

Enhanced Real-Time Face Detection in Challenging Environments: Leveraging YOLOv8 for Partially Occluded and Low-Light Conditions

Amina Shaikh^a, Imran Khan Keerio^{b*}, Mutiullah Shaikh^c, Asadullah Shah^a

^aKICT, International Islamic University, Malaysia

^bDepartment of Computer Science, Sindh Madressatul Islam University, Karachi, Pakistan

^cThe Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

* Corresponding author: Dr Imran Khan Keerio, Email: imran_keerio@yahoo.com

Received: 19 September 2025, Revised: 08 November 2025, Accepted: 24 December 2025, Published: 01 January 2026

KEY WORDS

YOLOv8
Face Detection
Partially Occluded Faces
Low-light Conditions
Deep Learning
Surveillance
Real-time Object Detection

ABSTRACT

Face recognition technology is considered a crucial component of modern surveillance frameworks; however, face recognition under unfavourable conditions, including partial obstruction and low-light scenarios, is still a prominent challenge. This paper analyzes how the deep learning model YOLOv8 can be used to solve these issues. The model was tested with a customized set of photos illustrating faces with different levels of occlusion (such as masks, helmets and scarves) and lighting. The researcher has classified the data into three categories 80% of the data is used in training, 10% of the data is used in testing, and 10% of the data is used in validation. The training process was taken 150 times, with critical measures such as 93% precision, 88.7% recall, an F1-score of 0.90, and a mean average precision (mAP@50) of 0.92. The model performed well in ordinary surveillance circumstances, but its detection efficacy decreased in cases of severe occlusions and low-light surroundings, as seen by a lower mAP@50-95 score of 0.512. YOLOv8 is compared with YOLOv7, RetinaFace, and Faster R-CNN, based on performance metrics. Given these challenges, YOLOv8 provides a solution for real-time face identification in security applications. Future research will focus on enhancing the model's resilience through sophisticated pre-processing approaches, multimodal detection, and transfer learning to improve performance in increasingly complex real-world scenarios.

1. Introduction

Currently, law enforcement agencies employ sophisticated surveillance systems, including facial recognition technology, for forensic and security purposes. Research in surveillance systems, including facial recognition, has made significant advancements; nonetheless, many gaps remain to be explored [1]. Photographs of faces captured in inadequate lighting, along with images of helmets and face masks, provide significant challenges for the existing system to differentiate. Current facial recognition algorithms, designed to differentiate images based on the eyes, nose, and mouth, may be inadequate in the previously listed situations.

In order to overcome these challenges, scientists are employing state-of-the-art computer science technology, including deep learning, which is among the most effective to solve the above challenges. Deep learning is an area of artificial intelligence that utilizes neural networks to recognize faces in photographs and videos. For this problem, various deep learning models are available, including one. YOLO, which is often used for facial recognition systems [24]. Moving on, there are other variants of the YOLO model, one of which, YOLO8, instantly recognizes an image in a single frame of video or image. This model delivers rapid and accurate face recognition, even with a large number of faces.

Deep learning, a subfield of AI, has significantly improved the accuracy of facial recognition using other models, such as Convolutional Neural Networks (CNNs). However, as previously discussed, the current version of YOLOv8 is utilized for real-time face recognition, making it more efficient in challenging scenarios such as low-light conditions and faces with helmets or masks. Although prior face detectors handle either occlusion or poor illumination, most do not evaluate both conditions simultaneously under real-time constraints, and they seldom report high-IoU performance (e.g., mAP@0.75–0.95) or latency on practical hardware. Methods like occlusion-aware detectors and strong baselines (e.g., RetinaFace/Faster R-CNN/YOLOv7) mainly report mAP@0.50 on well-lit benchmarks and have limited analysis for helmet/mask occlusions in low-light scenes. Public datasets also underrepresent “helmet/mask + low light” cases common in surveillance. Thus, the community lacks a systematic, high-IoU, real-time evaluation for jointly occluded and dimly lit faces. Identifying faces of persons with partially covered faces in poor lighting conditions is a challenging task. This study utilized the YOLOv8 model to evaluate its standard performance in the aforementioned demanding setting.

