

Data Augmentation for Object Detection: A Comparative Study of Transformation Techniques

Paper Outline

Aqsa Mohsin, Nasrul Huda

May 7, 2025

Abstract

This paper explores the role of data augmentation in enhancing the performance of object detection models. By systematically evaluating common augmentation methods—cropping, rotation, stretching, color adjustments, and flipping—we aim to identify which transformations provide the most significant performance improvements. A YOLOv5/Faster R-CNN model will be used across various controlled experiments, and results will be evaluated using standard metrics like mAP and IoU. Our findings aim to guide practitioners in selecting effective augmentations for real-world applications.

1 Introduction

Object detection models rely heavily on large annotated datasets, which are often expensive and time-consuming to create. Data augmentation offers a cost-effective alternative to synthetically enhance dataset size and diversity. This paper aims to evaluate different augmentation methods and determine their effectiveness in improving model performance. The focus is on identifying which augmentations provide the greatest benefit when applied individually and in combination.

2 Methodology

2.1 Dataset

The study will use a standard object detection dataset such as COCO or Pascal VOC. If needed, a small custom dataset will be created for controlled experimentation. All data will be annotated appropriately and preprocessed for training.

2.2 Model Architecture

The model used will be YOLOv5 or Faster R-CNN, chosen for their performance and popularity. The same model will be used across all experiments to ensure consistency.

2.3 Augmentation Techniques

We will implement and test the following transformations: cropping, rotation, stretching, color adjustments (brightness, contrast), and flipping. These will be applied using the Albumentations or torchvision library.

2.4 Experimental Design

Each augmentation will be tested individually, and in logical combinations. The same training hyperparameters (e.g., learning rate, batch size) will be used to ensure fair comparison. A baseline model trained without augmentation will serve as the control.

3 Evaluation Metrics

Performance will be measured using mean Average Precision (mAP) and Intersection over Union (IoU). Additional metrics like training time and generalization may also be considered.

4 Results

The results section will present quantitative metrics in tables and charts, and qualitative outputs through bounding box visualizations. Comparisons will be drawn between the baseline, individual augmentations, and combined setups.

5 Discussion

We will analyze which augmentations yielded the best improvements and explore any synergistic or diminishing effects when techniques are combined. Practical implications and trade-offs will also be discussed.

6 Conclusion

This paper will summarize the comparative performance of different augmentations and recommend the most effective strategies for real-world object detection. Future work may include auto-augmentation techniques.

References

- Singh, et al.: Data augmentation strategies in vision tasks. (2022)
- Zoph, B., et al.: Learning Augmentation Strategies via Reinforcement Learning. (2020)
- Anonymous: Recent Advances in Data Augmentation. (2024)