Trimble / Bilberry : AI Engineer technical exercise
Part 1 : Technical Exercice

Chappe Aslan
August 2023

Deep Learning Binary Classification Model using ResNet-101

1. Methodology

a. Data Verification and Augmentation Process for the Dataset

It was observed that the dataset is low volume and imbalanced, with a higher prevalence of instances belonging to the "Road" class compared to the "Field" class. This issue justified the methods used to verify the dataset and address its class imbalance.

Addressing the class imbalance and poor quantity applying data augmentation techniques, we aim to deal with the balance between the "Road" and "Field" classes while enhancing the dataset volume. The augmentation methods are RandomHorizontalFlip, ColorJitter and RandomRotation.

b. Model Design

The ResNet-101 [1] architecture was chosen to extract intricate features from images. ResNet-101 is renowned for its depth and capability to capture information at various scales.

Training and Optimization: The model was trained for 18 iterations using the Adam algorithm. Weighted BCE loss was employed as the loss function, with weights adjusted to account for class imbalance.

c. Model evaluation

Model evaluation was performed using metrics such as accuracy, F1-score, confusion matrix, Precision-Recall Curve and ROC-AUC curve. These metrics provided insights into the model's performance across both classes.

2. Results

After comprehensive training and evaluation, the model showcased promising performance. The strategic integration of the ResNet-101 architecture, weighted BCE loss, and data augmentation enhanced the model's ability to discriminate between classes. The best model obtained achieved an accuracy of 0.962.
Regarding the number of training epochs, we've observed that the accuracy tends to decrease after around 18 epochs. Beyond this point, overfitting becomes a concern, causing the accuracy to decline and fluctuate between lower values.

Depending on how the dual-class image is labeled, situations can arise where the F1 score equals 1, leading to the corresponding confusion matrix (Figure 1).

This model can also obtain a perfect Precision-Recall Curve (Figure 2) and ROC-AUC Curve (Figure 3) on this test dataset.

3. Discussion

Adding data augmentation techniques and using a weighted loss made the model stronger, especially when dealing with imbalanced classes. However, because of the dataset's nature, this model works well with single-labeled classes. But, it might not be as reliable when dealing with images labeled for two classes, like we can see with image 7 in the test dataset, where it predicted only one class.

Another approach to tackle the issue of class imbalance is dataset oversampling. Specifically, employing techniques like SMOTE (Synthetic Minority Over-sampling Technique) oversampling can be a valuable strategy in this context. Exploring these methods could serve as an alternative path to enhance the performance of this model.

Finally, this model is performing really well, but there's still potential for improvement by diving deeper into fine-tuning hyperparameters. We've already achieved great results, and by tweaking these settings further, we could take the performance to an even higher level.

4. Sources

[1] https://pytorch.org/vision/main/models/generated/torchvision.models.resnet101.html

