

Filter

1	0	-1
1	0	-1
1	0	-1

Vertical

1	1	1
0	0	0
-1	-1	-1

Horizontal

10	10	10			
10	10	10		0	
10	10	10			
			10	10	10
	0		10	10	10
			10	10	10

6x6

* vertical =

0	0	0	0
30	10	-10	-30
30	10	-10	-30
0	0	0	0

4x4

1	0	-1
1	0	-1
1	0	-1

1		-1
2		-2
1		-1

Sobel

3		-3
10		-10
3		-3

Scharr

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

use NN

Padding Stride

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

6x6 → 8x8
h x h pad=1

3x3

f x f

4x4

$n-f+1 \times n-f+1$
 $h+2p-f+1 \times h+2p-f+1$

Padding = 0
Padding = 1
Stride = 1

Stride = 2

p=0

s=2

3x3

f x f

3x3

$\frac{h+2p-f}{s} + 1 \times$
切り捨て

この値を代入は、Conv 後の stride と同じサイズにできる。

3	4	5
1	x	2

→

7	2	5
---	---	---

数字のテキストでは

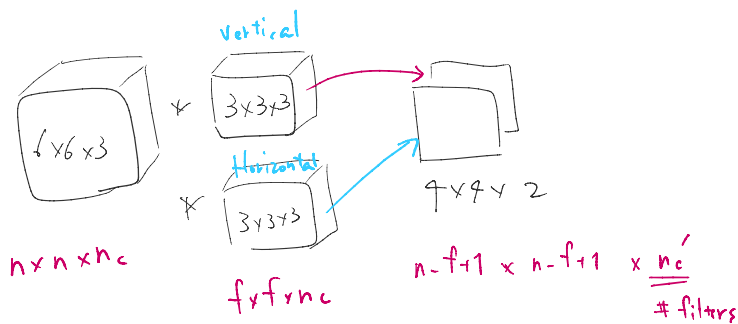
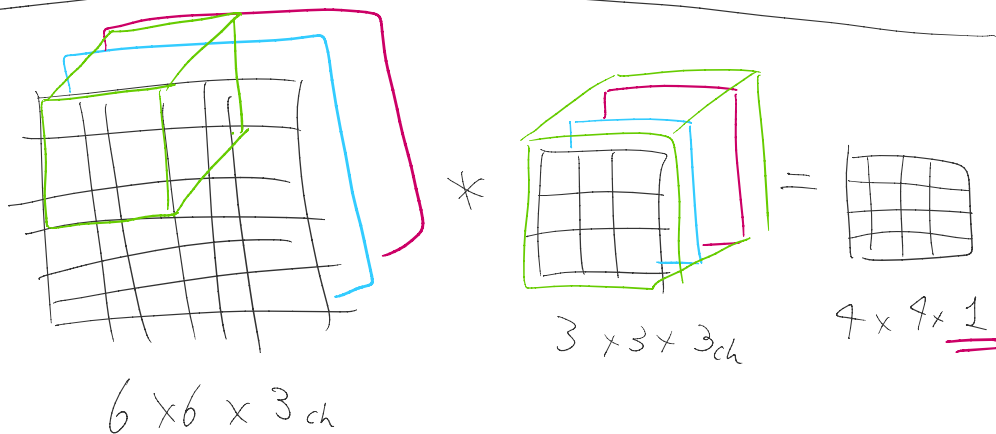
3	4	5
1	0	2
-1	9	7

filter 3x3

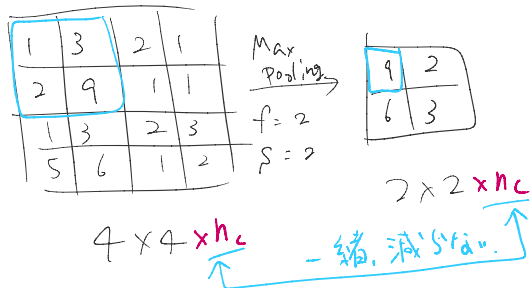
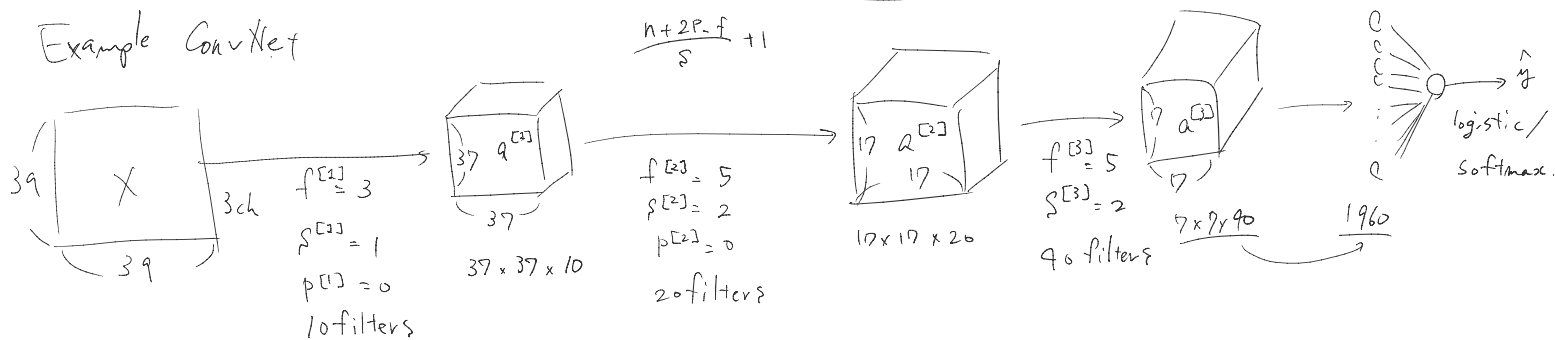


7	2	5
9	0	4
-1	1	3

数学的には
フィルタを変換した
ものを "Convolution" と呼ぶ。
しかし、Deep learning には、
フィルタをそのまま使うのを
"Convolution" と呼ぶ



Example ConvNet



Avg Pooling はあまり使われない。
f=2, s=2 は良く使われる
(だいたい半分サイズになる)