Webology (ISSN: 1735-188X)
Volume 19, Number 4, 2022

Hybrid Methodology With Convolutional Neural Networks And Data Augmentation For Aerial Image Object Detection For Illegal Minery

JIMMY ANDERSON FLÓREZ ZULUAGA[1], JOSÉ DAVID ORTEGA PABÓN[2], JUAN CAMILO DÍAZ PAZMIÑO[3], JHON FREDY ESCOBAR SOTO[4]

[1]Technological Develop for Defense Center-FAC/ Researcher Corporación Universitaria Remington. PHD Engineering , CvLAC 0001488683

[2]Technological Develop for Defense Center-FAC https://orcid.org/0000-0002-2743-9422 CvLAC 0001488627

[3]Technological Develop for Defense Center-FAC https://orcid.org/0000-0001-9927-7760 CvLAC 0001770683

[4]Ph.D Managements Researcher Corporación Universitaria Remington https://orcid.org/0000-0002-6826-6222 CvLAC: =0000974420

Abstract

Automated classification of aerial remote sensing images is essential in automation processes and machine learning fields. This article describes an experimental technology to detect objects of interest in multispectral aerial images. This technology aims to identify areas with illegal mining, an issue that affects both the natural environment and security in Colombia; this is accomplished through the combination of digital image processing (DIP) techniques. DIP is used to filter objects of non-interest, based on their shape and color, from convolutional neural network systems based on a pre-trained MobileNetV2 model; data-augmented metrics and data without augmented metrics were then used for the detection of the illegal mining phenomenon.

With classification systems based on artificial intelligence, DIP could improve the efficiency of the autonomous processing of high-resolution input images in terms of time and identification of false positives.

The application of digital image processing PDI with classification systems based on artificial intelligence could improve the efficiency in the results of autonomous processing of high-resolution input images, that could be used for detection process automatization.

Keywords Aerial images image processing, satelital multiespectral image processign, image classification, mobilnetv2, Image digital processing, illegal mining, artificial intelligence.

Introduction

Remote sensing of the Earth from aircraft and satellite use is widely used in environmental sciences. Their application is growing due to technological improvements (Cracknell and Hayes, 2007) and extensive area coverage achieved through the interpretation of spectrum portions for the generation of information.

These systems utilize sensors, such as high-resolution cameras, synthetic aperture radars, and satellite images, among other instruments. These tools generate valuable interpretations of high-resolution data and must be analyzed to support decision-making processes. The manual process for such analysis is difficult and expensive as it involves a significant amount of time from trained experts to generate an accurate understanding of the data. As a result, automation is a potential option to increase the efficiency of the remote sensing process.

In this sense, aerial images could be used for the generalized analysis of diverse types of documented phenomena. Various studies (Boadi et al., 2016; Leo Stalin and Gnanaprakasam, 2020; Muharram et al., 2020; Stalin and Kumar, 2021) have proven useful in analyzing a myriad of concerns. These situations include deforestation and pollution, cadastral maps preparation and improvement, building construction, legal and illegal mining control, significant changes in the natural environment and its effects, and the growth of desert systems, such as dunes, among other aspects. In addition, according to studies (Flórez Zuluaga et al., 2017, 2019), this information can be used to analyze meteorology in air safety and air traffic flow and for automated convective weather systems, among other applications.

Aerial photographs are also widely used for crop analysis and verification of fuel-polluting conditions on coastlines (Clark & McKechnie, 2020; Vashchenko & Shavykin, 2021). For example, deforestation in Colombia has primarily resulted from an increase in hectares used for illicit use and an increase in illegal mining activities registered throughout the national territory (Romero et al., 2020; Ibrahim et al., 2020; Rincón-Ruiz et al., 2016; Rubio et al., 2019) and are issues that affect territorial integrity. Therefore, the Colombian Air Force (FAC), towards achieving its mission, assigns resources to perform patrol activities with aircraft and remote sensors capable of taking aerial photographs with spatial resolutions greater than 0.2 meters per pixel. Based on the altitude, these photographs can cover areas of hundreds of square kilometers in a single mission. The analysis of this data currently involves full workdays performed by highly trained personnel knowledgeable in each case issue.

For data collection, FAC has Leica ADS 80 cameras. These cameras generate images in JPEG 2000 formats and, according to Marcelino et al. (2000), are one of the data formats customarily used to maintain integrity without losing image resolution. For data collection, FAC has Leica ADS 80 cameras. As a result, aerial images are encoded with up to 24-bit

color resolution at an average rate of 4.17 bits/pixel. In aerial images, a file of this type, depending on the resolution and the altitude at which the image was taken, could cover hundreds of square kilometers, obtaining verifiable data that generates valuable information.

Data processing gathered in an overflight can take days to generate information that contributes to determining objectives of institutional interest, such as illegal mining areas. This processing is performed by experts who need several years to develop the skills required to identify aspects of interest in the image effectively. The interpretation of such data becomes a costly process, considering the large volume of information available for the institution. These situations are currently addressed through automatic analysis and data mining techniques (Crespo Peremarch, 2014), in which issues are identified, and an approach to the attribute type required for automated data classification is proposed.

This document presents the technological approaches implemented by the FAC in the search for a more efficient and automated process to prevent significant information loss and is also cost/efficient. Both aspects were considered restrictions for a more efficient and automated process. JPEG 2000 files of an ADS sensor with 16 bits of resolution and four bands of information were used.

The study results will be a basis for developing new tools to help expert personnel determine areas affected by illegal mining and improve the efficiency of the interpretation process and the time needed to verify the images described.

Research Background

Human intelligence can detect objects or areas of interest in complex and noisy scenes, even when different sensors represent reality. This is currently known as salient object detection (SOD) and consists of selecting objects that attract an observer's attention (Wang et al., 2021). This is a widely used feature in searching for objects of interest in aerial or satellite images of high resolution. This technology is used in locations where one or few objects that generate contrast are expected due to the scene's characteristics. This, in turn, attracts the observer's interest in the same manner as it occurs through observations of rural or jungle areas via SOD sensors that detect changes in the topsoil.

