International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Object Detection in Low-Light Conditions Using YOLOV7

Ms. Hujaimah Nida

MTech Student, Computer Science, Jamia Hamdard (Deemed to be University)

ABSTRACT:

Reduced contrast, increased noise, and shifting lighting levels are some of the factors that make object recognition in low light circumstances extremely difficult. Accurate identification in these situations is essential for applications where reliable performance is essential, such as driverless cars, security monitoring, and surveillance systems. This work makes use of the cutting-edge YOLOv7 object recognition model, It is renowned for its remarkable real-time processing speed and precision. The model is trained and evaluated using the ExDark dataset, which comprises twelve different object categories captured in different low-light scenarios.Adaptive histogram equalization and picture normalization are two preprocessing techniques used to enhance image quality and boost detection efficiency before to feeding the data into the YOLOv7 model. Training is done using pre-trained weights from the COCO dataset, which has been specially adapted to recognize objects in low light. The model is evaluated using standard assessment criteria, such as mean average precision (mAP), recall, and precision. With a recall of 0.6556 and a precision of 0.7153, YOLOv7 performs better on the same dataset than previous iterations such as YOLOv5 and IA-YOLO, according to the experimental results. The advanced network architecture of YOLOv7, which incorporates trainable Bag-of-Freebies and E-ELAN modules to enable efficient Feature extraction in low light, is credited with the enhanced performance. The results highlight YOLOv7's potential for useful implementation in applications that need precise object detection in difficult illumination conditions. To further maximize real-time performance on edge devices, future improvements might include incorporating sophisticated picture augmentation techniques, image enhancement algorithms, and lightweight designs. The importance of continuously improving object detection technology is highlighted in this study in order to overcome the ongoing problems with low-light detection.

KEYWORDS: YOLOV7, deep learning, object detection, low-light video, ExDark dataset, Autonomous surveillance, Mean Average Precision (mAP).

INTRODUCTION:

Due to its ability to extract intricate patterns from input and produce very accurate models, deep learning has become a crucial technique in this model, particularly for object detection. Accuracy has greatly improved and computing efficiency has been optimized as CNN-based detection algorithms have developed over time. The high computational cost of Region-Based CNN (R-CNN), which creates candidate regions for object detection, is still a disadvantage. To improve speed and accuracy even further, Faster R-CNN incorporates a Region Proposal Network (RPN). By directly extracting region recommendations from feature maps, Fast R-CNN improves efficiency. However, by analyzing an entire image in a single forward pass, YOLO (You Only Look Once) significantly increases detection speed. Similar to this, Single Shot Detection (SSD) recognizes objects of various sizes by using a CNN backbone. From YOLOv1 to YOLOv7, the YOLO framework has improved feature extraction and recognition under difficult circumstances by employing techniques including bag of freebies, focal loss, and Extended efficient layer aggregation networks, or E-ELANs. Despite these improvements, it is still difficult to recognize objects in low light due to growing noise and deteriorating vision. To tackle this issue, the current study employs YOLOv7 for low-light object detection utilizing the ExDark dataset, which comprises a diverse collection of indoor and nighttime photographs. Experimental results show that YOLOv7 outperforms previous iterations and other CNN-based models under similar tough conditions. This document is formatted as follows: Section 2 provides a review of pertinent literature; Section 3 explains the YOLOv7 architecture and training process; Section 4 describes the dataset and methodology; Section 5 discusses the experimental results; and Section 6 provides the conclusion.

RELATED WORK:

In low-light conditions, such as poorly lit interior rooms, nighttime settings, and unfavorable weather conditions like fog, object identification has been the subject of numerous research. An enhanced Faster R-CNN was combined with Deep Convolutional Generative Adversarial Networks (DCGANs) in one method to improve object detection at night. To increase the accuracy of detection, we employed multi-scale Region of Interest (ROI) pooling and deep convolutional feature fusion. By improving image quality, a novel strategy called the Image-Adaptive YOLO (IA-YOLO) framework sought to improve detection performance.

