LAB 8 REPORT

A PREPRINT

Daniele cecca
Artificial Intelligence for Science and Technology
Milano Bicocca University
Supervised Learning

May 19, 2024

1 Introduction

This lab report presents the implementation and analysis of the Faster R-CNN object detection model. The objective was to evaluate the model's performance in detecting UNO cards within a variety of backgrounds. The report highlights the model's performance metrics, specifically the Intersection over Union (IoU) scores, and discusses potential challenges faced by the model in detecting certain images. Key insights are provided to understand the factors contributing to varying detection performance.

2 DataSet

The dataset comprises images of UNO cards placed on diverse backgrounds. Each image is annotated with ground truth bounding boxes for the UNO cards, serving as the reference for evaluating the model's performance.

3 Implementation

The Faster R-CNN model was implemented using a pre-trained ResNet-50 backbone. The following steps were taken:

- Data Preparation: The dataset was preprocessed, and annotations were formatted for compatibility with the Faster R-CNN framework.
- Model Training: The model was fine-tuned on the UNO card dataset to adapt the pre-trained weights to the specific task of detecting UNO cards.
- Inference and Evaluation: The trained model was used to make predictions on the test set, and performance metrics such as IoU and FPS were recorded.

4 Evaluation Metrics

- Intersection over Union (IoU): A metric that measures the overlap between the predicted bounding box and the ground truth box. Higher IoU indicates better detection accuracy
- Frames Per Second (FPS): A metric that measures the speed of the model during inference, indicating how many images the model can process per second.

5 Result

5.1 Performace Mentrics

- Average FPS: 9.493

- Highest IoU Score:0.93
- Lowest IoU Score:0.53

5.2 Visual Analysis

The following images showcase the model's performance with the highest and lowest IoU scores:

Figure 1: IoU

- features UNO cards on a high-contrast background, which allows the model to detect the cards accurately.
  IoU Score: 0.93
- The image includes UNO cards on a complex background, making it challenging for the model to accurately delineate the cards.
  IoU Score: 0.53

6 Discussion

The analysis of images with varying IoU scores provides insights into the model's detection capabilities. Here are the specific challenges encountered by the model:

The model struggles with images where the **background has patterns or textures that are similar to the UNO cards**. In the lowest IoU image, the intricate patterns and textures of the background create confusion, leading to inaccurate bounding boxes.

When UNO **cards overlap or are placed close together**, the model may struggle to distinguish between different cards. This results in lower IoU scores due to incorrect or merged bounding boxes.

Variations in lighting, such as shadows and reflections, can obscure parts of the cards, making detection difficult. Poor lighting reduces the visibility of the cards' edges and features, leading to errors in the predicted bounding boxes

Differences in the orientation and scale of the UNO cards affect detection performance. Cards that are tilted or viewed from an angle are harder for the model to recognize compared to cards that are aligned with the image plane.

6.1 Model Improvement

To enhance the model's performance, the following strategies could be considered:

- Data Augmentation: We could introduce more diverse training images with varying backgrounds, lighting conditions, and card arrangements to improve robustness. Techniques such as random rotations, scaling, and brightness adjustments can help the model generalize better to different scenarios.
- Advanced Techniques: We can implement techniques such as non-maximum suppression to refine bounding box predictions and context-aware models to improve detection accuracy in complex scenes. Additionally, incorporating attention mechanisms can help the model focus on relevant regions of the image.
- Additional Training Data: Increasing the amount and diversity of the training data can help the model learn to recognize UNO cards in a wider variety of contexts and conditions.

7 Conclusion

The Faster R-CNN model demonstrated robust performance with an average FPS of 9.493 and varying IoU scores across different images. The highest IoU score of 0.93 indicates the model's potential for accurate object detection, while the lowest IoU score of 0.53 highlights areas for improvement.

The challenges identified in this study—background complexity, card overlap, lighting conditions, and card orientation provide insights into areas where the model's performance can be enhanced. Background complexity poses a challenge because Faster R-CNN utilizes Region Proposal Networks (RPNs) to generate region proposals that likely contain objects, and when the background has complex patterns or textures similar to the UNO cards, the RPN may generate proposals that include parts of the background as well as the cards, reducing accuracy. Card overlap is problematic as the RPN may struggle with significantly overlapping objects, leading to incorrect or merged bounding boxes. Poor lighting conditions affect the model's performance because Faster R-CNN relies on feature maps extracted by convolutional layers, and obscured parts of the cards due to shadows or reflections make feature extraction less effective.

Variations in card orientation and scale also challenge Faster R-CNN because the model uses fixed-size anchors to propose regions, which may not align well with objects of varying orientations and scales. By understanding and addressing the factors contributing to lower IoU scores, we can develop more robust object detection models capable of performing well across a wide range of conditions and scenarios.

References

[1] Faster R-CNN (Paper) *Faster R-CNN*. Available at: `https://arxiv.org/abs/1506.01497`. Accessed 18 May. 2024.

[2] Faster R-CNN PyTorch *Faster R-CNN PyTorch*. Available at: `https://pytorch.org/vision/main/models/faster_rcnn.html`. Accessed 18 May. 2024.