Road and Field Image Classifier

Introduction:

This project aims to develop a two class classifier : Field & Road using the available data. To accomplish this task, a pre-trained VGG16 model with transfer learning was utilized to develop the classifier.

Data Preprocessing:

The dataset exhibits an imbalance in favor of the roads class, which contains a total of 108 images compared to 45 images for the field class. Additionally, the images come in a variety of sizes, with dimensions ranging between 165 and 2400. However, the majority of the images have resolutions that fall between 220 and 350 pixels, as shown in Figure 1.

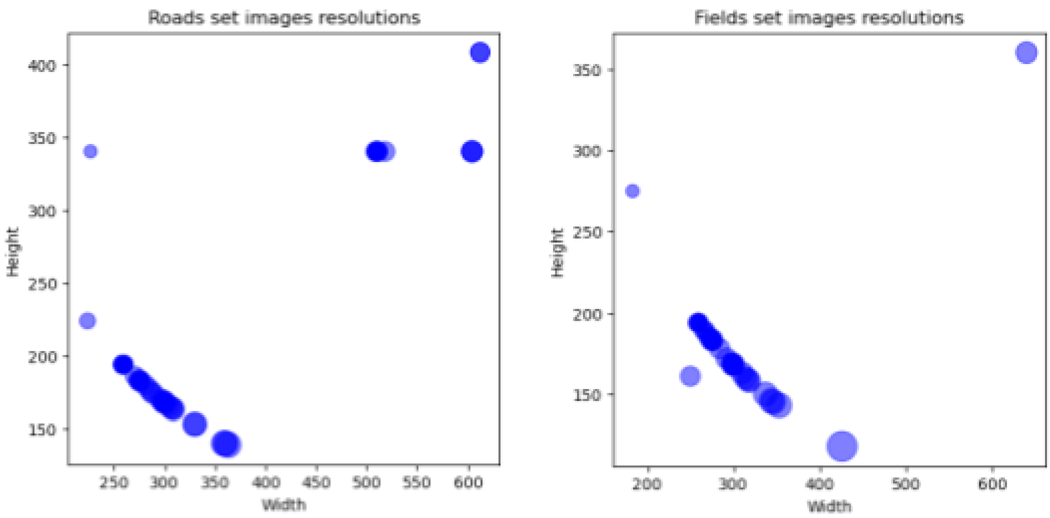

Fig 1 : Images resolutions distribution

To load and preprocess the dataset, the image_dataset_from_directory function from TensorFlow was used. This function creates a dataset from the images stored in the designated directory. The images were resized to 224x224 pixels, which is the default input size for vgg16 and also happens to be relatively close to the majority of the input sizes in the dataset, thus minimizing the loss of information during the resize process. In addition, pixel values were normalized to the range [0, 1]. Finally, the dataset was split into 80% training and 20% validation sets.

Model Architecture:

The VGG16 pre-trained model served as the base model with weights pre-trained on ImageNet. The top layers of the original VGG16 model were excluded, and a Flatten layer was added, followed by a Dense layer with 64 units and ReLU activation. To prevent overfitting, a Dropout layer with a rate of 0.1 was included. Finally, a Dense layer with a single output unit and sigmoid activation was added to obtain binary classification results. The model was compiled using the Adam optimizer with a learning rate of 1e-4 and binary cross-entropy loss. The evaluation metrics used were accuracy and F1 score.

Model Training and Evaluation:

The model was trained for 15 epochs with early stopping, monitoring the validation loss, and a patience of 3 epochs to prevent overfitting. TensorFlow's TensorBoard was used to visualize training progress.

Results:

The training and validation loss curves showed that the model converged well during training.

