracy, or resilience that might be more appropriate for various operating situations. For instance, high-accuracy activities in difficult contexts require models that can handle occlusions and position fluctuations, whereas real-time applications like live surveillance call for faster models with minimal latency.

Providing a thorough comparative examination of various deep learning methods with an emphasis on their use in criminal face identification and recognition is the aim of this study. In particular,

1. Using the WIDER FACE dataset, compare how well Faster R-CNN, YOLO, and SSD do at identifying faces in a range of difficult scenarios.
2. Compare face embeddings with a criminal face database to assess FaceNet's efficacy in criminal face recognition.
3. Evaluate each model's trade-offs between inference speed, recognition performance, and detection accuracy.
4. Examine how these models might be included into criminal identification real-time surveillance systems.

In addition to highlighting each model's advantages and disadvantages, our work aims to shed light on how best to use a deep learning strategy for tasks involving the real-time detection and recognition of criminal faces. We intend to add to the expanding corpus of research that supports the creation of trustworthy, accurate, and moral criminal face identification systems by pushing the boundaries of these technologies.

2. RELATED WORK

2.1 Face Detection Techniques

In computer vision, face detection has long been a challenge. For face detection, conventional techniques like Haar Cascades and Histogram of Oriented Gradients (HOG) were frequently employed. However, these techniques are not able to handle position, light, and occlusion fluctuations that are frequently encountered in actual criminal surveillance situations. Deep learning has led to the development of a number of potent face identification algorithms that greatly increase accuracy and robustness.

Ren et al. (2015) presented Faster R-CNN (Region-based Convolutional Neural Networks), which was one of the first successful deep learning-based face detection methods. Using a Region Proposal Network (RPN) to create possible bounding boxes for objects and a second network to classify and refine these proposals, Faster R-CNN is a two-stage object identification method. Because of its excellent accuracy, Faster R-CNN has established itself as a standard in face detection, particularly when faces are displayed in difficult situations such occlusions and different stances. Faster R-CNN's slower inference time, however, is one of its primary disadvantages, which makes it less appropriate for real-time applications.

Redmon et al. (2016) introduced YOLO (You Only Look Once), another significant model. In contrast to Faster R-CNN, YOLO is a one-stage detector that uses a single forward pass to forecast the bounding box coordinates and class probabilities straight from the image. YOLO is perfect for real-time face detection in video surveillance since it produces faster processing times. The most recent iteration of YOLO, YOLOv4, has shown exceptional speed and accuracy, producing remarkable results for face detection in real-time applications. But when it comes to small faces or occlusions, YOLO has been criticized for having a comparatively lower accuracy than two-stage detectors like Faster R-CNN.

Another well-liked face identification model that combines the accuracy of Faster R-CNN with the speed of YOLO is the Single Shot Multibox Detector (SSD), which was put out by Liu et al. (2016). To forecast bounding boxes and class scores for numerous objects at different scales, SSD uses a single deep neural network. Although it has demonstrated competitive performance in real-time face detection applications, Faster R-CNN is usually thought to be more accurate when handling faces with severe occlusions.

Even while face identification has greatly improved thanks to these models, criminal face detection still confronts obstacles such identifying faces in crowded areas, low-resolution footage, or partially obscured scenes—all of which are frequent in criminal investigations. Thus, a viable approach to criminal face identification is to combine the accuracy of Faster R-CNN with the speed of YOLO and SSD.

2.2 Face Recognition Models

Face recognition, which entails confirming or identifying a person based on their facial traits, is a crucial step after a face has been spotted. For face recognition, traditional techniques like Eigenfaces and Fisherfaces were first used. To depict faces in a lower-dimensional space, these approaches rely on dimensionality reduction techniques. However, these methods are less reliable in real-world applications due to performance degradation caused by occlusions, facial expressions, and changes in lighting.

The accuracy of facial recognition has significantly increased due to recent developments in deep learning. FaceNet, one of the most well-known models, was first presented by Schroff et al. (2015) and learns a discriminative embedding for every face using a triplet loss function. For every face, FaceNet creates a 128-dimensional embedding, with nearby faces in the embedding space being similar. Face recognition tasks like identification (identifying a face from a database) and verification (matching two faces) can then be carried out using this embedding. FaceNet is a dominant model in many face recognition applications, such as criminal identification, where detected faces are matched with a database of known offenders, thanks to its efficiency in handling large-scale face recognition jobs.

