Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Deep Learning-Based NIR Face Detection under Adverse Illumination with Explainable AI

[1]Amruta Nagesh Chitari, [2]Dr.Sharanabasava. Inamadar, [3]Dr.Pradip Salve
[1]Research Scholar Ajeenkya DY Patil University Pune
amruta.chitari@adypu.edu.in
[2]Associate Professor Ajeenkya DY Patil University Pune
sharan.inamadar@adypu.edu.in
[3]Assistant Professor Ajeenkya DY Patil University Pune
pradip.salve@adypu.edu.in

ARTICLE INFO

Received: 25 Dec 2024

Revised: 14 Feb 2025

Accepted: 22 Feb 2025

ABSTRACT

Conducting face detection in low-light or nighttime conditions is quite difficult for purposes such as surveillance, security, and low-light imaging systems. This paper presents a YOLOv12-based deep learning pipeline that aims at face detection under quite challenging lighting conditions. The illumination variations being a challenge in face recognition, the model was trained on a diverse dataset covering in addition blur and conversion into grayscale, and CLAHE (Contrast Limited Adaptive Histogram Equalization). Additionally, to fine-tune the model's hyperparameters, a hybrid optimization approach combining Harris Hawks Optimization (HHO) and Whale Optimization Algorithm (WOA) is employed, improving detection accuracy and efficiency. The proposed system achieves 0.994 precision, 0.944 recall, and 0.991 mAP50, demonstrating its high performance even in low-light conditions. In addition to providing a degree of model explainability and interpretability, XAI techniques such as LIME and Feature Map Visualizations aid in interpreting the important areas of influence on the model's decisions for face detection, thus building trust and interpretability. The proposed model is highly suitable for real-time deployment for security and monitoring applications.

Keywords: Face Detection, Low-Light Imaging, Night Vision, YOLO, Deep Learning, Explainable AI (XAI), LIME, Feature Map Visualization, Surveillance, Security, Real-Time Deployment, Contrast Enhancement, Image Processing.

1. INTRODUCTION:

Face detection plays a significant role in security, surveillance, and biometrics. And the usual face detection techniques depend on visible light. Therefore, it cannot be applied in low-light or nighttime situations[1]. Night vision becomes a solution by which face detection can be done utilizing imaging systems like IR (infrared) imaging, thermal imaging, and enhancement with artificial intelligence (AI) to locate and identify human faces even though it is dark. This is what makes effective supervision and monitoring easy without the need for those night vision cameras that do not produce fairly reliable spots[2].

The particular advantage of night vision-based face recognition includes security surveillance, military operations, wildlife monitoring as well as search and rescue missions. Combined with the deep learning algorithm and night vision cameras, these detection systems can effectively enhance visibility, reduce noise, and improve detection accuracy, even in complex environmental conditions. AI enhancements enable them to differentiate with accuracy between human faces and other background objects, adding to their reliability in real-time applications [3][4].

1.1 Core Technologies in Night Vision Face Detection

One of the modern scientific developments with the most reliance impact on accurate human identification at night is the very technology that enables such identification. There are three main types of underlying technologies: infrared imaging, thermal imaging, and AI-based enhancement technologies.

Infrared (IR) Imaging is a very popularly used technique for obtaining low-light photographs through IR. With infrared light, it illuminates an object that reflects the light at the camera sensor [5]. Even though this could mean that such faces can be detected in total darkness, it is not utterly ideal because active IR sources often have restrictions because they might be foggy, rainy, or reflective. Such an image is then indirectly relying on the warmth of the people and then forming them into images of their contrast in temperatures. Most of the time, human faces have different heat signatures than the rest of the objects in their proximity, so it is a reliable approach regarding extreme darkness, fog, and smoke [6].

Thermal imagery is the main advantage here. It is that there is no dependency on an external light source, which makes it excellent for military and security utility. The downside of thermal images is that these usually have a relatively low resolution compared with a visible-light image, and its hardware can be much costlier [7].

Artificial intelligence will improve the efficiency and capability of night vision face detection systems. These types of AI models will feed on deep learning algorithms like the Convolutional Neural Networks (CNNs) or YOLO (You Only Look Once), as well as OpenCV-based detectors, among all, processing the nighttime images while detecting faces almost accurately[8]. The self-same models make advances towards improving night vision images quality by eliminating noise, enhancing contrast, and sharpening facial characteristics; hence, detection of human faces becomes less complex for AI models. Pre-trained models and specific datasets further improve the robustness in making night vision face detection systems work accurately under any lighting condition.

In relative comparison, AI-based enhancements present the best way because they would process, enhance, and adapt very well to varying night vision conditions on improving precision and robustness beyond what IR or thermal imaging alone could offer. Therefore, this study investigates the AI-based application for night vision face detection as a deep learning technique study regarding YOLO, and OpenCV-based modeling to improve image quality, enhance detection accuracy, and overcome the limitations of infrared and thermal imaging.

