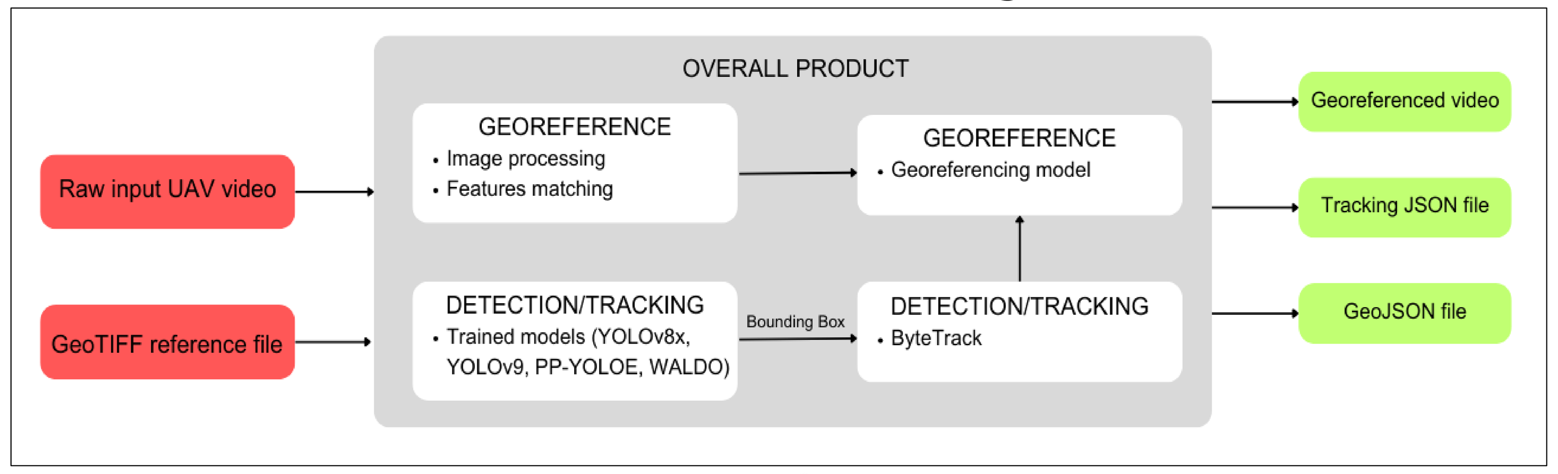
AUTOMATIC GEOREFERENCING FOR UAV VIDEO

PARKALYTICS

OBJECTIVE

Problem Statement: How can we extract traffic data and geographic information from UAV videos for traffic analysis?

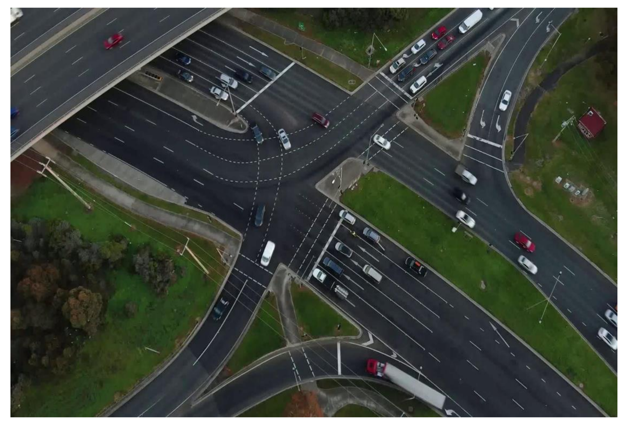
- ➢ Develop a system for automatic UAV video georeferencing and vehicle tracking.
- ➢ Run a convolutional neural network to match key geographical feature points between images.
- ➢ Train a YOLO series model with VisDrone and other custom datasets for vehicle detection and tracking.