To achieve more extensive coverage, it is typical to use optical-type remote sensing (RS) tools for high-resolution imaging (RSI) for civil and military purposes during observation missions; that is, dual-use tools that, at a subsequent stage, pursue the classification and detection of objects of interest. The analysis of various object types are conducted in the techniques implemented through this study due to an increased demand for services (Li et al., 2021; Li et al., 2020) that automate or facilitate object detection in these type of sources, including ships, vehicles, or people, combined with orientation and spatial resolution, and scale, categories, and background characteristics of the object.

Likewise, there is a demand for wide-area measurement systems for maritime, pre-shipping, deforestation, meteorology, and defense control, among others, to pursue more extensive control area coverage through remote sensing. Cheng and Han (2016) describe the challenges in aerial and satellite imagery analysis and provide a review of advances in the field, focusing primarily on general objects, such as airports, vehicles, and ships. They classify them using large method groupings, such as template matching-based, knowledge-based, object-based image analysis (OBIA), and machine learning-based methods.

Significant progress has been achieved in the artificial intelligence field regarding object classification in remote sensing images. However, this degree of processing is often limited by the lack of databases in objects of interest or by objective characterizations through labels that, in the case of uncommon objectives in the supervised systems, must be performed by experts who require large amounts of available data. According to Betti and Giraldez (2020) and Moreno Revelo (2020), supervised systems include those based on deep convolutional neural network (CNN) systems, require both training data and validation data, and have demonstrated high object recognition performance for object recognition and object detection with insufficient computational resources.

In addition to supervised techniques to analyze and measure the increase in agricultural, urban, and mining activities (along with other anthropogenic factors that modify the environment), digital image processing (DIP) techniques were used with satellite data and geographic information systems (GIS). These tools are used to study areas of interest (Manjunatha and Basavarajappa, 2017) where visual image interpretation techniques (VIIT) were combined with DIP to obtain better results. Significant advancements in scene analysis have been achieved in remote sensing image data analysis. Guan et al. (2020) propose a random access meta metric learning model consisting of two phases. The first phase is random episodic training introducing a DIP-based process on most classes that contain enough samples, the second performs fine-tuning with all classes to improve performance for classification of all categories. Paiva et al. (2019) describe digital processing techniques used for satellite images that provide satisfactory results in identifying areas with deforestation resulting from legal or illegal mining activities. These techniques are also used for cartography spatial analysis (Buzai et al., 2018) and the identification of environmental effects, such as pollution (Ingeniería, 2018; Manjunatha and Basavarajappa, 2017). According to Kyrkou and Theocharides (2019), other approaches, including computational efficiency, are based on deep learning and lightweight CNN systems. Because of their low computational requirements, these systems can even support control automation process operations in unmanned aircraft systems.

Likewise, Kyrkou and Theocharides (2019), and Martins and Cavassan Zaglia (2019) employed CNN with satellite data to detect deforestation in satellite images; another use is locating vehicles in aerial images, including traffic analysis and search and rescue in disaster

images. Furthermore, Li et al. (2017) describe these applications as ways to address multiclass classification procedures using imbalanced databases.

The category of CNN architecture depends on several factors, such as the category of image data or sensor, the computational capacity, and the existence of classification databases, among others. Regarding this, Hu et al. (2015) exhibit the use of deep CNNs for high-resolution remote sensing image systems (HRRS) through various convolutional layers, such as the compatibility layer. Simonyan and Zisserman (2015) describe CNNs based on high-performance computing systems, which allow an increased depth of the net using an architecture with tiny convolution filters (3 x 3) to improve system performance.

These CNN technologies are also used for Earth observation in multispectral satellite images (Podorozhniak et al., 2018). They can also serve to effectively recognize images of interest that are helpful in the detection of illegal mining activities from aerial data (Saavedra, 2021).

The detection of illegal mining is difficult because there are no databases with available references of particular characteristics to be used as a descriptor or baseline to train CNNs. The creation of a specific database could obstruct research in illegal mining since it implies the use of significant resources; that is, data that contribute to positive and negative sample training containing the object of interest, where such an object is not present. When there is insufficient data, alternative techniques, such as data augmentation to improve system metrics, are recommended, as described by Chen et al. (2018) and Okafor et al. (2017, 2018).

For the CNN method to be utilized, an analysis supported by the EAFIT University (EAFIT-HUMAT, 2021) was conducted at the Aerospace Technology Development Center for Defense (CETAD) of FAC. Mathematical models for detection and classification were validated with a private database consisting of 500 image croppings with objects of interest to evaluate the model performance. This performance assessment was based on deep learning. Among the evaluated models were: DenseNet169, DenseNet201, InceptionResNetV2, InceptionV3, Mobilent, MobileNetV2, ResNet101, ResNet152, ResNet50, VGG16, VGG19 and Xception. Among the results obtained, according to work metrics oriented to real-time processing, technologies such as MobileNetV2 have metrics that, compared to machine resources, are superior to other verified models.

Materials and Methods

This study aims to develop a methodology to allow the execution integration of automatic processes with the work of expert personnel in aerial image analysis. For this purpose, there is an advantage related to identifying areas with the presence of illegal mining in aerial images; it is a process that includes multiple repetitive tasks, and the effectiveness of CNNs has been demonstrated. They have been successfully used in repetitive processes that range from text recognition to analyzing objects of interest in aerial or satellite images (De Santiago et al., 2018; LeCun et al., 1998; Mateus de Souza et al., 2021; Radovic et al., 2017).

This work strives to identify the best possible technique combination to yield optimal results in areas with illegal mining in aerial images of rural and jungle areas of Colombia. Colombia is a country that is deeply affected by illegal mining, and threatens not only economic security but also the environment (Billon et al., 2020; Hilson, 2002; Jensen et al., 2021; Lara-Rodríguez, 2020).

