Deep Learning for Satellite Imaging

By

Hamza Abubakar
17BCE028

INSTITUTE OF TECHNOLOGY
NAAC ACCREDITED 'A' GRADE

DEPARTMENT OF COMPUTER ENGINEERING
Ahmedabad 382481

Deep Learning for Satellite Imaging

Seminar

Submitted in fulfillment of the requirements

For the degree of

Bachelor of Technology in Computer Engineering

By

Hamza Abubakar
17BCE028

Guided By
Prof. Neha Sisodiya
DEPARTMENT OF COMPUTER ENGINEERING

NIRMA UNIVERSITY
INSTITUTE OF TECHNOLOGY
NAAC ACCREDITED 'A' GRADE

DEPARTMENT OF COMPUTER ENGINEERING
Ahmedabad 382481

CERTIFICATE

This is to certify that the Seminar entitled "Deep Learning for Satellite Imaging" submitted by Hamza Abubakar (17BCE028), towards the partial fulfillment of the requirements for the degree of Bachelor of Technology in Computer Engineering of Nirma University is the record of work carried out by him/her under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination.

Prof. Neha Sisodiya
Assistant Professor
Department of Computer Engineering,
Institute of Technology,
Nirma University,
Ahmedabad

Dr. Madhuri Bhavsar
Dept. of Computer Engineering,
Institute of Technology,
Nirma University,
Ahmedabad

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to all those who provided me the responsibility to complete this Report. I acknowledge with thanks, the support rendered by Prof. Neha Sisodiya, under whose guidance I was able to complete the given task in the specified amount of time. I also appreciate the constructive suggestions given by my friends and classmates to further enhance the content of the Report.

At the home front, I am extremely grateful to all my Family Members and Relatives for the continuous support and encouragement they displayed during the course of completion of the Report.

ABSTRACT

In this Report, we discuss the Concept of Deep Leaning for Satellite Imaging. We begin by the History of the Technology followed by its definition, and a proper insight into the same. We discuss the various Technical Aspects for this Technology, along with the current work and future scopes in the field.

CONTENTS

INTRODUCTION

1.1 General

Deep Learning is subset of Machine Learning which is concerned with Algorithms which are inspired by the function and structure of the Brain, called **Artificial Neural Networks.**

1.2 Scope of Study

This Report is majorly focused on the definitions of Deep Learning, its supersets, related Terminologies and more. Also, it is focused on Satellite Imaging, Basic Concepts and Terminologies, use of the same for Deep Learning Applications, primarily focusing on Economic Housing as an Application.

1.3 Motivation

Over the last few Years, there has been a drastic change in Computer Technology, with various of them including Artificial Intelligence coming into play. Not only have these fascinated quite a lot of people, but these have also changed the way different things are. Ranging from small Factories to sending Rockets into outer space, everything is covered. In fact, Deep Learning Technology has helped ease a lot of day to day activities. Deep Learning in Satellite Imagery has really helped ease a lot of things. These include Poverty, Agriculture, security and defense, economic housing and what not.

OVERVIEW

1.1 Deep Learning

The term “Deep Learning” has been something that has been floating around for a while now. It is basically a much more specified type of Machine Learning, and in broader terms, that of Machine Learning.

In simple words., it is a Technology that basically takes in and analyses the data similar to how a human would. In fact, the entire concept of Deep Learning as been revolving around the Neural Network of the Human Brain.

1.2 What differentiates DL from ML?

Both these terms have been afloat in the Technological World since long, but it is a necessity to know how one differentiates from the other.

One thing to note here is that Deep Learning is actually Machine Learning.

1.2.1 Dependencies on Data

This is one of the fundamental things that differentiate between the two. Smaller Data Sets do not lead to Deep Learning Algorithms performing that efficiently. This is mainly due to the fact that Deep Learning Algorithms feed on a large amount of data to understand it perfectly. While on the other hand, Machine Leaning Algorithms perform in the opposite manner. These Algorithms with their handcrafted rules perform well when there is a smaller amount of data.

1.2.2 Dependencies on Hardware

Due to a greater complexity in their Algorithms, Deep Learning requires high-end Machines in contrast to Machine Learning which can adjust with the low-end devices. In fact, Deep

Learning Algorithms require high end GPUs which are an integral part of their processes. Deep Learning Algorithms inherently perform a large amount of matrix multiplication operations to get done with the work. Evidently, GPUs built for this purpose find an active usage here.

1.2.3 Feature Engineering

Basically, Feature Engineering is the process of putting the domain knowledge to create the feature extractors and in order to reduce the complexity of the data and make the patterns visible to make the learning algorithms work.

