Model Architectures Comparison

Baseline CNN (Custom Architecture)

Input
224×224×3

Conv2D + BN + ReLU
64 filters, 3×3

MaxPool 2×2

Conv2D + BN + ReLU
128 filters, 3×3

MaxPool 2×2

Conv2D + BN + ReLU
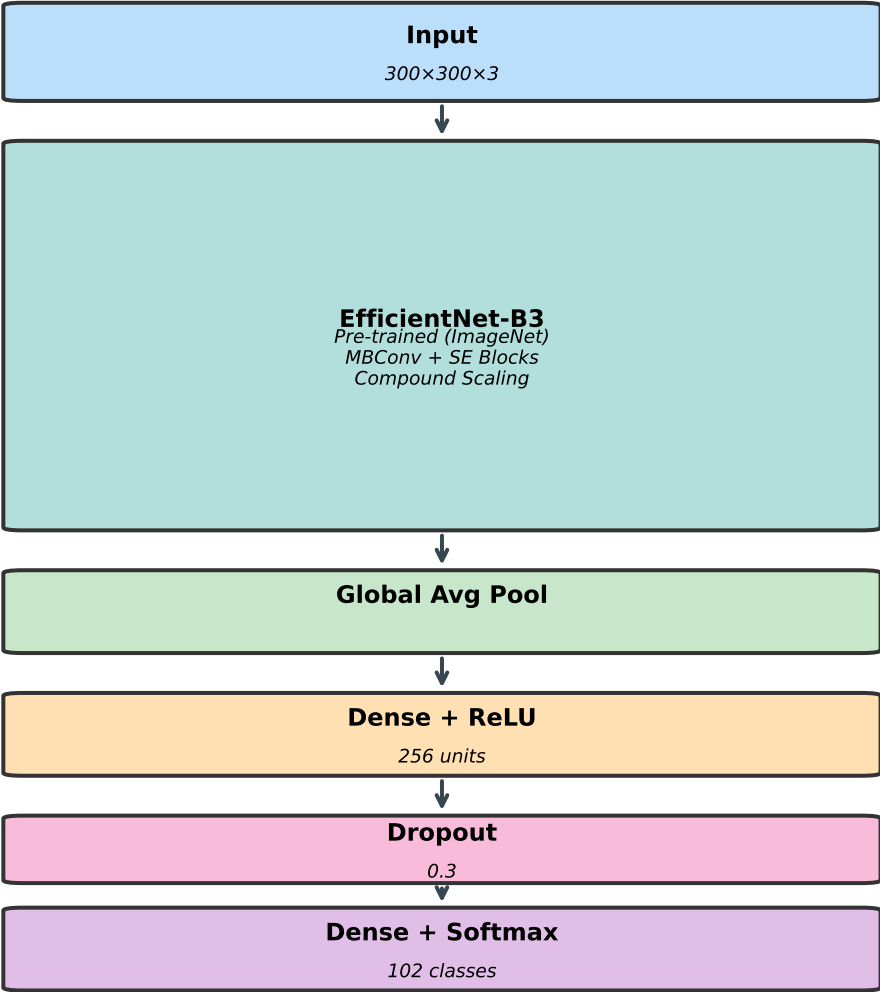*256 filters, 3×3*

MaxPool 2×2

Conv2D + BN + ReLU
512 filters, 3×3

Global Avg Pool

Dense + ReLU
512 units

Dropout
0.5

Dense + Softmax
102 classes
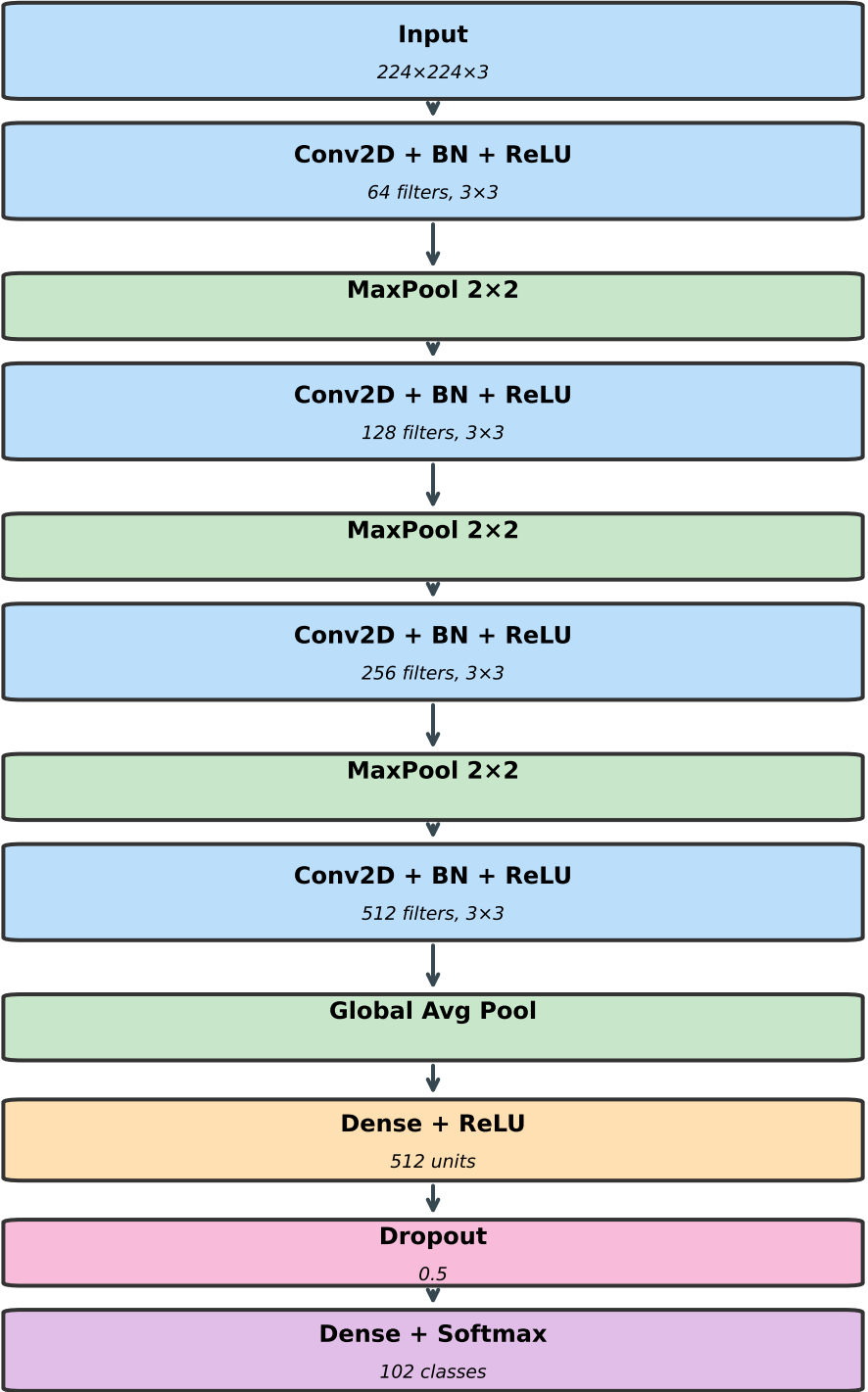
~2.5M Parameters

