RestNet50 Layers

|  |  |  |
| --- | --- | --- |
| Layer Name | Layer Type | Properties |
| input\_layer | Input | 224x224x3 RGB images |
| initial\_conv | Conv2D | 64 filters, 7x7 kernel, stride [2 2], same padding, ReLU |
| initial\_maxpool | MaxPooling2D | 3x3 pool size, stride [2 2] |
| residual\_stage1\_block1 | Conv2D ×3 | 3× Conv2D (1x1, 3x3, 1x1), total 9 Conv2D layers in stage 1 |
| residual\_stage2\_block2 | Conv2D ×3 | 4× Residual blocks × 3 Conv2D = 12 Conv2D layers |
| residual\_stage3\_block3 | Conv2D ×3 | 6× Residual blocks × 3 Conv2D = 18 Conv2D layers |
| residual\_stage4\_block4 | Conv2D ×3 | 3× Residual blocks × 3 Conv2D = 9 Conv2D layers |
| global\_avg\_pool | GlobalAveragePooling2D | Reduces spatial dimensions to 1D vector |
| dense\_1 | Dense | 256 units, ReLU activation |
| dropout | Dropout | 50% dropout rate |
| dense\_2 | Dense | 64 units, ReLU activation |
| output\_layer | Dense (Softmax) | 5 units (classes), Softmax activation |

MobileNetV2

|  |  |  |
| --- | --- | --- |
| Layer Name | Layer Type | Properties |
| input\_layer | Input | 224x224x3 RGB images |
| initial\_conv | Conv2D (1x1) | Expansion factor 1, followed by BatchNorm & ReLU6 |
| bottleneck\_block\_1 to 17 | Bottleneck Blocks ×17 | Each block has 3 layers: |
|  |  | - 1x1 Conv2D (expansion) |
|  |  | - 3x3 DepthwiseConv2D |
|  |  | - 1x1 Conv2D (projection) |
|  |  | Total from 17 blocks = 17 × 3 = 51 Conv2D layers |
| global\_avg\_pool | GlobalAveragePooling2D | Reduces spatial dimensions to a 1D vector |
| dense\_1 | Dense | 128 units, ReLU activation |
| dropout | Dropout | 40% dropout rate |
| output\_layer | Dense (Softmax) | 5 units (classes), Softmax activation |

EfficientNetB0

|  |  |  |
| --- | --- | --- |
| Layer Name | Layer Type | Properties |
| input\_layer | Input | 224x224x3 RGB images |
| initial\_conv | Conv2D (3x3, 32 filters) | Stride 2, followed by BatchNorm & Swish activation |
| mbconv\_stage\_1 | MBConv1 Block ×1 | 3x3 depthwise, 16 filters, includes 3 Conv2D layers |
| mbconv\_stage\_2 | MBConv6 Block ×2 | 3x3 depthwise, 24 filters, total 6 Conv2D layers |
| mbconv\_stage\_3 | MBConv6 Block ×2 | 5x5 depthwise, 40 filters, total 6 Conv2D layers |
| mbconv\_stage\_4 | MBConv6 Block ×3 | 3x3 depthwise, 80 filters, total 9 Conv2D layers |
| mbconv\_stage\_5 | MBConv6 Block ×3 | 5x5 depthwise, 112 filters, total 9 Conv2D layers |
| mbconv\_stage\_6 | MBConv6 Block ×4 | 5x5 depthwise, 192 filters, total 12 Conv2D layers |
| mbconv\_stage\_7 | MBConv6 Block ×1 | 3x3 depthwise, 320 filters, 3 Conv2D layers |
| final\_conv | Conv2D (1x1, 1280 filters) | Final convolution layer before custom head |
| global\_avg\_pool | GlobalAveragePooling2D | Reduces spatial dimensions to 1D vector |
| dense\_1 | Dense | 256 units, ReLU activation |
| dropout | Dropout | 50% dropout rate |
| output\_layer | Dense (Softmax) | 5 units (classes), Softmax activation |

