Report: Train your Network

Initial Network

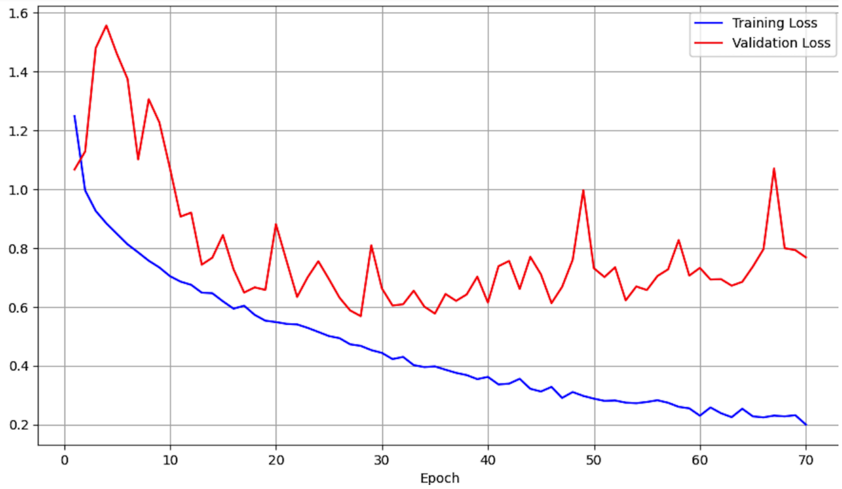
Model: "sequential"

```
┏━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━┳━━━━━━━━━━━━━━━━━━━━━━━━┳━━━━━━━━━━━━━━━┓
┃ Layer (type)                         ┃ Output Shape           ┃       Param # ┃
┡━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━╇━━━━━━━━━━━━━━━━━━━━━━━━╇━━━━━━━━━━━━━━━┩
│ rescaling (Rescaling)                │ (None, 150, 150, 3)    │             0 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ conv2d (Conv2D)                      │ (None, 150, 150, 24)   │           672 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ batch_normalization                  │ (None, 150, 150, 24)   │            96 │
│ (BatchNormalization)                 │                        │               │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ max_pooling2d (MaxPooling2D)         │ (None, 37, 37, 24)     │             0 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ dropout (Dropout)                    │ (None, 37, 37, 24)     │             0 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ conv2d_1 (Conv2D)                    │ (None, 37, 37, 32)     │         6,944 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ batch_normalization_1                │ (None, 37, 37, 32)     │           128 │
│ (BatchNormalization)                 │                        │               │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ max_pooling2d_1 (MaxPooling2D)       │ (None, 12, 12, 32)     │             0 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ dropout_1 (Dropout)                  │ (None, 12, 12, 32)     │             0 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ conv2d_2 (Conv2D)                    │ (None, 12, 12, 48)     │        13,872 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ batch_normalization_2                │ (None, 12, 12, 48)     │           192 │
│ (BatchNormalization)                 │                        │               │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ max_pooling2d_2 (MaxPooling2D)       │ (None, 6, 6, 48)       │             0 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ dropout_2 (Dropout)                  │ (None, 6, 6, 48)       │             0 │
├──────────────────────────────────────┼────────────────────────┼───────────────┤
│ flatten (Flatten)                    │ (None, 1728)           │             0 │
```

