ED5340 - Data Science: Theory and Practise

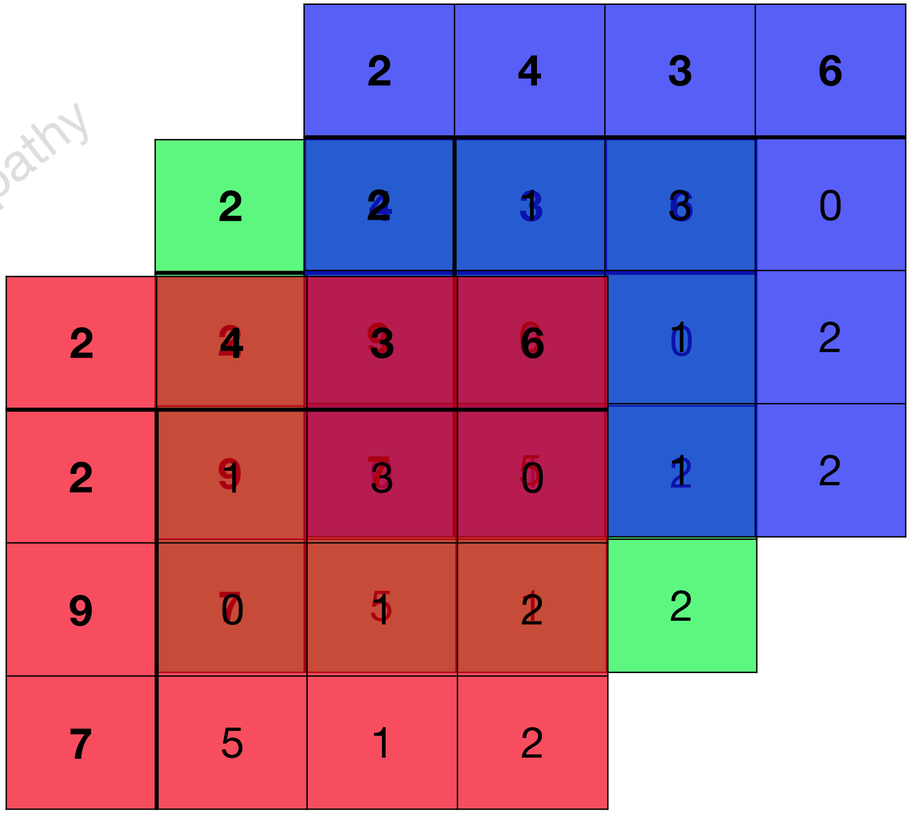
L28A - Popular Deep Networks

Ramanathan Muthuganapathy

Ramanathan Muthuganapathy (https://ed.iitm.ac.in/~raman)
Course web page: https://ed.iitm.ac.in/~raman/datascience.html
Moodle page: Available at https://courses.iitm.ac.in/

Three channels - RGB

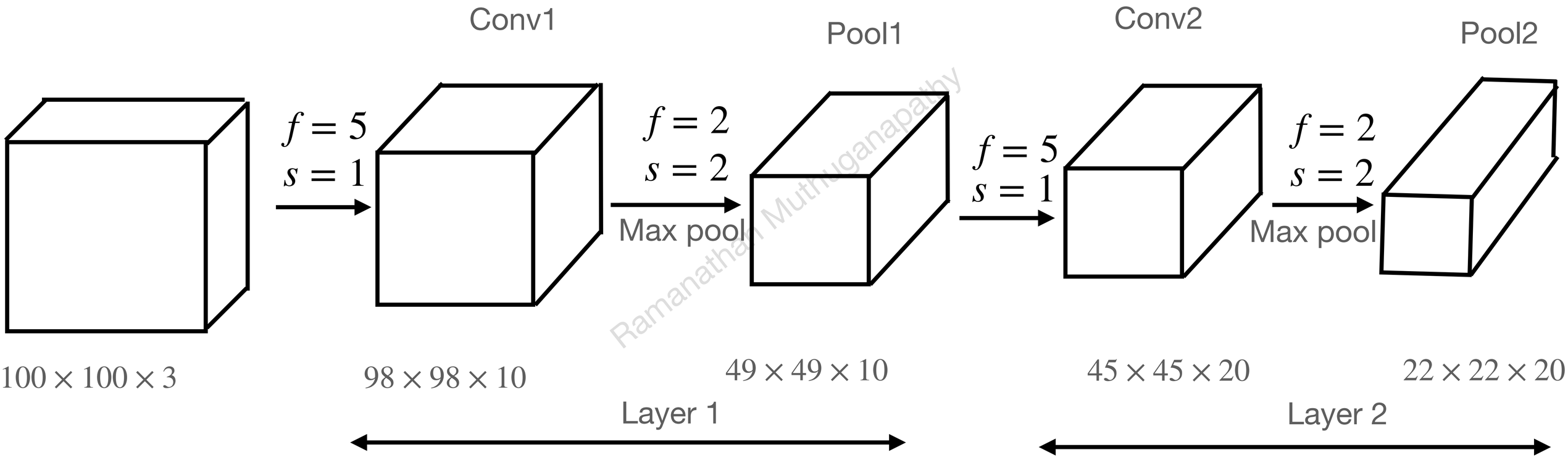
- RGB (Three channels, one for each colour)



