Enhancing Transparency in Real-Time Object Detection for Autonomous Systems Using Explainable AI

G Sarthak[1], K Nitesh[2],V Karthikeya[3],P Aryeshu[4], D Sai Charan[5],
Sampath Mounish[6],D Sankeerthana[7]
UG Students, Department of CSE[1,2,3,4,5,6,7]
GITAM (Deemed to be University), Visakhapatnam, India

Abstract: *The rapid advancement of Artificial Intelligence (AI) in critical applications such as autonomous driving and healthcare necessitates not only high accuracy but also transparency in decision-making. This paper introduces a framework that integrates Explainable AI (XAI) techniques into YOLOv8, a cutting-edge object detection model, to enhance interpretability in real-time environments. We employ Class Activation Mapping (CAM) methods, including Grad-CAM and HiResCAM, alongside Layer-wise Relevance Propagation (LRP) to create an explanation framework. By combining CAM and LRP, we leverage the strengths of both techniques, improving the clarity of model decisions without sacrificing performance. The proposed system maintains high detection accuracy while operating at 50-60 frames per second (FPS) on a dedicated GPU and approximately 24 FPS when executing XAI computations. Our results demonstrate that integrating these XAI techniques enhances the trustworthiness and safety of AI-driven systems, making them more reliable in autonomous driving and other critical applications.*

Keywords: YOLOv8, Explainable AI (XAI), Object Detection, Autonomous Systems, Class Activation Mapping (CAM), Layer-wise Relevance Propagation (LRP), Machine Learning