

Homework 6 – Memory Generation for the CNN Accelerator

Handout: 2025/11/17

Due: 2025/12/29

1. 使用Verilog HDL和Memory Compiler/Vivado BRAM來設計一個DNN硬體加速器，架構圖如Fig 1.

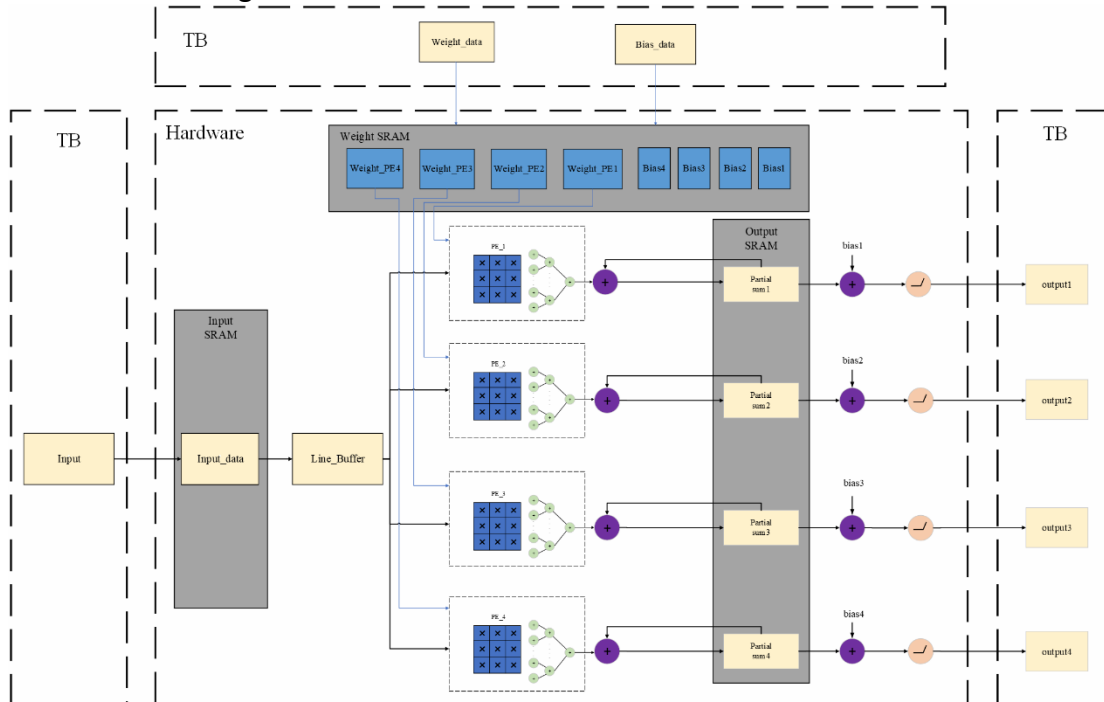


Fig 1. 整體硬體架構示意圖

2. 使用設計的DNN硬體加速器，運算Fig. 2的VGG第一層Convolution Layer，並將Input data進行Fig. 3的Padding後，使Convolution運算後的Output長和寬保持在56X56 透過Fig. 4的LineBuffer降低Input SRAM的讀取次數減少功耗，透過PE內的乘法器將Input和Weight進行相乘後，經過Adder Tree相加，並把Partial Sum存到Output SRAM，當輸出為最終輸出的時候，加上相對應的Bias後並經由Fig. 5的ReLU輸出，將輸出的每層Channel存成.bmp圖片檔，第一層Convolution輸出尺寸為56x56x64 共輸出64張.bmp圖片檔。
3. 提供檔案內容如下，資料擺放順序如Fig. 6，以下資料皆為16 Bit並為十六進制。
 1. Input (Input_data_hex.txt)