To perform this task, the FAC has trained qualified personnel to identify the distinctive aspects of areas affected by illegal mining (contamination by mercury, deforestation in regions with prominent sources of gold, and the presence of heavy machinery equipment, such as dredgers, dump trucks, and backhoes in rural or jungle areas) in aerial images. Experts spend a significant amount of time post-processing and searching for areas of interest to identify evidence of illegal mining activities. The sensor used to collect these images is a Leica ADS 80 with spatial resolutions between 0.13 to 0.50 meters per pixel and 16-bit resolution, with four reflectance bands of four channels: red (R), green (G), blue (B), and near-infrared (NIR). The NIR band of the electromagnetic spectrum has a wavelength range of 750 to 2500 nm and a frequency of 215 THz to 400 THz. Each file can contain information on sites ranging from 20 kilometers to hundreds of square kilometers. In Figure 1, the RGB information band and the NIR band are displayed.

Figure 1. Reflectance bands. Left: RGB bands. Right: NIR band

The first step of this study was the research background design to determine what technologies could contribute to identifying elements of interest in aerial imaging of jungle areas. As previously mentioned, techniques such as DIP are used to identify features, such as color, edges, and shapes, and different algorithms or CNN architectures are commonly used for classification.

Meetings with experts were then held to define the main characteristics of the areas of interest. Based on these aspects, an inventory of the technical properties of aerial images to be processed was created:

1. Prevalence of jungle areas in the areas of interest;
2. Ruptures in the upper part of the soil or localized deforestation within jungle areas;

3. A presence of industrial machinery, such as dump trucks, dredgers, and backhoes; and
4. In some regions, changes are related to the color of water sources due to pollution.

A process based on DIP and CNN technologies was selected to minimize the classification and processing time, using this information as a launching point to create a more cost/efficient process in terms of computational aspects. For this, the process was divided into two main models. The first model prepared approved data in the classification process in which the image areas that advance for an objective classification process are filtered and cropped. The second model identified objects of interest through automatic classification processes. Subsequently, based on the research background, experts stated that for the automated processes of determining objects of interest, discovering deforested areas and industrial machinery were vital as they were factors that would generate the most significant contrast in the image. With these elements, a hypothesis of using DIP as a filtering basis was proposed since the objects of interest are primarily found in areas with abundant vegetation, considering each region may have differences in the structuring of these filters. Likewise, these filters eliminate unnecessary information within the location, minimizing the amount of data moving through the classification process and reducing processing time.

For this purpose, it was determined that the filter definition, based on DIP techniques with range filters in the HSV (Hue, Saturation, Value), color model using each image's characteristics and according to the geographical area and the predominant characteristics of the land, enhances the process in terms of identifying false positives.

DIP Filtering

The DIP methodology uses the HSV model, which maps the RGB primary colors in dimensions the human eye can easily perceive and visualize. This model consists of three dimensions (hue, saturation, and value) that define the "color space." There are areas in aerial images with ranges of specific colors; for example, vegetation (green) in rural areas is throughout the national territory. Therefore, an algorithm was designed to select ranges to filter dense vegetation regions as the identification of yellow machinery is not relevant within the methodology and eliminates false positives from the classifications. One example of the results obtained in the processed images is presented below (Figure 2).

Figure 2. Ranges HSV

The algorithm consists of a minimum and maximum range function of colors to be filtered and the dilation value that defines the thickness of the points. This function is recurrent for each 75x75 pixel cutout generated in the mosaics. The objective is to filter croppings with a high vegetation index. The number of points within the HSV ranges is calculated for such an index. First, the clipping values are converted into HSV space; then, a binarized mask is generated using a dilation to expand the thickness of the vegetation areas, and; finally, the histogram is calculated to establish the number of points that determines if a clipping value contains vegetation regarding the percentage of vegetation previously shown (60%) (Figure 3).

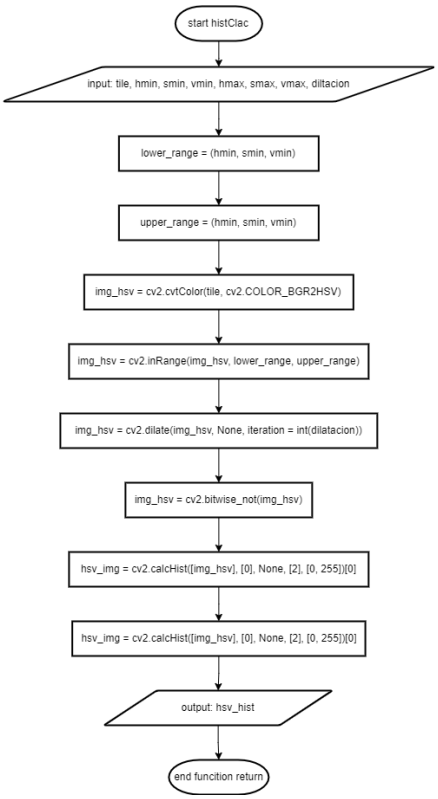

Figure 3. HSV Algorithm

Neural Networks

Regarding classification and identification issues, the literature reviewed showed that technologies derived from CNN yielded promising results (Fujita et al., 2017; Maloof et al., 2003; Qu et al., 2016). However, for other authors (Chen et al., 2018; Gupta et al., 2019), the flaw of these technologies is the existence of a previously classified and labeled data set. Since this data set with operational characteristics of industrial machinery in rural areas was not available, its creation was necessary. For this, according to the available data, it was determined that the objects of interest could be contained in cropped images from 250 x 250 to 75 x 75 pixels, for which the verification of results and classification were performed, obtaining better solutions with 75x75 pixel croppings.

Regarding detecting areas of interest, two types of classification models have been tested: the first classification model, a binary classification, determined the existence or non-existence of an object of interest from the categories “Machinery” and “Non-Machinery.” The second classification model was a multiclass model, in which the following classes were defined: 1) “Dump Truck,” 2) “Backhoe,” 3) “Dredge,” and 4) “Non-class.”

The completed database collected the initial coordinates of the cutout and class in a shapefile. The completed database collected the initial coordinates of the cutout and class in a shapefile. After reviewing over 200 aerial images, 10,000 clippings were generated, and, from this, the database was created consisting of 146 dredgers, 1,644 backhoes, 809 dump trucks, and 2,599 objects considered non-class, with a total of 5,198 elements classified, as seen in Figure 4.

