UJIAN AKHIR SEMESTER

EL-5105 VLSI untuk Pengolahan Sinyal Digital

School of Electrical Engineering and Informatics Institut Teknologi Bandung
Dosen : Prof. Trio Adiono, PhD
Hari : Kamis, 16 November 2023
Waktu : Take Home Test
Pengumpulan & Presentasi : Senin, 18 Desember 2023, Jam 10.00 WIB (laporan format pdf)

Pada UAS ini dilakukan perancangan Arsitektur prosesor AI berbasis CNN.

Salah satu algoritma perhitungan cepat yang dapat diterapkan pada *CNN* adalah Algoritma *Winograd*. Algoritma ini mengurangi jumlah multiplikasi melalui transformasi pada *input* dan *filter*. Untuk menghitung output *m* dengan filter *r* dituliskan *F(m, r)*, contoh: *F(2, 3)* didefinisikan sebagai vektor *input* $x = (x_0, x_1, x_2, x_3)$, *filter* $w = (w_0, w_1, w_2)$, dan vektor *output* $y = (y_0, y_1)$. Maka vektor *input* dan *filter* ditransformasikan dalam vektor berikut:

$$x' = (x_0 - x_2, x_1 + x_2, x_2 - x_1, x_1 - x_3)$$

$$w' = \left(w_0, \frac{w_0 + w_1 + w_2}{2}, \frac{w_0 - w_1 + w_2}{2}, w_2\right) \tag{II.1}$$

kemudian menggunakan persamaan berikut untuk menghitung *dot* produk,

$$\begin{aligned} f_1 &= (x_0 - x_2)w_0 \\ f_2 &= (x_1 + x_2)\frac{w_0 + w_1 + w_2}{2} \\ f_3 &= (x_2 - x_1)\frac{w_0 - w_1 + w_2}{2} \\ f_4 &= (x_1 - x_3)w_2 \end{aligned} \tag{II.2}$$

hasil akhir algoritma *Winograd* dari *F(2, 3)* adalah

$$y = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_0 & x_1 & x_2 \\ x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} f_1 + f_2 + f_3 \\ f_2 - f_3 - f_4 \end{bmatrix} \tag{II.3}$$

Dengan algoritma ini hanya membutuhkan 4 (empat) multiplikasi untuk komputasi output *y*. Sedangkan algoritma konvolusi standar menggunakan *2 x 3 = 6* multiplikasi. Secara umum, algoritma *Winograd* dapat dituliskan:

$$y = A^T[(Gw)e\ (B^T x)] \tag{II.4}$$

dengan *e* adalah *element wise multiplication*. *A*, *G*, dan *B* adalah konstanta matriks yang ditentukan oleh ukuran *input x* dan *filter w*.

Soal 1. Konstanta matriks untuk 1D *F(2,3)* adalah,

$$B^T = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{(II.5)}$$

$$A^T = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix}$$

Soal 2.

2D F(2×2, 3×3)

Output (m) = 2×2
Filter (r) = 3×3
Input = (m + r - 1) ×(m + r - 1) = 4×4
Multiplication = 16

$$y = A^T[(GwG^T)\odot(B^T xB)]A$$

Misalkan input (x)

x_0	x_1	x_2	x_3
x_4	x_5	x_6	x_7
x_8	x_9	x_{10}	x_{11}
x_{12}	x_{13}	x_{14}	x_{15}

Filter (w)

w0	w1	w2
w3	w4	w5
w6	w7	w8

Parameter matrix sama dengan 1D

dan parameter matriks

$$G = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

$$B^T = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix}$$

a. Transformasi Input xi = B^Tx

Proses kolom per kolom input

- Kolom input pertama (x0, x4, x8, x12)

$$x_i^1 = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_4 \\ x_8 \\ x_{12} \end{bmatrix} = \begin{bmatrix} x_0 - x_8 \\ x_4 + x_8 \\ -x_4 + x_8 \\ x_4 - x_{12} \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

