

# 基於Lenet-5之HLS實作

## 卷積神經網路硬體加速器設計

### HLS implementation based on Lenet-5 Design of Convolutional Neural Network Hardware Accelerator

學生：林裕閔、鄭筑瑄

#### 摘要

本專題實作高階合成(High-Level-Synthesis)卷積神經網路硬體加速器設計應用於Lenet-5的CNN Model，利用Vivado HLS和Vivado SDSOC做為開發平台，以及Zedboard做為開發板，透過FPGA運算速度快、功率消耗低等特性，提升Lenet-5對數字辨識的運算速度。

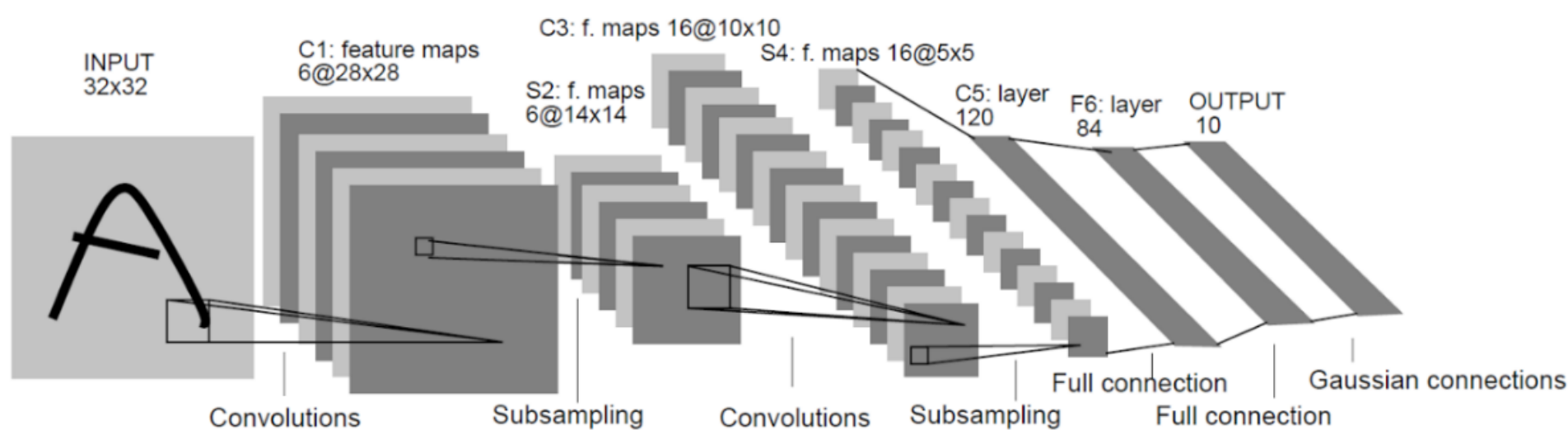
實驗結果，本專題提出兩種硬體加速器架構已在Xilinx Zedboard上實現。所使用的LUT資源分別為44997和18693個，工作頻率皆為100MHz，辨識速度皆比軟體的執行速度快捷。

