

中興大學電機系: 黃穎聰、范志鵬、賴永康、吳崇賓、林維亮; 逢甲大學電子系: 陳冠宏; 虎尾科大電機系: 林光浩

計畫介紹

適用於室內人群環繞的應用基礎上，發展自走載具相關的感測、導航、定位、以及AI運算晶片技術。

- 感測、導航技術強調能讓自走式載具以符合社交禮儀的方式閃避人群障礙，自主行進
- 能對行人以及各式障礙物做精準的分類，並且對每一目標進行行為預測
- AI晶片技術上提供載具在edge端執行感測、導航時最佳的運算效能
- 行控搭配定位技術，協助載具避開人流，獲得最佳化的路徑規劃。



Platform porting

總體目標

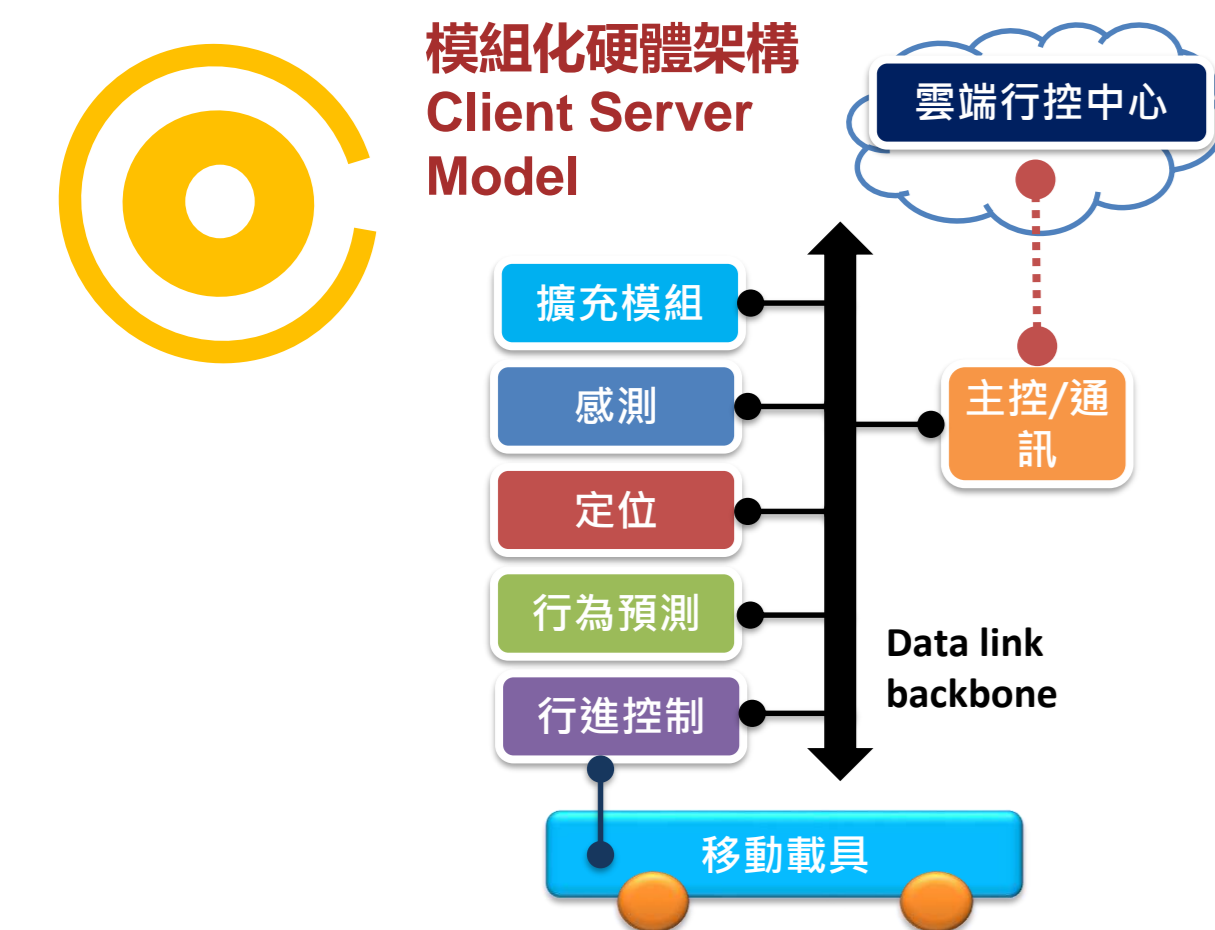
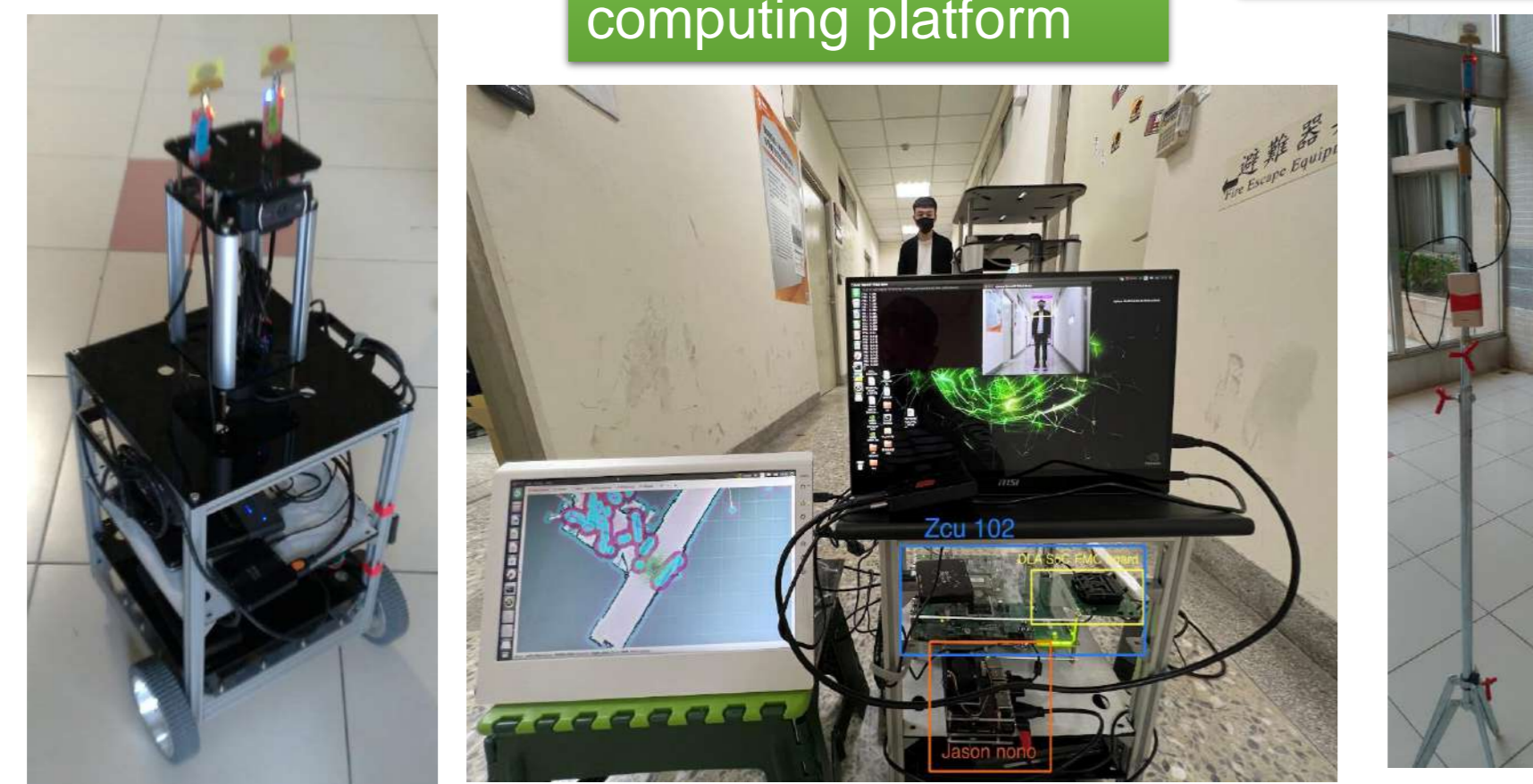