Coming to the differences, Machine Learning require the features to be identified by an expert and manually hand coded.

On the other hand, Deep Learning Algorithms automatically try to learn high-level features from the data. These put this technique a step ahead of their predecessors. As a result, the task of producing new features manually is taken care of themselves. Low-Level Features like points, edges and lines are detected by convolutional Neural Networks which then unify for High-End Features.

1.2.4 Approach to Problem Solving

Upon using a Traditional Machine Learning Algorithms, the problem is bifurcated into individual parts and then each one is solved individually, part by part.

Deep Learning, on the other Hand, involves an end to end process.

1.2.5 Execution Time

Deep Learning Algorithms require a large number of features that are to be extracted, and as a result these times a longer time to train. One particular type requires 2 weeks to train, when done from scratch. Machine Learning Algorithms require a much lesser time, which range from some seconds to a few hours.

When it comes to the testing time, the roles seem to have completely reversed. Deep Learning Algorithms, in fact take much lesser time to run, and Machine Learning Algorithms, take more time on increasing the data, for instance: K- nearest Neighbors.

1.2.6 Interpretability

Lastly, we have this as a factor which makes one think twice before applications in the Industry.

Machine Learning Algorithms run on a set of rules that reveal the criteria for the decision making performed. On the other hand, Deep Learning Algorithms the crisp rules that these use for the performance.

1.3 Satellite Imagery

These can be defined as the images that are taken from the outer space of earth of different other planets or celestial bodies which are captured by orbiting satellites that belong to the government and different businesses around the world.

1.3.1 Applications

These find various application that range from fishing, agriculture, regional planning to education, warfare and intelligence.

1.3.2 Data Characteristics

Satellite imagery in remote sensing boils down to four major categories, which are discussed below:

Temporal Resolution is the time passed between imagery collection locations for a given geographical area

Radiometric Resolution is achieved by the sensitivity of the image to the magnitude of the electromagnetic energy present

Spatial Resolution involves the number of pixels that go in the construction of an image

Spectral Resolution is the bifurcation of the different features and bands received into their respective categories

AUTOENCODERS

2.1 Definition

These are an unsupervised way of where the technique learns the representation for a particular set of data.

In simple terms, the aim of the Autoencoder is to initially take the data and compress it into a short code, followed by the uncompressing of the code to get something that closely matches the original data given.

This helps in Dimensionality Reduction of the data, for instance learning to ignore noise in the data.

We can say that it is a simple bottleneck representation of the original data given as input, a compressed knowledge representation.

Characteristics of an Ideal Autoencoder

- It is sensitive enough to the inputs so as to build a reconstruction.
- For the data that overfits or doesn't simply memorize, it is insensitive enough.

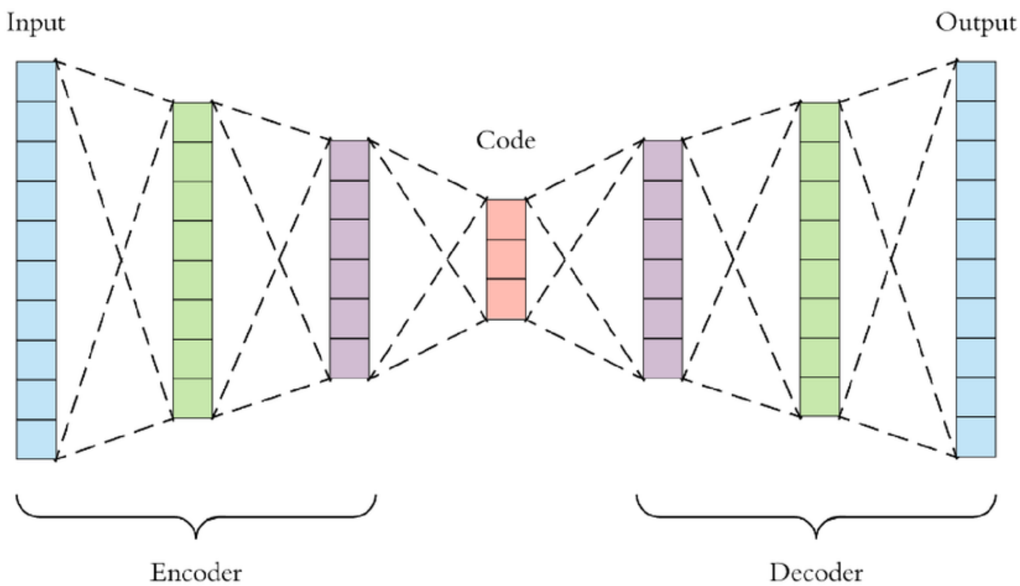
2.2 Process

The working of an Autoencoder is explained in simple terms below:

Initially, the technique will look for basic level features like points, corners etc. Advancing further, it will go for less basic features like lines and will go on to for the entire data.