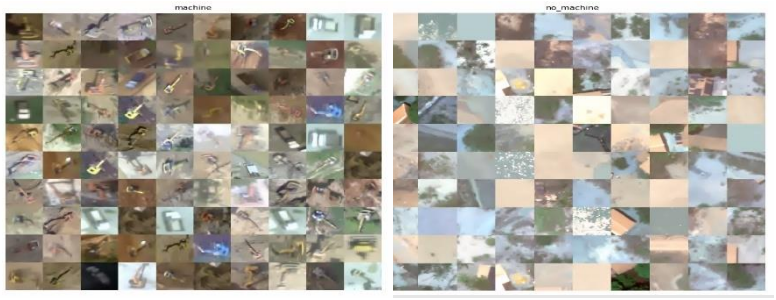

Figure 4. Sample of Generated Clippings

Based on this, an architecture based on the MobilNetV2 deep learning model was applied. This model had variations in the training conditions to verify if techniques such as data augmentation and filtering processes improve in some way the results obtained in the confusion matrix.

For network training and inference processes, a computer with an Intel® Xeon® Gold 6246 CPU@3.30GHz×48 processor with a Quadro RTX 8000/PCIe/SSE2 48GBX2 GPU and an Ubuntu 18.04 operating system was used.

Likewise, for the architecture comparison process and the analysis of results, each trained model was evaluated as a classifier using the cuts mentioned above. Regarding the correct classification percentage (accuracy), the metrics selected were defined as the percentage of hits performed. Losses (loss) are understood as the difference between the desired output and the output obtained by the system (De, 2018; Rangel Cortés et al., 2019). A total of 40 epochs without data augmentation and 40 epochs with data augmentation were performed on each architecture. Finally, a behavioral contrast between the best validation error (val_loss) and best validation accuracy (val_acc) was examined.

Data Augmentation

Data augmentation techniques help improve the performance and results of deep learning models by using new and different samples to train the network, allowing generalization when new data is available.

For this study, it was necessary to use data augmentation due to the small amount of data available and the exceptionally exhausting and expensive nature of the image collection process. For instance, there may be images in which no objects are found after careful left-to-right and top-to-bottom revision when performing a manual review. Data augmentation exhibited the same characteristics during each exercise and was based on rotation and projection techniques on the original image axes, more precisely regarding horizontal reflection, vertical reflection, and 90º clockwise rotations, as shown in Figure 5.

Although there are many applications, techniques, and filters, four were selected to conduct this study: Fliplr, Flipud, Rot90, CropToFixedSize (Figure 5), as these aerial images and the objects of interest can have any rotation type in the image space. Moreover, some images are captured at various heights, resulting in objects appearing with variations in size.

Figure 5 . Data Augmentation

Training

Transfer learning techniques were utilized to conduct this work in which a pre-trained neural network was employed, and the new data set was adapted by transferring or reusing learned features without the need of training a new network Transfer learning allows the use of pre-trained weights as a launching point with the possibility of reaching optimal performance levels in the model. The entire training process was performed with the following hyperparameters: batch_size of 128, and the Adam optimizer was employed with a learning rate of 0.001. With these values, four models were created.

Regarding the training process methodology used, the entire data set was divided into a ratio of 60% for training, 20% for tests, and 20% for validation. Although it was not mandatory to set these values, it must be designed for the models to have the capacity to classify a new image adequately. The training set was used to train and cause the CNN to learn the features

and patterns that the images contain. The validation set was a separate set, and its purpose was to validate the model's performance during training. Meanwhile, the test set is a particular set of data used to trial the model after training.

Binary Classification

For the binary classifier without data augmentation training, a sample of 5198 balanced cuts was utilized: 2599 balanced cuts of Machinery and 2599 of Non-machinery, respectively. For data augmentation and implementing a balanced model, there were 25,990 cuts from the data set, as shown in Table 1.

Table 1. Distribution of the Binary Classifier Sample with Data Augmentation

		Training	Validation	Tests	Total
Without Data Augmentation	Machinery	1559	520	520	2599
	Non-machinery	1559	520	520	2599
					5198
With Data Augmentation	Machinery	7797	2599	2599	12,995
	Non-machinery	7797	2599	2599	12,995
					25,990

Multiclass Classification

For the multiclass classification, the original data set was employed with the quantities arranged with data augmentation (Table 2), in which experts validated each category of cuts.

Table 2. Distribution of the Multiclass Classification Sample with Augmentation

		Training	Validation	Tests	Total
Without Data Augmentation	Dredge	88	29	29	146
	Backhoe	986	329	329	1644
	Dump Truck	485	162	162	809
	No Machinery-class	1559	520	520	2599
					5198

With Data Augmentation	Dredge	438	146	146	730
	Backhoe	4932	1644	1644	8220
	Dump Truck	2427	809	809	4045
	No Machinery-class	7797	2599	2599	12,995
					25,990

Figures 6-9 show the training graphs for the four models and provide a better understanding of each model and how the system learns with each epoch. A clear difference can be observed when the data augmentation technique is implemented as it merges more efficiently to avoid overfitting. On the contrary, when data augmentation is omitted, the validation curve is far from the training models, indicating that the models do not adjust to the data presented due to the small amount of data used.

