Name: Edgar Matoto

Week-1

# Transfer Learning Image Classification Algorithms​

## ResNet

ResNet (Residual Neural Network) is a deep learning architecture that allows training very deep networks effectively. It overcomes the problem of vanishing gradients by using shortcut connections called residual connections.

|  |  |
| --- | --- |
| **Pros** | * Enables training of very deep networks. * Residual connections facilitate learning and optimization. * Achieves state-of-the-art results in computer vision tasks. * Relatively easy to implement and train. |
| **Cons** | * Increased complexity and computational resources. * Potential overfitting with extremely deep architectures. * Limited impact on certain tasks or datasets. |
| **High-level Structure** | 1. Input Image 2. Initial Convolutional Layer and Max Pooling 3. Residual Blocks (multiple layers) 4. Global Average Pooling 5. Fully Connected Layers (classifier) 6. Output (class probabilities) |

## DensNet

DenseNet is a deep learning model known for its dense connectivity, where each layer is directly connected to every other layer in a feed-forward manner.

|  |  |
| --- | --- |
| **Pros** | * Strong feature reuse and propagation due to dense connectivity, leading to better gradient flow and information flow. * Reduces the number of parameters, making it memory-efficient and computationally efficient. * Combats vanishing gradient problem and encourages feature diversity. * Effective for transfer learning and achieving state-of-the-art results in image classification tasks. |
| **Cons** | * Higher memory requirements compared to traditional convolutional networks due to dense connectivity. * Increased computational complexity during training and inference. * Sensitivity to hyperparameter tuning. |
| **High-level Structure** | 1. Input image 2. Convolutional Layers for feature extraction 3. Dense Blocks with dense connectivity 4. Transition Layers for dimensionality reduction 5. Global Average Pooling to reduce dimensions 6. Fully Connected Layers for classification |

## VGG

VGG (Visual Geometry Group) is a widely used deep learning model known for its simple and uniform architecture.

|  |  |
| --- | --- |
| **Pros** | * Simple and easy-to-understand architecture. * Effective at learning local image features due to small filter sizes. * Achieves strong performance in image classification tasks. * Easy to implement and transfer to new datasets. |
| **Cons** | * Relatively deeper and requires more parameters compared to other architectures. * Higher computational and memory requirements. * Limited feature reuse due to repeated stacking of convolutional layers. * Prone to overfitting on smaller datasets. |
| **High-level Structure** | 1. Input image 2. Convolutional Layers with small filter sizes 3. Pooling Layers for downsampling 4. Fully Connected Layers for classification |

## Inception

The Inception model, also known as GoogLeNet, is a deep learning architecture famous for its inception modules that use multiple filter sizes within a layer.

|  |  |
| --- | --- |
| **Pros** | * Efficient use of computation by utilizing multiple filter sizes in parallel within each layer. * Captures both local and global features effectively. * Reduces the number of parameters compared to traditional architectures. * Achieves good performance while maintaining model compactness. |
| **Cons** | * Complex architecture that may be harder to interpret. * Increased computational requirements due to parallel operations. * Higher memory usage compared to simpler architectures. * Sensitive to hyperparameter tuning. |
| **High-level Structure** | 1. Input image 2. Inception Modules with parallel convolutional operations of different filter sizes 3. Pooling Layers for downsampling 4. Fully Connected Layers for classification |