Crop Classifications from Satellite Images

Abstract

The objective of this project was to harness satellite imagery of agricultural fields and leverage this data to construct a machine-learning model dedicated to the classification of crop types. The initial phase entailed acquiring satellite images and performing data preprocessing, aligning it with reliable ground truth data to prepare it for machine learning algorithms. Following each preprocessing method, we transformed the dataset into a flattened, pixel value-based structure and converted it into a CSV format. Subsequently, this structured dataset was utilized to train and evaluate both traditional machine learning algorithms and neural network-based models, to classify crop land images. The dataset can be found here

Data Sourcing

In this project, we focused on the farmlands of California's Fresno area, utilizing satellite images for crop classification. We developed two distinct datasets for this task. The satellite images were sourced from the National Agriculture Imagery Program (NAIP) for Fresno in 2018, accessed through earthexplorer. We used the USDA CDL crop-land labeled dataset as the truth standard for the same region and year. The CDL dataset spans 30 meters per land unit, while the NAIP images from EarthExplorer represent 1 meter per pixel.

We applied two data processing methods to label the datasets. In the first method, we resized and aligned the CDL data to match the NAIP images' spatial details. We ensured alignment accuracy by matching the CRS of the NAIP images to that of the CDL. We then extracted and labeled each pixel value from the NAIP images with the corresponding CDL class label in a loop. This resulted in a dataset with over 55-million-pixel value rows and their associated class labels. While extensive, this dataset was suitable only for traditional machine learning regression classifiers due to the loss of spatial context when flattening the images, making it unsuitable for neural networks that require full spatial information to detect patterns across images.

To overcome this limitation, we introduced a second approach. We first identified the top 5 classes, excluding the class label 255 to omit non-crop areas, and then iteratively aligned and cropped the NAIP images to match the CDL dimensions. This ensured that each NAIP pixel was mapped to a CDL pixel. We then compiled a subset of NAIP images with pixels all from the same CDL class. The final dataset combined these pixel values with their corresponding class labels for each class and NAIP image into a DataFrame. Utilizing 9 NAIP images from the area of interest, we assembled a dataset with dimensions (500000, 36), preserving the complete spatial information for CNN model compatibility. The data sourcing process is outlined below:

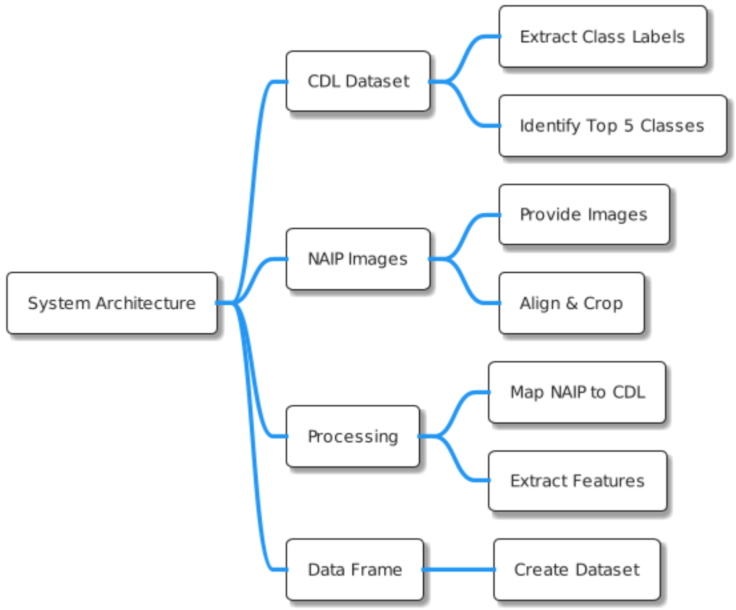

Fig 1: Date Processing Architecture

Method

We explored a variety of Neural Network models and Ensemble Learning Approach with multi-class classifier models and regression models such as Random Forest Classifier, Gradient Boosting Classifier and Logistic Regression. The model performance evaluation and selection process are as follows.

Model 1: Dense Neural Network

The Neural Network model exhibited a performance accuracy of 37%. This model employs fully connected layers, where each neuron is linked to every neuron in the preceding layer. While such an architecture is beneficial for certain types of data, it is less effective for image classification tasks, where spatial and hierarchical patterns are crucial.