Project flow chart

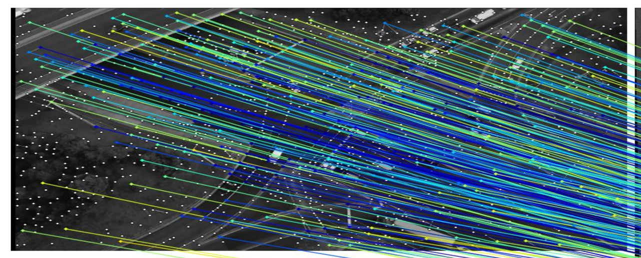
GEOREFERENCING

- OpenCV is used for video stabilization, image processing, homography calculations, and video output.
- SuperPoint and SuperGlue are open-source libraries for CNN-based feature detection and matching, which are used to match key outdoor features between a UAV video frame and a satellite image.

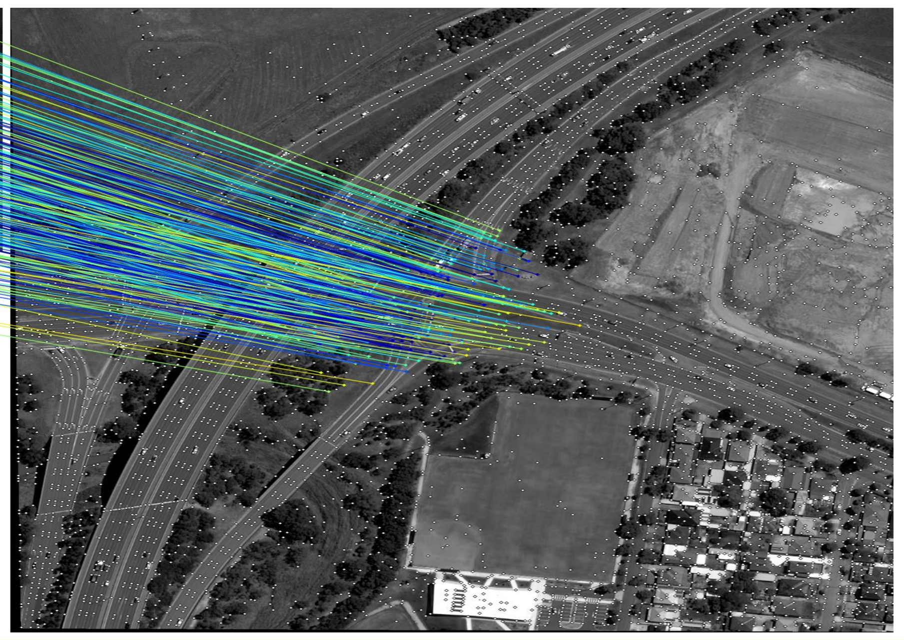


Input UAV Video Frame

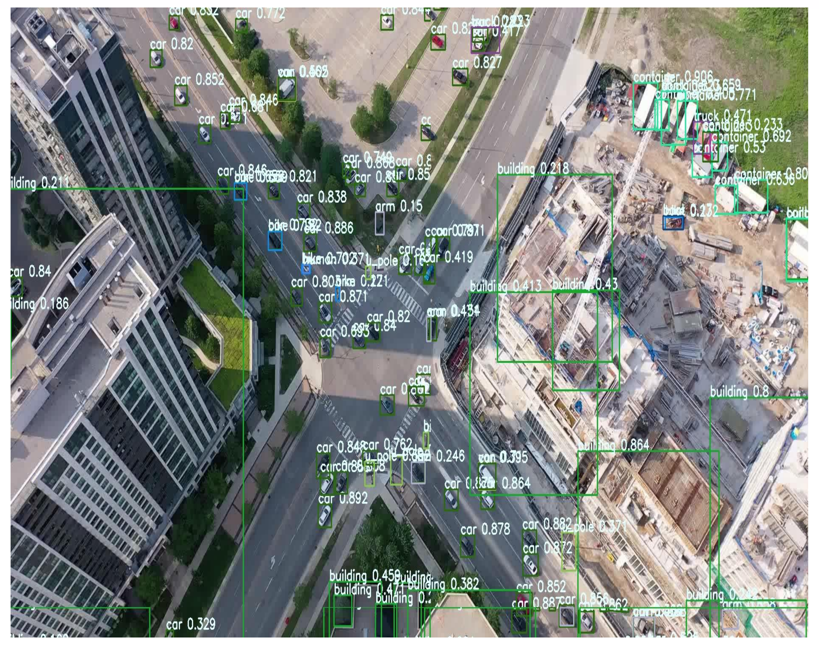
Input Satellite Image



SuperPoint is used to identify key feature points in the input images. SuperGlue is used to match feature points between the two images.



