Module 2: Building Blocks for Image Recognition

Video 5: VGG + Hands-on

Analytics Vidhya

Introduction

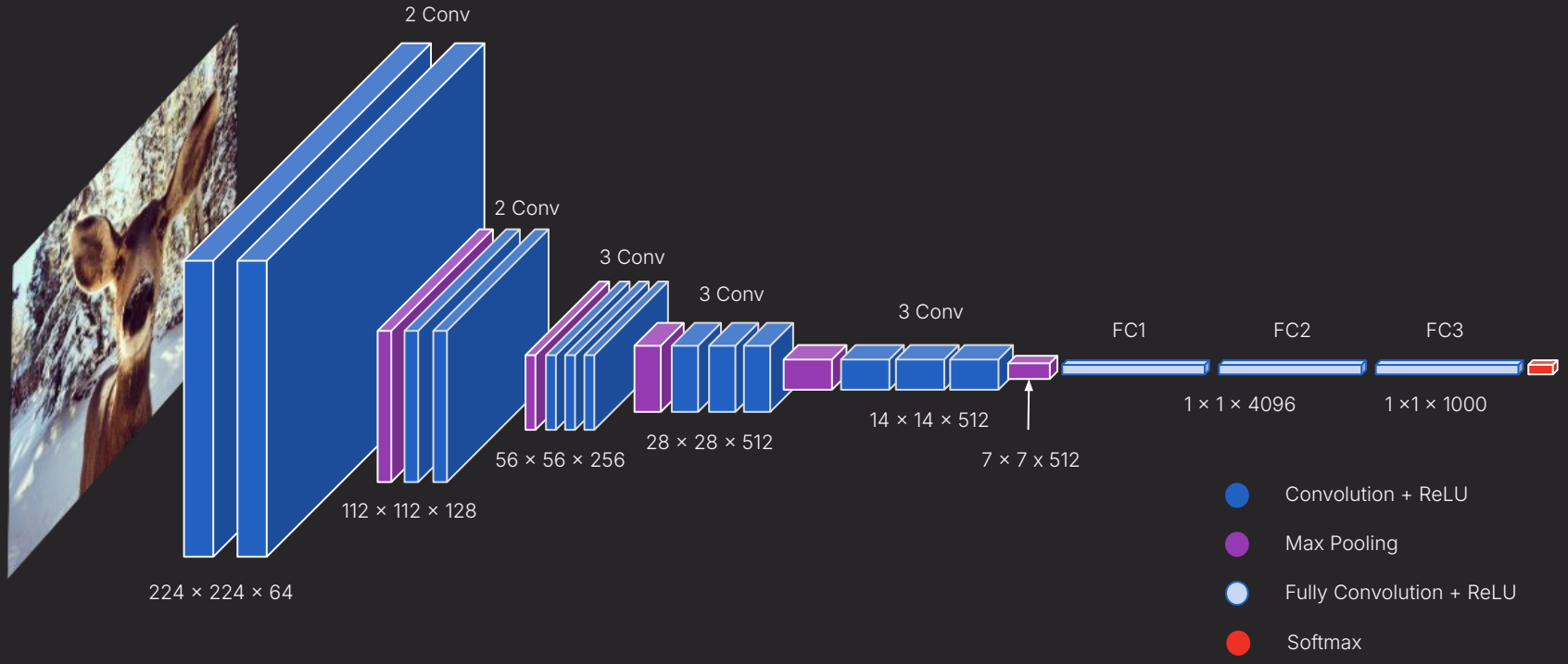
Developed by **Visual Geometry Group (VGG)** by Oxford University in **2014.**

Won ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014.

