Efficient sliding window by converting fully-connected layers into convolutions
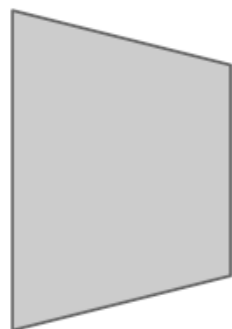

Image:
3 x 221 x 221

Convolution + pooling

Feature map:
1024 x 5 x 5

5 x 5 conv

5 x 5 conv

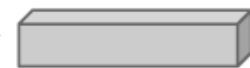4096 x 1 x 1

1 x 1 conv

1024 x 1 x 1

1 x 1 conv

Class scores:
1000 x 1 x 1

4096 x 1 x 1

1 x 1 conv

1024 x 1 x 1

1 x 1 conv

Box coordinates:
(4 x 1000) x 1 x 1