**Model Optimization and Tuning Phase Template**

|  |  |
| --- | --- |
| Date | 29 June 2025 |
| Team ID | SWTID1750316859 |
| Project Title | ASL - Alphabet Image Recognition |
| Maximum Marks | 10 Marks |

**Model Optimization and Tuning Phase**

The Model Optimization and Tuning Phase involves refining neural network models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

### Hyperparameter Tuning Documentation (8 Marks):

|  |  |
| --- | --- |
| **Model** | **Tuned Hyperparameters** |
| MobileNetV2 | class CFG:  batch\_size = 64 # Number of images processed before the model updates weights  img\_height = 128 # Height of each input image  img\_width = 128 # Width of each input image  epochs = 10 # Total number of training cycles over the entire dataset  num\_classes = 29 # Total number of output classes (e.g., 26 alphabets + 3 extra signs)  img\_channels = 3 # Number of color channels (3 for RGB images) |
| VGG16 | # Configuration for training  class CFG:      batch\_size = 64            # How many images to process in one step      img\_height = 64            # Height of input image      img\_width = 64             # Width of input image      epochs = 10                # Number of training rounds      num\_classes = 29           # Total ASL labels: A-Z + del + nothing + space      img\_channels = 3           # RGB image has 3 channels |

### 

### 

### 

### Final Model Selection Justification (2 Marks):

|  |  |
| --- | --- |
| **Final Model** | **Reasoning** |
| MobileNet-V2 | VGG16 is a deep convolutional network with ~138 million parameters, known for its simplicity but extremely large size and slow training. In this project, it took nearly **4 hours per epoch** and started with a very low accuracy of **0.03**, making it impractical despite being pretrained.  MobileNetV2, in contrast, is a lightweight model (~3.4 million parameters) optimized for speed and efficiency. It trained much faster (within minutes per epoch) and started with a significantly better accuracy (~18%). Its use of depthwise separable convolutions and inverted residuals makes it ideal for tasks like ASL recognition, where quick iteration and lower resource usage are essential.  Thus, VGG16 was discontinued, and MobileNetV2 was chosen for its superior performance, speed, and practicality. |