Land - Use Classification

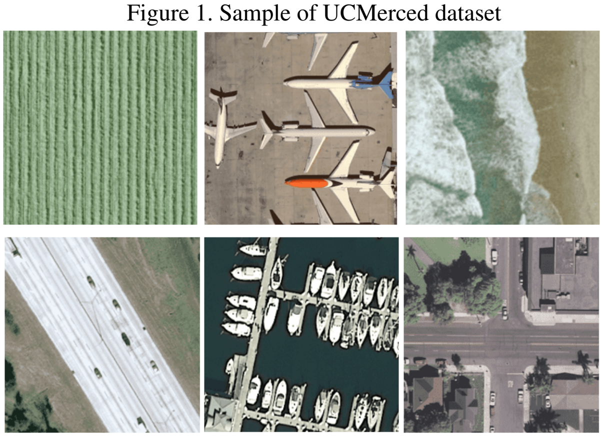
Garima Mehta
University of Minnesota - Twin Cities
mehta250@umn.edu

Shriya Rai
University of Minnesota - Twin Cities
rai00016@umn.edu

1. Introduction

The objective of the project is to classify land by using satellite images for various land-use types. The satellite images are taken from the UC Merced land-use data set, which is labelled data. The data set consists of 21 land-use classes containing a variety of spatial patterns, some with texture and/or color homogeneity and others with heterogeneous presentations. The data set is compiled from a manual selection of 100 images per class, each RGB image being approximately 256 256 pixels. The 21 land-use types include agricultural, airplane, baseball diamond, beach, buildings, chaparral, dense residential, forest, freeway, golf course, harbor, intersection, medium density residential, mobile home park, overpass, parking lot, river, runway, sparse residential, storage tanks, and tennis court classes. The approach used for this problem is implementing a Convolutional Neural Network. The challenge is to optimize the classification accuracy by tuning the hyperparameters, using data augmentation strategies, and novel representations for improved performance. Thus, the problem that our project aims to solve is essentially land use classification using state-of-the-art CNN architectures.

Figure 1. Sample of UCMerced dataset



1.1. Application

Resource planning: Resource planning deals with the allocation and utilization of resources to achieve maximum efficiency. Management of land is largely dependent on natural resources and land segregation for various activities. Soil quality, water availability, biodiversity and population density in an area are key factors for planning of resources. Modern approaches like land - use classification, which can be extended to real time classification not only determine appropriate land - use type but also provide decision makers with sustainable land resource management strategies that improve productivity and sustainability. Land-use classification will assist the decision makers in determining and putting into practice the best land-use management options for sustaining production.

Infrastructure planning: With the increase in population and the advent of modern architecture, infrastructure planning is highly important in developing nations. Land - use classification helps in determining the areas associated with residential (dense or medium), water, agriculture etc. which gives the planners the guidance to plan the various aspects of a city. A well-planned infrastructure strengthens the sustainability and livability of our cities and communities and helps in thriving in a flourishing environment. Land - use classification can also help in planning of factory installations as to minimize the damage caused by hazardous waste emissions from the same.

Disaster Management: Large scale natural disasters cause a great level of damage to property and lives, which leave a financial debt in the economy. There are various disaster managers that tend to mitigation strategies and work towards minimizing the damage done economically by planning effective ways. Land - use classification can help disaster managers in the pre and post planning, this includes the location and temporal information of the land before and after the disaster. Land - use classification can help in building structures in areas where the damage can be minimized or mitigated.

2. Methods and Technique

The workflow for land - use classification is as follows:

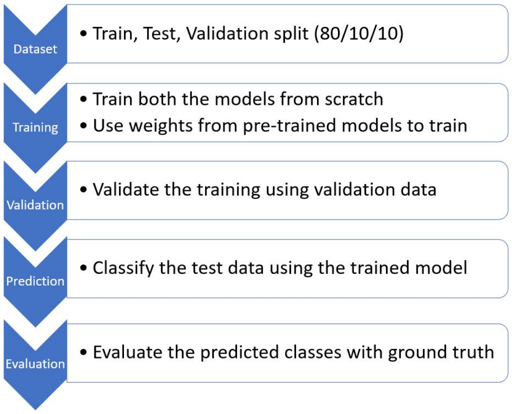


Figure 2. Workflow of the project

The dataset has 2100 satellite images which are divided into 21 classes, each with 100 images. The dataset is pretty clean hence no pre-processing of images was required. Entire data was split in 80/10/10 ratio with 80 training data, 10% validation data, and 10% testing data. The training data was further augmented using the ImageDataGenerator library provided by Keras. The image in the training set is rotated, flipped, and transformed by height and width as data augmentation helps in reducing overfitting and increasing generalization. Then the training and validation data is fed into the system. Using the validation data in the model is an essential thing as it helps in tuning the hyperparameters, or to assess whether the model is overfitting or not. All the models were trained for 300 epochs with a batch size of 64. Once the model is trained, the test data is fed in the system to classify the new images to their respective classes. Once the prediction for each of the classes is completed, the evaluation of the predicted classes is done using the ground truth.

2.1. Baseline Method

The baseline method that is employed in our project is inspired from the paper "Land Use Classification in Remote Sensing Images by Convolutional Neural Networks"[1]. The authors in the paper use GoogLeNet and CaffeNet for classifying these images. The baseline for our project was to use GoogLeNet architecture which gave an accuracy of 71.1% when model was trained from scratch and 89.5% when transfer learning with weights from Imagenet were used. GoogLeNet is a convolution network which utilises the concept of inception modules. These modules reduce the complexity of the 3-D filters of conventional architectures by means of a prior depth reduction phase. Due to reduced complexity, multiple filters can be used in parallel at different resolutions, To improve the effectiveness of the gradient backpropagation, given the depth of the network,GoogLeNet also employs auxiliary classifiers connected to intermediate layers. Thus, the advantages of this architecture are:

1. Different sized filters can be used at each layer therefore retaining spatial information.

2. Significantly reduces the number of free parameters of the network, making it less prone to overfitting and allowing it to be deeper.

The number of training parameters in GoogLeNet is very less compared to other Convolutional Neural Network architectures. For AlexNet, the number of training parameters is 60 million which has reduced to 4 million. The hypothesis that we aim to prove is that if we increase the number of training parameters, the model will learn more features from the images and thus will provide better results/accuracies.

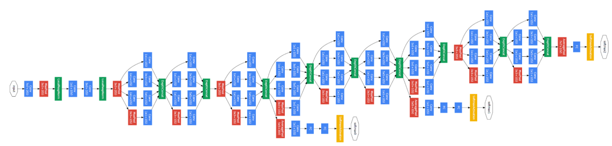


Figure 3. GoogLeNet Architecture

2.2. Proposed Method

The initial proposal for optimizing the classification accuracy is to use VGG16. The reason for using VGG16 is because it has 16 layers and higher trainable parameters than GoogLeNet, which is lesser than in GoogLeNet architecture. The method we used for training VGG16 was twofold:

1. Training the network from scratch using the images from the dataset.

2. Training the network using Transfer Learning (Imagenet weights) for the network.

The representation depth is beneficial for the classification accuracy, and that state-of-the-art performance on the ImageNet challenge dataset was the motivation to use VGG16 for this dataset as well.

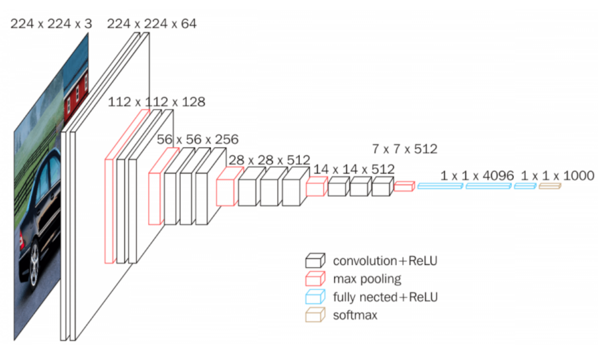


Figure 4. VGG16 Architecture

After the initial results, we planned to implement various other CNN based model architectures for obtaining accuracy that surpasses GoogLeNet architecture.The models that we tried were:

1. AlexNet
2. Xception
3. VGG19
4. InceptionResNetV2
5. MobileNet
6. MobileNetV2
7. DenseNet

3. Result and Analysis

The results from VGG16 were unsatisfactory and the accuracy for training the model from scratch was 55.6%. Though there were a large number of training parameters, the accuracy was not surpassing GoogLeNet and the reason was that the model was overfitting. The number of images in the training dataset was 1680 (80% of the entire dataset, 2100 images) and with the depth of VGG16, the model was prone to overfitting. The results from VGG16 when using a pretrained model (Imagenet weights) were surpassing the accuracy of GoogLeNet by 0.5%