Ramanathan Muthuganapathy

Example

Number of parameters



5 X 5 X 3 X 10 + 10 = 760

0

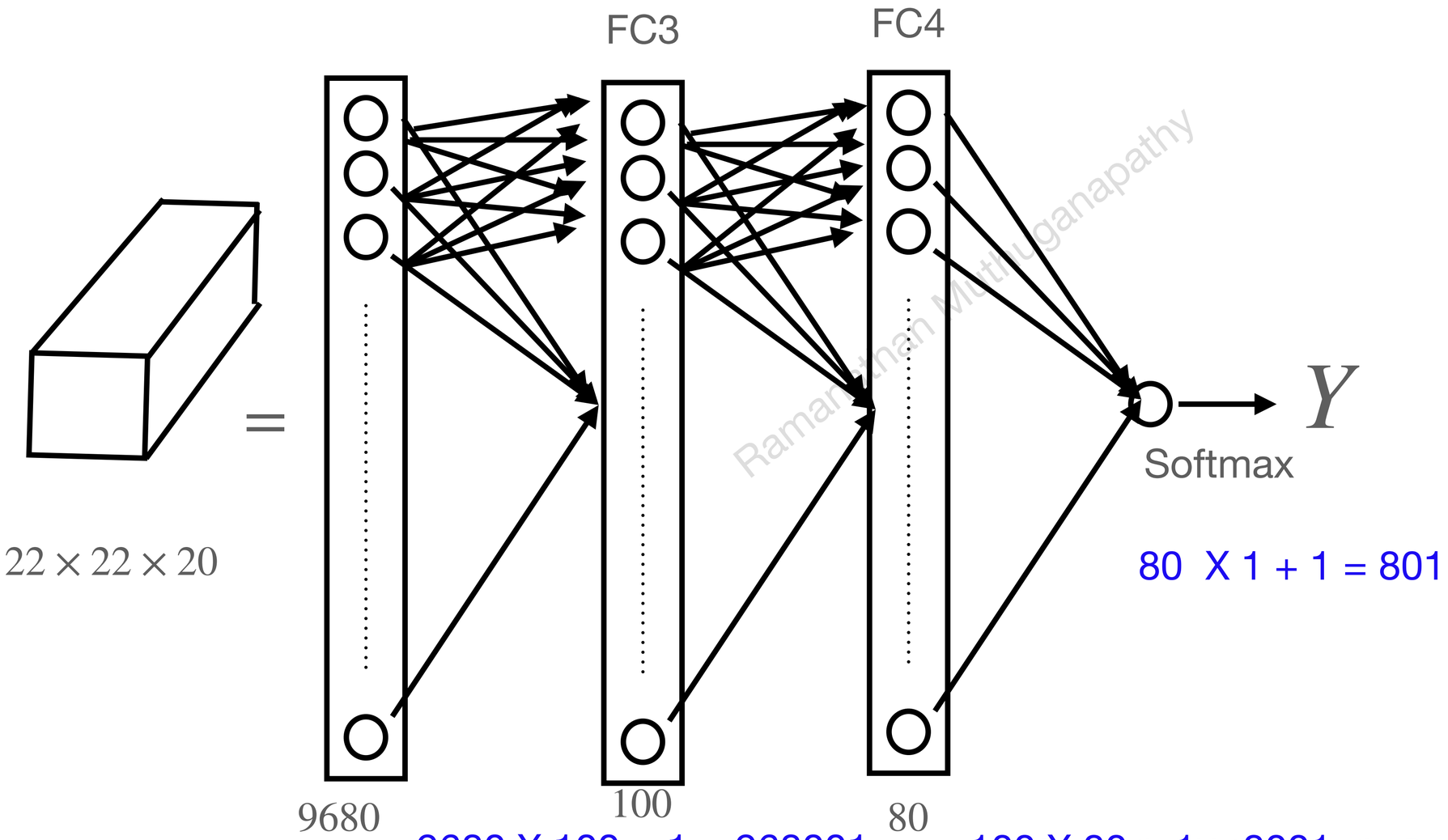
5 X 5 X 20 + 20 = 520

0

Ramanathan Muthuganapathy

Example

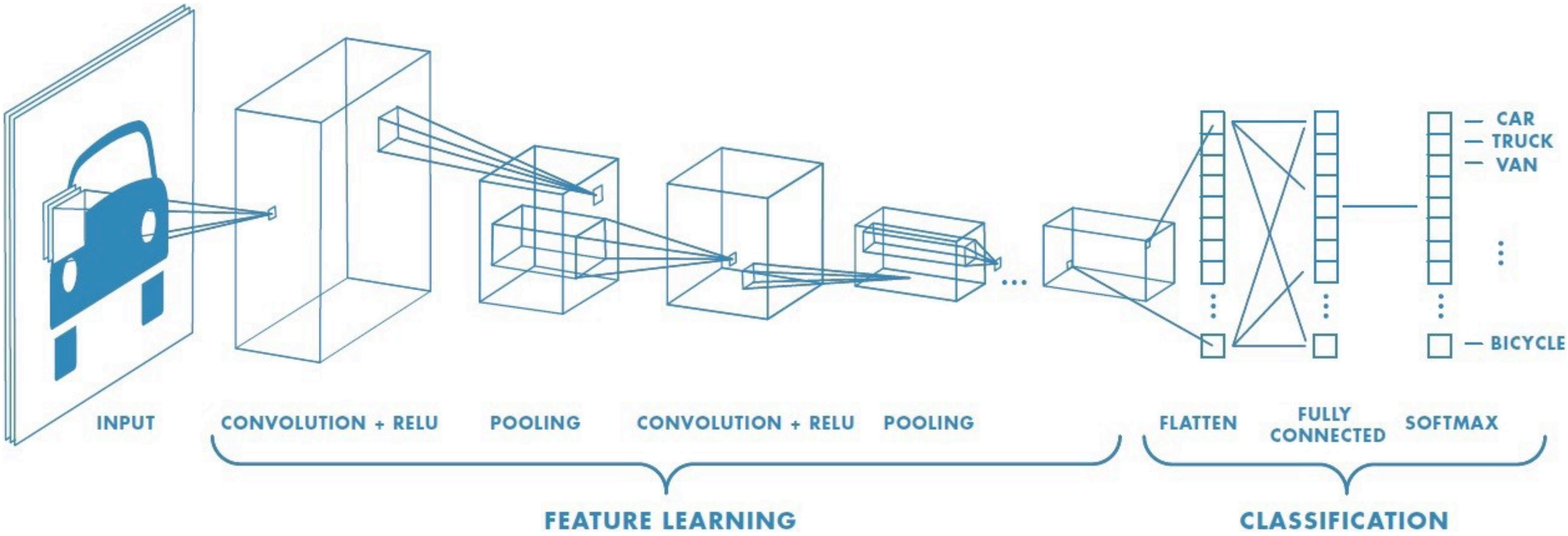
Number of parameters



Ramanathan Muthuganapathy

CNN

towardsdatascience.com



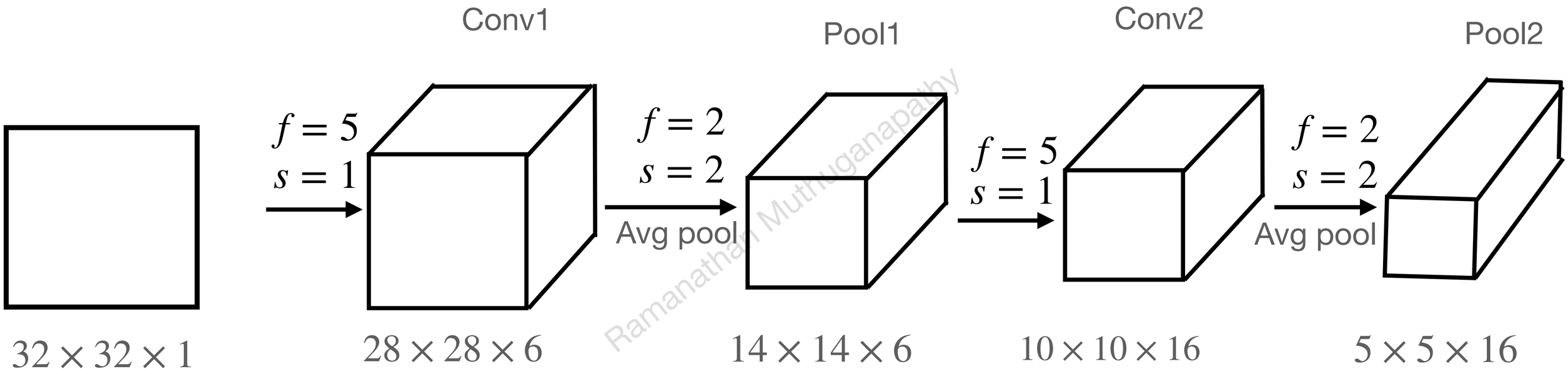
Popular networks

- LeNet
- AlexNet
- VGG-16
- ResNet
- InceptionNet (gogLeNet)