2. Related Works

2.1 Face Identification Using Deep Learning

The deep learning mask face recognition system classifies faces using a variety of deep learning algorithms. Deep learning, particularly convolutional neural networks, excels in image recognition, making it particularly useful for detecting face masks. In daily life, facial recognition is difficult due to blocked perspectives and shifting illumination. Occlusions can be caused by background components, such as shadows and reflections, as well as scene items like shields, helmets, and hands. Nighttime or dimly lit surveillance footage demonstrates that low light conditions make it difficult to see [7].

Traditional face recognition methods like Haar cascades and Support Vector Machines (SVMs) pose a lot of challenges. Moreover, they depend on the observable facial expression, and hence they are not effective in cases where faces are covered or in low-light situations [8].

2.2 YOLO for Real-Time Object Detection

YOLO, or “You Only Look Once,” is a fast and accurate method for detecting objects in real-time, utilizing deep learning. The YOLO approach simplifies object detection by eliminating the need for multiple forward passes through the network, instead

performing a single pass. It is ideal for real-time use, e.g., security and surveillance [24].

This study is closely connected to earlier research. A variety of experiments have been carried out to enhance the You Only Look Once framework for facial detection over several iterations. [10] Created a new object recognition method named YOLOv5 to enhance the accuracy of image segmentation. The results of the experiment indicate that the suggested algorithm structure achieves an accuracy of 98.5% in semantic segmentation and 95.4% in item recognition. The authors of [11] came up with a mobile platform that is aimed at identifying and sorting products in the aquaculture sector.

The suggested method employs the YOLOv4 detecting algorithm with a unique labelling system. Video analysis of the proposed method was successful, and the items that moved on the conveyor belt were correctly recognized, both randomly and in a given order and at a speed of 505.08 m/h. The success rate was 98.15%. One author has created a computerized system that can automatically be sorted and recognize various species of object. The webcam images have been created to create a new object detection network (FD_Net) that can discriminate among nine individual object species. This system is an improvement over the YOLOv7 algorithm in that 3x3 filters in depth-separable convolutions are used instead of Darknet53 in the bottleneck attention module of the extended feature extraction network. Exactly, 95.30% of the time.

The term YOLO has undergone a lot of development over time. Its most recent release, YOLOv8, has improved feature extraction, improved the network architecture, and increased performance. The progress of YOLOv8 is more evident in the detection of objects in tricky environments, including those concealed or those in low-light environments [13]. Past studies suggest that the predecessors of YOLO, including YOLOv5 and YOLOv7, were good at detecting objects, but not small ones or those partially obscured. YOLOv8 can solve these problems, thus it can be used as a dependable solution to face detection in real-time video systems when there are difficulties [14].

YOLOv8 is a more developed model which improves the functions of the previous versions of the YOLO as it adds new functions and makes the necessary changes to accelerate and increase flexibility. Being a single-stage object recognition system, YOLOv8 inherits the advantages of the earlier

versions of the systems, including YOLOv1 up to YOLOv7. In this scenario, YOLOv8 requires only a single feature extraction step to identify objects, which makes it faster than other methods that involve two stages [15, 16].

YOLOv8 introduces several updates, including the replacement of the C3 module with the C2f module as the backbone, and a change in the number of blocks per stage from [3, 6, 9, 3] to [3, 6, 3]. Additionally, the anchor base has been shifted to an anchor-free approach, inspired by TOOD and ppyoloe's YOLOv6. YOLOv8 offers excellent flexibility and is compatible with all earlier versions of the YOLO framework, enabling users to switch between versions with ease to evaluate their performance. If researchers are looking to explore the latest YOLO technology while retaining the option to use older models, YOLOv8 is a viable choice. Figure 1 illustrates the assembly of YOLOv8.