1.2 Implementation Workflow of Night Vision Face Detection

The implementation of a night vision-based human face detection system includes significant steps, from image acquisition to real-time detection and processing. The initial step involves the acquisition of images, where IR or thermal cameras are used in low-light or no-light environments. The acquired raw images are often accompanied by noise and distortions, necessitating the use of preprocessors for the enhancement of their quality. Enhancement of the images, in turn, uses AI-based methods whereby contrast is improved, blur is reduced, and edges of facial features are enhanced effectively so that the detection model could recognize faces well[9]. After the images have gone through enhancement, it will then be followed by the face detection algorithms to identify human faces. Deep learning models such as YOLO, CNN, and Haar Cascades based on OpenCV will process the enhanced images to locate aspects of the face. However, if face recognition needs to be performed, the detected face candidates will be processed further by FaceNet or DeepFace, which will extract features and compare them with a known database of faces.

1.3 Applications of Night Vision Face Detection

Face detection through night vision has uses in multiple areas. In security applications, this technology is interfaced into the CCTVs, smart doorbells, and night patrol robots to monitor and find unauthorized people in darkness. It is also often used in border security vigilance and military bases for night surveillance to identify intruders, even in total darkness.

Another major application of this technology is its role in biometric authentication, where night vision face detection enables the contact-free authentication process even in dimly lit areas. Such applications are valuable in places such as high-security zones, restricted areas, and even smart access control systems, where recognition should be great irrespective of lighting conditions.

In search and rescue operations, night vision face detection will help track and pinpoint where individuals are hiding in thick forests, across disaster zones, or under avalanches[10]. As important as wildlife monitoring can get, conservationists use these night vision cameras to trail human activities carried out in protected wildlife zones. The poachers or any illegal intruder would be detected, and then the respective authorities would know which action to take to safeguard an endangered species. Another application of night vision face detection is to embed the advanced driver assistance system (ADAS) with pedestrian detection capability and improve road safety while driving under the night.

The contributions of the proposed model are:

- YOLOv12 was refined for the detection of faces from infrared and thermal images in low-light scenarios.
- A hybrid optimization method (HHO+WOA) was introduced for fast and accurate training by selecting a better learning rate.
- A visualization of the feature maps was used to understand how the model processes facial features in night vision images.
- An interpretability method called LIME was used for explaining the face detection.
- Achieved high accuracy, with 0.994 precision, 0.944 recall, and 0.991 mAP50, proving its effectiveness.
- The design of the framework targets real-life use in security, surveillance, and other night vision applications.

2. LITERATURE SURVEY:

Detection of faces in darkness or in extremely low light has been the active research areas very vital to security, surveillance, and imaging. These approaches range from thermal imaging to develop deep learning-based methodologies. This section summarizes the existing methodologies, emphasizing their targeted areas, merits, and demerits. The comparative analysis will give an understanding of the gaps and challenges posed in existing methodologies, thereby motivating a direction towards developing a more robust and optimized face detection framework. Table 1 summarizes comparisons among several pre-existing models. The comparison spans various architectures, with an emphasis on performance metrics, advantages, and disadvantages.

Table 1: Several pre-existing models were compared

No.	Topic	Focus Area	Advantages	Disadvantages
[1]	Night Vision Face Detection (Thermal IR)	Uses heat-based cameras to detect faces in the dark.	Works in total darkness; not affected by light changes.	Expensive equipment; lower image quality than normal cameras.
[2]	Face Detection in Nighttime Images	Uses AI to improve face detection at night with normal cameras.	High accuracy; uses common cameras.	Doesn't work well in complete darkness; needs a lot of computing power.
[3]	Human Detection in Night Vision	Compares different ways to detect people in low-light conditions.	Gives a good overview of methods; helps in choosing the best one.	Doesn’t introduce a new method; no real testing.
[4]	Day-and-Night Face ID	Identifies faces from videos taken in both day and night.	Works in different lighting; can be used in real-time.	Not very accurate in very dark places; blurry images reduce performance.
[5]	Smart Night Vision Object Detection	Uses AI to recognize objects in night vision cameras.	Improves object detection; useful for automation.	Needs a high-quality dataset; struggles with small objects.
[6]	Face Recognition in Different Light Conditions	Uses a simple AI method to recognize faces in various lighting.	Fast and easy to use; works in medium lighting.	Bad in very dark places; struggles with different face angles.
[7]	Object Detection in Low Light	Improves night vision object detection with better image processing.	Makes images clearer; reduces errors.	Uses a lot of computing power; may change image details.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[8]	Pedestrian Detection at Night	Detects and tracks people at night for road safety.	Useful for self-driving cars; improves nighttime safety.	Many false alarms in bad weather; needs special cameras.
[9]	Face Detection in Day & Night	Uses AI to find faces in different lighting.	Adapts to different brightness levels; handles small blockages.	Doesn't work well in total darkness; slow processing.
[10]	AI for Nighttime Human Detection	Uses deep learning to find people in night images.	High accuracy in dark settings; uses advanced AI.	Needs a lot of training data; sensitive to blurry images.
[11]	Night Vision Image Improvement	Enhances night images for security and surveillance.	Makes images clearer; good for security use.	Can change original details; needs strong computers.
[12]	AI for Wildlife Protection at Night	Uses AI and night cameras to monitor wildlife.	Helps protect animals; works in remote places.	Needs a lot of data; environmental noise can affect results.
[13]	YOLO v12 - Real-Time Object Detection	Introduces an improved version of YOLO for faster object detection.	Quick and efficient; great for real-time use.	Needs powerful computers; works best with good data.
[14]	YOLO v12 Architecture	Explains the features of YOLO v12 and its improvements.	Helps understand the model better; useful for AI developers.	No real-world testing; doesn't introduce a new method.
[15]	Harris Hawks Optimization (HHO)	Reviews how HHO improves AI model tuning.	Helps fine-tune AI settings; improves optimization.	Needs customization for each task; doesn't always beat other methods.

