JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-15, No.-4, April (2020) pp 330-337 ISSN (Print) 0973-8975

OBJECT CLASSIFICATION IN HIGH RESOLUTION OPTICAL SATELLITE IMAGES BASED ON DEEP LEARNING TECHNIQUES

Wilver Auccahuasi[1], Percy Castro[2], Edward Flores[3], Fernando SernaquÉ[4], Sergio Arroyo[5], Javier Flores[6], Michael Flores[7], Felipe Gutarra[8], Nabilt Moggiano[9]

[1,2,3,4,5,6,7,8,9]Facultad de Ingenieria, Universidad Continental, Huancayo, Perú

[1]wauccahuasi@continental.edu.pe, [2]pcastro@continental.edu.pe, [3]eflores@continental.edu.pe, [4]fsernaque@continental.edu.pe, [5]sarroyo@continental.edu.pe, [6]jflores@continental.edu.pe, [7]mflores@continental.edu.pe, [8]fgutarra@continental.edu.pe, [9]nmoggiano@continental.edu.pe

https://doi.org/10.26782/jmcms.2020.04.00025

Abstract

The classification of objects that are present in the images or in the videos, is being developed progressively obtaining good results thanks to the use of Convolutional Networks, in this work we also use the convolutional networks for detection of objects that are present in high resolution satellite images, tests were carried out on ships that are on the high seas and in the ports, this classification is useful for monitoring the coasts, as well as for analyzing the dynamics of the ships can be applied in the search of ships, to cover this task of classifying ships in the spectral images, the use of high resolution satellite images of coastal areas and with a large number of ships is used, in order to build a set of images, containing images of the ships, in order to be used for training setting and testing of the convolutional network, a very particular configuration of the convolutional network caused by the particularity of high resolution satellite images is presented, the methodology developed indicating the procedures performed is also presented, a set of images containing 300 was built images of ships that are in the sea or are anchored in the ports, the results obtained in the classification using the convolutional networks are acceptable to be able to be used in different applications.

Keywords: Convolutional Networks, Satellite Image, Classification, High Resolution, Multispectral Image

I. Introduction

Convolutional networks are a very interesting variation of artificial neural networks. They are currently being widely used thanks to the availability of hardware that can be counted in these times. The best known convolutional networks are CNN.

We find jobs in many applications, such as in the extraction of automatic features from multispectral images; where CNN is used to label and identify palm cultivation units, resulting in a map of characteristics [II]. We find works where CNN and Virtual Reality are combined; using many data sources called multimodal data; where CNNs are used for image classification and virtual reality to create a 4D model [I]. In the classification of objects present in images and videos, such as pets, using CNN, obtaining a classification level of 90% [V]. Works are also presented, where images that are found in large volumes of databases are retrieved, using CNN, the results are better compared to conventional neural networks [IV]. CNNs are also used in vehicle detection, using videos produced by surveillance cameras, achieving acceptable results [III]. It is in this aspect that CNN networks can provide acceptable results in the classification of objects in high resolution satellite images, which is developed in the present work.

II. Materials and Methods

Satelital Image

A satellite image is a representation of data in the form of a matrix, formed by the optical instrument that is on board in space missions. This data matrix will depend on the resolution and characteristics of the optical instrument. They are formed by spectral bands, so they take the name of multispectral images. Among the bands are: the visible bands, red, green and blue.

Another characteristic of satellite images is the spatial resolution; this resolution is characterized by the equivalence of the pixel and its corresponding representation on the ground. This feature determines that optical instruments are classified as metric and sub metric. The equivalence in the metric is that one pixel in the image corresponds to more than one meter on the ground, and sub metrics indicate that one pixel in

The images used are in the range of sub metrics, so that in the image, objects can be observed. The ability to distinguish them will depend on the experience of the evaluator and the conditions of image acquisition. Having as a disadvantage the coverage in each acquisition. Sub-metric images have an approximate coverage of between 14 kilometers of average width and approximately 90 kilometers longer, compared to metric images that are 60 kilometers wide by approximately 300 kilometers long.

