Hindawi
Complexity
Volume 2019, Article ID 7206096, 11 pages
https://doi.org/10.1155/2019/7206096

Research Article

Secure UAV-Based System to Detect Small Boats Using Neural Networks

Moisés Lodeiro-Santiago,[1] **Pino Caballero-Gil**,[1] **Ricardo Aguasca-Colomo**,[2] **and Cándido Caballero-Gil**[1]

[1]*Departamento de Ingeniería Informática y de Sistemas, Universidad de La Laguna, Tenerife 38200, Spain*
[2]*Instituto Universitario SIANI-Edificio Central del Parque Cientfico y Tecnológico, Campus Universitario de Tafira, 35017 Las Palmas de Gran Canaria, Spain*

Correspondence should be addressed to Moisés Lodeiro-Santiago; mlodeirs@ull.edu.es

Received 17 October 2018; Revised 3 December 2018; Accepted 10 December 2018; Published 2 January 2019

Guest Editor: Magnus Johnsson

Copyright © 2019 Moisés Lodeiro-Santiago et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work presents a system to detect small boats (pateras) to help tackle the problem of this type of perilous immigration. The proposal makes extensive use of emerging technologies like Unmanned Aerial Vehicles (UAV) combined with a top-performing algorithm from the field of artificial intelligence known as Deep Learning through Convolutional Neural Networks. The use of this algorithm improves current detection systems based on image processing through the application of filters thanks to the fact that the network learns to distinguish the aforementioned objects through patterns without depending on where they are located. The main result of the proposal has been a classifier that works in real time, allowing the detection of pateras and people (who may need to be rescued), kilometres away from the coast. This could be very useful for Search and Rescue teams in order to plan a rescue before an emergency occurs. Given the high sensitivity of the managed information, the proposed system includes cryptographic protocols to protect the security of communications.

1. Introduction

According to research in the area of political geography, EU governments are immersed in a difficult battle against irregular migration [1]. This phenomenon was fuelled by the 9/11 attacks and is becoming identified as a "vector of insecurity," so some countries are using it to justify drastic acts of immigration measures [2]. On the other hand, the so-called Transnational Clandestine Actors [3] operate across national borders, evading state laws, becoming rich at the cost of the despair suffered by many people living in "poor countries" and violating their basic human rights. Thus, this scenario leads to catastrophic consequences most times, with innumerable loss of human lives, mainly because of the vulnerability of the means used to travel [4]. Data from the European External Borders Agency FRONTEX [5] indicate that between 2015 and 2016 more than 800,000 people irregularly passed through the Mediterranean to Europe seeking refuge [6]. The number of irregular immigrants who cross the sea increases every year compared to the number of them who do it on foot [7]. To face this situation, the EU has been investing more and more resources in the detection of these flows [5].

Various advances in research and various works have been recently presented that deal with the problem of irregular immigration, [8–12]. These works make use of various image processing techniques for the detection of people and boats in the sea, demonstrating that it is feasible to use technology in combination with UAV systems to face these problems.

This work presents a system to cope with the aforementioned problem, making use of a system based on a UAV for capturing several sequences of images with a smartphone on it. The UAV uses an optimal route planning system such as the one presented in [13] adapted to marine and coastal environments. These images are sent in real time through antennas using LTE/4G coverage to a remote cloud server, where they are processed by a Convolutional Neural Network

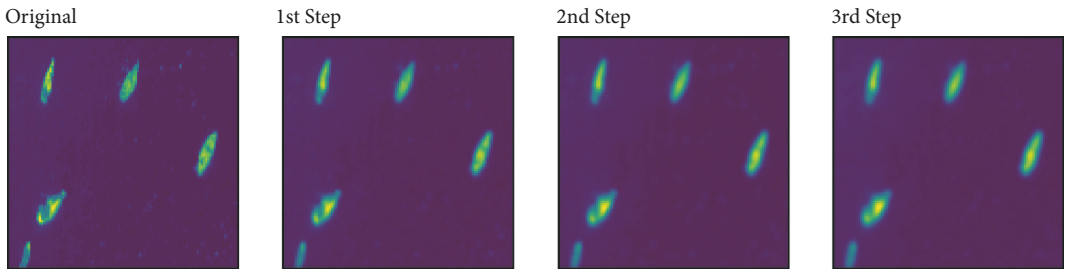


Figure 1: 3-step process on 5x5 kernels for noise removal.

(CNN) that has been previously trained to detect three types of objects: ships, pateras (or cayucos), and people (on land or at sea). These images may be used in a system for the detection and alert of various security and emergency. For this purpose, an Automatic Identification System (AIS) is used to compare each image of a detected shop, according to its GPS position, with a marine traffic database in order to find out whether it is a registered ship or not.

The security of the application against manipulation or attacks is structured in different levels depending on the used technology. For transmission via LTE (4G) in coastal areas with coverage, the SNOW 3G [14] algorithm is used for integrity protection and flow encryption [15]. Furthermore, in order to avoid image manipulation by inserting watermarks that may disable the ability to identify images [16], an algorithm has been designed that first adds white noise to the image and then compresses it using a JPEG compression [17]. This proposal not only prevents the attack but also, according to performed tests, increases the accuracy of the network.

Furthermore, to protect data transmission systems, an Attribute-Based Encryption (ABE) is used. In the bibliography, several proposals can be found that use ABE as a light cryptographic technique to deal with problems different from the one described in this work. On the one hand, in the paper [18], ABE is used to access scalable media where the complete subcontracting process returns plaintext to smartphone users. On the other hand, in [19], ABE is proposed to access health care records using a mobile phone with decryption process outsourced to cloud servers.

The present document is structured as follows: Section 2 discusses the use of neural networks, particularly convolutional ones. Section 3 defines the different stages of the proposed system and some experiments during data collection, training, and obtaining results. Section 4 describes the security layer, with emphasis on possible attacks and countermeasures applied to this type of system. Finally, Section 5 closes the paper with some conclusions.

2. Image Processing

Image processing is the first essential step of the proposed solution to the aforementioned problem. Image processing is a methodology that has been widely applied in the field of research for the identification of objects, tracking of objects, detection of diseases, etc. For many years in the field of artificial intelligence, neural networks have gained strength and, in image processing, the CNN have been used.

CNNs are a type of network created specifically for image and video processing. The relationship between CNNs and neural networks is quite simple because both have the same elements (neurons, weights, and biases). Mainly, the operation in these networks is based on taking the inputs and encoding some properties of the architecture CNNs passing the results from layer to layer in order to obtain classification data.

