Political Meme Classification Phase 4

Kate Arendes
Introduction to Deep Learning

March 20, 2022

Contents

Overview

In this phase of the project, the performance of the final model from the previous phase was improved through the use of data augmentation.

1 Effects of Data Augmentation

To augment the data, input images were altered via three forms of random transformation: rotation, zooming, and flipping (Figure 1). While these alterations theoretically should have helped the model, especially since there was a relatively small set of training images available, the improvements resulting from the experimentation in this phase were modest.

1.1 Rotation

After testing many different configurations, the most effective factor by which input data was randomly rotated was determined to be 0.2. Decreasing or increasing that factor by any amount slightly hindered the model's performance, and the only instance when the model outperformed the metrics of the previous phase was when the factor was set to this amount.

The configuration with a factor of 0.2 achieved a test loss of 0.4395, a test accuracy of 0.83, and a test precision of 0.8056 (Figure 4). Although the test accuracy was equal to that of the

Figure 1: A randomly selected liberal meme's appearance after rotation, zooming, and horizontal flipping.

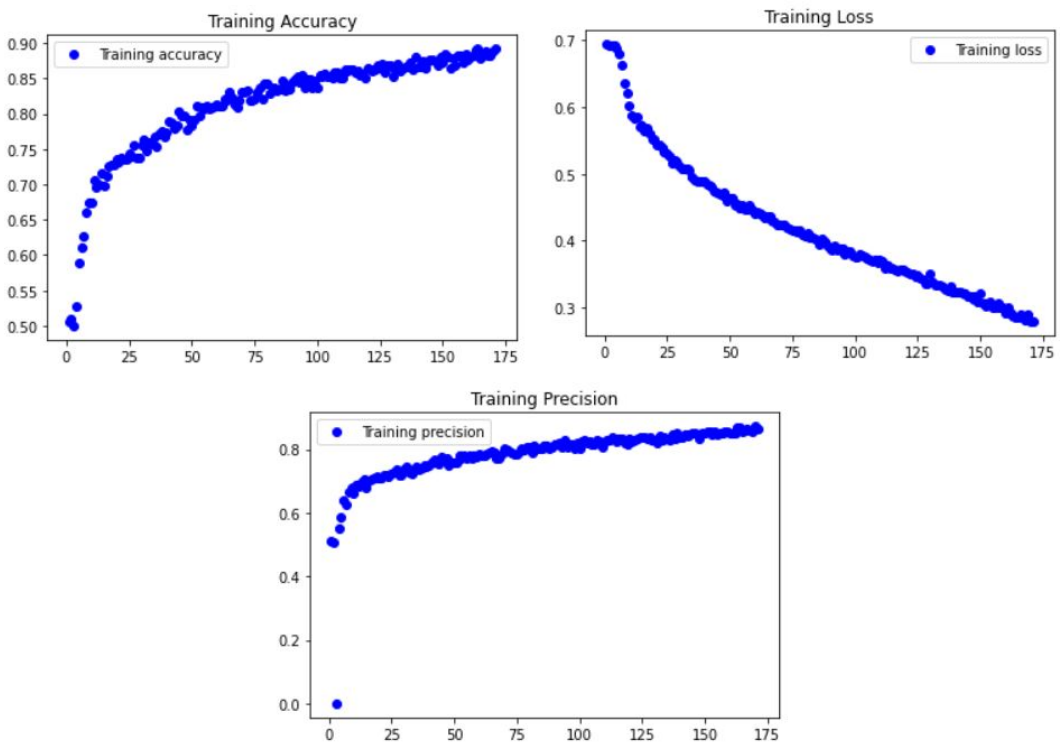

Figure 2: Training accuracy, loss, and precision for the best-performing random rotation configuration.

previous phase's model, the test precision was three points higher, and was the highest test precision achieved by any of the models employing data augmentation. The test loss of 0.4395 was also three points lower than that achieved by the previous phase's model.

1.2 Zooming

On its own, the introduction of random zooming to the input data greatly hindered the model's performance. While all configurations were ineffective, the least harmful involved randomly zooming with a factor of 0.19, which only managed to reach a validation accuracy of 0.625 as the loss erratically increased after epoch 35 (Figure 6). In testing, the model employing random zooming achieved an accuracy of 0.62, about twenty points lower than the best model from the previous phase (Figure 7).

