Vehicle Detection in Various Weather Conditions using YOLOv10 with Squeeze-and-Excitation Blocks

Parsania Ramya Parulbhai
Shubhranil Basak
Unnath Chittimalla
Velidanda Krishna Sai

August 23, 2024

Abstract

Vehicle detection under varying weather conditions presents significant challenges due to the diverse appearances and occlusions encountered. Our submission for the International Conference on Data, Electronics and Computing (ICDEC-2024) Vehicle Detection in Various Weather Conditions Challenge employs YOLOv10. After extensive experimentation with different YOLO versions and hyperparameter tuning via Optuna, YOLOv10 emerged as the most effective model, achieving notable performance metrics. Key innovations include the integration of Squeeze-and-Excitation blocks into the YOLO architecture, which enhanced the model's ability to differentiate between foreground and background. The fine-tuned YOLOv10 model was trained efficiently, with a compact file size of 31.9MB and a training duration of approximately 10-12 minutes on a free-tier Colab T4 GPU.

1 Introduction

Object detection, particularly vehicle detection [8] in challenging weather conditions, remains a crucial yet complex task in computer vision. Accurate vehicle detection [8] systems are vital for autonomous driving, traffic management, and safety applications. Traditional methods often struggle with the variability introduced by different weather conditions, making robust detection systems essential.

In the context of the International Conference on Data, Electronics and Computing (ICDEC-2024) Vehicle Detection in Various Weather Conditions Challenge (VDVWC Challenge) [2], our approach focused on leveraging YOLO (You Only Look Once) [4], a well-established object detection framework known for its speed and accuracy. YOLOv10 [9] was selected as the backbone model due to its superior performance across multiple trials.

Our methodology involved fine-tuning the YOLOv10 model using Optuna [5] for hyper parameter optimization. Initial trials with different YOLO versions and optimizers underscored the efficacy of YOLOv10 combined with the Adam optimizer over others. The integration of Squeeze-and-Excitation (SE) blocks [3] into the YOLO architecture aimed to enhance the model's attention and performance.

2 Squeeze-and-Excitation Block

Figure 1 below visually represents the flow of data through the Squeeze-and-Excitation (SE) Block [6]. The SE Block begins with an input tensor of shape (b, c, h, w), where b is the batch size, c is the number of channels, and h and w are the height and width. It performs global average pooling to reduce spatial dimensions to 1x1, then flattens the output to (b, c). A fully connected layer reduces the channel dimension to c/16, followed by ReLU activation. Another fully connected layer restores the dimension to c, and a sigmoid function normalizes the output. The tensor is reshaped and expanded to match the input's spatial dimensions, then multiplied element-wise with the original input, resulting in a re-calibrated output tensor of the same shape, emphasizing informative features.

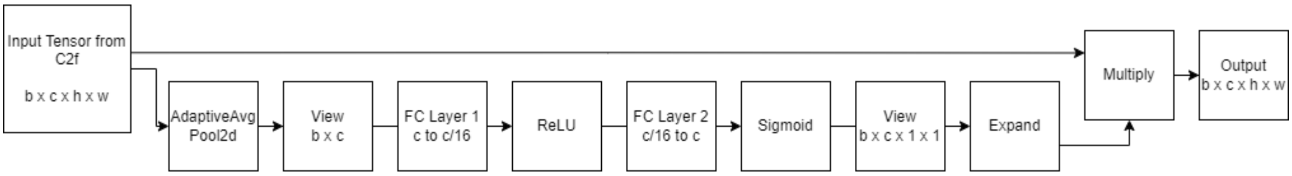

Figure 1: Squeeze and Excitation Block

3 Proposed Methodology

The core of our methodology revolves around YOLOv10 [9], a state-of-the-art object detection model. We utilized the pretrained YOLOv10m model and fine-tuned it for our specific vehicle detection task. Hyper parameter tuning was conducted using Optuna [5], which facilitated efficient exploration of parameter configurations to achieve optimal results.

To further improve model performance, we introduced SE blocks [3] into the YOLOv10 architecture. The SE blocks were incorporated after each pre-existing C2f Convolutional block, aiming to enhance the model's ability to focus on relevant features and suppress irrelevant ones. This modification was implemented by creating and integrating a new SE Block class into the YOLO code base, as shown in Figure 2 below, leading to a beneficial improvement in detection accuracy.

