Capstone Design Presentation
IMC

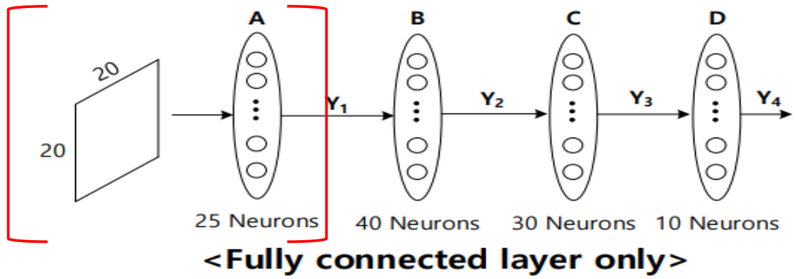
김원

Problem 1

Total Parameter(weight, bias) 구하기

1. 아래 그림과 같이 Network 를 구성할 때 Parameter 수를 구하라.

(모든 layer의 Activation 은 ReLU를 사용하며, Batch normalization 이 적용되어 있다)



<Fully connected layer only>

$$Parameters(at\ [\,]) = (400+1) * 25 = 10{,}025$$

<Input data>

$$\begin{bmatrix} x_1 & \cdots & x_{20} \\ \vdots & \ddots & \vdots \\ x_{380} & \cdots & x_{400} \end{bmatrix}$$ 20x20 행렬

Using row flatten(), flat[i]=$x_{\left[\frac{i}{N}\right].imodN}$

몫, 나머지

$$\mathbb{W}_A = \begin{bmatrix} w_{A1\,1} & \cdots & w_{A1\,25} \\ \vdots & \ddots & \vdots \\ w_{A400\,1} & \cdots & w_{A400\,25} \end{bmatrix}, \mathbb{B}_A = \begin{bmatrix} b_{A1} \\ \vdots \\ b_{A25} \end{bmatrix}$$

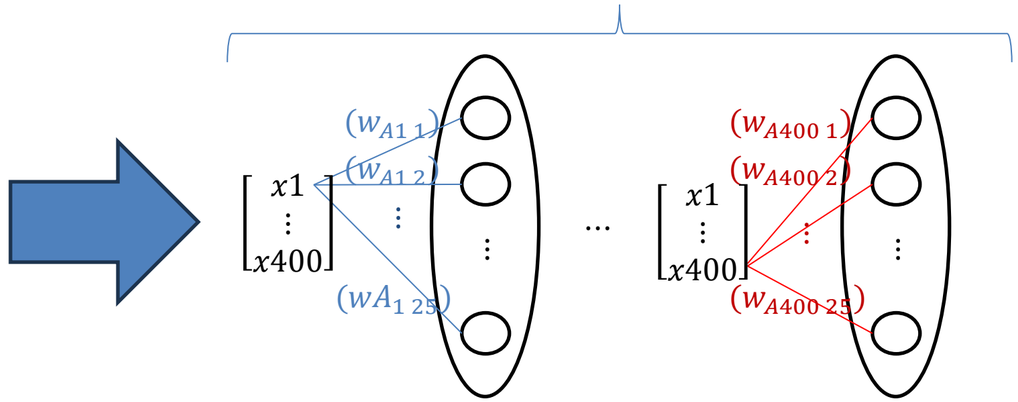
B벡터는 각각의 뉴런에 해당하는 바이어스

400x25 25x1

$$[x1 \quad \cdots \quad x400]$$ 1x400

Transpose

$$\mathrm{X} = \begin{bmatrix} x_1 \\ \vdots \\ x_{400} \end{bmatrix}$$ 400x1



$$\boldsymbol{Total\ Parameters =}$$
$$\boldsymbol{(400+1)*25}$$
$$\boldsymbol{+(25+1)*40}$$
$$\boldsymbol{+(40+1)*30}$$
$$\boldsymbol{+(30+1)*10}$$
$$\boldsymbol{= \underline{12{,}605개}}$$

*Total Parameters=(input+1)M1+(M1+1)M2+(M2+1)M3+(M3+1)M4

Problem 2, 3 (1/2)

곱셈, 덧셈 횟수 및 행렬 표현 구하기

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_{400} \end{bmatrix}, \ \mathbb{W}_A = \begin{bmatrix} w_{A1\,1} & \cdots & w_{A1\,25} \\ \vdots & \ddots & \vdots \\ w_{A400\,1} & \cdots & w_{A400\,25} \end{bmatrix}, \mathbb{B}_A = \begin{bmatrix} b_{A1} \\ \vdots \\ b_{A25} \end{bmatrix}, \ \mathbb{W}_B = \begin{bmatrix} w_{B1\,1} & \cdots & w_{B1\,40} \\ \vdots & \ddots & \vdots \\ w_{B25\,1} & \cdots & w_{B25\,40} \end{bmatrix}, \mathbb{B}_B = \begin{bmatrix} b_{B1} \\ \vdots \\ b_{B40} \end{bmatrix}$$

400x1 / 400x25 / 25x1 / 25x40 / 40x1

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A_1} \\ \vdots \\ \boldsymbol{A_{25}} \end{bmatrix} = \mathbb{W}_{\boldsymbol{A}}{}^{\boldsymbol{T}} \times \boldsymbol{X} + \mathbb{B}_{\boldsymbol{A}} = \begin{bmatrix} w_{A1\,1} & \cdots & w_{A400\,1} \\ \vdots & \ddots & \vdots \\ w_{A1\,25} & \cdots & w_{A400\,25} \end{bmatrix} \times \begin{bmatrix} x_1 \\ \vdots \\ x_{400} \end{bmatrix} + \begin{bmatrix} b_{A1} \\ \vdots \\ b_{A25} \end{bmatrix} = \begin{bmatrix} w_{A1\,1} * x_1 + \cdots + w_{A400\,1} * x_{400} + b_{A1} \\ \vdots \\ w_{A1\,25} * x_1 + \cdots + w_{A400\,25} * x_{400} + b_{A25} \end{bmatrix} = \begin{bmatrix} A_1 \\ \vdots \\ A_{25} \end{bmatrix}$$

25x400 / 400x1 / 25x1 / 25x1 / 25x1

A layer에 존재하는 각각의 뉴런에 저장된 값

$$w_{A1\,1} * x_1 + \cdots + w_{A400\,1} * x_{400} + b_{A1}$$

한 행당, 곱 = 400, 덧셈 = 400

A layer에서 총 곱 횟수 = 10,000회, 총 덧셈 횟수 = 10,000회

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B_1} \\ \vdots \\ \boldsymbol{B_{40}} \end{bmatrix} = \mathbb{W}_{\boldsymbol{B}}{}^{\boldsymbol{T}} \times \boldsymbol{Y_1} + \mathbb{B}_{\boldsymbol{B}} = \begin{bmatrix} w_{B1\,1} & \cdots & w_{B25\,1} \\ \vdots & \ddots & \vdots \\ w_{B1\,40} & \cdots & w_{25\,40} \end{bmatrix} \times \begin{bmatrix} f(A_1) \\ \vdots \\ f(A_{25}) \end{bmatrix} + \begin{bmatrix} b_{B1} \\ \vdots \\ b_{B40} \end{bmatrix} = \begin{bmatrix} w_{B1\,1} * f(A_1) + \cdots + w_{B25\,1} * f(A_{25}) + b_{B1} \\ \vdots \\ w_{B1\,40} * f(A_1) + \cdots + w_{B25\,40} * f(A_{25}) + b_{B40} \end{bmatrix} = \begin{bmatrix} B_1 \\ \vdots \\ B_{40} \end{bmatrix}$$

