ResNet-34

The model follows the standard ResNet34 architecture using residual blocks to prevent degradation in deep networks.

It includes:

- A stem convolution + max pooling

  It applies a common Conv2D → BatchNorm → ReLU pattern and used repeatedly throughout the model to apply convolutions with proper normalization and activation.

- Four residual stages

  Implements a residual block with optional downsampling. It contains two convolutional layers and a skip connection. Residual blocks help deep networks learn identity mappings, solving vanishing gradient issues and allowing better convergence in very deep architectures.

- A main classifier head
- An auxiliary classifier head from an intermediate layer

  The auxiliary output aids training by providing an additional gradient signal from an intermediate layer, improving convergence and reducing overfitting.

- A fully connected stack with dropout and batch normalization

  Boosts generalization as Batch normalization helps with training stability, while dropout reduces overfitting (low dropout rate, only 5% of the features are dropped during training, which is gentle regularization).

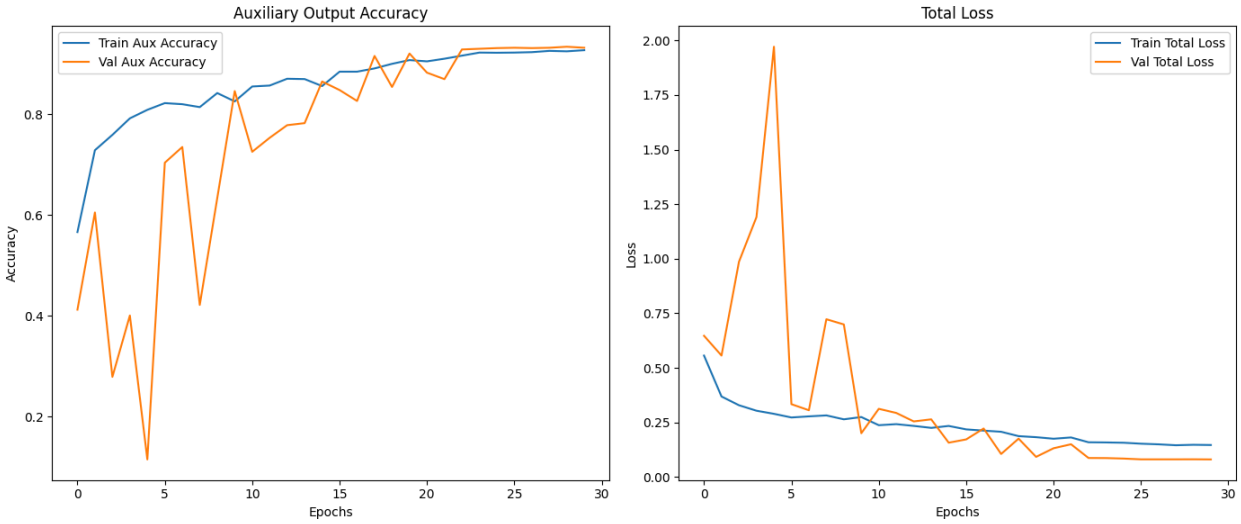
- Optimizer and Learning Rate Schedule

  Uses a cosine learning rate decay schedule with Adam optimizer, helps achieve smoother convergence by gradually reducing the learning rate. Adam provides adaptive gradients for better optimization performance.

- Losses and Compilation

  Compiles the model with different weights for main and auxiliary losses. Using an auxiliary loss with lower weight improves training by encouraging better feature learning early in the network.

- Training and Validation Performance of ResNet34 Model



Auxiliary Output Accuracy:

- shows how the model's accuracy for the auxiliary output evolved over 30 training epochs.
- Both curves increase and stabilize near 95%.
- The fluctuations of the validation accuracy could be due to the auxiliary output being less stable or the validation dataset being small or imbalanced.

Total Loss:

- Represents the total loss (sum of main and auxiliary losses) for both training and validation sets.
- The training loss steadily decreases as the model learns.
- The validation loss decreases overall but spiking sharply in early epochs. These spikes suggest overfitting or noise in the validation set.
- The validation loss stabilizes then, suggesting convergence.