Vehicle detection from very high resolution imagery using eCognition's deep learning tool set

Ellie Robertson
Airbus Defence and Space
6 Dominus Way, Meridian Business Park, Leicester, LE19 1RP, United Kingdom
E Ellie.Robertson@airbus.com W: www.intelligence-airbusds.com

Automated feature extraction using a convolutional neural network is a quick and easy way to enable statistical and visual analysis of satellite imagery. The deep learning can show visually where the vehicles are located and quantitatively how many are present within the imagery.

The convolutional model needs to be able to detect and extract the features of interest. These features need to be detected in variety of weather and climate conditions on the satellite imagery to enable a robust application.

The deep learning model has example samples from the satellite imagery inputted into it which tell it what is a vehicle, but also a background set of samples to clarify what is not a vehicle. These are included in the eCognition ruleset which then takes image chips of the samples for the deep learning model to use when conducting the feature extraction. The model is comprised of three layers, two of which are hidden layers with pooling and trained on the scenes with known samples. This model can then be applied to unknown imagery scenes to test whether vehicle detection is taking place. This convolutional model is particularly good at vehicle detection but there is confusion when specific vehicle types are categorised.

Future works includes the model being scaled up in size to enable a fully functioning robust application, there needs to be an increase in the number of image chip samples to improve accuracy and reliability, this will help with specific vehicle type identification. One way in which this can be done quickly is by rotations and flipping the image chip samples already collected.

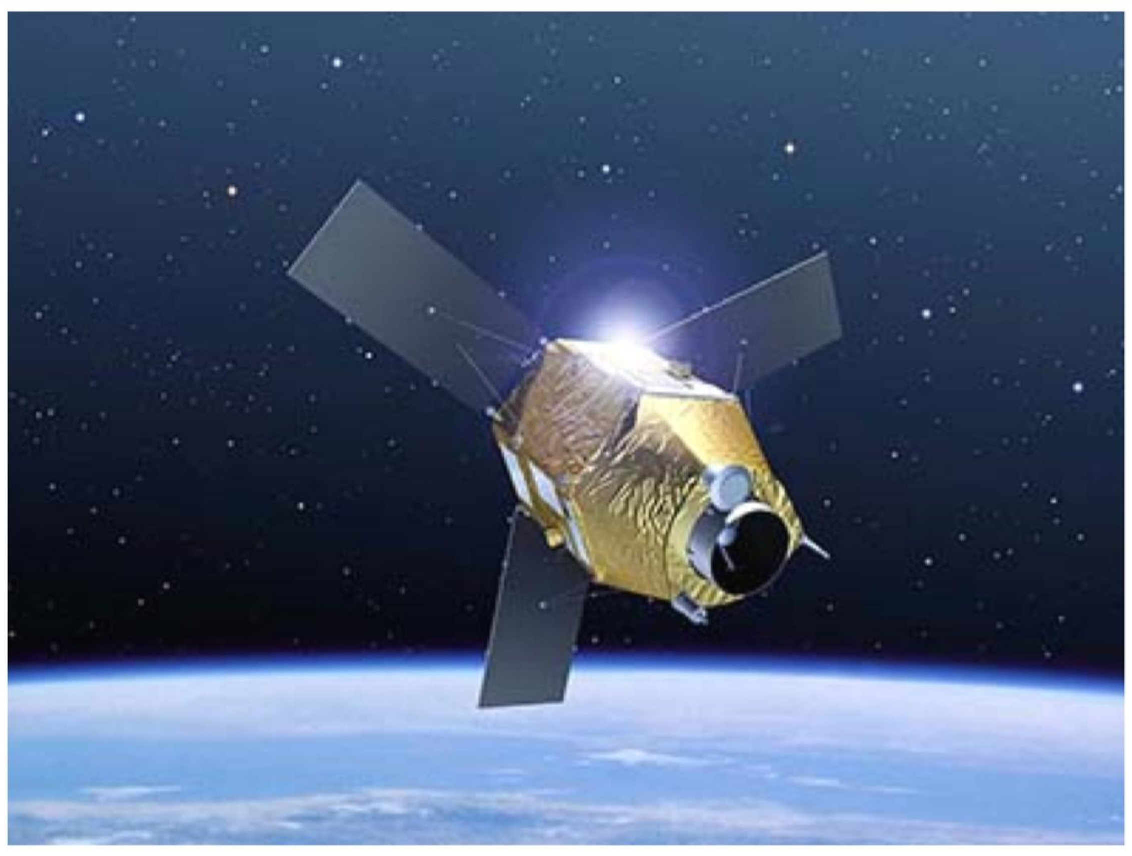

Figure 1: Pleiades 1A sensor

Introduction

A convolutional neural network is a deep learning technique which uses several layers of neural networks. This method is generally used for image recognition therefore it is useful for visual and spatial data such as satellite imagery, Tamta, K. et al (2015). Vehicle extraction is used here and from this statistics are taken from the results to allow analysis, such as concentration and movement of vehicles.

The imagery used in the convolutional neural network is Pleiades imagery which is collected by the Airbus Defence and Space Pleiades 1A and 1B sensors. The imagery has a high resolution of 0.5m containing the RGB and Infrared bands, Airbus Defence and space (2018)..