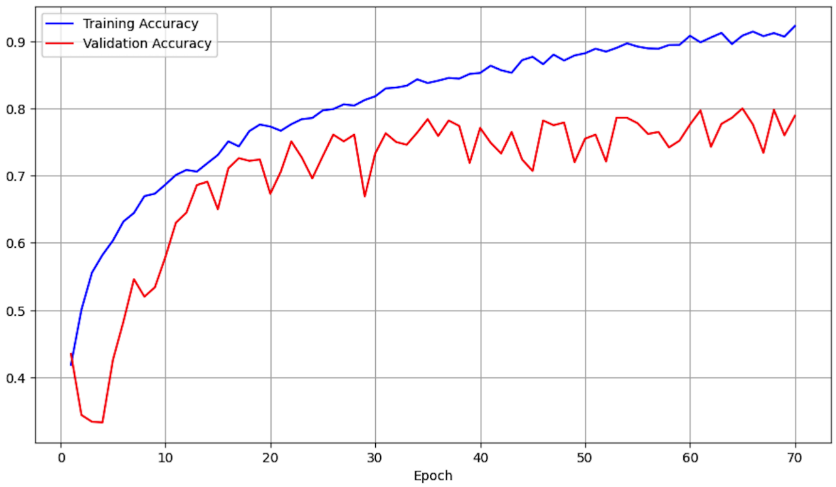
```
├────────────────────────────────────────────────────────────────────────
│ dense (Dense)                │ (None, 64)             │       110,656 │
├────────────────────────────────────────────────────────────────────────
│ dropout_3 (Dropout)          │ (None, 64)             │             0 │
├────────────────────────────────────────────────────────────────────────
│ dense_1 (Dense)              │ (None, 3)              │           195 │
└────────────────────────────────────────────────────────────────────────
```

Total params: 132,755 (518.57 KB)
Trainable params: 132,547 (517.76 KB)
Non-trainable params: 208 (832.00 B)

Plot Showing Training and Validation Loss as Function of Epoch



Plot Showing Accuracy Against the Training and Validation Sets as a Function of Epoch



The Accuracy and Loss of Best Learned Model (Obtained as the Model in Effect when Overfitting Begins) when Measured Against the Held-Back Test Set

Raw Training Output at Epoch 35 (Best Performance):

```
Epoch 35/70
32/32 ━━━━━━━━━━━━━━━━━━━━ 2s 61ms/step - accuracy: 0.8385 (83.85%) - loss: 0.4016 - val_accuracy: 0.7840 (78.40%) - val_loss: 0.5768
```

Raw Test Set Evaluation:

```
* Evaluating basic_model
30/30 ━━━━━━━━━━━━━━━━━━━━ 2s 72ms/step - accuracy: 0.7362 (73.62%) - loss: 0.8933
```

Report on Hyperparameter Optimization Strategy Including What Hyperparameter You Experimented With, How You Changed Them, and How This Affected Your Accuracy

Hyperparameter Optimization Strategy Report

The specific hyperparameter changes and impact has been included as following:

- Convolutional Layer Architecture:
    - Original: Three blocks with fixed filter sizes (32, 64, 64)
    - Modified: Progressive filter sizes (24, 32, 48)
    - Effect: The increase in filter sizes helped create a balanced feature extraction hierarchy, and allowed us to stay within the limit of 150,000 parameters (had 132,755 total parameters)
- Pooling Layer Strategy:
    - Original: Uniform pooling sizes (2, 2) across all layers
    - Modified: Progressive pooling sizes (4,4), (3,3), (2,2)
    - Effect: The larger initial pooling sizes helped improve computational efficiency.
- Dropout Implementation:
    - Original: Two dropout layers (0.25, 0.5)
    - Modified: Progressive dropout pattern (0.1, 0.2, 0.3, 0.4)
    - Effect: The gradual increase in dropout rates provided more regularization throughout the network, helping prevent overfitting.
- Architecture Changes:
    - Original: GlobalAveragePooling2D before dense layers
    - Modified: Switched to Flatten layer
    - Effect: Using Flatten preserved more spatial information, though at the cost of increased parameters (it went from ~63k parameters to ~130k parameters)
- Optimizer Selection:
    - Original: RMSprop optimizer
    - Modified: Adam optimizer (keeping same learning rate of 0.001)
    - Effect: The switch to Adam provided better convergence properties and more stable training.
- Impact on Performance:
    - Original model achieved ~69% test accuracy
    - Modified model achieved 73.62% test accuracy
    - Best validation accuracy improved from ~72% to 78.40%
    - Training became more stable with less fluctuation in validation metrics