1.3 Flipping

Similar to the application of random zooming, the inclusion of flipping on its own, horizontally and/or vertically, was harmful to the model, and provided poor results even with additional hyperparameter tuning. However, when combined with a model already utilizing the best factors for rotation (0.2) and zooming (0.19), applying flipping still rendered a relatively well-performing model.

Although vertical flipping proved to be more effective than horizontal flipping, using both forms

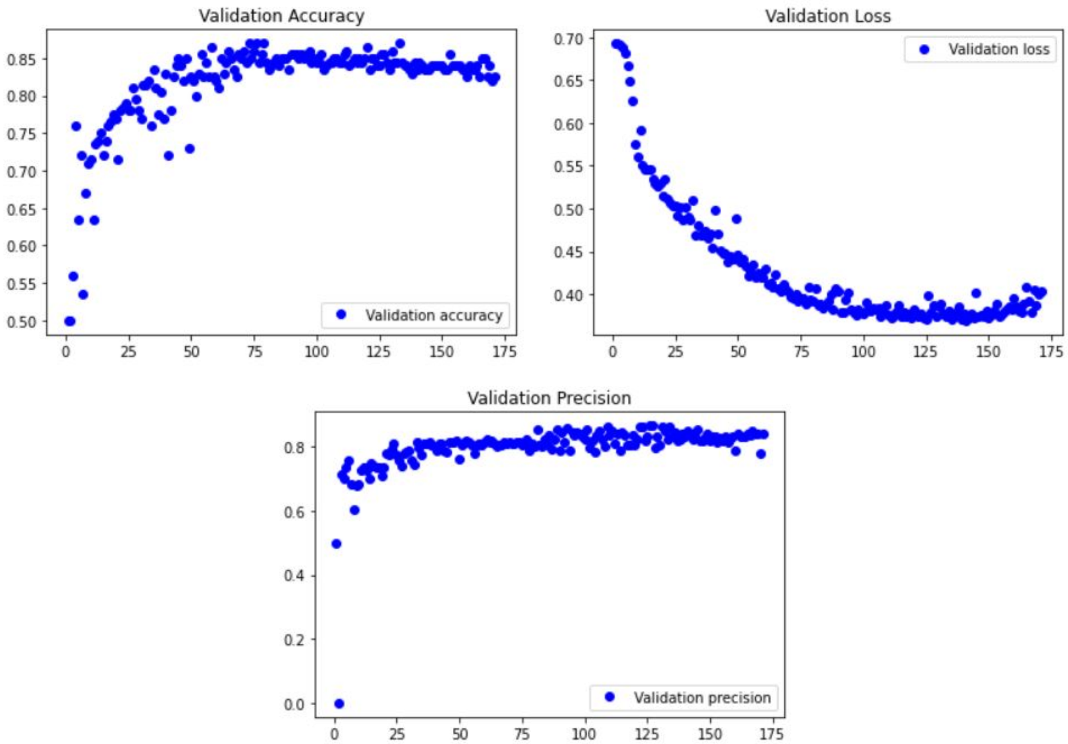

Figure 3: Validation accuracy, loss, and precision for the best-performing random rotation configuration.

```
50/50 [==============================] - 138s 3s/step - loss: 0.4395 - accuracy: 0.8300 - precision:
0.8056
```

Figure 4: Test metrics for the best-performing random rotation configuration.