40x25 / 25x1 / 40x1 / 40x1

B layer에 존재하는 각각의 뉴런에 저장된 값

$$w_{B1\,1} * f(A_1) + \cdots + w_{B25\,1} * f(A_{25}) + b_{B1}$$

한 행당, 곱 = 25, 덧셈 = 25

B layer에서 총 곱 횟수 = 1,000회, 총 덧셈 횟수 = 1,000회

$$Y_1 = f(A) = f(\mathbb{W}_A{}^T \cdot \mathbb{X} + \mathbb{B}_A)$$

Problem 2, 3 (2/2)

곱셈, 덧셈 횟수 및 행렬 표현 구하기

$$\mathbb{W}_C = \begin{bmatrix} w_{C1\,1} & \cdots & w_{C1\,30} \\ \vdots & \ddots & \vdots \\ w_{C40\,1} & \cdots & w_{C40\,30} \end{bmatrix}, \mathbb{B}_C = \begin{bmatrix} b_{C1} \\ \vdots \\ b_{C30} \end{bmatrix}, \mathbb{W}_D = \begin{bmatrix} w_{D1\,1} & \cdots & w_{D1\,10} \\ \vdots & \ddots & \vdots \\ w_{D30\,1} & \cdots & w_{D30\,10} \end{bmatrix}, \mathbb{B}_D = \begin{bmatrix} b_{D1} \\ \vdots \\ b_{D10} \end{bmatrix}$$

40x30 30x1 30x10 10x1

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C_1} \\ \vdots \\ \boldsymbol{C_{30}} \end{bmatrix} = \mathbb{W}_C{}^T \times \boldsymbol{Y_2} + \mathbb{B}_C = \begin{bmatrix} w_{C1\,1} & \cdots & w_{C40\,1} \\ \vdots & \ddots & \vdots \\ w_{C1\,30} & \cdots & w_{C40\,30} \end{bmatrix} \times \begin{bmatrix} f(B_1) \\ \vdots \\ f(B_{40}) \end{bmatrix} + \begin{bmatrix} b_{C1} \\ \vdots \\ b_{C30} \end{bmatrix} = \begin{bmatrix} w_{C1\,1} * f(B_1) + \cdots + w_{C40\,1} * f(B_{40}) + b_{C1} \\ \vdots \\ w_{C1\,30} * f(B_1) + \cdots + w_{C40\,30} * f(B_{40}) + b_{C30} \end{bmatrix} = \begin{bmatrix} C_1 \\ \vdots \\ C_{30} \end{bmatrix}$$

30x40 40x1 30x1 30x1 30x1

C layer에 존재하는 각각의 뉴런에 저장된 값

$$w_{C1\,1} * f(B_1) + \cdots + w_{C40\,1} * f(B_{40}) + b_{C1}$$

한 행당, 곱 = 40, 덧셈 = 40

C layer에서 총 곱 횟수 = 1,200회, 총 덧셈 횟수 = 1,200회

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{D_1} \\ \vdots \\ \boldsymbol{D_{10}} \end{bmatrix} = \mathbb{W}_D{}^T \times \boldsymbol{Y_3} + \mathbb{B}_D = \begin{bmatrix} w_{D1\,1} & \cdots & w_{D30\,1} \\ \vdots & \ddots & \vdots \\ w_{D1\,10} & \cdots & w_{D30\,10} \end{bmatrix} \times \begin{bmatrix} f(C_1) \\ \vdots \\ f(C_{30}) \end{bmatrix} + \begin{bmatrix} b_{D1} \\ \vdots \\ b_{D10} \end{bmatrix} = \begin{bmatrix} w_{D1\,1} * f(C_1) + \cdots + w_{D30\,1} * f(C_{30}) + b_{D1} \\ \vdots \\ w_{D1\,10} * f(C_1) + \cdots + w_{D30\,10} * f(C_{30}) + b_{D10} \end{bmatrix} = \begin{bmatrix} D_1 \\ \vdots \\ D_{10} \end{bmatrix}$$

10x30 30x1 10x1 10x1 10x1

D layer에 존재하는 각각의 뉴런에 저장된 값

$$w_{D1\,1} * f(C_1) + \cdots + w_{D30\,1} * f(C_{30}) + b_{D1}$$

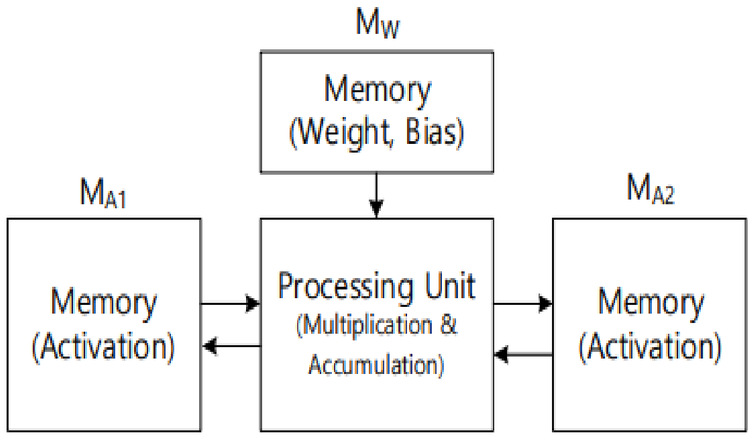
한 행당, 곱 = 30, 덧셈 = 30

D layer에서 총 곱 횟수 = 300회, 총 덧셈 횟수 = 300회

$$Y_2 = f(B) = f(\mathbb{W}_B{}^T \cdot Y_1 + \mathbb{B}_B),\ Y_3 = f(C) = f\left(\mathbb{W}_C{}^T \cdot Y_2 + \mathbb{B}_C\right),\ Y_4 = f(D) = f(\mathbb{W}_D{}^T \cdot Y_3 + \mathbb{B}_D)$$

