

# An Overview of Different Techniques to Provide Semantics to Remote Sensing Data

Sumedh Ghavat  
Information Technology VESIT  
Mumbai, India  
2017.sumedh.ghavat@ves.ac.in

Parth Kodnani  
Information Technology VESIT  
Mumbai, India  
2017.parth.kodnani@ves.ac.in

Harshita Singh  
Information Technology VESIT  
Mumbai, India  
2017.harshita.singh@ves.ac.in

Jayashree Hajgude  
Information Technology VESIT  
Mumbai, India  
jayashree.hajgude@ves.ac.in

**Abstract**—Due to the rapid advancements in the remote sensing field, there is an immense amount of data being generated. This data is raw. Hence, it requires semantics and one way of providing these semantics is land use and land cover classification. We present different techniques that can be used for the same. After analyzing the different classification models, it was observed that the machine learning models performed poorly as compared to the deep learning models. VGG19 gave the best accuracy of 97.64%. In order to provide semantics to remote sensing data, the different classification models are essential and this work can be further extended into diverse domains.

**Index Terms**—Remote Sensing, Convolutional Neural Network, Machine Learning, Deep Learning, Transfer Learning

## I. INTRODUCTION

There have been advancements in remote sensing technologies in the past years. Due to these advancements, there has been a stark rise in the amount of remote sensing data. This data has been made freely available under many projects which use satellite images from satellites like Sentinel-2, Landsat 8, etc [3]. Its applications are diverse and it can be used in a lot of industries like land use planning, mining, climate change detection, land use and land cover change detection, etc. Thus, this data is very useful. But these applications of remote sensing are restricted due to multiple reasons. One such reason is lack of semantics to the data because the data available is raw. Raw data cannot be used directly in many applications. One way of providing these semantics is land use and land cover classification. Land use and land cover classification classifies the satellite image to the particular class of land cover that it belongs to. Different land cover classes are residential, forest, industrial, sea, etc. Once the images have semantics associated to them, these various applications can be realised, which would help in solving the grave problems associated within our ecosystem.

As the world attempts to overcome problems regarding climate change, deforestation, global warming, etc., information regarding land use and land cover has become of utmost importance. Our objective is to guarantee utilization of land cover to create preconditions for an ideal living environment and ensure financially, socially and ecologically sustainable development.

This work presents the different machine learning and deep learning techniques that can be used for land use and

land cover classification. The different machine learning models like random forest classifier and support vector machines have been implemented. Pre-trained convolutional neural networks like ResNet, VGG and their variants have also been implemented to solve this problem. The performances of these models are analysed and compared.

## II. RELATED WORK

Patrick Helber, et al. in their work “EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification” have proposed EuroSAT which is a novel dataset of Sentinel-2 satellite images. The dataset comprises of 27,000 images, belonging to 10 different land cover classes. In their work, they have trained CNNs on this dataset, achieving an accuracy of 98.57% [2].

In the work “Classifying land cover from satellite images using time series analytics” by Patrick Schäfer, Dirk Pflugmacher, Patrick Hostert, Ulf Leser, they have used a time-series-based classification approach (WEASEL+MUSE) for land cover classification along with the conventional machine learning approach (Random Forests) [4]. The machine learning model (Random Forests) achieved an F1-score of 89.0% and the WEASEL+MUSE methodology had 89.6% F1-score. But, when the two models were combined together, the F1-score achieved was a little higher, of 91.1%.

Xuan Yang et al. [1] and Xin-Yi Tong et al. [5] in their papers have used datasets which comprise of images from the Chinese GF(GaoFen) satellites. Both the research papers use the same technology i.e. Deep Learning for large scale land use mapping but both have different approaches. In the paper “A Fast And Precise Method For Large-scale Land-use Mapping Based On Deep Learning”, there is optimization of the data tiling method and the structure of DCNN for the multi-channel data and the splicing edge effect, which are unique to remote sensing deep learning, and improve the accuracy of land-use classification [1]. Whereas in paper “Large-Scale Land Cover Classification in GaoFen-2 Satellite Imagery”, initially, the CNN classification is done, followed by segmentation and voting and finally multi-scale classification [5]. The authors Xuan Yang et al have achieved an accuracy of 80.8% for the 3-channel images [1], whereas Xin-Yi Tong et al. have achieved an overall accuracy of 95.71% and an average accuracy of 87.12% [5].



