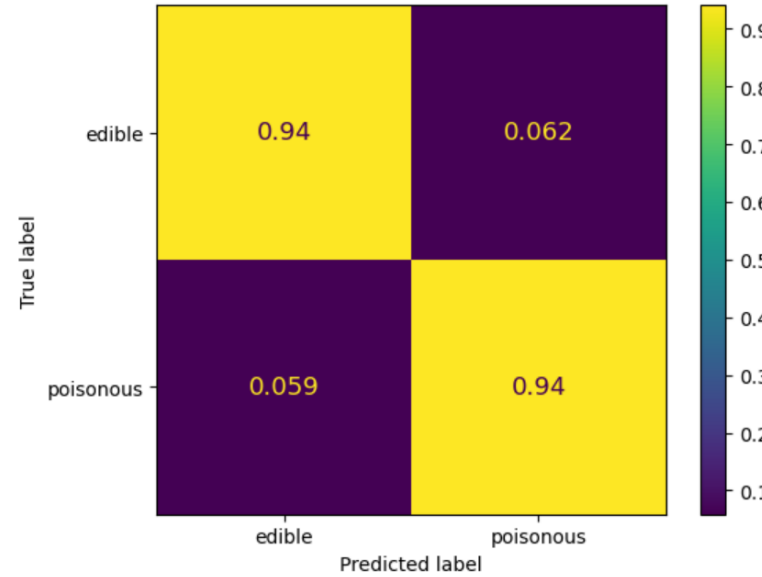
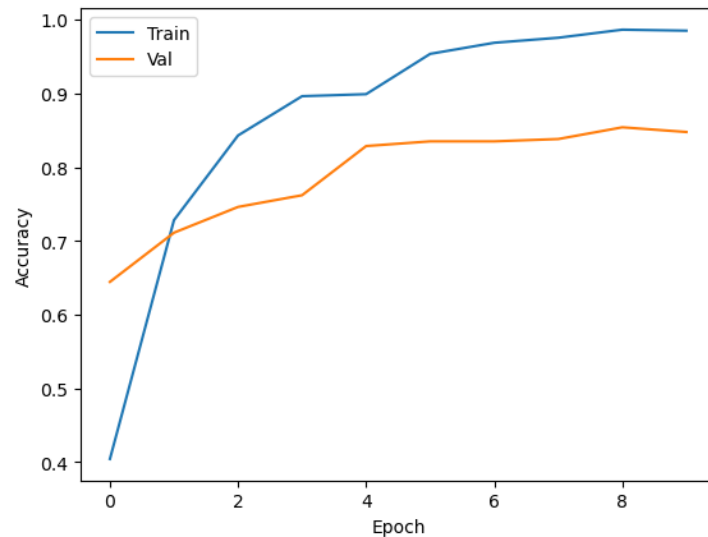
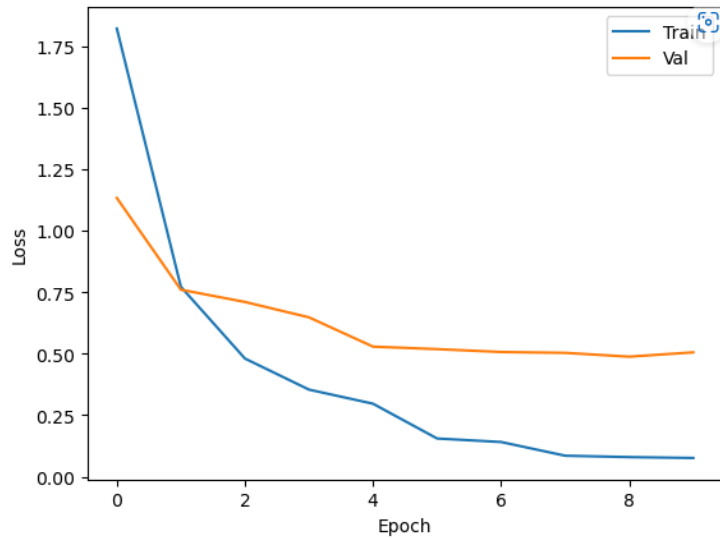