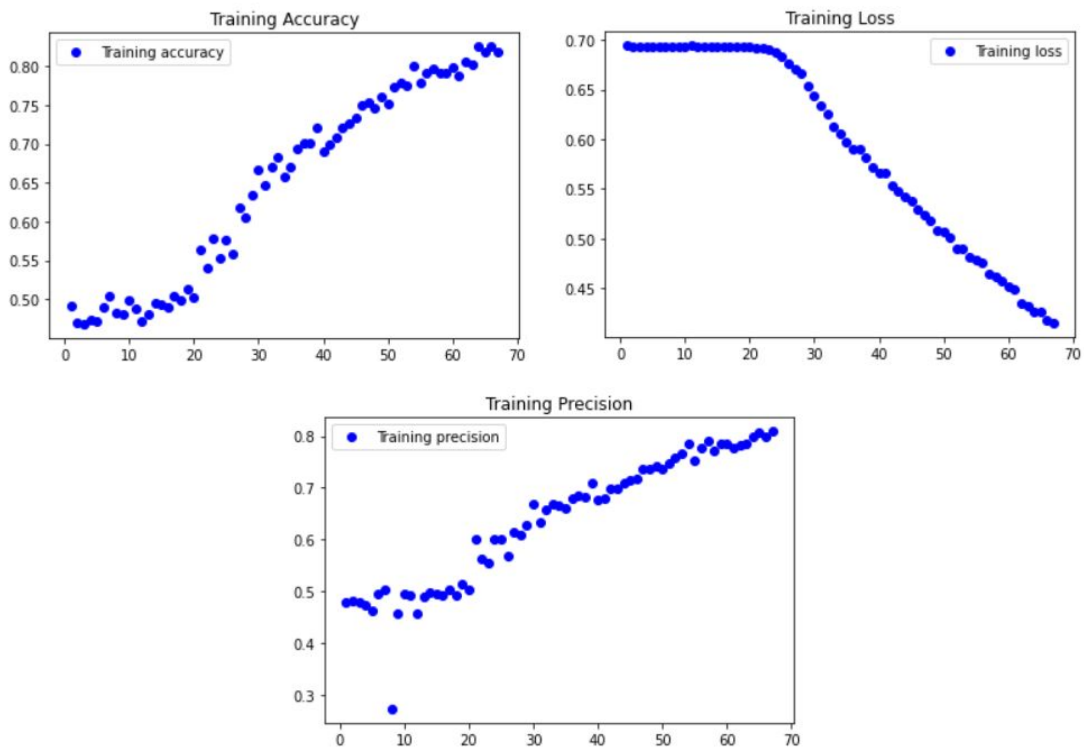

Figure 5: Training accuracy, loss, and precision for the best-performing random zoom configuration.

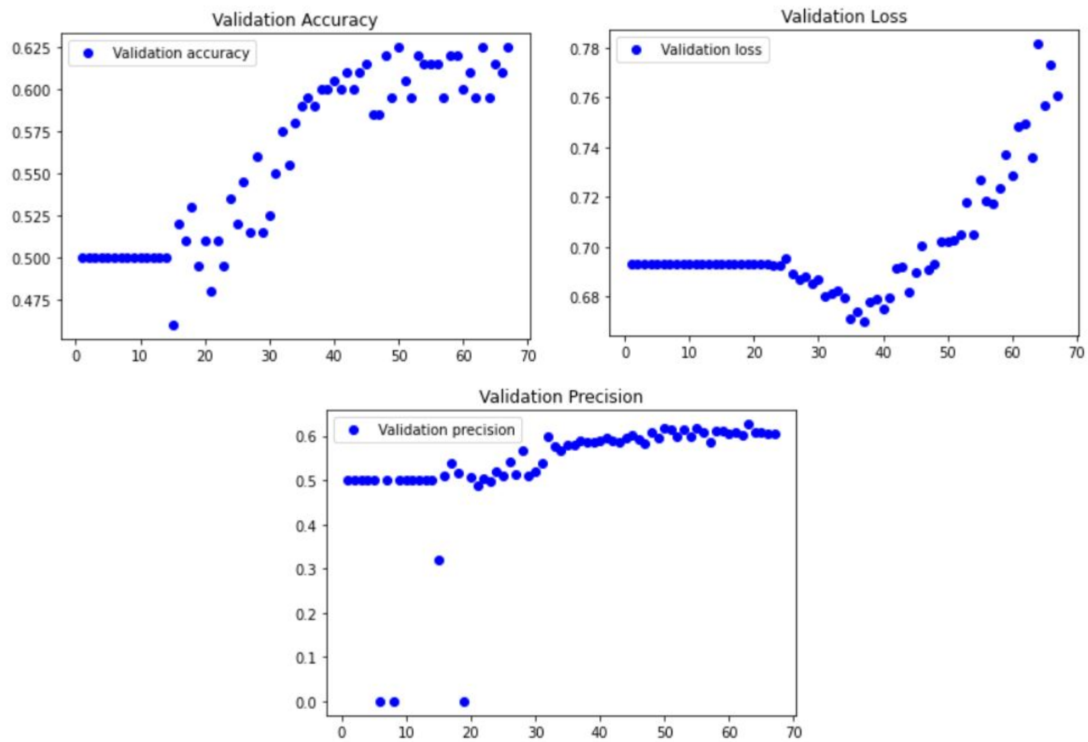

Figure 6: Validation accuracy, loss, and precision for the best-performing random zoom configuration.

```
50/50 [==============================] - 38s 773ms/step - loss: 0.6531 - accuracy: 0.6200 - precision:
0.6364
```

Figure 7: Test metrics for the best-performing random zoom configuration.