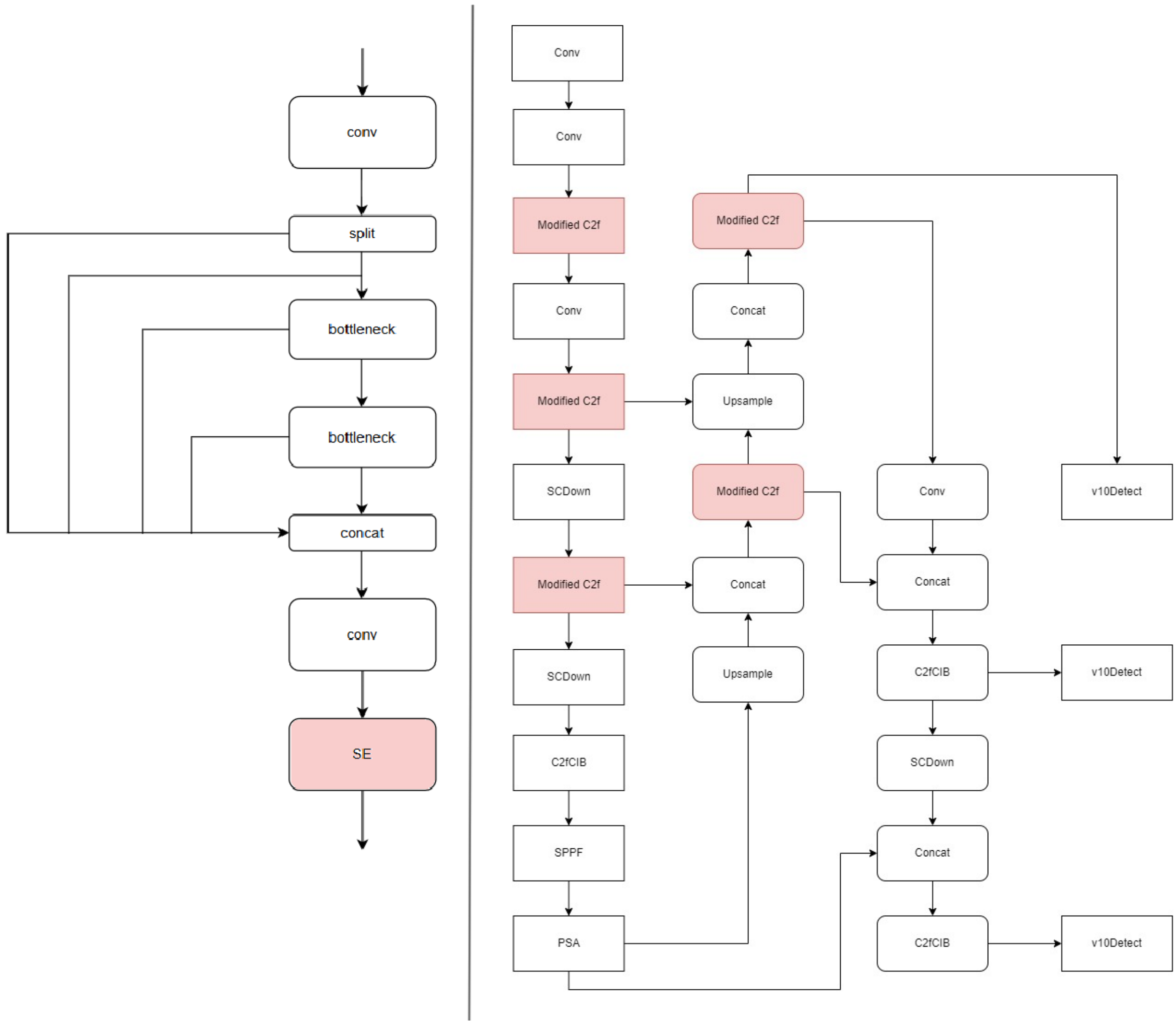

Figure 2: [Left] Modified C2f Block [Right] Modified YOLOv10 Architecture

4 Dataset

The **ICDEC'24 Vehicle Detection Dataset** [7] contains images of 14 vehicle types across 4 weather conditions - sunny day, clear night, rainy day and rainy night. The dataset includes 2,600 training images and 200 validation images, organized into `train`, `val`, and corresponding `labels` folders. Its clear organization aligns with standard object detection practices.

5 Results and Discussion

The model's performance was evaluated using Mean Average Precision (mAP) and Intersection over Union (IoU) metrics. Our fine-tuned YOLOv10 model [1] achieves an **mAP of 0.543 at 0.5 IoU and 0.274 at 0.5-0.95 IoU**, indicating its effectiveness in detecting vehicles across different weather conditions. With **precision and recall values of 0.530 and 0.493**, respectively, the model demonstrates a good balance between accurate detection and minimizing false positives. Results can be previewed from the GitHub repository [1].

6 Conclusion

Our approach effectively leverages YOLOv10 with novel enhancements to address the challenges of vehicle detection in diverse weather conditions. The use of Optuna for hyper parameter tuning and the incorporation of SE blocks contributes to improve detection. The compact and efficient model, trained in a short time frame, demonstrates the potential for practical deployment in real-world scenarios. Future work may explore additional architectural modifications and more diverse data to further enhance detection capabilities.

References

[1] YOLOGANg GitHub Repository for the Proposed Methodology and results, 2024.

[2] ICDEC'24 Vehicle Detection in Various Weather Conditions (VDVWC) Challenge, 2024.

[3] Jie Hu, Li Shen, Samuel Albanie, Gang Sun, Enhua Wu. Squeeze-and-Excitation Networks. 2018.

[4] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. YOLO, 2015.

[5] Optuna. Optuna: A Hyperparameter Optimization Framework, 2024.

[6] Papers with Code. Squeeze-and-Excitation Networks, 2018.

[7] Sourajit Maity. Dataset for the ICDEC'24 Challenge, 2024.

[8] Sourajit Maity, Avirup Bhattacharyya, Pawan Kumar Singh, Munish Kumar, Ram Sarkar. Last Decade in Vehicle Detection and Classification: A Comprehensive Survey, 2022.

[9] Ultralytics. YOLOv10, 2024.