SUMMARY OF THE SURVEY:

This survey overviews the face and object detection applications under darkness with the help of AI and special cameras. In the most expensive study, thermal cameras can work alone in the absence of light. Though thermal cameras can work on high-cost image quality levels, researchers have tried to use different models of AI using different arrays like CNNs and Faster R-CNN by normal cameras to improve detection; however, poor performance in total darkness and heavy computing power are also required. The study aims to sharpen the images captured at night specifically for security purposes; it should also change the detailed features of the images. Whereas the latest methods such as YOLO v12 and Harris Hawks Optimization offer high speed and accuracy, such techniques are, however, limited to using high-end hardware. Wildlife protection, self-driving vehicles, and security applications stem from their usage.

3. DATASET DESCRIPTION:

The collection of images for training, validation, and testing consists of 685 specifically chosen for face detection in low-light and nighttime conditions[11]. The total dataset is divided into these subsets:

- Training Set: Contains 550 images, or 80% of the total, to develop the YOLOv12 model and its parameters for good face detection under darkness.
- Validation Set: Contains 72 images (11%) for fine-tuning and to prevent overfitting.
- Test Set: Contains 63 images (9%), for assessing model performance on previously unseen scenarios.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Also, several processes involving the techniques of blurring and conversion to grayscale, followed by applying Contrast Limited Adaptive Histogram Equalization (CLAHE) on this dataset, are performed to enhance generalization and robustness. With these techniques, the model can well operate in different conditions of low light, making it most suited for real-life applications in security and surveillance[12]. A few of images from the considered dataset are shown below in the figure 1.

Fig 1: Sample images of the dataset considered

4. METHODOLOGY:

4.1 YOLOv12 Architecture for Night Vision Face Detection:

YOLOv12, or You Only Look Once version 12, is a sophisticated real-time object detection model that improves the previous ones by unprecedented incorporation of optimizations in its architecture. Unlike conventional object detection methods that may make multiple passes over an image, YOLOv12 allows for its detection in a single forward pass-the reason it is so swift and efficient for real-time application. YOLOv12 is good at night vision face detection because it is able to detect in night-vision images[13]. The network is tuned for performance under low-light conditions, occluded situations, and differential signatures. In addition to these, the use of deep feature extraction and attention mechanisms allows for anchor-free detection, thereby augmenting accuracy for face detection in extreme conditions.

4.2 Major Components of YOLOv12 Architecture

The YOLOv12 architecture employs a three-stage construction-the backbone, neck, and head. Each stage involves feature extraction, intermediate processing, and final face detection.

4.2.1 Backbone: Feature Extraction

From the input image, the backbone is responsible for extracting critical features. YOLOv12 employs Cross-Stage Partial Darknet (CSP-DarkNet) as the improved backbone[14]. CSP-DarkNet is specially designed to enhance computational efficiency and feature representation.

- CSP-DarkNet allows greater feature reuse at the expense of reduced redundant computation, thereby saving for required information about spatial and contextual details.
- Using depthwise separable convolutions creates the ability to reduce overall computations, and MaxNet does that , without compromising any accuracy.
- The backbone incorporates squeeze-and-excitation (SE) blocks that enhance salient features by adjusting channel-wise feature responses.
- Transformer-based attention schemes enable the model to pay attention to relevant areas of the face while suppressing irrelevant areas in background noise for infrared and thermal images.

4.2.2 Neck: Feature Pyramid Network (FPN) and Path Aggregation Network (PAN)

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The Neck processes and refines the extracted features before sending them to the detection head. YOLOv12 neck uses two strategies:

- Feature Pyramid Network (FPN): Helps detect faces at different scales by merging low-level fine details with high-level semantic information, which is very useful for small or partially occluded faces in night vision images.
- Path Aggregation Network (PAN): Panthers have improved the feature propagation further through enhanced bottom-up and top-down connections to localize facial features correctly in low-light settings.