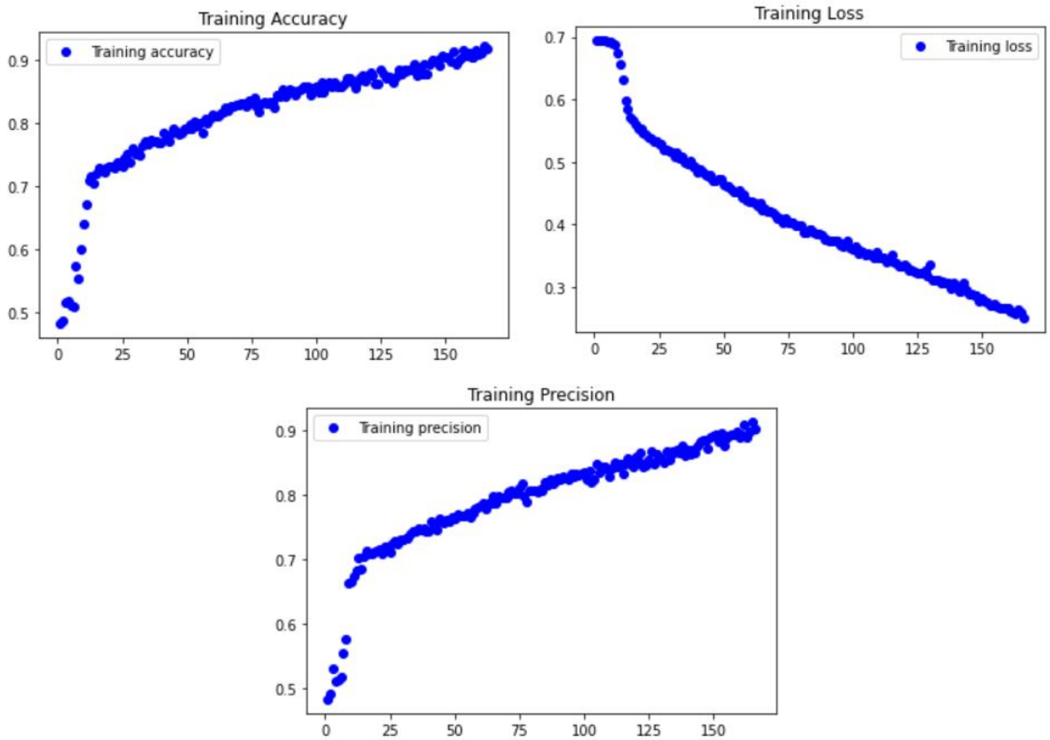

Figure 8: Training accuracy, loss, and precision for the best-performing random flip configuration.

simultaneously yielded the best results when any sort of flipping was done. Ultimately, the model implementing both types of flipping as well as rotation and zooming achieved a test accuracy of 0.825 and test precision of 0.7826, which was nearly equivalent to the model from the previous phase (Figure 10). Its test loss was 0.4338, three points lower than the model from the previous phase.

2 Best Data Augmentation Configuration

Taking the best-performing factors for rotation (0.2) and zooming (0.19), and testing a model using these configurations, a performance better than that of the model in the previous phase was attained (Figure 11).

2.1 Training and Validation

Compared to the model from the previous phase, the best-performing model employing data augmentation reached its greatest metrics approximately thirty epochs earlier, and thus ended training far earlier (Figure 12). The minimum validation loss, which was 0.3596, occurred at epoch 104, the maximum validation accuracy of 0.88 was seen at epoch 119, and the maximum validation precision hovered just under 0.85 from epoch 95 onward until maxing out at 0.8586 at epoch 118 (Figure 13). These validation metrics reflect that data augmentation made the model generally more accurate and precise throughout the training/validation process.