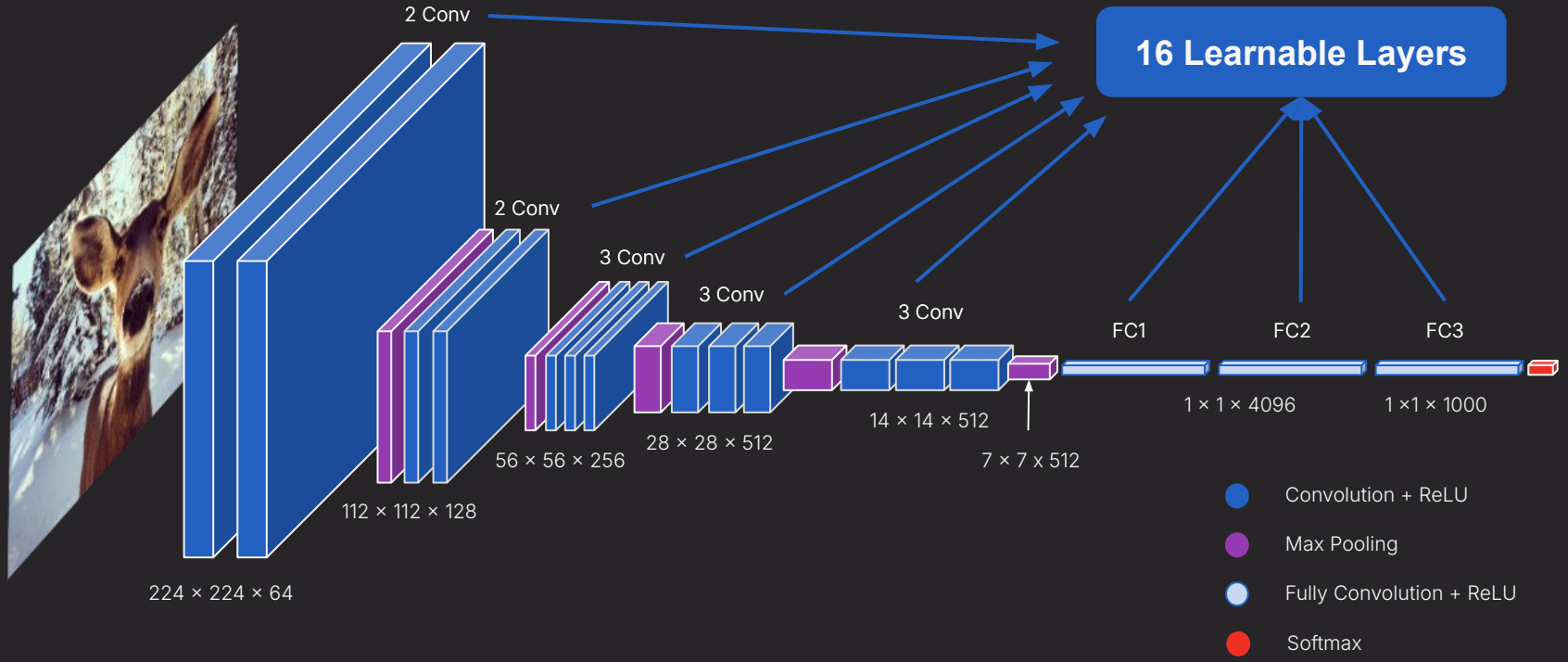
VGG Architecture exemplifies depth and simplicity in network design.

- Popular Variants: VGG16 and VGG19.