The special thing about is the mathematical convolution that is applied. A C convolution is a mathematical operation on two functions f and g to produce into a new function that represents the magnitude in which f is a superimposed and a transferred and inverted version of g. For example, the convolution of f and g is denoted by $f * g$ and is defined as the integral of the product of both functions after moving one of them at a distance of t. Thus, the C convolution is defined as $C = (f * g) = \int_{-\infty}^{\infty} f(\eta)g(t - \eta)d\eta$. In CNN, the first argument of the convolution usually refers to the input and the second argument refers to the kernel (a fixed-size matrix with positive or negative numerical coefficients, with an anchor point within the matrix that, as a general rule, is located in the middle of the matrix). The common output of applying a convolution with a kernel is treated as a new feature map H such that $H(x, y) = \sum_{i=0}^{M_i-1} \sum_{j=0}^{M_j-1} I(x + i - a_i, y + j - a_j)K(i, j)$.

Figure 1 illustrates an example of the evolution when applying a 3-step process on a 5x5 kernel with values of 0.04 for removing residual noise. Although at first glance there are no significant differences, the image becomes blurrier as the different steps (from 1 to 3) are applied (this can be seen better on the edges of the pateras).

The layers of compression or pooling layers are applied along the neural network to reduce the space of the representation by making use of a number of parameters and the same computation of the network. This process is applied independently in each step in depth within the network, taking as reference the inputs. It is also used for the reduction of data overfitting. There are different types of pooling although among the most best is the one known as Max-Pooling. In the Max-Pooling process, having as input an array of NxN and MxM grids is taken such that $M \subseteq \{1..N\}$. The resulting number of horizontal and vertical steps will determine the discard threshold for the new layers.

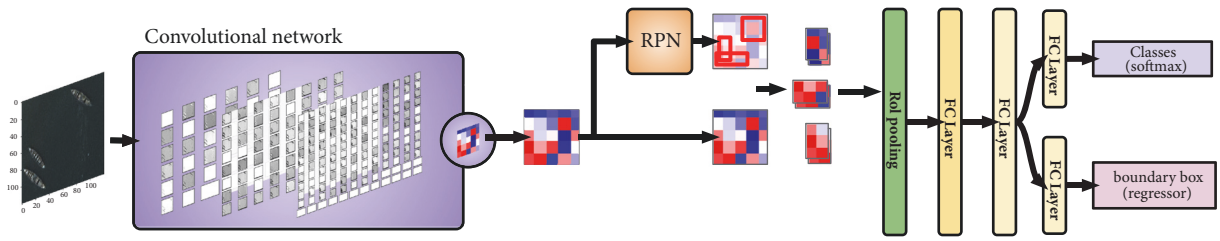


Figure 2: Faster R-CNN flow.

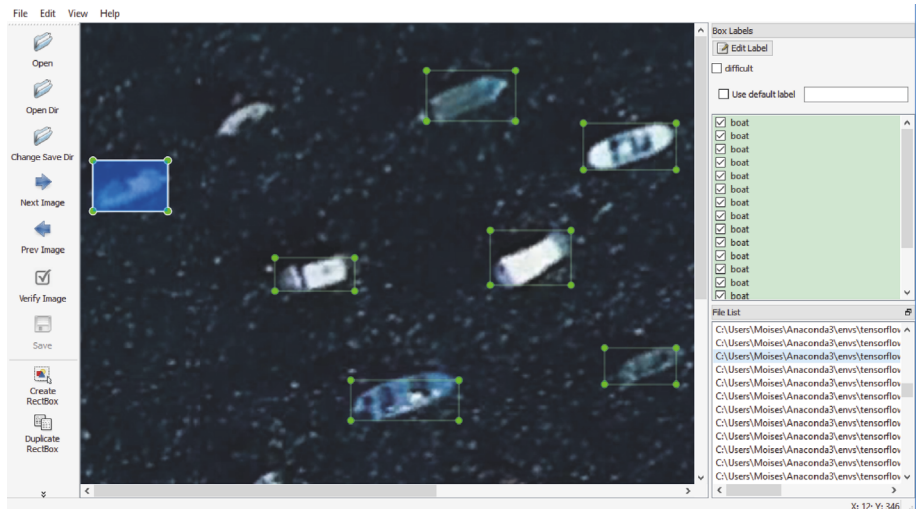


Figure 3: Classification tool.

In order to get the model used in this work performance improvements, modern object detector based on CNNs knowing as Faster R-CNN [20, 21]. This model depends in part on an external region used for selective search [22]. The Faster R-CNN model has a design similar to that of Fast R-CNN [23], so that it jointly optimises classification and bounding box regression task. Moreover, the proposed region is replaced by a deep learning network and the Region of Interest (RoI) is replaced by features maps. Thus, the new Region Proposal Network (RPN) is more efficient for the generation of RoIs because for every window location, multiple possible regions are generated based on a bounding box ratio. In other words, a visualisation is made on each location in the characteristics map, considering a number k of different boxes centred on it (a longer area, a fatter one, a longer one, etc.). This is shown in Figure 2 in an example, where a softmax classifier composes a Fully Connected (FC) Layer.

3. Proposed System

This section describes the procedures performed after the acquisition of the data, explaining the processing of the images as well as the detailed training process, providing information on each of the obtained results and discussing why to choose one or other result to continue the experiments. Finally, some conclusions results are provided through a demonstration image with correct classification ratios.

3.1. Data Collection and Classification. For the collection of images of pateras, due to the lack of accessibility to boats of the type patera or cayuco in a massive way, we have opted for the gathering of information from of an image, by using Google Earth software, always looking for a height with respect to sea level of 100 meters (height at which the drone would fly in the experiment) and maintaining a totally perpendicular view. After obtaining a dataset of 3,347 images corresponding to three classes of the problem, we opted for a classification of each of the objects of the various images, taking into consideration that an image can have one or several objects. According to the applied dataset, complexity, and capabilities and based on the available documentation, the majority of the authors refer to the Pareto Principle [24] as the most convenient. Thus, the ratio 80/20, which is the most used has been considered an adequate proportion for the train/test neural network.

For the classification of each image, we have used the software known as LabelImg (see Figure 3), created in Python, for supervised training. This software creates a layer that separates each image into different objects limited by a bounded box corresponding to the position (x, y), width, and height of the box. This process has been performed for the three types of object considered in this work.

As a result of this classification, the XML files corresponding to all the objects within the images are obtained. These XML files are used later in the training.