2. Weight (conv1_kernel_hex.txt)
3. Bias (conv1_bias_hex.txt)

Tabel 1. VGG-16 第一層CONV Layer

#Input Ch.	#Output Ch.	Input Size	Kernel Size	Stride

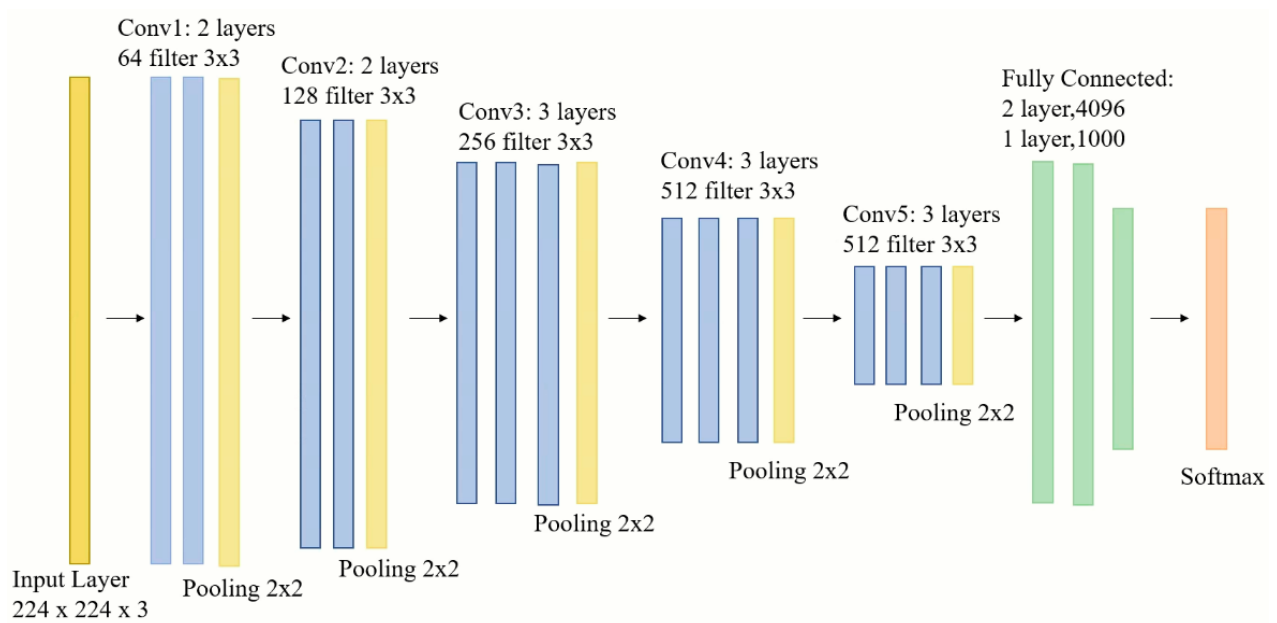


Fig 2. VGG-16

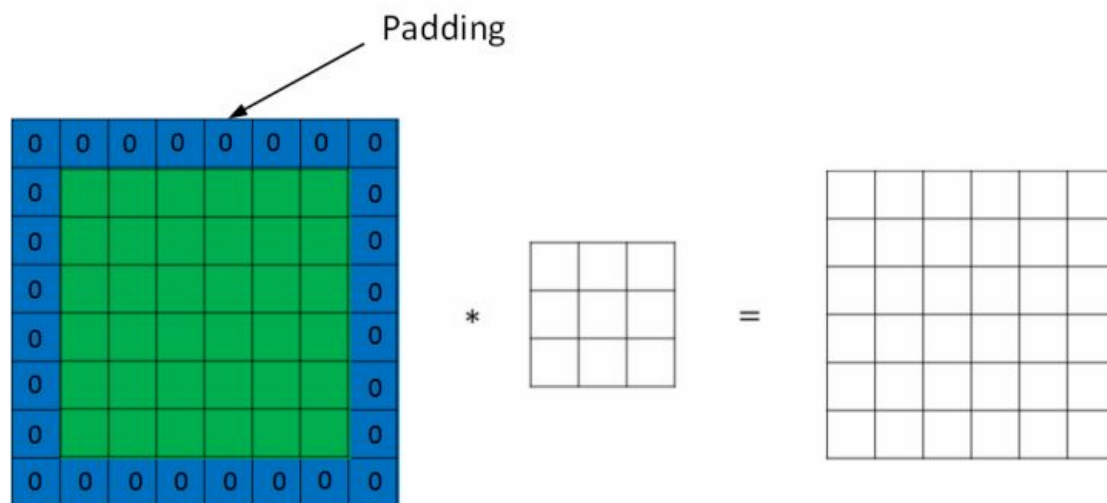


Fig 3. Padding



※ 〈注意〉：綠框部分是會需要被跳過的

Fig 4. LineBuffer