Goal & The recommended ML model

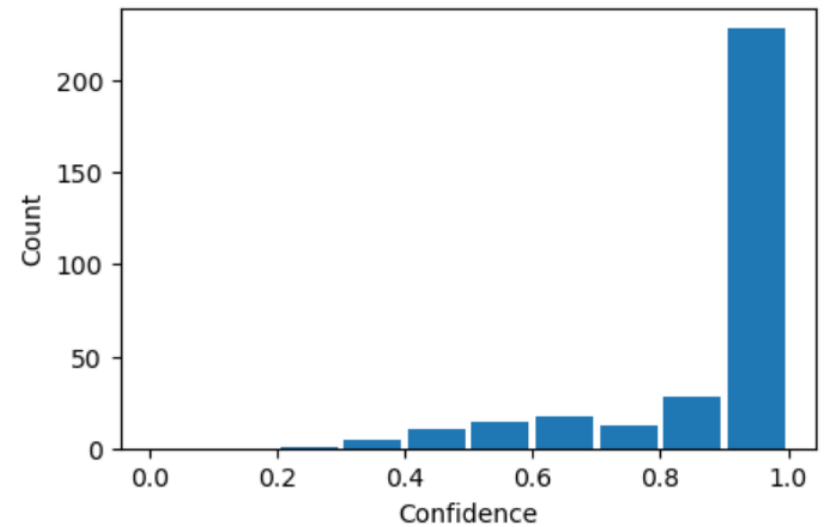
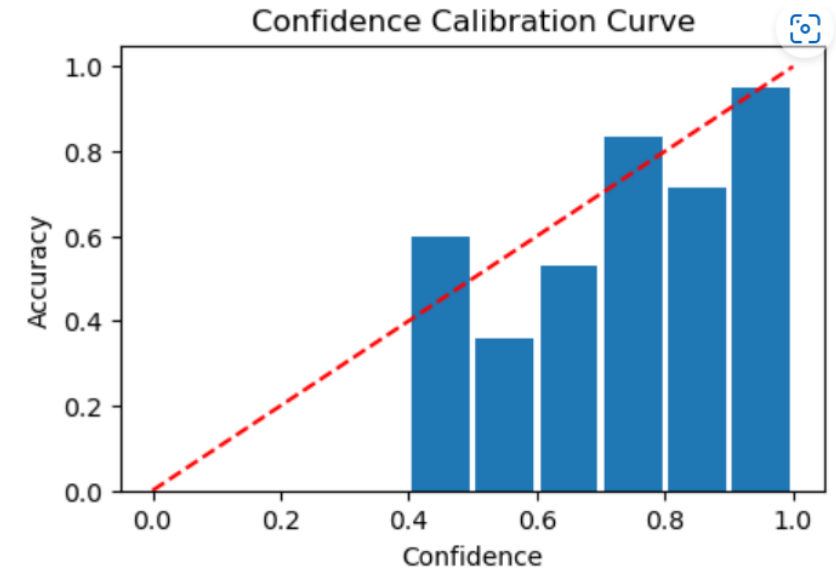
- **Goal:** Using Machine Learning Algorithm to classify fungi into 'edible' or 'poisonous' species from photos.
- **Recommended Model Architecture:** DINOv2 + Deep Neural Network (DNN)
 - **DINOv2** is a self-supervised foundation model used for extracting image features from photos. These features can then be applied to further tasks.
 - **Deep Neural Network (DNN)** is used as the classifier in this case. The DNN will take the labeled features extracted by DINOv2 as input, and through 3 layers, it will train and classify fungi species as either "edible" or "poisonous."
- **Data Preprocessing:**
 - To address the issue of class imbalance, we are using a **WeightedRandomSampler**.
 - Some **data augmentations** are applied to improve the generalization.

DINOv2 + DNN

Validation accuracy is 94.46%



True Positives (TP): 0.94
False Negatives (FN): 0.062
True Negatives (TN): 0.94
False Positives (FP): 0.059
For the Edible Class:
Precision: 93.2%, Recall: 93.8%