Problem 4

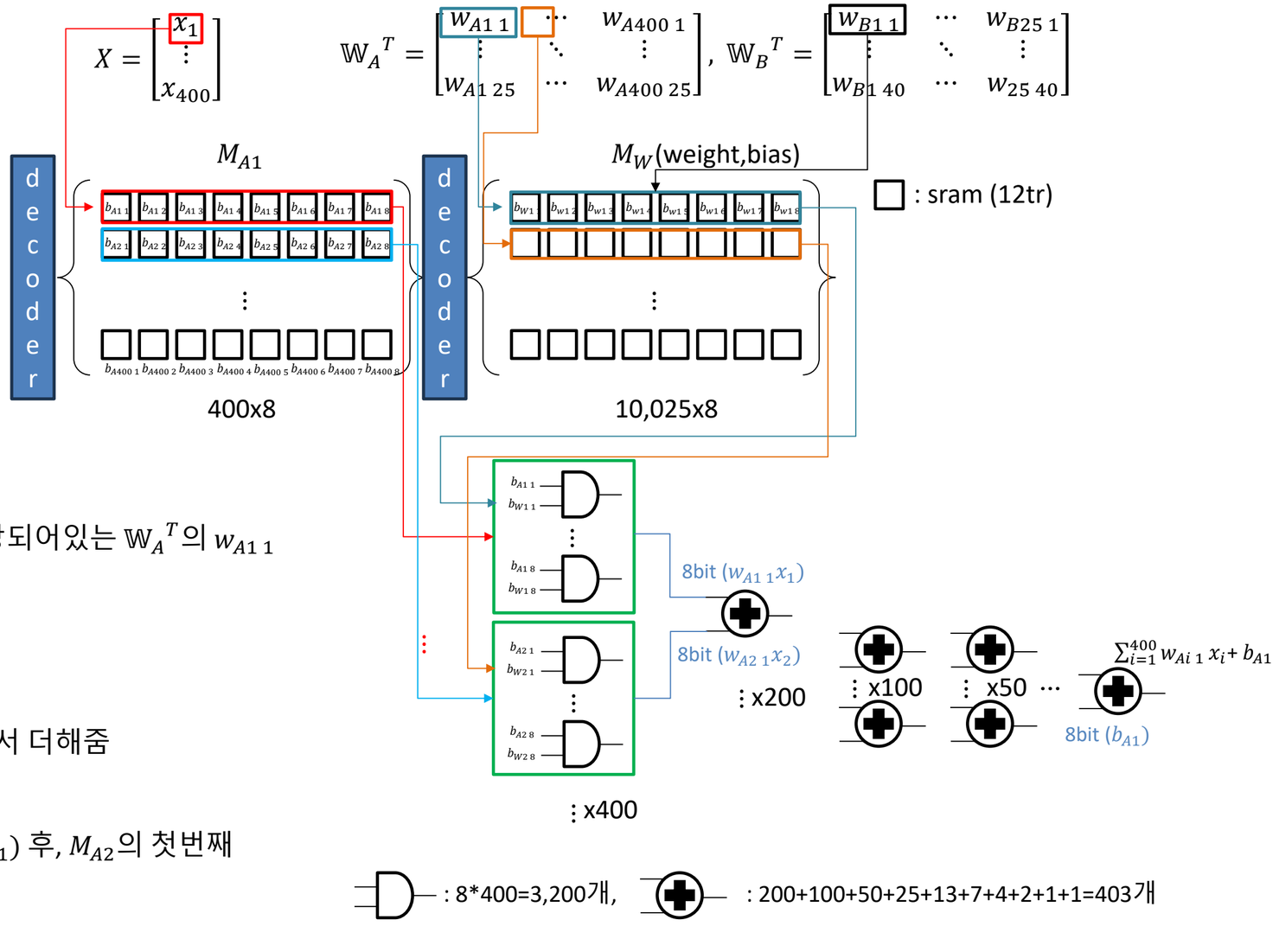
Hardware 크기 및 동작 원리



M_{A1} : 3,200bit, M_W : 80,200bit, M_{A2} : 240bit(30x8b)

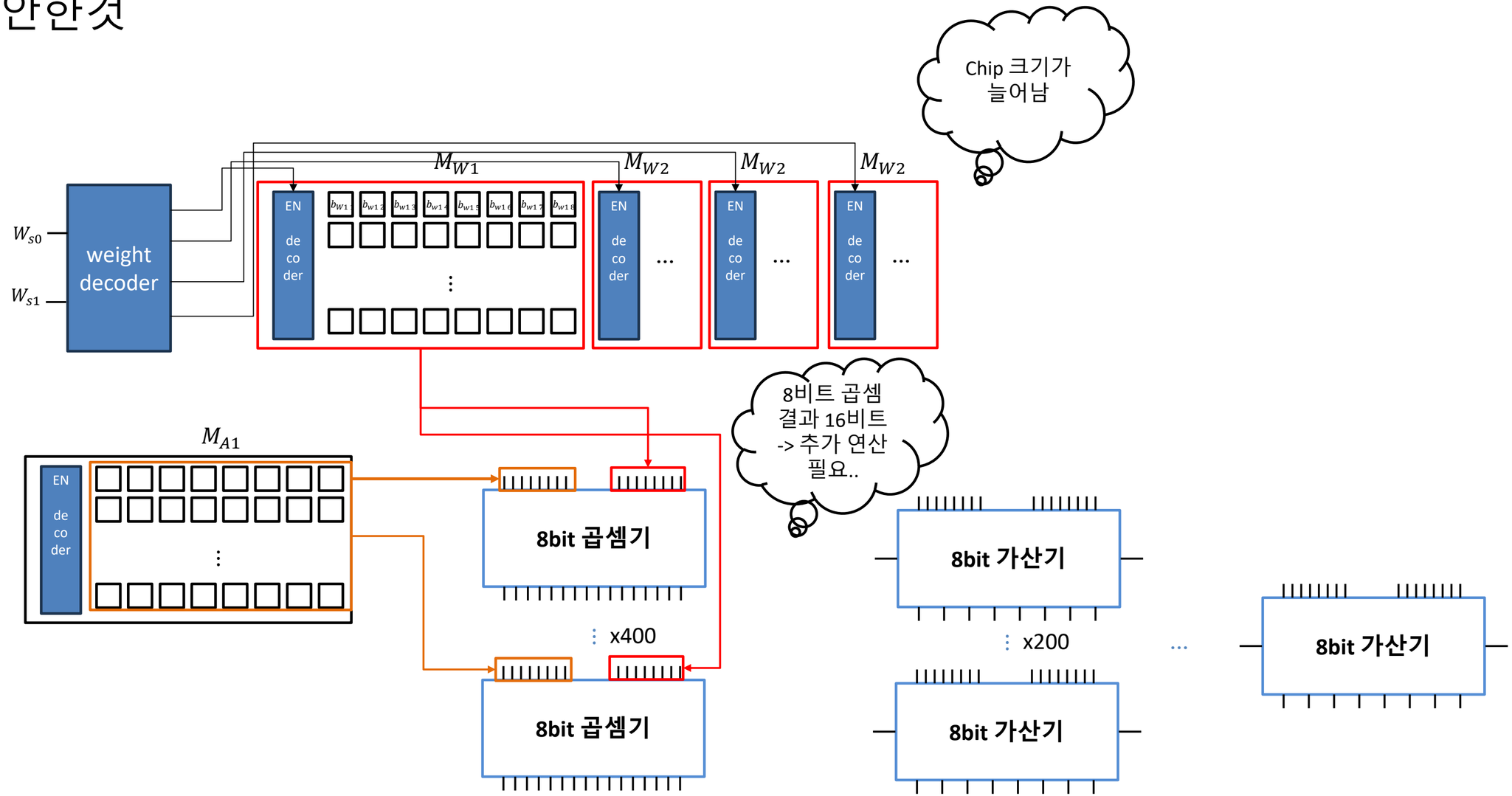
1. M_{A1}에 Input data(X [400x1]) write, M_w에 $\mathbb{W}_A{}^T$ [25x400] data write
2. Processing Unit(logic), M_{A1} 에 첫행에 저장된 x_1 (8bit) & M_W에 저장되어있는 $\mathbb{W}_A{}^T$의 $w_{A1\,1}$ 이 저장되어 있는 1행 8bit 을 read
3. 후 $\mathbb{W}_B{}^T$의 $w_{B1\,1}$값 write
4. 읽어온 bit를 각각 and 연산
5. 각각의 and 연산된 bit를 sum 해주기위해 8비트 가산기로 누적해서 더해줌
6. 마지막 가산기에서 첫번째 뉴런의 바이어스 값인 b_{A1} 을 더해줌
7. 마지막 가산기에서 나온 출력(A_1= $\sum_{i=1}^{400} w_{Ai\,1}\,x_i$+ b_{A1})을 ReLU, $f(A_1)$ 후, M_{A2}의 첫번째 행에 write
8. 2~7을 반복하여 A [25x1] 행렬을 M_{A2}에 write