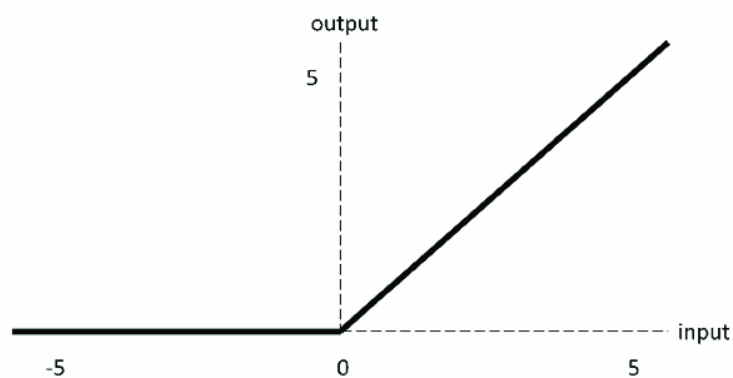


Fig 5. ReLU : 輸入小於等於0時，輸出則為0；輸入大於0時，輸出等於輸入。

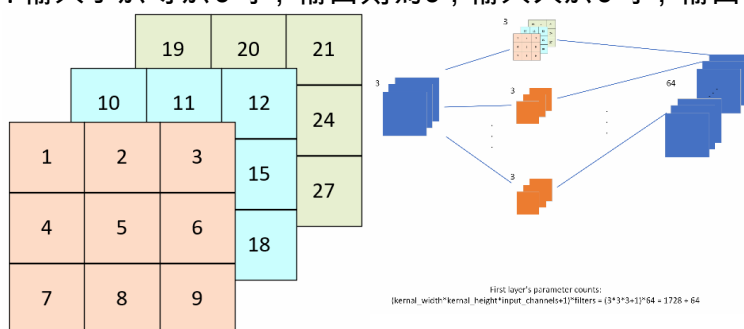
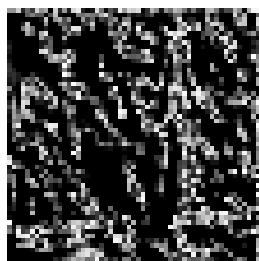
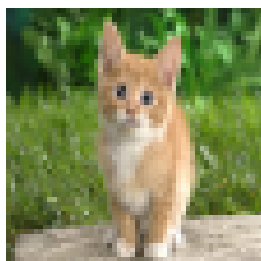


Fig 6. txt檔內Input Weight擺放順序



(a) (b)
Fig 7. (a)輸入圖片(先在TB將224x224轉為56x56)
和(b)Conv1 Channel 1的輸出圖片(56x56)