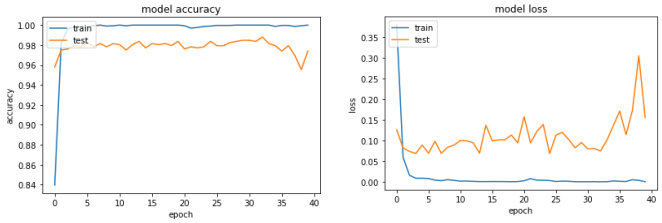

Figure 6. Model 1: Binary Classification without Data Augmentation

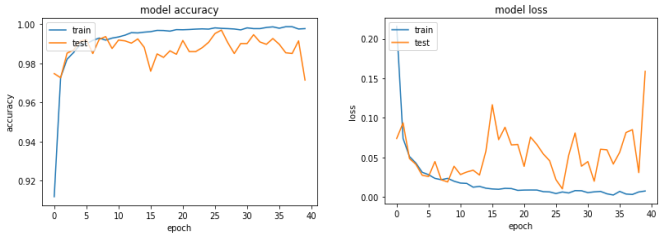

Figure 7. Model 2: Binary Classifier with Data Augmentation

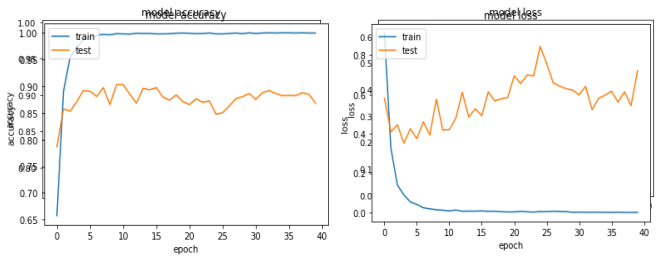

Figure 8. Model 3: Multiclass Classification without Data Augmentation

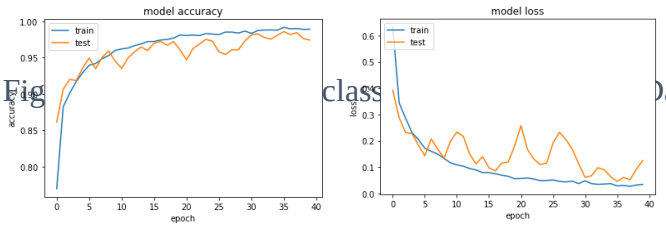

Fig 8. class Data Augmentation

Test Approach

Once the data set is defined and the CNNs are trained, a series of tests are proposed, and the following configurations are established:

1. Binary classification tests. Only the Machinery and Non-machinery categories were used for classification and training in these tests. The test objective was to determine image areas with possible objects of interest without defining the object type or its position within the cropping. Classification training tests with data augmentation techniques and without data augmentation were also developed to explore performance improvements and contribute to data augmentation in this classification process.

2. Multiclass classification tests. For these tests, labels were divided into the following classes: “Dredger,” “Backhoe,” “Dump truck,” and “No-class” to find an object of interest and perform the first automatic classification according to the previously defined categories.

In both settings, the tests were conducted with DIP filters and without DIP filters; also, in both cases, the classifier training was tested with and without data augmentation.

Experimental Results

The evaluation of the implemented inference process performance for classifying objects of interest in high-resolution aerial images was conducted through a simulation of the current process performed by experts. A set of 15 mosaics with objects of interest was employed for experiment validation. With these mosaics, a new data set was created, consisting of data as shown in Table 3. This was performed to compare the data obtained by the system and the manual labels determined by the experts, comparing the models acquired with and without data augmentation.

Table 3 1. Experimental Set

Binary Classifier		Multiclass Classifier	
Máquina	138	Draga	39
		Volqueta	36
No máquina	37513	Retro	63
		No máquina	37513

The results are presented as a confusion matrix. Experiment 1: Binary classification is presented in Figures 10-11 and in Table 4. Experiment 2: Multiclass classification is presented in Figures 12-13 and in Table 5.

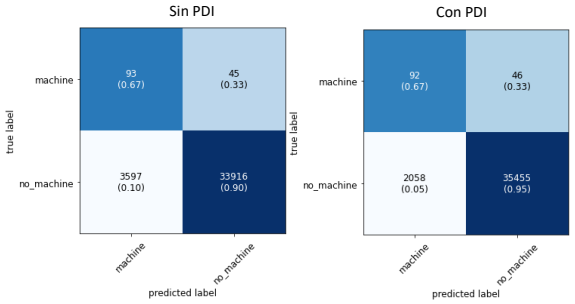

Figure 2 Model 1 Confusion Matrix with and without DIP

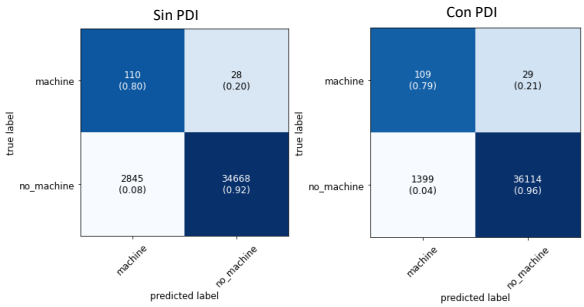

Figure 3. Model 2 Confusion Matrix with and without DIP

Table 4. Summary of Binary Classification Experimental Results

Metrics	Accuracy	Recall	F1-Score
Model 1 without PID filter in data set	0.51	0.79	0.50

Model 1 with PID filter in data set	0.52	0.81	0.53
Model 2 without PID filter in data set	0.52	0.86	0.52
Model 2 with PID filter in data set	0.54	**0.88**	**0.56**

