Project Title: **Convolutional Neural Network Accelerator for Image Denoising**	
Name: 陳奕帆	ID: 110062612
Name: 小山翼	ID: 110062526

Project Description:

1. 應用主題

We are going to apply DNN accerelator on one of the fundamental challenges in the field of image processing and computer vision, image denoising. The underlying goal in image denoising is to estimate the original image by suppressing noise from a noise-contaminated version of the image. This technique plays an important role in a wide range of applications such as image restoration, visual tracking, image registration, image segmentation, and image classification, where obtaining the original image content is crucial for strong performance **[1]**.

(a) (b)

Fig 1. noise-contaminated image (a) and the image after denoising (b)

2. 軟硬體架構規劃

Software

A. Quantization

In this project, we will try quantization aware training other than post quantization training which we use in the previous works. The reason is that the accuracy drop in a lossy process which we move to a lower precision from float can be minimized with the help of quantization aware training **[2]**.

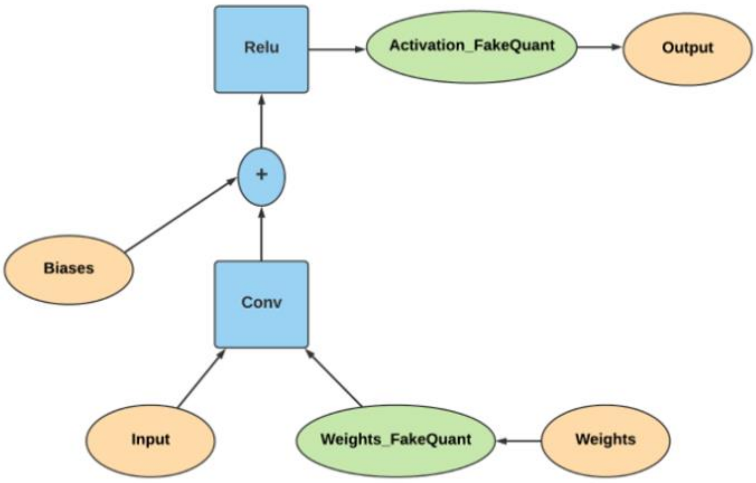

Fig 2. Visual Representation of a quantization aware training graph.

B. Denoise CNN

We are going to design a CNN model for image denoising **[3]**. The model architecture is showed in Fig .1. There are three types of layers. (1) Conv+ReLU: for the first layer. (2) Conv+BN+ReLU: for layers 2 ~ (D-1), where D is an user defined parameters. The batch normalization can make training faster and more stable by recentering and rescaling the layers' inputs. (3) Conv: for the last layer. Besides, we use residual learning to improve performanceby adding the input image with the last layer output.

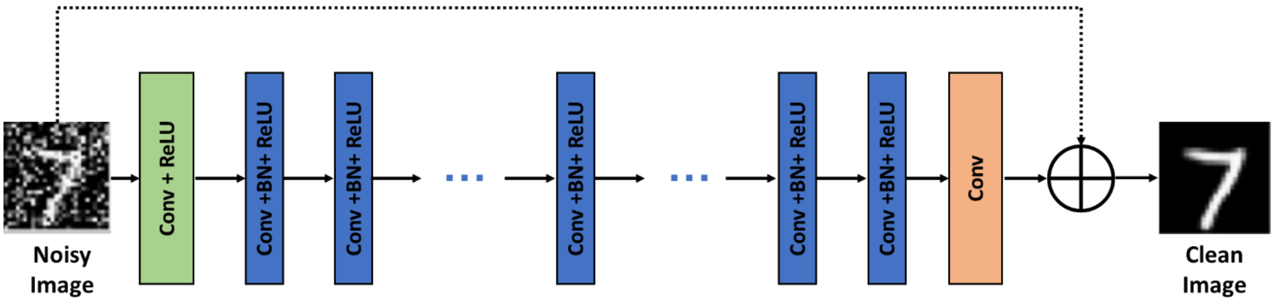

Fig 3. Denoise CNN architecture.