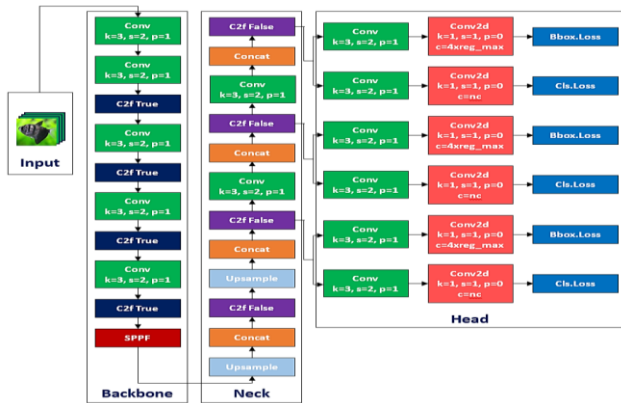


Fig. 1. Architecture of YOLOv8 [17]

Despite strong progress in occlusion-aware and low-light face detection, prior studies typically treat these stressors in isolation, benchmark mainly at AP@0.50, and seldom report latency/FPS on accessible hardware; moreover, public benchmarks underrepresent helmet/mask occlusions in dim scenes, limiting external validity for surveillance use [21]. Consequently, the field lacks a systematic, high IoU (mAP@0.50–0.95), real-time evaluation when occlusion and poor illumination co-occur. To address this gap, we investigate whether a single-stage detector (YOLOv8) can sustain precision–recall at higher IoUs while maintaining real-time throughput under joint stress [9]. A test protocol is curated that stratifies by occlusion ratio and luminance band, train with targeted illumination (e.g., gamma, noise) and occlusion augmentations, and report mAP@0.50–0.95, stratified PR curves, and end-to-end latency/FPS on commodity hardware, alongside ablations of the robustness components. This design turns a descriptive challenge into a concrete, measurable problem and establishes a practical baseline for occluded, low-light face detection.

3. Methods

Figure 2 illustrates the five stages of the occluded face detection methodology: data collection, data preprocessing, YOLOv8 model training, YOLOv8 model assessment, and YOLOv8 model testing. Data collection involves gathering photographs of partially obscured faces from the internet. Preprocessing involves standardizing image dimensions, annotating images with labels and bounding boxes, and performing data augmentation. Training on the YOLOv8 model entails the use of training data and the YOLOv8 model. The validation of the YOLOv8 model is carried out by evaluating the results of its training and checking its potential.



Fig. 2. Research design flow

3.1 Dataset Creation and Augmentation

The researcher developed a personalized dataset to replicate surveillance where individuals had half-covered faces in dark environments. The collection involved about 1,100 human face photos whereby some of them were wearing helmets and others face mask in low-lights. Researcher has been using images of various online datasets that are publicly available. In order to improve the dataset, augmentation techniques were applied to improve the model performance, some of them are as follows.

- Brightness modification: Emulated dim lighting and excessive exposure scenarios.
- Horizontal and vertical flips: Enhanced dataset variability.
- Noise injection: Emulated subpar camera performance in low-light conditions
- Random rotations and scaling: Guaranteed resilience to variations in facial orientation.

Moreover, the model was made resilient through synthetic occlusion and simulation techniques of illumination. Synthetic occlusions were created through the programmatically overlaying of masks, sunglasses and random geometric patterns on the face areas. Gamma correction and luminance factors were used to simulate low-light conditions in order to emulate night time surveillance footage. Noise injection was also introduced using Gaussian and salt-and-pepper noise to simulate the distortions of camera sensors. These artificial augmentations enhanced the capability of YOLOv8 to be generalized to unpredictable real-life conditions.

3.2 Model Training and Hyperparameter Tuning

In order to manage the computational burden, the researcher used Google Labs and trained the YOLOv8

model accordingly. The researcher has separated the dataset into the following: 80% training, 10% testing, and 10% validation. The model was trained on more than 150 iterations with the following hyperparameters:

- Batch size: 16
- Initial learning rate: 0.001
- Image size: 640x640 pixels (standard for YOLO models)

In the training process, the learning rate was adjusted by the researcher to prevent overshooting and easy convergence. The accuracy of the model was also evaluated by calculating F1 score and a mean Average Precision (mAP) at different levels of IoU.

3.2 YOLOv8 Model Evaluation

This section focuses on calculating the model's performance after training, considering various factors such as matrices and visualization.