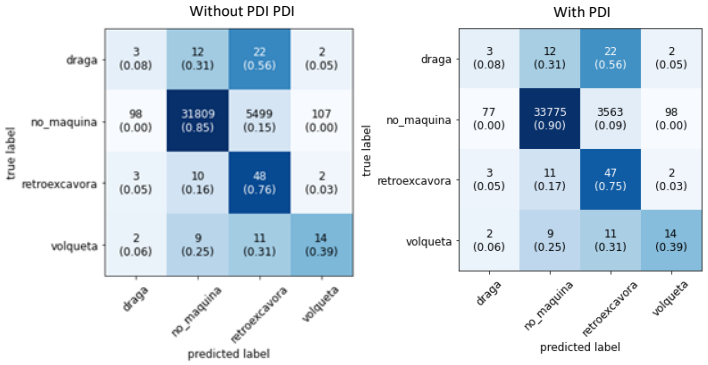

Figure 12. Model 3 Confusion Matrix with and without DIP

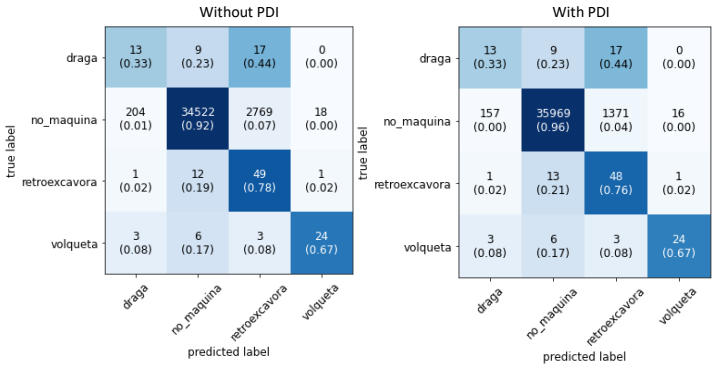

Figure 13. Model 4 Confusion Matrix with and without DIP

Table 5. Summary of Multiclass Classification Experimental Results.

Metrics	Accuracy	Recall	F1-Score
Model 3 without PID filter in data set	0.29	0.52	0.29
Model 3 with PID filter in data set	0.29	0.53	0.30
Model 4 without PID filter in data set	0.41	0.67	0.42

Model 4 with PID filter in data set	0.42	0.68	0.45

Discussion and Conclusions

The automated classification identified that the classification requires cuts of 75x75 pixels and implies a complementary process. Therefore, developing a PID-based system to perform this classification task to prevent an object of interest from being lost in the clipping process was necessary. For this purpose, a function to isolate a clipping area was based on an overlapping window covering all the regions of interest with full clips. However, the clipping process was slow and resulted in a large number of clippings being processed. Table 7 details the times and amount of clipping that originated from an image

The target areas are jungle and rural regions and are a common factor. Much of the resulting clippings correspond to these regions, which becomes too noisy for the process goal. Within the functionalities of the DIP, a series of filters were developed, based on the image histograms, to mitigate this noise and achieve greater efficiency in the process. These filters also eliminated colors corresponding to jungle areas from the image and recalculated cuts, limited to those in which data different from the frequencies of the predominant colors of reference in such regions were found.

By implementing these filters on previously analyzed images, the number of cuts is shown in Table 7. By implementing an integrated methodology, the results were evaluated by integrating the DIP process considering time as a new variable in the validation process, obtaining the results presented in Table 7.

In conclusion, DIP processes, when there are large areas in aerial or satellite images with similar conditions, can be useful for the application of filters that can help reduce the search area of objects of interest in these images. This in turn reduces the area that will be the input of object detection processes with artificial intelligence, avoiding the generation of false positives in areas where there is vegetation.

It can also be concluded that the use of data augmentation allows to improve the metrics, especially for objects that do not have a large number of images that allow adequate training of the systems, thus contributing to improve the performance of the system.

On the other hand, it can be concluded that the processes of detection of objects of interest, applied to cases of illegal mining through artificial intelligence, facilitate the search of areas automatically, so that the techniques and methods implemented generate cost-effective processes. efficient for the detection of areas with machinery so that they can be used by the authorities as inputs in the fight against this phenomenon.

Based on these results, the development of detection and filtering systems could continue, using artificial intelligence architectures for segmentation and the combination of techniques in order to obtain better results that allow determining areas of illegal mining.

Table 6. Inference times and trims

#	Image Name	Size (MB)	Cuts	Average Clipping Time (ms)	Half trim time (ms)	Inference Time (ms)	Average Inference Time (ms)	DIP cutou	DIP Inference Time (ms)	Average Inference Time (ms)
1	rec1_mision 24_linea11	25.6	1400	541.58	0.39	834.73	1.68	1090	636.06	1.71
2	rec1_mision 33_linea4	115.9	6834	2626.34	0.38	3781.6	1.80	3812	2088.2	1.82
3	rec1_mision 33_linea7	31.9	1855	717.37	0.39	1099.27	1.69	1580	966.24	1.64
4	rec1_mision 107_linea7	60.6	3526	1401	0.40	2028.61	1.74	3155	1841.66	1.71
5	rec1_mision 108_linea7	33.5	1947	734.90	0.38	1091.60	1.78	1724	932.57	1.84
6	rec2_mision 24_linea11	8.5	450	166.11	0.37	250.25	1.80	319	191.55	1.67
7	rec2_mision 33_linea4	44.9	2556	936.63	0.37	1427.72	1.79	1194	700.10	1.70
8	rec2_mision 107_linea7	60.6	3526	1831.43	0.38	1980.01	1.78	2019	1177.82	1.71
9	rec2_mision 108_linea7	83.2	4836	1831.43	0.37	2731.50	1.77	4647	2597.01	1.79
10	rec3_mision 33_linea4	20.7	1092	436.11	0.40	629.69	1.73	593	390.26	1.52
11	rec3_mision 108_linea7	65.5	3744	1393.85	0.37	2068.42	1.81	1595	861.13	1.85
12	rec4_mision 33_linea7	34.2	1995	771.36	0.39	1146.37	1.74	1968	1132.95	1.73
13	rec6_mision 33_linea4	20.7	1160	449.16	0.39	710.71	1.63	939	521.37	1.80

14	rec6_mision33_linea7	26.2	1470	606.85	0.41	890.48	1.65	819	456.94	1.79
15	rec8_mision33_linea4	23.1	1260	485.56	0.39	693.96	1.81	684	382.97	1.79
Total			**37651**	**14929.68**		**21364.92**		**26138**	**14876.83**	