In image data, localized feature extraction is vital, and dense layers may not effectively capture these features. This limitation is related to the lack of spatial hierarchy in dense networks, as opposed to convolutional layers used in more advanced models. The evaluation reports are given below

```
... 782/782 [==============================] - 0s 226us/step
              precision    recall  f1-score   support

           0       0.30      0.19      0.23      4948
           1       0.44      0.60      0.51      5007
           2       0.29      0.22      0.25      4969
           3       0.39      0.50      0.44      5035
           4       0.34      0.32      0.33      5041

    accuracy                           0.37     25000
   macro avg       0.35      0.36      0.35     25000
weighted avg       0.35      0.37      0.35     25000
```

Fig 2: Accuracy Report Dense Neural Network

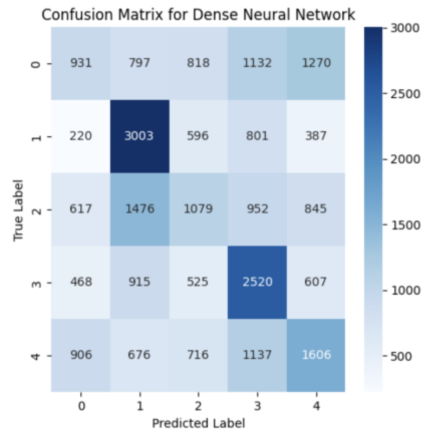

Fig 3: Confusion Matrix for Dense Neural Network

Model 2: 1D Convolutional Neural Network

The 1D Convolutional Neural Network, with an accuracy of 52.13%, marks an improvement over the Dense Neural Network. This model employs convolutional layers that are proficient at extracting spatial features from image data, using shared weights and focusing on local connectivity.

However, the architecture of this model might still be too simplistic to fully capture the complexity and diversity of the patterns present in the dataset. This is evident from its moderate performance, indicating potential limitations in dealing with complex image classification tasks. The evaluation reports are given below

```
... 782/782 [==============================] - 1s 738us/step
              precision    recall  f1-score   support

           0       0.40      0.31      0.35      4948
           1       0.58      0.75      0.65      5007
           2       0.38      0.34      0.36      4969
           3       0.74      0.69      0.71      5035
           4       0.47      0.51      0.49      5041

    accuracy                           0.52     25000
   macro avg       0.51      0.52      0.51     25000
weighted avg       0.52      0.52      0.52     25000
```

Fig 4: Accuracy Report 1D Convolutional Neural Network

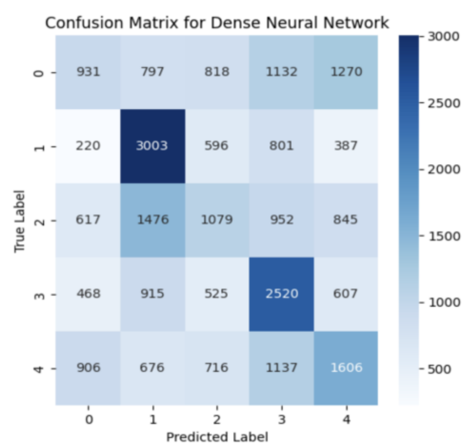

Fig 5: Confusion Matrix for 1D Convolutional Neural Network

Model 3: Advanced Neural Network with Residual and Inception-like Modules

Achieving an accuracy of 78.05%, this model outperforms the other models significantly. The integration of residual connections in this model aids in addressing the common issues in training deeper neural networks such as `vanishing gradient problem`. The inception-like modules enable the network to extract features at various scales, making it highly effective for image data where relevant features can vary significantly in size and shape. The model evaluation matrixes are given below