Ramanathan Muthuganapathy

LeNet

Le Cun et al. 1998, Gradient-based learning applied to document recognition

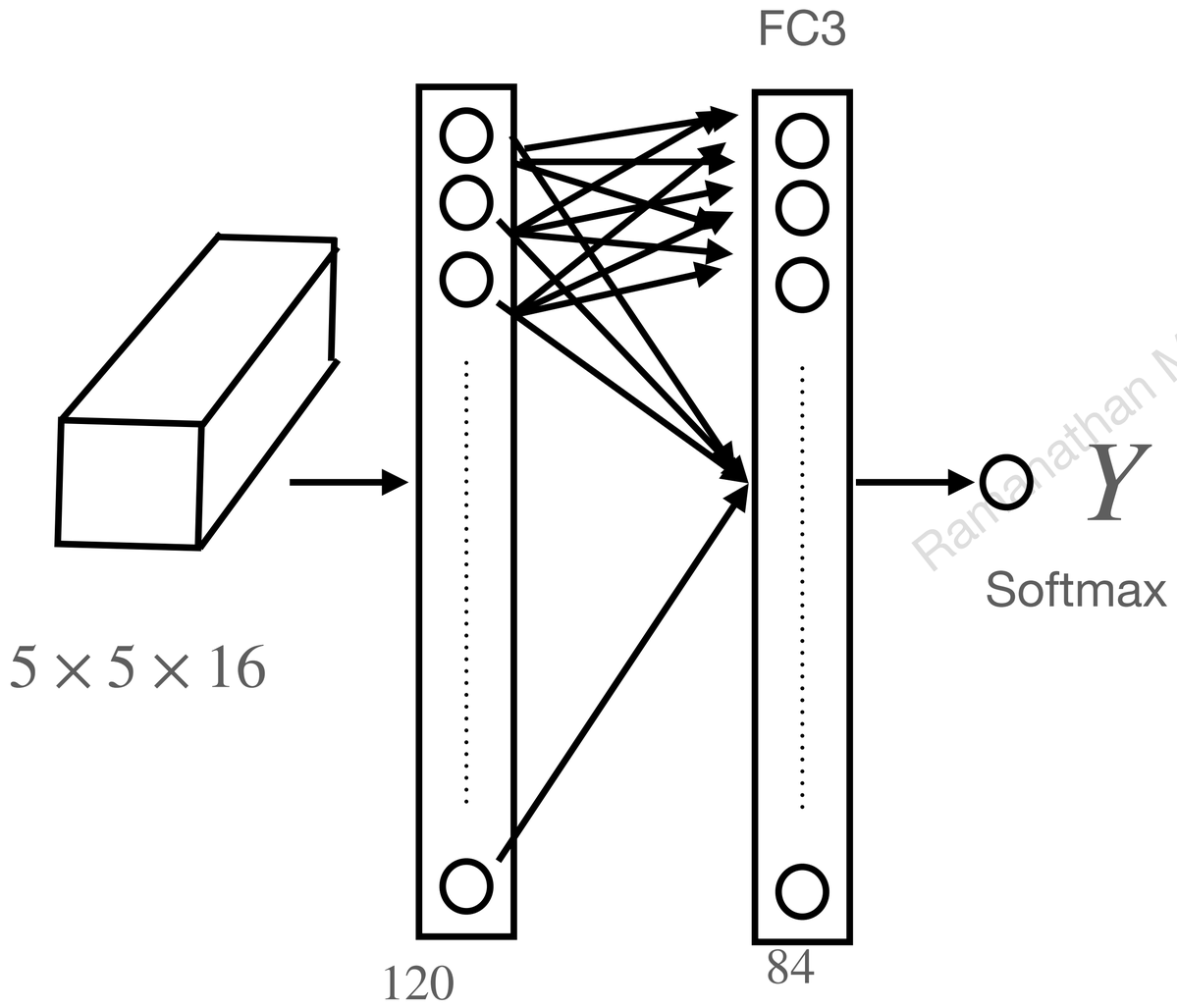


Ramanathan Muthuganapathy

5 X 5 X 16 + 20 = 420 0

LeNet

FCs

