Automatic Labeling of Objects on Footage Collected by Drones

Sandeep Kumar, Ashant Kumar, Abhishek Kumar
under the supervision of Dr. Chalavadi Vishnu

SEVA
SECURE AND EFFICIENT VISION FOR AUTONOMOUS SYSTEMS

भारतीय प्रौद्योगिकी संस्थान तिरुपति
TIRUPATI

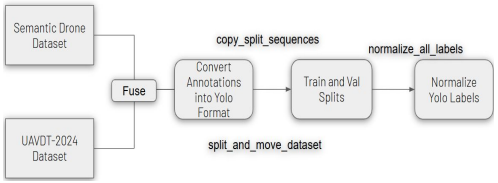
Introduction

Automatic Labeling of Objects on Footage Collected by Drones is an automated labeling pipeline for drone footage using YOLOv8 and segmentation fusion. It enables scalable and secure self-labeling of aerial environments.

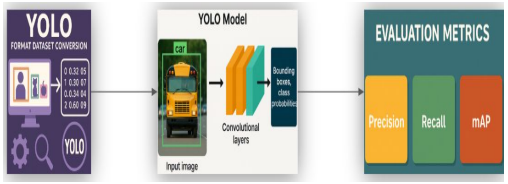
AIM

Design an autonomous, confidence-aware annotation pipeline using object detection and semantic segmentation to minimize manual labeling in UAV systems.

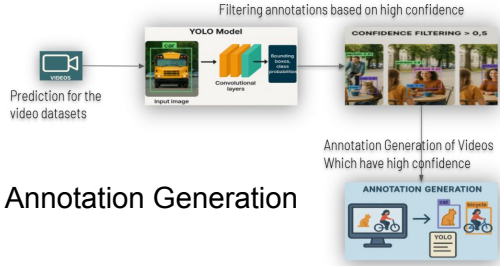
Methodology



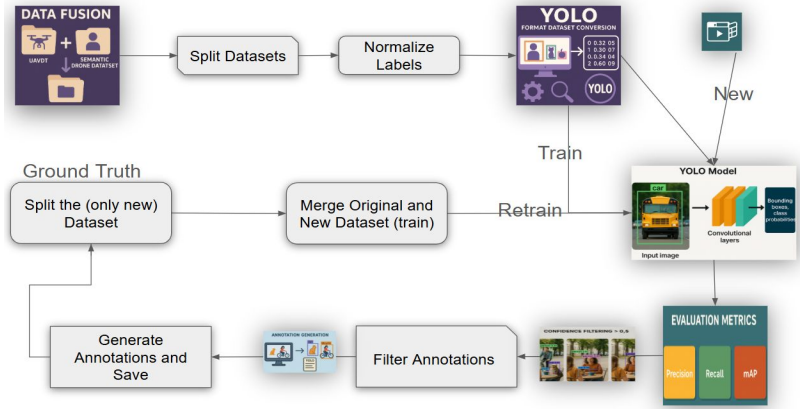
Dataset segregation



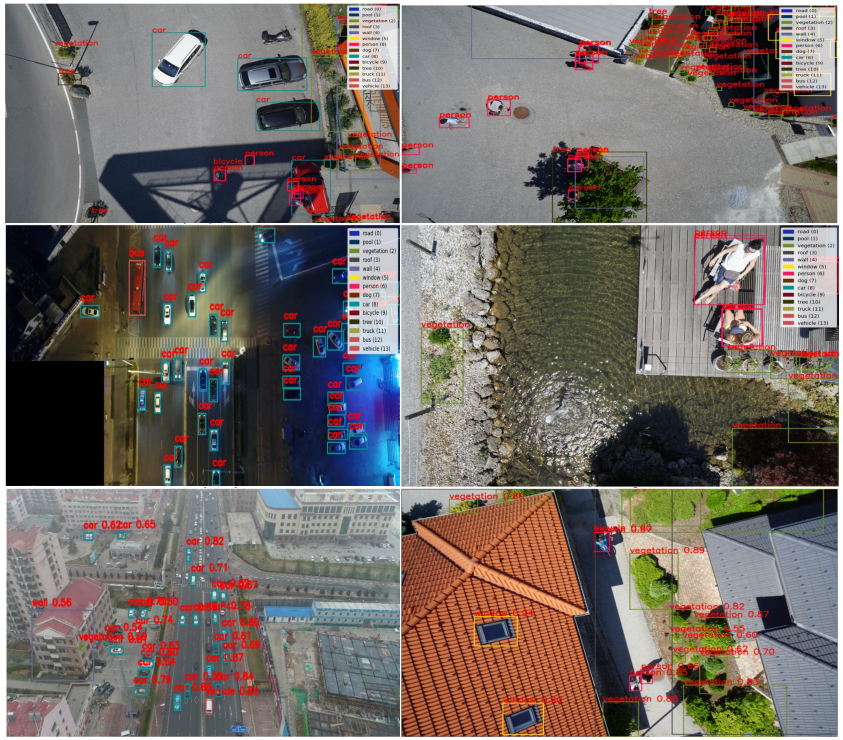
Model Training



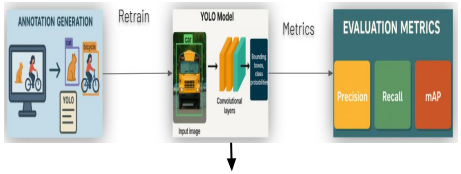
Annotation Generation



Full pipeline showing the steps involved in object detection, annotation, and retraining



Predicted Sample Annotations



Model Retraining

Results

Metric	Before	After	Diff	Trend
Box Loss (Train)	0.73327	0.98916	0.25589	Increase
Cls Loss (Train)	0.45441	0.70747	0.25306	Increase
DFL Loss (Train)	0.86014	0.94087	0.08073	Increase
Precision	0.64380	0.62733	-0.01647	Decrease
Recall	0.26823	0.53795	0.26972	Increase
mAP@0.5	0.28205	0.55362	0.27157	Increase
mAP@0.5:0.95	0.17731	0.39938	0.22207	Increase
Box Loss (Val)	2.02042	0.99684	-1.02358	Decrease
Cls Loss (Val)	2.42410	0.97288	-1.45122	Decrease
DFL Loss (Val)	1.05216	1.01282	-0.03934	Decrease

Changes in Metrics before and after Retraining

Class	Before	After	Diff	Trend
unlabeled	0.1855	0.0715	-0.1140	Decrease
pool	0.7305	0.0715	-0.6590	Decrease
vegetation	0.0798	0.0208	-0.0590	Decrease
roof	0.3768	0.0081	-0.3687	Decrease
wall	0.0513	0.0715	0.0202	Increase
window	0.1524	0.0715	-0.0809	Decrease
person	0.1884	0.2166	0.0282	Increase
dog	0.0029	0.0715	0.0686	Increase
car	0.2459	0.3107	0.0648	Increase
bicycle	0.1027	0.0715	-0.0312	Decrease
tree	0.3563	0.0636	-0.2927	Decrease
truck	0.0291	0.0715	0.0424	Increase
bus	0.0000	0.0715	0.0715	Increase
vehicle	0.0955	0.2094	0.1139	Increase

mAP@0.5:0.95 per class before and after Retraining

Conclusion

It delivers a secure, efficient, and scalable vision pipeline for autonomous UAV systems. It minimizes human involvement while improving detection performance through retraining.