

PROJECT REPORT

Land-Cover Classification

Vivin T
Praveen Raj M

Introduction

Due to converging influences, the data science and remote sensing communities have started to align in recent years. Deep learning (DL) is an umbrella concept that encompasses a variety of algorithm architectures based on neural networks. Multi-layer perceptrons, deep belief networks, stacked autoencoders, deep neural networks, and restricted Boltzmann machines are examples of these architectures. Since at least the late 1980s, neural networks have been used to classify satellite images. Since at least the late 1980s, neural networks have been used in satellite image classification and have been implemented in remote sensing software packages, typically with one or two hidden layers. Since data was costly and processing power was insufficient, the number of hidden layers used in satellite image classification remained limited. With the increased availability of Big Data and computational power at the turn of the century, this all changed, necessitating the use of more (i.e. deeper) secret layers and complex network architectures. Since 2015, DL has been used in a variety of applications including mapping land cover and crops, estimating crop yields, detecting oil palm trees and plant diseases, and more, with great accuracy. The aim of this project is to work on deep learning algorithms to classify the land use and land cover with a good accuracy score.

Data

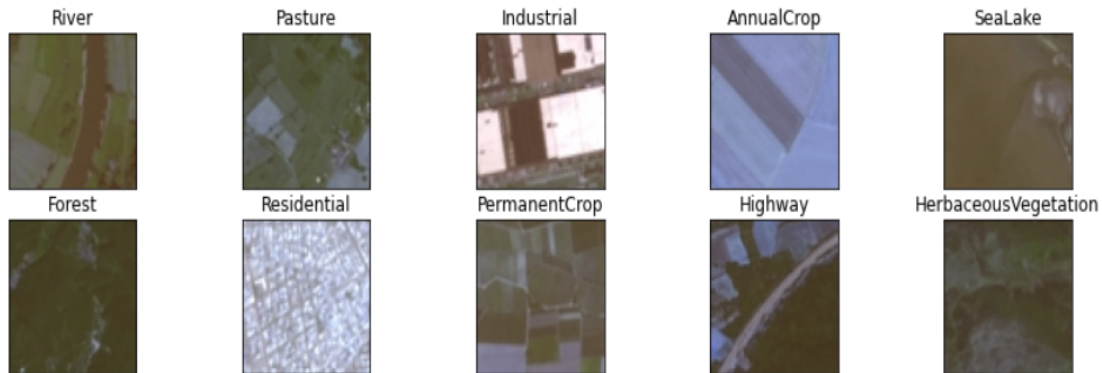
Remotely sensed image classification is a difficult process. The lack of accurately labelled ground truth datasets has hampered the development of classification in the remote sensing field. The Sentinel-2 Multispectral Instrument (MSI) consists of two satellites that observe the Earth at spatial resolutions of 10 m, 20 m, and 60 m. Among freely available satellite items, the 10 m spatial resolution stands out as the best. The original dataset contains 10 classes and 27000 labeled images in 64x64 format. The data from Sentinel-2 is multispectral, including 13 bands in the visible, near infrared, and shortwave infrared spectrum. These bands have varying spatial resolutions ranging from 10 to 60 metres, classifying images as high-medium resolution. Although other satellites with higher resolution (1m to 0.5 cm) are available, Sentinel-2 data is free and has a long revisit period (5 days), making it an excellent choice for monitoring land use. The inclusion of three red edge bands in the Sentinel-2 data, which can capture the heavy reflectance of vegetation in the near infrared portion of the electromagnetic spectrum (EMS), is another unique feature of the data .

Some popular datasets in remote sensing includes UC Merced (UCM) land use dataset introduced by Yang et al , manually created novel datasets such as the two benchmark datasets PatternNet and NWPU-RESISC45 ,Aerial Image Dataset (AID).

The above mentioned datasets, unlike the EuroSAT dataset, depend on commercial very-high-resolution and preprocessed photos.The fact of using commercial and preprocessed very-high-resolution image data makes these datasets unsuitable for real-world applications.Furthermore, while these datasets place a strong emphasis on increasing the number of classes covered, they have a low number of images per class.

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation Red Edge	0.705	20
Band 6 - Vegetation Red Edge	0.740	20
Band 7 - Vegetation Red Edge	0.783	20
Band 8 - NIR	0.842	10
Band 8A - Vegetation Red Edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