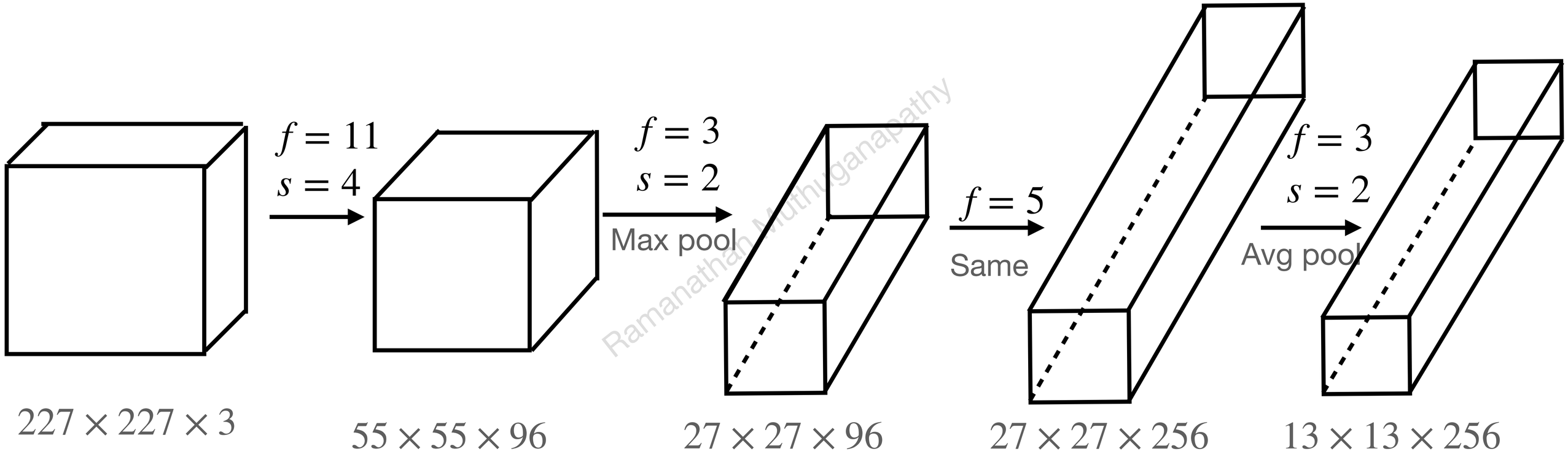


Ramanathan Muthuganapathy

60K Parameters

AlexNet

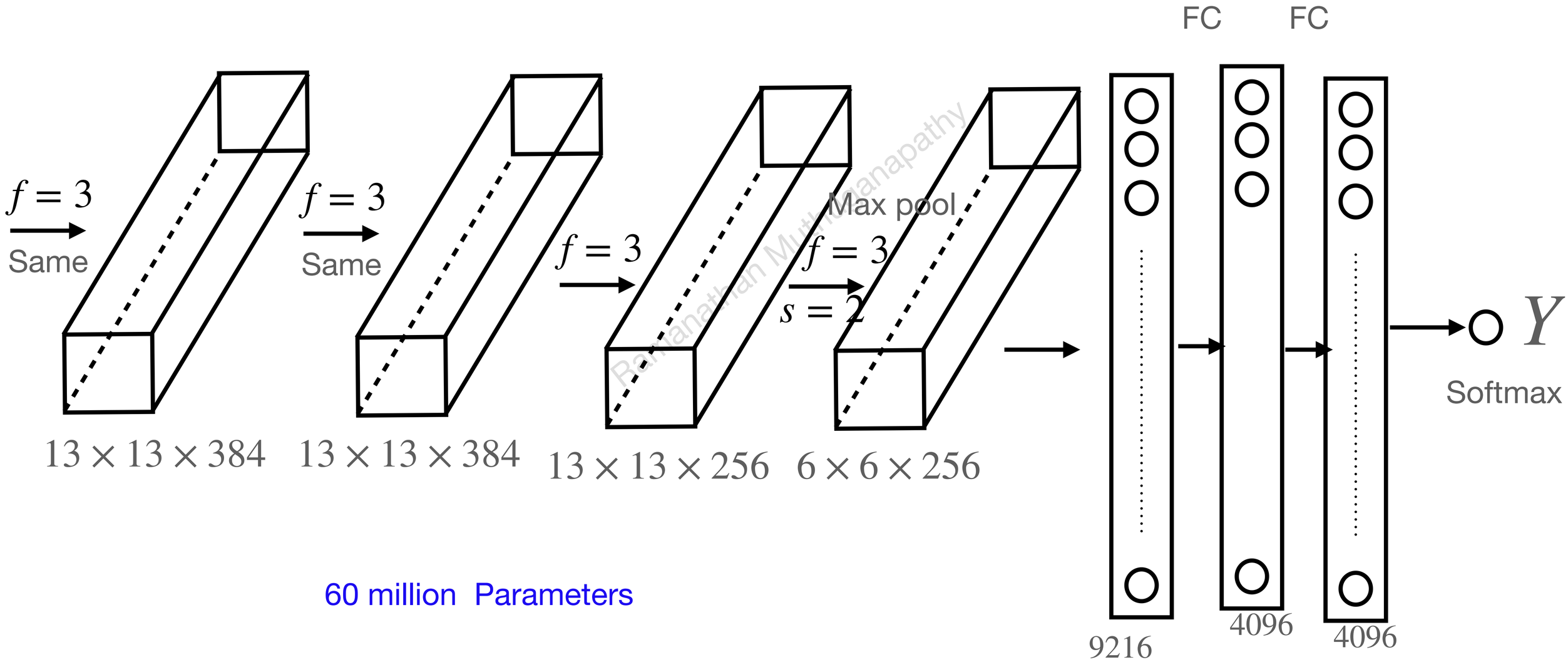
Krizhevsky et al. 2012, ImageNet classification with deep convolutional neural networks



