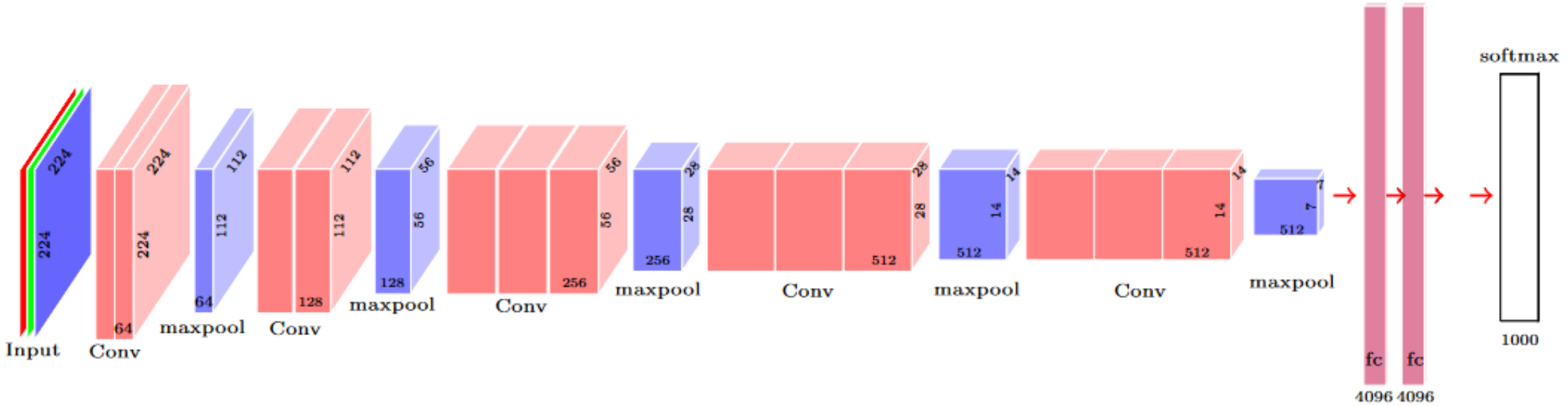
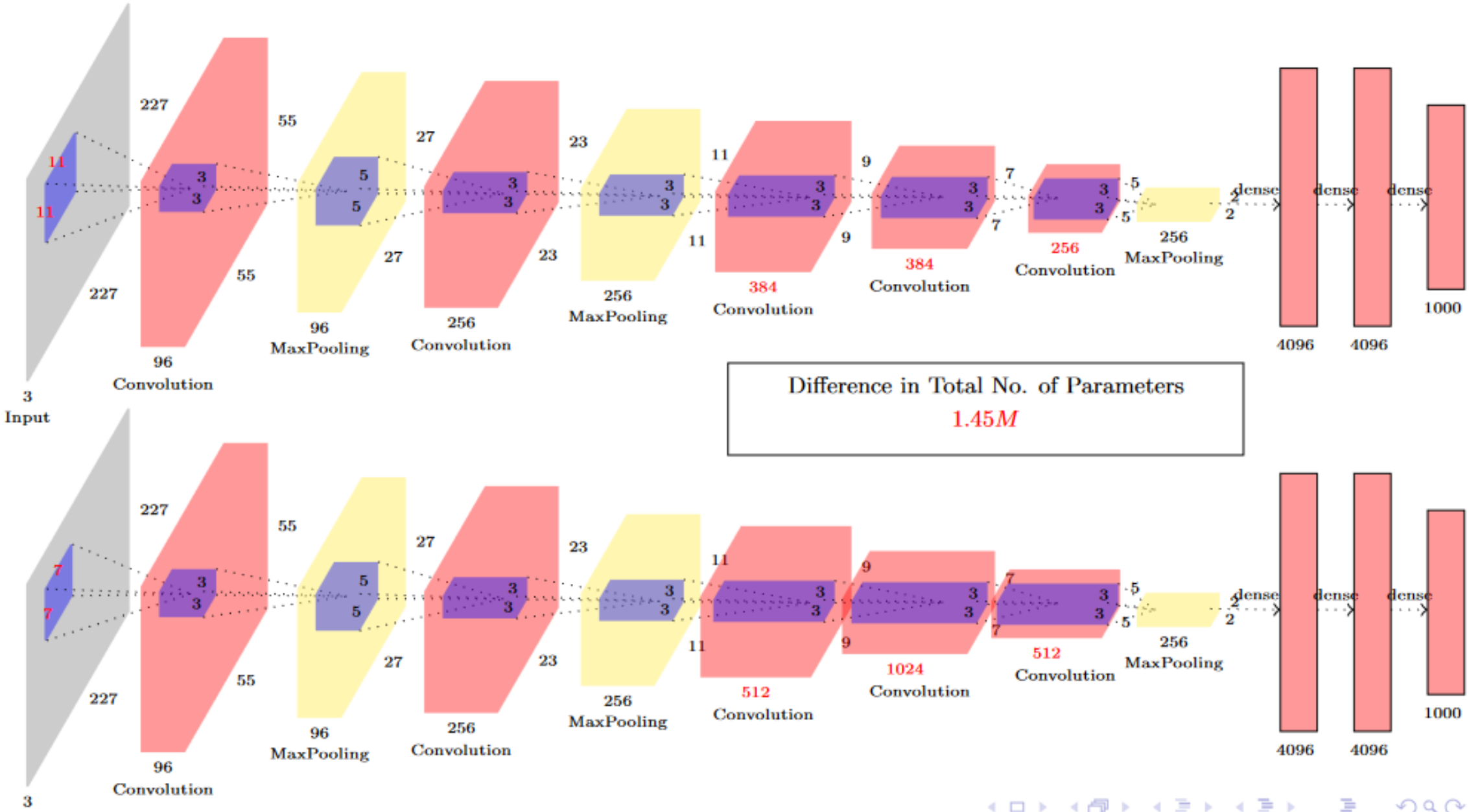
VGG NET

