

ULAB_CSE499_Template (5).pdf

by Md. Yusuf Hossain

Submission date: 17-Feb-2025 09:28PM (UTC+0600)

Submission ID: 2591144359

File name: ULAB_CSE499_Template_5_.pdf (5.27M)

Word count: 10917

Character count: 64630

Object Detection of Low Light Images Using Modified YOLOv8

by

Md Yusuf Hossain (201014089)
Rezowana Zayed Zahin (201014096)
Md. Rashed -Ul- Hoque (201014097)
Joyonti Rani Karmakar (191014052)

⁶
*Project/Thesis report (CSE499) submitted in partial fulfillment of the
requirements for the degree of*

Bachelor of Science in Computer Science and Engineering

Under the supervision of

Mohammed Ashikur Rahman, PhD



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
UNIVERSITY OF LIBERAL ARTS BANGLADESH
FALL 2024**

© Md Yusuf Hossain, Rezowana Zayed Zahin, Rashiq-Ull-Hoque, Joyonti Rani Karmakar
All rights reserved

DECLARATION

Project Title Y-ZA YOLOv8: Object Detection of Low Light Images
Authors Md Yusuf Hossain, Rezowana Zayed Zahin, Rashiq-Ul-Hoque, Joyonti Rani Karmakar
Student IDs 201014089, 201014096, 201014097, 191014052
Supervisor Mohammed Ashikur Rahman

5
We declare that this project/thesis report entitled *Title of Capstone Project* is the result of our own work except as cited in the references. The project/thesis report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Md Yusuf Hossain
201014089

13
Department of Computer Science and Engineering, ULAB

Rezowana Zayed Zahin
201014096

13
Department of Computer Science and Engineering, ULAB

Md. Rashiq -Ul- Hoque
201014097

13
Department of Computer Science and Engineering, ULAB

Joyonti Rani Karmakar
191014052

13
Department of Computer Science and Engineering, ULAB

Date: January 30, 2025



Department of Computer Science and Engineering
University of Liberal Arts Bangladesh
Mohammadpur, Dhaka - 1207

CERTIFICATE

This is to certify that the project/thesis report entitled **Y-ZA YOLOv8: Object Detection of Low Lights Images**, submitted by **Md Yusuf Hossain** (201014089), **Rezowana Zayed Zahin**² (201014096), **Rashiq-Ul-Hoque** (201014097), **Joyonti Rani Karmakar** (191014052) is a undergraduate student of the **Department of Computer Science and Engineering** has been examined. Upon recommendation by the examination committee, we hereby accord our approval of it as the presented work and submitted report fulfill the requirements for its acceptance in partial fulfillment for the degree of *Bachelor of Science in Computer Science and Engineering*.

Mohammed Ashikur Rahman, PhD, Asst. Professor
Dept. of CSE, University of Liberal Arts Bangladesh

²
Muhammad Golam Kibria, PhD, Professor
Dept. of CSE, University of Liberal Arts Bangladesh

Raihan Kibria, Senior Lecturer
Dept. of CSE, University of Liberal Arts Bangladesh

Suravi Akhter, Lecturer
Dept. of CSE, University of Liberal Arts Bangladesh

Place: Dhaka
Date: January 30, 2025

ACKNOWLEDGEMENTS

We would like to express our deep and sincere gratitude to our research supervisor, *Mohammed Ashikur Rahman*, for giving us the opportunity to conduct research and providing invaluable guidance throughout this work. His dynamism, vision, sincerity and motivation have deeply inspired us. He has taught us the methodology to carry out the work and to present the works as clearly as possible. It was a great privilege and honor to work and study under his guidance.

We are greatly indebted to our honorable teachers of the Department of Computer Science and Engineering at the University of Liberal Arts Bangladesh who taught us during the course of our study. Without any doubt, their teaching and guidance have completely transformed us to the persons that we are today.

We are extremely thankful to our parents for their unconditional love, endless prayers and caring, and immense sacrifices for educating and preparing us for our future. We would like to say thanks to our friends and relatives for their kind support and care.

Finally, we would like to thank all the people who have supported us to complete the project work directly or indirectly.

Md Yusuf Hossain , **University of Liberal Arts Bangladesh**
Rezowana Zayed Zahin , **University of Liberal Arts Bangladesh**
Rashiq-Ul-Hoque, **University of Liberal Arts Bangladesh**
Joyonti Rani Karmakar , **University of Liberal Arts Bangladesh**
Date: January 30, 2025

Dedicated to
my loving Mother
– Md Yusuf Hossain

Dedicated to
my loving Mother
– Rezowana Zayed Zahin

Dedicated to
my loving Mother
– Rashed-Ul-Hoque

Dedicated to
my loving Mother
– Joyonto Rani Karmakar

ABSTRACT

Object detection in low-light conditions is crucial where reliable performance under adverse lighting is essential. However, it presents significant challenges due to reduced visibility which hinder the extraction of meaningful features and degrade detection performance. To address these challenges, this study introduces an enhanced YOLOv8n model that integrates Efficient Channel Attention (ECA) with Global Context (GC). The GC module effectively captures long-range dependencies and models global contextual information, ensuring the model can focus on spatially significant regions within the image. In addition, the ECA module is incorporated to enhance channel-wise attention. This module selectively prioritizes the most relevant channels while maintaining computational efficiency by focusing on critical features even in low-contrast scenarios without introducing significant overhead. The improved YOLOv8n model is evaluated on the ExDark dataset, a well-established benchmark for low-light object detection tasks. Experimental results reveal that our proposed Y- ZA YOLO model achieves Mean Average Precision (mAP)@0.5 of 0.404 and detection accuracy compared to the mAP of 0.360 of the baseline YOLOv8n model. That signifies an improvement of 4.4%. Leveraging Global Context (GC) with Efficient Channel Attention (ECA) improves the robustness of the model YOLOv8n, enabling effective handling of low-light conditions in applications like surveillance, nighttime navigation, and autonomous driving.

Keywords: YOLOv8n, Global Context, Efficient Channel Attention, Object-Detection, Y- ZA YOLO

Contents

1	Introduction	1
1.1	Problem statement	2
1.2	Aims, objectives and Motivation	2
1.2.1	Aims	2
1.2.2	Objective	2
2	Literature Review	4
2.1	Literature Analysis	10
3	Methodology	12
3.1	Data Collection	13
3.1.1	Data Distribution	13
3.2	Data Preprocessing	15
3.2.1	Image Resizing	15
3.2.2	Data Split	15
3.3	YOLOv8n Model	16
3.3.1	Proposed model	20
3.4	Object Detection Web Application	26
3.5	Budget Analysis	29
4	Results and Analysis	30
4.1	Individual Model Performance	30
4.1.1	Results Analysis	30
4.1.1.1	Confusion Matrix	30
4.1.1.2	Precision-Confidence Curve	34
4.1.1.3	Recall-Confidence Curve	37
4.1.1.4	F1 Confidence Curve	40

4.1.1.5	Precision-Recall Curve	44
5	Conclusion	48
5.1	Social, Legal, Ethical, and Environmental Issues	48
5.2	Brief Summary	49
5.3	Future works	50
	References	52

List of Figures

3.1	Overview of Methodology	12
3.2	Pie Chart of Exdark Data Distribution	14
3.3	Bar Chart of Exdark Data Distribution	15
3.4	Adapted from "Improved Steel Surface Defect Detection Algorithm Based on YOLOv8"	17
3.5	Proposed Y-ZA YOLOv8 Model	21
3.6	Improved Version of Neck in YOLOv8	22
3.7	Architecture of ECA (Efficient Channel Attention)	23
3.8	Architecture of Global Context	25
3.9	Interface of Object Detection app	27
3.10	Low Light Image for Detection	28
3.11	Detected Objects (People & Bicycle)	28
4.1	Confusion matrix for YOLOv8	32
4.2	Confusion matrix for Y-ZA YOLOv8	32
4.3	Precision-Confidence Curve for YOLOv8	36
4.4	Precision-Confidence Curve for Y-ZA YOLOv8	36
4.5	Recall-Confidence Curve for YOLOv8	39
4.6	Recall-Confidence Curve for Y-ZA YOLOv8	39
4.7	F1 Confidence Curve for YOLOv8	42
4.8	F1 Confidence Curve for Y-ZA YOLOv8	42
4.9	Precision Recall Curve for YOLOv8	45
4.10	Precision-Recall Curve for Y-ZA YOLOv8	45

List of Tables

2.1	Analysis for GC, ECA and ECA+GC.	10
3.1	Analysis of Project Budget	29
4.1	Analysis of Performance for Epochs 10	47
4.2	Analysis of Performance for Epochs 50	47

Chapter 1

Introduction

In today's world, technology plays a vital role in addressing complex issues, including the challenge of object detection in low-light conditions. Low-light conditions often conceal important visual details, presenting major challenges for conventional detection models. Effective object detection plays a critical role in diverse applications, such as autonomous navigation, surveillance, and medical imaging.

This project presents a well-structured model aimed at addressing the complexities of low-light object detection through the integration of Efficient Channel Attention (ECA) with Global Context (GC). The GC module efficiently captures global contextual relationships, modeling long-range dependencies across feature maps to preserve critical information. The global context block is used to enhance our model. In sequence transformation, a set of key features for mapping to the output are connected to the input and output sequence using global dependence in a self-attention mechanism J.-Y. Ding et al. (2024).

Simultaneously, the ECA mechanism enhances local cross-channel interactions without dimensionality reduction, ensuring precise and efficient feature extraction. Efficient Channel Attention (ECA) is integrated into the neck to enhance channel-wise feature extraction while maintaining computational efficiency Pan et al. (2024). Together, these mechanisms provide a reliable and optimized solution to the challenges faced by low-light environments.

Low-light object detection is not constrained by geographical or situational boundaries as it has global importance. Creating significant challenges in efficiency and

accuracy, our project seeks to address these issues by designing a user-friendly, high-performance detection model capable of functioning effectively even in visually degraded environments. So to provide accessible solutions technology should overcome its limits. By integrating ECA and GC modules, we aim to make object detection efficient, scalable, and adaptable to diverse applications by showcasing the power of innovation in solving real-world challenges such as public safety, industrial automation, surveillance and night vision.

1.1 Problem statement

YOLOv8 is an advanced object detection model, capable of addressing challenges like low-light conditions. However, the detection accuracy of this model still requires improvement Kong & You (2024) highlighting the need for further investigation, including the integration of Efficient Channel Attention (ECA) with a spatial attention module Q. Wang et al. (2020).

1.2 Aims, objectives and Motivation

The development of YOLOv8 is guided by well-defined and impactful aims, highlighting our dedication to enhancing object detection performance in low-light conditions while responsibly utilizing advanced attention mechanisms.

1.2.1 Aims

To improve the YOLOv8 model for object detection in low-light conditions by addressing visibility challenges and integrating advanced attention mechanisms like Efficient Channel Attention (ECA) and Global Context (GC).

1.2.2 Objective

- To improve the YOLOv8 model to improve object detection accuracy and reliability in low-light conditions by addressing visibility challenges.

- To incorporate advanced attention mechanisms such as ⁹Efficient Channel Attention (ECA) and Global Context (GC) to enhance feature extraction and emphasize critical information.
- To achieve better performance and deployment in real-world applications.