### III. PROPOSED METHODOLOGY

The dataset used is EuroSAT. The different machine learning and deep learning models that are trained, tested and validated using this dataset are :

- Decision Tree
- Random Forest
- Extremely Randomised Trees
- Support Vector Machine
- ResNet50
- ResNet50V2
- ResNet152V2
- VGG16
- VGG19

#### A. About Dataset

EuroSAT consists of 27,000 images spanning over 10 classes of size 64x64 pixels. These images consist of 3 channels - Red, Green, Blue. Each class consists of approximately 2000 to 3000 images. The different land cover classes are visualised in Fig. 1. The dataset is available on <https://github.com/phelber/eurosat>.

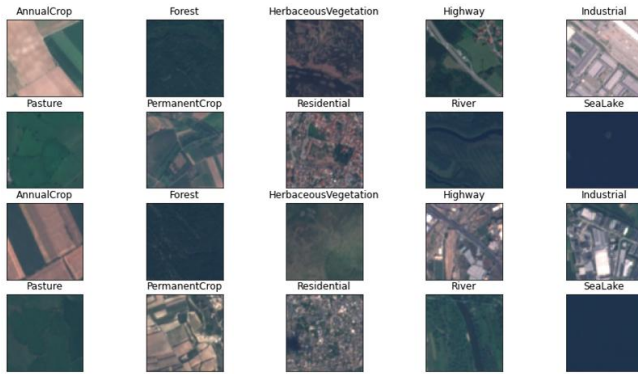


Fig. 1. Visualizations of Different Classes

#### B. Exploratory Data Analysis

Exploratory data analysis is performed on the dataset. Fig. 2 shows the class distribution of the dataset. As seen, the class size varies throughout the dataset, hence stratification will be required to be performed while splitting the dataset. To get deeper insights about the class distribution, a pie chart is plotted. It can be seen in Fig. 3.

#### C. Data Preprocessing

The dataset was split into training and testing sets. An 80-20 split was done. The first step in preprocessing was performed by rescaling the values of the training and testing sets between 0 and 1. Further, augmentation was performed on the training set. Various operations were performed to augment the image data. Horizontal and vertical flips were allowed on the images. A rotation range of 60 was set, which will rotate the pixels on the image between 0 and 60 degrees. The width and height shift range was set to 0.2 which will allow shift upto 20% of the width and height of the image in both horizontal and vertical directions. Image zooming was configured by setting the zoom range to 0.3. This lets the zoom range to be between 80% (zoom in) to 120% (zoom out). Images were allowed to be distorted along both horizontal and vertical axes upto 20%.

