SoC Final Report

Name	전상우, 장원	Student ID	2019023654, 2019097138
Title	2D CNN module(sobel, sharpening)		

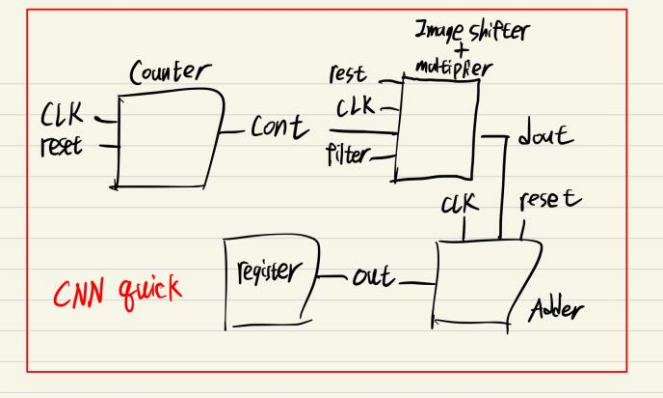
1. Theory:

   CNN(Convolutional Neural Network)은 주로 이미지 인식과 같은 시각적 데이터 처리에 사용되는 딥러닝 모델입니다. 여기에는 convolution, activation, pooling, fully connected layer가 있습니다. 간단한 CNN 2D layer를 만들기 위해 합/곱셈기, counter와 레지스터 파일을 이용하여 구성했습니다.

   Sobel 필터는 이미지 처리에서 경계 검출을 위해 사용되는 필터 중 하나입니다. 주로 gradient 계산을 통해 edge를 찾는 데 사용됩니다. 이 때 0,1,2 와 그 음수로 구성된 x, y 필터를 적용하여 rms을 결과로 사용합니다.

   Sharpening filter는 이미지의 선명도를 향상시키는 데 사용됩니다. 주로 이미지의 고주파 성분을 강조하여 세부적인 특징을 뚜렷하게 만들어줍니다. 이 때 -1과 가운데 성분에 9로 구성된 필터를 적용합니다.

2. 구현 방식: 지난 주에 만든 1D CNN을 사용하되 ,좀 더 간단하게 만들 수 있는 방법을 고안했습니다.



- 카운터: 카운터는 convolution을 위해 이미지(혹은 필터)의 column을 한 칸 씩 옮기고, 맨 끝 부분에 도달한다면 row를 한 칸 내리도록 설계했습니다.

```
else if(count % 30 == 0 && count !=0) begin

    count <= count + `filter_w - 1; // 30 + 3 - 1 = 32

end

else begin

    count <= count + 1;
```

- Image shifter + multiplier: 이미지의 convolution 필터 크기에 맞게 9x9로 정해놓고 옆으로 한 칸 씩 옮기게 했습니다. 이 때 각각의 cell을 filter의 cell과 바로 곱할 수 있게 했습니다. Count를 더하며 끝에 도달하면 밑으로 영역을 옮기도록 했습니다.

```
dout8 <= Image[0*`image_w + count]                 * filter[0:7];
dout7 <= Image[0*`image_w + 1 + count]             * filter[8:15];
dout6 <= Image[0*`image_w + `filter_w - 1 + count] * filter[16:23];
dout5 <= Image[1*`image_w + count]                 * filter[24:31];
dout4 <= Image[1*`image_w + 1 + count]             * filter[32:39];
dout3 <= Image[1*`image_w + `filter_w - 1 + count] * filter[40:47];
dout2 <= Image[2*`image_w + count]                 * filter[48:55];
dout1 <= Image[2*`image_w + 1 + count]             * filter[56:63];
dout0 <= Image[2*`image_w + `filter_w - 1 + count] * filter[64:71];
```