Chapter 2

Literature Review

A comprehensive study, enhancements to the YOLOv8 algorithm to improve steel surface detection has explored improvements that include the integration of the SimAM attention-free mechanism to enhance feature extraction efficiency and the optimization of the SPPF module to expand the receptive field, enabling better detection of small and large objects. Experimental results demonstrated significant increases in precision and recall, indicating the effectiveness of the improved YOLOv8 in addressing challenges related to real-time industrial defect detection. In future, there is still room for improvement in the detection accuracy of the improved YOLOv8 algorithm, necessitating further investigation Kong & You (2024).

An article conducted by Q. Wang et al. (2020) introduces a novel channel attention mechanism that aims at improving deep convolutional neural network (CNNs) performance while reducing computational complexity. Unlike SE-Net, ECA-Net removes fully connected layers and uses 1D convolution to enable local cross-channel interaction, preserving essential feature information without dimensionality reduction. The core advantages of ECA-Net include its lightweight nature, scalability, and computational efficiency, making it ideal for resource-constrained applications. The model demonstrates superior performance across tasks such as object detection, image classification, and semantic segmentation on benchmarks like ImageNet and COCO. Its lightweight design makes it ideal for resource-constrained applications, including edge computing and mobile vision. However, ECA-Net's lack of spatial attention limits its applicability in tasks requiring detailed spatial dependencies. Future work may focus on integrating spatial attention and exploring hybrid approaches to address these limitations while retaining the computational efficiency of

ECA-Net. Overall, ECA-Net provides a scalable and effective solution for enhancing CNN capabilities in real-world applications.

A detailed study, introduces an advanced global context modeling method that integrates GC block simply a better design reducing computational costs and enhancing performance. The lightweight property allows us to apply it to multiple layers in a backbone network to construct a global context network (GCNet). It achieves significant improvements in tasks like object detection, segmentation, image classification, and action recognition on benchmarks such as COCO, ImageNet, and Kinetics with minimal computational overhead. The model's reliance on pre-trained backbones like ResNet presents a limitation, potentially affecting adaptability in custom environments. Future work proposes further optimization and integration into other architectures for large-scale applications. Overall, GCNet offers practical scalability and notable advances in global context modeling. The attention mechanism plays a pivotal role in the field of deep learning and has achieved magnificent inventions in computer vision. They are used in object detection, image classification and segmentation to achieve great performance Cao et al. (2019).

A study conducted by He et al. (2023), proposes the ShuffDet backbone replacing the CSPDarkNet53 network to reduce complexity significantly and enhance feature extraction by implementing ECA (Efficient Channel Attention) with path aggregation feature pyramid network (PAFPN). Experimental results on the KITTI dataset demonstrate that ShuffYOLOX achieves a 65% increase in detection speed and a 34.57% reduction in parameter count compared to YOLOX, with a minor drop in accuracy. This makes it suitable for autonomous driving. The model also performs well in recognizing small objects and maintaining high detection precision, which are critical for safety in intelligent driving scenarios.

In another study, the integration of the Global Context (GC) block into YOLOv8 has enhanced its feature extraction and global information modeling capabilities through by introducing the YOLOv8+GC model. This advancement proved effective in reducing misdiagnosis rates in pediatric wrist fracture detection and achieved state-of-the-art performance. Future work could focus on addressing data imbalance, expanding the dataset variety and adapting the YOLOv8+GC model for

broader applications beyond pediatric wrist fractures to enhance generalizability Ju et al. (2024).

In a comprehensive study, the integration of the Global Context (GC) block in object detection models, such as GC-YOLOv3, enhances feature representation by modeling long-range dependencies in spatial and channel dimensions. By assigning weights to features based on relevance, the GC block improves detection accuracy without significantly increasing model parameters. GC-YOLOv3, which incorporates the GC block and learnable semantic fusion, achieves superior performance over YOLOv3. Future work may explore optimizing the GC block for scalability and adaptability to different network layers and applications Yang & Deng (2020).

In another study an improved SSD (Single Shot MultiBox Detector) model is introduced to enhance the efficiency and accuracy of robotic sorting and assembly in industrial settings. By replacing the VGG16 backbone with ResNet-50, the algorithm improves feature extraction and convergence speed. Additionally, it incorporates ECA-Net (Efficient Channel Attention) and ISAB (Improved Spatial Attention Block) attention mechanisms to optimize feature learning and improve detection performance. Evaluated on a specially constructed dataset of aliased multicategory hardware, the proposed model achieves a mean average precision (mAP) of 98.20% and a processing speed of 78 FPS, surpassing existing algorithms like Faster R-CNN, YOLOv4, and EfficientDet in accuracy and real-time efficiency. Despite its strengths, the paper highlights challenges such as labor-intensive dataset labeling and reliance on the backbone network, suggesting future research to address these limitations and extend its applicability Zhang et al. (2022).

An article by Chen et al. (2024) introduces an enhanced YOLOv8 model, CES-YOLOv8, to improve the detection of strawberry ripeness in complex agricultural environments. The study addresses challenges such as occlusion, varying lighting, and irregular fruit distribution by incorporating three key improvements: replacing parts of the YOLOv8 backbone with the ConvNeXt V2 module for better feature extraction, introducing the ECA (Efficient Channel Attention) mechanism to enhance feature focus, and using angle and distance compensation in the SIoU loss function to improve localization accuracy. Training on a dataset of augmented strawberry images captured under diverse conditions, the CES-YOLOv8 model

achieved notable improvements over the original YOLOv8, with 88.2% precision, 89.8% recall, a mean Average Precision (mAP50) of 92.1%, and an F1 score of 88.99%. These advancements enhance the model's ability to detect strawberries at various maturity levels accurately, even in complex environments, making it a valuable tool for automated harvesting. While the model demonstrates strong adaptability and performance, the study highlights the need for further research to extend its application to other crops and diverse agricultural scenarios.

In another paper, Yang et al. (2024) an improved lightweight model, FE-YOLO is introduced to enhance construction waste detection in complex environments. The model modifies the YOLOv8 architecture by replacing the C2f module with a Faster C2f module, which reduces computational complexity and parameters while maintaining high performance. Additionally, the Efficient Channel Attention (ECA) mechanism is integrated to improve feature extraction in the bottleneck layer. These enhancements were tested on a custom construction waste dataset, achieving 92.7%, a 3% improvement over YOLOv8n, along with a 12% reduction in parameters and a 13% decrease in computational load. This demonstrates the model's efficiency and accuracy, making it well-suited for real-time applications in construction waste management. Future research will focus on further generalization to other waste types and datasets.

In another study, enhancements are aimed to the YOLOv8n model at improving the detection of foreign objects, such as bird nests, kites, and balloons, on power transmission lines. These improvements address challenges like detecting small objects and maintaining high accuracy in real-time conditions. The study incorporates the Efficient Channel Attention (ECA) mechanism to strengthen feature extraction and introduces a small object detection module to better capture small target details. Experimental results show that the improved model achieves a mean Average Precision (mAP@50) of 93.98%, with higher recall and precision rates compared to the baseline YOLOv8n. The enhanced model balances accuracy and speed, making it suitable for deployment in real-world power line inspection scenarios. Future work includes further refinements to extend its applicability across different conditions and object types H. Wang et al. (2024).

In an article conducted by Zheng et al. (2024) presents an improved YOLOv8n-based model for detecting foreign object debris (FOD) on runways at airport. Key enhancements include a small-object detection layer, GhostConv modules for computational efficiency, the DWR module for better feature extraction, and the Inner-Weighted MPDIoUv3 loss function for improved regression accuracy. The model achieved a real-time detection speed of 42 FPS on embedded devices, making it highly efficient for practical applications. However, the study is limited by its reliance on a custom dataset (FOD-Z) and requires further integration with autonomous systems, such as drones or smart vehicles, for broader real-world deployment.

Another study addresses the issue of liquid retention on pipette tips, a critical factor in preventing cross-contamination and ensuring accurate experimental results in high-throughput liquid handling. The authors enhanced the YOLOv8 algorithm by integrating the GCNet and LSK attention modules into the backbone and C2f modules, improving the model's ability to capture global context and dynamically adjust to different backgrounds. Additionally, the SimAM module is incorporated into the neck network to enhance feature focus. Tested on a custom pipette tip dataset, the model achieved a 1.7% improvement in mAP@0.5, a 2% increase in the F1 score, and a 1.7% boost in precision compared to the baseline YOLOv8. While demonstrating superior detection capabilities for small and transparent liquids, the model faces limitations in computational efficiency, highlighting the need for further optimization for deployment in resource-constrained environments Yin et al. (2024).

C. Wang et al. (2024) propose an improved YOLOv8+ model for detecting and classifying strawberry ripeness to support efficient and intelligent harvesting. Enhancements include integrating an Efficient Channel Attention (ECA) module for improved feature extraction and adopting the Focal-EIOU loss function to address imbalances in classification. The system combines deep learning and image processing to determine ripeness levels—unripe, ripe, and fully ripe—by analyzing the red pixel percentage on strawberry centerlines. The model achieved 97.81% accuracy and an F1 score of 97.07%, demonstrating robust detection and classification in real-time agricultural environments. Limitations include the dependency on a

specific dataset and environmental conditions, which require further optimization for broader applicability.

An article introduces TIR-YOLO-ADAS, an enhanced YOLOv8-based framework for object detection in thermal infrared images, tailored for advanced driver assistance systems (ADAS). Key features include integrating Part of the Attention Mechanism (PAM) to improve feature representation and applying a focal loss function to reduce missed detections. Tested on the FLIR Thermal ADAS dataset, the framework improved mean Average Precision (mAP) from 32.24% to 34.42% and demonstrated superior detection of small and poorly visible objects. While promising for low-visibility conditions, challenges remain in addressing the limitations of thermal infrared imaging, such as low resolution and limited texture information, which may hinder broader application M. Ding et al. (2024).

Li & Shi (2024) propose YOLO-GE, a novel attention-enhanced underwater object detection model based on YOLOv8s. The model incorporates image enhancement to address poor underwater imaging conditions, adds high-resolution feature layers for small object detection, and integrates a GEBlock attention fusion module to suppress noise and capture long-range contextual information. Tested on datasets like UTDAC2020, YOLO-GE achieved significant improvements in detection accuracy for underwater environments, overcoming challenges like image blur, low contrast, and diverse object scales. Despite its high accuracy, the model's complexity and resource demands limit its real-time application.

In another study, the improved YOLOv8 model was designed for the detection of pediatric wrist fractures using X-ray images. The model integrates various attention modules, such as ResCBAM and ResGAM, to enhance feature extraction and accuracy. Evaluated on the GRAZPEDWRI-DX dataset, the YOLOv8-AM achieved a state-of-the-art mAP@0.5 of 65.8%, significantly improving over the baseline YOLOv8. However, limitations include the relatively modest enhancement offered by some attention modules, highlighting the need for further refinements in clinical applications Chien et al. (2024).

Yuan et al. (2024) present the YOLOv8-ACU model that enhances the version of YOLOv8-pose tailored for facial acupoint detection in Traditional Chinese Medicine.

The model incorporates the ECA attention mechanism for refined feature extraction, replaces the neck module with a lightweight Slim-neck, and uses the GIoU loss function for improved localization accuracy. The YOLOv8-ACU achieves a 9.3% reduction in GFLOPs, demonstrating its efficiency and suitability for mobile and embedded applications. However, its reliance on specific datasets limits its adaptability to broader clinical scenarios.