VGG16

Model	Accuracy
From Scratch	55.6%
Using Transfer Learning	90.0%

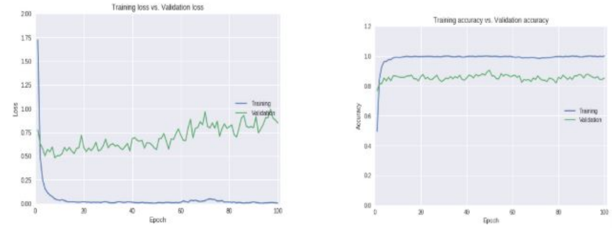
GoogLeNet

Model	Accuracy
From Scratch	71.42%
Using Transfer Learning	89.5%

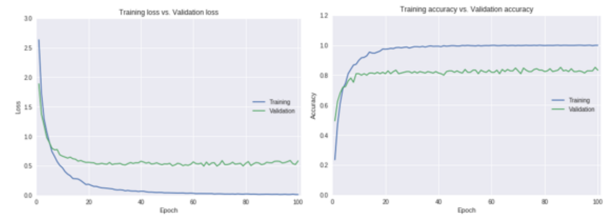
Figure 5. Comparison of VGG16 and GoogLeNet

Taking further steps, we started with implementing AlexNet from scratch and the model was again overfitting as it also has a large number of training parameters. The accuracy for this model was 57%, which bolstered our opinion of the model with a large number of parameters and depth to be prone to overfitting with smaller datasets. Hence, we used the other models with transfer learning only and below are the results for it.

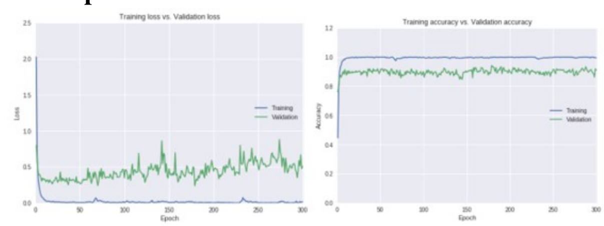
1. Xception - 82.1%



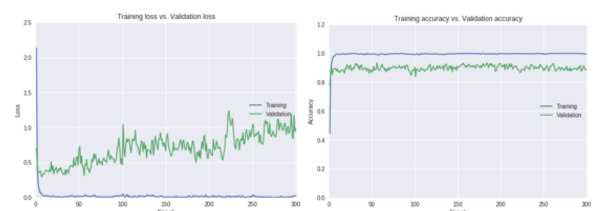
2. VGG19 - 81.9%



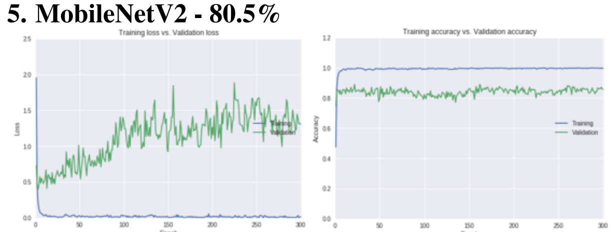
3. InceptionResNet - 91.9%



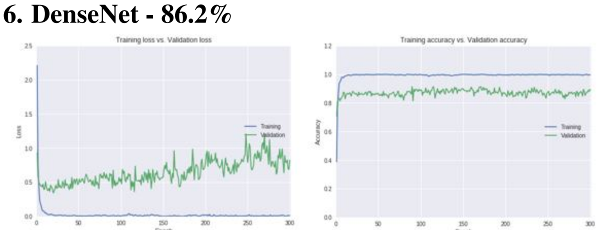
4. MobileNet - 88.1%



5. MobileNetV2 - 80.5%



6. DenseNet - 86.2%



4. Conclusion

As we see from the above table and the results for GoogLeNet and VGG16 model architectures, for the UC-Merced Land - Use dataset, inception modules was the highlight and the models with lesser number of training parameters and smaller depths performed better than the others. GoogLeNet, InceptionResNet and MobileNet performed better than the other models with relatively deeper networks and more trainable parameters. j