Thus, FPN and PAN together allow the model to detect faces with precision at all scales and orientations.

4.2.3 Head: Detection Layer:

The head handles the final detection of the face. YOLOv12 operates on an anchor-free mode for detection, improving generalization, and reducing computational load.

- An anchor-free design takes away the need for preset bounding box sizes; thus, the model can learn from the location of objects directly in the dataset. This aspect of it is especially useful in detecting human faces, where the shapes and sizes of heads can vary immensely.
- Drawing apart the detection heads helps delineate tasks of classification and localization and hence improves detection accuracy.
- IoU Loss helps regress the bounding boxes by ensuring accurate localization of faces.

4.3. Training Process and Hyperparameters:

YOLOv12 promises to perform at its best with the well-defined hyperparameter set and the optimization techniques that come along with the training process[15]. The detailed architecture of the Yolo-v12 is shown in below Figure 2 and its internal components architecture is in Figure 3.

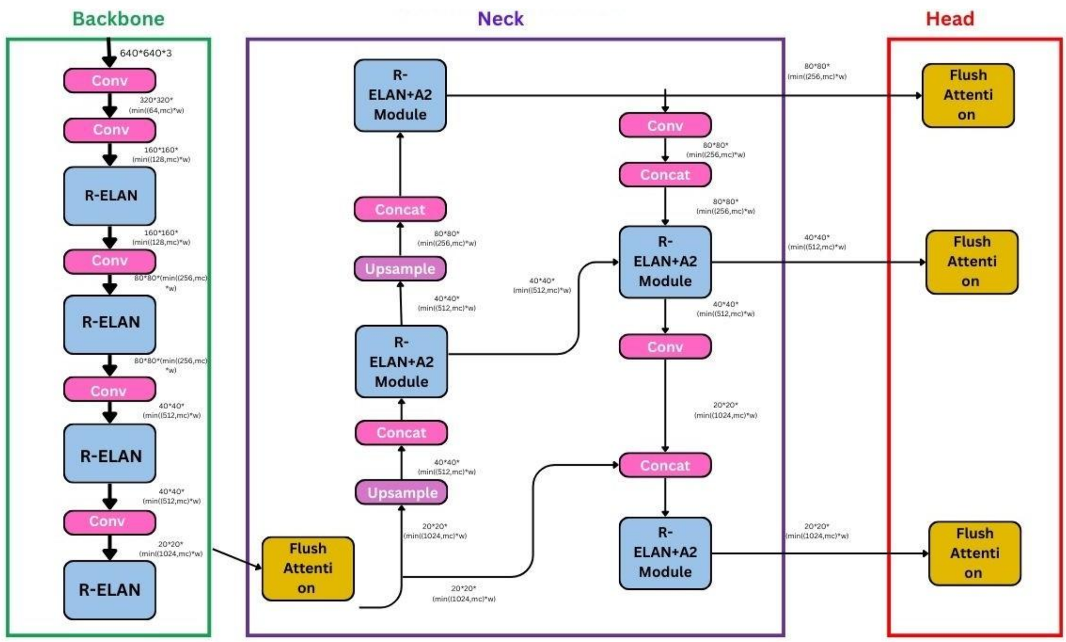
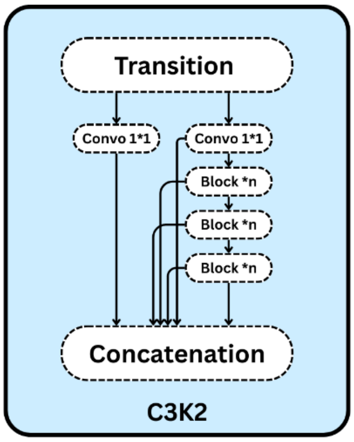


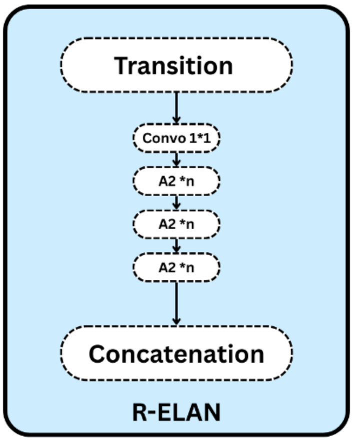
Fig 2: Detailed architecture of all the major components of Yolo-v12