2.1 Literature Analysis

This table (2.1) categorizes research contributions into three focus areas: Global Context (GC), Efficient Channel Attention (ECA) and their combination ECA+GC.

Authors	GC	ECA	ECA+GC
Yang and Deng (2020)	✓	✗	✗
Wang et al. (2020)	✗	✓	✗
Cao et al. (2020)	✓	✗	✗
He et al. (2023)	✗	✓	✗
Ju et al. (2024)	✓	✗	✗
You and Kong (2024)	✗	✗	✗
Zhang et al. (2022)	✗	✓	✗
Chen et al. (2024)	✗	✓	✗
Yang et al. (2024)	✗	✓	✗
H. Wang et al. (2024)	✗	✓	✗
Zheng et al. (2024)	✗	✓	✗
Yin et al. (2024)	✓	✗	✗
C. Wang et al. (2024)	✗	✓	✗
M. Ding et al. (2024)	✗	✓	✗
Li and Shi (2024)	✓	✗	✗
Chien et al. (2024)	✗	✓	✗
Yuan et al. (2024)	✗	✓	✗

Table 2.1: Analysis for GC, ECA and ECA+GC.

Studies like Yang and Deng (2020), Cao et al. (2020), Ju et al. (2024), Yin et al. (2024) and Li and Shi (2020) highlight the critical role of Global Context (GC) in delivering a global perspective. According to these papers, GC provides a comprehensive understanding of data, enhances feature extraction, and significantly improves the overall effectiveness of machine learning models by ensuring robust and scalable solutions.

Efficient Channel Attention (ECA) is utilized in studies such as Yang et al. (2024), C. Wang et al. (2024), Chien et al. (2024), Yuan et al. (2024), You and Kong (2024), Zhang et al. (2022), He et al. (2023), M. Ding et al. (2024), H. Wang et al. (2024), Zheng et al. (2024), Chen et al. (2024), Chien et al. (2024) which enhance feature refinement by focusing on channel-specific relationships. According to these papers, ECA is particularly valuable in tasks requiring efficient attention mechanisms, as it ensures precise feature prioritization without significant computational overhead.

The table emphasizes the extensive use of Global Context (GC) as a critical foundational approach for capturing broad, high-level feature representations. Meanwhile, Efficient Channel Attention (ECA) is increasingly recognized for its computational efficiency and scalability. The combination of ECA+GC shows significant promise in creating a balanced framework that integrates global and local feature extraction techniques.

Chapter 3

Methodology

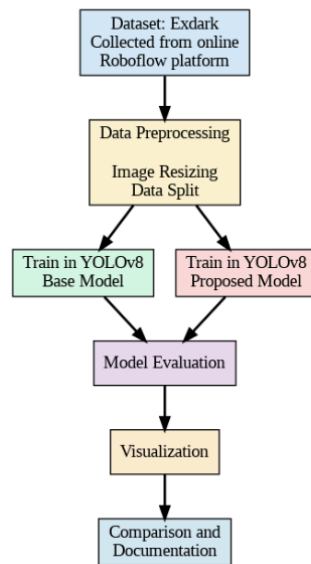


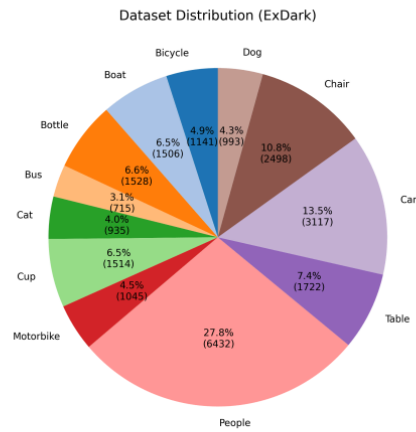
Figure 3.1: Overview of Methodology

3.1 Data Collection

In our study, we collected the Roboflow "ExDark" dataset, a reliable platform for accessing high-quality datasets. This dataset has 12 classes Bicycle, Boat, Table, People, Bottle, Dog, Motorbike, Cup, Cat, Car, Chair, Bus. The dataset was carefully selected to align with the requirements of our task, as it contains diverse images captured under extreme low-light conditions. Using the "ExDark" dataset, we ensured a rich variety of environmental scenarios and object categories, which are crucial for training and evaluating our model effectively. This ensured that our improved YOLOv8 model, with ECA and GC modules, is trained on data representative of real-world low-light scenarios, establishing a solid foundation for achieving robust and accurate object detection.

3.1.1 Data Distribution

The "ExDark" dataset is made up of a diverse range of object classes. The ExDark dataset is distributed across 12 object classes, providing a diverse set of images for robust object detection in low-light environments. The Dog class has the highest representation with 805 images, followed by Cat with 733 images and Boat with 679 images. Other well-represented classes include bicycle (650 images), chair (648 images) and car (637 images), ensuring significant coverage of everyday objects.



65
Figure 3.2: Pie Chart of Exdark Data Distribution

The People class comprises 609 images, while smaller objects such as Bottle (547 images) and Cup (519 images) add diversity to the dataset. Larger objects, such as Bus (526 images), and more compact items, such as Table (504 images) and Motorbike (502 images), contribute to the overall balance. This distribution ensures a comprehensive dataset for training and evaluating object detection models across a variety of object types and scenarios.

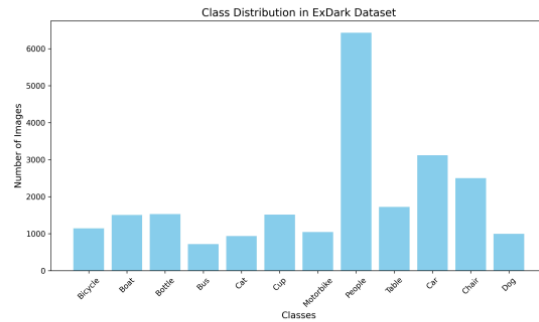


Figure 3.3: Bar Chart of Exdark Data Distribution

51

3.2 Data Preprocessing

Data preprocessing is an essential step to ensure that the dataset is well-structured, accurate, and ready for analysis. It involves addressing missing or incorrect data, modifying data types as needed, and preparing the dataset for use in machine learning models.

3.2.1 Image Resizing

The "ExDark" dataset has been obtained from Roboflow, a trusted platform for high-quality machine learning datasets. It features a diverse collection of images captured under extreme low-light conditions, making it perfectly suited to our research objectives. For our study, the images were resized to 640x640 pixels to align with the requirements of the YOLOv8n model.

3.2.2 Data Split

We partitioned the "ExDark" dataset into distinct training, validation, and testing subsets to ensure a robust evaluation of the model's performance. Adopting a 70:10:20 split ratio, 70% of the data was allocated for model training, 10% for

testing and the remaining 20% for validation. This strategic partitioning preserved a substantial training dataset while allowing for comprehensive validation and testing on unseen data. By carefully partitioning the dataset, we aimed to improve its reliability and diversity, creating an ideal setting for training and assessing our low-light object detection model. This approach facilitated efficient learning and ensured an accurate performance evaluation in challenging low-light conditions.

3.3 YOLOv8n Model

YOLOv8n, the lightweight variant of the YOLOv8 family, excels in object detection with its efficient architecture. In low-light conditions, its advanced feature extraction capabilities and adaptability enable it to identify objects with reduced visibility. Its speed and compact design make YOLOv8n a suitable choice for scenarios requiring quick and accurate detection in challenging lighting environments. Its compact design allows for faster inference speeds while maintaining reliable accuracy, making it an ideal choice for low-light object detection applications. YOLOv8n state-of-the-art performance and adaptability ensured a strong foundation for building a robust detection system. This model is the most lightweight and rapid in the YOLOv8 series, designed for environments with limited computational resources. So, we have selected YOLOv8n as the base model for its lightweight architecture and high efficiency in object detection tasks.

In YOLOv8n architecture there are three sections: Backbone, Neck and Head. In the backbone, a modified version of CSPDarknet 53 network is used Sohan et al. (2024). The Backbone is a deep learning architecture which acts as a feature extractor of the inputted image.

At Neck combines multi-scale features from the Backbone to enhance the detection capability of objects at different scales integrates features extracted from multiple layers of the Backbone module. Then Head generates the final output of the object detection model by predicting object classes and their corresponding bounding boxes.

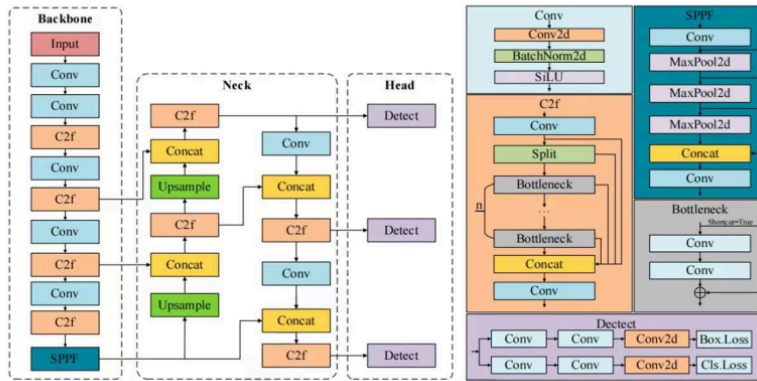


Figure 3.4: Adapted from "Improved Steel Surface Defect Detection Algorithm Based on YOLOv8"

The detailed discussion of the architecture are given below:

Backbone: It gradually reduces the spatial resolution of the input image while increasing the depth (number of feature channels) to capture more abstract representations of the image Timilsina (2024).

The details of Backbone are given below:

- **Input image:** The image dimensions are 512×512×3. The spatial resolution of an image is 512X512. The number of RGB is 3. Then the image is processed to the convolutional layer.
- **Convolutional Layer (Conv Layer) :** The fundamental building block of the architecture, comprising a Conv2d layer, a BatchNorm2d layer, and a SiLU activation function. Performs a 2D convolution operation on the input feature maps to extract spatial features. It helps in feature learning by applying filters to the input image or feature map Timilsina (2024).

The Convolutional layer has kernel size 3, stride 2 and padding 1. Here, k represents the number of filters or kernels, which determines the depth of the output feature map. Each filter is designed to capture distinct features

from the input, representing Stride refers to the step size at which the filter or kernel moves across the input. A larger stride decreases the spatial dimensions of the output feature map and p represents Padding is the process of adding a border of zeros around the input. It helps retain spatial information and allows for control over the dimensions of the output feature map Timilsina (2024). After the Conv2d layer, it was processed to the BatchNorm2d layer. Batch normalization is a method used in deep neural networks to enhance training stability and accelerate convergence. In convolutional neural networks (CNNs), the BatchNorm2d layer normalizes 2D inputs, which are usually the output of convolutional layers. It ensures that the values flowing through the network remain within a reasonable range, preventing problems during training. Finally, SiLU (the Sigmoid Linear Unit), also known as the Swish activation function, is frequently utilized in neural networks. The SiLU activation function is mathematically defined as:

$$SiLU(x) = x \cdot \sigma(x) \quad (3.1)$$

where $\sigma(x)$ is the sigmoid function, represented as:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3.2)$$

This function combines the input x with its sigmoid output to produce smooth and non-linear activation, which is effective for deep learning tasks. After completing the operations of the Conv layer, the process moves to the next Conv layer which follows the same mechanism. The output from this layer is then passed to the C2f layer for transformation.

- **C2f Layer:** At C2f layer, split the feature map into two parts: one feature that goes to the bottleneck layer and other goes directly to the concat layer. In the C2f layer, there are two parameter shortcuts and n Timilsina (2024).