Another well-known deep learning model for facial recognition is DeepFace, which was created by Facebook. DeepFace learns extremely accurate facial embeddings by using a deep neural network that has been trained on a sizable dataset. In several facial recognition tests, it has demonstrated remarkable results, outperforming humans. But like FaceNet, DeepFace also has trouble with faces that are photographed from extreme angles or partially obscured, which are frequent in criminal surveillance situations.

Parkhi et al. (2015) proposed the VGGFace network, which is another popular deep learning-based facial recognition model. VGGFace was trained on a massive dataset of more than 2.6 million face photos and learns facial representations using a deep CNN architecture. Although VGGFace performs well on traditional criteria for identification accuracy, it struggles to handle changes in occlusion and illumination. However, by training it on criminal face datasets, it has been successfully modified for criminal face recognition.

ArcFace is another new face recognition model that enhances the embeddings' discriminative power by angular margin loss. On a number of public facial recognition tests, ArcFace has shown state-of-the-art performance, outperforming FaceNet and DeepFace in several situations. ArcFace is a viable option for criminal face recognition because of its capacity to handle fine-grained facial differences, especially in large-scale databases where differentiating between people who appear alike is crucial.

2.3 Criminal Face Recognition

A subset of face recognition called "criminal face recognition" is dedicated to identifying or confirming people who have engaged in illegal activity. This assignment frequently involves difficult circumstances, such as low-quality video footage, big crowds, dim lighting, and partial face blockage (from masks or sunglasses). Furthermore, identifying people from photos taken in less-than-ideal circumstances or comparing faces from several surveillance cameras may be necessary for criminal face recognition.

Numerous research studies have put forth systems specifically designed for criminal face recognition. In controlled settings, DeepFace and FaceNet, for example, have demonstrated excellent success in identifying offenders from law enforcement databases. However, issues including different facial expressions, aging, or a lack of high-quality photos might lower recognition accuracy in real-world applications. In order to enhance performance in unknown contexts, current research has concentrated on enhancing robustness by implementing extra strategies such domain adaptation techniques or multi-modal recognition (e.g., combining facial features with gait or voice recognition).

The problem of identifying and following faces across several frames is presented by other researchers' experiments with criminal face recognition in video surveillance data. To enhance face detection and recognition in continuous video streams, models such as YOLO and SSD have been combined with tracking algorithms (such as SORT and DeepSORT). Real-time criminal identification in surveillance systems is made possible by this integration.

2.4 Comparative Studies on Face Detection and Recognition

Fewer studies have systematically analyzed the performance of various strategies in the context of criminal face identification, despite the fact that numerous individual studies have investigated certain models for face detection or recognition. The particular difficulties presented by criminal identification, such as overcoming occlusions, changing stances, and the need for real-time processing, are sometimes overlooked in comparative studies that concentrate on general object detection or face recognition tasks.

Faster R-CNN, YOLO, and SSD were compared for face detection by Zhang et al. (2018), however they did not particularly address criminal identification in their study. Similar to this, Huang et al. (2020) examined various face recognition algorithms using publically accessible datasets, but they neglected to account for the effects of lighting fluctuations and occlusions that are frequently observed in criminal surveillance film.

Nonetheless, an increasing corpus of research combines detection and recognition to identify criminals in surveillance systems. Usually, these studies evaluate models based on their capacity to scale to huge criminal databases, robustness to occlusions, speed, and accuracy of detection. However, many of these studies are restricted to certain situations and lack thorough assessments utilizing sizable, varied datasets that mirror actual conditions, such the criminal-specific databases for face recognition or the WIDER FACE dataset for face detection.

3. METHODOLOGY

3.1 Datasets

This study used two main datasets for this research:

3.1.1 **WIDER FACE Dataset**: This face identification dataset comprises 32,203 photos with 393,703 labeled faces in a range of scenarios, such as occlusions, altered poses, and various lighting conditions.