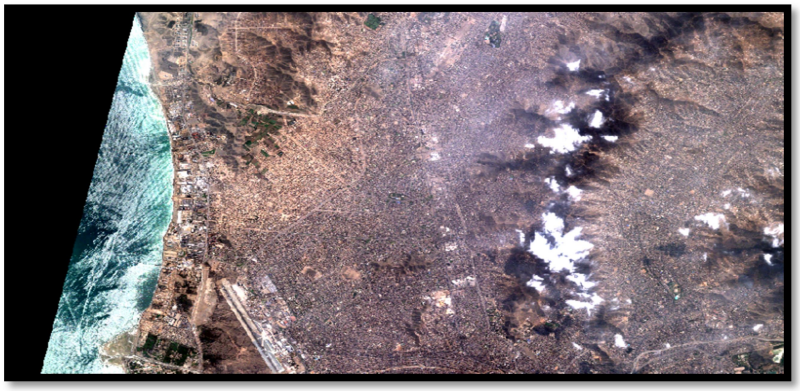


Fig. 1: Submetric satellite image, viewed from its original size.

In Figure 1 it can be seen that the images are in the combination of natural color, formed by the bands of the visual spectrum and ordered in the order of Red, Green and Blue. So it is also observed that the image is in its original view, where at first sight you cannot distinguish any object in the image.

High Resolution Satellite Image

When we work with sub metric images. A pixel corresponds to less than one meter on land, managing to observe objects that on average have a dimension of 5 meters, represented in the image between 6 and 7 pixels. By this characteristic the image is called high resolution image; because the image does not degrade as it becomes wide. Another important feature that the image has is its size, on average; an original image is between 3 gigabytes. The larger the length of the image, the larger the size in gigabytes. The high resolution image is shown below:

Fig. 2: Submetric satellite image, view the realizer of the zoom process.

In image 2, you can see the result of performing the zoom process, you can see that the image does not lose resolution and objects present in the image begin to appear. In the image you can see that you can distinguish roads, parks and vegetation areas grouped by green.

The image does not lose its resolution if we continue to perform the zoom process. When the zoom process is performed, objects within the image begin to discriminate as can be seen in the following image:

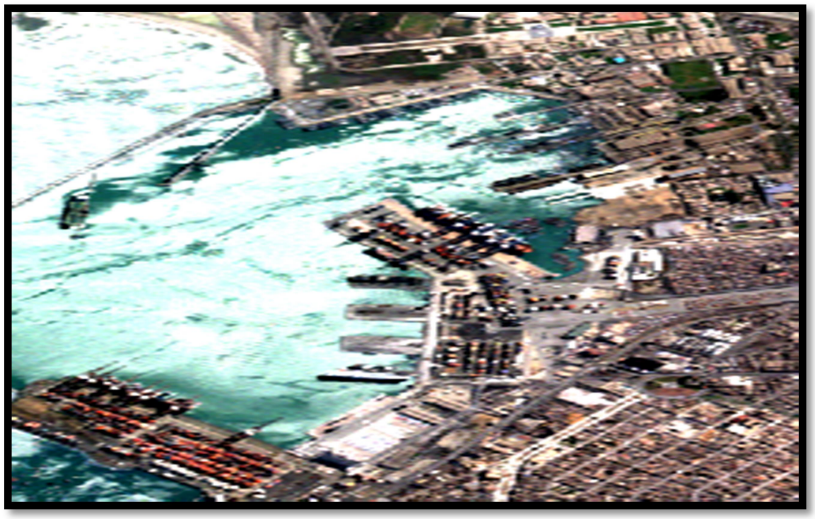


Fig. 3: Submetric satellite image, view to the realizer the zoom process increased.

In image 3, it is obtained by zooming to the image, where you can see you can already distinguish the houses, as well as the sea and objects that are on its coasts. Increasing the zoom the detail of the objects improve, as you can appreciate in the following figure:

Fig. 4: Submetric satellite image, where you can distinguish objects when performing the zoom process