The shortcut parameter is a true or false setting that shows whether the Bottleneck block inside the C2f block uses a shortcut connection. If shortcut=True, the Bottleneck block uses the shortcut; if shortcut=False, it does not. Here, n determines how many bottleneck blocks are used inside the C2f block. At the

end, the feature map from the bottleneck block and the split feature map are concatenated and inputted into a final convolutional block. In the C2f block the resolution of the feature map and the output channel is unchanged. Then it processes to the next Conv layer and this remains repeated until moving to the SPPF.

- **SPPF layer (Spatial Pyramid Pooling Fast layer):** The SPPF module is the improved version of SPP (Spatial Pyramid Pooling) Ju & Cai (2023). The SPPF Block consists of a convolutional layer followed by three MaxPool2d layers. The feature maps produced by each MaxPool2d layer are concatenated at the end and passed through another convolutional layer for further processing. The concept of Spatial Pyramid Pooling (SPP) involves dividing the input image into grids and pooling features independently from each grid cell. This approach allows the network to handle images of varying sizes effectively, capturing multi-scale information and making it useful for tasks like object recognition, where objects appear at different scales within an image Timilsina (2024). SPP-Fast simplifies this process to reduce computational demands Ju & Cai (2023). Instead of pooling features at multiple levels using different kernel sizes, it uses a single fixed-size pooling kernel. This adjustment reduces the computational complexity while providing a balance between speed and accuracy. The MaxPool2d layer plays a key role in reducing the spatial dimensions of the input tensor, lowering computational requirements and focusing on the most significant features Timilsina (2024). Max pooling retains only the maximum value from each region of the input tensor, discarding other values. In the MaxPool2d layer, this pooling operation is applied across both the height and width dimensions of the input. Then for further processing it passes to the upsample layer of the neck section.

Neck: The neck section obtains more feature information in better dense prediction. Due to too many convolutional networks, information will be lost. So a architecture is used called Feature Pyramid Network (FPN) and Path Aggregation Network (PAN). They are responsible for multi feature fusion to combine features from different layers of the network. Feature Pyramid Network handles upsampling the feature maps from top to bottom and Path Aggregation Network obtain the features

from bottom to top Ju & Cai (2023). Therefore, to guarantee accurate predictions for pictures of different sizes, these two feature outputs are combined.

¹⁹
Head: The final output generated by the object detection model is the prediction of the classes and bounding box of the objects by the head section. As an anchor-free model, YOLOv8n predicts an object's centre directly rather than using a known anchor box to determine its offset Ju & Cai (2023). Anchor-free detection speeds up complex post-processing operations that sort through candidate detections following inference by reducing the number of box predictions. There are two tracks in the Detect Block Timilsina (2024).

Bounding box predictions are made on the first track, while class predictions are made on the second. Two convolutional blocks and a single Conv2d layer are present in both tracks, providing the Bounding Box loss and Class Loss, respectively. There are three detect layer, first detect layer in the Head section specializes in detecting small objects that are inputted from the C2f layer in the Neck then second one in the Head section specializes in detecting medium-sized objects which is inputted from the C2f layer in the Neck and third in the Head section specializes in detecting objects that are inputted from the C2f layer Neck.

In summary, YOLOv8, the latest advancement in the YOLO series, sets a new standard for object detection with its anchor-free design, combining speed and accuracy across its variants. Utilizing convolutional and bottleneck layers with Spatial Pyramid Pooling Fast, it processes images efficiently for detection. Its backbone, neck, and head sections collaborate to extract features, upsample, and predict object classes and bounding boxes. With versatile models catering to different needs, YOLOv8 excels in performance and usability, making it a benchmark in object detection.

3.3.1 Proposed model

Y- ZA YOLOv8n Model

YOLOv8n has exhibited excellent performance in object detection with high speed, accuracy and efficiency. YOLOv8n is lightweight architecture which makes it more efficient to deploy and maintain, which is beneficial for both research and industry

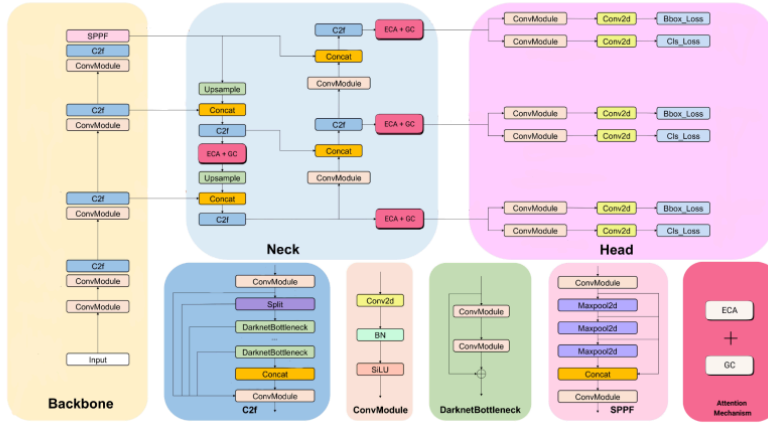


Figure 3.5: Proposed Y-ZA YOLOv8 Model

applications. Inspired by the significant contributions of YOLOv8n, we have proposed a Y-ZA YOLOv8n model to improve the ability to detect objects in low-light environments. Moreover, the Y-ZA YOLOv8n model excels at object detection under complex lighting conditions, offering improved accuracy and efficiency. We have integrated ECA (Efficient Channel Attention) and GC (Global Context) as attention modules after each of the four C2f layers in the Neck. The GC module captures long-range dependencies and relationships between all parts of the image. It enhances the contextual understanding of the entire scene. In low-light conditions, where objects often lack sharp edges or distinct local features, GC compensates by aggregating global spatial information and encoding scene-level contextual relationships (e.g., a cup on a table) to facilitate object recognition. Furthermore, the GC block integrates multi-scale features, enhancing the model's ability to analyze image content across different scales and resolutions. This functionality allows the model to effectively capture both fine details and broader contextual information within the image.

Also, the ECA module processes the feature map by analyzing the importance of each channel and prioritizing those that capture subtle edges, textures, or faint

patterns by improving the model's ability to detect objects in low-light images without adding significant computational overhead. Moreover, ECA avoids dimensionality reduction, preserving the original details of the feature map. This is essential for low-light object detection, as even small losses in feature details can lead to missed detections. Low-light conditions can cause objects to blend into the background or overlap visually. ECA helps distinguish these objects by enhancing the channels that highlight unique or distinguishing features. Integrating together GC and ECA provides a synergistic effect, with GC enhancing spatial awareness and ECA optimizing channel-wise feature importance. This dual improvement allows YOLOv8n to achieve state-of-the-art accuracy while remaining efficient and lightweight, making it an ideal choice for object detection in low light. By incorporating these mechanisms, YOLOv8n achieves a remarkable balance of precision, speed, and resource utilization, setting a new benchmark for object detection models.

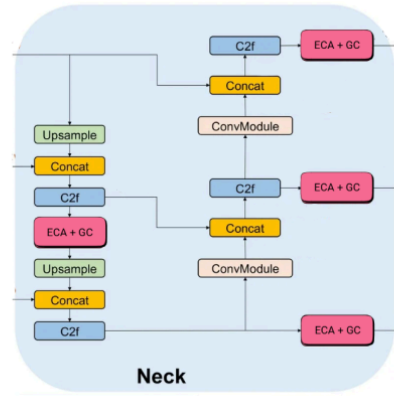


Figure 3.6: Improved Version of Neck in YOLOv8

In our proposed model we keep the network same as YOLOv8n. We have integrated ECA (Efficient Channel Attention) with GC (Global Context) after each of the four

C2f layers as attention modules in the Neck. Making a new innovative architecture for object detection in complex lighting conditions. The detailed architectures of the GC (Global Context) and ECA (Efficient Channel Attention) are discussed below.

ECA (Efficient Channel Attention) :

The Efficient Channel Attention (ECA) block in YOLOv8 is designed to improve feature map representation by emphasizing important channels while keeping the computational overhead minimal. Unlike traditional attention mechanisms that use fully connected layers for channel weighting, ECA employs a lightweight approach based on a 1D convolution with a carefully chosen kernel size Chen et al. (2024). The process begins with a global average pooling operation to aggregate spatial information for each channel, resulting in a global descriptor that summarizes the feature map. This descriptor is then passed through the 1D convolution, which efficiently captures cross-channel interactions and generates channel attention weights without introducing excessive parameters or complexity Chen et al. (2024). The computed attention weights are applied to the original feature map through element-wise multiplication, enhancing the most relevant channels while suppressing the less significant ones. This mechanism ensures that YOLOv8 effectively focuses on the critical features for object detection, balancing accuracy and computational efficiency.

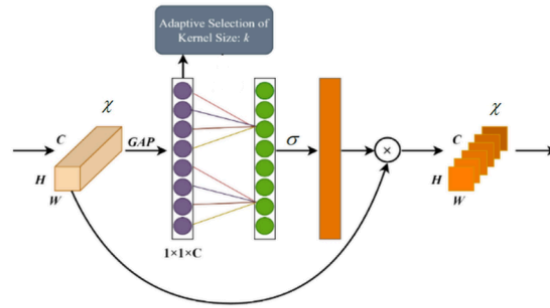


Figure 3.7: Architecture of ECA (Efficient Channel Attention)

In the Neck, the Efficient Channel Attention (ECA) mechanism plays a critical role in refining feature maps by emphasizing the most important channels. The process begins with Global Average Pooling (GAP), which compresses each channel's spatial dimensions into a single global descriptor, effectively summarizing the spatial information. This descriptor is then passed through a lightweight 1D convolution with a carefully chosen kernel size(k), followed by an activation function called sigmoid to capture dependencies between adjacent channels without introducing excessive parameters Yuan et al. (2024). The output of this operation is a set of attention weights that quantify the relative importance of each channel.

These attention weights are applied to the original feature map via element-wise multiplication, scaling each channel according to its relevance. By incorporating ECA, the Neck produces enriched multi-scale feature maps (e.g., 80x80, 40x40, 20x20), enabling the detection of objects of various sizes with improved accuracy. Its lightweight design ensures computational efficiency while improving the network's ability to focus on critical features, making it a key component for effective object detection.

GC (Global Context) : The GC block is a component of the neural network designed to enhance the ability of the model to understand and process global information. By capturing features from the entire image, this block strengthens the model's ability to comprehend the overall content of the image. Additionally, the GC block integrates multi-scale features, enhancing the model's capability to analyze image content across different scales and resolutions Yang & Deng (2020). So by incorporating a GC block the object can be detected effectively.

Here, the Global Context (GC) module in the Neck is designed to enhance feature maps by integrating global spatial information, making the network better equipped to handle complex object detection tasks. It begins with a global pooling operation, such as Global Average Pooling (GAP), which compresses the entire spatial region of the input feature map into a compact global descriptor, effectively summarizing the overall context of the image. This global descriptor acts as a summary of the image's spatial information and is passed through an attention mechanism to compute importance weights for each channel in the feature map.