The final layer will encode into a code that closely matches the original input.
Then comes the task of the Decoding Layer: It will construct the original data from the compressed code.

2.3 Architecture



1

When the input is given to the Autoencoder, it passes through the encoder, which is actually a fully connected ANN, which produces the code. This code then passes through the Decoder layer which does the work with the help of a similar ANN.
Generally, it is that the architecture of the Encoder and the Decoder are both the same. One important thing to note is that

there has to be the same dimensionality for both the input and the output.

Before training an Autoencoder, there are four hyperparameters that have to be taken care of, as given below:

- The number of codes in the middle layer, a short code results in greater compression.
- The number of layers in the Autoencoder, greater the number of layers, deeper is the autoencoder.
- The number of nodes that are present in each layer. Stacked Encoders usually look like a "Sandwich". There is a fact that there are increasing layers into the Encoder and decreasing Layers into the Decoder.
- The Loss Function

2.4 Types of Autoencoders

There are generally four types of Autoencoders, all of them discussed below:

2.4.1 Vanilla Autoencoder

This is the simplest form of the Autoencoder, having three layers, with a neural net and a hidden layer. Here, both the input and the output are the same.

2.4.2 Multilayer Encoder

In this type, we use 3 hidden layers in place of just one.

2.4.3 Convolutional Autoencoder

This type includes the usage of 3 Dimensional images instead of the 1-Dimensional inputs. The image that is given as the input is down sampled giving a compressed version of the images.

2.4.4 Regularized Autoencoders

While other Autoencoders are just limited to having the encoders and decoders shallow and a small code size, these allow to have more properties like copying its input to the output.

2.4.4.1 Sparse Autoencoder

Apart from identification, sparse Autoencoders help classify the input given. It has to also work on the classification features of the data set. Here, there a penalty added to the Loss Function, making the representation sparser as compared to the earlier ones.

2.4.4.2 Denoising Autoencoder

Here the reconstruction error term of the loss function is changed, making the Autoencoder learn this way. This is achieved by adding some noise of the input image and making the autoencoder learn it. This eventually results in a more robust representation of the data.

2.5 Need for Autoencoders

In an Autoencoder, it is not the conversion of the input to the output that proves to be useful, rather it is the latent representation *h* that are a matter of importance. By this process, the ANN is trained for the new data to come in.

CONVOLUTIONAL NEURAL NETWORK(CNN)

A CNN or ConvNet is a subset of Deep learning which is usually applied to Visual Imagery.

In a CNN, the input image is taken, the various features and objects are assigned weights and are differentiated from one another. Pre processing in a CNN is a much lesser requirement when compared to the traditional Algorithms.

While other Algorithms are found to be hand engineered, these have the ability to learn on their own with enough training.

In broader terms, the role of a CNN is to reduce the taken images into a form which are comparatively easier to process, whilst retaining the features are a necessity for obtaining a goof prediction.

3.1 Architecture

The Architecture of the CNN is simply based on that of the Neurons of the brain and was inspired by the structure of the visual cortex. As known. Individual Neurons are known to detect only a specified region of the visual fields. A collection of the many neurons constitutes the entire visual area.

3.2 Building Blocks

The Architecture of a CNN consists of a stack of different layers that transform the input into the output through a differentiable function.

3.2.1 Convolutional Layer

This is the main layer in a CNN. These consist of a set of learnable kernels, which include a small receptive field, but go on through the entire depth of the input volume. When going forward, every layer is convolved, and an activation map of the that filter in 1 Dimension is produced.

In the general sense, every entry that comes out as output can be visualized as of the neuron which itself is limited but then constructs the entire image when combined with others in the activation map.

3.2.2 Pooling Layer

Once the representation is obtained, the job of the pooling layer comes into play. It reduces the number of parameters and computation in the network. The pooling layer operates on each feature map independently. The most commonly used one is the max pooling layer.

3.2.3 Fully Connected Layer

These layers do the job of classifying the data into various classes. These connect every individual neuron present in one layer to every neuron in the other one.

Similar to the Multi-Layer Perception Network, the flattened matrix goes through a fully connected layer, thus resulting in a classification on the Images.