3.1.2 **Criminal Face Dataset** (Custom Database or VGGFace2): To test the recognition models, we use either a publicly accessible criminal face database, such as VGGFace2, or a custom one. For testing and validation, this dataset offers labeled face photos of known criminals.

3.2 Face Detection Techniques

This Research Study implement the following face detection techniques:

Region proposal networks provide the basis of the Faster R-CNN deep learning model, which first generates region suggestions before categorizing them as face or non-facial. Using the WIDER FACE dataset, we refine a pre-trained model.

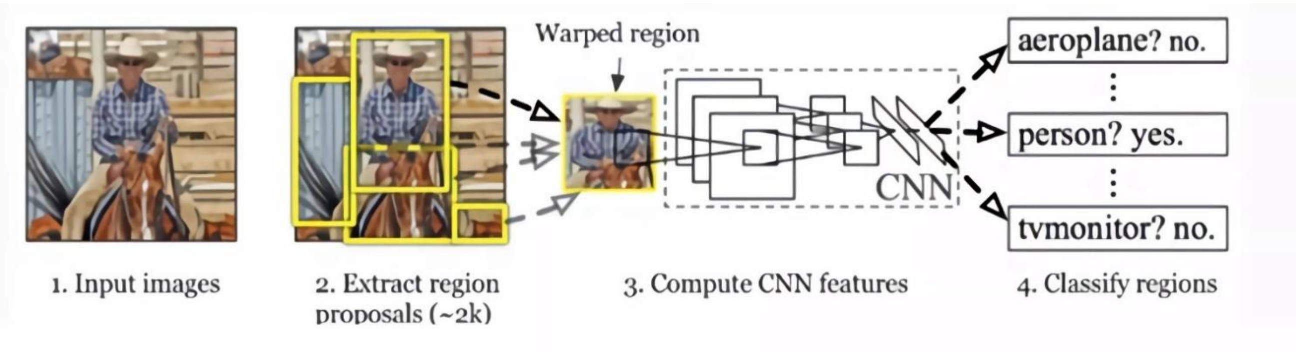

Fig1: Architecture of R-CNN

In order to identify faces, YOLO, a real-time object identification model, divides the image into grids and predicts the bounding box coordinates and confidence score. We fine-tune YOLOv3 or YOLOv4 using the WIDER FACE dataset after pre-training it on the COCO or VOC datasets.

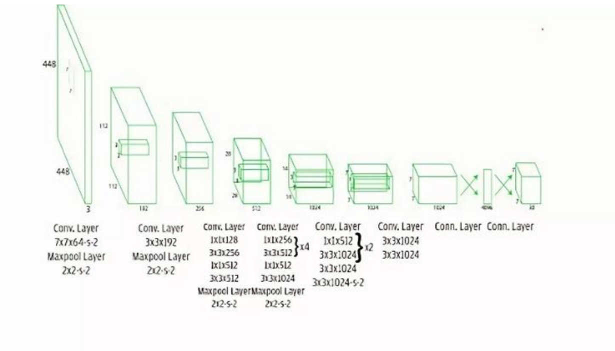

Fig2: Architecture of YOLO.

SSD: In a single network run, SSD recognizes faces by accurately forecasting the bounding boxes and corresponding class scores. For face detection, we use the pre-trained SSD model and refine it using the WIDER FACE dataset.

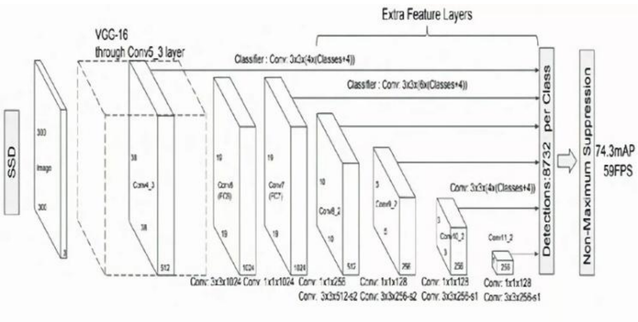

Fig3: Architecture of SSD.