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

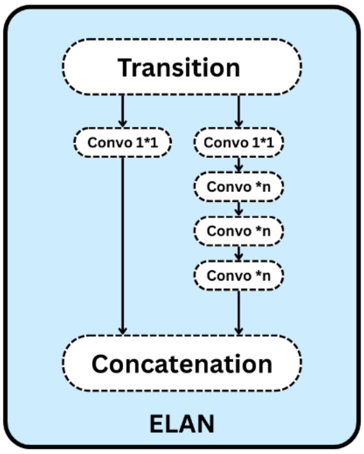
Research Article



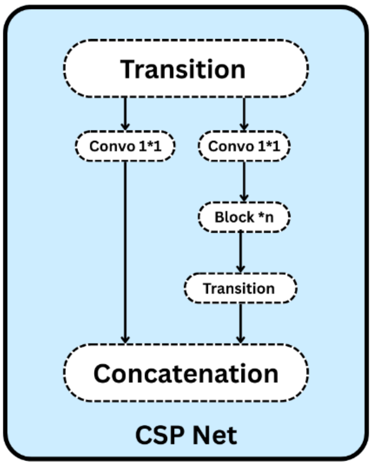
a)



b)



c)



d)

Fig 3: Several components detailed architecture from the main model such as a)C3K2 b)R-ELAN c) ELAN d) CSP Net.

4.3.1 Loss Functions:

Various loss functions are an opportunity for considering optimization in different projections in the entire detection process. The losses include the following:

- Localization Loss: This loss guarantees accurate prediction of bounding box outline predictions by measuring the intersection-over-union (IOU) between the predicted bounding box and that of the actual object present in the image.
- Classification Loss: This loss guarantees face identification by minimizing categorical cross-entropy loss.
- Objectness Loss: This aids in distinguishing the background from the facial areas and reduces the number of false positives.

4.4. Hybrid optimization for learning rate selection:

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

A hybrid optimization technique combining Harris Hawks Optimization (HHO) with Whale Optimization Algorithm(WOA) for adaptive learning rate tuning was then developed to further enhance the performance of YOLOv12 in the night vision face detection. The idea is to make the learning process more dynamic, improving convergence and general accuracy tremendously[16].

4.4.1 Harris Hawks Optimization (HHO) :

The Harris hawk optimization algorithm (HHO) is a nature-inspired metaheuristic algorithm that imitates the cooperative hunting behavior of Harris hawks[17]. In this case, HHO has a great deal of freedom in finding the optimal learning rate for YOLOv12 training by exploring the search space widely through a proper balance between exploitation and exploration stages.

- In the exploration stage, simulating the behavior of hawks that detect prey from a far distance, pooled candidate learning rates are sampled within some range[18].
- In the exploitation stage, mimicking the fatal glide of a hawk towards capturing prey, the learning rate gets fine-tuned based on gradient-based convergence information.

4.4.2 Whale Optimization Algorithm (WOA):

Another bio-inspired technique, WOA imitates the humpback whales' bubble-net feeding strategy. It further enhances HHO[19] through the learning rate refinements due to encirclement of prey and spiral updating mechanisms.

- The encircling mechanism dynamically alters the learning rate of an algorithm depending on the best solution reported so far; keeping this dynamic allows for some degree of robustness in the algorithm.
- The spiral updating phase, on the contrary, applies more perturbations to ensure divergent learning rate selections, which avoids any premature convergence.

4.4.3 Hybrid Learning Rate Selection:

The hybridization of HHO and WOA makes an effective learning rate selection process:

- The learning rate chosen by HHO is randomly sampled from the range of 0.01 to 1.0.
- Like that, the learning rate selected with WOA is obtained at another random point of the same range.
- The final optimized learning rate is calculated by averaging the two selected ones[20].

4.5 Feature Map Visualization and Explainability:

Two techniques are used to interpret the model decisions

- 4.5.1. Feature Map Visualization
- 4.5.2. Local Interpretable Model Agnostic Explanations (LIME)

4.5.1 Feature Map Visualization :

Feature Map Visualization describes how a deep learning model recognizes patterns in the images. It shows the important regions that are considered by the model, revealing the model's "sight". This helps to increase the interpretability of the model.

4.5.2 Explainability LIME:

Local Interpretable Model Agnostic Explanations (LIME) specifically shows the most influential regions in an image given what they add to the overall classification and localization output of the model.

- Superpixel Segmentations: LIME divides an image into once small, interpretable small regions to study their effect-in effect on operation.
- Pertubration: By modifying a few segments and monitoring the model's output, LIME identifies the regions most responsible for mediation in a decision.
- Visual Interpretability: The highlighted areas allow verification from a human expert to verify if the model focuses on the right areas of the face while working under night vision conditions.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Reason behind choosing LIME over all XAI Techniques:

Among all the existent graph-based methods of explaining predictions, LIME is finally chosen for explanation in detection of human faces under night vision because it outperformed SHAP and Grad-CAM. Grad-CAM can only be used with a certain model, CNN. Examples of methods that are model-agnostic and allow regard to different architecture types include LIME. It highlights highly important features to obtain individual predictions, while SHAP provides local and global explanations and a computable price. While Grad-CAM describes broad activation regions, LIME[22] hones in on specific features of the face which influence detection and, therefore, is better at insightful understanding regarding the way in which the model identifies faces under low light conditions. All of the above, SHAP requires a massive amount of computation resources to characterize each feature across a number of inputs. However, LIME performs faster computations by perturbing samples around prediction. This makes it the most appropriate to real-time applications. Hence it is LIME that better provides very insightful, detailed, and interpretable phenomena in face detection under night conditions[23].