- Kolom input kedua (x1, x5, x9, x 13)

$$x_i^2 = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_5 \\ x_9 \\ x_{13} \end{bmatrix} = \begin{bmatrix} x_1 - x_9 \\ x_5 + x_9 \\ -x_5 + x_9 \\ x_5 - x_{13} \end{bmatrix} = \begin{bmatrix} a_4 \\ a_5 \\ a_6 \\ a_7 \end{bmatrix}$$

- Kolom input ketiga (x2, x6, x10, x 14)

$$x_i^3 = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_6 \\ x_{10} \\ x_{14} \end{bmatrix} = \begin{bmatrix} x_2 - x_{10} \\ x_6 + x_{10} \\ -x_6 + x_{10} \\ x_6 - x_{14} \end{bmatrix} = \begin{bmatrix} a_8 \\ a_9 \\ a_{10} \\ a_{11} \end{bmatrix}$$

- Kolom input keempat (x3, x7, x11, x 15)

$$x_i^4 = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_3 \\ x_7 \\ x_{11} \\ x_{15} \end{bmatrix} = \begin{bmatrix} x_3 - x_{11} \\ x_7 + x_{11} \\ -x_7 + x_{11} \\ x_7 - x_{15} \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{13} \\ a_{14} \\ a_{15} \end{bmatrix}$$

Matriks baru xi

a_0	a_4	a_8	a_{12}
a_1	a_5	a_9	a_{13}
a_2	a_6	a_{10}	a_{14}
a_3	a_7	a_{11}	a_{15}

b. Transformasi Input V = xiB

Proses baris per baris input xi

- Baris input xi pertama (a0, a4, a8, a12)

$$V^1 = [a_0 \quad a_4 \quad a_8 \quad a_{12}] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$
$$= [(a_0 - a_8) \quad (a_4 + a_8) \quad (-a_4 + a_8) \quad (a_4 - a_{12})]$$
$$= [b_0 \quad b_1 \quad b_2 \quad b_3]$$

- Baris input xi kedua (a1, a5, a9, a13)

$$V^2 = [a_1 \quad a_5 \quad a_9 \quad a_{13}] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$
$$= [(a_1 - a_9) \quad (a_5 + a_9) \quad (-a_5 + a_9) \quad (a_5 - a_{13})]$$
$$= [b_4 \quad b_5 \quad b_6 \quad b_7]$$

- Baris input xi ketiga (a2, a6, a10, a14)

$$V^3 = [a_2 \quad a_6 \quad a_{10} \quad a_{14}] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$
$$= [(a_2 - a_{10}) \quad (a_6 + a_{10}) \quad (-a_6 + a_{10}) \quad (a_6 - a_{14})]$$
$$= [b_8 \quad b_9 \quad b_{10} \quad b_{11}]$$

- Baris input xi keempat (a3, a7, a11, a15)

$$V^4 = [a_3 \quad a_7 \quad a_{11} \quad a_{15}] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$
$$= [(a_3 - a_{11}) \quad (a_7 + a_{11}) \quad (-a_7 + a_{11}) \quad (a_7 - a_{15})]$$
$$= [b_{12} \quad b_{13} \quad b_{14} \quad b_{15}]$$

Matriks baru V

b_0	b_1	b_2	b_3
b_4	b_5	b_6	b_7
b_8	b_9	b_{10}	b_{11}
b_{12}	b_{13}	b_{14}	b_{15}

c. Transformasi Filter wi =Gw
Proses kolom per kolom filter w
- Kolom filter pertama (w0, w3, w6)

$$w_i^1 = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_0 \\ w_3 \\ w_6 \end{bmatrix} = \begin{bmatrix} w_0 \\ \left(\frac{w_0 + w_3 + w_6}{2}\right) \\ \left(\frac{w_0 - w_3 + w_6}{2}\right) \\ w_6 \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