- Adder: adder는 convolution으로 나온 결과들을 더해서 output의 cell에 저장할 수 있도록 했습니다. 이 때 sobel filter와 sharpening filter의 음수 부분에 주의하며 코드를 수정했습니다.

```
out <= -dout0 + dout1 + dout2
       -dout3 + dout4 + dout5
       -dout6 + dout7 + dout8;
```

sobel-x

```
- dout0 - dout1 - dout2
- dout3 + dout4 - dout5
- dout6 - dout7 - dout8;
```

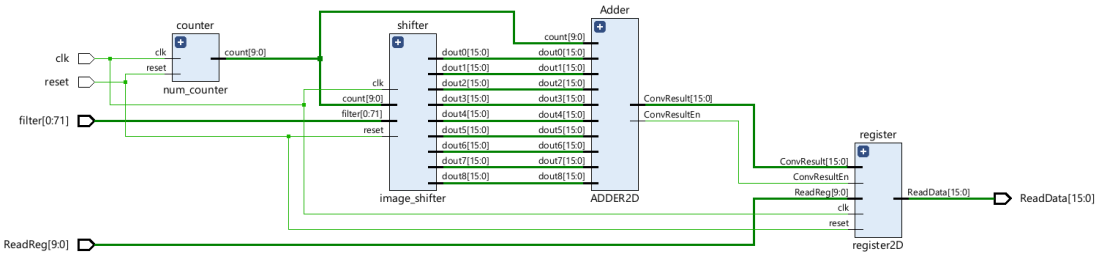
sharpening

- Register: 덧셈기에서 나온 결과를 레지스터와 메모리 파일에 저장하도록 했습니다.

```
memout[i] <= ConvResult;
$writememh("sharp_Output1.mem", memout);
```

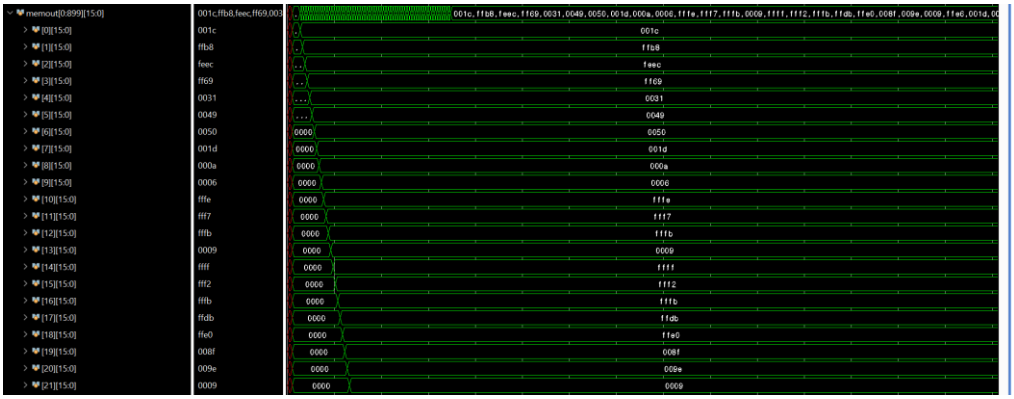
3. Simulation:
   - Sobel–x

     -회로도



초반에 저희가 구상한 회로대로 이미지를 받아서 카운터 로직에 따라 shift하고 그 합을 레지스터에 더할 수 있게 합성된 것을 볼 수 있었습니다.

-테스트벤치 시뮬레이션결과

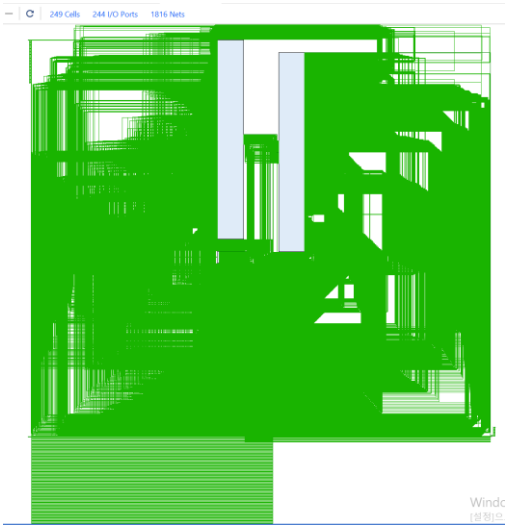


Convresult en값과 함께 초기 001e... 가 나왔습니다. 모두 output.mem과 값이 일치했습니다. 클락의 rising edge마다 조금씩 밀려서 계단형으로 출력되는 모습입니다.