<구조>



Problem 4

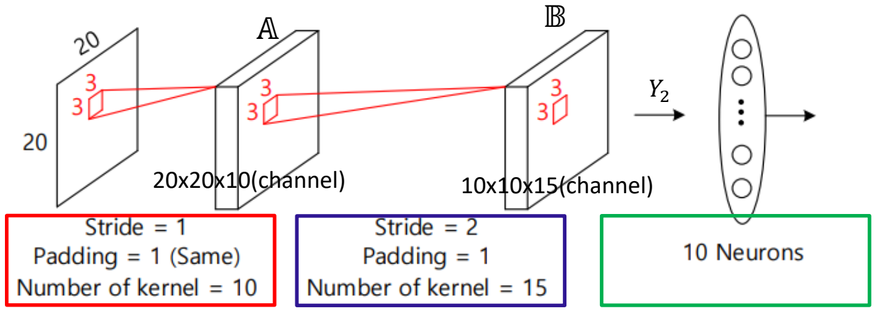
처음 고안한것



Problem 5

1) Total Parameter 구하기

5. 아래와 같은 Convolutional neural network 에 대해서 위 1~4 를 반복하라.



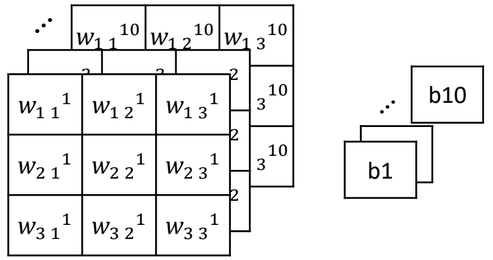
<Convolutional Neural Network>

Cf)

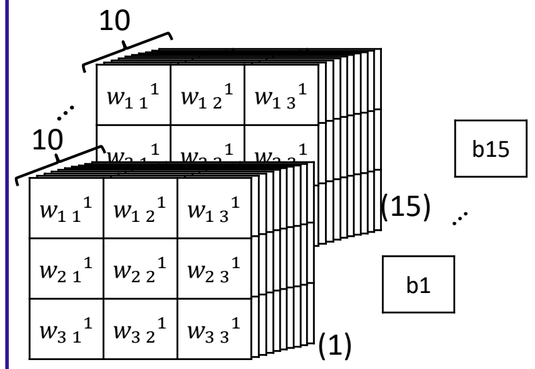
$$H_{out} = \frac{H_{in} + 2P - K}{Stride} + 1$$

$$W_{out} = \frac{W_{in} + 2P - K}{Stride} + 1$$

$$\boldsymbol{Total\ Parameters} = (3*3*1)*10 + 10 + (3*3*10)*15 + 15 + (1350+1)*10 = \underline{16{,}475개}$$



Parameters = (3x3x1)x10+10 = 100



Parameters = (3x3x10)x15+15 = 1,365

$$Y_2 = \begin{bmatrix} y_1 \\ \vdots \\ y_{1500} \end{bmatrix} \quad 1500x1$$

$$\mathbb{W} = \begin{bmatrix} w_{1\,1} & \cdots & w_{1\,10} \\ \vdots & \ddots & \vdots \\ w_{1500\,1} & \cdots & w_{1500\,10} \end{bmatrix}, \mathbb{B} = \begin{bmatrix} b_1 \\ \vdots \\ b_{10} \end{bmatrix}$$

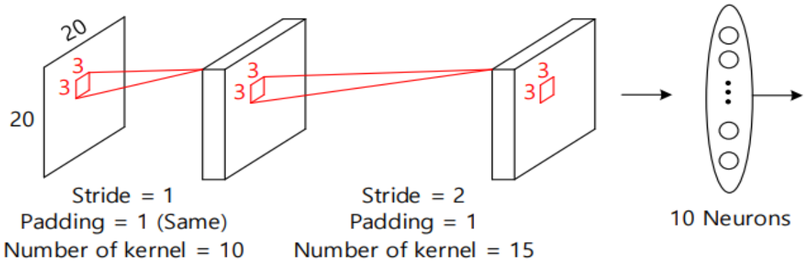
Parameters = (1500+1)x10 = 15,010

Parameters = (K x # of input channel) x # of kernel + # of kernel
- K : kernel size(kernel width * kernel length)

Problem 5

2,3) 곱셈, 덧셈 횟수 및 행렬 표현 구하기 (1/7)

5. 아래와 같은 Convolutional neural network 에 대해서 위 1~4 를 반복하라.