This method taught CNN-PP and YOLOv3 simultaneously in an end-to-end fashion and was evaluated on the ExDark dataset. Similarly, the Multi-Scale Domain Adaptive YOLO (MS-DAYOLO) model, which improves object identification in low light, was built using YOLOv3. Researchers used the ExDark dataset to validate its efficacy, showing that collaborative semantic learning methods may be used for object detection in difficult lighting and weather scenarios. Additionally, a particularly tailored version of Mask R-CNN was developed for pedestrian detection on poorly lit roadways.This method accelerated detection speed by enhancing the deep learning and eliminating the instance mask branch. Another strategy combined YOLOv7 with a CNN-based defogging algorithm to help with visibility issues and enhance road object recognition in foggy circumstances. A improved technique for identifying things in low light also involved using YOLOv5 to detect visually appealing images created by an image enhancement algorithm. An orthogonal tangent regularity network was also used in a unique technique called Multitask Auto Encoding Transformation (MEAT) to increase recognition accuracy in low light. This technique blends elements inspired by humans with machine vision.

EXISTING METHOD:

A deep learning-based object recognition system called YOLOv3 (You Only Look Once version 3) successfully strikes a balance between speed and accuracy. Its fully convolutional architecture, which is based on Darknet-53, features residual connections and multi-scale detection for enhanced performance. By forecasting bounding boxes over three different scales using feature maps at varying resolutions, YOLOv3 is able to accurately identify small, medium, and large objects. The approach improves detection accuracy by utilizing logistic regression and anchor boxes. Compared to two-stage detectors

like Faster R-CNN, YOLOv3 operates with a single forward pass, which significantly speeds it up while maintaining competitive accuracy.
Its grid-based approach predicts class probabilities and a number of bounding boxes for each cell. To ensure that only the most relevant predictions are retained, redundant overlapping detections are removed using Non-Maximum Suppression (NMS). The model is trained using either pre-learned weights from the COCO dataset or a custom dataset created for specific tasks. Because of its interaction with popular deep learning frameworks like PyTorch, TensorFlow, and Darknet, it may be deployed in a variety of ways. Even with the advancements in later versions like YOLOv4 and YOLOv5, YOLOv3 is still widely utilized due to its efficacy in real-time item recognition jobs.

PROPOSED METHODOLOGY:
To increase accuracy in low light, this study proposes an optimized object identification framework based on the YOLOv7 model. The proposed technique uses preprocessing techniques like noise reduction and histogram balancing to maximize image quality prior to detection with YOLOv7, which incorporates enhanced anchor boxes for object localization and tunable learning rates to optimize performance, is enhanced using the ExDark dataset. Data augmentation methods like brightness normalization and contrast correction are applied to further increase the model's adaptability. modified the non-maximum suppression and the confidence threshold (NMS) are two post-processing techniques to lower false positives. By implementing these enhancements, the proposed method seeks to outperform existing algorithms in object detection in challenging lighting conditions.

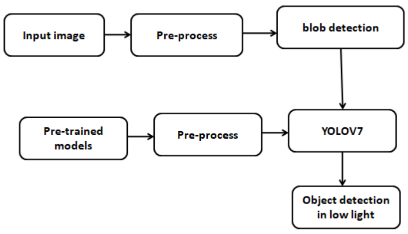


Figure1. Architecture model of project

Input Image: in this taking low light dataset for lowlight object detection, is used in this work to supply input photos to the YOLOv7 model. The 7,363 images in this dataset include precise bounding boxes labeled for 12 distinct item types, including people, tables, and animals. All images are resized to 640 by 640 pixels before processing in order to ensure that they comply with the YOLOv7 detection framework. Since most of the collection's photos were taken in low light, the visibility of objects is often limited, making accurate detection challenging. Preprocessing is crucial since the format and quality of the input photographs have a significant impact on the model's overall detection ability.
Pre-processing: Before the YOLOv7 network analyzes the input image, it goes through many preparation steps to enhance feature extraction. One of these processes is normalization, which involves scaling pixel values to a preset range, such as [0,1] or [-1,1], to ensure consistent model training. Additionally, to increase the model's robustness and scenario-adaptability, data augmentation techniques including rotation, flipping, and cropping are applied. The dataset annotations are also converted using the YOLO format, which systematically encodes each object with attributes such as

width, height, x-center, y-center, and class index. This preprocessing step is essential for optimizing image interpretation and ensuring that YOLOv7 accurately recognizes objects in the given dataset.