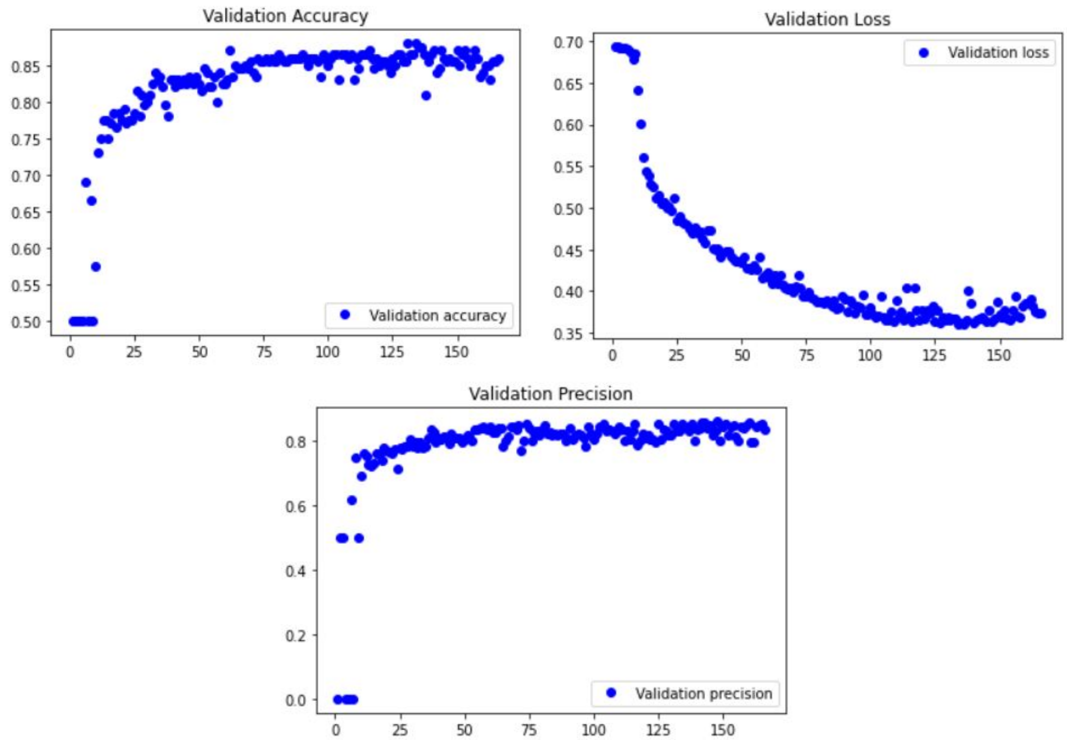

Figure 9: Validation accuracy, loss, and precision for the best-performing random flip configuration.

```
50/50 [==============================] - 41s 836ms/step - loss: 0.4338 - accuracy: 0.8250 - precision:
0.7826
```

Figure 10: Test metrics for the best-performing random flip configuration.

```
data_augmentation = keras.Sequential(
    [
        layers.RandomRotation(0.2),
        layers.RandomZoom(0.19)
    ]
)
```

Figure 11: The most effective data augmentation configuration found after experimentation.

