NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
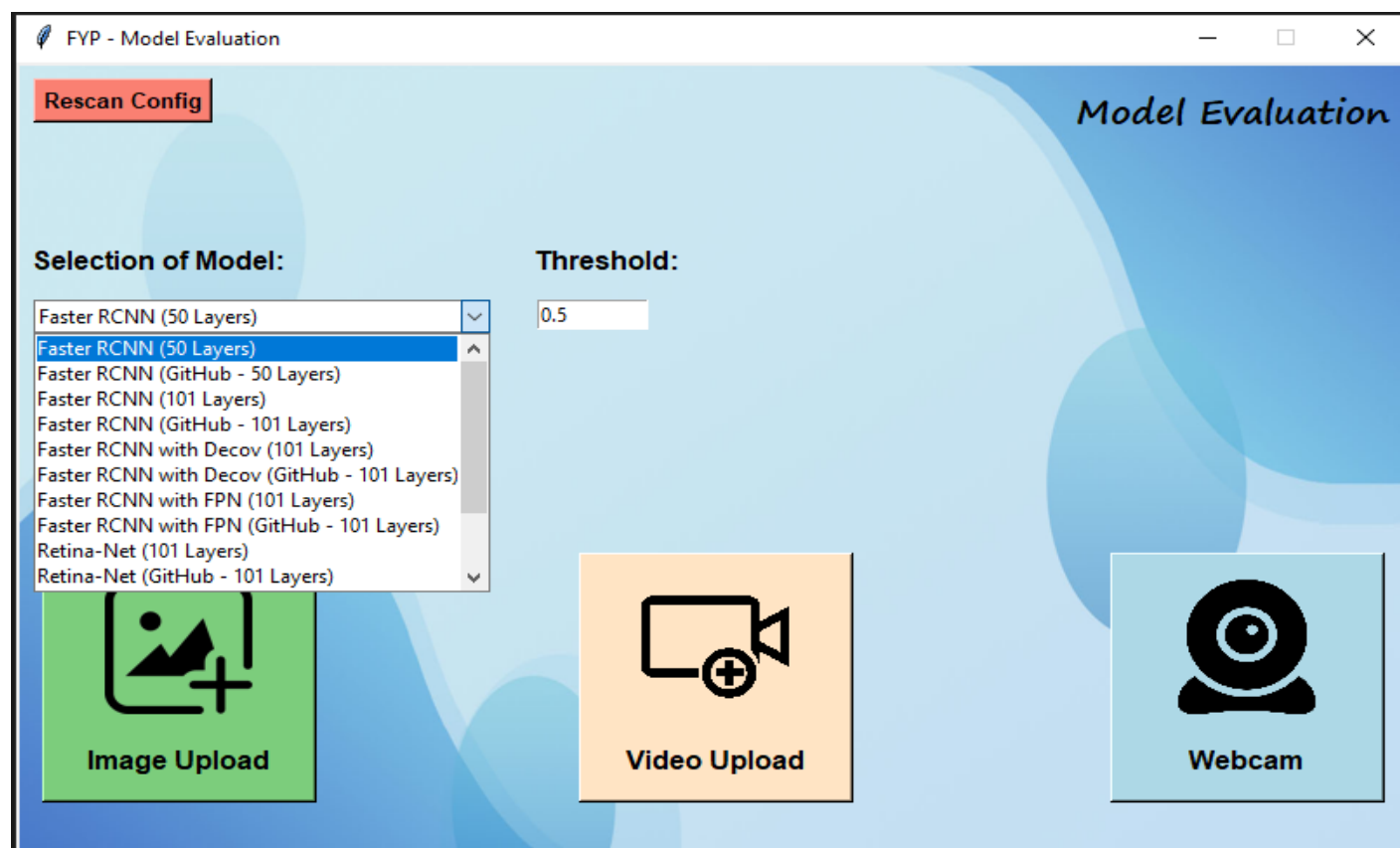College of Engineering