References

1. Betti, A. y Giraldez, N. R. (2020). Aplicación de técnicas basadas en aprendizaje profundo para la clasificación de imágenes satelitales y otras plataformas de observación terrestre (Proyecto final). Universidad Nacional de Mar del Plata.
2. Billon, P. Le, Roa-García, M. C. y López-Granada, A. R. (2020). Territorial peace and gold mining in Colombia: local peacebuilding, bottom-up development and the defence of territories. Conflict, Security & Development, 20(3), 303-333. doi: 10.1080/14678802.2020.1741937
3. Boadi, S., Nsor, C. A., Antobre, O. O. y Acquah, E. (2016). An analysis of illegal mining on the Offin shelterbelt forest reserve, Ghana: Implications on community livelihood. Journal of Sustainable Mining, 15(3), 115-119. doi: 10.1016/j.jsm.2016.12.001
4. Buzai, G. D., Lanzelotti, S. L., Paso Viola, L. F. y Principi, N. (2018). Cartografía analógica y digital para la delimitación regional y el análisis temático: aplicación a la cuenca del río Luján (Argentina). Revista de Geografía Norte Grande, (69), 99-119. https://doi.org/10.4067/S0718-34022018000100099
5. Chen, H. T., Liu, C. H. y Tsai,W. J. (2018). Data Augmentation for Cnn-Based People Detection in Aerial Images. 2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 1-6. doi: 10.1109/ICMEW.2018.8551501
6. Chen, S. A., Escay, A., Haberland, C., Schneider, T., Staneva, V., & Choe, Y. (2018). Benchmark Dataset for Automatic Damaged Building Detection from Post-Hurricane Remotely Sensed Imagery. 1–8. http://arxiv.org/abs/1812.05581
7. Cheng, G. y Han, J. (2016). A survey on object detection in optical remote sensing images. ISPRS Journal of Photogrammetry and Remote Sensing, 117, 11-28. doi: 10.1016/j.isprsjprs.2016.03.014.
8. Clark, A., & McKechnie, J. (2020). Detecting banana plantations in thewet tropics, Australia, using aerial photography and U-net. Applied Sciences (Switzerland), 10(6), 1–15. https://doi.org/10.3390/app10062017.
9. Cracknell, A. P. y Hayes, L. (2007). Introduction to remote sensing (2.a ed.). CRC Press. ISBN: 9780429121135. https://doi.org/10.1201/b13575
10. Crespo Peremarch, P. (2014). Análisis de minería de datos para la clasificación de imágenes aéreas (tesis de maestría). Universitat Politècnica de València, España.

11. De Santiago, V. A., Da Silva, L. A. R. y De Andrade Neto, P. R. (2018). Testing Environmental Models supported by Machine Learning. SAST '18: Proceedings of the III Brazilian Symposium on Systematic and Automated Software Testing, 3-12. doi: 10.1145/3266003.3266004
12. EAFIT-HUMAT. (2021). Documento privado resultado de consultoria. Asesoría para el diseño, desarrollo e implementacion de algoritmos de deep learning. Informe documentacion interna de la Fuerza Aérea Colombiana.
13. Florez Zuluaga, J. A., Vargas Bonilla, J. F. y Reina, J. (2017). Intelligent techniques for identification and tracking of meteorological phenomena that could affect flight safety. Ciencia y Poder Aéreo, 12, 24-35. https://doi.org/10.18667/cienciaypoderaereo.559
14. Florez Zuluaga, J. A., Ortega Pabón, D., Vargas Bonilla, J. F. y Quintero Montova, O. L. (2019). Meteorological Risk Early Warning System for Air Operations. 2019 IEEE International Symposium on Technology and Society (ISTAS), 2019, pp. 1-10, doi: 10.1109/ISTAS48451.2019.8938012
15. Fujita, A., Sakurada, K., Imaizumi, T., Ito, R., Hikosaka, S. y Nakamura, R. (2017). Damage detection from aerial images via convolutional neural networks. Proceedings of the 15th IAPR International Conference on Machine Vision Applications (MVA), 5-8. doi: 10.23919/MVA.2017.7986759
16. Guan, J., Liu, J., Sun, J., Feng, P., Shuai, T. y Wang, W. (2020). Meta metric learning for highly imbalanced aerial scene classification C. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4042-4046.
17. Gupta, R., Goodman, B., Patel, N., Hosfelt, R., Sajeev, S., Heim, E., Doshi, J., Lucas, K., Choset, H. y Gaston, M. (2019, junio). Creating XBD: A dataset for assessing building damage from satellite imagery. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 10-17.
18. Hilson, G. (2002). Small-Scale Mining in Africa: Tackling Pressing Environmental Problems With Improved Strategy. The Journal of Environment & Development, 11(2), 149-174. doi: 10.1177/10796502011002003
19. Hu, F., Xia, G. S., Hu, J. y Zhang, L. (2015). Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. Remote Sensing, 7(11), 14680-14707. doi: 10.3390/rs71114680
20. Ibrahim, E., Lema, L., Barnabé, P., Lacroix, P. y Pirard, E. (2020). Small-scale surface mining of gold placers: Detection, mapping, and temporal analysis through the use of free satellite imagery. International Journal of Applied Earth Observation and Geoinformation, 93(102194). https://doi.org/10.1016/j.jag.2020.102194
21. Angeles, Guillermo Raul; Geraldi, Alejandra Mabel; Marini, M. F. (2020). Procesamiento digital de imágenes Metodologías y técnicas (A. M. Geraldi (ed.); CONICET). https://ri.conicet.gov.ar/handle/11336/135234.