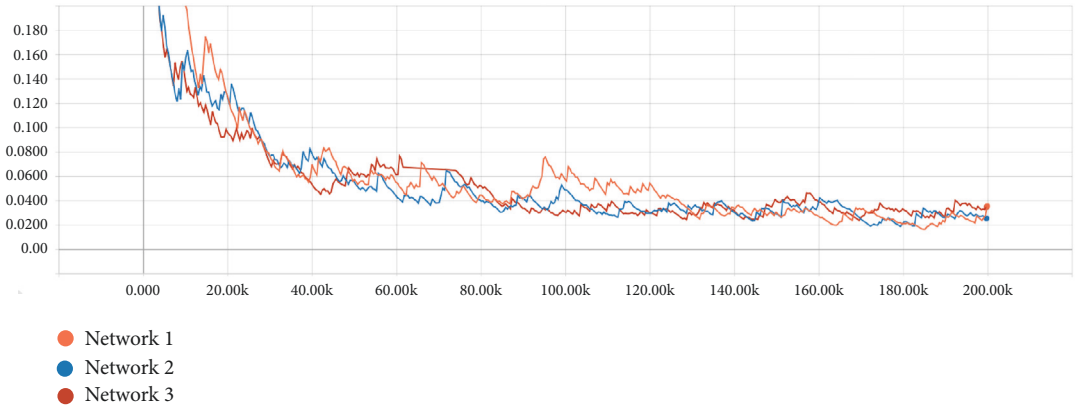


Figure 4: First trainings of CNN maintaining coefficients with different datasets.

Table 1: Object distribution for imbalanced networks.

Classes	Person	Patera	Boat
Train	1791	284	646
Test	466	96	64

Table 2: Classification Loss for Networks 1, 2, and 3.

Network	Classification Loss
1	0.03553
2	0.02554
3	0.03494

3.2. Training. The total time for each training stage of the neural network with the conditions described above has been an average of 16 hours using a GPU Nvidia 1050. The following guidelines were followed in order to obtain the best possible network for detection:

(i) Train the same neural network three times with the same learning coefficients, regulation coefficients and activation function. The reason for doing only three trainings instead of 5, 10, 30, or more is because there is no rule of thumb that shows an exact trend in the result of the network. Therefore, to rule out strange behaviours, each network was trained 3 times to see empirically that the three results (even starting from a random vector in the direction of the gradient descent) gave similar results. With this, false positives can be discarded when compared with other networks.

(ii) In all training sessions, a number of 200.0K iterations was established for each of the networks.

(iii) The reflected values have a tendency of 0.95, which means that they are not real values but that is the value of tendency in each instant calculated from previous values (so it can be higher or lower).

(iv) In each training, the network changed the dataset on the basis that the dataset has a total of 3347 images divided approximately between a ratio of 80% training and 20% for testing resulting in a distribution as shown in Table 1. For each training of each neural network, the initial set of training and test has been altered to demonstrate the efficiency of the neural network from different datasets. This type of randomness has been applied to demonstrate the functionality and efficiency of the system in methods based on stochastic decisions. Given that the applied methodology is stochastic and random, performing permutations on the dataset allows obtaining different results, which is used to obtain better datasets to be used as a basis for other trainings.

(v) The used programming language was Python 3 for the machine learning, and the TensorFlow software library for the neural network oriented environment [25].

The use of tools such as TensorFlow (among other frameworks for analysis in convolutional networks) has been widely used in recent years for the detection of patterns in images. One of the most notable current works is its use in medical environments to face deadly diseases such as cancer [26], which slightly improves the performance obtained by specialists in dermatology. Among others, neural networks have been used in the marine environment [27] to identify marine fouling using the same framework. Although according to several studies [28, 29] the use of unbalanced datasets in neural networks is detrimental, we have opted for an approach to a real problem where a balanced data network is not available. At the end of this section, a comparison was be made with a balanced network. It can be appreciated that although the results of the balanced network are better, they do not differ too much.

Once the three neural networks finished training (see Figure 4) the final results shown in Table 2 were obtained based on the Classification Loss (CL). The CL equation [21] is optimised for a multitask loss function and represented as $L(\{p_i\},\{t_i\}) = (1/Ncls)\sum_i L_{cls}(p_i, p_i^*) + \lambda(1/Nreg)\sum_i L_{reg}(t_i, t_i^*)$, where the terms of the equation represent the loss in classification over two classes (depending on whether the object exists or not), while the second term is

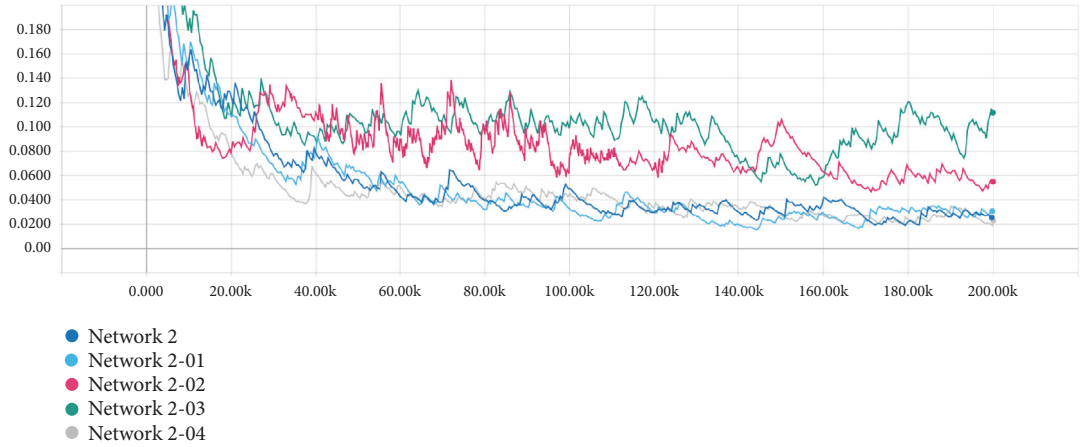


Figure 5: Sensitivity analysis based on data from Network 2.

Table 3: Learning rates for new trainings using dataset from Network 2.

Network	Learning Rate
2 - 01	0.003
2 - 02	0.00003
2 - 03	0.00001
2 - 04	0.001

the loss of regression of the bounding boxes where an object is found.

When interpreting Figure 4, it must be taken into consideration that the used parameter was the CL. This value is better the closer it gets to zero. Initially, the graph starts with discrete values between 0 and 1, where 1 is a total loss and 0 is a no loss.