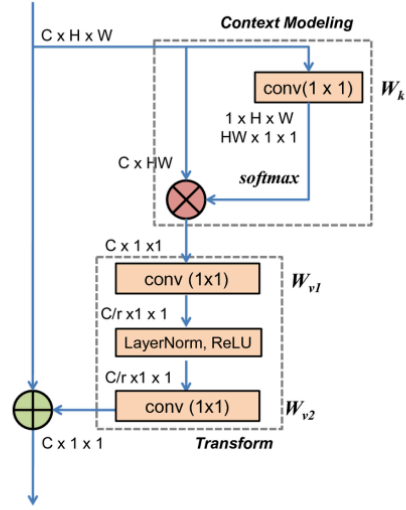


Figure 3.8: Architecture of Global Context

The attention mechanism identifies the most relevant channels containing critical global information while suppressing the less important ones. Once the global descriptor is refined, it is fused back into the original feature map through element-wise operations like addition or multiplication. This fusion step ensures that every spatial location in the feature map benefits from the global contextual information, enriching the feature map with a balance of local details (from the original feature map) and global awareness (from the GC module).

By combining these complementary features, the GC module enhances the Neck’s ability to represent objects of various scales and in diverse contexts, enabling more accurate predictions while being computationally efficient.

By combining these two mechanisms, the Neck benefits from both GC (Global Context) and ECA (Efficient Channel Attention). GC ensures that the feature maps include a high-level understanding of the entire image, while ECA enhances the

representation by prioritizing the most important channels. Together, they produce refined multi-scale feature maps that are rich in both local and global information, enabling the network to detect objects of various sizes and in diverse contexts more accurately and efficiently.

3.4 Object Detection Web Application

In this project, we develop an intuitive and interactive web application for real-time object detection using Streamlit, a powerful Python library for creating web applications. This application is designed to identify and label objects in images by leveraging the capabilities of advanced machine learning models. The application interface allows users to upload images by simply dragging and dropping them into the designated area, supporting popular formats like PNG, JPG and JPEG. Once an image is uploaded, the system processes it and highlights detected objects by drawing bounding boxes around them, complete with labels indicating the object's category (e.g., bicycle, chair, people). This functionality enables users to visualize and analyze objects in the image with ease.

The object detection model employed in this application is trained to recognize multiple classes, ensuring high accuracy in diverse scenarios. In the example displayed, the application effectively identifies multiple objects, including a person, food items, and a plant, showcasing its versatility in detecting items across varying environments. This tool is not only user-friendly but also highly practical, serving a wide range of applications such as surveillance, inventory management, and research. By integrating machine learning with a seamless web-based interface, the application ensures that users, regardless of technical expertise, can access advanced object detection technology in an accessible and efficient manner. Through the power of Streamlet, we have made object detection tools simple and accessible, allowing users to easily utilize advanced technology in their everyday tasks.

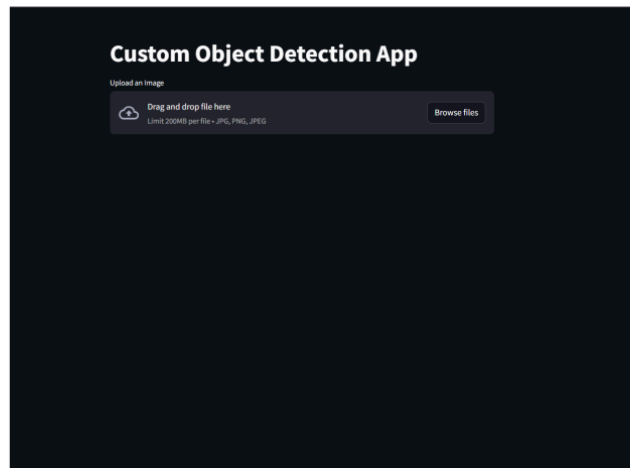


Figure 3.9: Interface of Object Detection app

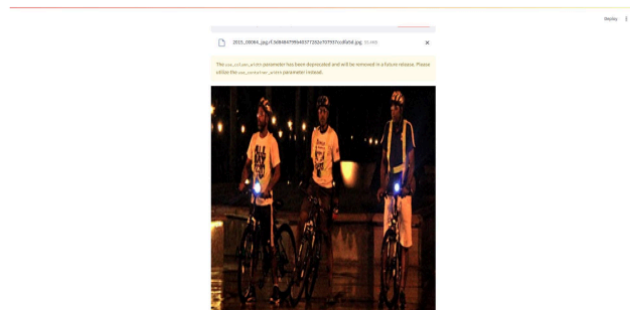


Figure 3.10: Low Light Image for Detection

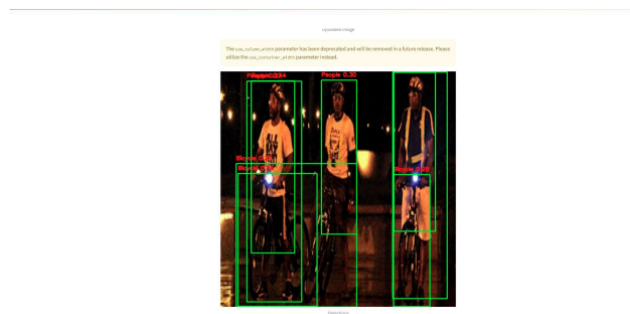


Figure 3.11: Detected Objects (People & Bicycle)

3.5 Budget Analysis

SL. No	Item	Price (BDT)
1	Google Colab Pro	3,000
2	GPU	25,000
3	RAM	10,000
4	Miscellaneous Cost	2,000
Total		40,000

Table 3.1: Analysis of Project Budget

Chapter 4

Results and Analysis

4.1 Individual Model Performance

We have compared the results between YOLOv8 and Y-ZA YOLOv8 in terms of confusion matrix, mAP@50, precision recall curve, recall confidence curve, precision confidence curve, and F1 confidence curve.

4.1.1 Results Analysis

4.1.1.1 Confusion Matrix

The confusion matrix provides a clear summary of the performance of an object detection model by highlighting accurate predictions, areas of confusion, and challenges in classifying certain objects or distinguishing background.

Metrics for the Confusion Matrix: (Figure 4.1)

True Positives (TP) : The diagonal elements of the matrix represent correctly classified instances for each class.

- People: 350 instances are correctly classified as "People."
- Car: 226 instances are correctly classified as "Car."
- Background: 823 instances are correctly classified as "Background."

False Positives (FP) : These are cases where the model incorrectly predicts a class that does not belong to the true label. For example:

- Chair: 11 cases predicted as "Chair" belong to other categories like "Table."
- Bottle: 7 cases predicted as "Bottle" actually belong to "Boat."
- People: 138 cases predicted as "People" belong to other classes.

False Negatives (FN) : These are ²³ cases where the model fails to predict a class that is present. For example:

- People: 103 instances of "People" are misclassified into other categories.
- Table: 263 instances of "Table" are misclassified into other categories.
- Dog: 96 instances of "Dog" are misclassified.

True Negatives (TN) : These represent ²³ cases where the model correctly identifies that a particular class is not present. Since this confusion matrix does not directly show TN for each class, TN values are calculated indirectly by considering all non-relevant predictions correctly excluded.

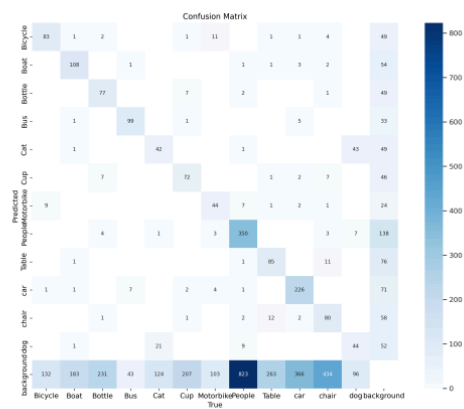


Figure 4.1: Confusion matrix for YOLOv8

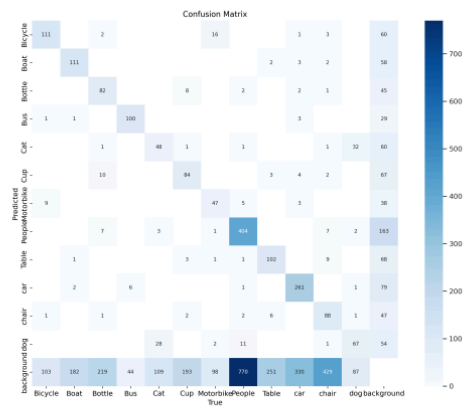


Figure 4.2: Confusion matrix for Y-ZA YOLOv8

True Positives (TP) : The diagonal elements of the matrix represent correctly classified instances for each class.(Figure 4.2)

- Bicycle: 111 instances correctly classified as "Bicycle."
- Boat: 111 instances correctly classified as "Boat."
- Bottle: 82 instances correctly classified as "Bottle."
- Bus: 100 instances correctly classified as "Bus."
- Cat: 48 instances correctly classified as "Cat."
- Cup: 84 instances correctly classified as "Cup."
- Motorbike: 47 instances correctly classified as "Motorbike."
- People: 404 instances correctly classified as "People."
- Table: 102 instances correctly classified as "Table."
- Car: 261 instances correctly classified as "Car."
- Chair: 88 instances correctly classified as "Chair."
- Dog: 67 instances correctly classified as "Dog."
- Background: 770 instances correctly classified as "Background."

False Positives (FP) : These represent cases where the model predicts a class that does not belong to the true label:

- Bicycle: 16 instances predicted as "Bicycle" belong to other categories (e.g., "Bottle").
- Boat: 3 instances predicted as "Boat" belong to other categories.
- People: 163 cases predicted as "People" belong to other classes.
- Chair: 47 cases predicted as "Chair" belong to other categories.

39

False Negatives (FN) : These represent cases where the model fails to predict the correct class:

- Bicycle: 103 instances of "Bicycle" are misclassified into other categories.
- Boat: 182 instances of "Boat" are misclassified.
- People: 193 instances of "People" are misclassified into other categories.
- Table: 251 instances of "Table" are misclassified into other categories.
- Dog: 87 instances of "Dog" are misclassified into other categories.

4.1.1.2 ³¹ Precision-Confidence Curve

The Precision-Confidence Curve illustrates the relationship between precision and confidence threshold for multiple object categories in an object detection model.

Class-Specific Precision Trends : (Figure 4.3)

- Each thin colored line represents a different object class (Bicycle, Boat, Bottle, Bus, Cat, Dog, etc.).
- ¹⁸ The thick blue line represents the overall precision trend across all classes.

Overall Model Performance:

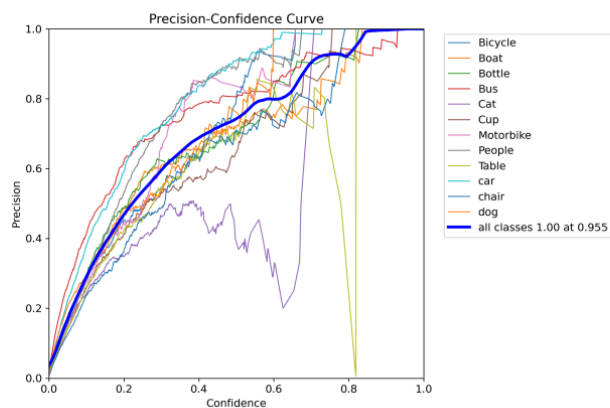
- At 0.955 confidence, the model reaches 100% precision, meaning that at this threshold, all detected objects are correctly classified.
- The precision generally increases with confidence as the model becomes more selective in its detections.
- Some classes exhibit irregular trends, with fluctuations at higher confidence levels.