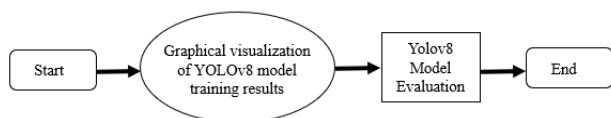


Fig. 3. Model evaluation chart for YOLOv8

The process of training models is depicted in figure 3. After the researcher trained the model, the training results of the model showed the accuracy plots and curves. Precision, recall, Mean Average Precision (mAP), Intersection over Union (IoU), and F1 score to quantify accuracy and performance were further analyzed to identify the overfitting and underfitting parameters [19]. The independent test dataset was used to analyze to guarantee that the ability of the model to adapt to new data and its performance are compared to the baseline models or older versions of YOLO. The researcher also applied the model to real-life scenarios to evaluate its performance by performing the required model adjustments, including the adjustment of hyperparameters or architectural modifications, to ensure the YOLOv8 model can be released.

3.4 YOLOv8 Model Selection

YOLOv8 was selected to identify real-time objects. It proves to be an expert at testing items with great Precision and offers rapid service. This model can be used in applications that demand rapid detection like surveillance systems [20]. YOLOv8 proved several advantages over its past version, some of which are listed below.

- Enhanced feature extraction: Improved management of diminutive items and obstructions [21].
- It Provides Faster processing for real-time applications [22].
- Improved architecture: It utilizes new technologies, such as the Bidirectional Feature Pyramid Network (BiFPN) and loss functions like CIOU, to enhance localization [23].

4. Experimental Results and Discussions

This section of the research paper describes the findings produced by the YOLOv8 model under partially obscured and low-light conditions. These findings incorporated various performance metrics, model forecasts, loss curves, confusion matrix analysis, and plots of attribute distributions, and are additionally addressed in the following sections. These results provide the strengths and weaknesses of the model.

4.1 YOLOv8 Model Evaluation Metrics

To determine typical performance measures such as Precision, Recall, F1-score, and Mean Average Precision (mAP), the YOLOv8 model was tested in a partially occluded environment. The data set includes snapshots of faces of differing sizes and degrees of occlusion, which was used to test the ability of the model to identify faces under diverse conditions, including faces with partial obstruction. The details of the relevant outcomes are presented in the following table.

Table 1

Model Evaluation Result Based on Key Metrics

Epoc h	Precisio n	Reca ll	F1- Scor e	mAP@5 0	mAP@5 0-95
150	0.93	0.887	0.90	0.92	0.512

The Precision of this model is 0.93%, and Recall score is 0.887%, both values of which are relatively high, meaning that the model is very accurate and effective in recognizing the greatest number of faces; nevertheless, it can be improved. The F1 -score of 0.90 shows a reasonable degree of Precision and recall. The model performs excellently by the score of mAP@50, though its mAP@50-95 of 0.512% indicates a weakness in efficacy as the IoU threshold increases, especially on heavily occluded faces. Based on the outcomes, the models exhibit a high level of efficacy in face detection, but further optimization is

necessary to enhance the accuracy in adverse environments.

This high performance of YOLOv8 in standard and moderately occluded status can be explained by its anchor-free detection mechanism, C2f backbone, Bidirectional Feature Pyramid Network (BiFPN), which jointly allows effective multi-scale fusion of features and precise localization of some facial features, including eyes and nose, even in partially covered cases [28, 29]. However, it works poorly in the case of extreme occlusion and low light since important discriminative features become perceptually blurred, leading to decreased feature-map activation in higher layers. When there is low light, the convolutional filters are not able to identify the clear gradient and edge information whereas when there is heavy occlusion, the areas of the face that are not visible cause class imbalance and decrease in the confidence scores. All this explains why the lower mAP@50-95 score (0.512) and the necessity to apply more illumination normalization, synthetic occlusion training, and multimodal inputs improve robustness in extreme conditions [30,31].

4.2 Confusion Matrix Analysis

The findings given in Figure 4 are associated with the confusion matrix that illustrates thorough evaluation of the classification performance of the YOLOv8 model through comparisons between the predicted and ground truth labels. The true positives, false positives and false negatives can be identified in the figure which allows understanding of the distribution of the model in detecting faces under varying conditions including partial occlusion. Confusion matrices evaluate the model's accuracy and its tendency to make errors of specific types, such as misclassifying faces or ignoring partially occluded ones.