마지막 값도 ffb6으로 output.mem과 결과가 일치했습니다.

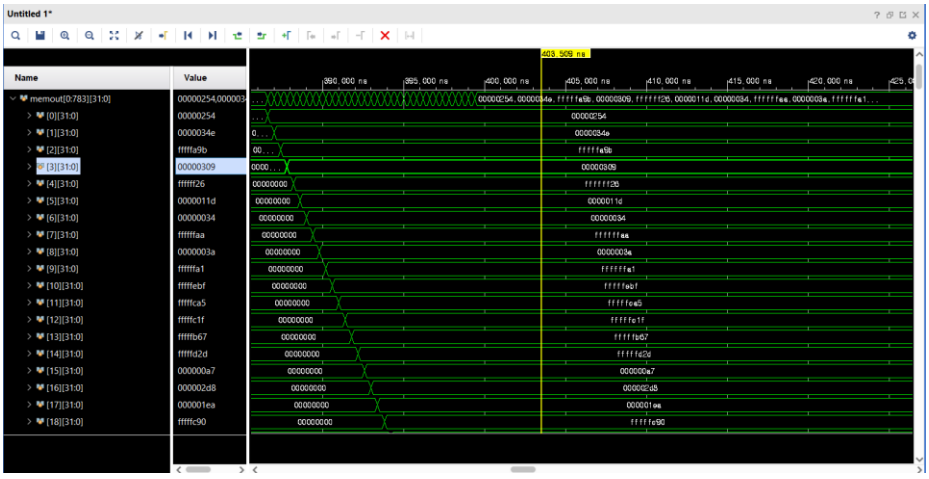
- Sharpening

-회로도

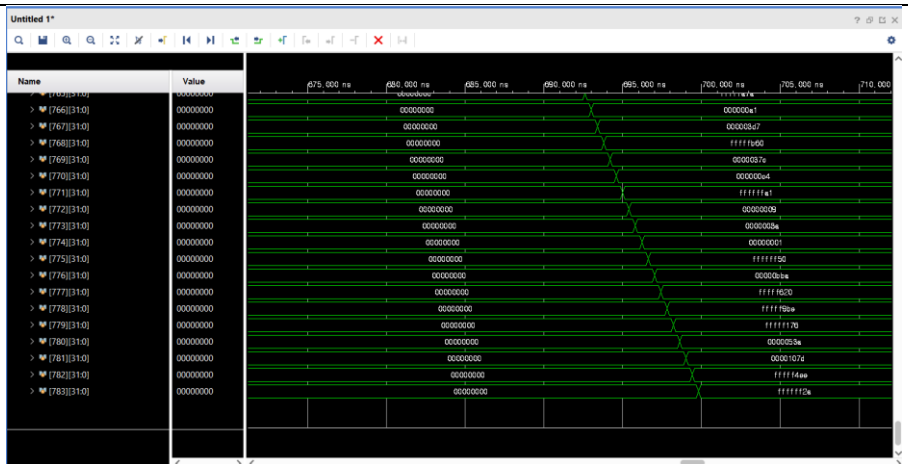


회로를 다소 난잡하게 설계하여... 최적화하는 데에는 실패했습니다.
결과에 목적을 맞춰 설계했지만 앞으로는 place and route에도 신경을 쓰도록 하겠습니다.