- Kolom filter kedua (w1, w4, w7)

$$w_i^2 = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_4 \\ w_7 \end{bmatrix} = \begin{bmatrix} w_1 \\ \left(\frac{w_1 + w_4 + w_7}{2}\right) \\ \left(\frac{w_1 - w_4 + w_7}{2}\right) \\ w_7 \end{bmatrix} = \begin{bmatrix} c_4 \\ c_5 \\ c_6 \\ c_7 \end{bmatrix}$$

- Kolom filter ketiga (w2, w5, w8)

$$w_i^3 = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_2 \\ w_5 \\ w_8 \end{bmatrix} = \begin{bmatrix} w_2 \\ \left(\frac{w_2 + w_5 + w_8}{2}\right) \\ \left(\frac{w_2 - w_4 + w_8}{2}\right) \\ w_8 \end{bmatrix} = \begin{bmatrix} c_8 \\ c_9 \\ c_{10} \\ c_{11} \end{bmatrix}$$

Matriks baru wi

c_0	c_4	c_8
c_1	c_5	c_9
c_2	c_6	c_{10}
c_3	c_7	c_{11}

d. Transformasi Filter U= wiG^T

Proses baris per baris filter wi

- Baris pertama filter wi (c0, c4, c8)

$$U^1 = \begin{bmatrix} c_0 & c_4 & c_8 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$
$$= \begin{bmatrix} c_0 & \left(\frac{c_0 + c_4 + c_8}{2}\right) & \left(\frac{c_0 - c_4 + c_8}{2}\right) & c_8 \end{bmatrix}$$
$$= \begin{bmatrix} d_0 & d_1 & d_2 & d_3 \end{bmatrix}$$

- Baris kedua filter wi (c1, c5, c9)

$$U^2 = \begin{bmatrix} c_1 & c_5 & c_9 \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$
$$= \begin{bmatrix} c_1 & \left(\frac{c_1 + c_5 + c_9}{2}\right) & \left(\frac{c_1 - c_5 + c_9}{2}\right) & c_9 \end{bmatrix}$$
$$= \begin{bmatrix} d_4 & d_5 & d_6 & d_7 \end{bmatrix}$$

- Baris ketiga filter wi (c2, c6, c10)

$$U^3 = [c_2 \quad c_6 \quad c_{10}]\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$

$$= \left[c_2 \quad \left(\frac{c_2 + c_6 + c_{10}}{2}\right) \quad \left(\frac{c_2 - c_6 + c_{10}}{2}\right) \quad c_{10}\right]$$

$$= [d_8 \quad d_9 \quad d_{10} \quad d_{11}]$$

- Baris keempat filter wi (c3, c7, c11)

$$U^3 = [c_3 \quad c_7 \quad c_{11}]\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$$

$$= \left[c_3 \quad \left(\frac{c_3 + c_7 + c_{11}}{2}\right) \quad \left(\frac{c_3 - c_7 + c_{11}}{2}\right) \quad c_{11}\right]$$

$$= [d_{12} \quad d_{13} \quad d_{14} \quad d_{15}]$$

Matriks baru U

d_0	d_1	d_2	d_3
d_4	d_5	d_6	d_7
d_8	d_9	d_{10}	d_{11}
d_{12}	d_{13}	d_{14}	d_{15}

e. Perkalian Titik M=V.U

b_0	b_1	b_2	b_3
b_4	b_5	b_6	b_7
b_8	b_9	b_{10}	b_{11}
b_{12}	b_{13}	b_{14}	b_{15}

Θ

d_0	d_1	d_2	d_3
d_4	d_5	d_6	d_7
d_8	d_9	d_{10}	d_{11}
d_{12}	d_{13}	d_{14}	d_{15}

=

$b_0\ d_0$	$b_1\ d_1$	$b_2\ d_2$	$b_3\ d_3$
$b_4\ d_4$	$b_5\ d_5$	$b_6\ d_6$	$b_7\ d_7$
$b_8\ d_8$	$b_9\ d_9$	$b_{10}\ d_{10}$	$b_{11}\ d_{11}$
$b_{12}\ d_{12}$	$b_{13}\ d_{13}$	$b_{14}\ d_{14}$	$b_{15}\ d_{15}$