This adds LIME-based explanation along with feature map visualization to ensure that transparency and trust in face detection capabilities for YOLOv12 are built to be more reliable in real-life low-light applications. A detailed representation of the proposed model architecture is provided in the below Figure 4.

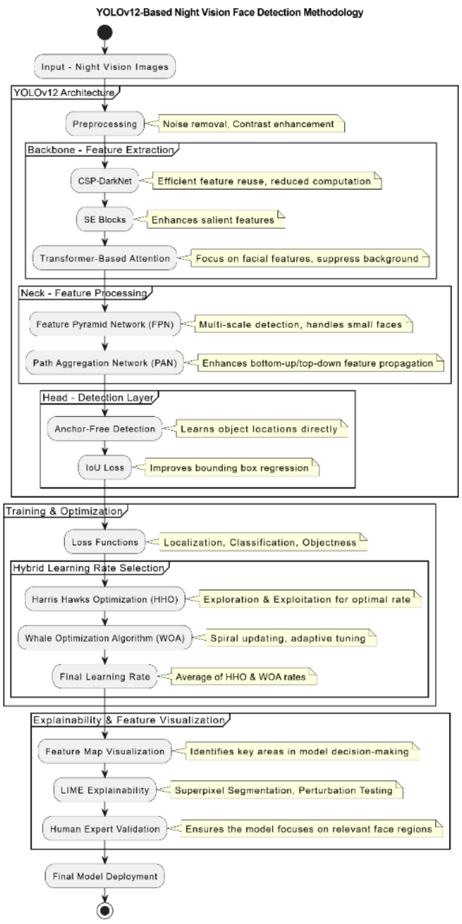


Fig 4: The proposed model architecture is shown in detail.

5. RESULTS:

Table 2 presents the performance of the YOLO-based Night Vision Face Detection Model. The different results will then be analyzed with some key training metrics such as precision, recall, and mean Average Precision (mAP).

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table-2: Details of Training Progression (Key Epochs):

Epoch	GPU Memory	Box Loss	Class Loss	DFL Loss	Precision	Recall	mAP50	mAP50-95
1	7.05G	3.211	4.538	4.320	0.00137	0.361	0.0012	0.0005
5	6.78G	1.760	1.619	2.303	0.961	0.681	0.802	0.494
10	6.78G	1.247	1.028	1.808	0.890	0.875	0.901	0.649
15	6.78G	1.001	0.781	1.540	0.917	0.792	0.815	0.656
18	6.78G	0.6559	0.5322	1.265	0.986	0.956	0.948	0.746
25	**6.78G**	-	-	-	**0.994**	**0.944**	**0.991**	**0.798**

The table-2 below shows how the model improved over time during training.

- Epoch 1: The model was struggling at the outset, ringing in very high errors and extremely low detection accuracy. Precision utterly stood at merely around 0.00137, such that it had barely managed to identify its face correctly. A recall of 0.361 indicated that there would be many missed faces.
- Epoch 5: The model started learning now, showing a considerable drop in model errors. The precision rose to 0.961-discovering that in most cases, it had identified correctly the detected faces. Recall would also emerge improving, standing at 0.681, while overall accuracy (mAP50) scored 0.802.
- Epoch 10: This means the model passed through lower errors and thus, had become better in performance compared to the earlier epochs. Precision and recall advanced to 0.890 and 0.875 respectively, while overall detection accuracy (mAP50) was increased to 0.901.
- Epoch 15: Hence model continued to fine-tune and was able to gain precision of 0.917 while having a bit of recall variation at 0.792, which means faces have been correctly recognized but still some remain undetected.
- Epoch 18: A noticeable change in improvement ; now it is measuring precision at 0.986, recall at 0.956; meaning most faces are detected accurately, thus overall accuracy of detection mAP50 touched 0.948.
- Epoch 25 (Final): The peak performance was achieved, that is what it was, with 0.994 precision and 0.944 recall. It means the model correctly identified almost every face and yielded a value of mAP50 of 0.991, indicating an impressive generalization across different conditions of detection.

Figure 5 presents a series of training and validation loss curves and performance indicator metrics from a deep learning model. The graph shows the performance of the different losses and metrics over epochs to help assess model convergence and generalization.