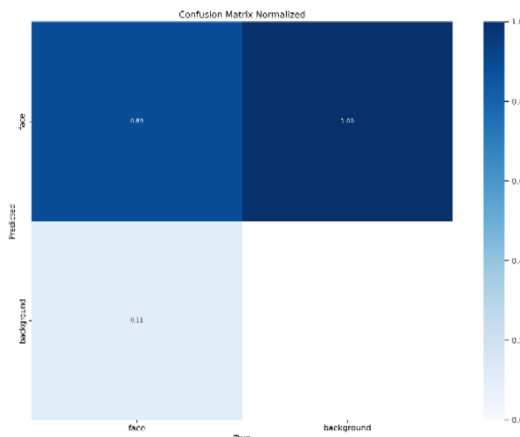


Fig. 4. Confusion Matrix for YOLOv8

The above figure indicates that the actual positive rate of 0.89 in face detection signifies that the model correctly identifies most faces. Conversely, the background detection rate is low (0.11), with

infrequent cases of non-faces being identified as faces. The model is reliable in distinguishing between faces and non-faces, as demonstrated by a perfect score of 1.00 in identifying background areas. The findings of YOLOv8, in general, demonstrate a strong effectiveness in face recognition even in challenging conditions, such as partial occlusion and insufficient illumination.

4.3 YOLOv8 Training and Validation Loss Curves

Figure 5 displays the training and validation loss curves, along with the assessment metrics. The figure illustrates the key features of the YOLOv8 model, with a particular emphasis on face detection, particularly for partially occluded faces in low-light conditions.

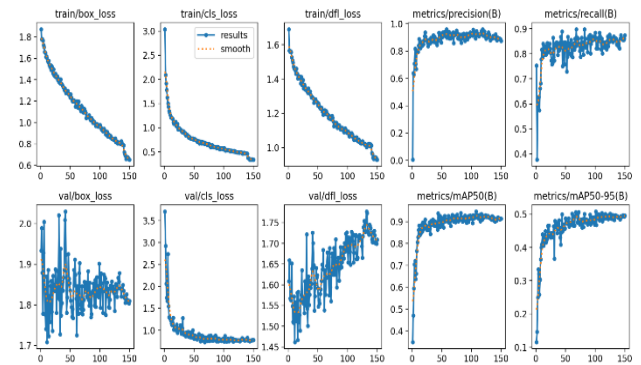


Fig. 5. Training and Evaluation Curves of YOLOv8

The above graph displays several key results, including box loss, classification loss, detection loss, Precision, recall, and Mean Average Precision (mAP). The selected Model YOLOv8 provides results that show a consistent decrease in train/box_loss, train/cls_loss, and train/dfl_loss over time, indicating that the model exhibits progressive improvement through iterative learning. Moving on to the metrics for Precision (B) and recall (B), curves show an upward trend, highlighting the model's enhanced ability to recognize faces more accurately with reduced false alarm rates. The results of mAP@50(B) and mAP@50-95(B) demonstrate a considerable improvement in the models' overall performance despite the challenging environment, including occlusion and low light. On the other hand, the validation/loss graphs exhibit prominent instability in validation losses, particularly in cases where insufficient lighting is present during the validation process. Apart from these issues, the model shows stable results with improved detection and Precision.

4.4 Bounding Box Attribute Analysis

Figure 6 presents a visual comparison of the main features in pairs: x-coordinate, y-coordinate, width, and height of the bounding boxes recognized by the model under low-light conditions and partial face coverage.