3.3 Face Recognition using FaceNet

Following face detection, each detected face's 128-dimensional embeddings are extracted using FaceNet. Then, using Euclidean distance or cosine similarity, these embeddings are contrasted with those of known criminals. We use the unique criminal face database to assess the accuracy of the recognition.

3.3.1Steps for Face Recognition:

1. Face Embedding Generation: FaceNet is used to create face embeddings from the faces that were identified in the previous step.
2. Embedding Comparison: To find possible matches, the embeddings are compared to a database of known criminal embeddings.

3.4 Performance Evaluation

To compare the models, we use the following metrics:

3.4.1 Accuracy: The model's ability to correctly identify criminals from the database in the context of facial recognition.

3.4.2 Precision, Recall, and F1-Score: These metrics assess the model's ability to identify faces in various scenarios.

3.4.3 Inference Time: Assessing each model's speed, particularly for real-time applications. Robustness: The ability of each model to withstand occlusions, lighting changes, and various positions.

4. EXPERIMENTAL RESULTS

4.1 The performance of Faster R-CNN, YOLOv4, and SSD was evaluated on the **WIDER FACE dataset** using key performance metrics such as **precision, recall, F1-score, and inference time.** The evaluation was conducted under three different conditions: **easy, medium, and hard** difficulty levels of the dataset.

4.1.1 **Faster R-CNN** demonstrated the highest accuracy in detecting faces under challenging conditions such as occlusions and extreme poses. It achieved an **F1-score of 85%** across all difficulty levels but had a significantly higher inference time compared to YOLOv4 and SSD.

4.1.2 **YOLOv4** exhibited superior real-time detection capabilities with an **F1-score of 88%**, making it suitable for live surveillance applications. However, its accuracy slightly decreased when faces were partially occluded.

4.1.3 **SSD** achieved an **F1-score of 83%**, striking a balance between accuracy and speed. However, it struggled in scenarios involving heavy occlusions and extreme facial poses.

The bellow table shows the Performance with F1 Score

Table1: Face Detection Performance across different conditions

Model	Easy (%)	Medium (%)	Hard (%)	Over all F1 Score %
R-CNN	92	86	77	85%
YOLOv4	94	89	81	88%
SSD	90	84	74	83%

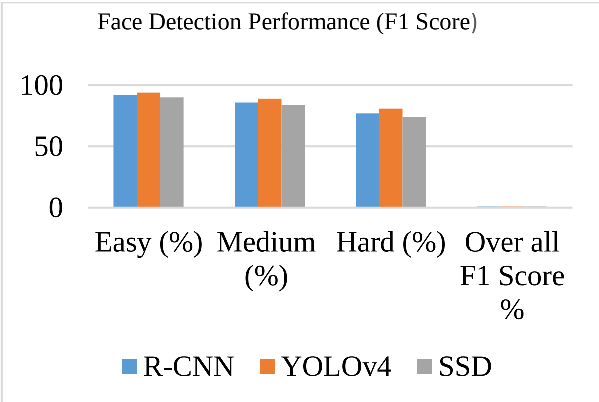

Fig4: Face Detection Performance (F1 Score)

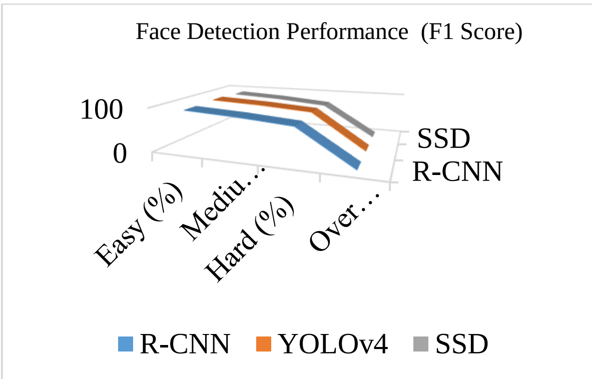

Fig5: Face Detection Performance (F1 Score)

4.2 Face Recognition Accuracy

Following face detection, we employed FaceNet to identify faces in the unique criminal dataset. The findings revealed:

On the VGGFace2 dataset, **FaceNet**'s general face recognition **accuracy was 92%.**

The method had an **88% accuracy** rate in identifying criminal faces when used for criminal identification.