Considering these results, the next step in obtaining an improved network was to copy the data from the best network (number 2) and perform a sensitivity analysis on 4 new trainings varying the training coefficients. In the image of Figure 5 we can see the behaviour of the different networks (including the original network number 2). The variation of the coefficients was of multiplicative type, altering the different coefficients of learning rate according to the distribution shown in Table 3.

The learning rate is a measure that represents the size of the vector that is applied in the descent of the gradient when applying the partial derivatives. On the one hand, if the learning rate is very large, the steps will be larger and will approach a solution faster. However, this can be a mistake because it could jump without coming to a good approximation to the solution. On the other hand, if it is very small, it will take longer to train but it will come up with a solution. That is why the study was carried out with different learning rates, to check which learning rates come closest to a good solution in less time. Thus, using these results, the best final coefficient (see Table 4) with respect to the classification loss was the Network 2-04 (grey line in Figure 5), with a

Table 4: CL for Networks 1, 2, 3, and 4 trained with original Network 2.

Network	Classification Loss
2 (original)	0.02554
2 - 01	0.03075
2 - 02	0.05487
2 - 03	0.06422
2 - 04	0.02310

Table 5: CL for Networks 1 and 3 using the parameters of best Network 2-04.

Network	Classification Loss
2 - 04 (best)	0.02310
1 with 0204 coefficients	0.02559
3 with 0204 coefficients	0.03043

Table 6: Balanced training and test dataset.

Classes	Person	Patera	Boat
Train	300	300	300
Test	80	80	80

coefficient lower than 0.02310 which means that in 97.8% cases it produces correct classifications.

Afterwards, the best parameters of the best Network (2-04) were exported to the initial sets of networks 1 and 3 to see if a better result could be obtained by applying the coefficients of the best network so far (see Figure 6). In that image we can see how, although for a short time, the best network is still Network 2 with the fourth training (2-04). However, with these parameters, Networks 1 and 3 (1-0204 and 3-0204) improve slightly with respect to the initial Networks 1 and 3 values (see Table 5 and Figure 7).

After having obtained a result that is feasible in terms of experiments, we decided to make an analysis on a neural network with balanced data (80% training and 20% test), this time having the following random distribution of images (see Table 6).

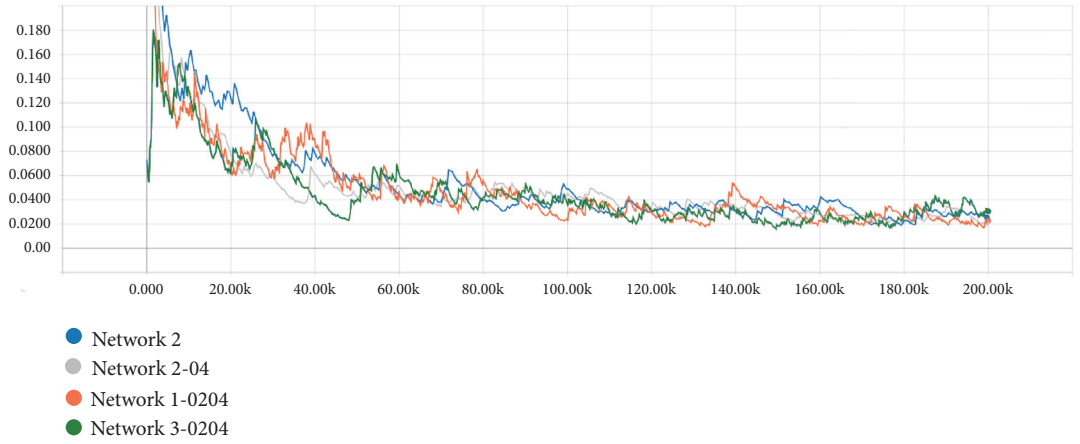


Figure 6: Sensitivity analysis of Networks 1 and 3 with data from Network 2-4.

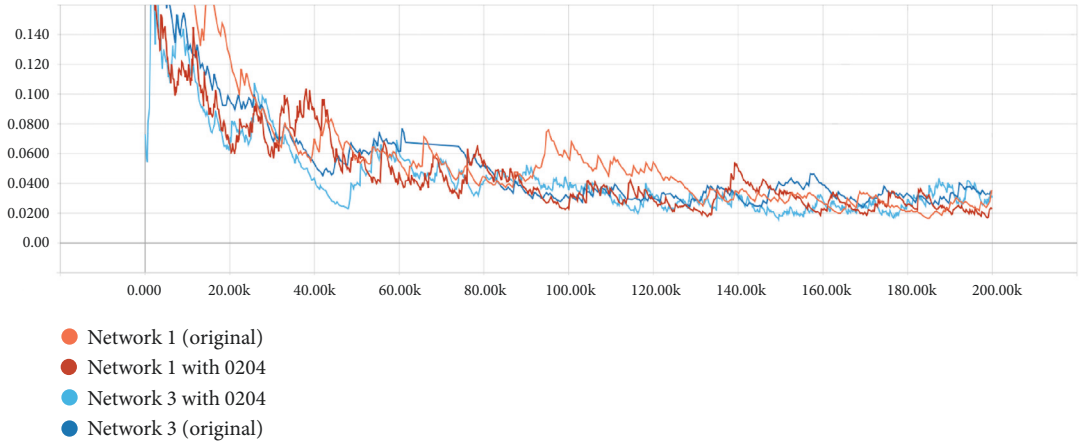


Figure 7: Comparison of original Networks 1 and 3 with respect to Network 2-04.

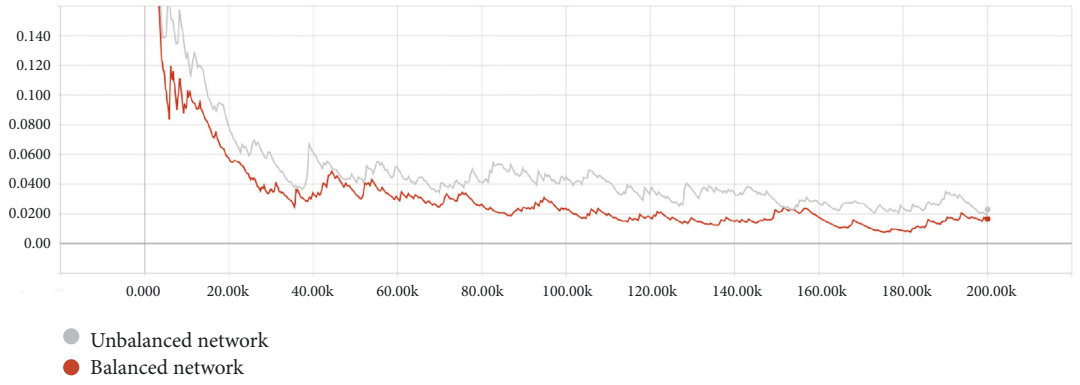


Figure 8: Balanced and unbalanced network.