-테스트벤치 시뮬레이션 결과



초기 값 00000245부터 끝까지 conv2의 값과 일치하는 것을 확인할 수 있었습니다.
특이사항으로 conv1을 계산하고, 그것을 다른 mem1구조에 저장하고 나서 같은 방식으로 계산을 진행할 수 있도록 두 모듈을 합쳤습니다.



마지막 부분 역시 conv2의 값과 일치하였습니다.

4. 합성 결과 설명

- Sobel-x

1. Slice Logic

Site Type	Used	Fixed	Prohibited	Available	Util%
Slice LUTs	6965	0	0	41000	16.99
LUT as Logic	6965	0	0	41000	16.99
LUT as Memory	0	0	0	13400	0.00
Slice Registers	14699	0	0	82000	17.93
Register as Flip Flop	14699	0	0	82000	17.93
Register as Latch	0	0	0	82000	0.00
F7 Muxes	2559	0	0	20500	12.48
F8 Muxes	948	0	0	10250	9.25

-총 41000개의 LUTs 중 6965개가 사용되었으며, 이는 16.99%의 사용률을 보입니다.
-메모리로 사용된 LUTs는 없으며, 전체 13400개가 사용 가능합니다.
-플립플롭으로 사용된 레지스터는 14699개로, 82000개의 17.93%를 차지합니다.
-멀티플렉서(F7, F8)도 각각 12.48%, 9.25%의 사용률을 보입니다.

3. Memory

Site Type	Used	Fixed	Prohibited	Available	Util%
Block RAM Tile	0	0	0	135	0.00
RAMB36/FIFO*	0	0	0	135	0.00
RAMB18	0	0	0	270	0.00

Mem 파일을 이용하고 Bram은 이용하지 않아서 util = 0%된 것 같습니다.

4. DSP

```
+-----------+------+-------+------------+-----------+-------+
| Site Type | Used | Fixed | Prohibited | Available | Util% |
+-----------+------+-------+------------+-----------+-------+
| DSPs      |    0 |     0 |          0 |       240 |  0.00 |
+-----------+------+-------+------------+-----------+-------+
```

-DSP는 사용되지 않았습니다. 논리연산보다는 CNN이 단순한 연산으로 처리되기 때문인 것으로 생각됩니다. 이 경로가 주로 LUT, MUX, IBUF 및 OBUF를 사용하여 구현되었기 때문일 수 있습니다.