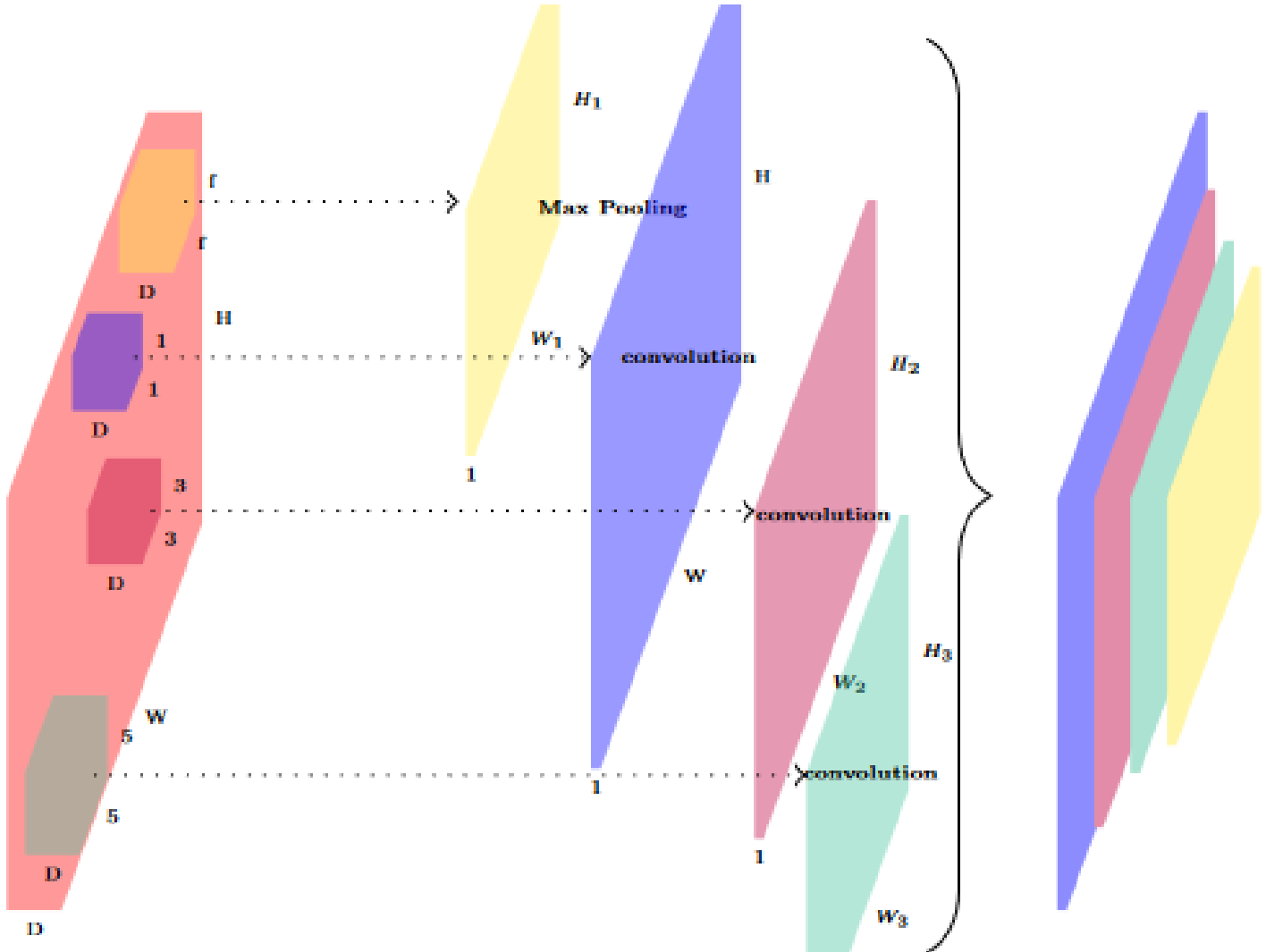


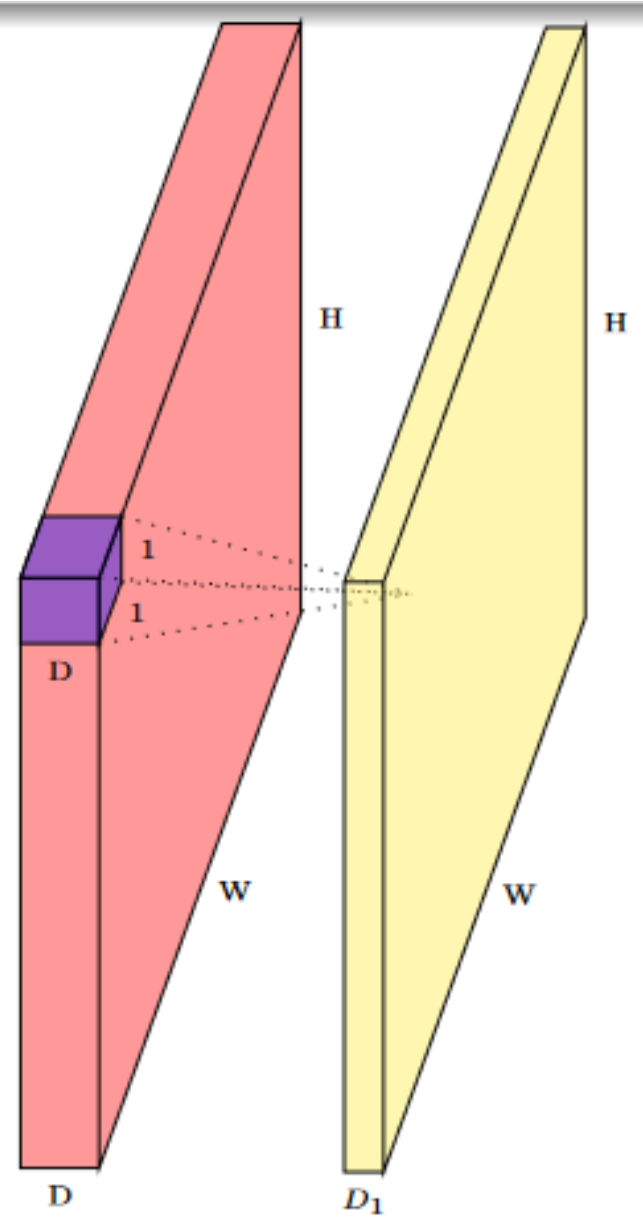
- Kernel size is 3×3 throughout
- Total parameters in non FC layers $= \sim 16M$
- Total Parameters in FC layers $= (512 \times 7 \times 7 \times 4096) + (4096 \times 4096) + (4096 \times 1024) = \sim 122M$
- Most parameters are in the first FC layer ($\sim$ 102M)