After hours of training with the balanced network, we got a better result than with what had been the best detection network until now (see Figure 8). The results shown in Table 7 mean that the network trains well with these training coefficients and it even improves the results with a balanced network type (although in a real environment it is difficult to find it).

3.3. Results. In order to check the efficiency of the best neural network obtained in the previous section, different random frames have been extracted from a video showing different scenarios where pateras and people are seen from a real drone (see Figure 9). It should be noted that all these frames have never been previously seen by the neural network (not even in the testing stage), but are completely new to the neural

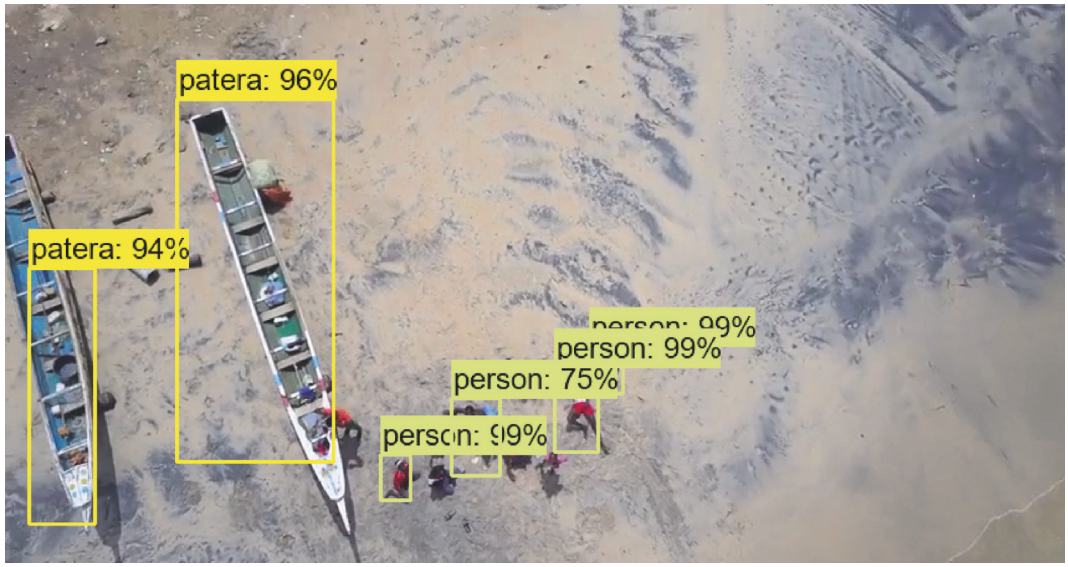


Figure 9: Image detection test from validation dataset.

Table 7: CL for Network 2-04 compared to the new balanced network.

Network	Classification Loss
Unbalanced Network 2-04	0.02310
Balanced network	0.01658

network. This set of frames is known as validation dataset. On the one hand, as a main result, the proposal produces a correct classification of boats and pateras between 94 and 96 % (although these ratios can vary from 92 to 99 % depending on the frame). On the other hand, a correct classification index for people has been obtained, which is around 98-99 %, although, in certain frames (a video has thousands of frames), this ratio can drop to 73 % due to interference with other objects in the video and the environment.

From the obtained results, we conclude that the defined procedure based on the Faster R-CNN proposed for training can be successfully used to detect boats, people and pateras.

4. Security

In a system, like the defined above whose the results can be the difference between saving a human being saving or not it is essential to have the appropriate mechanisms to ensure that the information is not modified or accessed by illegitimate parties. It is for this reason, a study of possible attack vectors related to neural networks for image detection and problems in wireless communications has been performed, paying special attention in adversarial and Man in the Middle attacks.

4.1. Adversarial Attacks. Neural networks are one of the most powerful technological algorithms in the field of artificial intelligence. Among the various networks we can find some specifically oriented to image detection (as seen throughout this work). Sometimes, the simple behaviour of a network fed with inputs (pixel's images) where the output is a type of classification can lead to error, so that it can be inferred that the network does not act correctly. An adversarial attack [30] is a type of attack within the rising field of artificial intelligence consisting in introducing an imperceptible perturbation that leads to an increased probability of taking the worst possible action.

In the case analysed in this work, this attack involves using a type of images that can be supplied to the network that though represent a certain type of object (for example a ship), for the network they mean something else (like a dog, a toaster.).

In environments where there are thousands or millions of types of classes and classifications it could be a problem. That is the case, for example of Google's Inception V3 [31], could be used to alter the driving of an autonomous vehicle that uses this type of network for altering the images of its environment by applying stickers [16] on traffic signs for the purpose of changing the maximum speed in a road.

The way in which this type of attacks act is through the excitation of the neural network inputs through the inclusion of new figures or noise (generally not perceptible to the human eye) making modifications in the input image (with gradient descent and back propagation techniques) making the network suffer something similar to an optical illusion.

The answer to the question of what this type of attack is looking for is how to maximise the error that can be achieved by entering erroneous information. That is to say, to do the opposite that the neural network expects to do this to minimise the error with the input parameters, all this, taking into account the fact that a formula must be applied to minimise the difference between the added disturbance and the original image with respect to the human eye.

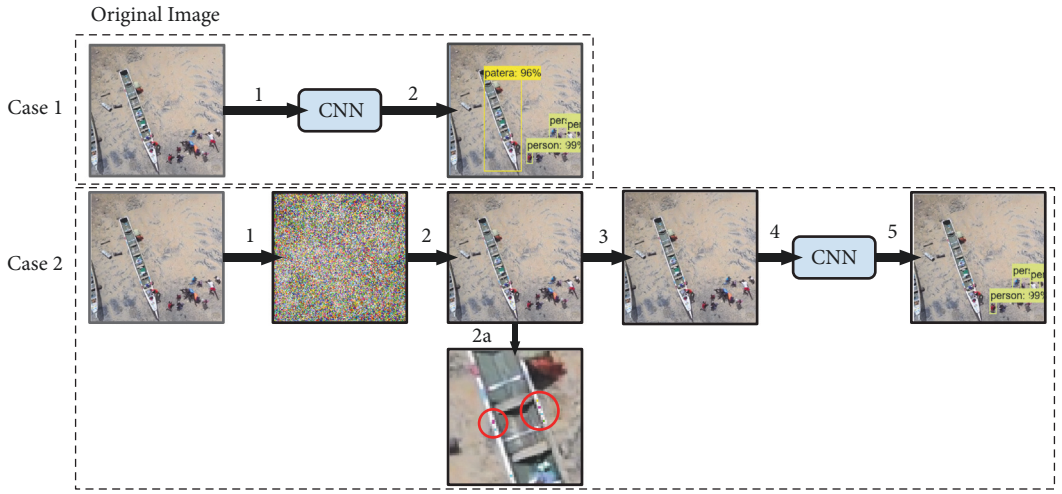


Figure 10: Adversarial attack proof of concept.