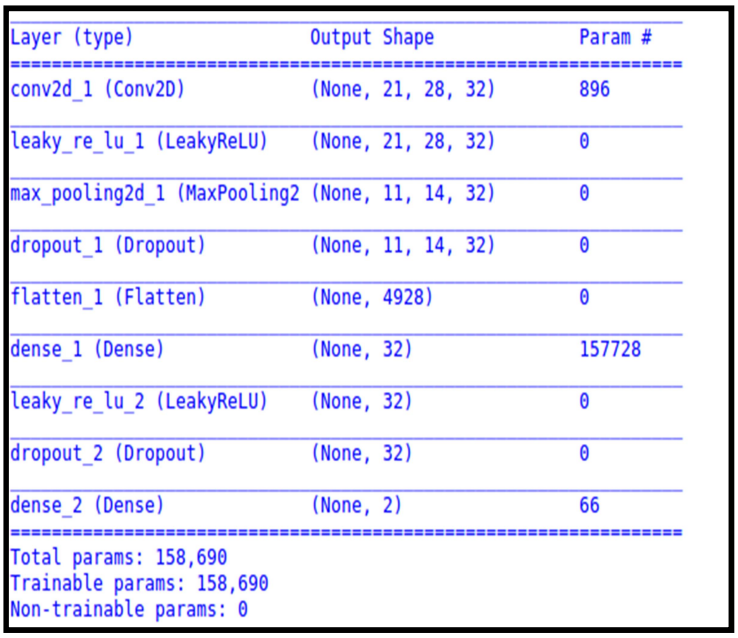
In figure 4, it can be seen that the detail of the objects in the image improves, also this increase in the zoom in the image, this is not distorted, you can see the boats that are on the coast of the sea and the Houses present in the image.

Convolutional Networks

The convolutional networks are a variant to the neural networks, with the characteristic of having many internal layers that perform a specific task to be able to improve in the classification. The convolutional network that was worked is characterized by basic layers that are described below:

- A "convolution" layer at the beginning of the network.
- A pooling layer at the exit of the convolution layer
- A flatten layer at the exit of the pooling layer
- A "dense" layer at the exit of flatten layer
- A second layer "dense" at the exit of the dense layer

Figure 5 shows the structure of the convolutional network with its respective layers described, as well as the parameters generated at the time of training.



Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 21, 28, 32)	896
leaky_re_lu_1 (LeakyReLU)	(None, 21, 28, 32)	0
max_pooling2d_1 (MaxPooling2	(None, 11, 14, 32)	0
dropout_1 (Dropout)	(None, 11, 14, 32)	0
flatten_1 (Flatten)	(None, 4928)	0
dense_1 (Dense)	(None, 32)	157728
leaky_re_lu_2 (LeakyReLU)	(None, 32)	0
dropout_2 (Dropout)	(None, 32)	0
dense_2 (Dense)	(None, 2)	66

Total params: 158,690
Trainable params: 158,690
Non-trainable params: 0

Fig. 5: Convolutional network model created

As a computational tool that was worked for the creation of the convolutional network, the use of Python language was used with the TensorFlow and Keras libraries.

Dataset Creation

The creation of the database of images. It corresponds to the fundamental part of the proposal presented in the present work. Having a large number of images, helps fundamentally, so that the artificial intelligence technique can have good results. The procedure for creating the image base corresponds to manual work. For this purpose, the greatest amount of high-resolution satellite images, containing information corresponding to the ships, was used. The job is to be able to detect and classify these objects in the original images. One of the characteristics of the investigation is to be able to perform the detection first on ships and then be able to continue with other objects; achieving a multi classifier, using the different techniques provided by Artificial Intelligence.

Fig. 6: image nesque conformanel dataset creado.

To increase the number of images in our data set, we used the technique of data augmentation; the objective is to increase the number of images. Using the following criteria such as image rotation and size increase.

Proposed Methodology

The proposed methodology is related to the reading and detection of ships present in high resolution satellite images. An image bank with tagged ships was implemented. The first thing to consider is that the satellite image has the original format. The working format of satellite images is GEOTIFF. So it is very difficult to open such files. The procedure is to convert to a different format such as: BMP, JPG and PNG. It is recommended to use the maximum special resolution of the image so as not to lose details of the original image, the "BMP" format recommended. You should also consider that the image in GEOTIFF has a pixel density of "16" bits. When converted to BMP, the conversion is done to 8 bits. This decrease in pixel density is necessary to work with the image. With the images in BMP format we proceed to crop and create the bank of images that correspond to the boats. With this image bank the classification is carried out using convolutional networks. The technique used to navigate the original image is the "sliding window". The flow chart of the proposal is shown in Figure 7.