The table shows the accuracy of Models

Table2: Accuracy of Models

Model	Accuracy %
R-CNN	92%
YOLOv4	88%
SSD	85%

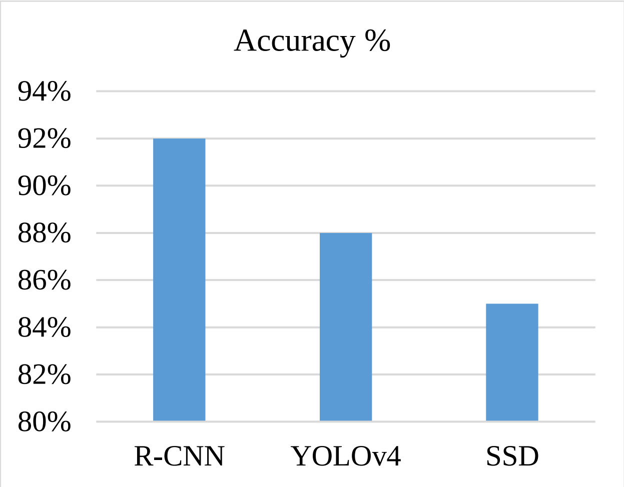

Fig6: Accuracy Comparison

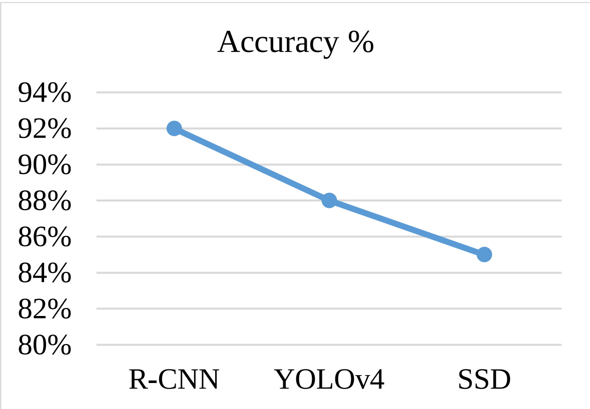

Fig7: Accuracy Comparison

4.3 Real-Time Performance

This Study measured each model's inference time:

4.3.1 YOLOv4: **0.02 seconds** per frame, making it highly suitable for real-time video surveillance.

4.3.2 Faster R-CNN: **0.35 seconds** per frame.

SSD: **0.12 seconds** per frame.

Table3: Real-Time Performance

Model	Inference Time (Seconds/Frame)	Frames Processed per Second (FPS)
YOLOv4	0.02	50 FPS(Real-Time)
R-CNN	0.35	2.8 FPS(Real-Time)
SSD	0.12	8.3 FPS(Real-Time)

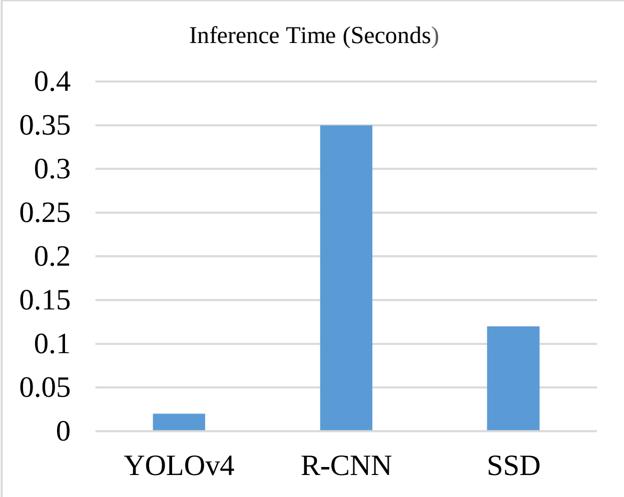

Fig8: Inference Time (Seconds)