In the neural network that has been presented in this work, the number of classes has been limited to classify a total of 3 types of objects (ships, pateras, and people) so the margin of error within the possible classification could mean a sort of security system against this type of attack. Because of this, it can be said that, in a controlled environment, this type of attack would have no effect on the proposed system.

However, as a proof of concept, an adversary attack has been created that could modify the behaviour of our network. To do this, we have taken a random frame from a video sequence where we can see a whole patera, a part of another one and people (who could be castaways) in the sand. In Figure 10 it is possible to appreciate two main cases:

Case 1.

(1) Starting from the frame extracted from the video, it has been processed directly by our neural network.

(2) As a result, we have been obtained a detection of the patera with an index of 0.96 and of the people with an index variant between 0.98 and 0.99.

Case 2.

(1) Based on the same starting image seen in Case 1, training has been carried out with a different neural network to the original one. With this, we demonstrate that adversarial attacks also fulfill a transition property that can affect other networks. The result of this step is the generation of an image with noise. The noise shown in the image has been modified by enlarging the brightness of the image in 10 steps because the original was a black image with little visible noise.

(2) By applying the original noise to the initial image, a new resulting image is obtained that, with the naked eye, as can be seen in the image 2a of Figure 10, it has some pixels different from the original image.

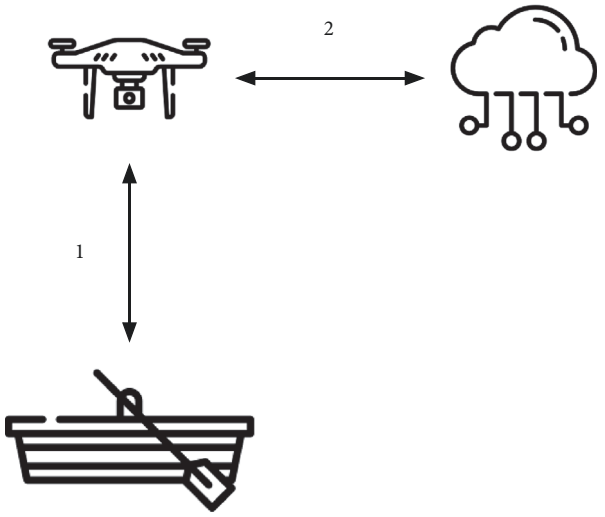


Figure 11: Attacking vectors.

(3) To soften the effect appreciated in point 2, a series of mathematical operations are applied to each pixel to soften the textures and obtain a finished image.

(4) The image generated in step 3 is sent to the neuronal network for the detection of pateras.

(5) Finally it can be seen that by applying this new image, which at first sight is the same as the original, the system does not detect the patera.

4.2. Attack Vectors. In a possible scenario where an attacker wants to bypass the security measures that have been implemented, he/she could follow one of the following two ways (see Figure 11).

(1) As discussed in the previous section, there is a type of attack called an adversarial attack that is designed

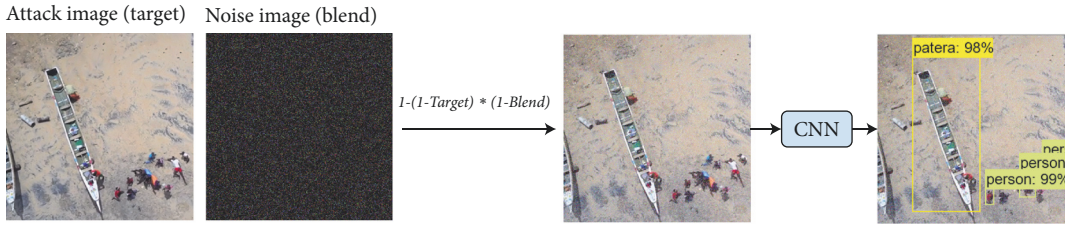


Figure 12: Random Gaussian noise blending.

to confuse the neural network. The aforementioned technique that has been used in a real environment of the inclusion of stickers [16] could be applied to the pateras in order to avoid the drone control by pretending the patera look like an unrecognised object or other object. Among the possible countermeasures to mitigate the attack, we have

(i) JPEG compression method: This method is based on the hypothesis that the input image (i.e., the one taken by the drone) can be manipulated by the aforementioned attack so that the generated image has a noise that confuses the network. For the removal of this malicious noise it is possible to go for an 85% compression using a JPEG compression format [17] that will make the embedded noise blur, while maintaining the basic characteristics of shape in the image.

(ii) Noise Inclusion: The drone could have a simple internal image manipulation system to apply a Gaussian random noise so that the noise is imperceptible in the image before being sent to a server for processing. To do this we use an image of noise previously generated (or created in the moment) and then apply the formula of the blending method known as "screen" described with the formula: $1 - (1 - Target) * (1 - Blend)$. The advantage of this compared to the method described above is that the loss of image quality is not affected (depending on the weight and size of the noise). However, it could include a slightly visible noise (see Figure 12).

(2) Man in the Middle attack (MITM) is a sort of attack where an attacker is placed between sender and receiver. In this case the sender would be the drone and the receiver the server that will do the image processing through the neural network. The communication media can vary depending on the coverage in the area of emission. It is always a wireless connection like 2, 3, 4, or even 5G. In this case, the attacker can intercept the signal with the image in order to modify it on the fly including the necessary noise to make the image undetectable. To deal with such attacks, the system protects the security of the communication system through the cryptographic scheme described in the following section.