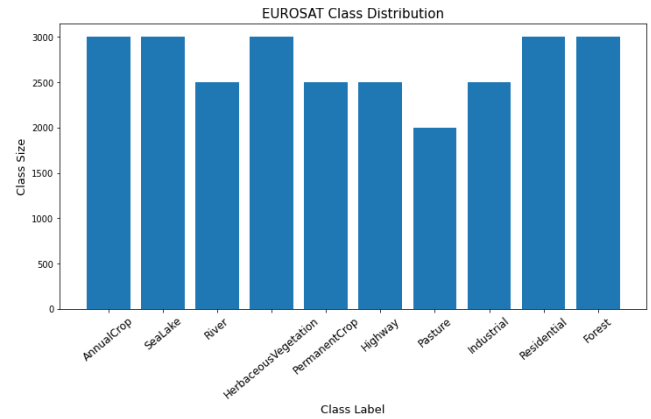


Fig. 2. Plot of Class Distributions of the Dataset

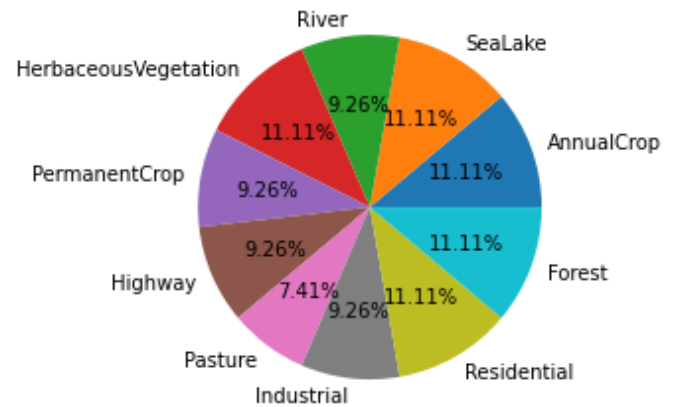


Fig. 3. Pie Chart of Class Distributions of the Dataset

#### D. Applying Machine Learning

This part of the experimental setup includes application of machine learning algorithms on the preprocessed dataset. Algorithms used were:

- Decision Tree
- Random Forest
- Extremely Randomised Trees
- Support Vector Machine

#### E. Applying Deep Learning

We have used transfer learning to solve the various problems we faced while using traditional machine learning algorithms on our dataset. The initial layers of the models are kept in their pre-trained state with their weights frozen, while the final layers are trained to generalize the features of the new dataset. ResNet: ResNet, short for Residual Network, is a very deep artificial neural network which uses skip connections, helps solve the problem of vanishing gradients which in turn makes its deep architecture possible. The variants of ResNet which were used are:

- ResNet50
- ResNet50V2
- ResNet152V2

VGG: VGG, short for Visual Geometry Group, follows a rather simple architecture where images are passed through a stack of convolution layers. It uses smaller filters like 3x3 instead of the larger ones for the convolution operations. The convolution operations are responsible for increasing the

depth and max pooling handles the job of reducing the volume of the input image. The variants of VGG which were used are:

- VGG16
- VGG19

The deep learning models were loaded without their final fully connected layers, thus allowing us to add new final layers including the output layer according to our needs. The input dimensions of the image were specified to customize the models for our dataset. The output layer is of 10 nodes for 10 classes of the dataset. A softmax activation function was used with the output layer which gave the probabilities for the individual classes. For the ResNet variants, one fully connected layer of 2048 nodes was used after the convolution and pooling layers followed by the output layer. A dropout of 0.2 was given to the same to prevent overfitting. And for the VGG variants, two fully connected layers of 2048 nodes were used instead of one. A similar dropout of 0.2 was applied to both of them. The models compiled used an Adam optimizer initialized with a learning rate of 0.01. The models were pre-trained for several epochs and then re-trained end to end after recompilation with a learning rate of 0.0001. Various callbacks that were used in this process include checkpoints to save the weights with the highest accuracy, early stopping to terminate the training process after certain epochs if there were no developments and learning rate reducer to periodically reduce the value of the learning rate.

#### IV. RESULTS & ANALYSIS

The accuracy of the different machine learning and deep learning models are shown in Table. I. Initially, classification was done using the various machine learning models. As seen, the Support Vector Machine algorithm gave the highest accuracy out of all the machine learning models of 66.46%. Comparatively, Decision Tree, Random Forest and Extremely Randomised Trees algorithms did not perform well. The overall performance of the machine learning algorithms was not very good on the extensive satellite imagery dataset.

Deep Learning has been proven effective over traditional machine learning algorithms in various computer vision related tasks. Among the ResNet variants, ResNet50 performed the best and gave an accuracy of 96.77%. Its loss and accuracy graph has been shown in Fig. 4. Whereas, amongst the VGG variants, VGG19 gave the best result of 97.64%. Its loss and accuracy graph has been shown in Fig. 5. Overall, the VGG variants performed better than the ResNets as well as the machine learning models.