Objective

The aim of the feature extraction tool is to detect vehicles automatically through a deep learning algorithm to create a commercial product. Statistics about volume and distribution of the vehicles will be able to be extracted allowing patterns and future predictions to be made.

Figure 2: Examples of manually picking samples of trucks and cars to import into the deep learning model.

Method

1. Samples of vehicles split into the following categories were collected manually.
   a. Cars
   b. Trucks
   c. Background
2. The samples needed to be collected over a variety of weather and climate conditions and various locations. This data was recorded in a shapefile within ArcGIS.
3. The sample shapefiles are imported into eCognition and a chessboard segmentation over these samples is performed. A chessboard segmentation was chosen due to its simplicity of splitting the samples into equal-sized objects. All three samples are segmented in this way to give them all an equal chance of being picked up by the model.
4. 'Generate labelled samples' function within eCognition is used to generate image chips to be imported into the model for training, Trimble (2018). In this case a sample count of 2000 with a sample patch of 22 pixels.
5. The deep learning model is created, comprising of three layers with two hidden layers. The first hidden layer has a kernel size of 13 and 40 feature maps. The second hidden layer has a kernel size of 4 with 12 feature maps. Both layers have max pooling, Trimble (2018).
6. The model was trained on 25 scenes in groups of five at five different locations. This was done to allow varied weather and climate conditions.
7. Finally the model is tested blind on scenes it was not trained upon to test how well it extracted the categories of vehicles.

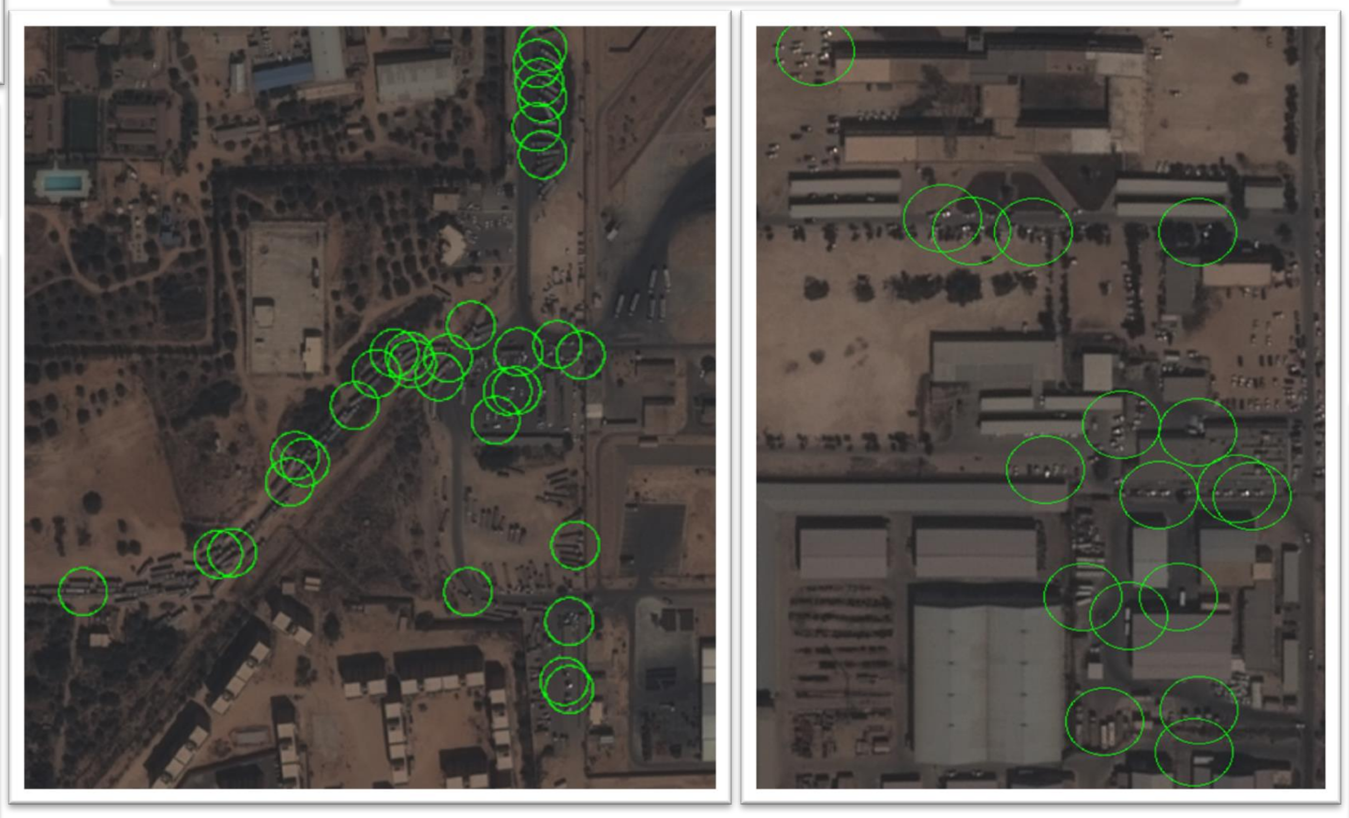

Figure 3: Some trucks are correctly extracted by the deep learning algorithm, however there are some errors with the algorithm extracting cars. This a confusion due to the algorithm thinking a car is the cab of the truck. The circle represent the trucks being extracted and counted.