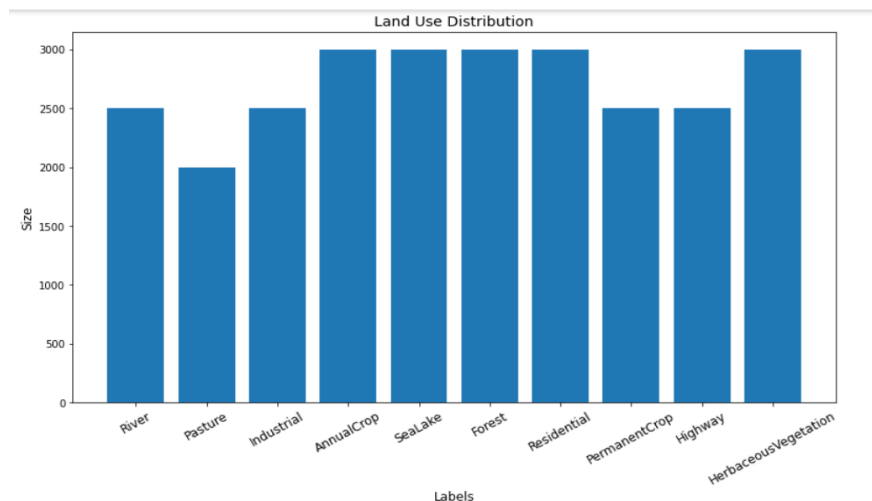
Sentinel-2 Bands, Wavelength, and Resolution



Method

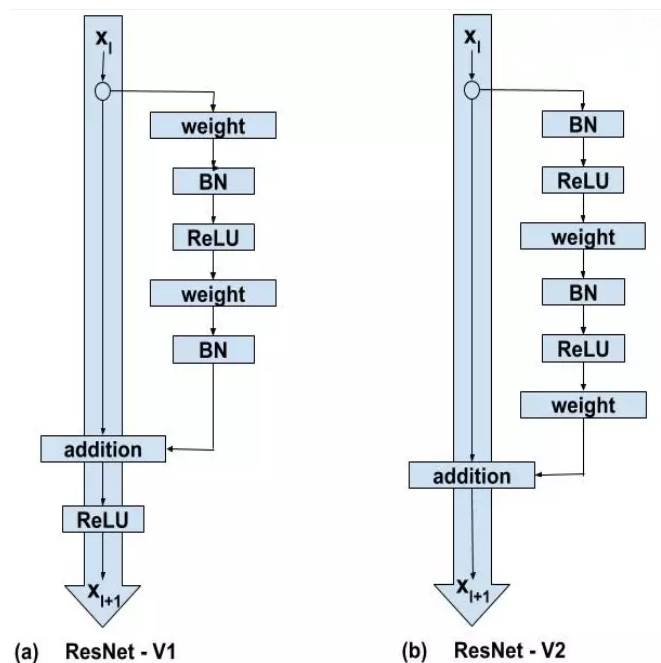
Data pre- processing and Exploratory data analysis

This dataset contains over 27000 images spread through 10 class labels, which were collected by the Sentinel-2A satellite in 64x64 format. Initially only RGB are used out of 13 spectral bands of these hyperspectral images. Basic exploratory analysis is performed on the data. There is not much data pre-processing involved. The images are resized to 64x64. After resizing the images data augmentation is performed to get better results. This dataset of images taken in a specific set of circumstances, But the model, on the other hand, should operate in a number of situations, such as various orientations, locations, scales, and brightness levels. Data augmentation is a technique that allows us to greatly increase the variety of data available by training our neural network with additional synthetically modified data. Data augmentation techniques such as cropping, padding, and horizontal flipping are commonly used techniques.



Model Used

The type of DL architecture implemented in this project is a transfer learning model known as Resnet 50V2. ResNet50V2 is a modified version of ResNet50 that performs better than ResNet50 and ResNet101 on the ImageNet dataset. The core idea of ResNet is based on skip connection, that is it allows to take activation from one layer and feed it to the future layer. Consider the problem of vanishing gradient, when it goes deeper some neurons will not contribute anything because their weights have reduced significantly. But now if activation is brought from the earlier layer and added to the current layer before activation, now it will certainly contribute. In ResNet50V2, a modification was made in the propagation formulation of the connections between blocks, as opposed to ResNetV1 where convolution layer is used first and then batch normalization. Batch Normalization and ReLU activation function with a softmax output classification function were used in this project. ReLU for short is a piecewise linear function, i.e. $y = \max(0, x)$, that will output the input directly if it is positive, otherwise, it will output zero. Batch normalisation is a technique for training extremely deep neural networks that standardises the inputs to each layer for each mini-batch, minimising the number of training epochs needed.



Model training and hyper-parameter optimization

The optimizer used here is Adam. Adam is a replacement optimization algorithm for stochastic gradient descent for training deep learning models. Adam combines the best properties of the AdaGrad and RMSProp algorithms to provide an optimization algorithm that can handle sparse gradients on noisy problems. There are two call-backs used in this project - early stopping and model check point. A callback is a powerful tool to customize the behavior of a Keras model during training, evaluation, or inference. Early stopping is a method that allows you to specify an arbitrarily large number of training epochs and stop training once the model performance stops improving on the validation dataset. ModelCheckpoint callback is used to save the best model observed during training as defined by a chosen performance measure on the validation dataset. After setting up these hyperparameters the model has to be fine-tuned.

