# Highlights in Science, Engineering and Technology EMIS 2024 Volume 119 (2024)

## Abstract Object detection has always been a major focus in the field of computer vision, significantly progressing with the advent of deep learning. This paper reports recent advances in object detection and analyses various methodologies through the lens of dual-stage, single-stage, transformer-based, and key point detection approaches. The strengths and weaknesses of each method are quantitatively compared across common datasets. This study also addresses current challenges there still exist, and anticipates future directions.

## Introduction The introduction of AI technologies in everyday applications like autonomous vehicles has magnified the importance of accurate and reliable object detection systems. This paper discusses the progression of object detection technologies, their impact on the automotive industry, and broader implications for urban development and intelligent transportation systems. Essential to these technologies is the accuracy and real-time capability of object detection methods which directly influence the operational success of such autonomous systems.

## Method ### 2.1. Two-stage Detection Approaches Such approaches first generate region proposals and then classify these regions. Popular two-stage methods like Faster R-CNN demonstrate high precision and flexibility in handling diverse datasets. However, they can struggle with small objects and require significant computational power due to their complex architectures.

### 2.2. Single-stage Detection Approaches These methods, exemplified by YOLO and SSD, aim for real-time detection without the intermediate proposal step. They trade off some accuracy for speed, making them suitable for applications requiring instant processing.

### 2.3. Key Point-based Detection Approaches These methods detect key points of objects to infer their position. CornerNet and CenterNet are prime examples, which mitigate the dependency on anchor boxes and improve localization but still face challenges with clutter and small objects.

### 2.4. Transformer-based Detection Approaches Transformers have been adapted for object detection with the DETR model, promoting simpler and potentially more effective detection frameworks. Yet, issues with convergence speed and smaller object detection persist, highlighting areas for future research.

## Experiment ### 3.1. Datasets and Metrics The performance was assessed using standard datasets like Pascal VOC and COCO, employing metrics such as average precision and mean average precision to evaluate different object detection methods.

### 3.2. Performance Comparison \*\*Table 1. Accuracy (mAP) comparison between different representative methods\*\*

| Method | Category | VOC2007 | COCO |  
|:----------------|:-----------------------|---------:|-----:|  
| Fast RCNN | Two-stage method | 70.0 | 19.7 |  
| Faster RCNN | | 73.2 | 21.9 |  
| SSD | Single-stage method | 81.6 | 26.8 |  
| YOLOv4 | | 81.8 | 43.5 |  
| Retina-Net | | 79.5 | 36.2 |  
| RefineDet | | 79.2 | 39.1 |  
| CornerNet | Key point-based method | 79.5 | 42.5 |  
| CenterNet | | 78.0 | 41.8 |  
| FCOS | | 78.8 | 44.7 |  
| Deformable DETR | Transformer-based method | 79.5 | 52.3 |  
| Swin Transformer| | 81.0 | 57.7 |

## Discussion Focused discussion on ongoing issues including multi-domain detections, challenges with small object detections, and the need for lightweight, high-precision networks.

## Conclusion This paper consolidates various advancements and ongoing challenges in the field of object detection, providing a comprehensive look into the methodologies and their performance on benchmark tests. Future directions highlighted include addressing smaller object detection and improving model efficiency for faster real-time applications.

## References 1. Fast R-CNN. Girshick R. 2015. 2. Faster R-CNN: Towards real-time object detection. Ren S.Q., et al. arXiv:1506.01497. 3. Efficient non-maximum suppression, Neubeck A. and Van Gool L. 2006. 4. RoI pooling based fast multi-domain convolutional neural networks, Qin Y., et al. 2016.