22. Jensen, M., Combariza Bayona, D. A. y Sripada, K. (2021). Mercury Exposure among E-Waste Recycling Workers in Colombia: Perceptions of Safety, Risk, and Access to Health Information. International Journal of Environmental Research and Public Health, 18(17). doi: 10.3390/IJERPH18179295
23. Kyrkou, C. y Theocharides, T. (2019). Deep-Learning-Based Aerial Image Classification for Emergency Response Applications Using Unmanned Aerial Vehicles. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 517–525. doi: 10.1109/CVPRW.2019.00077
24. Lara-Rodríguez, J. S. (2020). How institutions foster the informal side of the economy: Gold and platinum mining in Chocó, Colombia. Resources Policy, 74, 101582. doi: 10.1016/j.resourpol.2020.101582
25. LeCun, Y., Bottou, L., Bengio, Y. y Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2323. doi: 10.1109/5.726791
26. Leo Stalin, J. y Gnanaprakasam, R.C. P. (2020). Application of Unmanned Aerial Vehicle for Mapping and Modeling of Indian Mines. Journal of the Indian Society of Remote Sensing, 48(6), 841-852. doi: 10.1007/s12524-020-01118-3
27. Li, B., Xie, X., Wei, X. y Tang, W. (2021). Ship detection and classification from optical remote sensing images: A survey. Chinese Journal of Aeronautics, 34(3), 145-163. doi: 10.1016/j.cja.2020.09.022
28. Li, C., Cong, R., Guo, C., Li, H., Zhang, C., Zheng, F. y Zhao, Y. (2020). A parallel down-up fusion network for salient object detection in optical remote sensing images. Neurocomputing, 415, 411-420. doi: 10.1016/j.neucom.2020.05.108
29. Li, F., Li, S., Zhu, C., Lan, X. y Chang, H. (2017). Cost-Effective Class-Imbalance Aware CNN for Vehicle Localization and Categorization in High Resolution Aerial Images. Remote Sensing, 9(5), 494. doi: 10.3390/rs9050494
30. Manjunatha, M. C. y Basavarajappa, H. T. (2017). Land Use Land Cover Classification Analysis in Chamarajanagara Taluk, Southern Tip of Karnataka State, India using Geo-informatics. Journal of Environmental Science, Computer Science and Engineering & Technology, 6(3), 209-224. doi: 10.24214/jecet.a.6.3.20924
31. Marcelino, M. W., Gormish, M. J., Bilgin, A. y Boliek, M. P. (2000). An Overview of JPEG-2000. Data Compression Conference Proceedings, 523-541.
32. Martins, F. F. y Cavassan Zaglia, M. (2019). Application of convolutional neural network to pixel-wise classification in deforestation detection using PRODES data. GEOINFO, ¡20 años después!, 57, 57-65.
33. Mateus de Souza, M., se Santiago, V. A., Körting T. S., Leonardi R., de Freitas M. L. (2021). Deep Convolutional Neural Network for Classifying Satellite Images with Heterogeneous Spatial Resolutions. En O. Gervasi et al. (eds.), Computational Science and Its Applications – ICCSA 2021. ICCSA 2021. Lecture Notes in Computer Science, vol. 12955. doi: 10.1007/978-3-030-87007-2_37

34. Muharram, F. W., Ibrahim, F., Setyawan, N., Meissarah, P. y Wulandari, Y. S. (2020). Adquisición de fotografía aérea de pequeño formato de la zona de dunas de arena de Parangtritis utilizando vehículo aéreo no tripulado de ala fija y punto de control terrestre integrado. ACRS 2020 – 41.ª Conferencia Asiática sobre Teledetección, febrero de 2021.
35. Okafor, E., Schomaker, L. y Wiering, M. A. (2018). An analysis of rotation matrix and colour constancy data augmentation in classifying images of animals. Journal of Information and Telecommunication, 2(4), 465-491. doi: 10.1080/24751839.2018.1479932
36. Okafor, E., Smit, R., Schomaker, L. y Wiering, M. (2017). Operational data augmentation in classifying single aerial images of animals. 2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA), 354-360. doi: 10.1109/INISTA.2017.8001185
37. Paiva, P. F. P. R., de Lourdes Pinheiro Ruivo, M., da Silva Júnior, O. M., de Nazaré Martins Maciel, M., Braga, T. G. M., Nogueira de Andrade, M. M., dos Santos Junior, P.C., … Ferreira, B. M. (2019). Deforestation in protect areas in the Amazon: a threat to biodiversity. Biodiversity and Conservation, 29, 19-38. doi: 10.1007/S10531-019-01867-9
38. Podorozhniak, A., Lubchenko, N., Balenko, O., & Zhuikov, D. (2018). Neural network approach for multispectral image processing. 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), 978–981. https://doi.org/10.1109/TCSET.2018.8336357
39. Radovic, M., Adarkwa, O. y Wang, Q. (2017). Object Recognition in Aerial Images Using Convolutional Neural Networks. Journal of Imaging, 3(2). doi: 10.3390/jimaging3020021
40. Rincón-Ruiz, A., Correa, H. L., León, D. O. y Williams, S. (2016). Coca cultivation and crop eradication in Colombia: The challenges of integrating rural reality into effective anti-drug policy. International Journal of Drug Policy, 33, 56-65. https://doi.org/10.1016/j.drugpo.2016.06.011
41. Romero, N., Ríos, J. y Güiza, L. (2020). Desafíos del Estado colombiano en torno al aprovechamiento ilícito de oro y los cultivos de uso ilícito en la Amazonía: estudio de caso de San José del Fragua (Caquetá). Estudios Socio-Jurídicos, 22(2), 291-317.
42. Saavedra, Santiago, The response of illegal mining to revealing its existence (September 29, 2021). Available at SSRN: https://ssrn.com/abstract=3933128 or http://dx.doi.org/10.2139/ssrn.3933128https://doi.org/http://dx.doi.org/10.2139/ssrn.3933128

43. Simonyan, K. y Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 1-14.
44. Stalin, J. L. y Kumar, K. S. (2021). Application of UAV Remote Sensing Technology for Sand Quarry Volumetric Audit and Environmental Impact Assessment: A Case Study Done in Neyvasal Sand Quarry, Cuddalore District, Tamil Nadu—India. Journal of the Indian Society of Remote Sensing, 49, 179-191. doi: 10.1007/s12524-020-01160-1
45. Subcommittee On Western Hemisphere, Transnational Crime, Civilian Security, Democracy, Human Rights, A. G. W. I. (2020). Illicit mining: threats to u.s. National security and international human rights. http://www.govinfo.gov
46. Vashchenko, P. S., & Shavykin, A. A. (2021). Aerial Photography For Assessing The Sensitivity Of The Coast To Oil Spills (On The Example Of The Kola Bay). Natural Volatiles & Essential Oils, 8(4), 5917–5930.
47. Wang, W., Lai, Q., Fu, H., Shen, J., Ling, H. y Yang, R. (2021). Salient Object Detection in the Deep Learning Era: An In-depth Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence. doi: 10.1109/TPAMI.2021.3051099