C. Residual learning

Consider a clean image $\mathbf{x}$ corrupted by noise $\mathbf{v}$. Noisy image $\mathbf{y} = \mathbf{x} + \mathbf{v}$. Our goal is to recovering $\mathbf{x}$ can be formulated as finding a parametric function $\hat{\mathbf{x}} = \mathbf{F}(\mathbf{y}; \boldsymbol{\theta})$. Instead of directly learning $\hat{x}$, we are going to learn the residual mapping $\mathbf{H}(\mathbf{y}; \boldsymbol{\theta}) \approx -\mathbf{v}$. Therefore, the denoising parametric model based on learning can be reformulated as $\boldsymbol{F}(\boldsymbol{y}; \boldsymbol{\theta}) = \boldsymbol{H}(\boldsymbol{y}; \boldsymbol{\theta}) + \boldsymbol{y}$. Several studies show that the residual mapping is much easier to be learned than the original unreferenced mapping. Therfore, we use residual path in our model to improve training speed and the performance.

Hardware

A. Accelerator Architecture using Row Stationary

Our accelerator is consisted of one weight SRAM, one activation SRAM, and one PE Array. Each PE calculates 1D-primitive convolution. The 5x4 PE Array can processed the convolution of 8x8 input feature map at one time. We use row stationary **[5]** to maximize data reuse. Filters are reused horizontally, input activations are reused diagonally, and output activations are using vertically.