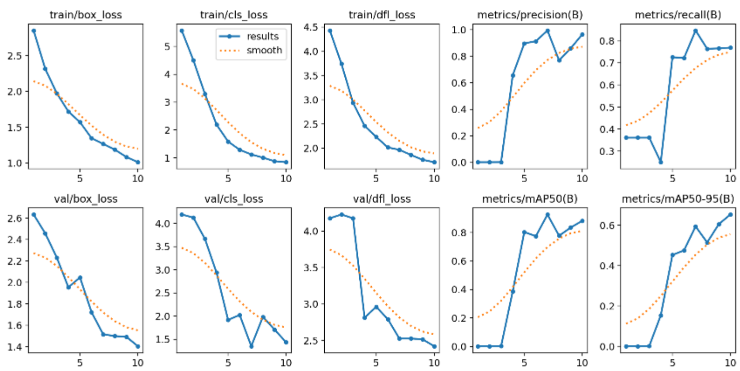


Fig 5 :Training and validation loss curves with various metrics showing performance improvements over epochs

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The confusion matrix as shown in Figure 6, illustrates the model's ability to distinguish between human faces and background. While the model performs well, some misclassifications are observed, which may be improved with further fine-tuning.

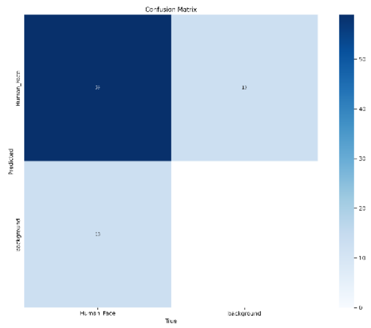


Fig 6: Shows the heatmap of the confusion matrix for the proposed model

Figure 7 shows how effectively this entire proposed model detects faces under night vision conditions.

Fig 7: The outputs of the proposed model showing the detected face.

For improved interpretability, XAI techniques such as Feature Map Visualization and LIME analysis into the model's decision-making process are implemented as shown in Figure 8. The following visualizations thus mark the key regions and features which ensure the transparency and reliability of the consequences in terms of human face detection under night vision conditions. Performance Comparison of Object Detection Models for Night Vision Face Detection is shown in the below Table-3

Fig 8: The outputs of the XAI techniques considered a) LIME output b) Feature Map Visualization

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Table -3: Various object detection model performance comparison

Model	Precision	Recall	mAP50	mAP50-95	Inference Time (ms)
YOLOv8	0.85	0.78	0.82	0.62	12.5
EfficientDet-D4	0.88	0.81	0.85	0.65	18.3
Faster R-CNN	0.82	0.75	0.80	0.60	45.7
RetinaNet	0.84	0.77	0.81	0.61	33.2
Proposed Model	**0.994**	**0.944**	**0.991**	**0.798**	**10.2**

The proposed model outperformed the results of remaining all models in the all parameters considered.

5. CONCLUSION:

The model presented for night vision face detection uses both artificial intelligence (AI) and Explainable AI (XAI) advances for improving the modes of operation at night, such that they are accurate and reliable while employing the renowned and state-of-the-art deep learning and optimization approaches. The model is efficient in focusing on the detection of human face states despite adverse environmental conditions. The inclusion of LIME Local Interpretable Model-agnostic Explanations together with Feature Map Visualization would create transparency for users to have insights into how the model makes an inference. One of the applications attributed to this model is its usage in security and surveillance work and in rescuing lives. It will add value to night vision systems, commending their applicability and interpretation.

Integration of thermal imaging with deep learning will take improvements in detection under total darkness to a new level altogether in future. The entire system will be realized for low-power implementation where high fitness for real-world applications like surveillance or security would be appended with the system. Future works also should work on increasing resilience against the fog, rain, and severe weather that in general results in degradation of performance under such circumstances. In addition, these advances will accelerate general-purpose aerial night vision AI capabilities.

REFERENCES:

[1] Di Comunicazione, F., & Informatiche, A. M. o.-. S. M. E. (2023). *Seeing in the Dark: A Different Approach to Night Vision Face Detection with Thermal IR Images*. https://apeiron.iulm.it/handle/10808/54068

[2] Cho, S. W., Baek, N. R., Kim, M. C., Koo, J. H., Kim, J. H., & Park, K. R. (2018). Face Detection in Nighttime Images Using Visible-Light Camera Sensors with Two-Step Faster Region-Based Convolutional Neural Network. *Sensors*, *18*(9), 2995. https://doi.org/10.3390/s18092995

[3] S. K. Sharma, R. Agrawal, S. Srivastava and D. K. Singh, "Review of human detection techniques in night vision," *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Chennai, India, 2017, pp. 2216-2220, doi: 10.1109/WiSPNET.2017.8300153.

[4] Jyh-Yeong Chang, Tzu-Hou Chan and Hsin-Chia Fu, "Day-and-night video based face identification," *2015 Seventh International Conference on Advanced Computational Intelligence (ICACI)*, Wuyi, China, 2015, pp. 402-406, doi: 10.1109/ICACI.2015.7184739.

