

Supplementary Document



Figure 1: Real Dataset

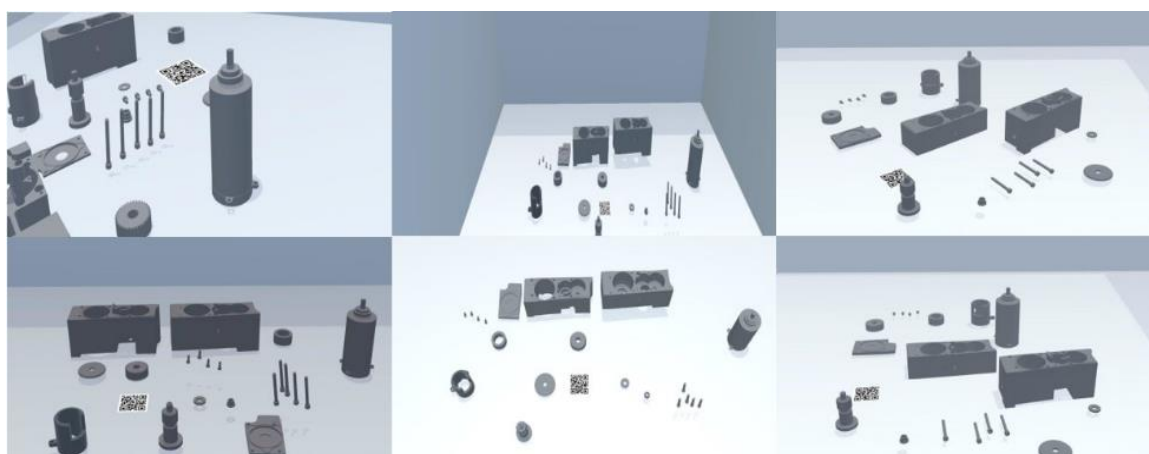


Figure 2: Synthetic Data Generated in Unity

Table 1: Compare different version of Yolo model with custom dataset

Training Dataset Type	YOLO version	Dataset Size	Test Data	mAP @0.5	Epochs
Real	v5m	600	Real	0.755	30
	v7	600	Real	0.856	30
	v8m	600	Real	0.887	30
Synthetic	v5m	600	Real	0.410	30
	v7	600	Real	0.429	30
	v8m	600	Real	0.512	30
Mixed (50-50)	v5m	600	Real	0.759	30
	v7	600	Real	0.809	30

	v8m	600	Real	0.889	30
--	-----	-----	------	--------------	----

Table 2: Interactive commands

Button Name	Voice Command	Process Initiation
Start QR	Start QR	Initialize QR code tracking
Stop QR	Stop QR	Stop QR code tracking
Photo Capture	Photo Capture	Click the picture using HoloLens camera
Object Tracking	Object Tracking	Initialize the object tracking algorithm
Next	Next	Get to the next assembly step information
Previous	Previous	Go back to the previous assembly step information
IP	Connection to server	Establish connection between server and HoloLens

Table 3: Hardware Specification and detection quality

Models Trained	Model Parameters	FPS (Cycle Time)
YOLOv5m	21.2M	RTX 3090 – 13 FPS RTX 3060 – 5 FPS Jetson Nano – 0.5 FPS
YOLOv7	36.9M	RTX 3090 - 9FPS RTX 3060 – 4 FPS Jetson Nano – 0.3 FPS
YOLOv8m	25.9M	RTX 3090 – 11 FPS RTX 3060 – 4 FPS Jetson Nano – 0.3FPS