TABLE I. ACCURACY DETAILS OF DIFFERENT CLASSIFICATION MODELS

Algorithm	Accuracy
Decision Tree	45.59%
Random Forest	62.32%
Extremely Randomised Trees	62.70%
Support Vector Machine	66.46%
ResNet50	96.77%
ResNet50V2	94.11%
ResNet152V2	95.05%
VGG16	97.48%
VGG19	97.64%

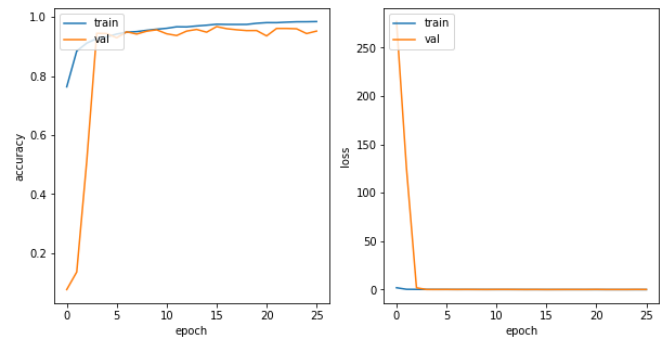


Fig. 4. Accuracy and Loss Graphs for ResNet50

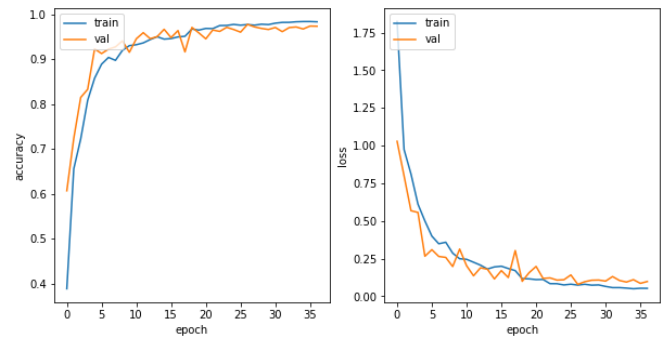


Fig. 5. Accuracy and Loss Graphs for VGG19

#### V. FUTURE SCOPE

The implementation of these state-of-the-art deep Convolutional Neural Networks offer a broad range of possible applications in different domains. These various applications include:

- Urban planning
- Natural resource management
- Biodiversity protection
- Disaster management
- Precise Target detection for military applications

#### VI. CONCLUSION

In this work, we have presented different classification methods that can be used to provide semantics to remote sensing data. The models are trained and tested on the novel dataset EuroSAT. Inferring from the work, machine learning models performed poorly as compared to deep learning models. VGG19 gave the highest accuracy of 97.64%. The different classification models are essential in providing semantics to the remote sensing data. This work can be further extended into diverse domains and many real life applications can benefit from this.

#### REFERENCES

- [1] Xuan Yang, Zhengchao Chen, Baipeng Li, Dailiang Peng, Pan Chen, Bing Zhang. 2019, "A Fast And Precise Method For Large-scale Land- use Mapping Based On Deep Learning", IEEE International Geoscience And Remote Sensing Symposium.
- [2] Patrick Helber, Benjamin Bischke, Andreas Dengel, Damian Borth. 2019, "EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
- [3] Nataliia Kussul, Andrii Shelestov, Mykola Lavreniuk, Igor Butko, Sergii Skakun. 2015, "Deep Learning Approach For Large Scale Land Cover Mapping Based On Remote Sensing Data Fusion", IEEE International Geoscience and Remote Sensing Symposium.

- [4] Patrick Schafer, Dirk Pflugmacher Patrick Hostert, Ulf Leser. 2018, "Classifying land cover from satellite images using time series analytics", EDBT/ICDT Joint Conference.
- [5] Xin-Yi Tong, Qikai Lu, Gui-Song Xia and Liangpei Zhang. 2018, "Large-Scale Land Cover Classification in Gaofen-2 Satellite Imagery", IEEE International Geoscience & Remote Sensing Symposium.