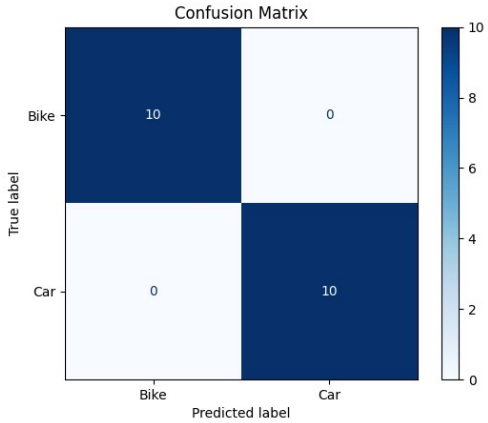
Karim Abu-Shams and Paula Pina

MODEL REPORT:

KEY PERFORMANCE VISUALIZATIONS

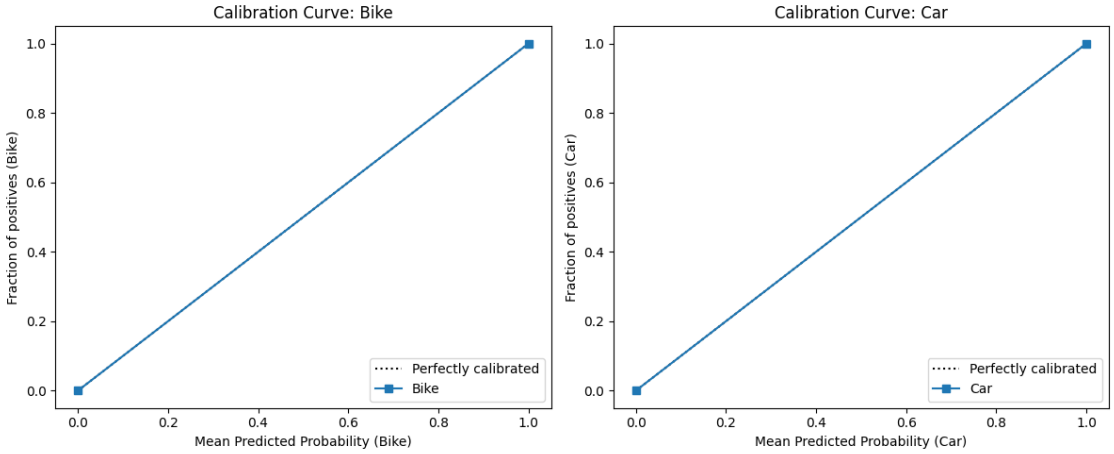
Confusion Matrix

As we can see, the model perfectly classifies each example of the validation set into their true class. This might be because we're using a very powerful pre-trained model (VGG11) for a very simple problem.



Calibration Curve

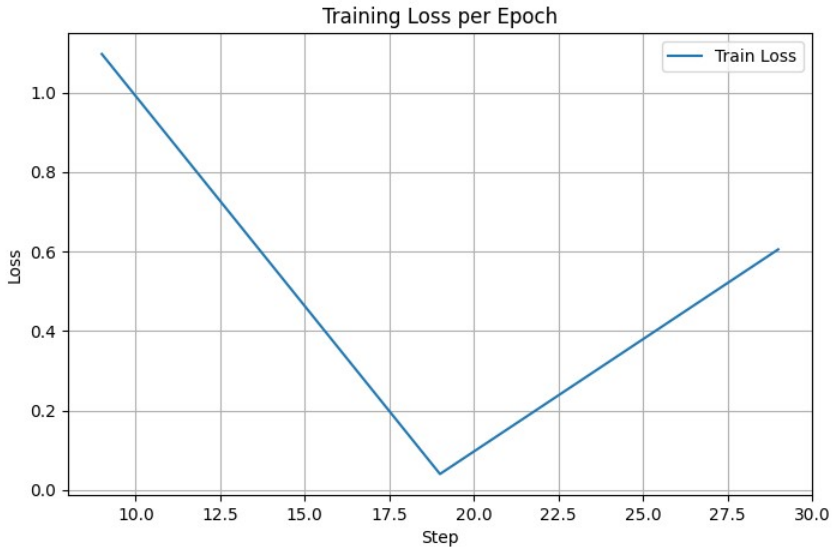
A perfectly calibrated model is one whose predicted probabilities match the actual likelihood of outcomes. For example, if the model predicts a 70% chance of a car, then in reality, 70% of such predictions are indeed cars.



In this case, the calibration curve lies exactly on the diagonal, indicating that the model's predicted probabilities are very reliable. When it predicts 80% confidence for a class, that class truly occurs roughly 80% of the time. This suggests the model not only distinguishes between cars and bikes correctly but also assigns probability estimates that can be trusted for downstream decisions.

Training Loss per Epoch

Here we see that the training loss first decreases but then decreases over our 3 epochs. It isn't a good result, since we always want the error to decrease as epochs go by. Still, since we only have 3 epochs and our dataset is quite small, we still get really good results.



Validation Accuracy per Epoch

We see that validation accuracy starts at 0.95 and goes up to 1.0 after just one epoch. In the end, our model perfectly classifies our validation data.