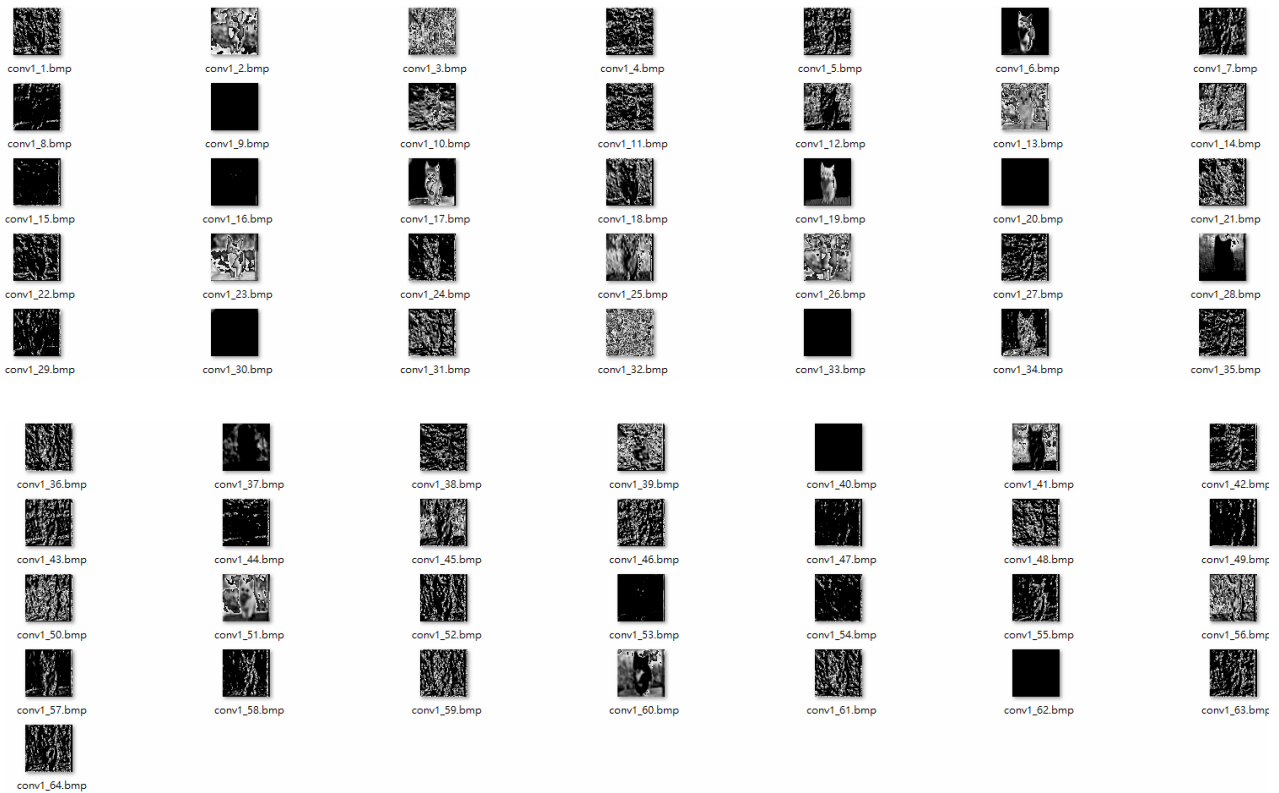


Fig 8. Convolution第一層輸出圖片

Report Requirement

1. Design Compiler (70%)

檔案類須含有:

- I. Testbench(10%)
- II. Verilog RTL code & Gate-level code(20%)
 - SRAM(10%)
 - LineBuffer (2.5%)
 - PE (2.5%)
 - Adder Tree(2.5%)
 - ReLU(2.5%)

III. Memory Compiler (5%)

以下全都有繳交才有5分，否則0分。

-Input SRAM.v & Input SRAM.db 檔

-Weight SRAM.v & Weight SRAM.db 檔

-Output SRAM.v & Output SRAM.db 檔

IV. Image

-Conv1 image *64(10%)

word報告須含有:

V. 架構圖解釋(5%)

VI. Area資訊和critical path資訊(10%)

VII. 心得(10%)

2. Xilinx Vivado(30%)

Project name需為: HDL_HW6_MXXXXXXXXXX

檔案類須含有:

I. HDL_HW6_MXXXXXXXXXX.xpr.zip (5%)

II. 需使用到BRAM IP(10%)

III. xdc(5%)

IV. wcfg(5%)

word報告類須含有:

V. Project Summary-Overview截圖(5%)

以上打包成HDL_HW6_MXXXXXXXXXX.zip壓縮檔並繳交