VGG 16

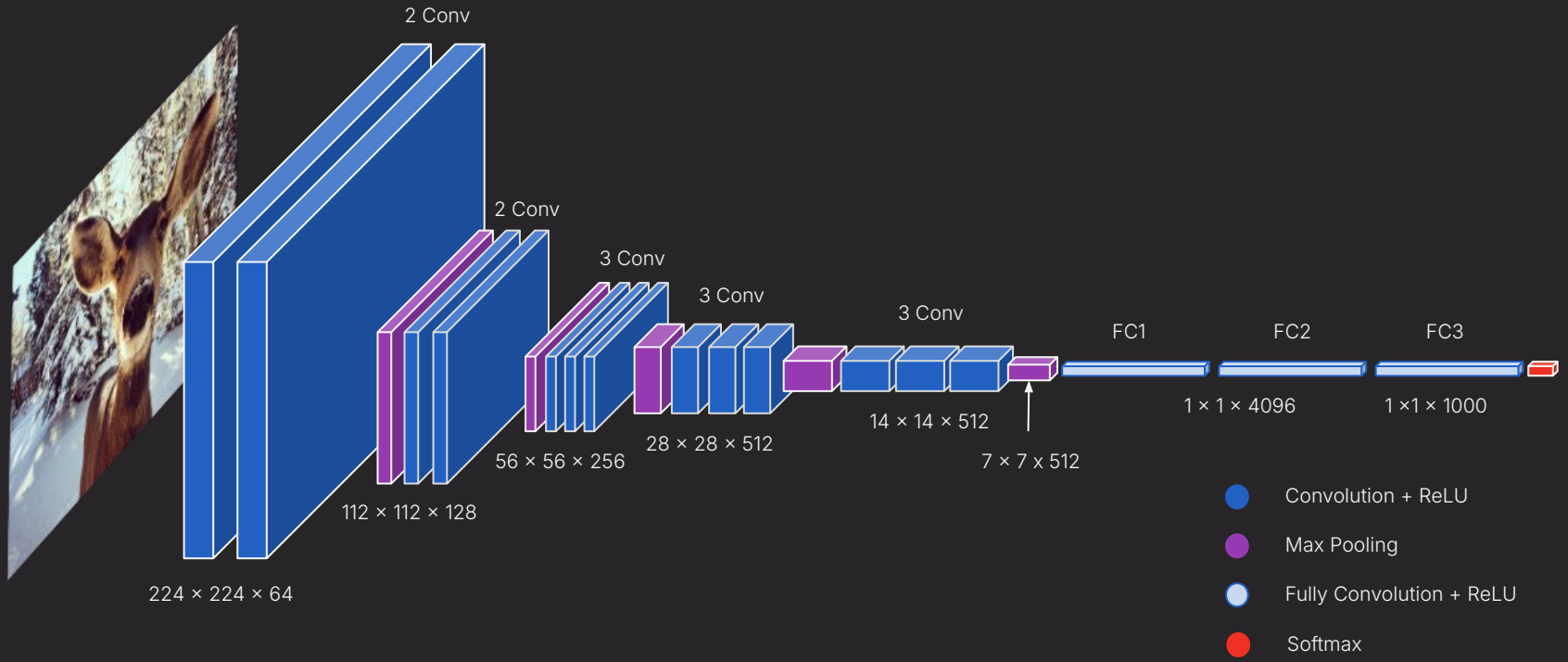

2 Conv
2 Conv
3 Conv
3 Conv
3 Conv
FC1
FC2
FC3
224 × 224 × 64
112 × 112 × 128
56 × 56 × 256
28 × 28 × 512
14 × 14 × 512
7 × 7 × 512
1 × 1 × 4096
1 ×1 × 1000
Convolution + ReLU
Max Pooling
Fully Convolution + ReLU
Softmax

VGG 16


2 Conv
2 Conv
3 Conv
3 Conv
3 Conv
16 Learnable Layers
FC1
FC2
FC3
224 × 224 × 64
112 × 112 × 128
56 × 56 × 256
28 × 28 × 512
14 × 14 × 512
7 × 7 x 512
1 × 1 × 4096
1 ×1 × 1000
Convolution + ReLU
Max Pooling
Fully Convolution + ReLU
Softmax