3.2.4 ReLu Layer

This is a short form of the Rectified Linear Unit, effectively reducing the negative values from the activation map matrix my converting them to zero. Without affecting the receptive fields of the Convolution Layer, it increases the non-Linear functions.

3.2.5 Loss Layer

The Loss Layer does he job of specifying how the training penalizes the deviation between the actual and the predicted output of the data.

DEEP LEARNING FOR SATELLITE IMAGERY

In the recent years, there has been a tremendous advancement in the Technology, the same in fields like Artificial Intelligence and Satellite Imagery have helped solve many different problems in addition to proving unlimited applications.

When combining these two, there a lot od issues that can be addressed all over the world. One can pinpoint any desired location on the earth and get the satellite images with ease. Not only this, one can even obtain past images of some area and compare it with the present ones to obtain the information as to how that area has changed over time. What makes this even more interesting is the fact that, we can apply DL Algorithms to all these data (Particularly Images) and predict the future outcomes.

4.1 Architecture

Initially, the data is obtained from a trusted data source, which can be either a government handled or a private source.

Next, there has to be preprocessing; large resources are required, hence one can either develop their own solutions in clouds or available platforms or can go for large resources.

An API can be used to get the Pre-Processing images when required.

When the prepared images are finally available, one can train the network and save the model. This is followed by running a batch of API and label the images to store them.
Once the images are prepared, the dashboard common to all can be prepared and use it or an API to request the image, run the model and eventually present the results.

4.2 Business Applications

There are quite a few Business Applications which prove to be quite profitable when combined with this Technology.

Agriculture

Farmers can use the Technology to keep an eye over their crops, predict natural calamities and weather that would prove to be an important factor in the determination of the crop yield.

Real Estate

The Estate companies can look for the Housing Prices for a particular geographical region.

Finance and Insurance

There can be a forecast of the supply of goods, which can be estimated on the basis of the containers delivering them

HOUSE PRICE PREDICTION BY DEEP LEARNING FOR SATELLITE IMAGERY

Satellite Images have been surfacing for a while now. There are number of Applications and problems that have been taken care of using this technology. One such Application is the Prediction of House Prices using this Technology.

While there are numerous high-quality images that are easily available through satellites which can be related to the prices of houses all around the world, there are very unstructured and can not be used to extract meaningful data from.

Adding to this Technology, Deep Learning can be incorporated for a smarter house prediction. Still, there has to be abundant data to train these models, else it becomes altogether a lot more challenging.

4.1 Process

It is started with a trained CNN along with a large classification data-set which will help detect low level image features such as corners and edges. The knowledge gained from this model is trained and is made ready to perform a new task.

To accomplish this objective, a dataset of house prices from a desired geographical region is combined with a dataset of satellite images of the same. The input would include a feature vector along with the image of the same.

The output which is obtained from the CNN is just a number, which is the price o the house in Dollars.

After pretraining, the last two layers of the CNN are deleted and the feature vector is obtained for the image. A number of fully connected layers are then combined together to get the feature vector for the Image.

4.2 Dataset and Features

4.2.1 Real Estate Properties

The Real Estate assessment data is obtained from various public services. These include the number of floors, area, the number of bedrooms and more.

4.2.2 High Resolution Satellite Images

Along with the above features that are used with the Models, the Satellite Images are obtained from the various public platforms that have been captured by orbiting satellites.

4.2.3 Preprocessing of Data

Before the data is sent to the Model to be trained, there are a number of preprocessing that has to be taken care of. Additional unwanted features such as ID of the person have to be omitted from the data set, so as to efficiently and rapidly train the data.

4.3 Training of the Deep Convolutional Network

Once the Features and the Input Images are obtained, it is time to train the Model and obtain the calculated results.

CONCLUSION

All through out the Report, we have seen the Technology that can affect and transform millions of lives all around the world. Deep Learning when combined with Satellite Imagery comes handy in a lot of Applications like Poverty, Agriculture and more. These will prove to be helpful for applications like protection against natural calamities, in addition to proving to be profitable for a number of businesses all around the world.

REFERENCES

Information on Satellite Imagery
https://en.wikipedia.org/wiki/Satellite_imagery
Autoencoders and its types
https://www.jeremyjordan.me/autoencoders/
Guide to Data Networks [1]
https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
Applying Deep Learning
https://towardsdatascience.com/applied-deep-learning-part-3-autoencoders-1c083af4d798

Sina Jandhagi & Hoormazd Rezai: House Price Prediction using Satellite Imagery, Stanford
Arcith, Swati, Raghu, Madhukar: Beyond Spatial Regressive Models, Watson Research Center