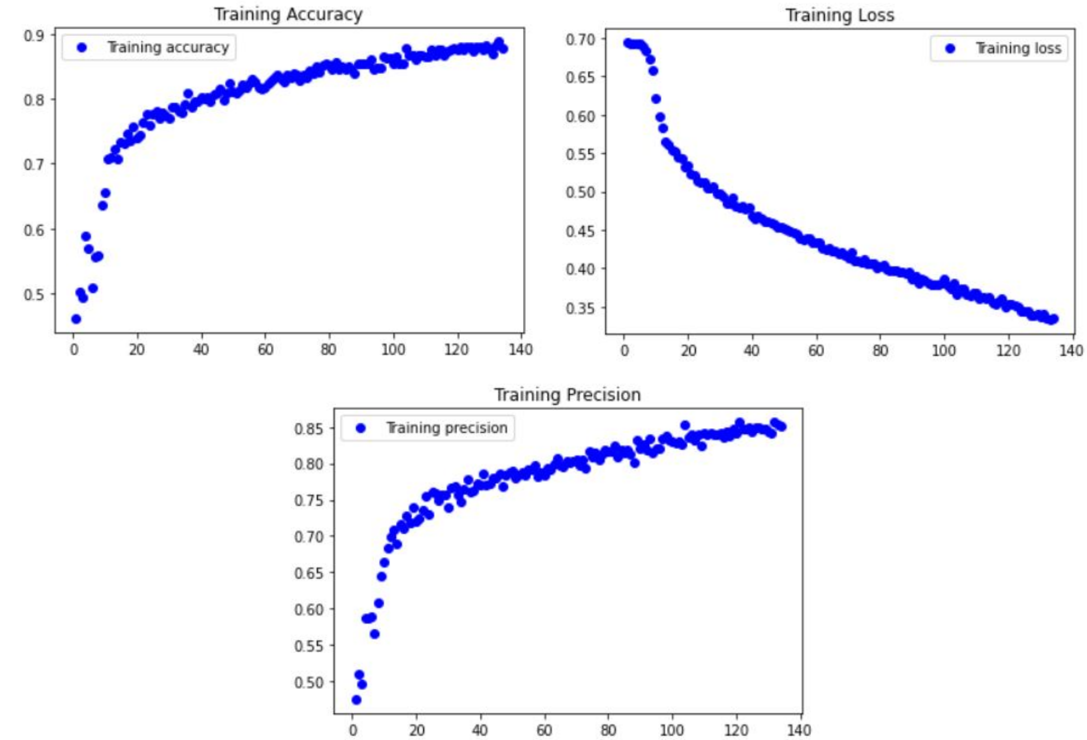

Figure 12: Training accuracy, loss, and precision of the best-performing model.

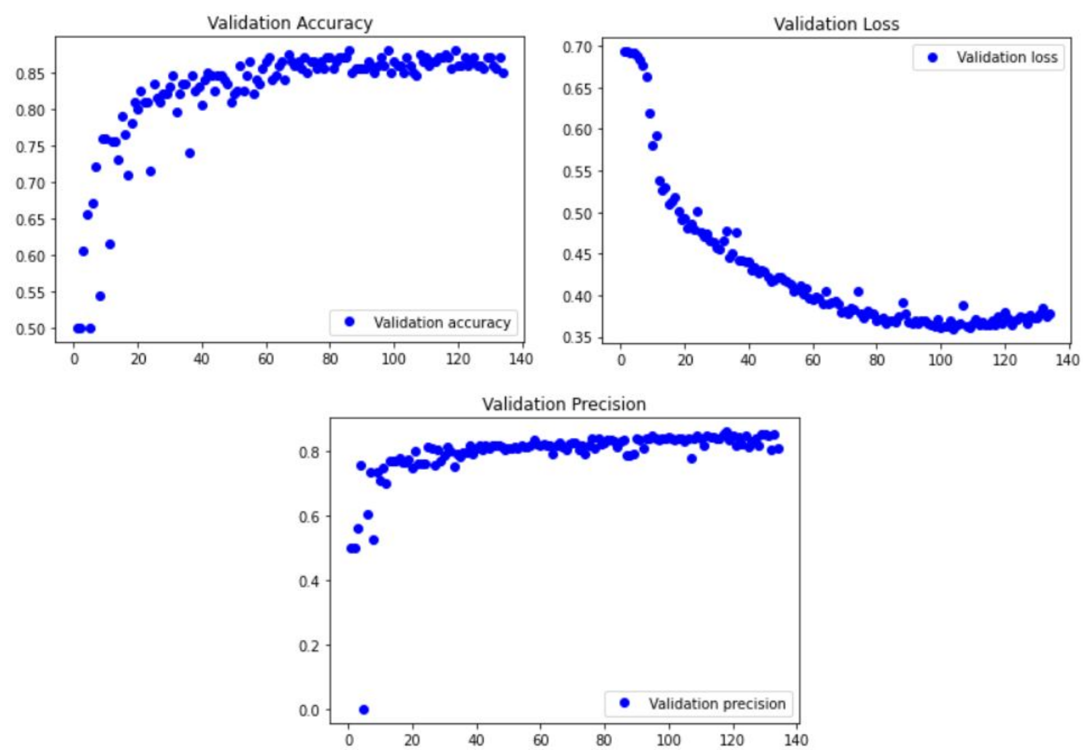

Figure 13: Validation accuracy, loss, and precision of the best-performing model.

```
50/50 [==============================] - 61s 1s/step - loss: 0.4349 - accuracy: 0.8400 - precision: 0.
7982
```

Figure 14: Results of testing the best-performing model.

Augmentation Form	Test Accuracy	Test Loss	Test Precision
Previous Model (No Augmentation)	0.8300	0.4675	0.7895
Rotation	0.8300	0.4395	0.8056
Zooming	0.6200	0.6531	0.6364
Flipping (with Rotation and Zooming)	0.8250	0.4338	0.7826
Final Model (Flipping and Rotation)	0.8400	0.4349	0.7982

Figure 15: Results from all experimentation during this phase.

2.2 Evaluation

The test metrics of the final model demonstrate a slight improvement compared to those of the model that did not utilize data augmentation. The test accuracy of 0.84 and test precision of 0.7985 were each only one point higher than the previous phase's best, while the test loss of 0.4349 was about three points lower (Figure 14). Ideally, with the introduction of normalization techniques, the performance gains of the next phase's model will be more pronounced.

3 Table of Results

See Figure 15.