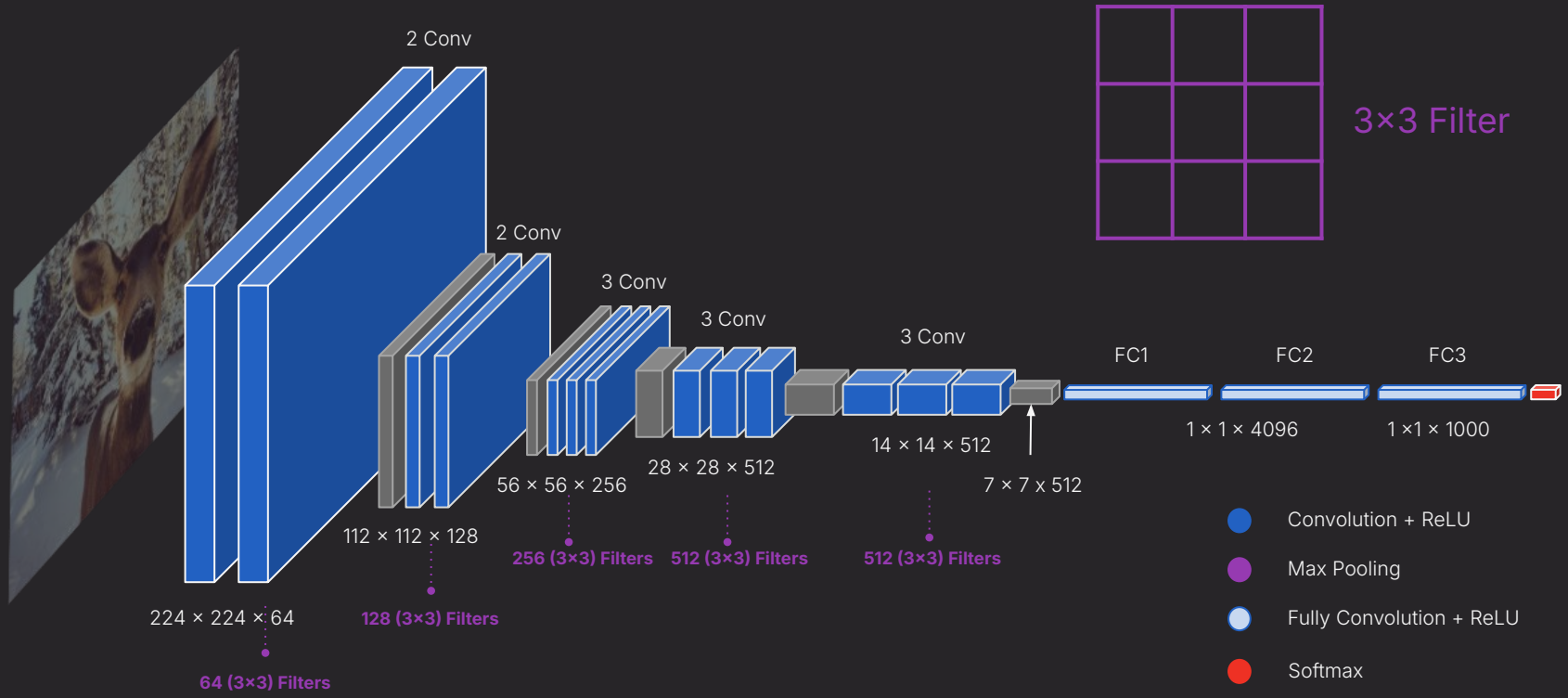
VGG 16

- 13 Convolutional Layers and 3 Fully Connected Layer.