Object Detection
With Deep Learning in Real World Scenario

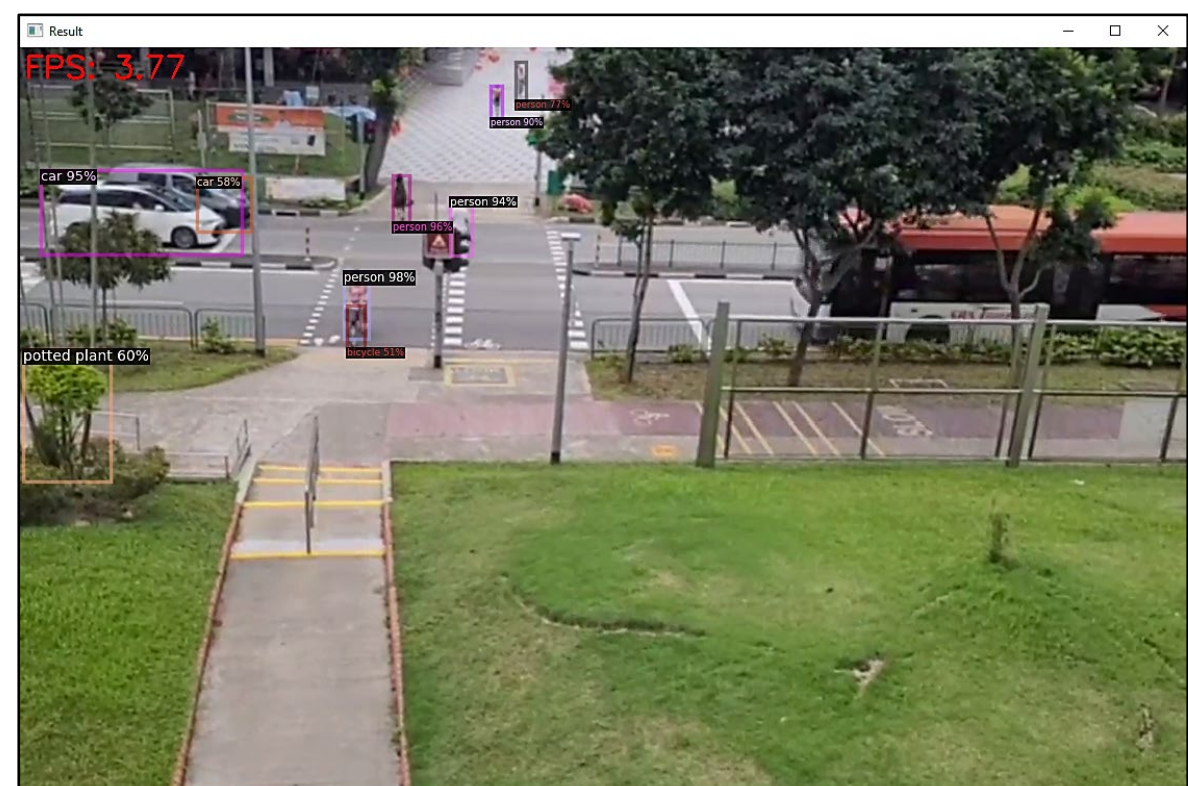Student: Lim Yen Yong   Supervisor: A/P Lu Shijian

Project Objectives:

The objective of this project is to evaluate existing object detection models such as Faster R-CNN, Retina-Net, YOLOv7, and Deformable Detection Transformer and perform different fine-tuning methods and data augmentation to determine if there are any improvements and the consistency of it. The result helps to provide insights into the strengths and weaknesses of these models which can be used to guide future research efforts toward improving object detection performance. A software was also developed to generate the output based on the different models in the real-world scenario.

Results:

Model	Epochs	Without Fine-Tuning					With Fine-Tuning					Improvement
		mAP	AP^{50}	AP^s	AP^m	AP^l	mAP	AP^{50}	AP^s	AP^m	AP^l	mAP
Baseline Faster R-CNN (50 Layers)	~2	22.55	42.01	10.10	26.23	32.08	23.51	43.07	10.34	27.64	34.56	+ 0.96
Faster R-CNN (101 Layers)	~2	25.02	44.80	11.17	**28.69**	35.56	**25.70**	45.79	11.77	**29.22**	**37.54**	+ 0.68
Faster R-CNN with Decov (101 Layers)	~2	25.10	**45.70**	11.37	28.39	34.07	25.12	**46.34**	10.95	28.80	35.62	+ 0.02
Faster R-CNN with FPN (101 Layers)	~2	**25.31**	45.14	**14.69**	28.05	32.01	24.52	45.36	**13.01**	27.77	31.63	- 0.79
Retina-Net (101 Layers)	~2	24.03	39.12	12.06	27.29	30.15	23.99	40.10	12.24	27.37	30.97	- 0.04
YOLOv7	12	21.21	32.68	9.85	22.39	30.05	23.93	36.31	12.14	26.28	31.62	+ 2.72
Deformable DETR (50 Layers)	6	24.31	40.73	9.44	25.73	**36.42**	25.14	43.50	11.63	28.67	34.95	+ 0.83

Overall Result (Fine-Tuning & Data Augmentation)   Faster R-CNN with FPN (101 Layers)