Project Proposal - ECE 176

Minchan Kim
Halıcıoğlu Data Science Institute
A17278478
mik042@ucsd.edu

Andy Ho
Department of Mathematics
A17539521
anh034@ucsd.edu

Abstract

Object detection models struggle in adverse conditions like fog and glare, limiting their real-world reliability. This project introduces a dynamic adaptation framework that classifies environmental conditions in real time and selects the most suitable detection model for inference. Using augmented datasets from Cityscapes, we train specialized sub-models for foggy and glaring conditions, enabling adaptive switching based on a lightweight classifier. Our approach improves detection accuracy under varying conditions with minimal latency overhead, demonstrating the potential for real-time adaptability in vision-based applications.

1 Problem Definition

1.1 Motivation

Object detection models are widely used in applications such as autonomous vehicles, surveillance systems, and robotics. However, their reliability decreases significantly in adverse environmental conditions like fog and glare. These conditions reduce visibility and alter image characteristics, making it challenging for models to perform accurately. Addressing this issue is critical for ensuring the safety and functionality of systems that operate in diverse, real-world environments.

1.2 Key Parts of the Problem

- **Adverse Conditions**: Models must adapt to dynamic and unpredictable scenarios, such as fog, glare, or sudden changes in lighting.
- **Lack of Real-Time Adaptability**: Existing models lack the ability dynamically adjust their parameters or architectures during inference to handle changing conditions.
- **Efficiency Trade-offs**: Although robust solutions exist, they often come at the cost of increased computational complexity, making them impractical for real-time applications.

1.3 Understanding the Problem

The primary challenge lies in designing a system that can efficiently identify and adapt to varying environmental conditions in real time. This involves creating a lightweight condition classifier and leveraging specialized sub-models to optimize performance across conditions.

2 Tentative Method

2.1 Detailed Structure of the Method

- **Conditional Classifier**: A lightweight CNN trained to classify environmental conditions such as "Clear," "Foggy," or "Glaring."
- **Sub-Model Specialization**: Separate object detection sub-models (based on YOLOv5) trained on datasets specific to each condition.
- **Dynamic Adaption Framework**: During inference, the condition classifier determines the current environment and selects most appropriate sub-model for object detection.

2.2 Reasons for Choosing This Method

This approach allows for both modularity and adaptability:

- Modular design simplifies adding support for new conditions (e.g., rain, snow).
- By training specialized models, each sub-model can be optimized for a specific condition, improving accuracy compared to a single generalized model.
- Using a lightweight condition classifier minimizes overhead, ensuring real-time performance.

2.3 Strengths of the Method

- **Scalability**: Can easily extend to handle more environmental conditions.
- **Efficiency**: Optimized sub-models reduce unnecessary computation during inference.
- **Flexibility**: The dynamic framework adapts to varying conditions without requiring expensive retraining.

3 Experiments

3.1 Datasets

The primary dataset used in this project will be the **Cityscapes** dataset, as it contains high-resolution urban scene images with annotations for object detection. In addition, we will be augmenting this data using **Albumentations** to simulate diverse environmental conditions.

The data will be formatted in the following ways:

- Input: Images in JPEG/PNG format (1024x2048 resolution).
- Labels: Bounding box annotations in COCO JSON format.

3.2 Experiments and Their Purpose

In our project, we highlight four crucial experiments. The description and purpose of each are as follows:

1. **Condition Classifier Training**: Train the condition classifier to recognize "Clear," "Foggy," and "Glaring" conditions. The purpose is to ensure high classification accuracy to reliably guide sub-model selection.
2. **Sub-Model Training**: Fine-tune YOLOv5 for each specific condition using augmented datasets. The purpose is to optimize detection accuracy for each environment.
3. **Dynamic Adaption Testing**: Evaluate the complete framework, including condition classification and sub-model switching, on a mixed dataset with varying conditions. The purpose is to measure the framework's adaptability and compare its performance to a single generalized model.
4. **Latency Evaluation**: Measure the time required for condition classification and model switching. The purpose is to ensure the framework meets real-time performance requirements.

4 References

M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 3213–3223.

J. Redmon and A. Farhadi, "YOLOv5: You Only Look Once for Real-Time Object Detection," *arXiv preprint arXiv:1506.02640*, 2020. [Online]. Available: `https://arxiv.org/abs/1506.02640`.

A. Buslaev, A. Parinov, E. Khvedchenya, V. I. Druzhinin, and A. A. Kalinin, "Albumentations: Fast and Flexible Image Augmentations," *Information*, vol. 11, no. 2, p. 125, 2020. [Online]. Available: `https://albumentations.ai`.