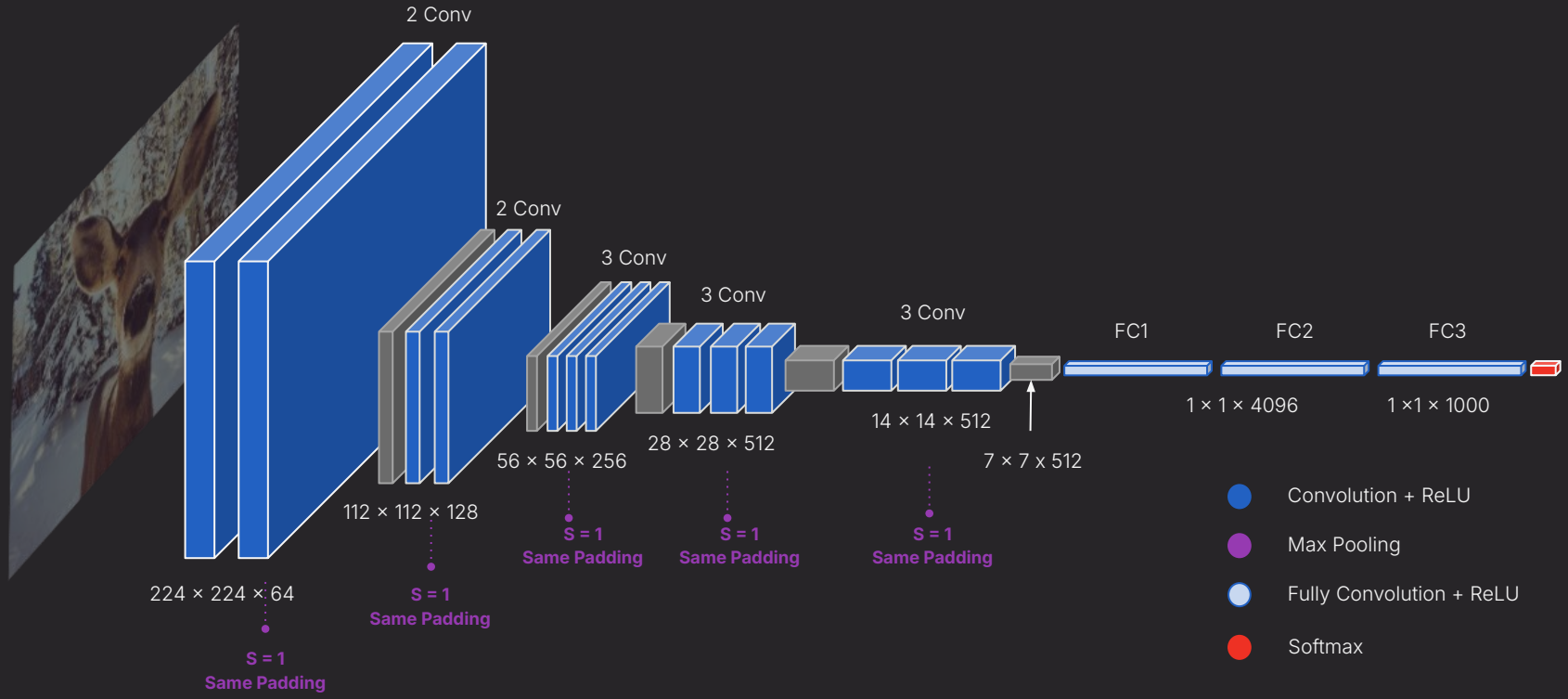
VGG 16

- Uses **3×3** filter throughout.