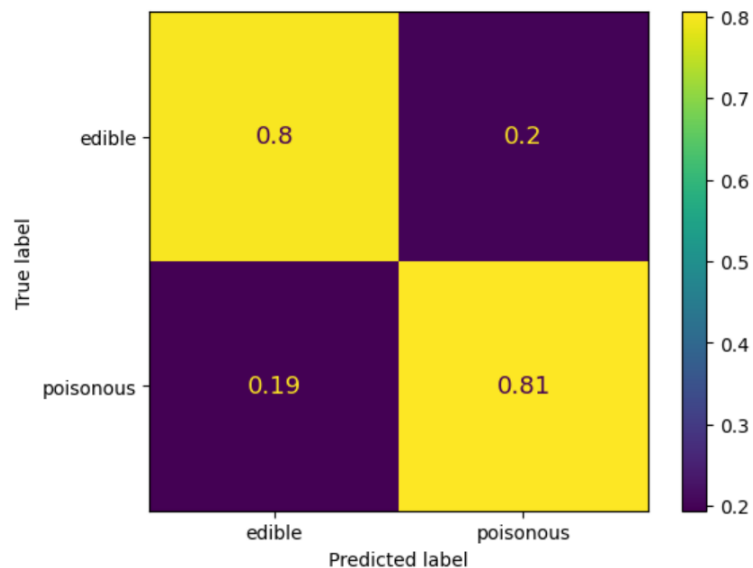
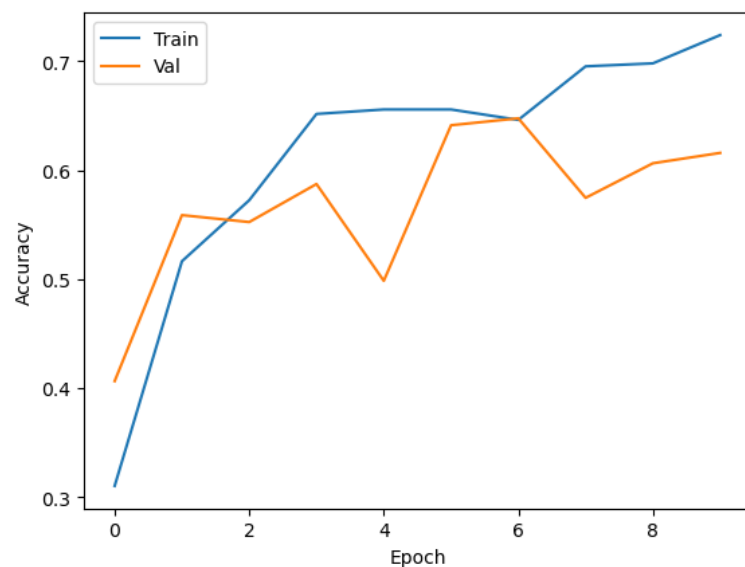
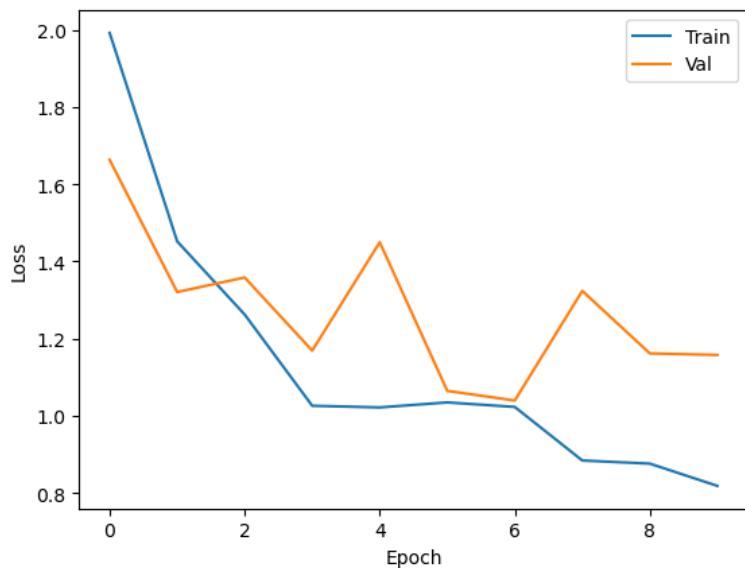