#### Lenet-5模型

Convolution(卷積層):kernel為5x5的卷積運算。

Pooling(池化層):使用Average pooling。

Fullyconnect(全連接層):將運算完的數據展平。

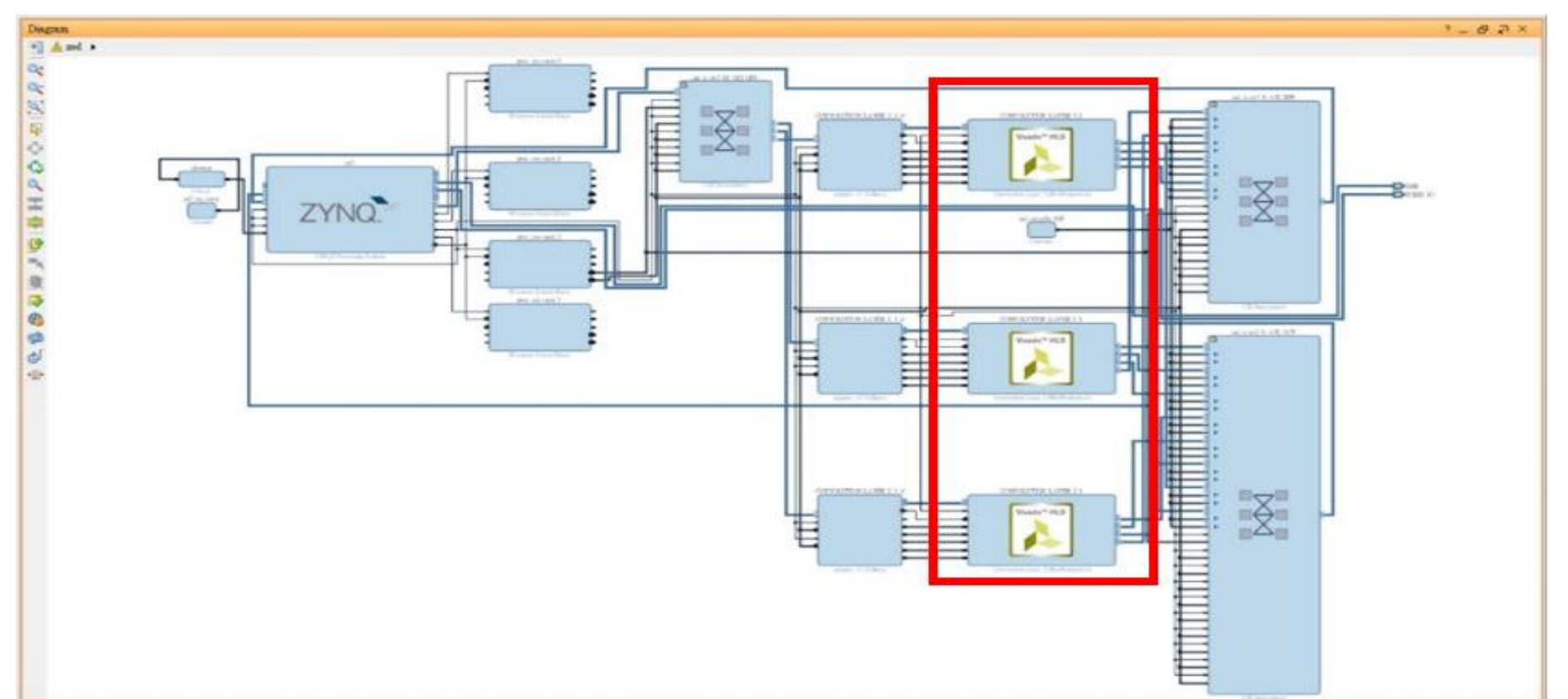


#### 實驗成果

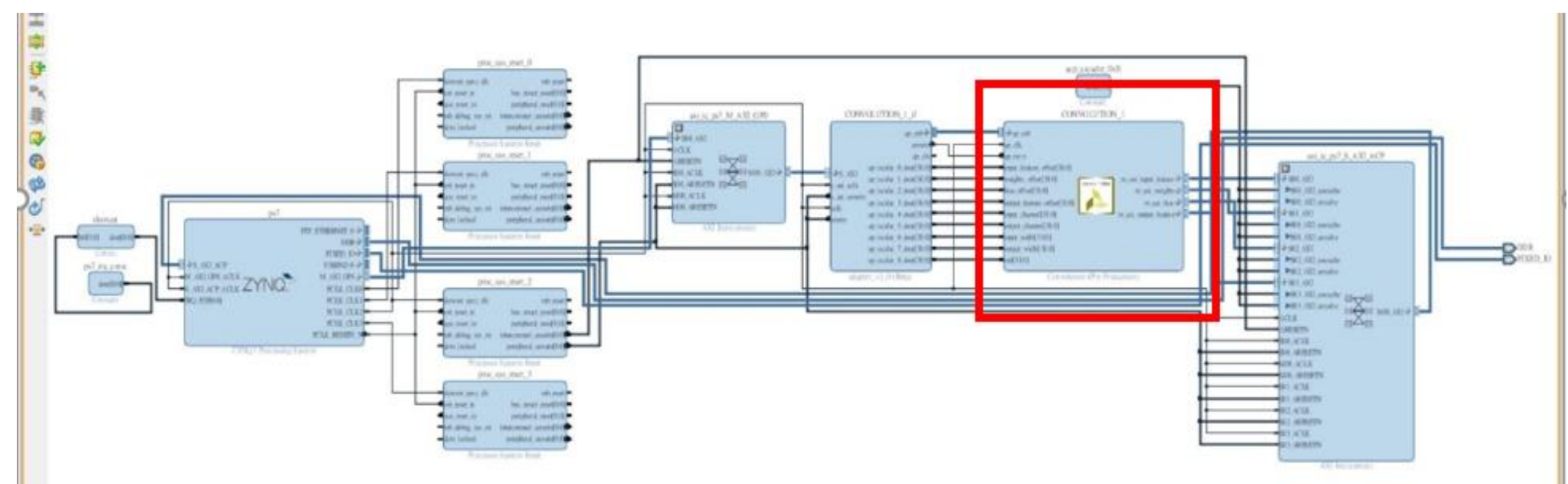
軟體設計研究結果:

```
batch : 1 test img num : 10000
-----
Read MNIST DATA...
MNIST DATA is loaded
Read MNIST Label...
MNIST Label is loaded
Load models
model loaded
SW test start
SW test completed
accuracy : 9863/10000
Test Completed
SW execution time : 334.382 seconds
C1 : 108.74 seconds
C2 : 160.048 seconds
C3 : 47.0172 seconds
```