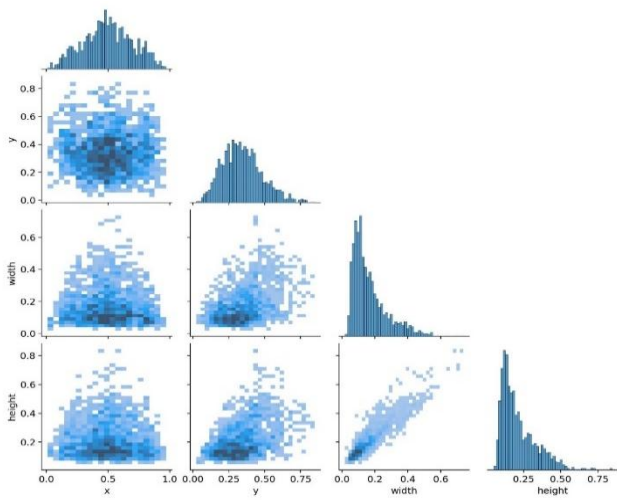


Fig. 6. Bounding Box Attribute Distribution for YOLOv8

The matrix provides a broad view of how traits that are linked together show the model's ability to detect faces in a harsh environment. Here, the diagonal histogram shows the distribution of each property. The histogram above presents numerical results for each property, illustrating their distribution and density. According to the x-coordinate of the histogram, the horizontal placement of identified faces shows that they are mostly grouped in certain parts of the picture. The parameters that models use to identify faces in difficult circumstances, like partially occluding faces, are height and width of histograms. In other parameters, like the x- and y-coordinates of the scatterplot, the correlation of the vertical and horizontal position of the objects, like helmets, in the defined frame are given.

The above matrix is an important tool to assess the effectiveness of YOLOv8 in processing partially obscured faces. The performance of the model as indicated by the above-mentioned matrix is according to the trend of the scatterplots, especially when the faces are covered by a mask like a helmet or a scarf. In addition, the figure also demonstrates the trends in the model output, like clustering of bounding boxes with a specific size or area, which shows that the system can be improved in case of occlusions, and when the lighting conditions vary.

The features of model performance may be analyzed through the analysis of those bounding box properties created by YOLOv8 in a scatterplot matrix, which assesses the capacity of the model to operate and identify faces in extremely low-lighting conditions.

4.5 YOLOv8 Model Prediction

Figure 6 illustrates the results obtained from the YOLOv8 model experiment on faces partially covered in low light.



Fig. 7. Sample Image Prediction of the YOLOv8 Model

In each picture, a known face is surrounded by a box that displays a confidence level (between 0.3 and 0.9), indicating the model's accuracy. In different situations, the model provides significant performance. It can accurately identify faces even when they are partly hidden, such as when people wear hats or scarves, and it can also do this in low-light conditions. The confidence scores indicate the model's level of certainty in various situations. A higher confidence score (around 0.8 to 0.9) indicates that the detections are more likely to be correct.

In comparison, a lower confidence score (around 0.3 to 0.6) suggests that the face is either completely hidden or poorly lit, which may result in less accurate detections. Despite these challenges, YOLOv8 demonstrates its ability to accurately detect faces in various real-life scenarios. However, the model has some issues with extreme occlusions, which are clear in pictures that get lower confidence scores.

4.6 Comparative Evaluation with Other Models

To contextualize YOLOv8's performance, the results were compared with other well-established object detection models, including Faster R-CNN, RetinaFace, and YOLOv7. All models were trained and tested using the same dataset and hyperparameters where applicable.

Table 2 presents a summary of key performance metrics (Precision, Recall, F1-score, and mAP@50).

Table 2

Comparative Performance of YOLOv8 with Contemporary Models

Model	Precision	Recall	F1-score	mAP@50	mAP@50-95	Ref
Faster R-CNN	0.88	0.83	0.85	0.89	0.46	[25]

RetinaF ace	0.91	0.85	0.8 8	0.90	0.49	[2 6]
YOLO v7	0.92	0.86	0.8 9	0.91	0.50	[2 7]
YOLO v8 (Proposed)	0.93	0.88	0.9 0	0.92	0.512	[2 8]

The result of the comparisons proves that YOLOv8 is more precise and faster than its predecessors in partially occupied and low-light settings. However, the mAP@50-95 values indicate that both the Faster R-CNN and RetinaFace are competitive on the heavily obscured instances, which may open the chance of future studies to consider a hybrid model.

5. Conclusion

Based on the findings achieved by using YOLOv8 model to identify faces in partially covered masks and low-light scenarios, one can assume that the model has enough face recognition features and has the capacity to solve real-life problems of surveillance systems. The results obtained using F1-scores, as per the provided environment (low light with occluded mask), further strengthen the narrative. It can be a challenging task for other models used in face recognition systems to generate this type of accuracy. Moving on to another parameter, such as confusion vectors and performance measures of YOLOv8, shows its best performance as compared to its previous versions for object detection. The decreasing number in the mAP@50-95 result indicates that the model exhibits good performance in a challenging situation, such as low light. In summary, for real-world applications, YOLOv8 is a highly suitable choice, particularly for law enforcement agencies in forensic investigations, as it effectively performs in challenging environments and lays a solid foundation for future progress in object detection technology.

