OoDDINO:A Multi-level Framework for Anomaly Segmentation on Complex Road Scenes

Yuxing Liu
College of Computer Science and Artificial Intelligence, Southwest Minzu University
Chengdu, Sichuan, China
lyxlyx_47@outlook.com

Ji Zhang
College of Computer Science and Artificial Intelligence, Southwest Minzu University
Chengdu, Sichuan, China
Engineering Research Center of Sustainable Urban Intelligent Transportation, Ministry of Education, China
jizhang901@gmail.com

Xuchuan Zhou
College of Computer Science and Artificial Intelligence, Southwest Minzu University
Chengdu, Sichuan, China
xczhou@swun.edu.cn

Jingzhong Xiao*
College of Computer Science and Artificial Intelligence, Southwest Minzu University
Chengdu, Sichuan, China
21700013@swun.edu.cn

Huimin Yang
College of Computer Science and Artificial Intelligence, Southwest Minzu University
Chengdu, Sichuan, China
yhm653750@gmail.com

Jiaxin Zhong
College of Computer Science and Artificial Intelligence, Southwest Minzu University
Chengdu, Sichuan, China
zjxzjx5611@outlook.com

Abstract

Anomaly segmentation aims to identify Out-of-Distribution (OoD) anomalous objects within images. Existing pixel-wise methods typically assign anomaly scores individually and employ a global thresholding strategy to segment anomalies. Despite their effectiveness, these approaches encounter significant challenges in real-world applications: (1) neglecting spatial correlations among pixels within the same object, resulting in fragmented segmentation; (2) variability in anomaly score distributions across image regions, causing global thresholds to either generate false positives in background areas or miss segments of anomalous objects. In this work, we introduce OoDDINO, a novel multi-level anomaly segmentation framework designed to address these limitations through a coarse-to-fine anomaly detection strategy. OoDDINO combines an uncertainty-guided anomaly detection model with a pixel-level segmentation model within a two-stage cascade architecture. Initially, we propose an Orthogonal Uncertainty-Aware Fusion Strategy (OUAFS) that sequentially integrates multiple uncertainty metrics with visual representations, employing orthogonal constraints to strengthen the detection model's capacity for localizing anomalous regions accurately. Subsequently, we develop an Adaptive Dual-Threshold Network (ADT-Net), which dynamically generates region-specific thresholds based on object-level detection outputs and pixel-wise anomaly scores. This approach allows for distinct thresholding strategies within foreground and background areas, achieving fine-grained anomaly segmentation. The proposed framework is compatible with other pixel-wise anomaly detection models, which act as a plug-in to boost the performance. Extensive experiments on two benchmark datasets validate our framework's superiority and compatibility over state-of-the-art methods. Source code is available at: https://github.com/OoDDINO/OoD-DINO.

CCS Concepts

• **Computing methodologies** → **Image segmentation**.

Keywords

Anomaly Segmentation; Open-set Object Detection; Adaptive Dual-Threshold Network

ACM Reference Format:
Yuxing Liu, Ji Zhang, Xuchuan Zhou, Jingzhong Xiao, Huimin Yang, and Jiaxin Zhong. 2025. OoDDINO:A Multi-level Framework for Anomaly Segmentation on Complex Road Scenes . In *Proceedings of the 33rd ACM International Conference on Multimedia (MM '25), October 27–31, 2025, Dublin, Ireland*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3746027.3754710

1 Introduction

Semantic segmentation, as a foundational task in computer vision, aims at classifying each pixel into predefined visual categories [6, 35]. Despite remarkable advances, existing segmentation methods are primarily restricted to recognizing objects within pre-established training distributions, limiting their applicability to open-set environments. In real-world, open-set contexts, particularly safety-critical domains such as autonomous driving [40, 41], segmentation models inevitably encounter out-of-distribution (OoD) or anomalous objects not represented in training sets. The

*Corresponding author: Jingzhong Xiao.

MM '25, Dublin, Ireland

ACM ISBN 979-8-4007-2035-2/2025/10
https://doi.org/10.1145/3746027.3754710