4.3. Attribute-Based Encryption. In the proposal described in this paper, an encryption is used to protect from unauthorised attackers the confidentiality of the database of the images captured from a smartphone on a UAV, which are labelled with the date when the image was taken, the GPS location of the photograph along with other selected metadata. Smartphones are less powerful than other systems in computations such as image transmission, key generation, and information storage and encryption. In order to reduce the overload of the security protocol, we propose the use of a light cryptographic technique. In addition, to offer the remote server the ability to securely examine all the images captured by UAVs in a region, an Attribute-Based Encryption is proposed. This is a type of public-key encryption in which private keys and encrypted texts depend on certain attributes, and decryption of encrypted text is only accessible to users with the satisfactory attribute configuration. In the proposal described in this document, the used attributes are related to date/time, geopositioning location, linked UAV, etc., so that the private key used in the remote server is restricted to be able to decipher encrypted texts whose attributes coincide with the policy of attributes linked to the UAVs it controls. This private key can be used to decrypt any encrypted text whose attributes match this policy but have no value in deciphering others. This means that each operator in a remote server has a set of UAVs assigned to him/her, so the images captured by any UAV cannot be decrypted either by an unauthorised attacker or by a server operator unrelated to that UAV. Since the used encryption is public-key encryption, its security is based on a mathematically hard problem, and security holds even if an attacker manages to corrupt the storage and obtain any encrypted text. The operations associated with the proposal involve the following phases.

(1) Setup phase: this phase is where the algorithm takes the implicit security parameter to generate the Public Key (PuK) and Master Key (MaK).

(2) KeyGen phase: in this phase, a trusted part generates a Transformation Key (TrK) and Private Key (PrK) linked to the smartphone, which are used to decrypt the information sent from it.

(3) Encrypt Phase: in this phase, the smartphone encrypts the image using PuK and MaK before sending it to the remote server.

(4) Transformation phase: this phase is where the remote server performs a partial decryption operation of the

encrypted data using TrK to transform the encrypted text into a simple encrypted text (partially decrypted) before sending it to the operators. If the operator's attributes satisfy the access structure associated with the encrypted text, he/she can use the decryption phase to retrieve the plaintext from the transformed ciphertext.

(5) Decryption phase: as the transformation phase transforms the encrypted text into a simple encryption, finally, the server operator uses this phase to retrieve the plaintext of the transformed ciphertext, using the PrK.

5. Conclusions

In this work, a novel proposal has been defined to provide a solution to the problem of the detection of small boats, which are used many times by irregular immigration. For this purpose, a Convolutional Neuronal Network has been created, specifically trained for the detection of three types of objects: boats, people and pateras. This system is used in coordination with a UAV that sends the signals via wireless connection (LTE) to a server that will be responsible for processing the image in the neural network and detecting if it is an anomalous situation. This work describes and includes several security systems that allow us to guarantee the stability of the data so that they cannot be altered either before or after being sent. As a complement to protect data transmission systems using the ABE algorithm, a novel mechanism has been implemented to mitigate adversarial attacks by overlapping Gaussian noise to the possible attacking image noise. In addition, to discard false positives, a compendium of the GPS coordinates of the UAV is made with an AIS system of geolocalised ships. The main contribution is a light neural network with a high rate of detection of objects (reaching up to 99% accuracy), which would be a great help for Search And Rescue or border patrol teams in case of having to perform a rescue. A study with thousands of frames could be done to see the detection ratio and the accuracy of each object, to determine which object is better detected.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Research was supported by the Spanish National Cybersecurity Institute (INCIBE) under the INCIBEC-2015-02492 and INCIBEI-2015-27338 grants and by the Spanish Ministry of Economy and Competitiveness, the FEDER Fund, and the CajaCanarias Foundation, under TEC2014-54110-R and DIG02-INSITU Projects. The financing granted to the ULL by the Ministry of Economy, Industry, Commerce and Knowledge, co-financed by the European Social Fund by 85%, is gratefully acknowledged.

References

[1] H. van Houtum, "Human blacklisting: The global apartheid of the EU's external border regime," *Environment and Planning D: Society and Space*, vol. 28, no. 6, pp. 957–976, 2010.

[2] L. Vives, "Unwanted sea migrants across the EU border: The Canary Islands," *Political Geography*, vol. 61, pp. 181–192, 2017.

[3] P. Andreas, "Redrawing the Line: Borders and Security in the Twenty-first Century," *International Security*, vol. 28, no. 2, pp. 78–111, 2003.

[4] F. J. de Lucas Martn, "Muertes en el mediterráneo: inmigrantes y refugiados, de infrasujetos de derecho a amenazas para la seguridad," *Quaderns de la Mediterrània= Cuadernos del Mediterráneo*, vol. 22, pp. 272–277, 2015.

[5] Frontex, *Frontex - European Border And Coast Guard Agency*, 2018, https://frontex.europa.eu/.

[6] UNHCR, *The Un Refugee Agency*, 2018, http://www.unhcr.org/.

[7] P. Soddu, *Ceuta and melilla. security, human rights and frontier control*, Institut Europeu de la Mediterrània (eds) IEMED Mediterranean Yearbook Med, pp. 212–214, 2006.

[8] M. Díaz-Cabrera, J. Cabrera-Gámez, R. Aguasca-Colomo, and K. Miatliuk, "Photogrammetric analysis of images acquired by an uav," *International Conference on Computer Aided Systems Theory*, pp. 109–116, Springer, 2013.

[9] A. M. Klimkowska and I. Lee, "A prealiminary study of ship detection from UAV images based on color space conversion and image segmentation," in *Proceedings of the 4th ISPRS International Conference on Unmanned Aerial Vehicles in Geomatics, UAV-g 2017*, pp. 189–193, Germany, September 2017.

[10] T. Giitsidis, E. G. Karakasis, A. Gasteratos, and G. C. Sirakoulis, "Human and fire detection from high altitude UAV images," in *Proceedings of the 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, PDP 2015*, pp. 309–315, Finland, March 2015.

[11] S. Freitas, C. Almeida, H. Silva, J. Almeida, and E. Silva, "Supervised classification for hyperspectral imaging in UAV maritime target detection," in *Proceedings of the 2018 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pp. 84–90, Torres Vedras, April 2018.

[12] W. Huo, Y. Huang, J. Pei, Q. Zhang, Q. Gu, and J. Yang, "Ship Detection from Ocean SAR Image Based on Local Contrast Variance Weighted Information Entropy," *Sensors*, vol. 18, no. 4, p. 1196, 2018.

[13] V. San Juan, M. Santos, and J. M. Andújar, "Intelligent UAV Map Generation and Discrete Path Planning for Search and Rescue Operations," *Complexity*, vol. 2018, Article ID 6879419, 17 pages, 2018.

[14] J. Molina-Gil, P. Caballero-Gil, C. Caballero-Gil, and A. Fúster-Sabater, "Software implementation of the SNOW 3G generator on iOS and Android platforms," *Logic Journal of the IGPL. Interest Group in Pure and Applied Logics*, vol. 24, no. 1, pp. 29–41, 2016.