<Convolutional Neural Network>

1. Kernel size = 3x3, stride = 1에 맞게 padding이 적용된 input data의 3x3 capture 행렬을 $\mathbb{X}_f(i,j)$행렬 형태로 만듬 (im2col)
2. $\mathbb{X}_f(i,j)$행렬을 row flatten 후, Transpose
3. 각 행의 요소를 flatten

<Input data>

$$\mathbb{X} = \begin{bmatrix} x_{1\,1} & \cdots & x_{1\,20} \\ \vdots & \ddots & \vdots \\ x_{20\,1} & \cdots & x_{20\,20} \end{bmatrix}$$

20x20x(1channel) 행렬

Padding=1 적용

$$\mathbb{X}_p = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & x_{1\,1} & \cdots & x_{1\,20} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & x_{20\,1} & \cdots & x_{20\,20} & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

22x22x(1channel) 행렬

Using row flatten(), flat[i]=$x_{\left[\frac{i}{N}\right].imodN}$

$$X_f(0,0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & x_{1\,1} & x_{1\,2} \\ 0 & x_{2\,1} & x_{2\,2} \end{bmatrix} \rightarrow [0\ 0\ 0\ 0\ x_{1\,1}\ x_{1\,2}\ 0\ x_{2\,1}\ x_{2\,2}]$$ 1x9

$$X_f(0,1) = \begin{bmatrix} 0 & 0 & 0 \\ x_{1\,1} & x_{1\,2} & x_{1\,3} \\ x_{2\,1} & x_{2\,2} & x_{2\,3} \end{bmatrix} \rightarrow [0\ 0\ 0\ x_{1\,1}\ x_{1\,2}\ x_{1\,3}\ x_{2\,1}\ x_{2\,2}\ x_{2\,3}]$$ 1x9

$$\vdots \qquad\qquad \vdots$$

$$X_f(19,19) = \begin{bmatrix} x_{19\,19} & x_{19\,20} & 0 \\ x_{20\,19} & x_{20\,20} & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow [x_{19\,19}\ x_{19\,20}\ 0\ x_{20\,19}\ x_{20\,20}\ 0\ 0\ 0\ 0]$$ 1x9

Cf) $X_f(q,p) = \begin{bmatrix} X_p(q,p) & X_p(q,p+1) & X_p(q,p+2) \\ X_p(q+1,p) & X_p(q+1,p+1) & X_p(q+1,p+2) \\ X_p(q+2,p) & X_p(q+2,p+1) & X_p(q+2,p+2) \end{bmatrix}$

$$\mathbb{X}_f(i,j) = \begin{bmatrix} X_f(0,0) & X_f(0,1) & \cdots & X_f(0,19) \\ X_f(1,0) & X_f(1,1) & \cdots & X_f(1,19) \\ \vdots & \vdots & \ddots & \vdots \\ X_f(19,0) & X_f(19,1) & \cdots & X_f(19,19) \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} flatten(X_f(0,0)) \\ flatten(X_f(0,1)) \\ \vdots \\ flatten(X_f(19,19)) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & x_{1\,1} & x_{1\,2} & 0 & x_{2\,1} & x_{2\,2} \\ 0 & 0 & 0 & x_{1\,1} & x_{1\,2} & x_{1\,3} & x_{2\,1} & x_{2\,2} & x_{2\,3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{19\,19} & x_{19\,20} & 0 & x_{20\,19} & x_{20\,20} & 0 & 0 & 0 & 0 \end{bmatrix}$$ 400x9

$$= \begin{bmatrix} X_{f1} \\ X_{f2} \\ \vdots \\ X_{f400} \end{bmatrix} = \mathbb{X}_f$$

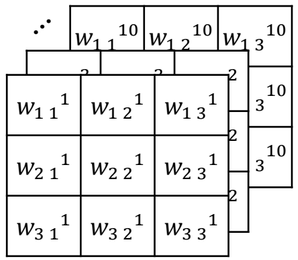
Cf) $X_{f1} = [0 \quad 0 \quad 0 \quad 0 \quad x_{1\,1} \quad x_{1\,2} \quad 0 \quad x_{2\,1} \quad x_{2\,2}]$

Problem 5

2,3) 곱셈, 덧셈 횟수 및 행렬 표현 구하기 (2/7)

1. Kernel 도 input data 처럼 flatten 하여 10x9 행렬 만듬
   이때, 행은 # of kernels 수와 같음

<kernel(weight)>



Using row flatten(), flat[i]=$w_{\left[\frac{i}{N}\right].imodN}$

몫, 나머지

$$\begin{bmatrix} {w_{1\,1}}^{1} & {w_{1\,2}}^{1} & {w_{1\,3}}^{1} \\ {w_{2\,1}}^{1} & {w_{2\,2}}^{1} & {w_{2\,3}}^{1} \\ {w_{3\,1}}^{1} & {w_{3\,2}}^{1} & {w_{3\,3}}^{1} \end{bmatrix} \rightarrow [\,{w_{1\,1}}^{1}\ {w_{1\,2}}^{1}\ {w_{1\,3}}^{1}\ {w_{2\,1}}^{1}\ {w_{2\,2}}^{1}\ {w_{2\,3}}^{1}\ {w_{3\,1}}^{1}\ {w_{3\,2}}^{1}\ {w_{3\,3}}^{1}]$$

1x9

$$\vdots \qquad\qquad \vdots$$

$$\begin{bmatrix} {w_{1\,1}}^{10} & {w_{1\,2}}^{10} & {w_{1\,3}}^{10} \\ {w_{2\,1}}^{10} & {w_{2\,2}}^{10} & {w_{2\,3}}^{10} \\ {w_{3\,1}}^{10} & {w_{3\,2}}^{10} & {w_{3\,3}}^{10} \end{bmatrix} \rightarrow [\,{w_{1\,1}}^{1}\ {w_{1\,2}}^{1}\ {w_{1\,3}}^{1}\ {w_{2\,1}}^{1}\ {w_{2\,2}}^{1}\ {w_{2\,3}}^{1}\ {w_{3\,1}}^{1}\ {w_{3\,2}}^{1}\ {w_{3\,3}}^{1}]$$

1x9

$$\mathbb{W} = \begin{bmatrix} {w_{1\,1}}^{1} & {w_{1\,2}}^{1} & {w_{1\,3}}^{1} & {w_{2\,1}}^{1} & {w_{2\,2}}^{1} & {w_{2\,3}}^{1} & {w_{3\,1}}^{1} & {w_{3\,2}}^{1} & {w_{3\,3}}^{1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ {w_{1\,1}}^{10} & {w_{1\,2}}^{10} & {w_{1\,3}}^{10} & {w_{2\,1}}^{10} & {w_{2\,2}}^{10} & {w_{2\,3}}^{10} & {w_{3\,1}}^{10} & {w_{3\,2}}^{10} & {w_{3\,3}}^{10} \end{bmatrix}$$

10x9

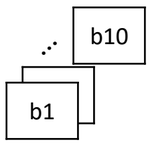
Transpose

Transpose

$$\mathbb{W}^{T} = \begin{bmatrix} {w_{1\,1}}^{1} & {w_{1\,1}}^{2} & \cdots & {w_{1\,1}}^{10} \\ \vdots & \vdots & \ddots & \vdots \\ {w_{3\,3}}^{1} & {w_{3\,3}}^{2} & \cdots & {w_{3\,3}}^{10} \end{bmatrix} = [W^{1} \quad W^{2} \quad \cdots \quad W^{10}] \qquad \text{Cf) } W^{i} = \begin{bmatrix} {w_{1\,1}}^{i} \\ \vdots \\ {w_{3\,3}}^{i} \end{bmatrix}$$

9x10

<kernel(bias)>



$$B = \begin{bmatrix} b^{1} \\ \vdots \\ b^{10} \end{bmatrix},\ B^{T} = [b^{1} \quad \cdots \quad b^{10}]$$