在人群環繞環境下閃避障礙，自主移動

第三代自走載具以及系統整合

ASIC + FPGA + GPU computing platform

UWB anchor



智慧感測與辨識技術

- 影像式物件分類
- 行人臉部方向判別與行為預測(危害評估)
- 相位陣列式超音波掃描雷達(人群空隙偵測)
- 影像/光達/雷達訊號多重資料融合

AI邊緣計算處理器晶片

- 新一代深度學習運算架構(DLA)
- 動態重組與多工平行運算技術
- DLA發展與驗證環境
- CNN模型自動硬體架構映射與最佳化技術

室內定位系統

- 基於超寬頻訊號之室內定位系統與車輛跟隨系統
- 單視覺影像比對式之室內定位系統
- 結合光達、室內輔助定位與慣性導航的SLAM技術

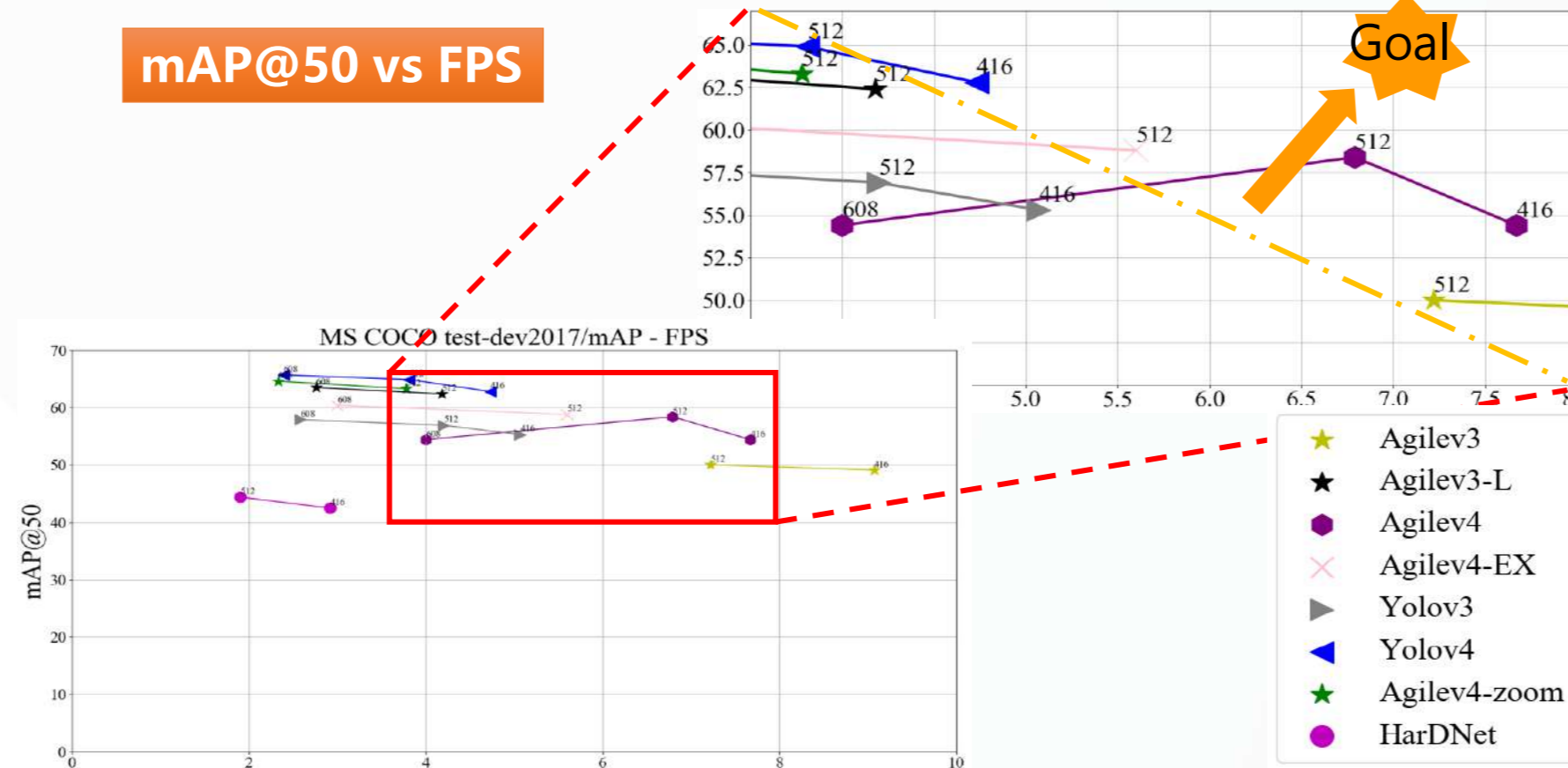
具載導航控制與系統整合實作

- 階層式、考慮社交禮儀與人流之路徑規劃與導航模型
- 虛擬平台上的導航模型訓練與加速
- 自走載具製作
- 感測、定位、導航在FPGA + GPU異質平台上的整合

成果亮點

兩階段式高效能物件偵測、追蹤辨識系統與行人資訊萃取技術

- Stage 1 全新影像式AI物件偵測模型架構Agilev4，可針對人群聚集的室內環境進行多種物件辨識
- 在mAP@50 vs FPS，精準度與運算速度乘積複合指標以及速度-精確度-功率消耗綜合指標評比皆優於現有的AI物件偵測模型YOLOv4
- 獨創之Group-of-Picture (GoP)-mode加速技術，運算速度加速達12.3X
- Stage 2 輕量化分類器可萃取包括人臉方向，性別，衣著，大人/小孩等行人資訊，只需要4 MFLOPS/object運算量



| Model | GFLOPS | mAP@50 | FPS | Power (W) | Score |
|---------------|--------|--------|------|-----------|-------|
| Agilev4-leaky | 49.36 | 58.4 | 7.23 | 6.612 | 4199 |
| Agilev4-mish | 49.36 | 56.5 | 6.79 | 6.612 | 3748 |
| Agilev3 | 49.5 | 50.0 | 7.22 | 6.241 | 3471 |
| YOLOv4 | 91.09 | 64.9 | 3.81 | 8.706 | 1965 |
| YOLOv3 | 99.79 | 56.9 | 4.20 | 9.784 | 1584 |