Blob Conversion: After preprocessing, the image is converted into a blob, a binary large object format suitable for deep learning models. The blob conversion ensures that each image maintains the same batch format and 640x640 resolution for efficient GPU processing. In order to preserve uniformity across all training and validation images, pixel values are normalized in this stage. The blob model, which reduces computing costs while preserving essential image characteristics for accurate detection, enables YOLOv7 to analyze photos more efficiently.

YOLOv7 Model Detection: YOLOv7 is utilized in this work for object recognition in low light because it is quicker and more accurate than earlier YOLO versions. The latest YOLOv7 version, which is officially supported, employs advanced training techniques known as the "bag of freebies," can improve the model's capacity for low-level feature learning. This feature is quite useful when it comes to identifying items in low light. Using the ExDark dataset, the study assesses YOLOv7 and contrasts its results with the most advanced low-light detection techniques. This comparison assesses YOLOv7's performance under challenging lighting conditions, highlighting its potential for improved object detection.

E-ELAN model :

This model incorporates the model network backbone network to optimize memory economy and improve gradient flow during back propagation. The main goal of its architecture is to make learning more efficient by minimizing the distance gradients must travel between layers. By implementing a more advanced version of the ELAN computational block, E-ELAN enhances the network's learning capacity while maintaining the integrity of gradient propagation. Important techniques like expansion, shuffling, and merging cardinality allow the model to progressively acquire complex features without affecting gradient flow. These architectural improvements lead to improved object identification job performance, particularly under difficult conditions.

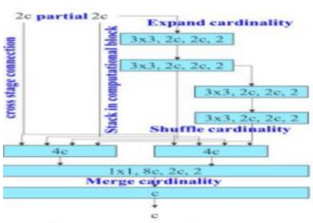


Figure 1. The architecture of E-ELAN[14]

Trainable Bag-of-Freebies:

The network head in YOLOv7 generates the final detection outputs. The lead head and the auxiliary head are the two main parts of the architecture. While the lead head is responsible for producing the final object detection results, the auxiliary head facilitates the training procedure by enhancing feature learning in the intermediate layers. This dual-head configuration facilitates improved gradient propagation across the network, which enhances detection accuracy and training stability. These two elements working together guarantees that the model efficiently improves its predictions, which improves performance on object identification tasks.

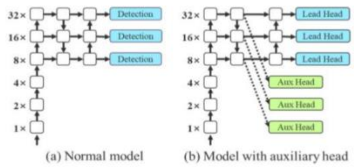


Figure 2: YOLOv7's auxiliary and lead heads in comparison to standard single-head models

Dataset, training and validation process:

To work with the low light video based object detection, specialized datasets such as NightOwls for pedestrian identification and Dark Face for nighttime facial recognition have been developed. The largest publicly available collection is ExDark, which consists of 7,363 images categorized into 12 item classes. For YOLOv7 training, the data set divided into two types train and test with 80 and 20 percentages for validation. To accommodate YOLOv7, the annotation format was changed. 50 training epochs were conducted using PyTorch with pre-trained COCO weights. The model performed well in low-light identification, as seen by its precision score of 0.8253 and recall of 0.7956.

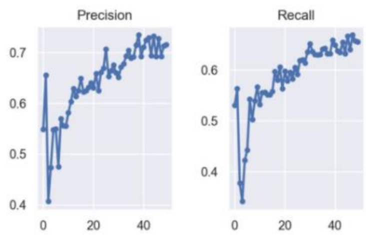


Figure 3 The precision and recall values with epoch during the ExDark dataset training of YOLOv7 are displayed

The YOLOv7 model is used to perform object detection after the blob has been generated. To enhance feature extraction and optimize gradient flow, especially in low-light photos, this is a feature of YOLOv7. Additionally, the model employs Trainable Bag-of-Freebies techniques to enhance its ability to learn complex low-light features. The architecture consists of two prediction heads: an auxiliary head that aids in training and a lead head that makes final predictions. The model generates bounding boxes, class labels, and confidence ratings after image processing. Performance examination using key measures output parameters. mAP@0.5 was 0.725, precision was 0.823, and recall was 0.713. These results show that YOLOv7 performs better than other models and earlier versions of YOLO when it comes to recognizing things in low light.

