

Deep Learning-based Object Recognition for Wide-Area Surveillance

Student(s)

Furkan Güner
Gürkan Soykan
Berke Işın Can

Faculty Member(s)

Müjdat Çetin
Ertunç Erdil

Company Advisor(s)

Fatih Kahraman
Gökhan Özbülak
Said Aydemir



ABSTRACT



Figure 1. Aerial image – Object Recognition

Wide-area surveillance, in particular detection and recognition of objects/structures in wide-areas is a very important issue in terms of security and planning. Wide-areas must be scanned and the structures that they have must be specified in order to evaluate the cost of damage in case of natural disasters such as earthquakes, landslides, avalanches, etc. With the help of such information, we can identify the most damaged areas which require immediate help therefore we can act faster in case of emergency.

Considering the amount of lands that need to be specified and tagged, human power is dramatically inadequate to complete this task. For this reason, we need computer tools that can help us solve this problem. This project focuses on deep learning techniques for solving wide-area object recognition problems.

OBJECTIVES

The main goal of the project was developing a generic object recognition model that automatically learns through examples based on special objects entered by experts via a web-based annotation tools.

The goal is to use Deep Learning models designed for object recognition. It is then intended to utilize transfer learning and special topologies for modeling and probing. One of the main goals is to establish a cross check system by finding similarities of marked objects from users.

OBJECT RECOGNITION

Visual	Result	Comments
	Guard	False
	chimpanzee, chimp, Pan, troglodytes	False
	Siamese cat, Siamese	True

Figure 2. A simple example of successful object recognition. Those images were used in Keras (deep learning library) and we put a small fraction of results here.

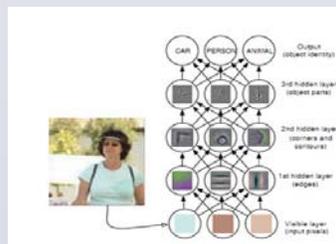


Figure 3. Object Recognition Steps.

Object recognition is a process for identifying a specific object in a digital image or video. Object recognition algorithms rely on matching, learning, or pattern recognition algorithms using appearance-based or feature-based techniques.

Common techniques include edges, gradients, Histogram of Oriented Gradients (HOG), Haar wavelets and linear binary patterns.

In aerial imagery, we can see many different objects with various appearance. Thus, we should get multiple features from images and recognize objects by using machine learning / classification methods acting on their features.

LEARNING TYPES

Machine Learning is an application of artificial intelligence (AI) that provides the ability to automatically learn and improve from experience without being explicitly programmed. In this project, the aim was to model a generic classification problem, namely object recognition for wide-areas. One can consider several different learning types, including:

- 1) Supervised Learning : It means that learning algorithm benefits from a labelled dataset which is called training dataset in order to learn a classifier.
- 2) Unsupervised Learning : Data are provided with no a priori knowledge of the hidden patterns that they contain.
- 3) Semi- Supervised Learning : Knowledge is provided from one data collection in order to analyze related data collections.
- 4) Reinforcement Learning : It requires learning by interacting with environment. The algorithm receives rewards by performing correctly and penalties for performing incorrectly.

Deep Learning is the cluster of artificial intelligence functions or models that imitates the working principles of the human brain in processing data and creating patterns for use in decision making. It is a subset of machine learning in Artificial Intelligence (AI) that has networks which are capable of learning from data that is unstructured or potentially unlabeled. Deep Learning algorithms can learn the representative and discriminative features in a hierarchical manner from data. An algorithm is considered to be deep if the input data are passed through a series of nonlinearities or nonlinear transformations before they reach the output. These transformations mostly occur in the neural networks.

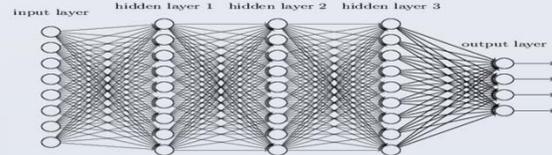


Figure 4. Neural Network Connections.

Neural Networks are the series of algorithms that attempt to identify underlying relationships in a set of data by using a process that mimics the way the human brain operates. Neural networks have the ability to adapt to changing input so the network produces the best possible result .

ONGOING & PLANNED WORK ON THE PROJECT

Currently, code / models and libraries such as Keras, TensorFlow, Torch etc. presented in open source for object detection / identification have been examined. Thereafter, studies have focused on Theano, TensorFlow, Caffe, Keras. First trials were on a machine which does not have GPU support for these frameworks. Subsequently, the team started to work on a machine having GPU support. However, currently there are some problems with installation of the drivers and software.

In order to be working with these frameworks we dedicated one of our computers as a workspace based on Ubuntu and another machine which is dedicated to the team by BİLGEM.

We have worked on a pretrained model named VGG-16 with TensorFlow and Keras. The model is tested with various test images. Based on the results the team prepared a report regarding to the success rate of the model.

The tasks which are needed to be completed before the project's final report are as follows: Establishing the conceptual model on Enterprise Architect (EA) and obtaining the approval of the Project Consultant in BİLGEM. After that, preparation of an Architectural Design Description Document describing the basic architectural components of the system is needed. Then, encoding of the specified method in the relevant development environment will be performed. Following that, the team is going to prepare a report of classification achievements on given images. Then, labeling satellite images with the objective to be informed later for the specified object is going to be done. Finally, following the labeling process, detection will be performed for newly arrived images with a network that has been used in the literature and existing models (which will not require training from scratch).

REFERENCES

- DetectNet: Deep Neural Network for Object Detection in DIGITS
Bengio, Y., Goodfellow, I. J., & Courville, A. (2015). Deep learning. *Nature*, 521, 436-444.
- Gleason, J., Nefian, A. V., Bouyssounousse, X., Fong, T., & Bebis, G. (2011, May). Vehicle detection from aerial imagery. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on* (pp. 2065-2070). IEEE.