可擴增與模組化之低功耗AI硬體加速器

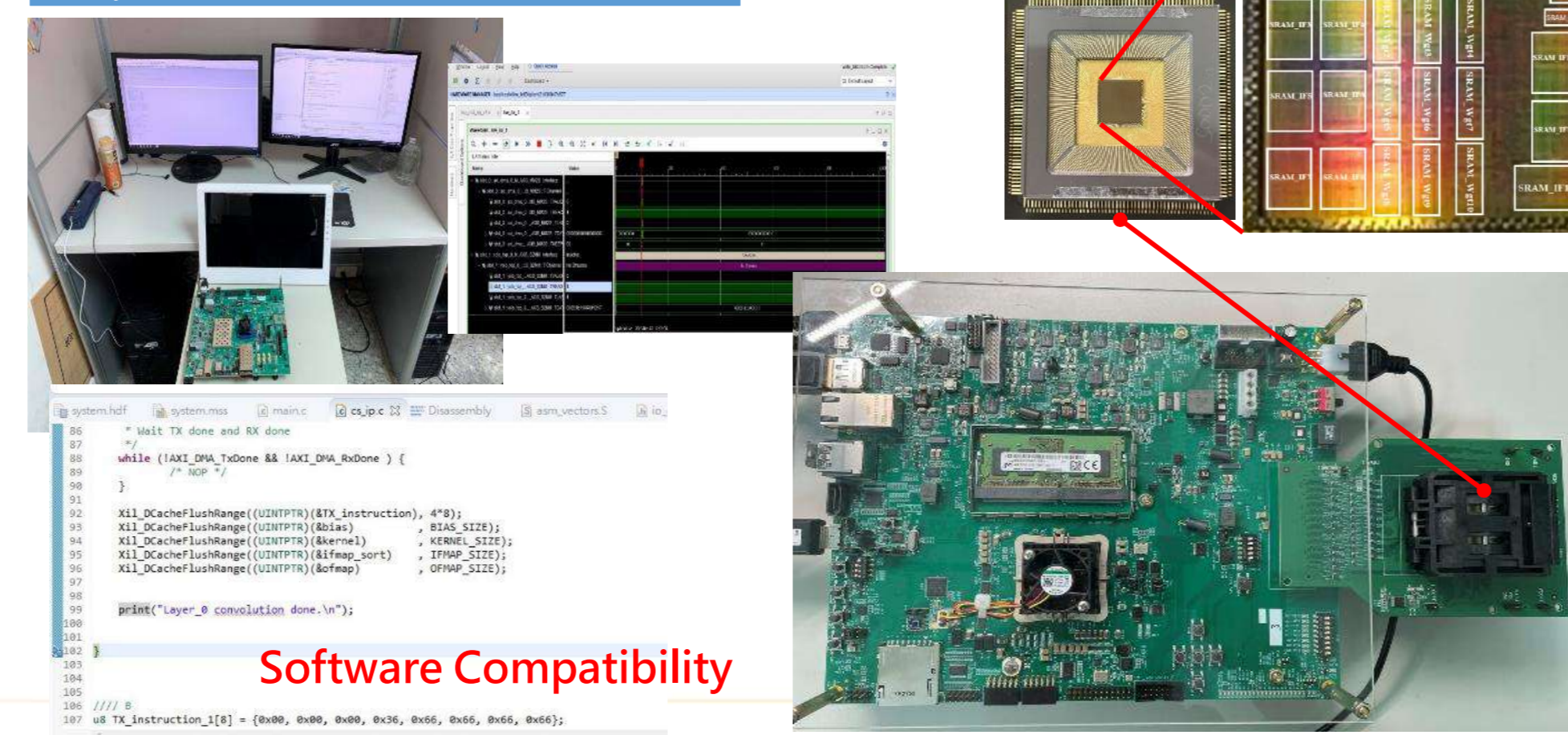
- 可重組運算架構，可模組化成36X的PEs，最高支援2160PEs
- 可組態成高運算吞吐量(2160PEs, 864GOPS@200MHz)或低功耗(36PEs, 3.58GOPS@50MHz)獨立IP。
- Hetero Layer Fusion頻寬最佳化技術，最高可在YoloV2上節省43%頻寬
- IC在40nm實現上，其Area efficiency，在不計入Sparsity運算時為參考文獻之2.29X，計入50%Sparsity運算，為參考文獻1.58X~2.91X