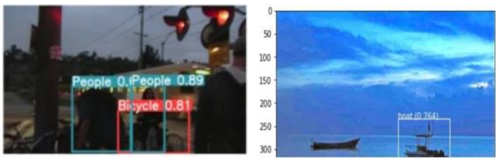
RESULTS AND DISCUSSION:

A confusion matrix that displays the detection accuracy for every item class and the mean average precision (mAP) at a classification threshold of 0.5 is used to evaluate the performance of the YOLOv7 model. The mAP measure, a widely used standard for assessing object detection models, enables comparisons with the findings of other investigations. According to the confusion matrix for the ExDark dataset, which displays class-wise detection findings, the "table" category had the lowest accuracy while the "humans" category had the highest because backdrop pieces were incorrectly categorized as tables. Similarly, some objects, such as cats, were sometimes mistaken for dogs because of their visual resemblance. The detection findings from the ExDark dataset further support the YOLOv7 model's low-light capability.

International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

An example of low-illumination objects detection. Our detector have achieved amazing results in some common scenes of low-illumination.



To assess YOLOv7's effectiveness in low light, its performance was compared with three advanced deep-learning-based detection methods using the ExDark dataset. The first, known as Image Adaptive YOLO (IA-YOLO), employs preprocessing techniques including defogging and sharpening before using YOLOv3 for detection. The second is the traditional YOLOv5 as a baseline. With a shared-weight detection head and an additional feature learning module designed for low-light settings, the third is an improved YOLOv5 variation. In terms of recall and overall performance, YOLOv7 beat IA-YOLO and traditional YOLOv5, even though the upgraded YOLOv5 had somewhat higher mAP and precision scores. These results suggest that YOLOv7 reduces false negatives while potentially increasing false positives.

Method	mAP @ 0.5	Precision	Recall
Image adaptive Yolo3	0.403	--	--
YoloV5	0.642	0.672	0.573
Improved Low-Light YoloV5	0.698	0.715	0.629
YoloV7(Our Work)	0.725	0.823	0.713

Table 1: YOLOv7's Low-Light Detection Performance Comparison Using ExDark Dataset This dataset, a sizable collection of photos taken in low light, was used to assess YOLOv7's performance in low light. model higher detection accuracy than YOLOv5 demonstrated the effect of important enhancements like the E-ELAN backbone, compound scaling, auxiliary head, and re-parameterized convolution. These improvements allowed the model to extract useful low-level properties even though it wasn't made with low-light detection in mind. Using image enhancing techniques to highlight important information in low light can yield even better results.

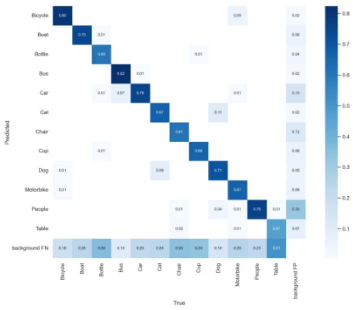


Figure 4: Confusion matrix for ExDark dataset classifications based on YOLOv7

Research on the latest developments in deep learning-based image augmentation methods has shown promise for enhancing object detection in difficult scenarios. To improve YOLOv7's detection capabilities, future research could combine it with deep learning-based picture augmentation techniques. By enhancing feature visibility and contrast, preprocessing input photos can assist YOLOv7 achieve higher accuracy, especially in low-light conditions.

CONCLUSION AND FUTURE WORK:

The YOLOv7 model demonstrated good performance in low-light object recognition with a mAP 0.5 of 0.725, a precision of 0.823, and a recall of 0.713 in experiments conducted on the ExDark dataset. Under challenging lighting conditions, the E-ELAN design, trainable Bag-of-Freebies techniques, and advanced preprocessing techniques increased its effectiveness. According to the results, YOLOv7 works better than earlier iterations of YOLO and other low-light detection models, making it a dependable option for uses including nighttime monitoring, autonomous driving, and security surveillance. Still, there is room for development. To increase detection accuracy, future studies could concentrate on using cutting-edge low-light improvement techniques. Additionally, YOLOv7's detection of small or partially obscured objects may be improved by fine-tuning it on bigger low-light datasets and adding attention techniques.

Its useful applications across a range of real-world settings could be expanded by optimizing real-time deployment for mobile and edge devices.