Why I recommend this model

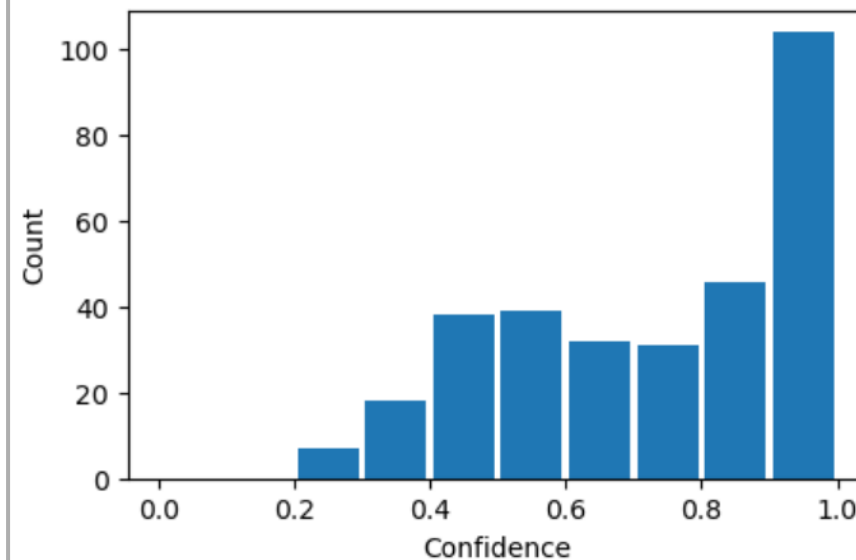
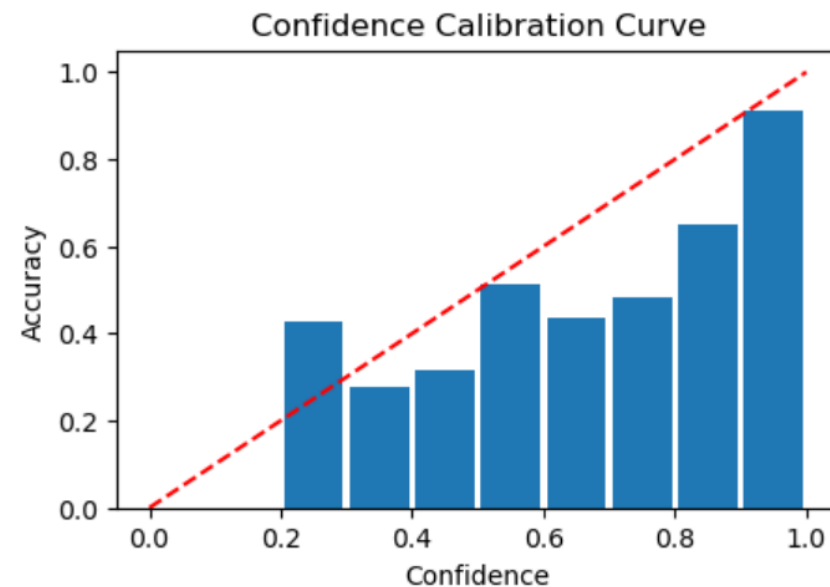
- Let's compare with Resnet18.
- Resnet18 is a supervised pretraining model, its structure includes the feature extraction from the labeled data, followed by a linear classifier.

ResNet18

Validation accuracy is 73.35%



True Positives (TP): 0.8
False Negatives (FN): 0.2
True Negatives (TN): 0.81
False Positives (FP): 0.19
For the Edible Class
Precision: 77.9%, Recall: 80.0%



Conclusion

- The Dino + Deepnet model outperforms the Resnet model on all major performance metrics, including precision, recall, and ground truth/false positive rate. The confusion matrix of the Dino + Deepnet model shows that it classifies the "edible" and "other" categories more accurately, with a lower false positive rate.
- Performance: The Resnet model shows an obvious overconfidence problem, which may be due to overfitting or the probability output of the model is not properly calibrated. Although the Dino + Deepnet model also shows high confidence, its confidence calibration curve is closer to the ideal curve, indicating that its degree of overconfidence is low, and the confidence of the model prediction is more consistent with its accuracy.

Limitation

- Slight Overfitting: As shown in the loss and accuracy plots, there is a small gap between the training and validation curves, indicating potential overfitting to the training data.
- Calibration: From the confidence calibration curve, the model is slightly overconfident, meaning it may overestimate the confidence of its predictions.
- Memory limitations: I use 10 epoch and batch size 10, so if you can have a bigger memory on your device, can be trained for more epoch and bigger batch size
- Data imbalance: although I have used the WeightedRandomSampler to account for imbalance data, but if the numbers of data in each species are balance in the beginning, the performance can be more correct.

Recommendations

- **1. Regularly Monitoring and Calibration**
 - Regularly calibrate the model's confidence levels to ensure they the true predictive accuracy.
- **2. Expanding the Dataset**
 - Can Consider augmenting the dataset with more samples and diverse, to further enhance the model's generalization capabilities.
- **3. Expand the memory**