Variability Across Object Classes:

- Certain classes (Bus, Dog, Bicycle) maintain high precision across a broad range of confidence values.
- Other classes (Table, Cat) show significant fluctuations and steep declines at higher confidence levels, indicating inconsistent detection performance.

Precision-Confidence Trade-off:

- Higher confidence thresholds improve precision but reduce the number of detections, potentially missing some true positives.
- Lower confidence thresholds increase the number of detections but introduce more false positives, lowering precision.



15
Figure 4.3: Precision-Confidence Curve for YOLOv8

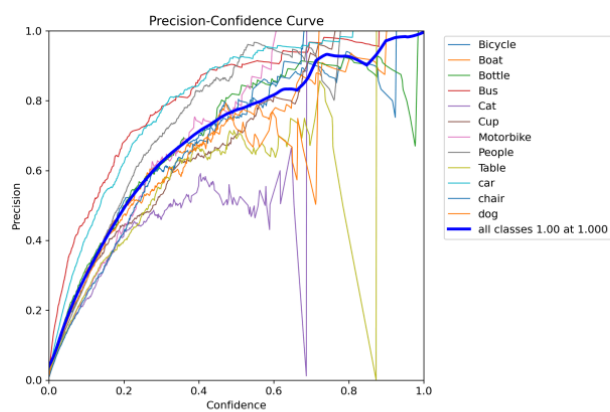


Figure 4.4: Precision-Confidence Curve for Y-ZA YOLOv8

Class-Specific Precision Trends: (Figure 4.4)

- Each thin colored line represents a different object category (Bicycle, Boat, Bottle, Bus, Cat, Dog, etc.).
- The thick blue line represents the overall model precision trend.

Overall Model Performance:

- At 1.000 confidence, the precision reaches 100%, meaning all detections at this confidence level are correct.
- Precision generally increases with confidence, as expected, but some classes exhibit irregular patterns.

Variability Across Object Classes:

- Some classes (Bus, Dog, Bicycle) achieve high precision across a wide range.
- Other classes (Table, Cat) have fluctuations and a steep drop at higher confidence levels, indicating inconsistency in predictions.

Precision-Confidence Trade-off:

- Higher confidence thresholds result in higher precision but reduce the number of detected objects.
- Lower confidence thresholds lead to lower precision due to more false positives.

⁴⁶ 4.1.1.3 Recall-Confidence Curve

The Recall-Confidence Curve visualizes how recall changes with varying confidence thresholds for multiple categories of object in an object detection model.

Class-Specific Recall Trends: (Figure 4.5)

- Each thin colored line represents a different object class (Bicycle, Boat, Bottle, Bus, Cat, Dog, etc.).

- The thick blue line shows the macro-averaged recall-confidence trend across all classes.
- Some classes, such as Bus and Dog, maintain higher recall for a broader confidence range, whereas others drop off more steeply.

Overall Model Performance:

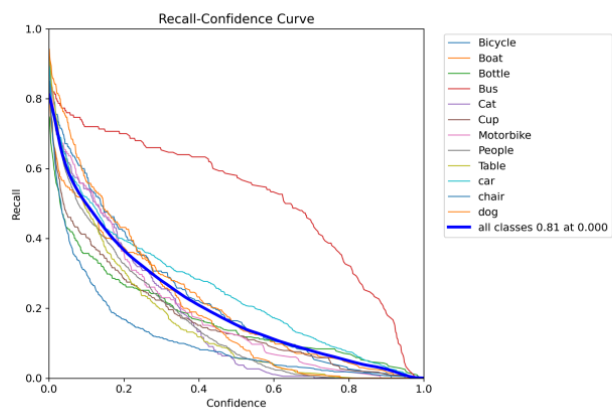
- At 0.000 confidence, the model achieves 81% recall, meaning 81% of all true objects are detected when no confidence threshold is applied.
- As the confidence threshold increases, the recall gradually decreases since only high-confidence detections are retained.

Recall-Confidence Trade-off:

- Lower Confidence : Higher Recall but More False Positives At low confidence thresholds (0.0 - 0.2), the model detects most objects (high recall), but it also includes incorrect detections (low precision).
- Higher Confidence : Lower Recall but More Precise Detections As confidence increases, only highly confident detections remain, reducing recall but increasing precision.

Performance Differences between Classes:

- Some classes (Bus, Dog) retain higher recall across different confidence values.
- Other classes (Table, Cat) show a steeper decline, meaning the model struggles to detect them consistently at high confidence.



15
Figure 4.5: Recall-Confidence Curve for YOLOv8

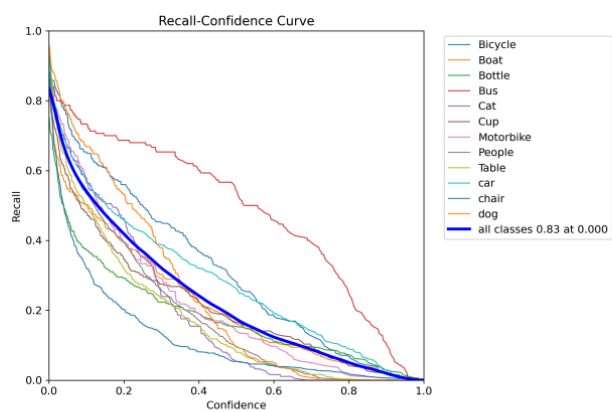


Figure 4.6: Recall-Confidence Curve for Y-ZA YOLOv8

Class-Specific Recall Trends: (Figure 4.6)

- Each thin colored line represents a different object class (Bicycle, Boat, Bottle, Bus, Cat, Dog, etc.).
- The thick blue line represents the overall model recall trend.

Overall Model Performance:

- At 0.000 confidence, recall is 83%, meaning 83% of true objects are detected without applying a confidence threshold.
- ³ As the confidence threshold increases, the recall gradually decreases as only high-confidence detections are retained.

Performance Differences between Classes:

- Some classes (Bus, Dog) maintain higher recall over a wide confidence range.
- Other classes (Table, Cat) experience a steep decline, suggesting poor recall performance at higher confidence levels.

Recall-Confidence Trade-off:

- Lower confidence thresholds increase recall but also introduce more false positives.
- Higher confidence thresholds reduce false positives but result in missing true detections, decreasing recall.

⁸ 4.1.1.4 F1 Confidence Curve

The F1-confidence curve for YOLOv8 provides a detailed understanding of how the model performs across various confidence thresholds when detecting objects.

Interpretation of the Curve : (Figure 4.7)

Low Confidence Threshold :

- At low thresholds, the model considers nearly all ⁵⁰ predictions, resulting in high recall (many objects detected) but low precision (many false positives).

- F1 scores are generally low here because false positives dominate.

Peak of the Curve:

- The peak F1 score for each line indicates the optimal confidence threshold for the best balance between precision and recall for that class.
- The peak of the blue line indicates the model's overall optimal threshold for object detection across all classes.

High Confidence Threshold :

- At high thresholds, only the most confident predictions are considered, resulting in high precision (few false positives) but low recall (misses many true objects).
- F1 scores decrease because the model becomes overly selective and misses relevant detections.

Per-Class Variations:

- Classes like "Dog" may have higher peaks or more distinct curves, showing better performance for these classes at certain thresholds.
- Flatter or lower curves indicate classes where the model struggles to balance precision and recall, potentially due to overlapping features or insufficient training data.

Optimal confidence threshold :

The optimal confidence threshold for deployment is around 0.186 (as shown for "all classes") where the F1 score is maximized at 0.40.

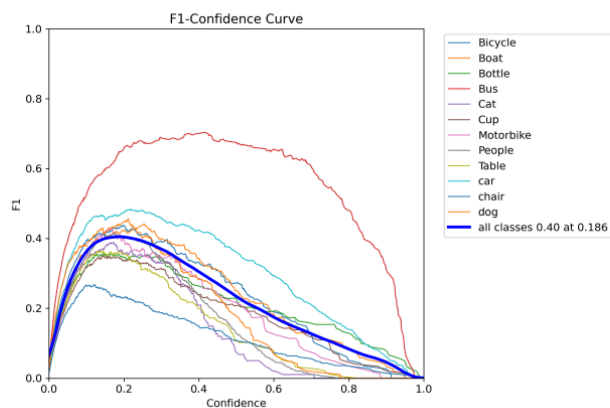


Figure 4.7: F1 Confidence Curve for YOLOv8

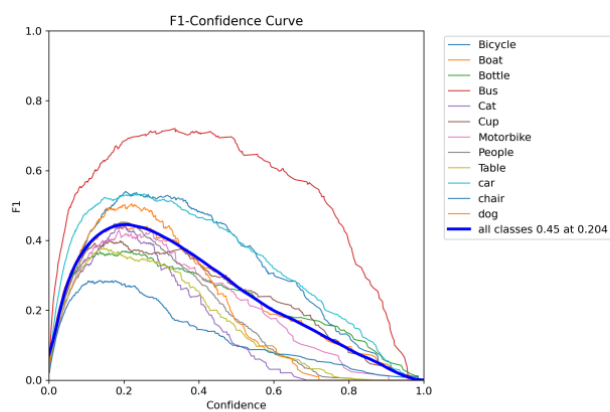


Figure 4.8: F1 Confidence Curve for Y-ZA YOLOv8

Interpreting the Curve : (Figure 4.8)

Low Confidence Threshold :

- At low thresholds, the model accepts nearly all predictions, leading to high recall (detecting most objects) but low precision (many false positives).
- F1 scores are generally low here due to the abundance of false positives.

Peak F1 Scores :

- The peak of each ³³ curve represents the optimal confidence threshold for that class, where the balance between precision and recall is maximized.
- For the blue line, this occurs at 0.204, where the overall model performs best across all classes with an F1 score of 0.45.

High Confidence Threshold :

- At higher thresholds, the model becomes stricter, allowing only the most confident predictions.
- This leads to high precision (fewer false positives) but low recall (many true objects are missed). F1 scores drop because of this trade-off.

Per-Class Performance :

- Classes like "Dog" (red line) have a distinctly high peak, indicating the model performs well for these categories at a specific confidence threshold.
- Classes with flatter or lower peaks may indicate challenges such as insufficient training data or difficulty distinguishing objects.

Insights for Object Detection

Optimal Confidence Threshold :

The optimal confidence threshold for deployment is approximately 0.204, where the F1 score reaches its peak at 0.45, providing the best balance between precision and recall across all classes.

18 4.1.1.5 Precision-Recall Curve

20 A Precision-Recall (PR) curve is a graph that shows the trade-off between precision (accuracy of positive predictions) and recall (ability to identify all positive instances) at different confidence thresholds.

Interpreting the Curve :(Figure 4.9)

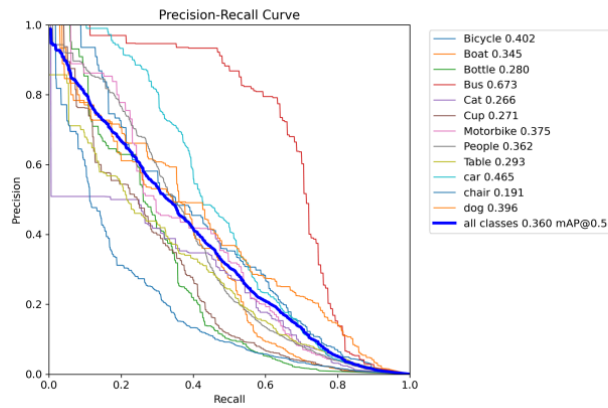
36 Trade-Off Between Precision and Recall :

- At high recall values (right side of the curve), precision decreases because the model includes more predictions, increasing the chance of false positives.
- At high precision values (left side of the curve), recall drops because the model becomes stricter, leading to more false negatives.