Ramanathan Muthuganapathy

AlexNet

FCs



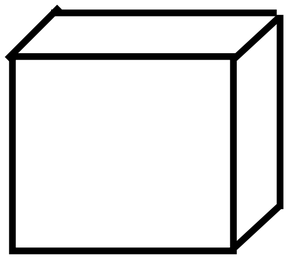
60 million Parameters

Ramanathan Muthuganapathy

VGG-16

Simonyan & Zisserman 2015, Very deep convolutional networks for large-scale image recognition

$f = 3$, $s = 1$ Same Max pool $f = 2$, $s = 2$



$224 \times 224 \times 3$ → Conv 64 X 2 → $224 \times 224 \times 64$ → Pool → $112 \times 112 \times 64$ → Conv 128 X 2 → $112 \times 112 \times 128$

→ Pool → $56 \times 56 \times 128$ → Conv 256 X 3 → $56 \times 56 \times 256$ → Pool → $28 \times 28 \times 256$ → Conv 512 X 3 → $28 \times 28 \times 512$

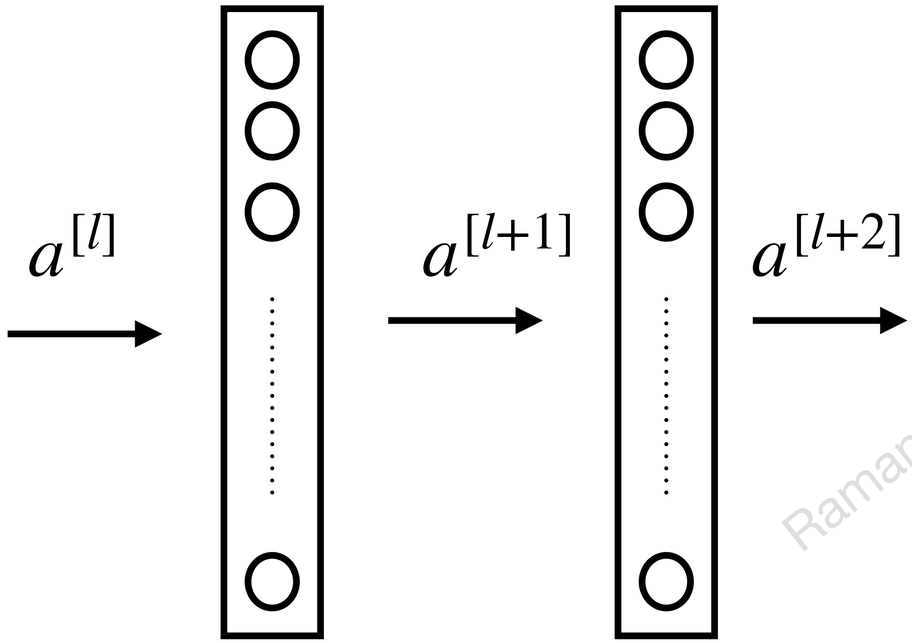
→ Pool → $14 \times 14 \times 512$ → Conv 512 X 3 → $14 \times 14 \times 512$ → Pool → $7 \times 7 \times 512$ → FC 4096 → FC 4096 → Softmax

138 million Parameters

Ramanathan Muthuganapathy

ResNet

He et al. 2015, Deep residual networks for image recognition



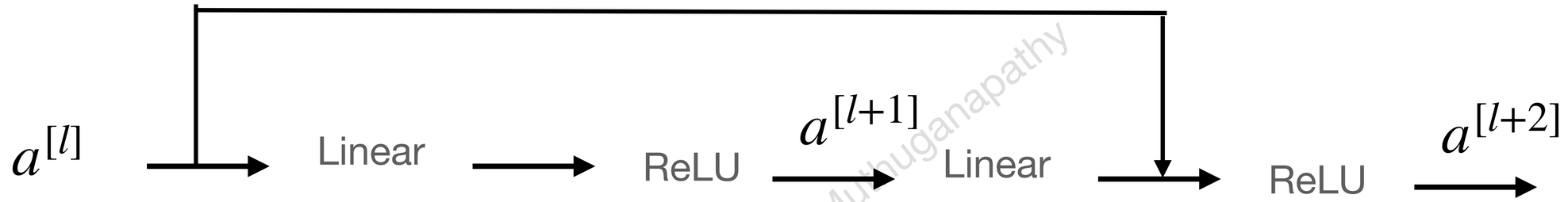
Ramanathan Muthuganapathy

$a^{[l]}$ → Linear → ReLU → $a^{[l+1]}$ → Linear → ReLU → $a^{[l+2]}$

$$a^{[l+2]} = g(z^{[l+2]})$$

ResNet

Key Idea - Skip Connection (short cut)