DLA prototype & Verification platform

1st generation chip

2nd generation DLA design

FPGA Hardware/Software Integration System and Verification Platform



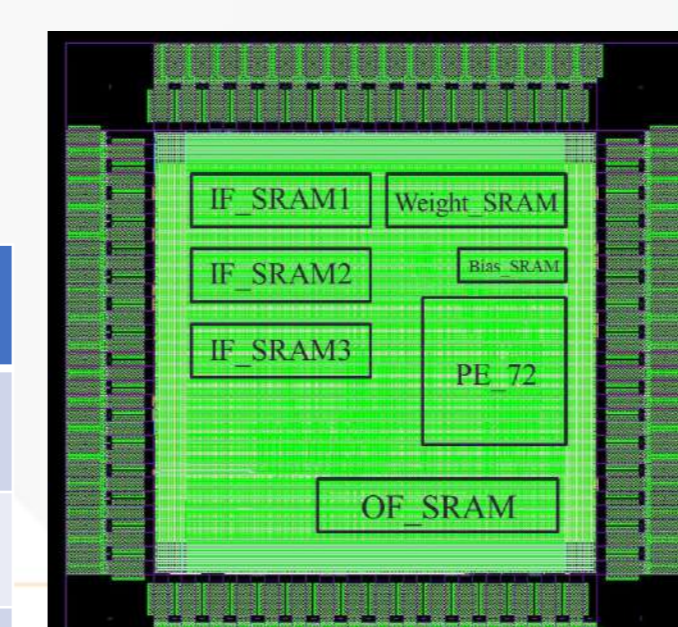
Software Compatibility

IC Integration System and Verification Platform

- PE extensible, 36X PEs
- SRAM size : 173M (25%, vs. 1st DLA)
- Network Sparsity Supported
- Bandwidth: 46%, vs. 1st DLA
- PE utilization rate: 95%(72PE), 85%(288PE)

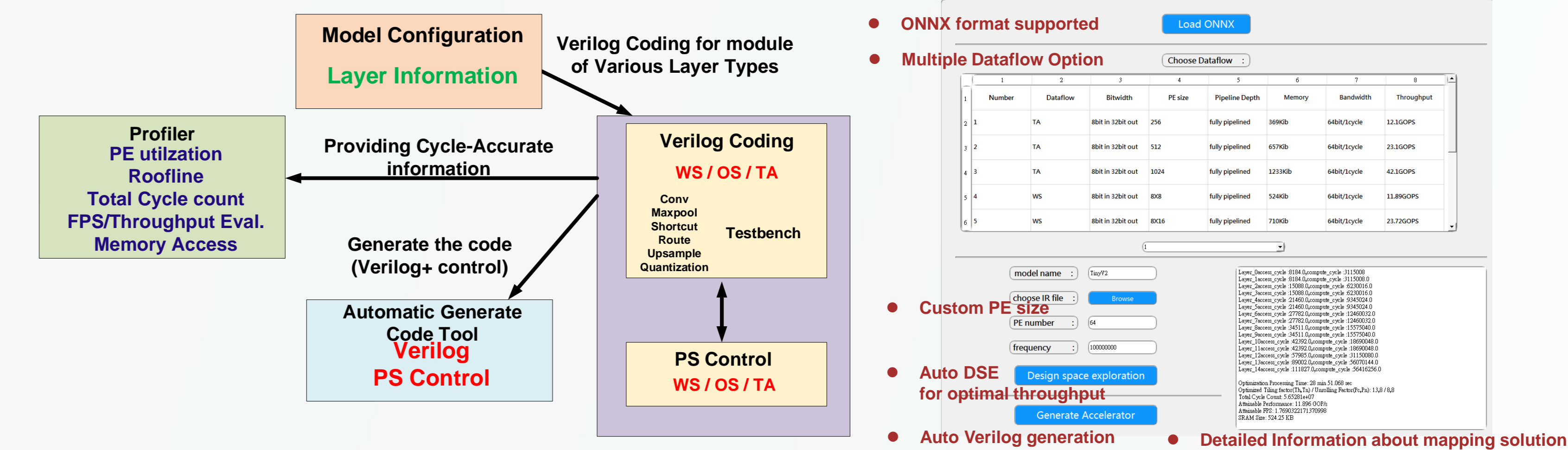
40nm Chip

| | |
|------------------|---------------------------|
| Core Frequency | 200MHz |
| Die size | 1340x1340 um ² |
| Core Utilization | 52% |
| I/O pad | 128 |



客製化DNN硬體加速器生成與模型映射、效能評估工具

- 自動生成基於卷積神經網路之硬體架構之Verilog code工具: 可支援Convolution, Pooling, Fully-Connection, Shortcut, Route, Depth-wise Convolution, Up-sampling等組合之神經網路
- 可生成3種運算核心硬體架構之Verilog code
- 可視效能指標分析工具(Profiler):分析PE使用率、Roofline factor、Total cycle count、FPS評估、各層DRAM/SRAM access的情況