5.타이밍

```
  Location             Delay type                Incr(ns)  Path(ns)    Netlist Resource(s)
-------------------------------------------------------------------    -------------------
  M26                                                  0.000     0.000 r  ReadReg[0] (IN)
                       net (fo=0)                      0.000     0.000    ReadReg[0]
  M26                  IBUF (Prop_ibuf_I_O)            0.864     0.864 r  ReadReg_IBUF[0]_inst/O
                       net (fo=3600, routed)          20.881    21.745    register/ReadReg[0]
  SLICE_X50Y123        LUT6 (Prop_lut6_I4_O)           0.053    21.798 r  register/ReadData_OBUF[4]_inst_i_401/O
                       net (fo=1, routed)              0.000    21.798    register/ReadData_OBUF[4]_inst_i_401_n_0
  SLICE_X50Y123        MUXF7 (Prop_muxf7_I0_O)         0.121    21.919 r  register/ReadData_OBUF[4]_inst_i_314/O
                       net (fo=1, routed)              0.000    21.919    register/ReadData_OBUF[4]_inst_i_314_n_0
  SLICE_X50Y123        MUXF8 (Prop_muxf8_I1_O)         0.054    21.973 r  register/ReadData_OBUF[4]_inst_i_142/O
                       net (fo=1, routed)              1.094    23.068    register/ReadData_OBUF[4]_inst_i_142_n_0
  SLICE_X46Y99         LUT6 (Prop_lut6_I5_O)           0.156    23.224 r  register/ReadData_OBUF[4]_inst_i_57/O
                       net (fo=1, routed)              0.000    23.224    register/ReadData_OBUF[4]_inst_i_57_n_0
  SLICE_X46Y99         MUXF7 (Prop_muxf7_I0_O)         0.121    23.345 r  register/ReadData_OBUF[4]_inst_i_20/O
                       net (fo=1, routed)              0.000    23.345    register/ReadData_OBUF[4]_inst_i_20_n_0
  SLICE_X46Y99         MUXF8 (Prop_muxf8_I1_O)         0.054    23.399 r  register/ReadData_OBUF[4]_inst_i_8/O
                       net (fo=1, routed)              2.431    25.829    register/ReadData_OBUF[4]_inst_i_8_n_0
  SLICE_X33Y42         LUT5 (Prop_lut5_I0_O)           0.156    25.985 r  register/ReadData_OBUF[4]_inst_i_3/O
                       net (fo=1, routed)              0.000    25.985    register/ReadData_OBUF[4]_inst_i_3_n_0
  SLICE_X33Y42         MUXF7 (Prop_muxf7_I1_O)         0.129    26.114 r  register/ReadData_OBUF[4]_inst_i_1/O
                       net (fo=1, routed)              2.849    28.964    ReadData_OBUF[4]
  U17                  OBUF (Prop_obuf_I_O)            2.517    31.481 r  ReadData_OBUF[4]_inst/O
                       net (fo=0)                      0.000    31.481    ReadData[4]
  U17                                                                 r  ReadData[4] (OUT)
-------------------------------------------------------------------    -------------------
```

```
report_timing: Time (s): cpu = 00:00:08 ; elapsed = 00:00:06 . Memory (MB): peak = 3564.762 ; gain = 0.000
```

-경로 누적 지연: 입력에서 출력까지의 총 누적 지연은 31.481 ns입니다.
-사용된 자원:
여러 LUT, MUX, IBUF, OBUF가 사용되었습니다.
주요 지연 원인은 SLICE_X50Y123 및 SLICE_X33Y42에서 발생했습니다.

mem 파일로 아웃풋을 따로 받는 상황에서, 개별적으로 output을 어떻게 설정해야 하는지 잘 몰라서 따로 지정하지 않았습니다. 그 결과 slack에 대한 분석이 빠진 것 같습니다.

- Sharpening

1. Slice Logic

Site Type	Used	Fixed	Prohibited	Available	Util%
Slice LUTs*	10398	0	0	41000	25.36
LUT as Logic	7428	0	0	41000	18.12
LUT as Memory	2970	0	0	13400	22.16
LUT as Distributed RAM	2970	0			
LUT as Shift Register	0	0			
Slice Registers	12784	0	0	82000	15.59
Register as Flip Flop	12784	0	0	82000	15.59
Register as Latch	0	0	0	82000	0.00
F7 Muxes	2495	0	0	20500	12.17
F8 Muxes	948	0	0	10250	9.25

-LUT 사용량: 전체 41000개 중 10398개 사용 (25.36% 사용률)
-이 중 논리로 사용된 LUT는 7428개입니다. (18.12% 사용률)
-메모리로 사용된 LUT는 2970개입니다. (22.16% 사용률)
-분산 RAM으로 사용된 LUT로는 2970개를 사용했습니다.
-시프트 레지스터로 사용된 LUT는 없습니다. (0%)
-플립플롭으로 사용된 레지스터는 12784개입니다. (15.59% 사용률)

2. Memory

Site Type	Used	Fixed	Prohibited	Available	Util%
Block RAM Tile	0	0	0	135	0.00
RAMB36/FIFO*	0	0	0	135	0.00
RAMB18	0	0	0	270	0.00

-Block RAM Tile: 사용하지 않았습니다. (0%)
-RAMB36/FIFO: 사용하지 않았습니다. (0%)
-RAMB18: 사용하지 않았습니다. (0%)
이 또한 mem파일로 지정해서 그런 것 같습니다.