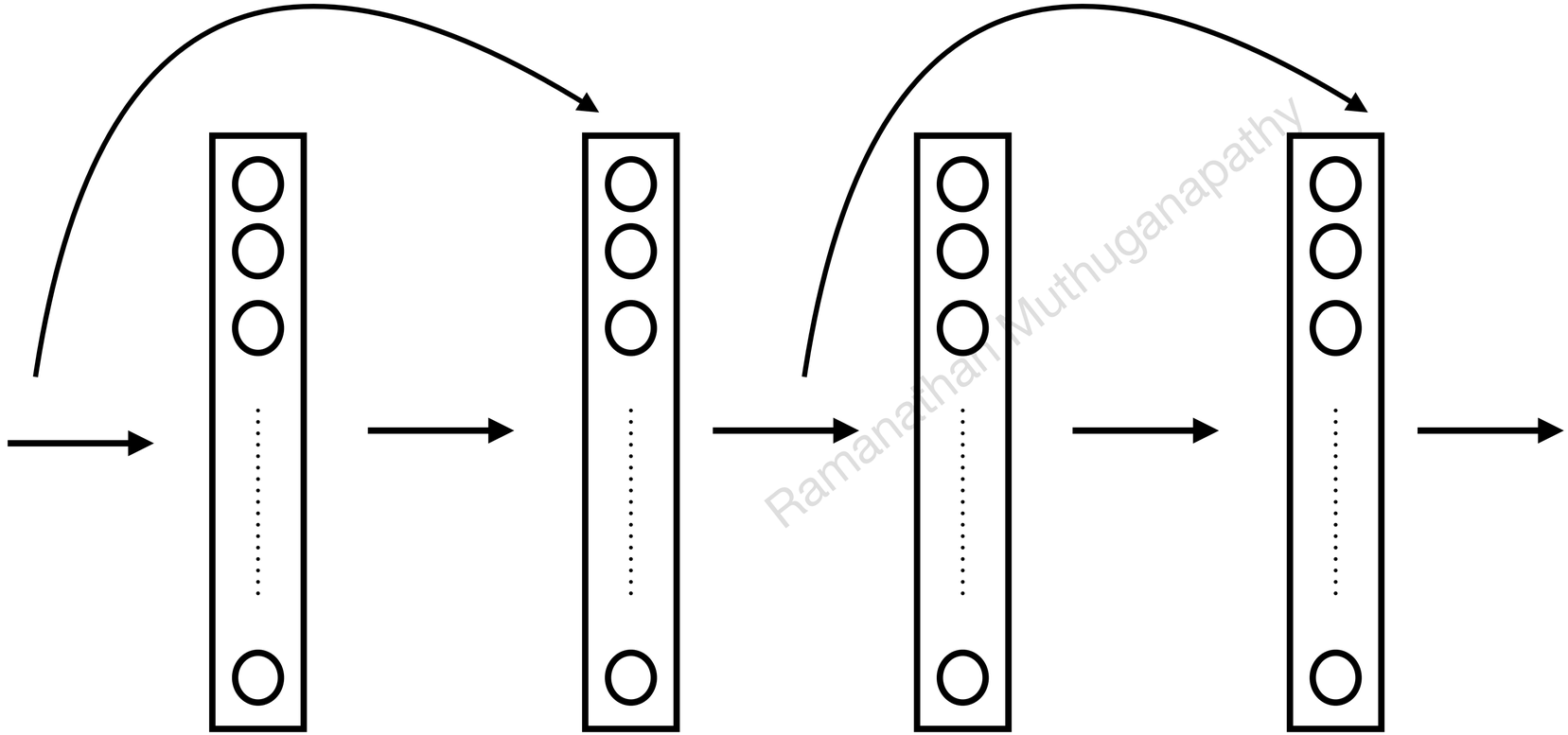


Ramanathan Muthuganapathy

$$a^{[l+2]} = g(z^{[l+2]} + a^{[l]})$$

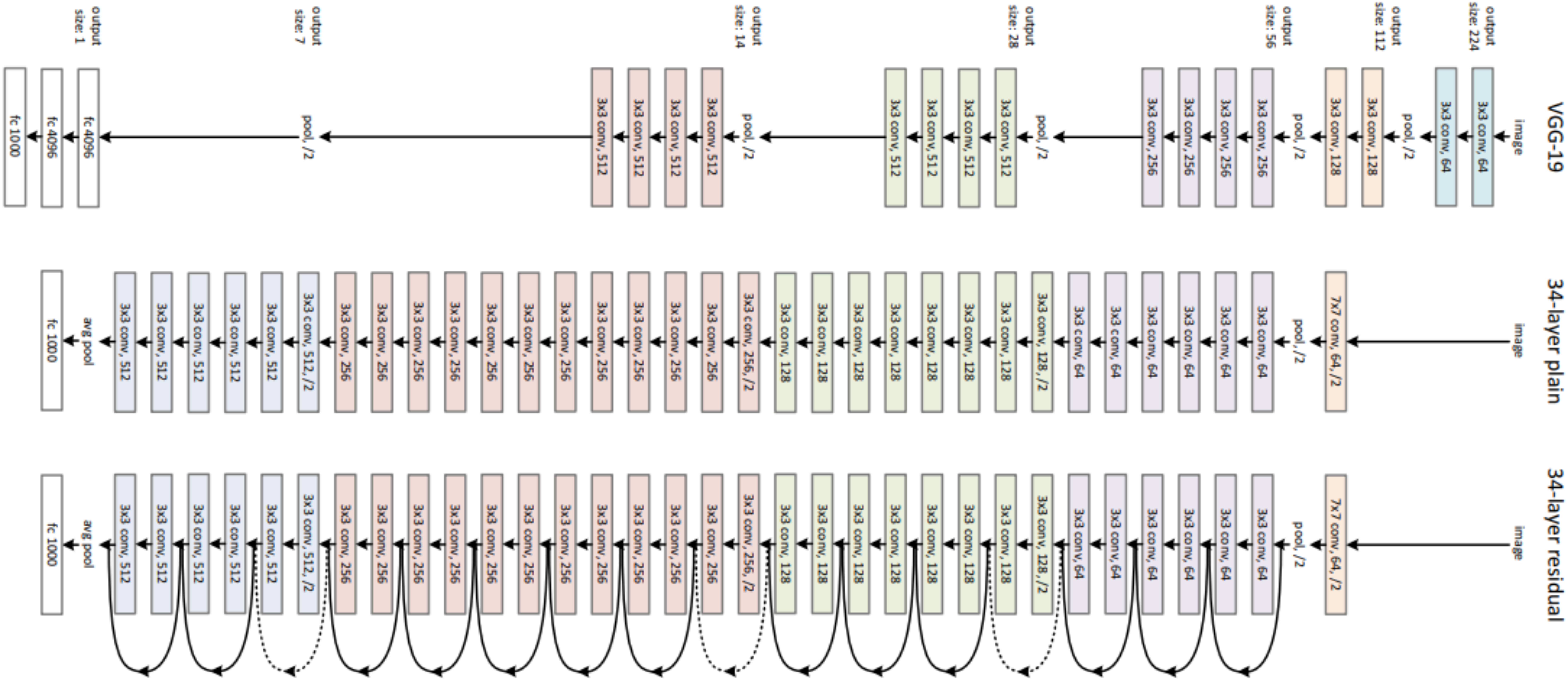
ResNet

Skip Connection



Ramanathan Muthuganapathy

VGG-19 and ResNet



Convolution

1 X 1

2	**4**	**3**	**6**
2	1	3	0
9	0	1	2
7	5	1	2

*

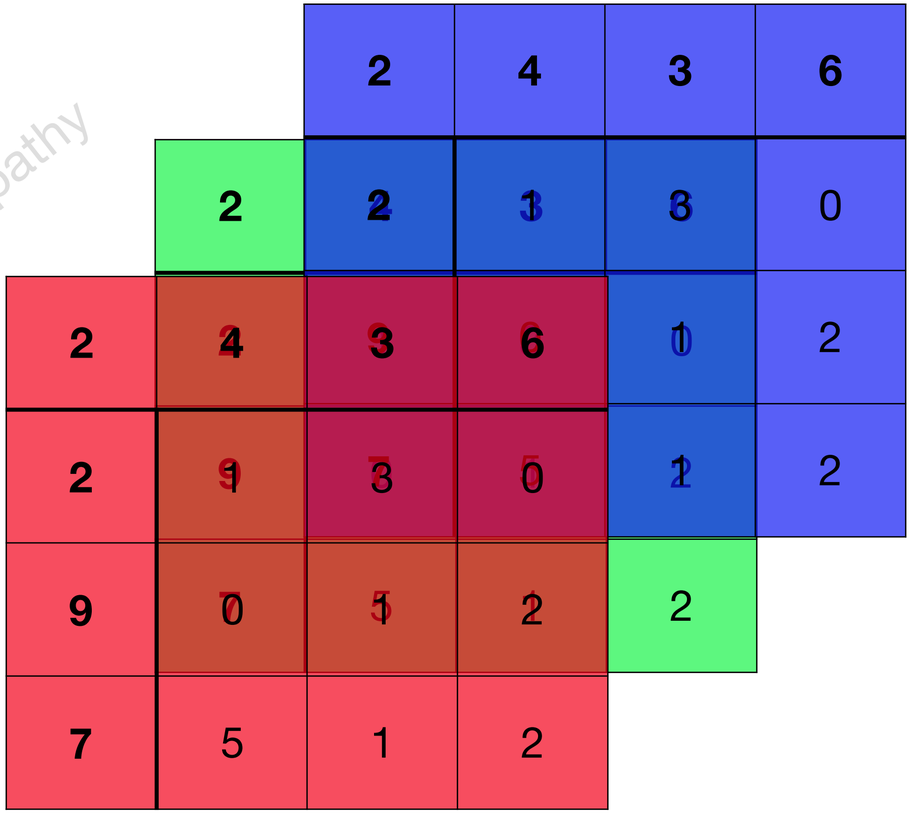
1

Ramanathan Muthuganapathy

=

Three channels - RGB

- RGB (Three channels, one for each colour)



Ramanathan Muthuganapathy

Three channels - RGB

- RGB (Three channels, one for each colour)