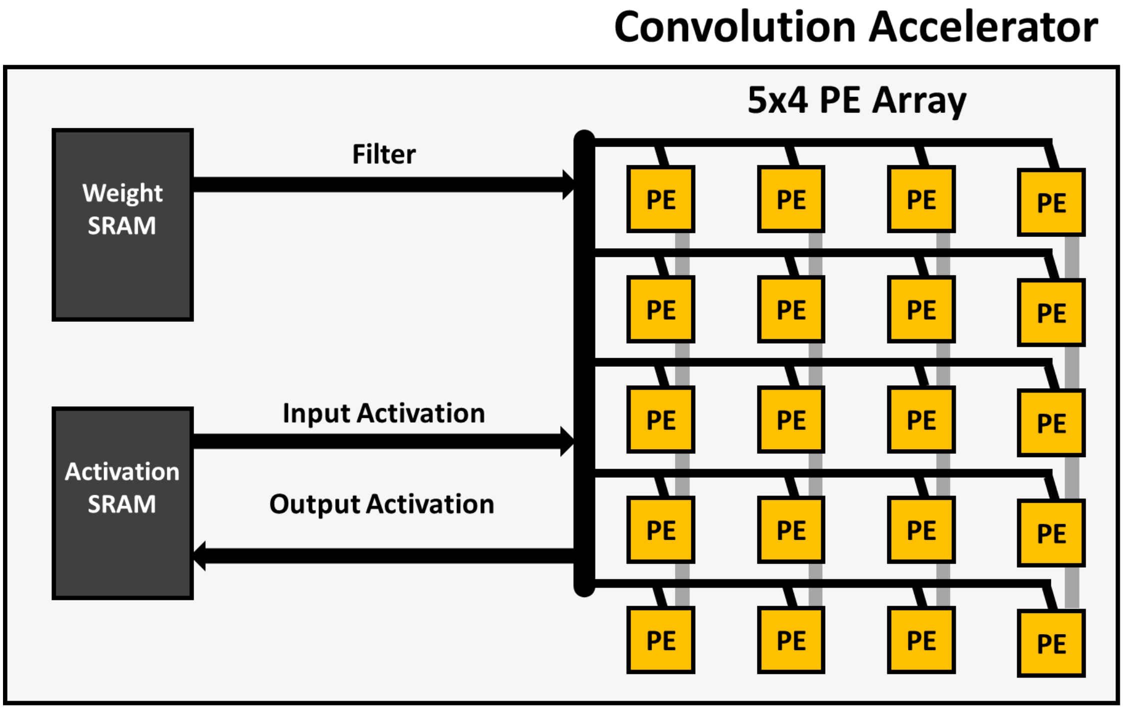

Fig 4. Convolution Accelerator Architecture

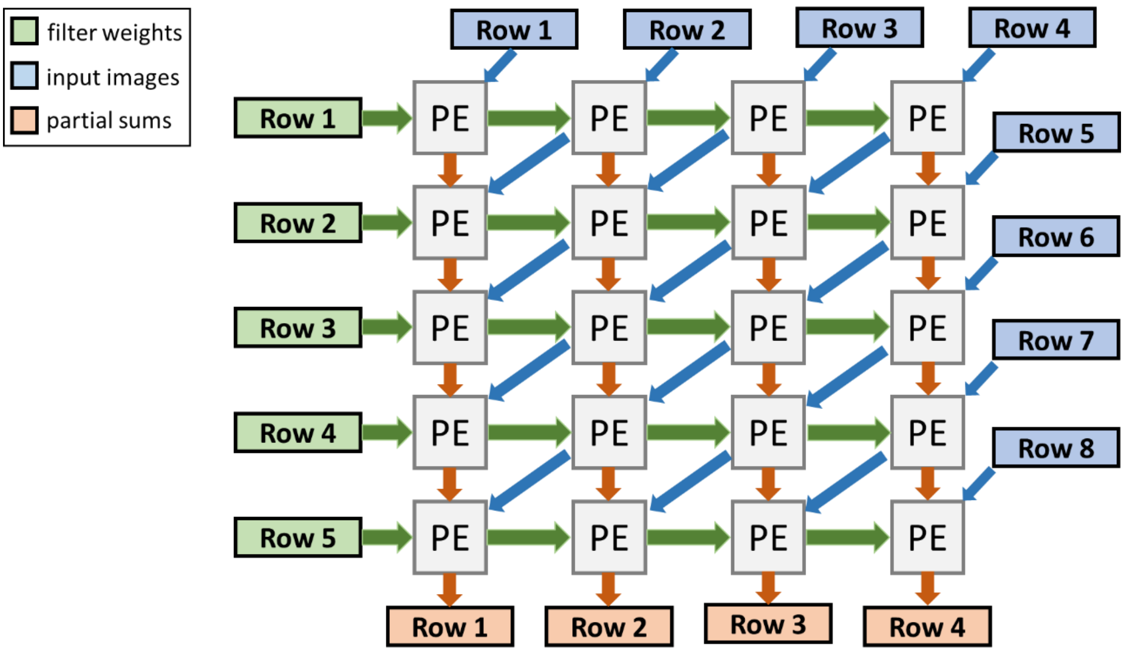

Fig 5. 5x4 PE Array

B. Ring streaming dataflow

Fig. 4 shows the difference between normal streaming dataflow and our ring streaming dataflow **[4]**. For normal streaming dataflow, when the input pixel reaches the end of current row, it returns to the front of next row. As a result, latencies appear across the rows. Conversely, the proposed ring streaming dataflow reduces unnecessary computations to improve computational performance. As the input pixel reaches to the end of current row, it moves down to the next row and returns to the front of that row. Hence, no latency is produced across the rows and computational performance is improved.

(a) Normal streaming dataflow (b) Proposed ring streaming dataflow

Fig 6. The illustration of (a) normal streaming dataflow and (b) proposed ring streaming dataflow.

3. 預期效果

Our goal is to recover clean image from noisy image. This is a regression problem. Therefore, we will use mean absolute error(MAE) and mean square error(MSE) as metrices. We are going to minimize MAE and MSE to let predicted images more similar to original images. Besides, we will use peak signal-to-noise ratio(PSNR) as metric, which is commonly used to quantify reconstruction quality for images and videos. Besides, we will use the clock rate, cycle count, area, power consumption to evaluate the performance of our accelerator. We will compare the denoise model running in accelerator with running in bare metal C program.

3. 参考資料

[1] Image Denoising | Vision and Image Processing Lab, https://uwaterloo.ca/vision-image-processing-lab/research-demos/image-denoising

[2] Inside Quantization Aware Training | by Vaibhav Nandwani, nside Quantization Aware Training | by Vaibhav Nandwani, https://towardsdatascience.com/inside-quantization-aware-training-4f91c8837ead

[3] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, Lei Zhang, “Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising”, https://arxiv.org/abs/1608.03981

[4] L. -C. Hsu, C. -T. Chiu and K. -T. Lin, "An Energy-Aware Bit-Serial Streaming Deep Convolutional Neural Network Accelerator," 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 4609-4613, doi: 10.1109/ICIP.2019.8803526.

[5] Y. -H. Chen, T. Krishna, J. S. Emer and V. Sze, "Eyeriss: An Energy-Efficient Reconfigurable Accelerator for Deep Convolutional Neural Networks," in IEEE Journal of Solid-State Circuits, vol. 52, no. 1, pp. 127-138, Jan. 2017, doi: 10.1109/JSSC.2016.2616357.