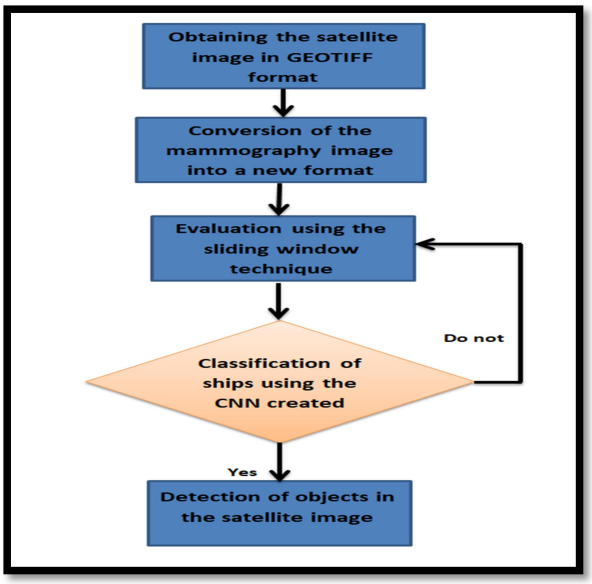


Fig. 7: Flow chart of the proposal.

III. Results

The results obtained at the time of the detection of the ships, using the created convolutional network, using the images of our created database. Show us in a first result the detection performance of a 0.94 which indicates a high level of image recognition, as can be seen in figure 8.

```
32/36 [=======================>....] - ETA: 0s
36/36 [==============================] - 0s 244us/step
('Test loss:', 0.229321559270223)
('Test accuracy:', 0.9444444444444444)
```

Fig. 8: Result of the classification of the convolutional network.

In figure 8, it can be seen that the convolutional network created allows obtaining a 94% performance in a first classification, using the 300 images of our dataset. For this analysis that uses 150 images for training and the other 150 images to perform the tests. This choice of the 150 images is made randomly, without using any selection criteria.

IV. Conclusions

When we work with high resolution satellite images; We work with special images, because the objects to be analyzed are immersed in a super image, due to their large size and number of pixels. One of the technical problems, from the image format to the pixel density. In most cases, you should consider performing

conversions so that you can work with the techniques that provide us with artificial intelligence. The original format of the satellite images is the GEOTIFF, managing to work at its maximum resolution. Adding the geographic information and in most cases the pixel density is 16 bits. Having to convert it to another compatible format obtaining information loss. Therefore, it is important to analyze the format of the target image so that it does not influence the result of a classification. It is recommended to work in the format that retains the maximum resolution of the image such as the BMP format. Assuming the loss of information when converting the pixel density of the image from 16 to 8 bits.

With the use of artificial intelligence techniques and with an emphasis on convolutional networks. It is being used in the classification of objects; in different types of images for different applications. To improve the classification, the structure of the convolutional network can be improved, increasing the number of layers. Among the recommended layers are the use of different filters such as the medium filter, to improve the resolution and presentation of the image. Also layers with edge detection filters, to separate the objects.

As a final conclusion, we can indicate that with the increase in the amount of images in our image base; There is a high probability of improving the performance results of the convolutional network. New network structures can also be designed to evaluate performance improvement so that we can have optimal results.

References

I. Maiwald, F., Bruschke, J., Lehmann, C., &Niebling, F. (2019). A 4D information system for the exploration of multitemporal images and maps using photogrammetry, web technologies and VR/AR. Virtual Archaeology Review, 10(21), 1-13.

II. Peña, A., Bonet, I., Manzur, D., Góngora, M., &Caraffini, F. (2019, June). Validation of convolutional layers in deep learning models to identify patterns in multispectral images. In 2019 14th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

III. Riveros, L., & Raquel, E. (2018). Detección de vehículos con aprendizajeprofundo en Cámara de Vigilancia.

IV. Sánchez Santiesteban, S. (2018). Recuperación de imágenesporcontenidousandodescriptoresgeneradosporRedesNeuronalesConvolucionales. RevistaCubana de CienciasInformáticas, 12(4), 78-90.

V. Weinstein, B. G. (2018). Scene-specific convolutional neural networks for video-based biodiversity detection. Methods in Ecology and Evolution, 9(6), 1435-1441.