Class-Specific Insights :

- Classes like "Bus" (0.673) and "Car" (0.465) show strong performance, with high precision and recall over a broad range of thresholds.
- Low-performing classes such as "Chair" (0.191) and "Cat" (0.266) have flatter or shorter curves, indicating challenges in detecting these objects accurately.

52 In conclusion the PR curve provides a detailed evaluation of the model's detection performance for each class and overall. The aggregate mAP score of 0.360 indicates moderate performance, with room for improvement in certain classes.



37
Figure 4.9: Precision Recall Curve for YOLOv8

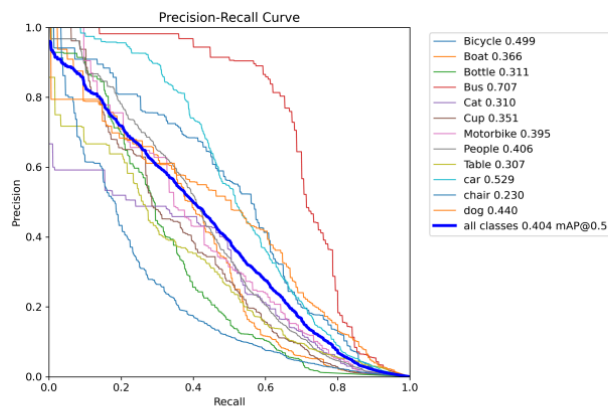


Figure 4.10: Precision-Recall Curve for Y-ZA YOLOv8

Interpreting the Curve :(Figure 4.10)

36

Trade-off Between Precision and Recall :

- At high recall values (right side of the graph), precision decreases because the model includes more predictions, which may increase false positives.
- At high precision values (left side of the graph), recall decreases because the model becomes stricter, missing more true positives.

Class-Specific Performance :

High-Performing Classes :

- The "Bus" class achieves the highest mAP (0.707), indicating robust detection accuracy and completeness.
- The "Car" class also performs well with an mAP of 0.529.

Low-Performing Class :

- The "Chair" class has a lower mAP (0.230), indicating challenges in detecting these objects accurately.
- Other low-performing classes, such as "Cat" (0.310) and "Table" (0.307), may need further optimization.

In conclusion, this precision-recall curve highlights the model's strengths while the overall mAP of 0.404 shows better performance.

Comparison of Performance :

Model	Epochs	Batch	mAP@0.5
YOLOv8	10	2	0.0696
Y-ZA YOLOv8	10	2	0.0709

Table 4.1: Analysis of Performance for Epochs 10

Model	Epochs	Batch	mAP@0.5
YOLOv8	50	8	0.360
Y-ZA YOLOv8	50	8	0.404

Table 4.2: Analysis of Performance for Epochs 50

Chapter 5

Conclusion

⁵ 5.1 Social, Legal, Ethical, and Environmental Issues

The development of Y-ZA YOLOv8, an improved object detection model designed for low-light environments using Efficient Channel Attention (ECA) and Global Context (GC) modules, carefully considers social, legal, ethical, and environmental responsibilities to ensure its safe and responsible application.

Social Implications

Y-ZA YOLOv8 addresses pressing societal needs for improved safety and security in low-light conditions. Its primary applications, such as nighttime surveillance, accident prevention, and public infrastructure monitoring, are designed to enhance safety and security in both urban and rural settings. The model's high detection accuracy contributes to societal well-being by preventing incidents caused by poor visibility or detection failures.

The model also supports the accessibility of advanced technology by optimizing performance for edge devices, enabling its deployment in resource-constrained areas, and thereby addressing technology gaps in underserved communities.

Ethical Responsibility

As engineers, ensuring public safety and environmental sustainability is paramount. The Y-ZA YOLOv8 model integrates design for safety principles, documented through a comprehensive Hazards and Failure Analysis (HFA). This analysis identified potential risks, such as false detections in low-light conditions, which could lead to critical failures in applications like surveillance, autonomous vehicles, or

healthcare monitoring. These risks were mitigated through the implementation of ECA and GC modules, which enhance detection accuracy and robustness, even in challenging environments.

Additionally, the ethical development of the model prioritized fairness and bias reduction. The training datasets were curated to ensure balanced representation across diverse scenarios, preventing unintended biases in the model's predictions. Energy-efficient training practices were adopted to minimize the environmental impact, reflecting a commitment to sustainable AI development.

Environmental Impact

Environmental sustainability was a key consideration during the development of Y-ZA YOLOv8. The incorporation of ECA and GC modules significantly reduces computational complexity, thereby lowering energy consumption during both training and inference. This optimization minimizes the carbon footprint of the model, aligning with global sustainability goals. The model was designed for compatibility with energy-efficient hardware, such as edge devices, reducing reliance on power-intensive cloud infrastructure. These measures ensure that the model not only performs efficiently but also operates in an environmentally sustainable manner.

The Y-ZA YOLOv8 model demonstrates a commitment to ethical, legal, and environmental standards. By incorporating safety measures, adhering to international legal requirements, addressing societal needs, and minimizing environmental impact, this enhanced model offers a responsible and innovative solution for object detection in low-light environments.

5.2 Brief Summary

The project "Y- ZA YOLO: An Improved model of YOLOv8 for object detection in low-light images" is focused to improve YOLOv8's performance in low-light object detection by giving influence in feature extraction and contextual understanding. While YOLOv8 is efficient and fast, it struggles in dim environments due to inadequate channel prioritization and limited spatial awareness, leading to reduced accuracy.

To overcome these challenges, the project incorporates Efficient Channel Attention (ECA) for better feature prioritization and Global Context (GC) for enhanced spa-

tial understanding. These changes improve detection accuracy but also introduce challenges like increased model complexity and higher resource requirements. The project seeks to optimize these trade-offs, delivering a robust, efficient model suitable for real-world low-light detection tasks, including surveillance, autonomous systems and more.

5.3 Future works

The proposed Y-ZA YOLO model which integrates ⁹Efficient Channel Attention (ECA) and Global Context (GC) modules to enhance object detection in low-light images, has shown promising results. However, there are several avenues for further exploration and improvement:

- **Integration of Advanced Image Enhancement Techniques :** Further research can focus on incorporating advanced image preprocessing methods such as adaptive histogram equalization or deep learning-based low-light enhancement algorithms.
- **Real-Time Performance Optimization :** While the model exhibits high accuracy, optimizing it for real-time applications, particularly on edge devices with limited computational power remains a priority.
- **Cross-Domain Generalization :** Investigating the model's robustness across various low-light scenarios, such as underwater environments or night-time urban scenes, can further validate its versatility and effectiveness.
- **Incorporation of Transformer Architectures :** Future versions of the Y-ZA YOLO model could explore hybrid frameworks, integrating transformer-based modules to capture long-range dependencies and further enhance detection accuracy.
- **Dataset Expansion and Diversity :** Expanding the training dataset to include more diverse and challenging low-light scenarios can help the model generalize better to real-world conditions. Efforts can also focus on creating synthetic datasets with realistic low-light effects to augment training.

- **Privacy-Preserving Object Detection** :As object detection technologies become more pervasive, ensuring user privacy is paramount. Future work can explore integrating privacy-preserving techniques such as differential privacy or federated learning to ensure sensitive visual data is processed securely without compromising detection accuracy.

References

- Cao, Y., Xu, J., Lin, S., Wei, F., & Hu, H. (2019). Gcnet: Non-local networks meet squeeze-excitation networks and beyond. In *Proceedings of the ieee/cvf international conference on computer vision workshops* (pp. 0–0).
- Chen, Y., Xu, H., Chang, P., Huang, Y., Zhong, F., Jia, Q., . . . Liu, S. (2024). Ces-yolov8: Strawberry maturity detection based on the improved yolov8. *Agronomy*, 14(7), 1353.
- Chien, C.-T., Ju, R.-Y., Chou, K.-Y., Lin, C.-S., & Chiang, J.-S. (2024). Yolov8-am: Yolov8 with attention mechanisms for pediatric wrist fracture detection. *arXiv preprint arXiv:2402.09329*, 2.
- Ding, J.-Y., Jeon, W.-S., Rhee, S.-Y., & Zou, C.-M. (2024). An improved yolo detection approach for pinpointing cucumber diseases and pests. *Computers, Materials & Continua*, 81(3).
- Ding, M., Guan, S., Liu, H., & Yu, K. (2024). Tir-yolo-adas: A thermal infrared object detection framework for advanced driver assistance systems. *IET Intelligent Transport Systems*, 18(5), 822–834.
- He, Q., Xu, A., Ye, Z., Zhou, W., & Cai, T. (2023). Object detection based on lightweight yolox for autonomous driving. *Sensors*, 23(17), 7596.
- Ju, R.-Y., & Cai, W. (2023). Fracture detection in pediatric wrist trauma x-ray images using yolov8 algorithm. *Scientific Reports*, 13(1), 20077.
- Ju, R.-Y., Chien, C.-T., Lin, C.-M., & Chiang, J.-S. (2024). Global context modeling in yolov8 for pediatric wrist fracture detection. *arXiv preprint arXiv:2407.03163*.
- Kong, H., & You, C. (2024). Improved steel surface defect detection algorithm based on yolov8. *IEEE Access*.
- Li, Q., & Shi, H. (2024). Yolo-ge: An attention fusion enhanced underwater object detection algorithm. *Journal of Marine Science and Engineering*, 12(10), 1885.

- Pan, Y., Xiao, X., Hu, K., Kang, H., Jin, Y., Chen, Y., & Zou, X. (2024). Odn-pro: An improved model based on yolov8 for enhanced instance detection in orchard point clouds. *Agronomy*, 14(4), 697.
- Sohan, M., Sai Ram, T., Reddy, R., & Venkata, C. (2024). A review on yolov8 and its advancements. In *International conference on data intelligence and cognitive informatics* (pp. 529–545).
- Timilsina, A. (2024). *Yolov8 architecture explained*. Retrieved from <https://abintimilsina.medium.com/yolov8-architecture-explained-a5e90a560ce5> (Accessed: 2025-01-28)
- Wang, C., Wang, H., Han, Q., Zhang, Z., Kong, D., & Zou, X. (2024). Strawberry detection and ripeness classification using yolov8+ model and image processing method. *Agriculture*, 14(5), 751.
- Wang, H., Luo, S., & Wang, Q. (2024). Improved yolov8n for foreign-object detection in power transmission lines. *IEEE Access*.
- Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., & Hu, Q. (2020). Eca-net: Efficient channel attention for deep convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 11534–11542).
- Yang, Y., & Deng, H. (2020). Gc-yolov3: You only look once with global context block. *Electronics*, 9(8), 1235.
- Yang, Y., Li, Y., & Tao, M. (2024). Fe-yolo: A lightweight model for construction waste detection based on improved yolov8 model. *Buildings*, 14(9), 2672.
- Yin, Y., Lei, J., & Tao, W. (2024). Detection of liquid retention on pipette tips in high-throughput liquid handling workstations based on improved yolov8 algorithm with attention mechanism. *Electronics*, 13(14), 2836.
- Yuan, Z., Shao, P., Li, J., Wang, Y., Zhu, Z., Qiu, W., . . . Han, A. (2024). Yolov8-acu: improved yolov8-pose for facial acupoint detection. *Frontiers in Neurobotics*, 18, 1355857.
- Zhang, Y., Liang, J., Lu, Q., Luo, L., Zhu, W., Wang, Q., & Lin, J. (2022). A novel efficient convolutional neural algorithm for multi-category aliasing hardware recognition. *Sensors*, 22(14), 5358.
- Zheng, L., Chen, X., & Zheng, L. (2024). Pgdig-yolo: a lightweight method for airport runway foreign object detection. *Journal of Electronic Imaging*, 33(4), 043014–043014.