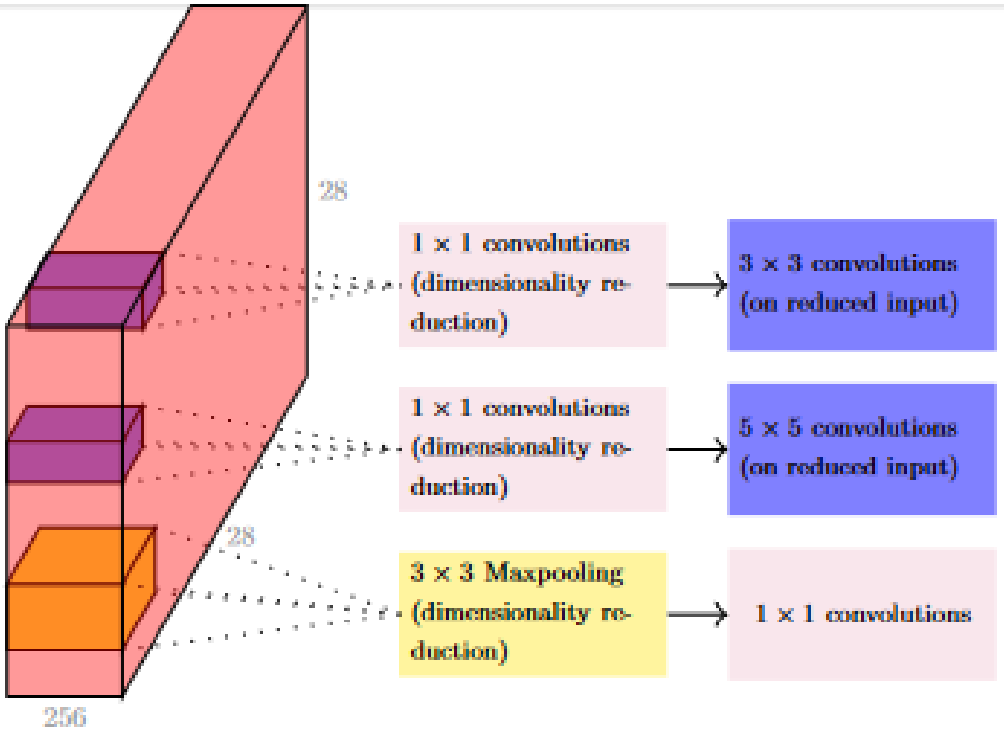
- **Alex NET**
- **ZF NET**

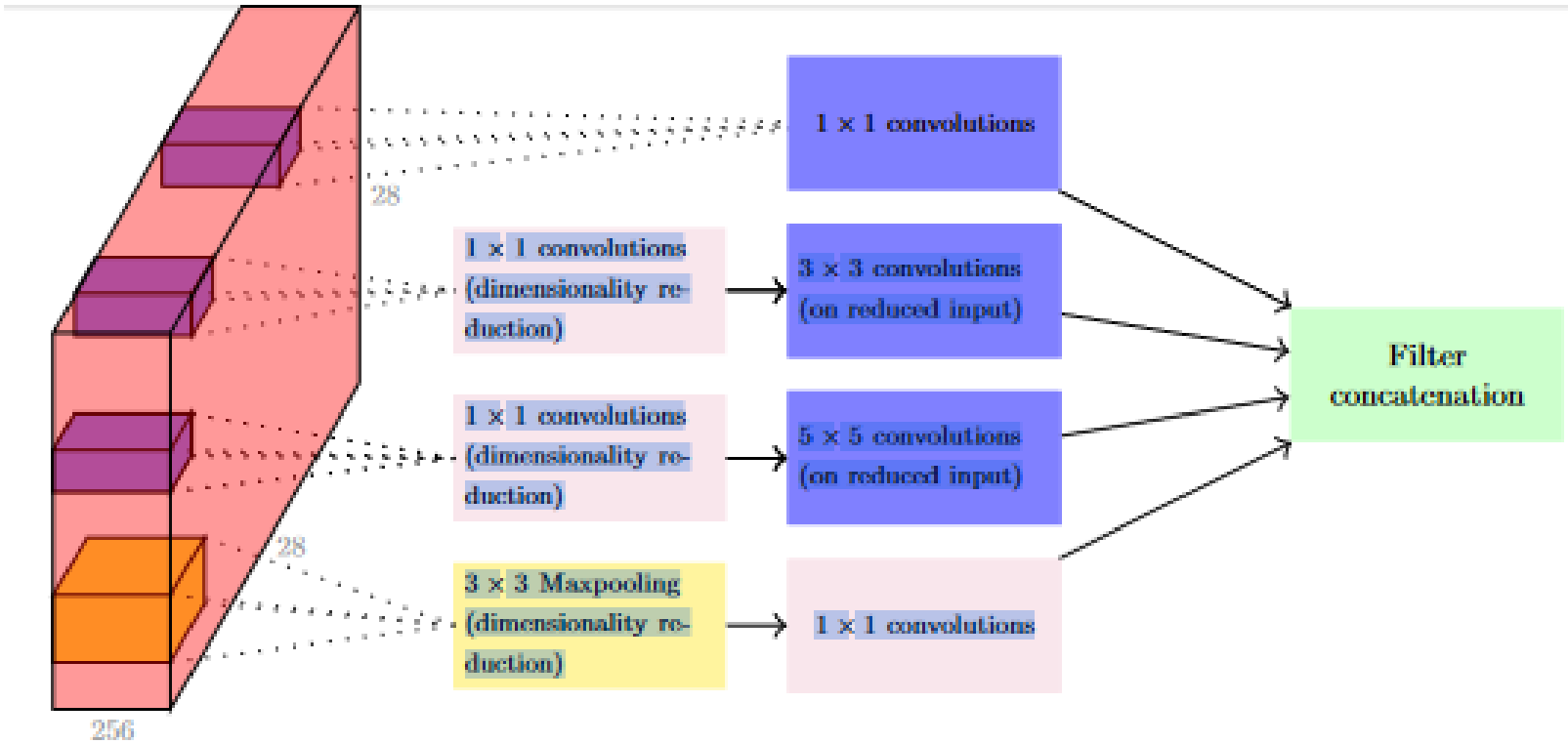

3
Input
227
11
11
227
96
Convolution
55
3
3
55
96
MaxPooling
27
5
5
27
256
Convolution
23
3
3
23
256
MaxPooling
11
3
3
11
384
Convolution
9
3
3
9
384
Convolution
7
3
3
7
256
Convolution
5
3
3
5
256
MaxPooling
2
2
dense
4096
dense
4096
dense
1000
Difference in Total No. of Parameters
1.45M
3
Input
227
7
7
227
96
Convolution
55
3
3
55
96
MaxPooling
27
5
5
27
256
Convolution
23
3
3
23
256
MaxPooling
11
3
3
11
512
Convolution
9
3
3
9
1024
Convolution
7
3
3
7
512
Convolution
5
3
3
5
256
MaxPooling
2
2
dense
4096
dense
4096
dense
1000