REFERENCES:

1. Girshick, R., Donahue, J., Darrell, T. & Malik, J. Rich feature hierarchies for accurate object detection and semanticsegmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, 580–587 (2014).
2. Girshick, R. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision, 1440–1448 (2015).
3. Ren, S., He, K., Girshick, R. & Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks.
4. Adv. neural information processing systems 28 (2015).
5. Liu, W. et al. Ssd: Single shot multibox detector. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14, 21–37 (Springer, 2016).
6. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, 779–

788 (2016).8/10
7. Redmon, J. & Farhadi, A. Yolo9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition, 7263–7271 (2017).
8. Redmon, J. & Farhadi, A. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767 (2018).
9. Bochkovskiy, A., Wang, C.-Y. & Liao, H.-Y. M. Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 (2020).
10. Yao, Z., Cao, Y., Zheng, S., Huang, G. & Lin, S. Cross-iteration batch normalization. In Proceedings of the IEEE/CVFconference on computer vision and pattern recognition, 12331–12340 (2021).
11. Ding, X. et al. Repvgg: Making vgg-style convnets great again. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 13733–13742 (2021).
12. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).
13. Howard, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprintarXiv:1704.04861 (2017).
14. Tarvainen, A. &Valpola, H. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Adv. neural information processing systems 30 (2017).
15. Neumann, L. et al. Nightowls: A pedestrians at night dataset. In Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part I 14, 691–705 (Springer, 2019).
16. Yang, W. et al. Advancing image understanding in poor visibility environments: A collective benchmark study. IEEE Transactions on Image Process. 29, 5737–5752 (2020).
17. Loh, Y. P. & Chan, C. S. Getting to know low-light images with the exclusively dark dataset. Comput. Vis. Image Underst. 178, 30–42 (2019).
18. Yolov5 github release. https://github.com/cs-chan/Exclusively-Dark-Image-Dataset. Accessed: 2013-08-17.
19. Lin, T.-Y. et al. Microsoft coco: Common objects in context. In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, 740–755 (Springer, 2014).
20. Kaur, M., Kaur, J. & Kaur, J. Survey of contrast enhancement techniques based on histogram equalization. Int. J. Adv. Comput. Sci. Appl. 2 (2011).
21. Rasheed, M. T., Guo, G., Shi, D., Khan, H. & Cheng, X. An empirical study on retinex methods for low-light image enhancement. Remote. Sens. 14, 4608 (2022).
22. Ren, W. et al. Low-light image enhancement via a deep hybrid network. IEEE Transactions on Image Process. 28, 4364–4375 (2019).
23. Shen, L. et al. Msr-net: Low-light image enhancement using deep convolutional network. arXiv preprint arXiv:1711.02488 (2017).
24. Li, C. et al. Low-light image and video enhancement using deep learning: A survey. IEEE transactions on pattern analysis machine intelligence 44, 9396–9416 (2021).
25. Wang, K. & Liu, M. Z. Object recognition at night scene based on dcgan and faster r-cnn. IEEE Access 8, 193168–193182 (2020).

26. Liu, W. et al. Image-adaptive yolo for object detection in adverse weather conditions. In Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, 1792–1800 (2022).
27. Hnewa, M. & Radha, H. Multiscale domain adaptive yolo for cross-domain object detection. In 2021 IEEE International Conference on Image Processing (ICIP), 3323–3327 (IEEE, 2021)
28. Huang, S.-C., Le, T.-H. & Jaw, D.-W. Dsnet: Joint semantic learning for object detection in inclement weather conditions. IEEE transactions on pattern analysis machine intelligence 43, 2623–2633 (2020).
29. Lai, K., Zhao, J., Liu, D., Huang, X. & Wang, L. Research on pedestrian detection using optimized mask r-cnn algorithm in low-light road environment. In Journal of Physics: Conference Series, vol. 1777, 012057 (IOP Publishing, 2021).
30. Qiu, Y., Lu, Y., Wang, Y. & Jiang, H. Idod-yolov7: Image-dehazing yolov7 for object detection in low-light foggy traffic environments. Sensors 23, 1347 (2023).
31. Wang, J. et al. Research on improved yolov5 for low-light environment object detection. Electronics 12, 3089 (2023).
32. Cui, Z. et al. Multitask aet with orthogonal tangent regularity for dark object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 2553–2562 (2021).