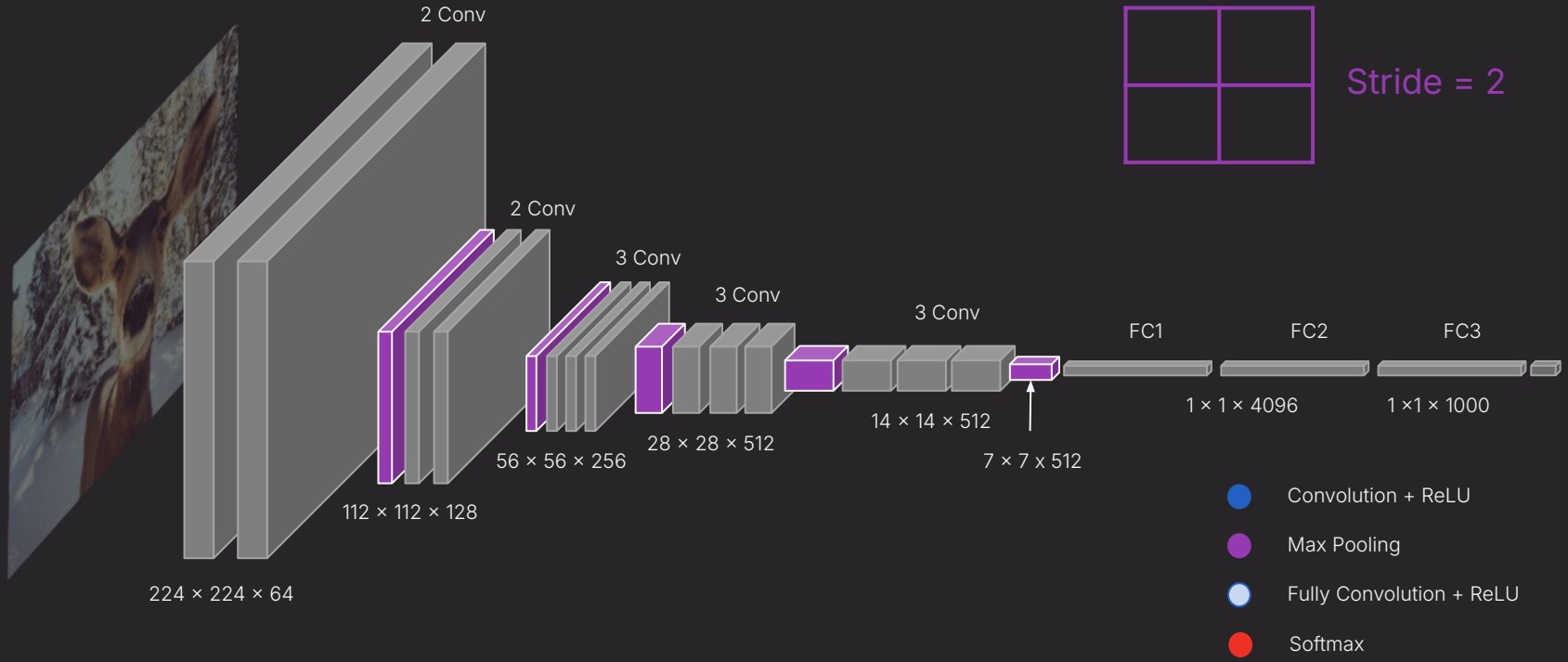
VGG 16

- Uses a **Stride of 1** and **same padding** throughout.