Results

The neural network classifies trucks and cars easily due to them being quite distinct in shape and size. However there are some errors present within the final results, which seemed to depend on the climate of the area and how bright the imagery is. One apparent error present within the results is that when a truck cab is left without a trailer component it can be misidentified as a car. There are ways in which this can be greatly improved.

Discussion

Overall the feature extraction algorithm successfully extracts vehicles in two categories: cars and trucks. The algorithm extracts the features by circling them to highlight their location along with counting the quantity. To make more use of this data the imagery can be sectioned off and the model applied to individual sections. This will then allow analysis of volume and potential change of areas of interest.
There are some errors present within the final results, these include some small objects of similar colour and shape to cars and trucks along with some misidentification.
There are a few ways in which the model can be improved. One way is simply by adding more samples into the dataset. eCognition already does image augmentation by expanding out from the sample chips created by passing over the sample and expanding out from the point of the manual sample selected. However a much quicker and efficient way would be to do image augmentation by synthesis samples through flipping and rotating the current collected samples, and this will allow the model to see these samples in more ways whilst vastly increasing the amount of samples available for the model to be trained on. we can even do image augmentation synthesis through contrast too. This in turn will allow the deep learning tool to learn more about recognising cars and trucks which will have an overall effect of improving its accuracy, Nielsen (2017).
Finally the model and application take a while to run with eCognition which is not ideal when turning it into a commercial product. To enable this stage to happen the toolset needs to be converted over to another platform to enable it to be more efficient and quicker.

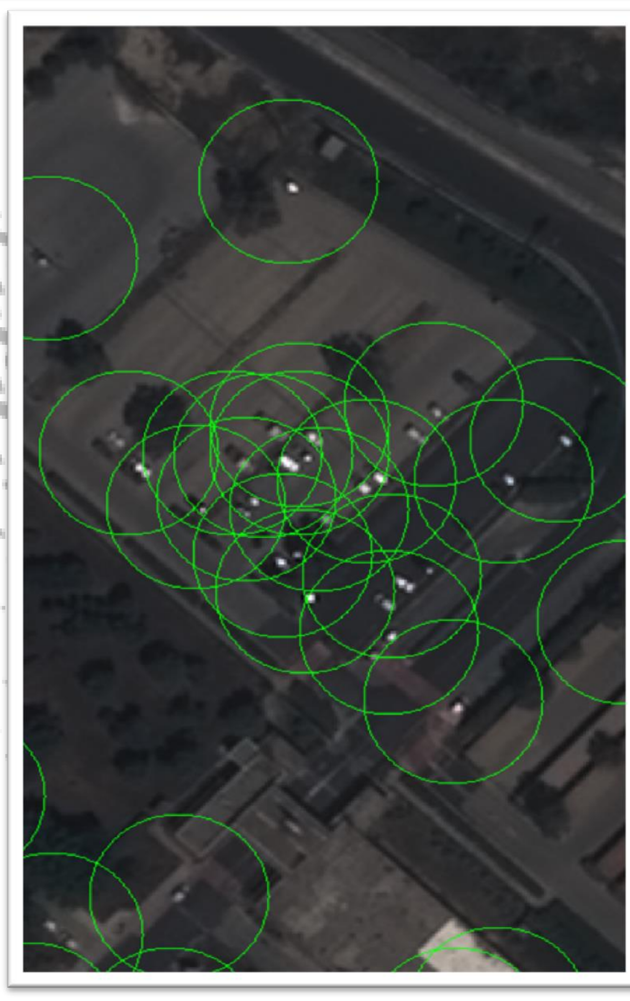

Figure 4: Cars in a car park being correctly identified by the algorithm. The circle is identifying each car individually and making a count for the statistical value.

Conclusion

The deep learning model within eCognition is an easy and accessible way to conduct neural networks to create a prototype, whilst gaining results relatively quickly. The convolutional neural network using eCognition's deep learning functionalities is a good way to extract features within imagery relatively quickly. However to turn this idea into a fully-fledged robust application, due to its size, it needs more processing power than eCognition can offer, as very deep neural networks need very powerful GPUs. Especially, due to more samples and image chips needing to be created to improve the overall accuracy of the model, particularly useful in helping to distinguish a truck cab to a car.

References

Tamta, K., Bhadauria, H.S., Bhadauria, A.S., 2015. Object-orientated approach of information extraction from high resolution satellite imagery. IOSR Journal of Computer Engineering (IOSR-JCE). PP 47-52.
AIRBUS DEFENCE AND SPACE. Satellite data [online]. Anon. 2018 [viewed 22nd September 2018]. Available http://www.intelligence-airbusds.com/satellite-data/
Trimble, 2018. eCognition developer: Tutorial 7- Convolutional neural networks in eCognition.
Nielsen, M., 2017. Neural networks and deep learning. Determination Press.