Georeferencing Output



WALDO Model: Resized Output Results

OBJECT DETECTION AND TRACKING

Models: Several object detection models have been trained and evaluated which includes WALDO (Parkalytics's current model), YOLOv8x, YOLOv9e, and PP-YOLOE-SOD+. Utilizing ByteTrack as a multi-object tracking model for tracking purpose. For all models, input images were resized to 640x640 pixels to fit the model input size.

Dataset & Preprocessing: A customized dataset is used by combining the VisDrone dataset and UAV videos from Parkalytics. The dataset comprises six label classes: bicycle, bus, car, human, motorbike, and truck. During early development progress, the models misclassified vans as cars due to similar appearances from an aerial perspective. To resolve, the van and car classes were merged into a single 'car' class.

Data Augmentation: Data augmentation techniques were used for model training that includes rotation, shear, noise addition, and contrast adjustment. However, the models trained on the augmented dataset did not perform well during evaluation, suggesting further refinement of the augmentation strategies may be necessary.

Model	Frames/second	Parameter (M)	mAP50
PP-YOLOE-SOD	16 fps	52.2	0.571
YOLO v8x	41 fps	68.2	0.545
YOLO v9e	45 fps	57.3	0.520
YOLO v9e data aug	45 fps	57.3	0.253
WALDO	30 fps	36.6	0.482

Table of evaluation statistics of detection models

* All models are evaluated using a video provided by Parkalytics that is not used for training. Evaluations are run on a desktop 2080Ti GPU

Overall System Output

The aerial footage is mapped onto satellite reference imagery, precisely registering the detected objects to their real-world geographic locations. Detected objects are visualized with center of bounding boxes (red dot), unique tracking IDs (green numbers), and object class labels (red numbers). Vehicle velocities (blue numbers) are computed by analyzing the transformations of tracked bounding box centers between frames.

Current Final Product Implemented with YOLOv8

FUTURE OBJECTIVES

- Development of vehicle heat maps and dedicated pedestrian tracking capabilities.
- Refinement of data augmentation strategies and retraining the YOLOv9 model on augmented datasets.
- Expansion of the training dataset with increased UAV footage captured specifically for vehicle and pedestrian detection from aerial perspectives, reducing reliance on generic datasets.
- Collect UAV video shot on severe weather or low light condition and train the models on those videos to enhance robustness.

REFERENCES

[1] G. Jocher, A. Chaurasia, and J. Qiu, "YOLOv8," Ultralytics, 2024. [Online]. Available: https://github.com/ultralytics/ultralytics. [Accessed: Jan. 10, 2024].
[2] S. Sturges, "WALDO," Version 2.5, 2024. [Online]. Available: https://github.com/stephansturges/WALDO. [Accessed: Mar. 6, 2024].
[3] P.-E. Sarlin et al., "SuperGlue: Learning Feature Matching with Graph Neural Networks," arXiv:1911.11763 [cs.CV], 2020. [Online]. Available: https://arxiv.org/abs/1911.11763. [Accessed: Feb. 28, 2024].
[4] G. Bradski et al., "OpenCV," Version 4.9, 2024. [Online]. Available: https://github.com/opencv/opencv. [Accessed: Mar. 10, 2024].
[5] Y. Zhang et al., "ByteTrack: Multi-Object Tracking by Associating Every Detection Box," arXiv:2203.08448 [cs.CV], 2022. [Online]. Available: https://github.com/ifzhang/ByteTrack. [Accessed: Mar. 6, 2024].
[6] G. Jocher, A. Chaurasia, and J. Qiu, "YOLOv9," Ultralytics, 2023. [Online]. Available: ultralytics/docs/en/models/yolov9.md at main · ultralytics/ultralytics (github.com). [Accessed: May 1, 2024].
[7] P. Zhu et al., "Detection and tracking meet drones challenge," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 11, pp. 7380–7399, 2021, IEEE.

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

ADVISERS: JENQ-NENG HWANG, CHENG-YEN YANG, ATHARVA MATTAM

SPONSORS: MICHAEL THOREAU, JOSEPH FLETCHER (PARKALYTICS)