5.1 Future recommendation

According to future suggestions, the YOLOv8 model can further be tested by presenting it with a severe low-light condition so its actual strength can be analyzed, in face recognition, by adding additional pre-processing techniques such as increased light or personalized noise reduction algorithms to enhance its recognition capability in cases where there is very low light. In addition, its performance may be further improved by providing extra training to the model with bigger datasets or fine-tuning data produced through augmented reality (AR) simulation. The model may be further enhanced by utilizing multi-

model face recognition with YOLOv8, supplemented with infrared or depth sensors, to ensure consistent performance in evening conditions. This kind of advanced development can make Face recognition systems more efficient and flexible, particularly in practical applications. Moreover, future research should focus on improving YOLOv8's resilience through practical enhancements. First, preprocessing techniques such as contrast-limited adaptive histogram equalization (CLAHE), noise filtering, and illumination correction should be explored to improve detection in severely low-light environments. Second, adopting multimodal detection by combining RGB images with infrared or depth sensors can provide additional spatial and thermal cues, helping the system identify partially occluded faces more effectively. Third, transfer learning using pretrained weights from large facial or object-detection datasets can significantly reduce training time while enhancing model generalization. Integrating these methods can yield a more robust, real-time facial detection pipeline for surveillance and forensic applications.

Competing Interests: The Authors do not have any competing interests

Funding Information: This research work is not funded by any public or private organization

Author Contribution: Amina Shaikh, the study's primary author, conducted data analysis using the dataset. The second author, Imran Khan Keerio, wrote the text; the third author, Mutiullah Shaikh, designed the study; and the fourth author, Prof. Dr. Asdullah Shah, examined and approved the final version.

Data Availability Statement: A customized dataset was created to simulate surveillance in which faces were partially covert in low-light settings, and can be provided upon request

References

- [1] H. T. Rauf, M. I. U. Lali, S. Zahoor, S. Z. H. Shah, A. U. Rehman, and S. A. C. Bukhari, "Visual features based automated identification of fish species using deep convolutional neural networks," *Comput. Electron. Agric.*, vol. 167, p. 105075, 2019.
- [2] S. Cui, Y. Zhou, Y. Wang, and L. Zhai, "Fish detection using deep learning," *Appl. Comput. Intell. Soft Comput.*, vol. 2020, no. 1, p. 3738108, 2020.
- [3] K. M. Knausgård, A. Wiklund, T. K. Sjørdalen, K. T. Halvorsen, A. R. Kleiven, L. Jiao, and M. Goodwin, "Temperate fish detection and classification: a deep learning based approach," *Appl. Intell.*, vol. 52, no. 6, pp. 6988–7001, 2022.