3. DSP

Site Type	Used	Fixed	Prohibited	Available	Util%
DSPs	13	0	0	240	5.42
DSP48E1 only	13				

3. DSP
DSPs: 13 사용 (5.42% 이용률)
DSP48E1 only: 13 사용

5. Timing Report

```
Timing Report

Slack:                    inf
  Source:                 nc2/count_2_reg[3]/C
                            (rising edge-triggered cell FDCE)
  Destination:            r2/memout_reg[0][0]/D
  Path Group:             (none)
  Path Type:              Max at Slow Process Corner
  Data Path Delay:        15.686ns  (logic 10.778ns (68.710%)  route 4.908ns (31.290%))
  Logic Levels:           19  (CARRY4=8 DSP48E1=5 FDCE=1 LUT1=1 LUT3=1 LUT6=2 RAMD64E=1)

    Location             Delay type                Incr(ns)  Path(ns)    Netlist Resource(s)
  -------------------------------------------------------------------    -------------------
                         FDCE                         0.000     0.000 r  nc2/count_2_reg[3]/C
                         FDCE (Prop_fdce_C_Q)         0.269     0.269 r  nc2/count_2_reg[3]/Q
                         net (fo=60, unplaced)        0.619     0.888    nc2/count_2[3]
                         LUT3 (Prop_lut3_I0_O)        0.153     1.041 r  nc2/mem1_reg_r9_640_703_6_8_i_2/O
                         net (fo=120, unplaced)       0.625     1.666    is/mem1_reg_r9_896_959_15_15/DPRA3
                         RAMD64E (Prop_ramd64e_RADR3_O)
                                                      0.053     1.719 r  is/mem1_reg_r9_896_959_15_15/DP/O
                         net (fo=1, unplaced)         0.665     2.384    is/mem1_reg_r9_896_959_15_15_n_0
                         LUT6 (Prop_lut6_I0_O)        0.053     2.437 r  is/ConvResult_21_i_17/O
                         net (fo=1, unplaced)         0.665     3.102    is/ConvResult_21_i_17_n_0
                         LUT6 (Prop_lut6_I0_O)        0.053     3.155 r  is/ConvResult_21_i_1/O
                         net (fo=15, unplaced)        0.584     3.739    is/ConvResult_22__0[15]
                         DSP48E1 (Prop_dsp48e1_A[29]_P[1])
                                                      3.255     6.994 f  is/ConvResult_21/P[1]
                         net (fo=1, unplaced)         0.584     7.577    is/ConvResult_21_n_104
                         LUT1 (Prop_lut1_I0_O)        0.053     7.630 r  is/ConvResult_20_i_78/O
                         net (fo=1, unplaced)         0.000     7.630    is/ConvResult_20_i_78_n_0
                         CARRY4 (Prop_carry4_S[1]_CO[3])
                                                      0.324     7.954 r  is/ConvResult_20_i_16/CO[3]
                         net (fo=1, unplaced)         0.000     7.954    is/ConvResult_20_i_16_n_0
                         CARRY4 (Prop_carry4_CI_CO[3])
                                                      0.058     8.012 r  is/ConvResult_20_i_15/CO[3]
                         net (fo=1, unplaced)         0.000     8.012    is/ConvResult_20_i_15_n_0
                         CARRY4 (Prop_carry4_CI_CO[3])
                                                      0.058     8.070 r  is/ConvResult_20_i_14/CO[3]
                         net (fo=1, unplaced)         0.000     8.070    is/ConvResult_20_i_14_n_0
                         CARRY4 (Prop_carry4_CI_CO[3])
                                                      0.058     8.186 r  is/ConvResult_20_i_12/CO[3]
                         net (fo=1, unplaced)         0.000     8.186    is/ConvResult_20_i_12_n_0
                         CARRY4 (Prop_carry4_CI_CO[3])
                                                      0.058     8.244 r  is/ConvResult_20_i_11/CO[3]
                         net (fo=1, unplaced)         0.000     8.244    is/ConvResult_20_i_11_n_0
                         CARRY4 (Prop_carry4_CI_CO[3])
                                                      0.058     8.302 r  is/ConvResult_20_i_10/CO[3]
                         net (fo=1, unplaced)         0.000     8.302    is/ConvResult_20_i_10_n_0
                         CARRY4 (Prop_carry4_CI_O[1])
                                                      0.220     8.522 r  is/ConvResult_20_i_9/O[1]
                         net (fo=1, unplaced)         0.584     9.106    is/ConvResult_20_i_9_n_6
                         DSP48E1 (Prop_dsp48e1_C[29]_PCOUT[47])
                                                      1.807    10.913 r  is/ConvResult_20/PCOUT[47]
                         net (fo=1, unplaced)         0.000    10.913    is/ConvResult_20_n_106
                         DSP48E1 (Prop_dsp48e1_PCIN[47]_PCOUT[47])
                                                      1.452    12.365 r  is/ConvResult_20__0/PCOUT[47]
                         net (fo=1, unplaced)         0.000    12.365    is/ConvResult_20__0_n_106
                         DSP48E1 (Prop_dsp48e1_PCIN[47]_PCOUT[47])
                                                      1.452    13.817 r  is/ConvResult_20__1/PCOUT[47]
                         net (fo=1, unplaced)         0.000    13.817    is/ConvResult_20__1_n_106
                         DSP48E1 (Prop_dsp48e1_PCIN[47]_P[0])
                                                      1.286    15.103 r  is/ConvResult_20__2/P[0]
                         net (fo=784, unplaced)       0.584    15.686    r2/D[0]
                         FDRE                                       r  r2/memout_reg[0][0]/D
```