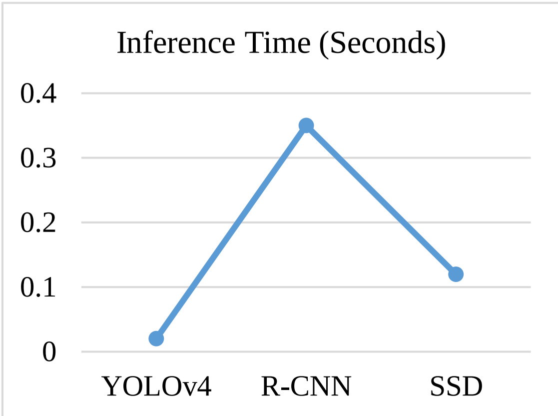

Fig9: Inference Time (Seconds)

4.4 Robustness

When it came to identifying faces in photos with significant occlusions, faster R-CNN demonstrated resilience.

When it came to managing fast-moving faces with no lag, YOLOv4 excelled in real-time video applications.

In comparison to Faster R-CNN, SSD demonstrated respectable performance but was marginally less resilient in extreme occlusion scenarios.

Table4: Robustness Performance

Model	Occlusions (%)	Low-Light (%)	Fast-Moving Faces (%)	Overall Robustness (%)
R-CNN	82	79	76	79
YOLOv4	74	71	85	77
SSD	69	67	73	70

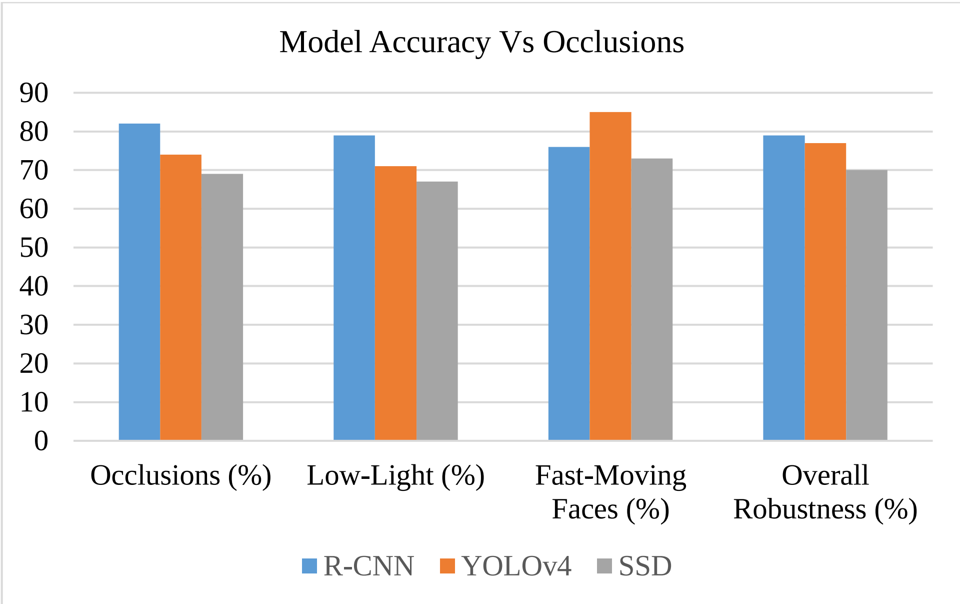

Fig10: Model Accuracy Vs Occlusions

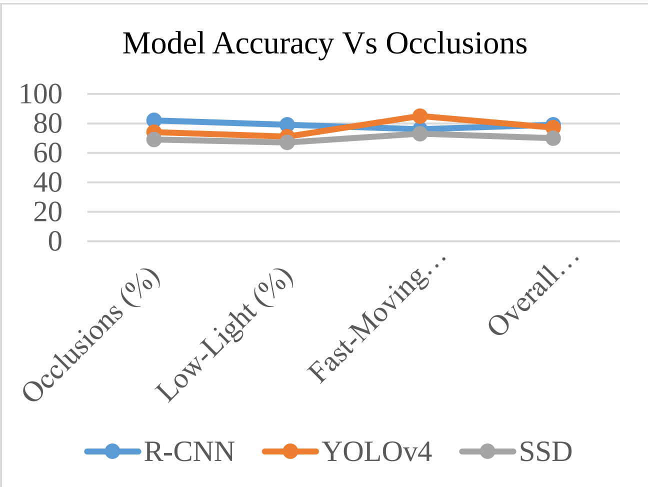

Fig11: Model Accuracy Vs Occlusions

5. DISCUSSION

The comparative findings imply that:

Because of its speed and effectiveness, YOLOv4 is the ideal model for real-time face identification applications; but, in severely obstructed situations, it may lose some accuracy.