1x10

Broadcasting

$$B' = \begin{bmatrix} b^{1} & \cdots & b^{10} \\ \vdots & \ddots & \vdots \\ b^{1} & \cdots & b^{10} \end{bmatrix}$$

400x10

Problem 5

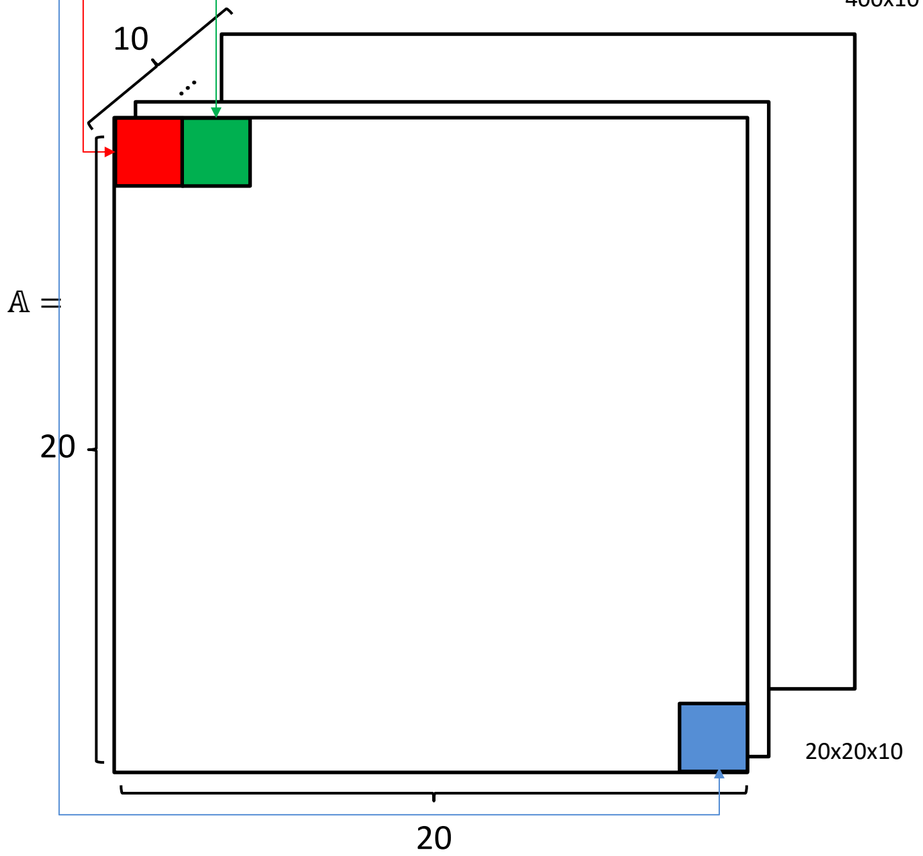
2,3) 곱셈, 덧셈 횟수 및 행렬 표현 구하기 (3/7)

$$\mathbb{X}_f \times \mathbb{W}^T + B' = \begin{bmatrix} 0 & 0 & 0 & 0 & x_{1\,1} & x_{1\,2} & 0 & x_{2\,1} & x_{2\,2} \\ 0 & 0 & 0 & x_{1\,1} & x_{1\,2} & x_{1\,3} & x_{2\,1} & x_{2\,2} & x_{2\,3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{20\,20} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} {w_{1\,1}}^1 & {w_{1\,1}}^2 & \cdots & {w_{1\,1}}^{10} \\ \vdots & \vdots & \ddots & \vdots \\ {w_{3\,3}}^1 & {w_{3\,3}}^2 & \cdots & {w_{3\,3}}^{10} \end{bmatrix} + \begin{bmatrix} b^1 & \cdots & b^{10} \\ \vdots & \ddots & \vdots \\ b^1 & \cdots & b^{10} \end{bmatrix} = \begin{bmatrix} 0 * {w_{1\,1}}^1 + \cdots + x_{2\,2} * {w_{3\,3}}^1 + b1 & \cdots & 0 * {w_{1\,1}}^{10} + \cdots + x_{2\,2} * {w_{3\,3}}^{10} + b10 \\ 0 * {w_{1\,1}}^1 + \cdots + x_{2\,3} * {w_{3\,3}}^1 + b1 & \cdots & 0 * {w_{1\,1}}^{10} + \cdots + x_{2\,3} * {w_{3\,3}}^{10} + b10 \\ \vdots & \ddots & \vdots \\ x_{20\,20} * {w_{1\,1}}^1 + \cdots + 0 * {w_{3\,3}}^1 + b1 & \cdots & x_{20\,20} * {w_{1\,1}}^{10} + \cdots + 0 * {w_{3\,3}}^{10} + b10 \end{bmatrix}$$

400x9 9x10 400x10 400x10

$$= \begin{bmatrix} X_{f1} \times W^1 + b1 & \cdots & X_{f1} \times W^{10} + b10 \\ X_{f2} \times W^1 + b1 & \cdots & X_{f2} \times W^{10} + b10 \\ \vdots & \ddots & \vdots \\ X_{f400} \times W^1 + b1 & \cdots & X_{f400} \times W^{10} + b10 \end{bmatrix} = \mathbb{A} = \begin{bmatrix} {A_1}^1 & \cdots & {A_1}^{10} \\ {A_2}^1 & \cdots & {A_2}^{10} \\ \vdots & \ddots & \vdots \\ {A_{400}}^1 & \cdots & {A_{400}}^{10} \end{bmatrix}$$

400x10



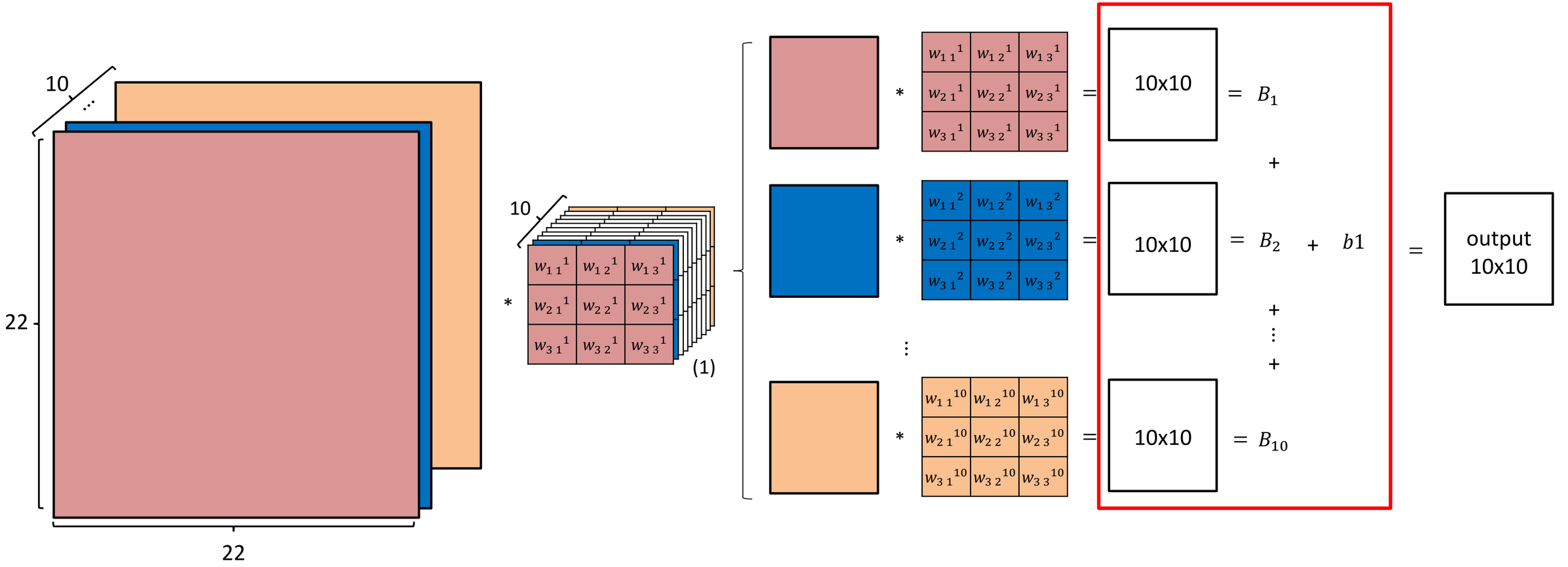
A의 한 채널 당, 곱 횟수 = 9*400회, 덧셈 횟수 = 9*400회