Slack: inf (무한대)

Source: nc2/count_2_reg[3]/C (rising edge-triggered cell FDCE)

Destination: r2/memout_reg[0][0]/D

Path Type: Max at Slow Process Corner

Data Path Delay: 15.686ns

Logic Delay: 10.778ns (68.710%)

Route Delay: 4.908ns (31.290%)

-synthesis결과 요약:

이 FPGA 디자인에서는 주로 LUT와 레지스터가 사용되고 있으며, 전체 자원의 사용률은 비교적 낮습니다. 전체적으로 자원 사용률이 여유가 있어 추가적인 로직을 추가할 수 있는 공간이 충분합니다. Memory (RAM) 자원은 사용되지 않았습니다. DSPs 자원은 일부 사용되었습니다 (5.42%).

Timing 경로에서 slack이 inf로 표시되었으며, 이는 특정 경로에서 타이밍 경로가 완전히 설정되지 않

았거나 타이밍 분석 도구에서 경로를 제대로 인식하지 못했음을 의미합니다. slack이 inf로 표시된 이유를 조사하여 필요한 타이밍 제약 조건을 설정하거나 경로 설정을 확인하여 최적의 회로로 개선할 수 있을 것 같습니다. 팀원들이 예상하기에는 output에 값이 제대로 나오게 설정하지 않았음을 원인으로 꼽았습니다.

6. 팀원의 역할

기본적으로 모든 코드와 시뮬레이션을 같이 진행했습니다. 특별하게 문제 해결에 대한 아이디어를 냈던 부분에 대해서만 간략하게 설명하겠습니다.

- 전상우: 이미지를 필터에 맞게 shifting 하는 부분 설계, 필터의 성분이 음수일 때 오류가 난 것을 확인하고 덧셈기 로직 수정. Sharpening의 1, 2단 레이어 설계

- 장원: convolution을 위해 shifting하는 카운터 설계, 각 cell마다 곱하여 합하는 방식 제시, 연산 결과를 레지스터 파일에 저장하는 부분 설계. Sharpening의 1, 2단 레이어 결합.