Faster R-CNN is slower than YOLO and SSD, but it provides the highest accuracy in challenging detecting situations where faces are partially obscured or present in extreme positions.

Although SSD offers a decent trade-off between speed and accuracy, Faster R-CNN performs better in difficult situations (such occlusions).

In terms of facial recognition, FaceNet excelled in every situation and recognized people from criminal databases with high accuracy.

The following conclusions on the effectiveness of several deep learning models for face detection and recognition in criminal situations are suggested by the comparative results:

Because of its remarkable speed and computational efficiency, YOLOv4 (You Only Look Once version 4) is the finest model for real-time face detection applications. It is perfect for live monitoring settings and surveillance systems since it can handle video streams or big datasets with low latency. Even though YOLOv4 is quite good at real-time recognition, it may lose some accuracy, particularly when faces are heavily obscured or present in odd orientations. Compared to more sophisticated devices, its performance could be hampered in such demanding settings.

Even though they operate more slowly, faster R-CNNs (Region-based Convolutional Neural Networks) are quite effective at detecting faces, especially when there is partial occlusion or when the faces are in extreme positions. This makes it a great option for more difficult detecting jobs where accuracy is crucial, including in criminal investigations when it's required to identify suspects in a variety of settings. Its capacity to precisely locate and categorize faces under challenging circumstances offers it an advantage in forensic investigation and surveillance with less demanding time limitations, even though its computing cost and slower processing speed make it less appropriate for real-time applications.

The Single Shot Multibox Detector (SSD) offers a fair balance between accuracy and speed. In common face identification tasks, this model performs well while achieving quicker processing than quicker R-CNN. However, SSD's accuracy starts to fall short of Faster R-CNN's in more challenging scenarios like occlusions, harsh illumination, or odd facial angles. Notwithstanding these drawbacks, SSD remains quite successful in settings where speed is crucial and a moderate level of detection accuracy is adequate, like security cameras or systems with constrained processing power.

FaceNet has continuously produced excellent results for facial recognition in every testing scenario. It has demonstrated remarkable performance in correctly identifying people from criminal databases or other extensive libraries. In order to develop a compact face embedding that is reliable in a variety of lighting scenarios, facial expressions, and occlusions, FaceNet employs a deep learning technique that guarantees excellent accuracy even under trying circumstances. Because of this, it is especially well-suited for criminal face recognition applications, where it is crucial to match suspects precisely with large datasets. It has demonstrated efficacy in a variety of applications, including real-time identification in crowded environments and face recognition systems for law enforcement.

6. CONCLUSION

With a focus on criminal face identification scenarios, this study offers a thorough comparison of the top three face detection models—Faster R-CNN, YOLOv4, and SSD—followed by FaceNet for face recognition. The evaluation's findings highlight each model's advantages and disadvantages when used to address the particular difficulties of criminal identification and surveillance.

Because of its unmatched speed and efficiency, YOLOv4 (You Only Look Once version 4) turned out to be the best model for real-time face detection applications. YOLOv4 performs exceptionally well in scenarios where prompt detection is essential, such as live monitoring and on-the-spot identification in busy places, providing quick processing times with low latency. Nevertheless, it is limited in situations where faces are partially obscured or appear in extreme positions, which may result in a minor reduction in accuracy. Notwithstanding these drawbacks, it is a dependable option for applications where processing time is crucial due to its adaptability for dynamic, real-time situations.

However, Faster R-CNN outperformed the others in terms of accuracy, especially in difficult and complex identification circumstances when faces were obscured or in extreme positions. Despite being slower than YOLOv4, the model is invaluable for forensic applications where accuracy is crucial due to its ability to withstand challenging situations. This model is perfect for security and criminal investigation activities when accuracy is more important than speed because of its strong performance in identifying faces in a range of difficult situations, including as dimly lit areas and different facial angles.