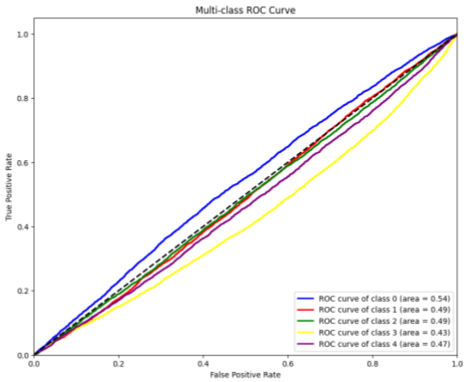

Fig 6: ROC Curve Unique classes

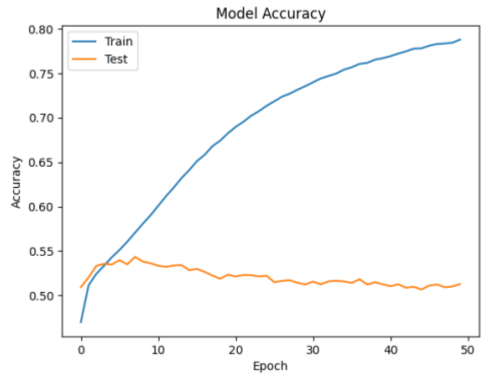

Fig 7: Model Accuracy Graph on Test and Train Data

The combination of various types of layers and modules provides the model with a higher capacity to learn nuanced patterns and relationships, making it particularly suitable for complex tasks like image classification. Although the predicting accuracy on test data of the model is from 55%-50%. The underlying reasons should be aligned with the number of NAIP images used (9) which is low comparatively. The learning accuracy of the model will significantly increase if we can use all satellite image of the Fresno region of California state available for 2018. It will require a significant amount of computational power if it is done through downloading the images from the https://earthexplorer.usgs.gov/ and fitting into the data-processing pipeline. We have searched for API to fetch NAIP images of a particular region in a given timeline but could not find such facilities available. The API based facility is available for Sentinel Satellite images which requires certain government or educational institute authorization.

Model 4: Ensemble Learning Approach

An ensemble learning approach, combining RandomForest and GradientBoosting classifiers, achieved an accuracy of 52.94%. While ensemble methods are known for improving performance by improvising the scores of multiple models.

Logistic Regression is a linear model and might not capturing the complex relationships present in the predictions from the base models (RandomForest, GradiantBoosting Classifiers), especially in the context of image data. The evaluation matrix is given below

Final Accuracy:	precision	recall	f1-score	support
61	0.75	0.71	0.73	100000
69	0.46	0.55	0.50	100000
75	0.41	0.30	0.35	100000
176	0.46	0.42	0.44	100000
204	0.55	0.67	0.61	100000
accuracy			0.53	500000
macro avg	0.53	0.53	0.52	500000
weighted avg	0.53	0.53	0.52	500000

Fig 7: Classification Report Ensemble Model

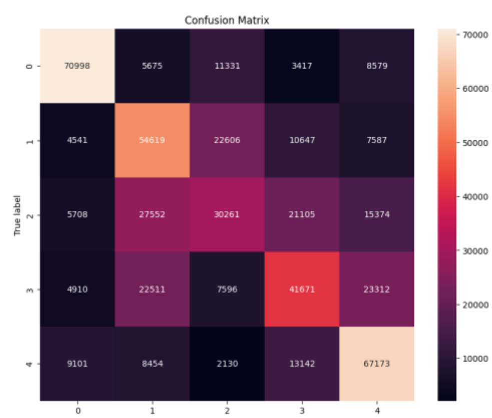

Fig 8: Confusion Matrix Ensemble Model

Conclusion
Model 3, with its advanced architecture, was found to be the most effective for the given image classification task, outperforming the other models in terms of accuracy and capability to capture complex patterns in the data.

This study highlights the potential of using advanced neural network architectures, such as those incorporating residual and inception-like modules, in effectively handling the intricacies of image classification tasks.

Challenges
During the project the most challenging part was to source the data and pre-processing. First challenge was to locate the AOI region to be extracted from CDL and NAIP image datasets. To overcome the challenge, we figured out the 4 corner's latitude and longitude of the AOI region and extracted the CDL data. Whereas we had to go through 100 of NAIP AOI region's images and selected 9 relevant images to use for model training. While pre-processing the dataset we faced challenges due to dimensional, CRS miss match between the two data sources which is overcome with CRS conversion and NAIP images alignment and cropping in reference with CDL image. The computational challenge was faced while training the Neural Network models on the large dataset. The training time was much long. Lastly, the crucial challenge is to improve the train accuracy of the Model 3. The scope of work to resolve the challenge is divided into two parts

1. Adding all the NAIP images of area of interest from the year 2018 into the data processing and dataset creating pipeline.
2. I have classified the crop types based on NDVI index; the indexed dataset is having avg_ndvi as an identifier of the class type. Using the Neural Network models on the refined dataset will increase the test and train accuracy.
3. Using more advanced Neural Network models to be used for the Crop Classification task:
    a. Recurrent Neural Net (LSTM)
    b. Transformer
    c. Multi-scale RestNet

References

1. Revised Dataset: GDrive
2. 2017- Rose M. Rustowicz," Crop Classification with Multi-Temporal Satellite Imagery"
3. Rußwurm, M., & Körner, M. (2019). Self-attention for raw optical Satellite Time Series Classification. *arXiv*. https://arxiv.org/abs/1910.10536v3

Steps to run the project:

1. Create a python env or use conda environment for package management and install the required project mainly GDAL, OSGEO, Tensorflow, Keras.
2. For data processing approach 1 run data_process_forenso.ipynb
3. For data processing approach 2 and the selected approach run data-preprocessing.ipynb
4. Model selection on output dataset of approach 1 run model_selection.ipynb
5. Model selection on final dataset run Crop_classification_DL_model_redefined.ipynb