Xception

|  |  |  |
| --- | --- | --- |
| Layer Name | Layer Type | Properties |
| input\_layer | Input | 224x224x3 RGB images |
| entry\_conv1 | Conv2D (3x3, 32 filters) | Stride 2, followed by BatchNorm & ReLU |
| entry\_conv2 | Conv2D (3x3, 64 filters) | Stride 1, followed by BatchNorm & ReLU |
| entry\_block\_1\_to\_3 | SeparableConv2D ×2 + Conv2D ×1 | 3 blocks × (2 Separable + 1 Conv2D) = 9 layers |
| middle\_flow\_blocks | SeparableConv2D ×24 | 8 blocks × 3 SeparableConv2D each |
| exit\_block | SeparableConv2D ×2 + Conv2D ×1 | Final exit block with 2 SeparableConv2D + 1 Conv2D = 3 layers |
| global\_avg\_pool | GlobalAveragePooling2D | Reduces spatial dimensions to 1D vector |
| dense\_1 | Dense | 256 units, ReLU activation |
| output\_layer | Dense (Softmax) | 5 units (classes), Softmax activation |

InceptionV3

|  |  |  |
| --- | --- | --- |
| Layer Name | Layer Type | Properties / Description |
| input\_layer | Input | Accepts 224x224x3 RGB images |
| conv1 | Conv2D (3x3, 32 filters) | Stride 2 |
| conv2 | Conv2D (3x3, 32 filters) | Stride 1 |
| conv3 | Conv2D (3x3, 64 filters) | Stride 1 |
| conv4 | Conv2D (3x3, 80 filters) | Stride 1 |
| conv5 | Conv2D (3x3, 192 filters) | Stride 2 |
| maxpool1 | MaxPooling2D (3x3, stride 2) | After initial Conv2D layers |
| inception\_A\_blocks | Inception Modules (Type A) | 3 blocks × 9 Conv2D each = 27 layers |
| inception\_B\_blocks | Inception Modules (Type B) | 5 blocks × 7 Conv2D each = 35 layers |
| inception\_C\_blocks | Inception Modules (Type C) | 3 blocks × 8 Conv2D each = 24 layers |
| aux\_maxpool\_1 | MaxPooling2D | Grid reduction between blocks |
| aux\_maxpool\_2 | MaxPooling2D | Grid reduction between blocks |
| global\_avg\_pool | GlobalAveragePooling2D | Reduces spatial dimensions to 1D vector |
| dense\_1 | Dense | 256 units, ReLU activation |
| dropout | Dropout | 0.5 rate for regularization |
| output\_layer | Dense (Softmax) | 5 units (for class prediction), softmax activation |

Swin Transformer

|  |  |  |
| --- | --- | --- |
| Layer Name | Layer Type | Properties / Description |
| input\_layer | Input | Accepts 224x224x3 RGB images |
| patch\_embedding | Conv2D (Patch Embedding) | 4x4 patches, stride 4, output: 96 channels |
| stage1\_block1 | Swin Transformer Block | 96 channels, resolution: 56×56 |
| stage1\_block2 | Swin Transformer Block | 96 channels |
| patch\_merging1 | Patch Merging | Reduces to 28×28, increases to 192 channels |
| stage2\_block1 | Swin Transformer Block | 192 channels |
| stage2\_block2 | Swin Transformer Block | 192 channels |
| patch\_merging2 | Patch Merging | Reduces to 14×14, increases to 384 channels |
| stage3\_block1-6 | Swin Transformer Blocks (6 blocks) | 384 channels |
| patch\_merging3 | Patch Merging | Reduces to 7×7, increases to 768 channels |
| stage4\_block1 | Swin Transformer Block | 768 channels |
| stage4\_block2 | Swin Transformer Block | 768 channels |
| global\_avg\_pool | GlobalAveragePooling2D | Reduces spatial dimensions to 1D vector |
| dense\_1 | Dense | 256 units, ReLU activation |
| dropout | Dropout | 0.5 dropout rate for regularization |
| output\_layer | Dense | 5 units (class prediction), softmax activation |

