CNN-EIA 2022-06-28

Analyzer Report (Meomry Capcity)

This report generated by Convolutional Neural Network Inference Analyzer (CNN-IA) to summarize the analysis needed to reach the optimal memory capacity for mlp_fc3_batch16 using restricted schedule space.

Memory Architecture:

	L0	L1	L2
Capacity	256	131072	536870912
Access cost	0	7	200
Static cost	0	0	0
Parallel count	256	1	1
Parallel mode	1	0	0
Parallel cost	2	0	0

Precision : 16

Minimum utilization : 0.0%

Outputs can be buffered by mac : 0

Replication to improve utilization : True

Glossary:

- Memory Levels: (L0, L1, L2)
The smallest index the nearest to CPU.

- Loop Notations: (FX, FY, OX, OY, OC, IC, ON)

FX : FILTER WIDTH
FY : FILTER HEIGHT
OX : OUTPUT WIDTH
OY : OUTPUT HEIGHT
OC : OUTPUT CHANNEL
IC : INPUT CHANNEL

ON : BATCH

CNN-EIA 2022-06-28

Exploration Table

L0-SIZE	L0-COST	L1-SIZE	L1-COST	L2-SIZE	L2-COST	TOTAL
16.0	0.05	16384.0	3.84	536870912.0	200.0	70842440.0
32.0	0.1	16384.0	3.84	536870912.0	200.0	57562240.0
64.0	0.2	16384.0	3.84	536870912.0	200.0	50153840.0
128.0	0.4	16384.0	3.84	536870912.0	200.0	48412240.0
256.0	0.8	16384.0	3.84	536870912.0	200.0	50497040.0
16.0	0.05	32768.0	4.8	536870912.0	200.0	74990600.0
32.0	0.1	32768.0	4.8	536870912.0	200.0	59790400.0
64.0	0.2	32768.0	4.8	536870912.0	200.0	51230000.0
128.0	0.4	32768.0	4.8	536870912.0	200.0	49373200.0
256.0	0.8	32768.0	4.8	536870912.0	200.0	49451600.0
16.0	0.05	65536.0	6.0	536870912.0	200.0	73775800.0
32.0	0.1	65536.0	6.0	536870912.0	200.0	56175600.0
64.0	0.2	65536.0	6.0	536870912.0	200.0	47975200.0
128.0	0.4	65536.0	6.0	536870912.0	200.0	44374400.0
256.0	0.8	65536.0	6.0	536870912.0	200.0	43572800.0
16.0	0.05	131072.0	7.5	536870912.0	200.0	75549800.0
32.0	0.1	131072.0	7.5	536870912.0	200.0	59199600.0
64.0	0.2	131072.0	7.5	536870912.0	200.0	45324200.0
128.0	0.4	131072.0	7.5	536870912.0	200.0	42735900.0
256.0	0.8	131072.0	7.5	536870912.0	200.0	42154300.0

Optimal Cost

[42154300.0 pJ]

MEM: L0 L1 L2

SIZE: 256.0 131072.0 536870912.0 COST: 0.8 7.5 200.0