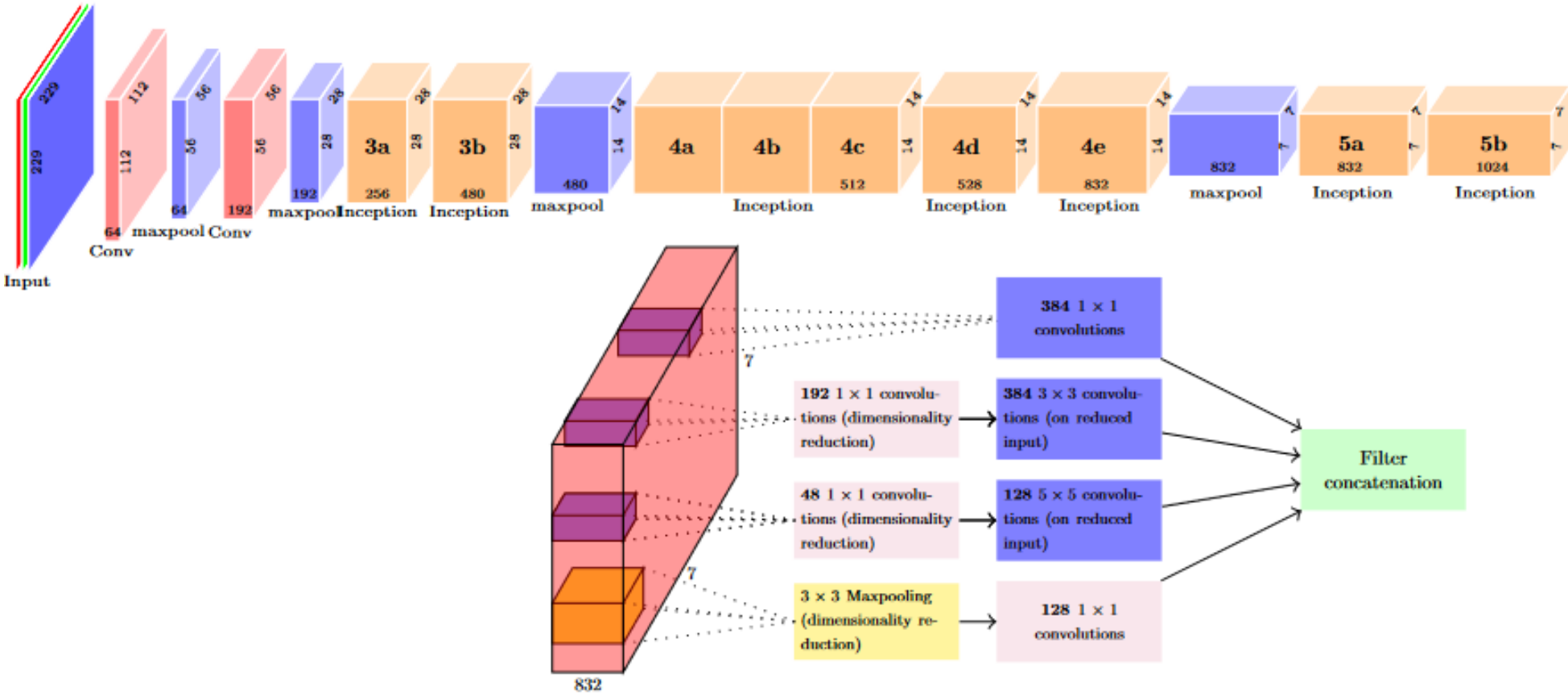
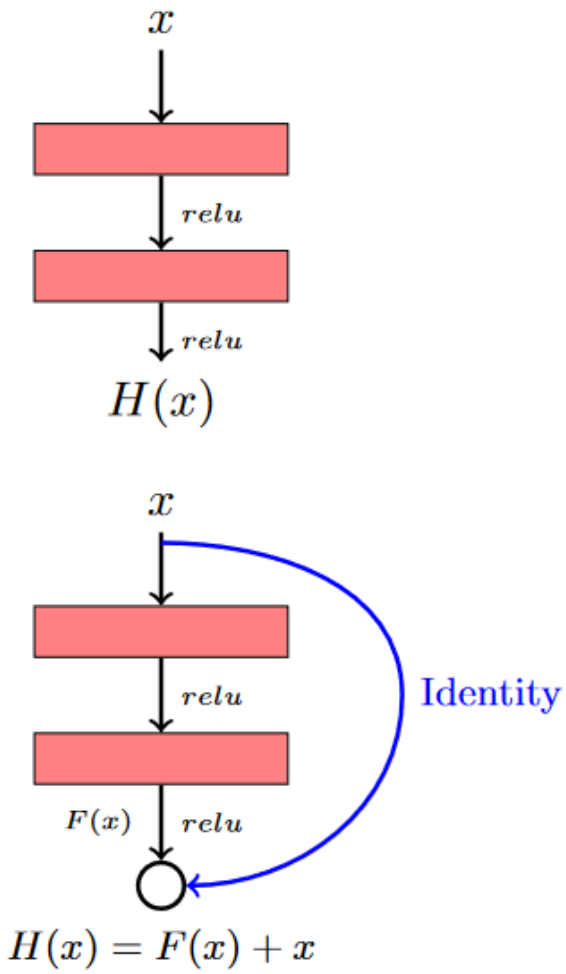
Google Inception net




H
H
1
1
D
W
W
D
D_1


28
28
256
1 × 1 convolutions (dimensionality reduction)
3 × 3 convolutions (on reduced input)
1 × 1 convolutions (dimensionality reduction)
5 × 5 convolutions (on reduced input)
3 × 3 Maxpooling (dimensionality reduction)
1 × 1 convolutions


28
28
256
1 × 1 convolutions
1 × 1 convolutions (dimensionality reduction)
3 × 3 convolutions (on reduced input)
1 × 1 convolutions (dimensionality reduction)
5 × 5 convolutions (on reduced input)
3 × 3 Maxpooling (dimensionality reduction)
1 × 1 convolutions
Filter concatenation


Input
Conv
maxpool
Conv
maxpool
Inception
Inception
maxpool
Inception
Inception
Inception
maxpool
Inception
Inception
3a
3b
4a
4b
4c
4d
4e
5a
5b
384 1 × 1 convolutions
192 1 × 1 convolutions (dimensionality reduction)
384 3 × 3 convolutions (on reduced input)
48 1 × 1 convolutions (dimensionality reduction)
128 5 × 5 convolutions (on reduced input)
3 × 3 Maxpooling (dimensionality reduction)
128 1 × 1 convolutions
Filter concatenation
832

Res NET