- [4] D. Xu and Y. Wu, "Improved YOLO-V3 with DenseNet for multi-scale remote sensing target detection," *Sensors*, vol. 20, no. 15, p. 4276, 2020.
- [5] Y. H. Lee and Y. Kim, "Comparison of CNN and YOLO for object detection," *J. Semicond. Display Technol.*, vol. 19, no. 1, pp. 85–92, 2020.
- [6] C. Dewi, B. J. Tsai, and R. C. Chen, "Shapley additive explanations for text classification and sentiment analysis of Internet Movie Database," in *Proc. Asian Conf. Intell. Inf. Database Syst.*, Singapore, 2022, pp. 69–80.
- [7] J. Wang, Y. Yuan, and G. Yu, "Face attention network: An effective face detector for the occluded faces," *arXiv preprint arXiv:1711.07246*, 2017.
- [8] Z. Zhao, K. Hao, X. Ma, X. Liu, T. Zheng, J. Xu, and S. Cui, "SAI-YOLO: A lightweight network for real-time detection of driver mask-wearing specification on resource-constrained devices," *Comput. Intell. Neurosci.*, vol. 2021, no. 1, p. 4529107, 2021.
- [9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.
- [10] Q. Liu et al., "A multitask model for real-time fish detection and segmentation based on YOLOv5," *PeerJ Comput. Sci.*, vol. 9, p. e1262, 2023.
- [11] H. Deng, J. Zhang, L. Chen, and M. Cai, "Improved mask wearing detection algorithm for SSD," *J. Phys. Conf. Ser.*, vol. 1757, no. 1, p. 012140, 2021.
- [12] H. Malik, A. Naeem, S. Hassan, F. Ali, R. A. Naqvi, and D. K. Yon, "Multi-classification deep neural networks for identification of fish species using camera captured images," *PLoS One*, vol. 18, no. 4, p. e0284992, 2023.
- [13] G. Jocher, A. Chaurasia, and J. Qiu, "YOLO by Ultralytics," 2023.
- [14] M. Hussain, "YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection," *Machines*, vol. 11, no. 7, p. 677, 2023.
- [15] T. Wu and Y. Dong, "YOLO-SE: Improved YOLOv8 for remote sensing object detection and recognition," *Appl. Sci.*, vol. 13, no. 24, p. 12977, 2023.
- [16] J. Zheng, H. Wu, H. Zhang, Z. Wang, and W. Xu, "Insulator-defect detection algorithm based on improved YOLOv7," *Sensors*, vol. 22, no. 22, p. 8801, 2022.
- [17] F. Prameswari, H. Octafiani, and T. Haryanto, "You Only Look Once (YOLOv8) for fish species detection," in *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 1359, no. 1, p. 012023, 2024.
- [18] P. Mishra, A. Biancolillo, J. M. Roger, F. Marini, and D. N. Rutledge, "New data preprocessing trends based on ensemble of multiple preprocessing techniques," *TrAC Trends Anal. Chem.*, vol. 132, p. 116045, 2020.
- [19] X. Luo and S. Pradhan, "Evaluation metrics," in *Anaphora Resolution: Algorithms, Resources, and Applications*. Berlin, Heidelberg: Springer, 2016, pp. 141–163.
- [20] G. Silpalatha and T. S. Jayadeva, "Accelerating fast and accurate instantaneous segmentation with YOLO-v8 for remote sensing image analysis," *Remote Sens. Appl. Soc. Environ.*, vol. 37, p. 101502, 2025.
- [21] Z. Liu, R. R. D. Abeyrathna, R. M. Sampurno, V. M. Nakaguchi, and T. Ahamed, "Faster-YOLO-AP: A lightweight apple detection algorithm based on improved YOLOv8 with a new efficient PDWConv in orchard," *Comput. Electron. Agric.*, vol. 223, p. 109118, 2024.
- [22] Y. Wang and J. He, "A rapid concrete crack detection method based on improved YOLOv8," *IEEE Access*, 2025.
- [23] M. Safaldin, N. Zaghdien, and M. Mejdoub, "An improved YOLOv8 to detect moving objects," *IEEE Access*, 2024.
- [24] S. Pan, Z. Chang, and Z. Chang, "WD-1D-VGG19-FEA: An efficient wood defect elastic modulus predictive model," *Sensors*, vol. 24, no. 17, p. 5572, 2024.
- [25] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," In *Advances in neural information processing systems*, vol.28, 2015.
- [26] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "Retinaface: Single-shot multi-level face localisation in the wild," In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, (pp. 5203-5212), 2020.

- [27] C. H. Kang, and S. Y. Kim, ‘Real-time object detection and segmentation technology: an analysis of the YOLO algorithm,” JMST Advances, vol. 5, no. 2, p. 69-76, 2023.
- [28] T. Wu, and Y. Dong, “YOLO-SE: Improved YOLOv8 for remote sensing object detection and recognition,” Applied Sciences, vol. 13, no. 24, 12977, 2023.
- [29] Z. Liu, R. R. D. Abeyrathna, R. M. Sampurno, V. M. Nakaguchi, and T. Ahamed, Faster-YOLO-AP: A lightweight apple detection algorithm based on improved YOLOv8 with a new efficient PDWConv in orchard. Computers and Electronics in Agriculture, 223, p. 109118, 2024.
- [30] C. Shorten, and , T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning. Journal of big data, vol. 6, no. 1, p.1-48, 2019.
- [31] L. Wang, Y. Chang, and K. Wang, “Dual-Branch Multimodal Fusion Network for Driver Facial Emotion Recognition,” Applied Sciences, vol. 14, no. 20, p. 9430, 2024.