5. Additional Data

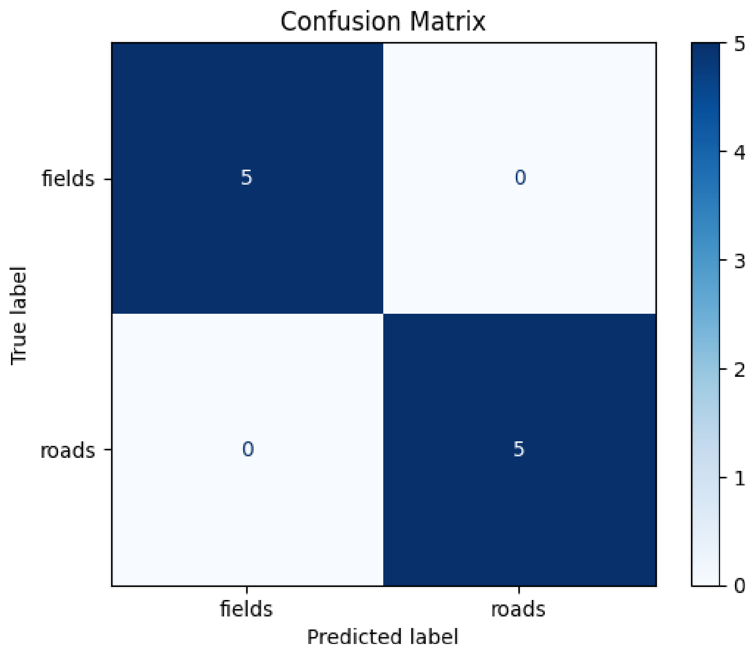

Figure 1. F1 score and Confusion Matrix

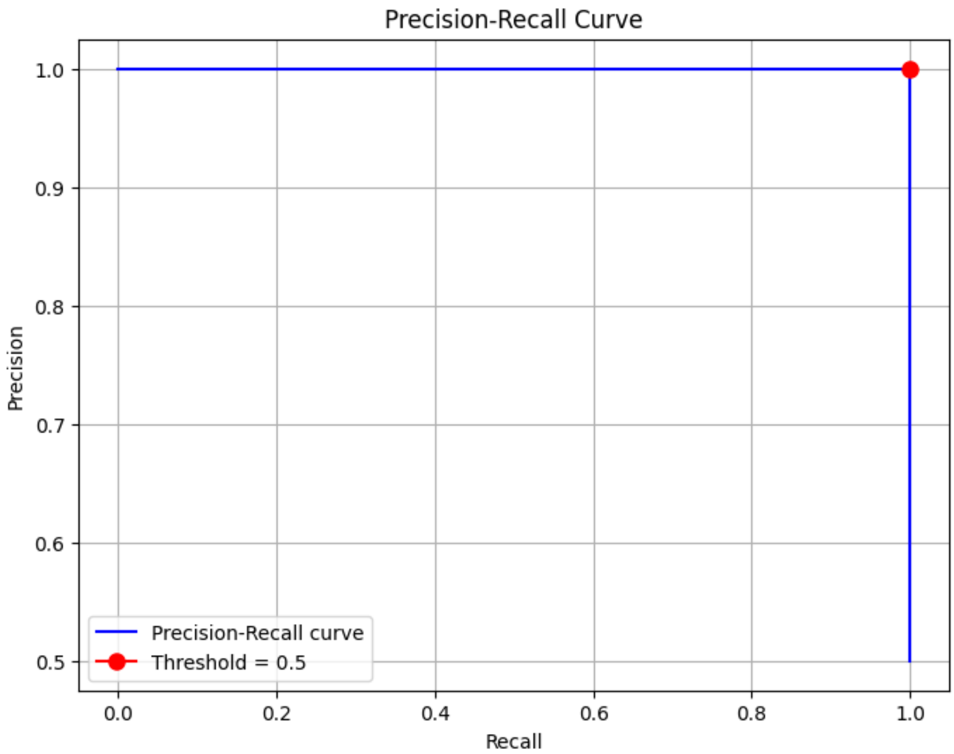

Figure 2. Precision-Recall Curve

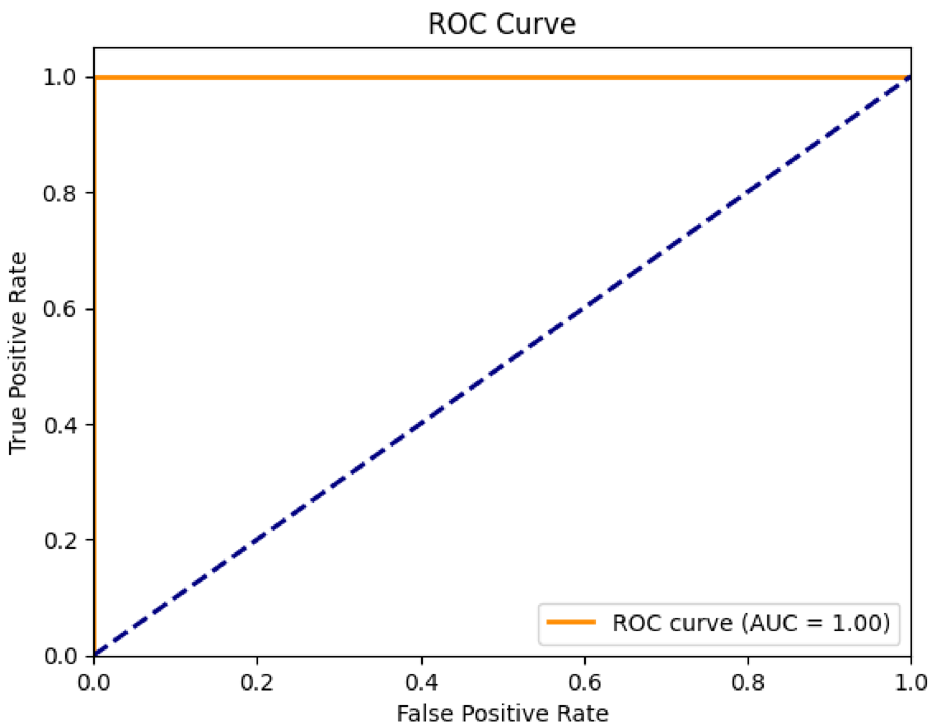

Figure 3. ROC-AUC Curve