총 곱, 덧셈 횟수 = 9*400*10 = 36,000회

Problem 5

2,3) 곱셈, 덧셈 횟수 및 행렬 표현 구하기 (4/7)

A의 10개의 channel 과 첫번째 kernel의 10개의 channel 과 각각 대응해서 컨볼루션, 각 결과를 Sum 해주면, 한 개의 Output channel 생성



Problem 5

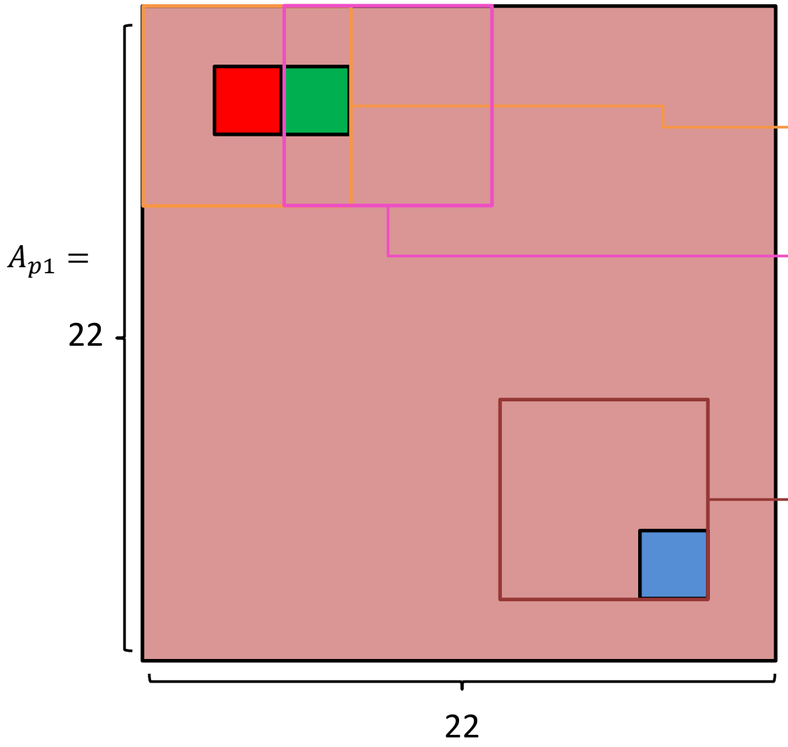
2,3) 곱셈, 덧셈 횟수 및 행렬 표현 구하기 (5/7)

$$\mathbb{A}=\begin{bmatrix} A_1{}^1 & \cdots & A_1{}^{10} \\ A_2{}^1 & \cdots & A_2{}^{10} \\ \vdots & \ddots & \vdots \\ A_{400}{}^1 & \cdots & A_{400}{}^{10} \end{bmatrix}$$

$$A_{f1} \times W_1{}^{T(1)} = B_1{}^{(1)} = \begin{bmatrix} 0 * w_{1\,1}{}^1 + 0 * w_{1\,2}{}^1 + \cdots + A_{22}{}^1 * w_{3\,3}{}^1 \\ \vdots \\ A_{358}{}^1 * w_{1\,1}{}^1 + A_{359}{}^1 * w_{1\,2}{}^1 + \cdots + A_{400}{}^1 * w_{3\,3}{}^1 \end{bmatrix}$$

100x9 9x1 100x1

A의 채널 1 만 봤을때, Padding = 1, Stride = 2



Using flatten(), flat[i]= $x_{[\frac{i}{N}].imodN}$

몫. 나머지

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & A_1{}^1 & A_2{}^1 \\ 0 & A_{21}{}^1 & A_{22}{}^1 \end{bmatrix} \rightarrow \underset{1x9}{\left[0\ 0\ 0\ 0\ A_1{}^1\ A_2{}^1\ 0\ A_{21}{}^1\ A_{22}{}^1\right]}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ A_2{}^1 & A_3{}^1 & A_4{}^1 \\ A_{22}{}^1 & A_{23}{}^1 & A_{24}{}^1 \end{bmatrix} \rightarrow \underset{1x9}{\left[0\ 0\ 0\ A_2{}^1\ A_3{}^1\ A_4{}^1\ A_{22}{}^1\ A_{23}{}^1\ A_{24}{}^1\right]}$$

$$\vdots \qquad\qquad \vdots$$

$$\begin{bmatrix} A_{358}{}^1 & A_{359}{}^1 & A_{360}{}^1 \\ A_{378}{}^1 & A_{379}{}^1 & A_{80}{}^1 \\ A_{398}{}^1 & A_{399}{}^1 & A_{400}{}^1 \end{bmatrix} \rightarrow \underset{1x9}{\left[A_{358}{}^1 A_{359}{}^1 A_{360}{}^1 A_{378}{}^1 A_{379}{}^1 A_{380}{}^1 A_{398}{}^1 A_{399}{}^1 A_{400}{}^1\right]}$$

첫번째 커널의 채널 1 만 봤을때

$$\times \begin{bmatrix} w_{1\,1}{}^1 & w_{1\,2}{}^1 & w_{1\,3}{}^1 \\ w_{2\,1}{}^1 & w_{2\,2}{}^1 & w_{2\,3}{}^1 \\ w_{3\,1}{}^1 & w_{3\,2}{}^1 & w_{3\,3}{}^1 \end{bmatrix}$$