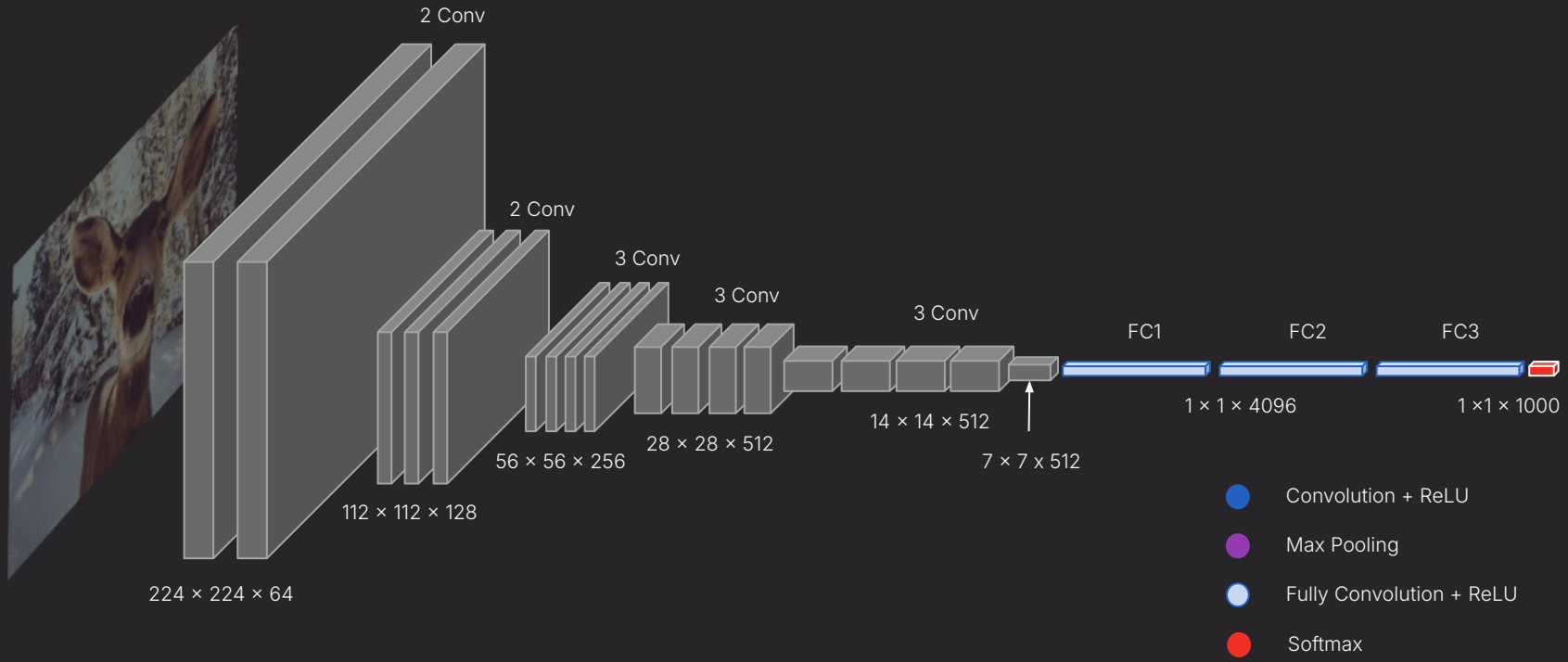
VGG16

- **5 Max Pooling layers** that acted as downsampling filters.



VGG16

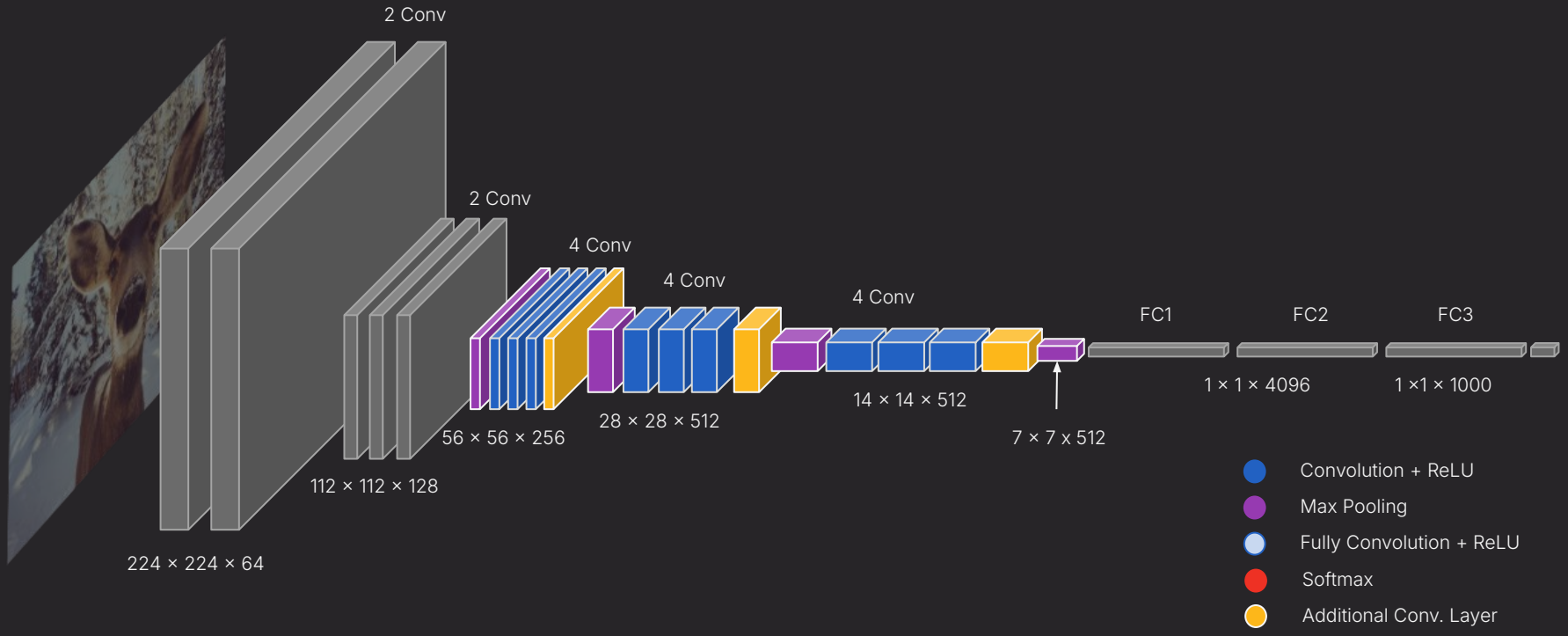
- **3 fully connected layers** culminate into a **softmax layer** that classified images.



Analytics Vidhya

VGG19

- VGG19 features the same VGG16 architecture but with **3 additional convolutional layers**.



Analytics Vidhya

UpNext: Hands-on