

國立清華大學資訊工程系 114 學年度上學期專題報告分工表

| | | | |
|-----------|---------------------------------------------------------------------------------------------|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 專題名稱 | FPGA-Based Acceleration of TinyLLaVA Model Inference via High-Level Synthesis on Alveo U280 | | |
| 參加競賽或計畫 | 參加對外競賽 | 參與其他計畫 | ■無參加對外競賽或任何計畫 |
| 學號 | 姓名 | 分配占比 | 分配內容 |
| 111060013 | 劉祐廷 | 34% | <ol style="list-style-type: none"> 1. 將 Phi-2 從 Python 改寫成 C++ 2. 驗證 Phi-2 C++ 模型的正確性 3. 透過 High-Level Synthesis 加速 Phi-2 模型推論 4. 說明影片製作、海報製作 5. 撰寫書面報告 (Phi-2 相關內容、研究動機、結語、硬體設備介紹、執行環境介紹) |
| 111060016 | 林芷韻 | 33% | <ol style="list-style-type: none"> 1. 將 SigLIP 從 Python 改寫為 C++ 2. 驗證 SigLIP C++ 模型的正確性 3. 撰寫 SigLIP Host 與 Kernel 初版架構 4. 撰寫書面報告 (SigLIP 和 Connector 相關內容、模型介紹) |
| 111060019 | 黃子恩 | 33% | <ol style="list-style-type: none"> 1. 整合 Phi-2 與 SigLIP，撰寫 TinyLLava 的整體框架 2. 針對各 kernel 內部結構及運算邏輯進行優化，提升運算效率 3. 設計並實作記憶體存取策略的最佳化方案，降低頻寬瓶頸 4. 撰寫書面報告 (SigLIP 和 Connector 相關內容、模型介紹) |

| | | | |
|------|-------------------|--------|--|
| 學生簽名 | 劉祐廷 林芷韻 黃子恩 | 指導教授簽名 | |
|------|-------------------|--------|--|