Flatten and T

Problem 5

2,3) 곱셈, 덧셈 횟수 및 행렬 표현 구하기 (6/7)

$$A_{f1} \times {W_1}^{T(1)} = {B_1}^{(1)} = \begin{bmatrix} 0 * {w_{1\,1}}^1 + 0 * {w_{1\,2}}^1 + \cdots + {A_{22}}^1 * {w_{3\,3}}^1 \\ \vdots \\ {A_{358}}^1 * {w_{1\,1}}^1 + {A_{359}}^1 * {w_{1\,2}}^1 + \cdots + {A_{400}}^1 * {w_{3\,3}}^1 \end{bmatrix}$$

100x1

$$B = \begin{bmatrix} \sum_{i=1}^{10} {A_f}^i \times {W^{iT(1)}} + b1 & \cdots & \sum_{i=1}^{10} {A_f}^i \times {W^{iT(15)}} + b15 \end{bmatrix} = [B^{(1)} \quad \cdots \quad B^{(15)}]$$

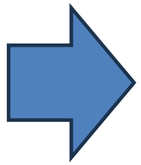
100x15

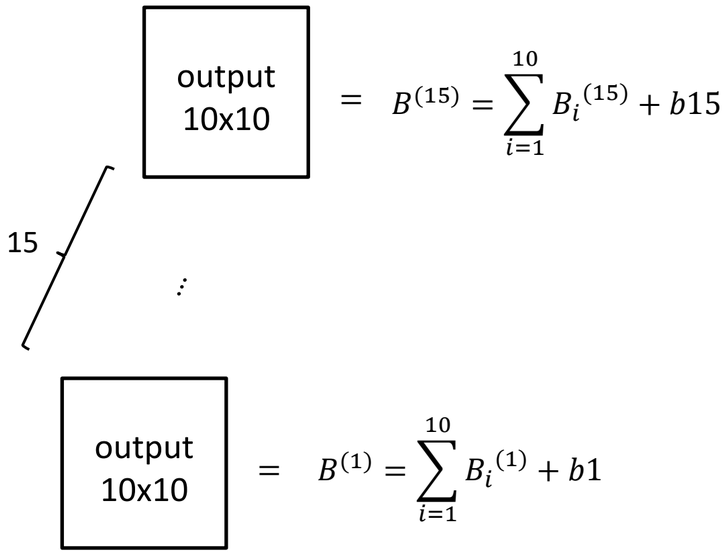
output 10x10 $= B^{(1)} = \sum_{i=1}^{10} {B_i}^{(1)} + b1$

$= \sum_{i=1}^{10} A_{fi} \times {W_i}^{T(1)} + b1$

곱 : 9*400*10 = 36,000회

덧셈 : 9*400*10 = 36,000회





B의 한 채널 당, 곱 횟수 = 9*100*10회, 덧셈 횟수 = 9*100*10 회

총 곱, 덧셈 횟수 = 9*100*10*15 = 135,000회

Problem 5

2,3) 곱셈, 덧셈 횟수 및 행렬 표현 구하기 (7/7)

$$\mathbb{W}_C = \begin{bmatrix} w_{C1\,1} & \cdots & w_{C1\,10} \\ \vdots & \ddots & \vdots \\ w_{C1500\,1} & \cdots & w_{C1500\,10} \end{bmatrix}, \mathbb{B}_C = \begin{bmatrix} b_1 \\ \vdots \\ b_{C10} \end{bmatrix}, B = [B^{(1)} \quad \cdots \quad B^{(15)}]$$

1500x10 　　10x1 　　100x15

Using col flatten(), flat[i]=$x_{imodN.\left[\frac{i}{N}\right]}$

$$\begin{bmatrix} B^{(1)} \\ \vdots \\ B^{(15)} \end{bmatrix} = \begin{bmatrix} 0 * {w_{1\,1}}^{1} + 0 * {w_{1\,2}}^{1} + \cdots + {A_{22}}^{1} * {w_{3\,3}}^{1} \\ \vdots \\ {A_{358}}^{10} * {w_{1\,1}}^{15} + {A_{359}}^{10} * {w_{1\,2}}^{15} + \cdots + {A_{400}}^{10} * {w_{3\,3}}^{15} \end{bmatrix} = \boldsymbol{Y_2} = \begin{bmatrix} y_1 \\ \vdots \\ y_{1500} \end{bmatrix}$$

1500x1 　　1500x1

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C_1} \\ \vdots \\ \boldsymbol{C_{10}} \end{bmatrix} = {\mathbb{W}_C}^{\boldsymbol{T}} \times \boldsymbol{Y_2} + \mathbb{B}_C = \begin{bmatrix} w_{C1\,1} & \cdots & w_{C1500\,1} \\ \vdots & \ddots & \vdots \\ w_{C1\,10} & \cdots & w_{C1500\,10} \end{bmatrix} \times \begin{bmatrix} y_1 \\ \vdots \\ y_{1500} \end{bmatrix} + \begin{bmatrix} b_{C1} \\ \vdots \\ b_{C10} \end{bmatrix} = \begin{bmatrix} w_{C1\,1} * y_1 + \cdots + w_{C1500\,1} * y_{1500} + b_{C1} \\ \vdots \\ w_{C1\,10} * y_1 + \cdots + w_{C1500\,10} * y_{1500} + b_{C10} \end{bmatrix} = \begin{bmatrix} C_1 \\ \vdots \\ C_{10} \end{bmatrix}$$

10x1500 　　1500x1 　　10x1 　　10x1 　　10x1

C layer에
존재하는 각각의
뉴런에 저장된 값

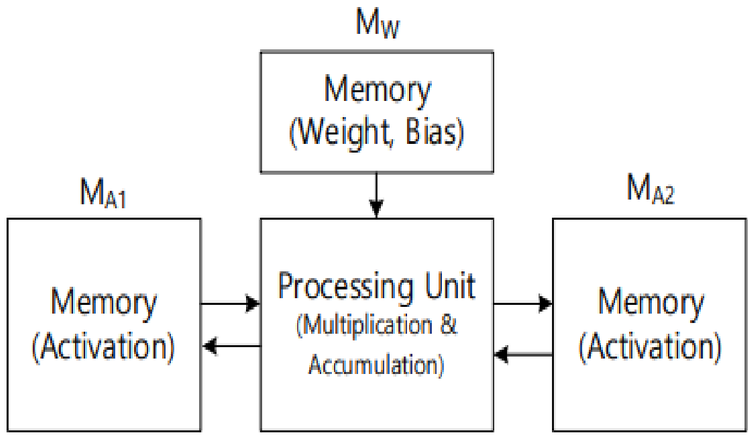
$w_{C1\,1} * y_1 + \cdots + w_{C1500\,1} * y_{1500} + b_{C1}$

한 행당, 곱 = 1,500, 덧셈 = 1,500

C layer에서 총 곱 횟수 = 15,000회, 총 덧셈 횟수 = 15,000회

Problem 5

Hardware 크기 및 동작 원리



M_{A1} : 12,000bit(1,500x8b),

M_W : 120,080bit(15,010x8b),

M_{A2} : 32,000bit(400x10x8b)

<구조 >

TBD

감사합니다!