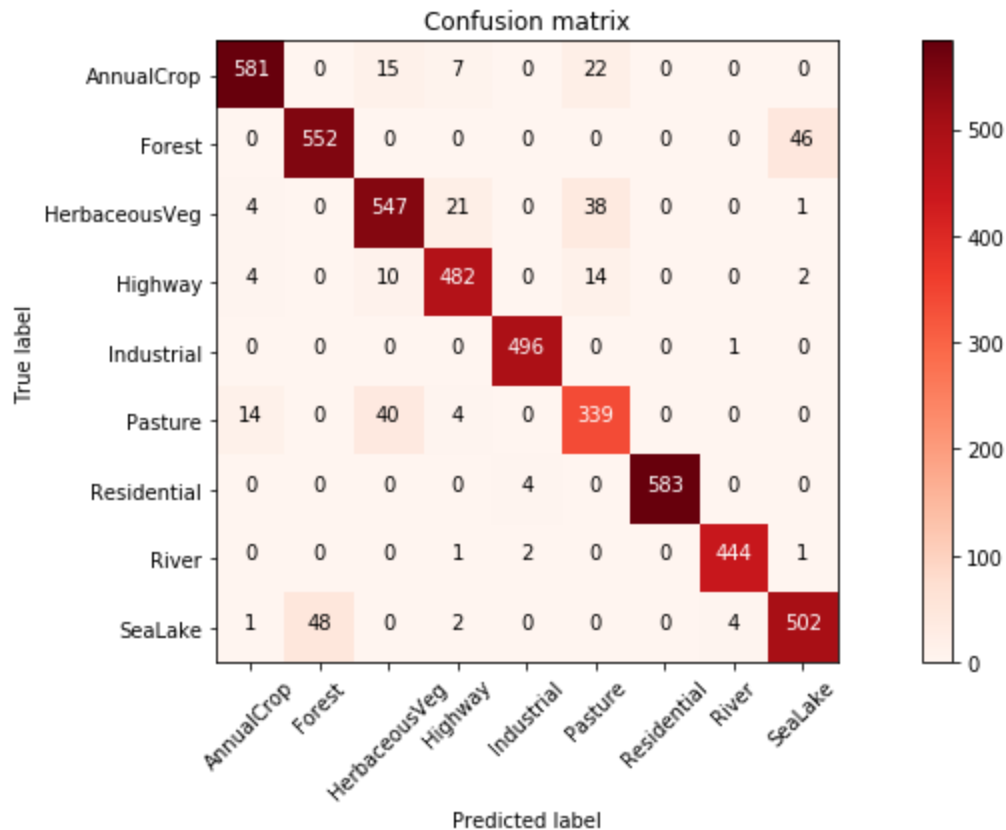
Fine-tuning is a way of applying or utilizing transfer learning. Specifically, fine-tuning is a process that takes a model that has already been trained for one given task and then tunes or tweaks the model to make it perform a second similar task. The model is trained twice, first the dense layers are trained, then the whole network is retrained.

Accuracy assessment

The accuracy of the algorithm was assessed using a number of different metrics.

The highest accuracy obtained for our model is 0.95, which is lower than the accuracy recorded in the original paper using the dataset (0.98). A Classification report is used to measure the quality of predictions of the classification algorithm, i.e. how many predictions are True and how many are False. More specifically, True Positives, False Positives, True negatives and False Negatives are used to predict the metrics of a classification report as shown below.

	Precision	Recall	F-Score	Support
AnnualCrop	0.879154	0.970000	0.922345	600.0
Forest	0.989831	0.973333	0.981513	600.0
HerbaceousVegetation	0.957192	0.931667	0.944257	600.0
Highway	0.946667	0.852000	0.896842	500.0
Industrial	0.954724	0.970000	0.962302	500.0
Pasture	0.937811	0.942500	0.940150	400.0
PermanentCrop	0.917864	0.894000	0.905775	500.0
Residential	0.967690	0.998333	0.982773	600.0
River	0.957746	0.952000	0.954865	500.0
SeaLake	0.986689	0.988333	0.987510	600.0



Forest and SeaLake are two of the classes that the model found difficult to differentiate, as seen in the Confusion matrix. A close examination of the images of these two classes reveals that even the human eye has difficulty distinguishing them.

Conclusion

This model has been trained with a large number of training images, and it can be seen that the model is capable of accurately classifying the land-use distribution. This model is a significant step towards using the vast amount of satellite data available in deep learning. The proposed dataset can be leveraged for multiple real-world Earth observation applications. Possible applications are land use and land cover change detection or the improvement of geographical maps.

You can see the entire notebook [here](#).

References

<https://arxiv.org/pdf/1709.00029.pdf>

<https://www.tandfonline.com/doi/full/10.1080/15481603.2019.1650447>