系統設計(一)HLS合成結果:

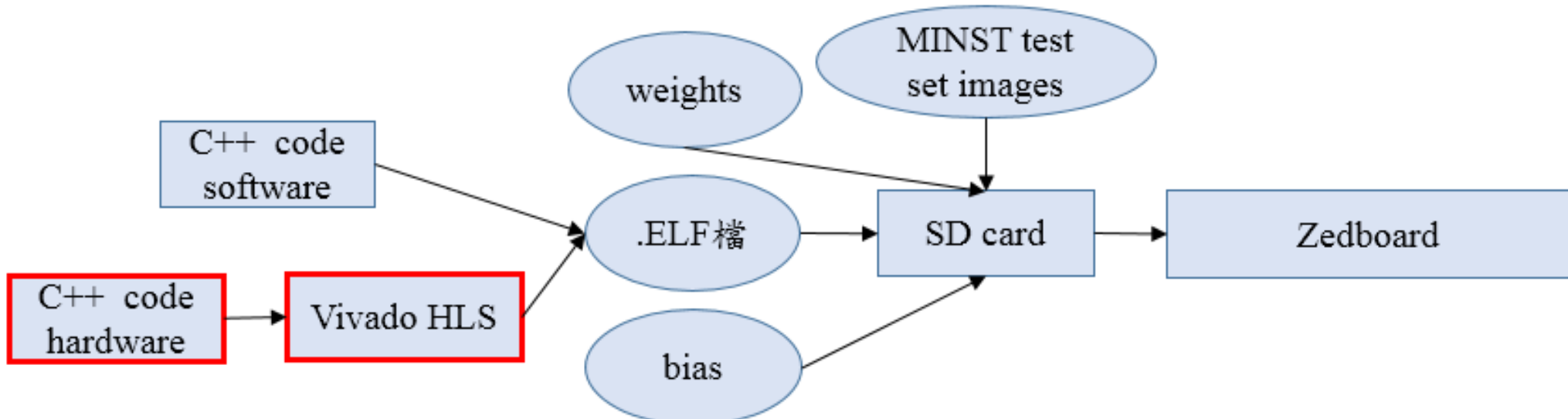


系統設計(二)HLS合成結果:



#### 系統方法與設計

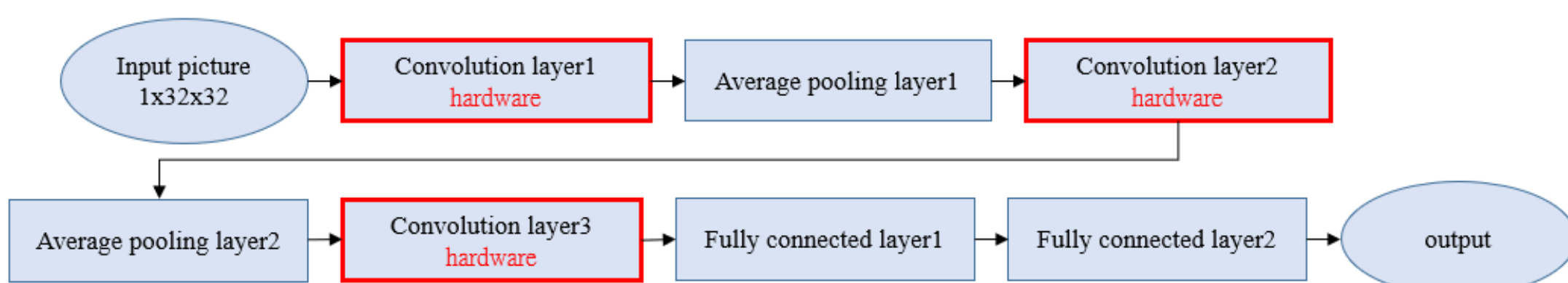
軟體與硬體設計實作流程:



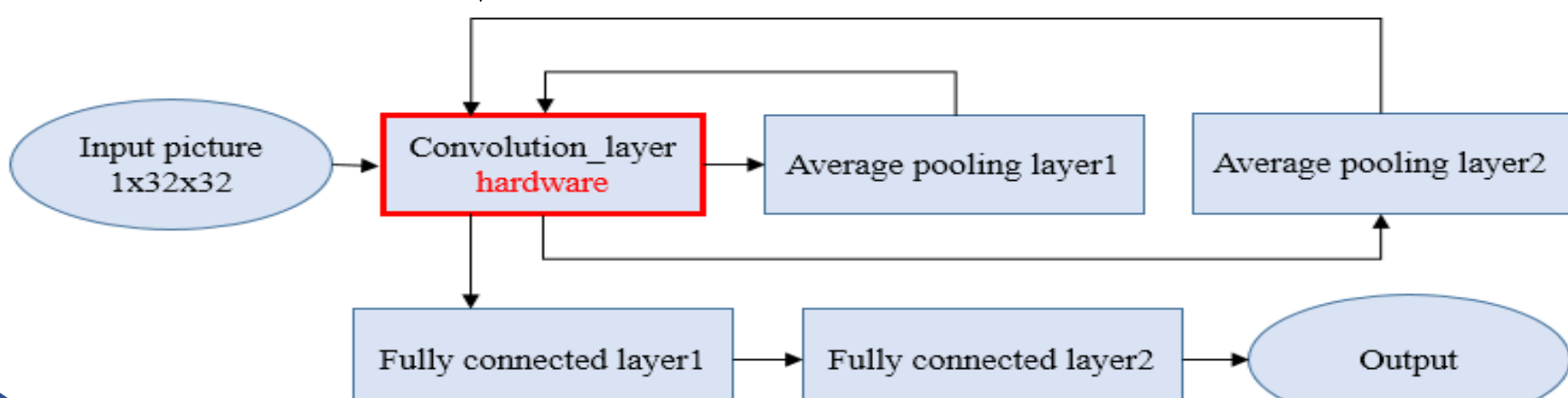
軟體設計結果分析:

Address	% Exclusive	Function	File
00013a...	46.2	CONVOLUTION_LAYER_2_SW	image_convolution_sw.h
000137...	28.3	CONVOLUTION_LAYER_1_SW	image_convolution_sw.h
00013e...	13.7	CONVOLUTION_LAYER_3_SW	image_convolution_sw.h
b6d493...	4.31	expm1f	
b6d4a0...	2.70	tanhf	
000140...	1.98	FULLY_CONNECTED_LAYER_1_S...	image_fullyconnected_sw

硬體系統設計(一):



硬體系統設計(二):



合成報告及辨識速度比較:

	執行時間	BRAM_18K	DSP48E	FF	LUT	DRAM (Kbits)
系統設計(一)	25.93 秒	190	205	31169	44997	291.776
系統設計(二)	166.41 秒	165	70	34228	18693	153.536
系統設計(二) 相較於 系統設計(一)	增加為 6.42 倍	減少 13.16%	減少 65.85%	增加 9.81%	減少 58.46%	減少 47.38%

兩種CNN硬體加速器優缺點比較:

	優點	缺點
系統設計(一)	相較軟體及系統設計(二)執行時間較短，數字辨識速度較快。	相較系統設計(二)占用較多硬體資源以及 DRAM 的使用。
系統設計(二)	相較軟體執行時間較短，且使用較少的硬體資源，並減少 DRAM 的使用。	相較於系統設計(一)執行時間較久，數字辨識速度較慢。