=

f_0	f_1	f_2	f_3
f_4	f_5	f_6	f_7
f_8	f_9	f_{10}	f_{11}
f_{12}	f_{13}	f_{14}	f_{15}

f. Transformasi Output Mi = A^TM
Proses kolom per kolom matriks M
- Kolom pertama matriks M (f0, f4, f8, f12)

$$M^1 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} f_0 \\ f_4 \\ f_8 \\ f_{12} \end{bmatrix} = \begin{bmatrix} f_0 + f_4 + f_8 \\ f_4 - f_8 - f_{12} \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \end{bmatrix}$$

- Kolom kedua matriks M (f1, f5, f9, f13)

$$M^2 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_5 \\ f_9 \\ f_{13} \end{bmatrix} = \begin{bmatrix} f_1 + f_5 + f_9 \\ f_5 - f_9 - f_{13} \end{bmatrix} = \begin{bmatrix} g_2 \\ g_3 \end{bmatrix}$$

- Kolom ketiga matriks M (f2, f6, f10, f14)

$$M^3 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} f_2 \\ f_6 \\ f_{10} \\ f_{14} \end{bmatrix} = \begin{bmatrix} f_2 + f_6 + f_{10} \\ f_6 - f_{10} - f_{14} \end{bmatrix} = \begin{bmatrix} g_4 \\ g_5 \end{bmatrix}$$

- Kolom keempat matriks M (f3, f7, f11, f15)

$$M^4 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} f_3 \\ f_7 \\ f_{11} \\ f_{15} \end{bmatrix} = \begin{bmatrix} f_3 + f_7 + f_{11} \\ f_7 - f_{11} - f_{15} \end{bmatrix} = \begin{bmatrix} g_6 \\ g_7 \end{bmatrix}$$

Matriks baru Mi

g_0	g_2	g_4	g_6
g_1	g_3	g_5	g_7

g. Transformasi Output Y = MiA
Proses baris per baris dari matriks Mi
- Baris pertama matris Mi (g0, g2, g4, g6)

$$Y^1 = \begin{bmatrix} g_0 & g_2 & g_4 & g_6 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} (g_0 + g_2 + g_4) & (g_2 - g_4 - g_6) \end{bmatrix} = \begin{bmatrix} y_0 & y_1 \end{bmatrix}$$

- Baris kedua matris Mi (g1, g3, g5, g7)

$$Y^2 = [g_1 \quad g_3 \quad g_5 \quad g_7]\begin{bmatrix}1 & 0\\1 & 1\\1 & -1\\0 & -1\end{bmatrix} = [(g_1 + g_3 + g_5) \quad (g_3 - g_5 - g_7)] = [y_2 \quad y_3]$$

Matriks Output

y_0	y_1
y_2	y_3

Pertanyaan:

Buatlah Arsitektur prosesor untuk melakukan keuda fungsi menghitung y diatas.

Data input dan template dibuat random bernailai 1 sd 8.

Tugas:

1. Buatlah DFG dan blok diagram data path algoritma diatas
2. Rancanglah arsitektur parallel untuk komputasi diatas (Systolic, Parallel, Pipeline atau arsitektur lainnya)
3. Verifikasi arsitektur dengan simulasi Verilog atau perhitungan manual menggunakan matlab.
4. Bandingkan hasilnya dengan perhitungan perkalian langsung?
5. Hitung berapa clock dibutuhkan untuk menyelesaikan semua proses.
6. Hitung throughput arsitektur tersebut
7. Berapa percepatan menggunakan arsitektur parallel usulan saudara.
8. Berapa pengurangan jumlah multiplikasi
9. Berapa kebutuhan multiplier dan adder untuk masing-masing arsitektur

Catatan:

- Jika ada pertanyaan, mohon disampaikan via MS Teams.