Ramanathan Muthuganapathy

2	**4**	**3**	**6**
2	1	3	0
9	0	1	2
7	5	1	2

2	**4**	**3**	**6**
2	1	3	0
9	0	1	2
7	5	1	2

2	**4**	**3**	**6**
2	1	3	0
9	0	1	2
7	5	1	2

1 X 1 Convolution on RGB

2D Convolution

2	**4**	**3**	**6**
2	1	3	0
9	0	1	2
7	5	1	2

2	**4**	**3**	**6**
2	1	3	0
9	0	1	2
7	5	1	2

2	**4**	**3**	**6**
2	1	3	0
9	0	1	2
7	5	1	2

Ramanathan Muthuganapathy

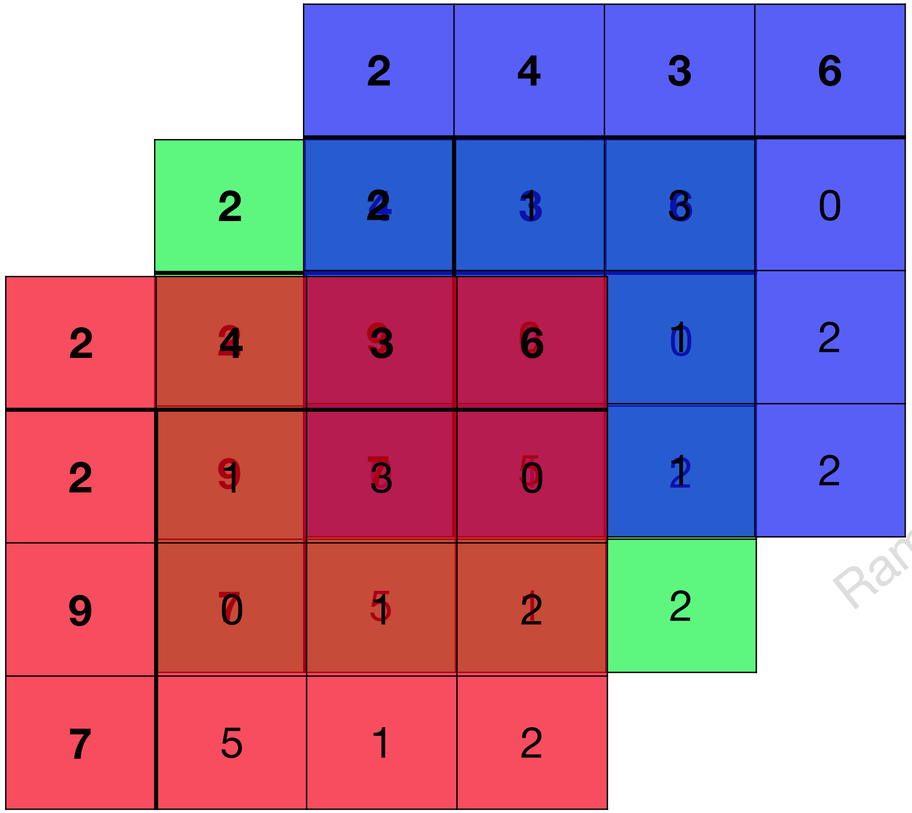
1

1

1

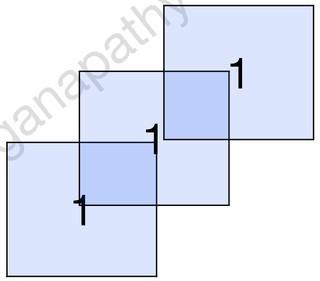
1 X 1 Convolution on RGB

2D Convolution (Stride = 1)



