

PhD Seminar Talk-I

CoSTAA YOLO: Convolutional Swin Transformer with Attention and Anchor box Optimization on YOLO V7 for Robust Pedestrian Detection

Sukesh Babu V S (CS22D0001), Research Guide: Dr. Rahul Raman

Department of CSE, IITDM Kancheepuram

24th April 2026

ABSTRACT

Pedestrian detection is a crucial element in modern computer vision applications, such as autonomous driving, traffic management, and public safety systems. While existing object detection frameworks achieve high accuracy under standard conditions, their performance often deteriorates in challenging situations that involve occlusion, high crowd density, scale variation, and visually similar objects. These limitations largely stem from a lack of diversity in current training datasets. In this work, we tackle these challenges by introducing the Enriched CamPed dataset, which is specifically designed to capture a wide range of real-world pedestrian scenarios. Based on this dataset, we propose a robust pedestrian detection framework called CoSTAA, which enhances detection performance through improved feature representation and attention mechanisms. We evaluate our proposed approach on multiple benchmark datasets, including CrowdHuman and PennFudan, and demonstrate that it outperforms state-of-the-art methods. This seminar will present the motivation behind our work, the methodology we employed, and the key findings, emphasizing its potential to improve reliability in real-world pedestrian detection systems.