

|  |  |
| --- | --- |
| Simon Chaisouang  shchaiso@uark.edu  010878218 | Venkata Kumari  vg021@uark.edu 011088021 |

CSCE 50103: Full Stack Deep Learning

Spring 2025

Homework #3

Late Days: 3

**License Plate Detection Using DETR**

**1. Introduction**

In this assignment, we trained a Detection Transformer (DETR) model to localize license plates in vehicle images. The goal was to build a robust pipeline for detecting license plates, which is fundamental for automated license plate recognition (ALPR) systems. This task serves as the first stage in a two-stage pipeline that facilitates license plate character recognition.

**2. Dataset Overview**

The dataset, derived from the UC3M-LP dataset, contains 2,547 unique vehicles with annotations in the COCO format. It features environmental variations such as different lighting conditions, angles, and occlusions to improve model robustness. Acknowledgment is given to the UC3M-LP dataset as the source.

**3. Methodology**

**3.1 Model Configuration**

We utilized the DETR framework provided in the assignment. The following hyperparameters were optimized for better performance:

* **Learning Rate:** Adjusted for stable training convergence.
* **Batch Size:** Increased to enhance gradient estimation.
* **Number of Epochs:** Extended for model generalization.
* **Augmentation:** Applied techniques like random cropping, flipping, and brightness adjustments to improve robustness.

**3.2 Data Augmentation**

To further improve the performance and robustness of the license plate detection model, we apply two different data augmentation techniques into the training pipeline. Both implementations are mainly contained in the transformer.py file. To mimic moving vehicles, a motion blur is applied by using a horizontal blur from OpenCV. To help the model generalize to different environmental conditions, weather-based augmentations were also implemented such as: rain (white streaks), fog (Gaussian blur), and glare (bright circular light spots). These augmentations are applied at random through make\_coco\_transform() to improve generalization within the model under different weather and motion conditions.

**3.3 Training Process**

Training was conducted in Google Colab using a T4 GPU. The dataset was preprocessed, and the DETR codebase was configured to train on the provided license plate detection dataset. The training process lasted several hours.

**3.4 Model Definitions**

**Model-1:**

The model was trained with the following configuration:

* Batch size:2
* Number of epochs:210

**Model-2:**

The model was trained with the following configuration:

* Batch size:2
* Number of epochs:300

**Model-3 (Data Augmentation):**

For this model we use the same configurations as Model 1 for comparison to see the extent of which our data augmentations techniques have improved the model’s performance.

**4. Results and Analysis**

**4.1 Model Performance**

The model was evaluated using the following metrics:

* **Average Precision (AP):** Measures the precision-recall trade-off.
* **Precision and Recall:** Assesses model reliability and sensitivity.

The DETR model performed well under diverse conditions but exhibited limitations with heavily occluded license plates and extreme lighting variations.

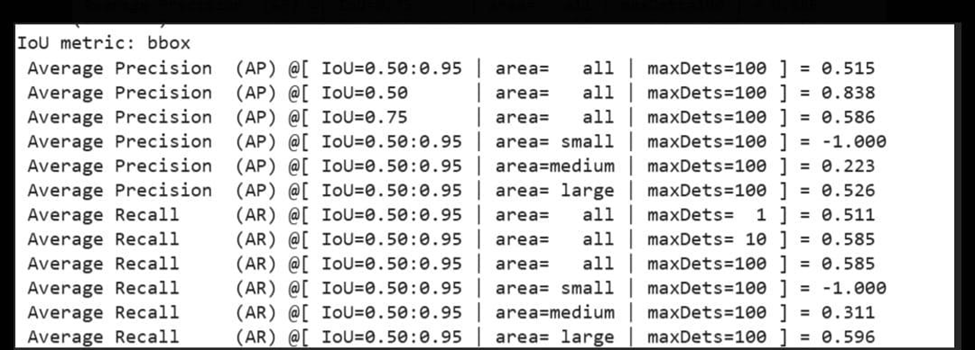
|  |  |
| --- | --- |
|  |  |
| **Figure 1.1 Confidence Score of Model 1 on Test Image 1** | **Figure 1.2 Confidence Score of Model 3 (Data Augmentation) on Test Image 1** |

Looking at the confidence scores of both model 1 and model 3, the model has shown significant improvement after applying data augmentation techniques to the training pipeline in comparison to before. This shows that the model was able to generalize the data better from generating more synthetic data from a limited dataset.

A close-up of a text

Description automatically generated

**Figure 2.1 Table of mean Average Precision (mAP) of Model 1**



**Figure 2.2 Table of mean Average Precision (mAP) of Model 2**

A screenshot of a computer code

AI-generated content may be incorrect.

**Figure 2.3 Table of mean Average Precision (mAP) of Model 3 (Data Augmentations)**

Although previously it was shown that model 3 was able to make more confident predictions than model 1 after applying data augmentation, model 3 did slightly worse than model 1 in terms of mean Average Precision scores. This means that model 1 did slightly better in detecting objects than model 3. However, looking at model 2 the model does notably better in detecting objects than model 1. This shows how optimizing our model with better hyperparameters increases its robustness in object detection.

**4.2 Inference on Real-World Images**

The trained model was tested on real-world license plate samples. While it successfully localized most license plates, failure cases were observed due to:

* Partial occlusions.
* Blurred or low-resolution plates.

**5. Discussion and Potential Improvements**

* **Data Augmentation:** Introduce more diverse augmentations to simulate real-world conditions.
* **Model Refinement:** Experiment with DETR variants, such as Deformable DETR, for better handling of small objects like license plates.
* **Post-Processing:** Implement strategies like Non-Maximum Suppression (NMS) to reduce false positives.
* **Dataset Expansion:** Include images from other regions with varying license plate formats.

**6. Conclusion**

This assignment demonstrated the application of DETR for license plate detection. Despite its strong performance, challenges in occlusion and extreme conditions highlight areas for refinement. With further optimizations, the model can achieve greater robustness and generalization for real-world ALPR systems. From our experiments, data augmentation techniques assist in furthering the model’s performance in making more confident predictions while fine tuning our model with more optimized hypermeters led to higher performance in object detection.