ResNet50 (Transfer Learning)

Input
224×224×3

Skip Conn.

ResNet50 Backbone
Pre-trained (ImageNet)
Layers 1-45: FROZEN

AdaptiveAvgPool
(1,1)

Dense + ReLU
512 units

Dropout
0.3

Dense + Softmax
102 classes
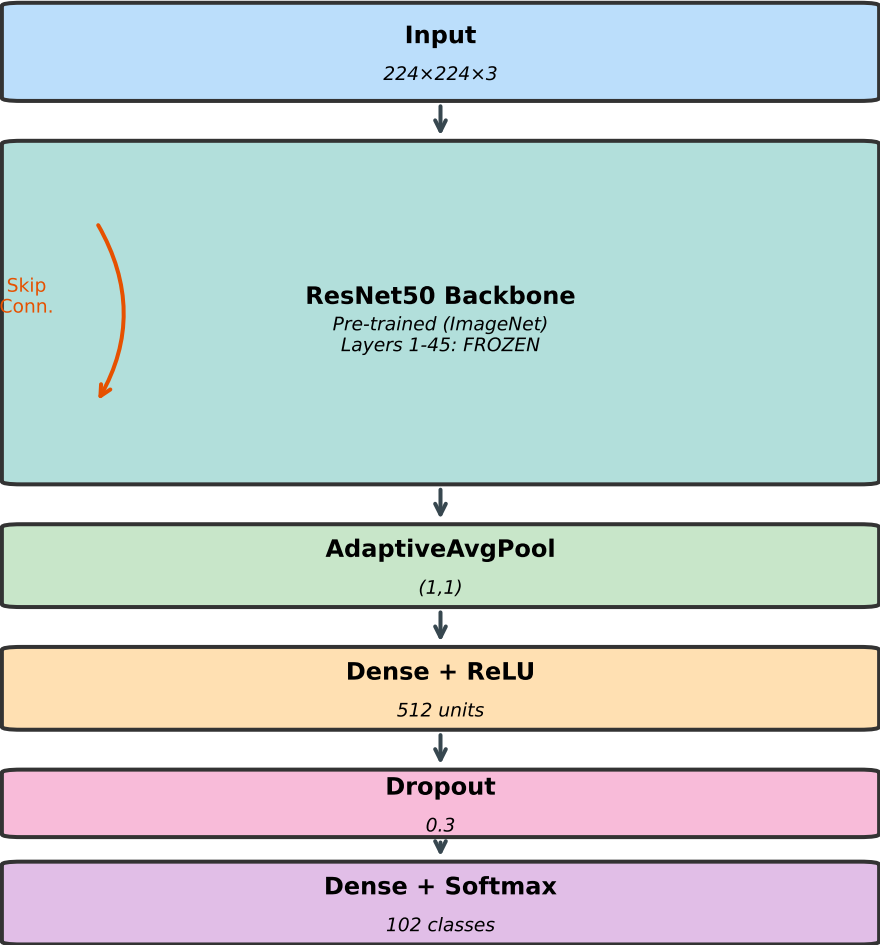
~23.5M Params (2M trainable)

EfficientNet-B3 (Transfer Learning)

Input
300×300×3

EfficientNet-B3
Pre-trained (ImageNet)
MBConv + SE Blocks
Compound Scaling

Global Avg Pool

Dense + ReLU
256 units

Dropout
0.3

Dense + Softmax
102 classes

~10.7M Parameters

Legend: Convolution | Pooling | Pre-trained Backbone | Fully Connected | Dropout | Output