CNN model

|  |  |  |
| --- | --- | --- |
| **Layer Name** | **Layer Type** | **Properties** |
| imageinput | Image Input | 224x224x3 images with zero-center normalization |
| conv\_1 | Convolution | 64 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| relu\_1 | ReLU | ReLU activation |
| conv\_2 | Convolution | 64 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| maxpool\_1 | Max Pooling | 2x2 pool size, stride [2 2], padding [0 0 0 0] |
| conv\_3 | Convolution | 128 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| relu\_2 | ReLU | ReLU activation |
| conv\_4 | Convolution | 128 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| maxpool\_2 | Max Pooling | 2x2 pool size, stride [2 2], padding [0 0 0 0] |
| conv\_5 | Convolution | 256 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| relu\_3 | ReLU | ReLU activation |
| conv\_6 | Convolution | 256 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| conv\_7 | Convolution | 256 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| maxpool\_3 | Max Pooling | 2x2 pool size, stride [2 2], padding [0 0 0 0] |
| conv\_8 | Convolution | 512 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| relu\_4 | ReLU | ReLU activation |
| conv\_9 | Convolution | 512 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| conv\_10 | Convolution | 512 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| maxpool\_4 | Max Pooling | 2x2 pool size, stride [2 2], padding [0 0 0 0] |
| conv\_11 | Convolution | 512 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| relu\_5 | ReLU | ReLU activation |
| conv\_12 | Convolution | 512 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| conv\_13 | Convolution | 512 filters, 6x6 kernel, stride [1 1], padding [2 2 2 2] |
| maxpool\_5 | Max Pooling | 2x2 pool size, stride [2 2], padding [0 0 0 0] |
| flatten | Flatten | Converts feature maps to 1D vector |
| fc\_1 | Fully Connected | 256 units |
| dropout | Dropout | 30% dropout rate |
| fc\_2 | Fully Connected | 64 units |
| softmax | Softmax | Softmax activation for multi-class classification |
| classoutput | Classification Output | Cross-entropy loss |

VGG16

|  |  |  |
| --- | --- | --- |
| Layer Name | Layer Type | Properties |
| input\_layer | Input | 224x224x3 RGB images |
| conv\_block1\_conv1 | Conv2D | 64 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| conv\_block1\_conv2 | Conv2D | 64 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| maxpool\_block1 | MaxPooling2D | 2x2 pool size, stride [2 2] |
| conv\_block2\_conv1 | Conv2D | 128 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| conv\_block2\_conv2 | Conv2D | 128 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| maxpool\_block2 | MaxPooling2D | 2x2 pool size, stride [2 2] |
| conv\_block3\_conv1 | Conv2D | 256 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| conv\_block3\_conv2 | Conv2D | 256 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| conv\_block3\_conv3 | Conv2D | 256 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| maxpool\_block3 | MaxPooling2D | 2x2 pool size, stride [2 2] |
| conv\_block4\_conv1 | Conv2D | 512 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| conv\_block4\_conv2 | Conv2D | 512 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| conv\_block4\_conv3 | Conv2D | 512 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| maxpool\_block4 | MaxPooling2D | 2x2 pool size, stride [2 2] |
| conv\_block5\_conv1 | Conv2D | 512 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| conv\_block5\_conv2 | Conv2D | 512 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| conv\_block5\_conv3 | Conv2D | 512 filters, 3x3 kernel, stride [1 1], same padding, ReLU |
| maxpool\_block5 | MaxPooling2D | 2x2 pool size, stride [2 2] |
| global\_avg\_pool | GlobalAveragePooling2D | Reduces spatial dimensions to 1D vector |
| dense\_1 | Dense | 256 units, ReLU activation |
| dropout | Dropout | 50% dropout rate |
| dense\_2 | Dense | 64 units, ReLU activation |
| output\_layer | Dense (Softmax) | 5 units (classes), Softmax activation |