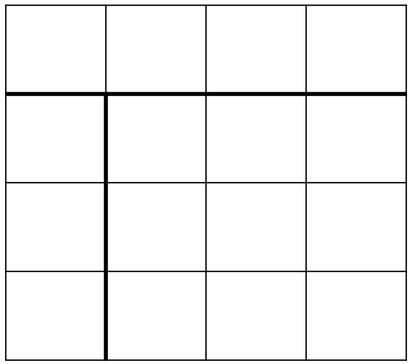
n X n X 3

*



1 X 1X 3

=



Ramanathan Muthuganapathy

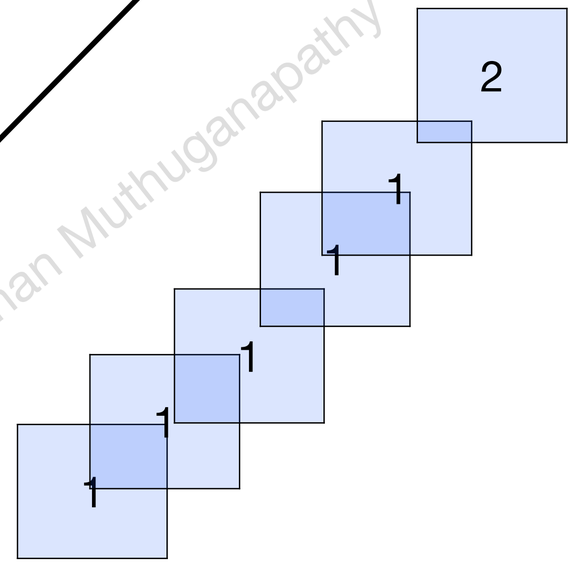
1 X 1 Convolution on RGB

2D Convolution (Stride = 1)

2	**4**	**3**	**6**
2	1	3	0
9	0	1	2
7	5	1	2

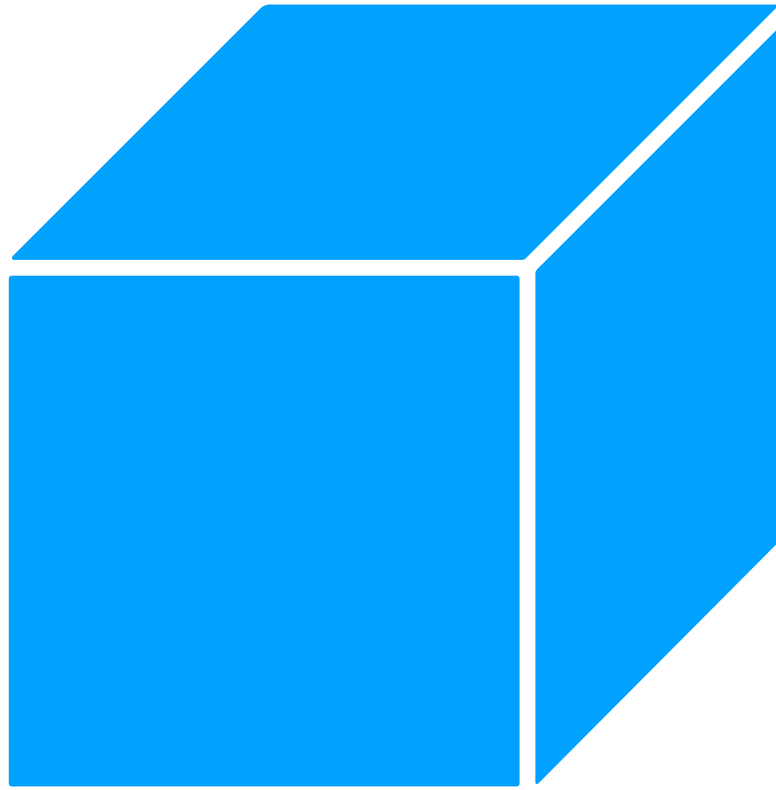
n X n X 256

*



32

=



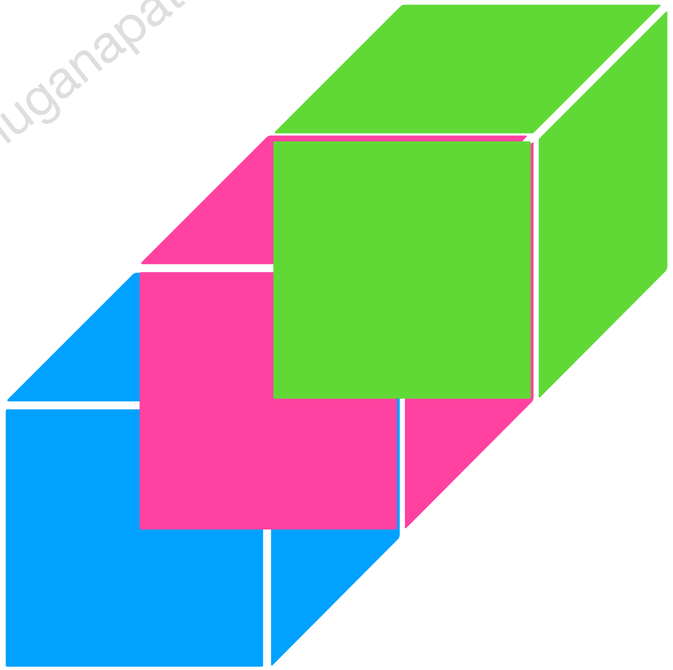
n X n X 32

Ramanathan Muthuganapathy

Inception module / Network

- 1 X 1, 3 X 3, etc..
- Stacking of volumes
- Pooling (‘same’)
- Several such modules in inception network
- Huge number of parameters
- See the paper for more details!

Ramanathan Muthuganapathy



Inception Network - Motivation

https://knowyourmeme.com/photos/531557-we-need-to-go-deeper