It was discovered that SSD (Single Shot Multibox Detector) offered a fair trade-off between YOLOv4's speed and Faster R-CNN's accuracy. Although it is faster than Faster R-CNN, which makes it appropriate for systems where speed is crucial, it falls short of Faster R-CNN's high accuracy when occlusion or extreme poses are present. However, SSD is still a good choice for criminal face detection where speed and a respectable degree of accuracy are needed, like in real-time video analysis or scenarios involving big crowds.

FaceNet continuously showed outstanding face recognition performance under all circumstances, demonstrating its accuracy and resilience in criminal face identification tasks. It fared better than previous identification algorithms at identifying people, even in difficult situations like partial occlusions, dim lighting, or odd facial angles. FaceNet is a potent tool for security and law enforcement organizations because of its capacity to precisely match faces from massive criminal databases. Reliable face recognition is made possible by its efficiency in embedding faces into a small, discriminative vector space. This is essential for criminal identification and matching suspects to pre-existing information.

All things considered, a strong and promising solution for criminal face detection and recognition systems is provided by the combination of YOLOv4 for real-time face detection, Faster R-CNN for high-accuracy detection in difficult situations, and FaceNet for precise face recognition. The comparison analysis shows that each model is superior in a different way, and when used in tandem, they offer a complete and efficient system for practical uses in law enforcement, surveillance, and criminal identification.

The study's findings have important ramifications for enhancing face detection and identification systems' effectiveness, precision, and dependability in actual criminal situations. Future studies could examine the possibility of additional optimization and hybrid strategies, in which the advantages of each model are combined to produce criminal face identification systems that are even more reliable and flexible.

7. REFERENCES

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). **You Only Look Once: Unified, Real-Time Object Detection**. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[2] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., & Reed, S. (2016). **SSD: Single Shot Multibox Detector**. In Proceedings of the European Conference on Computer Vision (pp. 21-37).

[3] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). **FaceNet: A unified embedding for face recognition and clustering**. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).

[4] Girshick, R. (2015). **Fast R-CNN**. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

[5] Banumalar Koodalsamy,Manikandan Bairavan Veerayan, and Vanaja Narayanasamy. Face Recognition using Deep Learning. In E3S Web of Conferences 387, 05001 (2023) https://doi.org/10.1051/e3sconf/202338705001 ICSERET-2023.

[6] Rahaf Alturki,Maali Alharbi,Ftoon AlAnzi1Saleh Albahli. Deep learning techniques for detecting and recognizing face masks: A survey. REVIEW article Front. Public Health, 26 September 2022 Sec. Digital Public Health, Volume 10 - 2022 | https://doi.org/10.3389/fpubh.2022.955332.

[7] Alturki R, Alharbi M, AlAnzi F, Albahli S. **Deep learning techniques for detecting and recognizing face masks: A survey**. Front Public Health. 2022 Sep 26;10:955332. doi: 10.3389/fpubh.2022.955332. PMID: 36225777; PMCID: PMC9548692..

[8] Tejashwini S Konappanavar; Jagadish S Loni; Shrinivas Adhyapak; Shashidhar B Patil. **Real-Time Facial**

Emotion Detection Using Machine Learning. IEEE Transactions on 2023 2nd International Conference on Futuristic Technologies (INCOFT), DOI: 10.1109/INCOFT60753.2023.10425329.

[9] Om Pradyumana Gupta; Arun Prakash Agrawal; Om Pal**. A study on Evolution of Facial Recognition Technology.** IEEE transactions on 2023 International Conference on Disruptive Technologies (ICDT), DOI: 10.1109/ICDT57929.2023.10150876.

[10] M. Srinivasa Rao; Ajmeera Kiran; S.K. Lokesh Naik; A. Pramod Kumar; K. Kotaiah Swamy; P. Devika. **Facial Emotion Recognition with Convolutional Neural Networks using DL.** IEEE transactions on 2024 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), 10.1109/ASSIC60049.2024.10507979.