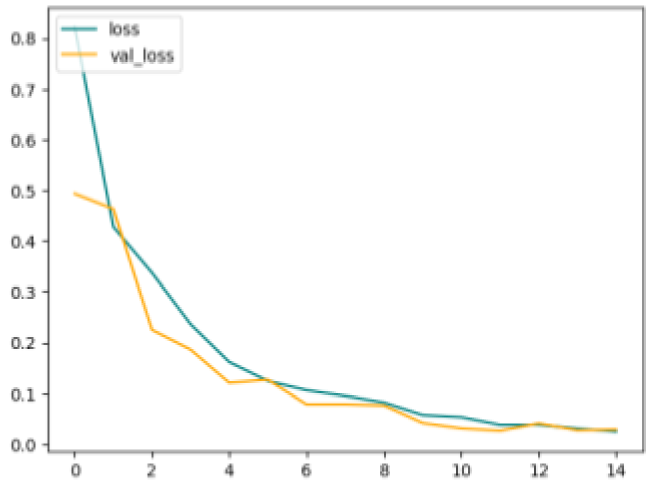

Fig 2 : Train and validation losses

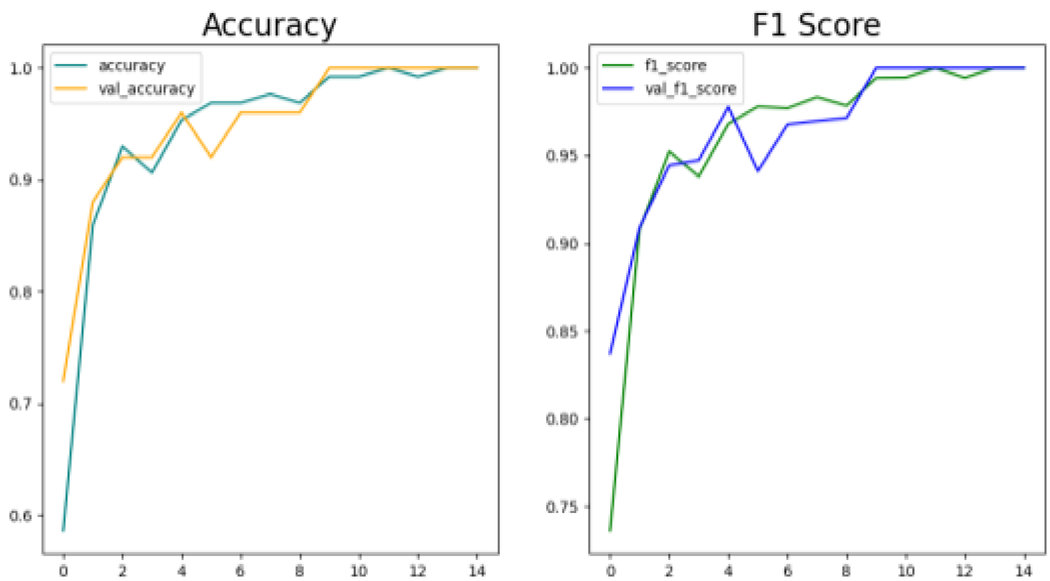

Fig 3 : Accuracy and F1 score

The accuracy and F1 score curves also demonstrated good performance as shown in figure 3.

The provided test set was labeled, and the model was tested on it. The predictions were then visualized, as shown in Figure 4. The visualized predictions appear to be correct for all images.

Fig 4 : Test set prediction results

The model was evaluated on the test dataset and achieved perfect accuracy (1) and F1 score (1).

Conclusion and Possible Improvements

The VGG16-based image classifier for roads and fields achieved satisfactory results in terms of accuracy and F1 score. This project demonstrated the effectiveness of transfer learning and pre-trained models in addressing image classification problems with limited data. However, there is still room for improvement. For instance, other pre-trained models such as ResNet, Inception, or EfficientNet could be experimented with for transfer learning. Additionally, data augmentation techniques could be utilized to expand the size and diversity of the training data. It would also be worthwhile to explore different model architectures, such as adding more layers, adjusting dropout rates, using different activation functions, or incorporating Spatial Pyramid Pooling (SPP) to better handle varying input image sizes. Finally, to improve inference time, mixed precision or exporting the model to faster frameworks such as ONNX, tensorrt, or tflite could be considered.