Model Development Phase

Date	10th July 2024
Team ID	SWTID1720195938
Project Title	**CovidVision: Advanced COVID-19 Detection from Lung X-Rays with Deep Learning**
Maximum Marks	5 Marks

Model Selection Report

In the model selection report for future deep learning and computer vision projects, various architectures, such as CNNs or RNNs, will be evaluated. Factors such as performance, complexity, and computational requirements will be considered to determine the most suitable model for the task at hand.

Model Selection Report:

Model	Description
VGG16	The VGG16 model is a convolutional neural network (CNN) architecture that was developed by the Visual Geometry Group (VGG) at the University of Oxford. VGG16 is highly effective in image classification tasks and has achieved top results in competitions such as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). VGG16 consists of 16 weight layers, including 13 convolutional layers and 3 fully connected layers. The model uses small 3x3 convolution filters throughout the network, which simplifies the design and makes it easy to implement. Recommended GPUs for training and inference include NVIDIA Tesla V100, A100, RTX 3090, or similar high-performance GPUs. Systems with at least 32-64 GB of RAM are recommended for training to handle the large memory requirements. High-speed storage (SSD/NVMe) is beneficial for faster data loading and saving checkpoints.

CNN Model	A typical CNN can achieve high accuracy on standard image classification tasks. Performance can be tailored by adjusting the depth and number of filters in the network. A typical CNN might consist of several convolutional layers followed by pooling layers, and one or more fully connected layers. For example, a simple CNN might have 5-10 convolutional layers. The number of parameters in a CNN can range from a few hundred thousand to several million, depending on the architecture. This is significantly fewer than VGG16, which has 138 million parameters. Recommended GPUs for training and inference include NVIDIA Tesla T4, V100, RTX 2080, or similar mid-range to high-end GPUs. Systems with 16-32 GB of RAM are recommended for training mid-sized CNNs. High-speed storage (SSD/NVMe) is beneficial for faster data loading and saving checkpoints.
RESNET 50	ResNet-50 is known for its high accuracy in image classification tasks. It was a significant improvement over previous models when introduced, especially in terms of handling the vanishing gradient problem due to its residual connections. ResNet-50 consists of 50 layers, including convolutional layers, batch normalization layers, ReLU activations, and residual blocks. ResNet-50 has approximately 25.6 million parameters, which is significantly fewer than VGG16 but still substantial. Recommended GPUs for training and inference include NVIDIA Tesla V100, A100, RTX 3090, or similar high-performance GPUs. Systems with at least 32-64 GB of RAM are recommended for training to handle the large memory requirements. High-speed storage (SSD/NVMe) is beneficial for faster data loading and saving checkpoints.
Xception	Xception is known for its high accuracy on image classification tasks. It outperforms many traditional models by using depthwise separable convolutions, which are more efficient and allow for deeper architectures. Xception consists of 71 layers, making it deeper than many traditional CNNs but still manageable due to the efficiency of its architecture. Recommended GPUs for training and inference include NVIDIA Tesla V100, A100, RTX 3090, or similar high-performance GPUs. Systems with at least 32-64 GB of RAM are recommended for training to handle the large memory requirements. High-speed storage (SSD/NVMe) is beneficial for faster data loading and saving checkpoints.