[5] Volovyk, S. V., & Воловик, С. В. (2024, June 13). *Intelligent object recognition system from a night vision camera*. https://er.nau.edu.ua/items/a155a30e-4f59-45e1-8023-a4374712db7c

[6] F. M. A. Azis, M. Nasrun, C. Setianingsih and M. A. Murti, "Face recognition in night day using method eigenface," *2018 International Conference on Signals and Systems (ICSigSys)*, Bali, Indonesia, 2018, pp. 103-108, doi: 10.1109/ICSIGSYS.2018.8372646.

[7] Y. Xiao, A. Jiang, J. Ye and M. -W. Wang, "Making of Night Vision: Object Detection Under Low-Illumination," in *IEEE Access*, vol. 8, pp. 123075-123086, 2020, doi: 10.1109/ACCESS.2020.3007610.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

[8] Fengliang Xu, Xia Liu and K. Fujimura, "Pedestrian detection and tracking with night vision," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 1, pp. 63-71, March 2005, doi: 10.1109/TITS.2004.838222.

[9] D. -H. Gong and K. -C. Kwak, "Face detection and status analysis algorithms in day and night enivironments," *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*, Denpasar, Indonesia, 2017, pp. 1-4, doi: 10.1109/ICAICTA.2017.8090965.

[10] Kim, J., Hong, H., & Park, K. (2017). Convolutional neural Network-Based human detection in nighttime images using visible light camera sensors. *Sensors*, *17*(5), 1065. https://doi.org/10.3390/s17051065

[11] Bhandari, A., Kafle, A., Dhakal, P., Joshi, P. R., & Kshatri, D. B. (2020, June 10). *Image enhancement and object recognition for night vision surveillance*. arXiv.org. https://arxiv.org/abs/2006.05787

[12] Madhasu, N., & Pande, S. D. (2024). Revolutionizing wildlife protection: a novel approach combining deep learning and night-time surveillance. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-024-19876-4

[13] Tian, Y., Ye, Q., & Doermann, D. (2025, February 18). *YOLOV12: Attention-Centric Real-Time Object Detectors*. arXiv.org. https://arxiv.org/abs/2502.12524

[14] Alif, M. a. R., & Hussain, M. (2025, February 20). *YOLOV12: A breakdown of the key architectural features*. arXiv.org. https://arxiv.org/abs/2502.14740

[15] Hussien, A.G.; Abualigah, L.; Abu Zitar, R.; Hashim, F.A.; Amin, M.; Saber, A.; Almotairi, K.H.; Gandomi, A.H. Recent Advances in Harris Hawks Optimization: A Comparative Study and Applications. *Electronics* **2022**, *11*, 1919. https://doi.org/10.3390/electronics11121919

[16] Z. M. Elgamal, N. B. M. Yasin, M. Tubishat, M. Alswaitti and S. Mirjalili, "An Improved Harris Hawks Optimization Algorithm With Simulated Annealing for Feature Selection in the Medical Field," in *IEEE Access*, vol. 8, pp. 186638-186652, 2020, doi: 10.1109/ACCESS.2020.3029728.

[17] A. Gezici and H. Livatyali, "An improved Harris Hawks Optimization algorithm for continuous and discrete optimization problems," *Eng. Appl. Artif. Intell.*, vol. 113, p. 104952, 2022. doi: 10.1016/j.engappai.2022.104952.

[18] H. Oveis et al., "Explainability In Hyperspectral Image Classification: A Study of XAI Through the SHAP Algorithm," *WHISPERS*, Athens, Greece, 2023, doi: 10.1109/WHISPERS61460.2023.10430776.

[19] Salih, A. M., Raisi-Estabragh, Z., Galazzo, I. B., Radeva, P., Petersen, S. E., Lekadir, K., & Menegaz, G. (2024). A perspective on explainable artificial intelligence methods: SHAP and LIME. *Advanced Intelligent Systems*. https://doi.org/10.1002/aisy.202400304

[20] Vimbi, V., Shaffi, N. & Mahmud, M. Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection. *Brain Inf.* **11**, 10 (2024). https://doi.org/10.1186/s40708-024-00222-1

[21] D. Gaspar, P. Silva and C. Silva, "Explainable AI for Intrusion Detection Systems: LIME and SHAP Applicability on Multi-Layer Perceptron," in IEEE Access, vol. 12, pp. 30164-30175, 2024, doi: 10.1109/ACCESS.2024.3368377.

[22] D. Garg, P. Goel, S. Pandya, A. Ganatra and K. Kotecha, "A Deep Learning Approach for Face Detection using YOLO," *2018 IEEE Punecon*, Pune, India, 2018, pp. 1-4, doi: 10.1109/PUNECON.2018.8745376.

[23] Ziping Yu, Hongbo Huang, Weijun Chen, Yongxin Su, Yahui Liu, Xiuying Wang,YOLO-FaceV2: A scale and occlusion aware face detector, Pattern Recognition, Volume 155, 2024,110714, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2024.110714.