適用於人群環繞環境並遵守社交禮儀的定位與導航與技術

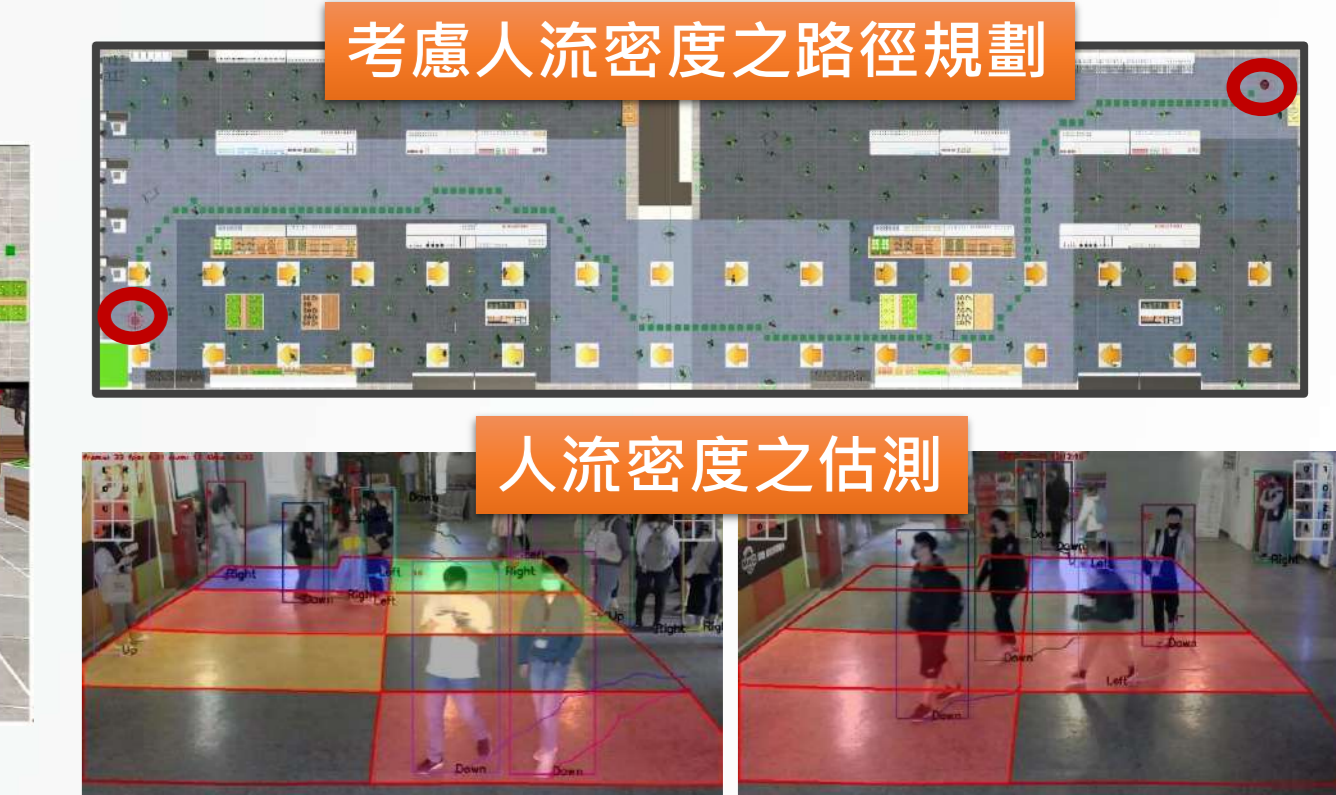
- 以遊戲引擎Unity在虛擬環境模擬並製造虛擬資料，加速導航模型的訓練
- 遵循社交禮儀的導航
- 行人行進軌跡預測，預先調整行進路徑減少避障等待
- 危險行為評估(小孩、視線方向)，動態調整行進時與行人間的安全距離
- 依區域人流、密度，來決定廣域導航路線



模擬超市



閃避行人更改路徑



考慮人流密度之路徑規劃

人流密度之估測



行人軌跡預測



面對面狀態出現、偵測、與地圖上標為禁航區



人群環繞環境場域實測

學餐穿堂

專題展



社交禮儀導航

台中高鐵

台中火車站