

# Sparse Convolution

## Simple Usage

See `main.cpp`.

## Files and Directories

```
1 |— acc_function
2 |   |— AcceleratorFunction.h
3 |   └─ winograd
4 |       |— WinogradFunction_1D.h
5 |       └─ WinogradFunction.h
6 |— CMakeLists.txt
7 |— feature_map
8 |   |— DirectFeatureMap.h
9 |   |— FeatureMap.h
10 |  |— Im2colFeatureMap.h
11 |  └─ RegularSparseFeatureMap.h // Regular sparse feature map (***HERE!**)
12 |— kernel
13 |   |— DirectKernel.cpp
14 |   |— DirectKernel.h
15 |   |— Im2colKernel.cpp
16 |   |— Im2colKernel.h
17 |   └─ Kernel.h
18 |— main.cpp
19 |— output_map
20 |   └─ OutputMap.h
21 |— resources
22 |   └─ pointcloud.npy
23 |— util
24 |   └─ GetTime.h
```

# Result and Analysis

- Output Channel: 64

Direct Conv	Im2col	Im2col with Winograd	Sparse
0.915951	0.697502	0.584971	0.160801
0.903722	0.692903	0.576382	0.160899
0.904667	0.702545	0.577662	0.15749
0.908973	0.693076	0.579219	0.157934
0.911708	0.703896	0.581641	0.158873
0.910727	0.691902	0.576274	0.156495
0.902105	0.711141	0.582059	0.157247
0.905044	0.711783	0.597845	0.158104

- Output Channel: 128

Direct Conv	Im2col	Im2col with Winograd	Sparse
1.80745	1.36492	1.18049	0.316016
1.76588	1.37583	1.16024	0.323528
1.77178	1.39502	1.14257	0.319032
1.78286	1.38325	1.16363	0.314616
1.78322	1.39397	1.30103	0.315658
1.7917	1.3722	1.30384	0.315082
1.78269	1.35794	1.32393	0.312747
1.78361	1.3626	1.3176	0.312864

- Output Channel: 256

Direct Conv	Im2col	Im2col with Winograd	Sparse
3.59123	2.75523	2.3069	0.625079
3.55964	2.76718	2.32161	0.632573
3.56235	2.78616	2.31645	0.620093
3.66558	2.73761	2.26556	0.61666
3.51596	2.71245	2.26299	0.619295
3.52357	2.71904	2.26883	0.618134
3.52284	2.72193	2.26711	0.619413
3.5313	2.71624	2.26579	0.618611

- Output Channel: 512

Direct Conv	Im2col	Im2col with Winograd	Sparse
7.07761	5.46572	4.56377	1.25613
7.06851	5.48121	4.5649	1.2567
7.07422	5.50153	4.53772	1.25476
7.07425	5.4713	4.54637	1.2469
7.05984	5.42913	4.55577	1.25162
7.07253	5.46255	4.56222	1.25147
7.09781	5.4592	4.52476	1.24712
7.04763	5.44672	4.54061	1.24799

Inference time is linearly corresponding to the amount of output channel because for every output channel the convolution operation repeats one more time. Sparse convolution is 63% faster than Winograd on average.