Model Training Report

Your Name

May 12, 2025

1 5.2 Summary Report

1.1 Final Accuracy and Class-wise Metrics

The model was trained across multiple configurations, and the following metrics were obtained:

- **Best Validation Accuracy:** 0.8616 from the second ResNet-18 model (with changeable parameters).
- **Final Test Accuracy:** 52% (on the test dataset after training the best model).

For the finetuning experiments, the following metrics were achieved:

- **Experiment 1:**
  - **Learning Rate:** $1e^{-4}$, **Batch Size:** 64, **Hidden Units:** 256, **Dropout:** 0.2
  - **Train Accuracy:** 88.00%, **Validation Accuracy:** 81.00%
  - **Precision:** 0.7896, **Recall:** 0.7761, **F1-score:** 0.7630
- **Experiment 2:**
  - **Learning Rate:** $1e^{-4}$, **Batch Size:** 64, **Hidden Units:** 1024, **Dropout:** 0.5
  - **Train Accuracy:** 86.66%, **Validation Accuracy:** 83.25%

- **Precision:** 0.8355, **Recall:** 0.8169, **F1-score:** 0.8112

- **Experiment 3:**
  - **Learning Rate:** $1e^{-3}$, **Batch Size:** 128, **Hidden Units:** 256, **Dropout:** 0.5
  - **Train Accuracy:** 57.09%, **Validation Accuracy:** 64.40%
  - **Precision:** 0.5734, **Recall:** 0.5598, **F1-score:** 0.5358

- **Final Model (with Early Stopping, Stronger Data Augmentation, Scheduler, L2 Regularization):**
  - **Epochs:** 15, **Train Accuracy:** 81.72%, **Validation Accuracy:** 83.01%
  - **Train Loss:** 0.0817, **Validation Loss:** 0.8130

- **EfficientNet-B0 Model:**
  - **Final Accuracy:** This model was worse than the simple overfitting models since it took a longer time to run and achieved similar validation accuracies, but on the flipside, the validation and train accuracies were close

1.2 Challenges Faced During Training and How They Were Overcome

Several challenges were encountered during the training process:

- **Overfitting:** The model initially showed signs of overfitting, especially when training on the smaller datasets. To counter this, early stopping, stronger data augmentation, and regularization techniques such as L2 regularization were employed.

- **Suboptimal Learning Rates:** During some experiments, the learning rate was either too high or too low, causing the model to either converge too slowly or not converge at all. The learning rate scheduler and adjustments to batch sizes helped address this issue.

- **Training Time:** The model took a considerable amount of time to train with multiple experiments due to the complexity of the models and the dataset size. This was mitigated by using a more powerful GPU and optimizing the batch size.

1.3 Design Choices and Future Improvements

The following design choices were made during the training process:

- **Model Architecture:** The ResNet-18 architecture was chosen due to its performance in image-based tasks and its ability to be fine-tuned effectively.
- **Hyperparameter Tuning:** Hyperparameters such as learning rate, batch size, and dropout rate were tuned using a grid search approach to find optimal values for each configuration.
- **Data Augmentation:** Stronger data augmentation techniques were applied to improve generalization and reduce overfitting.
- **Regularization:** L2 regularization was used to reduce overfitting by penalizing large weights.

Future improvements include:

- **Model Complexity:** Trying more advanced models such as ResNet-34 or ResNet-50 to further improve accuracy.
- **Data Augmentation Strategies:** Investigating additional data augmentation methods, such as cutout or MixUp, to improve model robustness.
- **Experimenting with Other Optimizers:** Exploring optimizers like AdamW or RMSProp to potentially improve convergence.