[15] I. Santos-González, A. Rivero-García, P. Caballero-Gil, and C. Hernández-Goya, "Alternative communication system for emergency situations," in *Proceedings of the 10th International Conference on Web Information Systems and Technologies, WEBIST 2014*, pp. 397–402, Spain, April 2014.

[16] K. Eykholt, I. Evtimov, E. Fernandes et al., "Robust physical-world attacks on deep learning visual classification," in *Proceedings of the in IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1625–1634, 2018.

[17] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, 1992.

[18] K. V. C. Ganesan, "Healthcare monitoring solution with decryption outsourcing by parallel computing in cloud," *Innovative Research in Computer and Communication Engineering*, vol. 2, pp. 56–64, 2014.

[19] S. Yu, *Data sharing on untrusted storage with attribute-based encryption [Ph.D. thesis]*, Worcester Polytechnic Institute, 2010.

[20] J. Huang, V. Rathod, C. Sun et al., "Speed/accuracy trade-offs for modern convolutional object detectors," in *Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, pp. 3296–3305, USA, July 2017.

[21] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems*, pp. 91–99, 2015.

[22] K. E. A. Van De Sande, J. R. R. Uijlings, T. Gevers, and A. W. M. Smeulders, "Segmentation as selective search for object recognition," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '11)*, pp. 1879–1886, November 2011.

[23] R. Girshick, "Fast R-CNN," in *Proceedings of the 15th IEEE International Conference on Computer Vision (ICCV '15)*, pp. 1440–1448, December 2015.

[24] V. Pareto, *Cours d'économie politique*, Librairie Droz, Second edition, 1964.

[25] M. Abadi, P. Barham, J. Chen et al., "Tensorflow: a system for large-scale machine learning," in *Proceedings of the in Symposium on Operating Systems Design and Implementation*, vol. 16, pp. 265–283, 2016.

[26] A. Esteva, B. Kuprel, R. A. Novoa et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

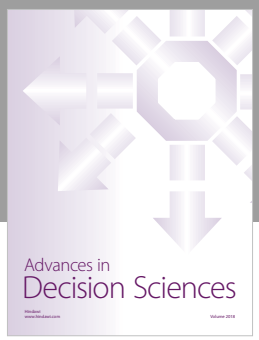
[27] C. S. Chin, J. T. Si, A. S. Clare, and M. Ma, "Intelligent Image Recognition System for Marine Fouling Using Softmax Transfer Learning and Deep Convolutional Neural Networks," *Complexity*, vol. 2017, Article ID 5730419, 9 pages, 2017.

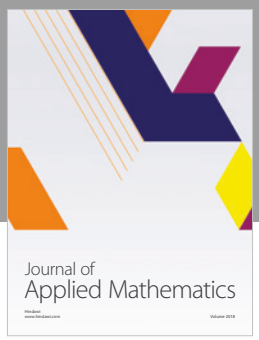
[28] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," *Neural Networks*, vol. 106, pp. 249–259, 2018.

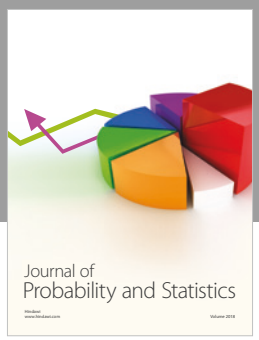
[29] L. Mathews and G. Steri, "Learning from imbalanced data," in *Encyclopedia of Information Science andTechnology*, pp. 1825–1834, IGI Global, Fourth edition, 2018.

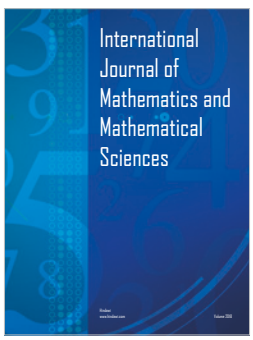
[30] T. B. Brown, D. Mané, A. Roy, M. Abadi, J. Gilmer, and D. Mané, "Adversarial patch," *Computer Vision and Pattern Recognition*, 2017.

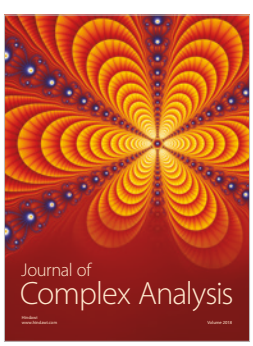
[31] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, pp. 2818–2826, July 2016.

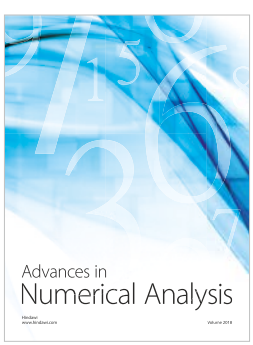


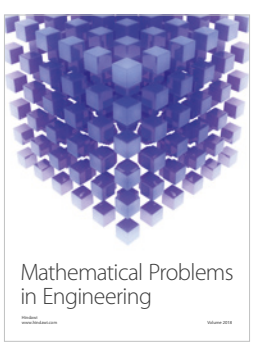


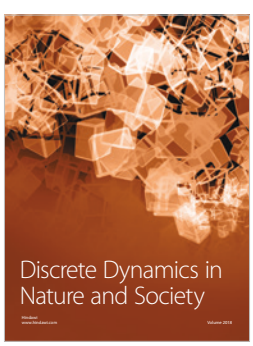


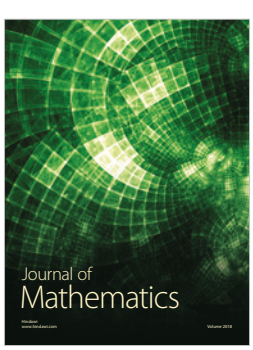


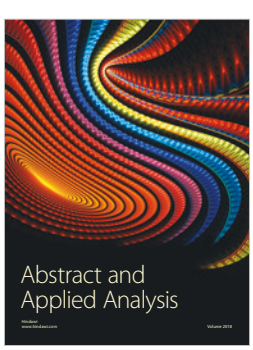


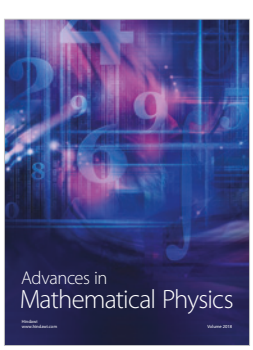












Hindawi

Submit your manuscripts at
www.hindawi.com