ORIGINALITY REPORT

22%	16%	13%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	abintimilsina.medium.com Internet Source	3%
2	Submitted to University of Liberal Arts Bangladesh Student Paper	3%
3	www.mdpi.com Internet Source	2%
4	arxiv.org Internet Source	1%
5	tr.overleaf.com Internet Source	1%
6	Submitted to Bilkent University Student Paper	1%
7	www.frontiersin.org Internet Source	1%
8	Onifade, Temilola O.. "An Optimized Object Detection Model for Unmanned Aerial Vehicles", University of Louisiana at Lafayette, 2024 Publication	<1%
9	www.researchsquare.com Internet Source	<1%
10	www.arxiv-vanity.com Internet Source	<1%

11 Xiaofeng Lian, Maomao Kang, Li Tan, Xiao Sun, Yanli Wang. "LSSMask: a lightweight semantic segmentation network for dynamic object", Signal, Image and Video Processing, 2025

Publication

12 Submitted to Telkom University

Student Paper

13 Jin, Yincheng. "Facial Gesture Detection Using Ubiquitous Acoustic Sensing", State University of New York at Buffalo, 2023

Publication

14 Hongwei Li, Jianzhi Huang, Zenan Gu, Deqiang He, Junduan Huang, Chenglin Wang. "Positioning of mango picking point using an improved YOLOv8 architecture with object detection and instance segmentation", Biosystems Engineering, 2024

Publication

15 Submitted to Macquarie University

Student Paper

16 Jingfeng Shan, Lapo Miccinesi, Alessandra Beni, Lorenzo Pagnini, Andrea Cioncolini, Massimiliano Pieraccini. "A Review of Foreign Object Debris Detection on Airport Runways: Sensors and Algorithms", Remote Sensing, 2025

Publication

17 Submitted to Monash University

Student Paper

18 Submitted to Liverpool John Moores University

19	Submitted to Anna University Student Paper	<1 %
----	---	------

20	Submitted to Instituto Polit�cnico de Bragan�sa Student Paper	<1 %
----	--	------

21	Yanpu Yin, Jiahui Lei, Wei Tao. "Detection of Liquid Retention on Pipette Tips in High-Throughput Liquid Handling Workstations Based on Improved YOLOv8 Algorithm with Attention Mechanism", Electronics, 2024 Publication	<1 %
----	---	------

22	sel.ist.osaka-u.ac.jp Internet Source	<1 %
----	--	------

23	Sejeong Lee, Yejin Kong, Taesam Lee. "Development of Deep Intelligence for Automatic River Detection (RivDet)", Remote Sensing, 2025 Publication	<1 %
----	---	------

24	assets-eu.researchsquare.com Internet Source	<1 %
----	---	------

25	ebin.pub Internet Source	<1 %
----	-----------------------------	------

26	"Artificial Intelligence in China", Springer Science and Business Media LLC, 2024 Publication	<1 %
----	--	------

27	J. L Febin Daya, S. Graceline Jasmine, Utku Kose. "Digital Farming and Smart Agriculture for Sustainable Future", CRC Press, 2025 Publication	<1 %
----	--	------

28 Jaiteg Singh, S B Goyal, Rajesh Kumar Kaushal, Naveen Kumar, Sukhjit Singh Sehra. "Applied Data Science and Smart Systems - Proceedings of 2nd International Conference on Applied Data Science and Smart Systems 2023 (ADSSS 2023) 15-16 Dec, 2023, Rajpura, India", CRC Press, 2024
Publication

29 Submitted to Leiden University
Student Paper

30 Li Wang, Fengfan Jiang, Feiyang Zhu, Lei Ren. "Enhanced Multi-Target Detection in Complex Traffic Using an Improved YOLOv8 with SE Attention, DCN_C2f, and SIoU", World Electric Vehicle Journal, 2024
Publication

31 Submitted to Oxford Brookes University
Student Paper

32 Rui-Yang Ju, Weiming Cai. "Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm", Scientific Reports, 2023
Publication

33 Submitted to Gitam University
Student Paper

34 Yukun Ma, Yajun Wei, Minsheng Ma, Zhilong Ning, Minghui Qiao, Uchechukwu Awada. "DCP-YOLOv7x: improved pest detection method for low-quality cotton image", Frontiers in Plant Science, 2024
Publication

35 Submitted to Cerritos College
Student Paper

36 Goodnews Michael, Essa Q. Shahra, Shadi Basurra, Wenyan Wu, Waheb A. Jabbar. "Real-Time Pipeline Fault Detection in Water Distribution Networks Using You Only Look Once v8", Sensors, 2024

Publication

37 Submitted to Indian Institute of Technology IIT Palakkad

Student Paper

38 Submitted to Multimedia University

Student Paper

39 Submitted to University of Stirling

Student Paper

40 Vasu Tammisetti, Georg Stettinger, Manuel Pegalajar Cuellar, Miguel Molina-Solana. "Meta-YOLOv8: Meta-Learning-Enhanced YOLOv8 for Precise Traffic Light Color Detection in ADAS", Electronics, 2025

Publication

41 Submitted to iitj

Student Paper

42 Amjad Rehman Khan, Tanzila Saba. "Explainable Artificial Intelligence in Medical Imaging - Fundamentals and Applications", CRC Press, 2025

Publication

43 Submitted to National College of Ireland

Student Paper

44 Submitted to University of Greenwich

Student Paper

45 Submitted to University of Lancaster

<1 %

46

Submitted to University of Westminster

Student Paper

<1 %

47

Yizhong Yang, Yexue Li, Maohu Tao. "FE-YOLO: A Lightweight Model for Construction Waste Detection Based on Improved YOLOv8 Model", Buildings, 2024

Publication

<1 %

48

erepo.uef.fi

Internet Source

<1 %

49

nemertes.library.upatras.gr

Internet Source

<1 %

50

SRIGANESH SRIHARI, HON WAI LEONG. "A SURVEY OF COMPUTATIONAL METHODS FOR PROTEIN COMPLEX PREDICTION FROM PROTEIN INTERACTION NETWORKS", Journal of Bioinformatics and Computational Biology, 2013

Publication

<1 %

51

aclanthology.org

Internet Source

<1 %

52

Submitted to Cyryx College, Maldives

Student Paper

<1 %

53

diglib.eg.org

Internet Source

<1 %

54

peerj.com

Internet Source

<1 %

55

Das, Shabal. "Advanced Drone Detection: Countering Optical Camouflage Strategies

<1 %

Using Visual Detection Systems", Idaho State University, 2024

Publication

56	Yue Ou, Tian Xu, Jizhuang Fan, Biying Xu, Hegao Cai, Jie Zhao. "Vision-based quality evaluation method towards automated penetrant testing", NDT & E International, 2025	<1 %
----	--	------

Publication

57	doaj.org	<1 %
----	---	------

Internet Source

58	doras.dcu.ie	<1 %
----	---	------

Internet Source

59	iopscience.iop.org	<1 %
----	---	------

Internet Source

60	liujiyuan13.github.io	<1 %
----	---	------

Internet Source

61	Ameya Sharma, Divya Dheer, Vivek Puri, Abdulrhman Alsayari, Shadma Wahab, Prashant Kesharwani. "Insights of biopolymeric blended formulations for diabetic wound healing", International Journal of Pharmaceutics, 2024	<1 %
----	---	------

Publication

62	C.H. Chen. "Signal and Image Processing for Remote Sensing", CRC Press, 2024	<1 %
----	--	------

Publication

63	Gianmarco Goycochea Casas, Zool Hilmi Ismail, Mathaus Messias Coimbra Limeira, Antonilmar Araújo Lopes da Silva et al. "Automatic Detection and Counting of	<1 %
----	---	------

Stacked Eucalypt Timber Using the YOLOv8 Model", Forests, 2023

Publication

64	Jemal Abate, Araba Aman, Dima Adem. "Machine Learning-Based Crop Yield Prediction for Sustainable Agriculture in Eastern Ethiopia", Springer Science and Business Media LLC, 2024	<1 %
----	---	------

Publication

65	Nunes, Gonalo Nuno Gomes Preto Ferreira. "Organizao Social no Peixe Invasor Australoheros Facetus: Hormonas e Comportamento", Universidade do Algarve (Portugal), 2024	<1 %
----	---	------

Publication

66	Wenyu Zhu, Shanwei Niu, Jixiang Yue, Yangli Zhou. "Multiscale wildfire and smoke detection in complex drone forest environments based on YOLOv8", Scientific Reports, 2025	<1 %
----	--	------

Publication

67	Xiaorui Wang, Chao Zhang, Zhenping Qiang, Chang Liu, Xiaojun Wei, Fengyun Cheng. "A Coffee Plant Counting Method Based on Dual-Channel NMS and YOLOv9 Leveraging UAV Multispectral Imaging", Remote Sensing, 2024	<1 %
----	---	------

Publication

68	parkplus.io	<1 %
----	---	------

Internet Source

69	rd.uffs.edu.br	<1 %
----	---	------

Internet Source

70	ris.utwente.nl	<1 %
----	---	------

Internet Source

71 Hadi Sedigh Malekroodi, Seon-Deok Seo, Jinseong Choi, Chang-Soo Na, Byeong-il Lee, Myunggi Yi. "Real-time location of acupuncture points based on anatomical landmarks and pose estimation models", Frontiers in Neurorobotics, 2024
Publication

72 Momina Liaqat Ali, Zhou Zhang. "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection", Computers, 2024
Publication

73 Yue Cao, Jiarui Xu, Stephen Lin, Fangyun Wei, Han Hu. "Global Context Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020
Publication

74 www.ncbi.nlm.nih.gov
Internet Source

75 Maoyu Zhu. "Improvement and Research on Object Detection Algorithm for Wind Turbine Blade Defect Scenarios Based on YOLOv8", Journal of Computer Science Research, 2024
Publication

76 Samsul Ariffin Abdul Karim, Anand J. Kulkarni, Chin Kim On, Mohd Hanafi Ahmad Hijazi. "Intelligent Systems of Computing and Informatics", CRC Press, 2024
Publication

77 Yuanyuan Wang, Haiyang Tian, Tongtong Yin, Zhaoyu Song, Abdullahi Suleiman Hauwa, Haiyan Zhang, Shangbing Gao, Ligu Zhou.

"The transmission line foreign body detection algorithm based on weighted spatial attention", Frontiers in Neurorobotics, 2024

Publication

78

Zijian Yuan, Pengwei Shao, Jinran Li, Yinuo Wang, Zixuan Zhu, Weijie Qiu, Buqun Chen, Yan Tang, Aiqing Han. "YOLOv8-ACU: improved YOLOv8-pose for facial acupoint detection", Frontiers in Neurorobotics, 2024

Publication

<1%

Exclude quotes On

Exclude matches Off

Exclude bibliography On