

Design of an Artificial Neural Network to Detect Obstacles on Highways through the Flight of an UAV

Dariel A. Islas Guzmán^{1,2}, J. Rodrigo Córdova Alarcón¹,
Adrián Alcántar Torres², Mario A. Mendoza Bárcenas¹

¹ Instituto Politécnico Nacional, Centro de Desarrollo Aeroespacial,
Mexico

² Instituto Politécnico Nacional, Escuela Superior de Física y Matemáticas,
Mexico

dariel.islas@hotmail.com, jcordovaa@ipn.mx,
adrian.alcantar.torres@gmail.com, mmendozab@ipn.mx

Abstract. Due to several risks that involve travelling in highways, such as vehicle collision, natural disasters and other obstacles, the need of implementing an early warning system which detects obstacles to provide security has been growing in the last years. Artificial neural networks based systems have been successfully applied in obstacle detection through image processing and recognition. To address this issue, a Multilayer Feed-forward Network (MFN) was designed to detect obstacles in highways from a zenith perspective. *Backpropagation* algorithm is used to supervise the training of the proposed neural network by minimizing the square error function via descending gradient criteria. Performance analysis was performed by using binarized, grayscale and RGB images with constant size and considering several contexts. The proposed neural network was trained and tested offline. In a next stage of the project, this neural network will be implemented into the on-board computer of unmanned aerial vehicles (UAV's) as an early warning platform.

Keywords: Artificial neural network, image recognition, backpropagation, obstacle detection.

1 Introduction

The improvement of highway driving safety by reducing the number and severity of incidents has become a great challenge in recent years [9]. Obstacles on highways such as overturned cars, mudslides, landslides, lack of road sections and fallen trees, represent a danger for drivers.

Nowadays vehicles have been equipped with technological tools, which are able to determine the presence of an obstacle on the road and thus transmit a warning to the road users, with a certain degree of reliability. For example

in [15,11,3] the obstacle detection is performed through stereoscopic vision. Other obstacle detection systems use radar-based methods such as ClearWayTM, developed by Navtech Radar. Also, autonomous and semi-autonomous land vehicle systems have been developed to accomplish this objective [20,5,13].

Recently the UAV's are affordable platforms for the implementation of several applications and opens up opportunities to develop an efficient and effective obstacle detection system for highways. Obstacle detection task performed during the flight of an UAV can be raised as a problem of pattern recognition, which is one of the main topics in computational vision and artificial intelligence. In order to solve the pattern recognition problem, the development of an algorithm which assigns any image in a representative class is performed [4,16].

There are a large number of papers which present several types of obstacle detection methods. Stereo vision, laser radar and computer vision are the most common methods of obstacle detection [23]. In [17], a Multi-Layer Perceptron (MLP) to detect traffic incidents from changes of quantitative data (traffic quantity per minute, velocity, occupancy rate) was proposed. They confirmed its effectiveness by adding enough combinations of similar average change patterns in order to increase the recognition rate and reduce the number of false detections. In [23] a Convolutional Neural Network it was proposed to detect contours of obstacles in streets. In [13] an unmanned ground vehicle to detect obstacles during off-road autonomous navigation by stereo vision was presented. Referring to [14], an obstacle detection system based on vision by the flight of a UAV has been implemented by segmenting sky and non-sky images and finding obstacles in segmenting sky afterwards.

The images for the detection of obstacles in previous work, are from an aerial perspective, while the problem from an overhead perspective has not been addressed. Therefore, we refer to an obstacle as any visible object in a highway from zenith perspective in this work in the next paragraphs.

In this work, the development of a neural network based system for detecting obstacles on highways is presented. We designed and analyzed the performance of the neural network by modifying its topology and using three different types of images: binarized, gray scaled and RGB images. With the obtained results, an obstacle detection system can be implemented into the on-board computer of an UAV by pattern recognition, which is being developed in our institute, seeking high degree of reliability, low operating cost and flexibility for use in any type of computational infrastructure for the prevention of highways accidents.

2 Artificial Neural Networks and Pattern Recognition

Pattern recognition is the mechanism of relating things and classifying them into different categories by taking and explaining decisions [21]. To solve pattern recognition problem, several approaches has been developed and implemented into a great variety of applications [2]. In this work, neural networks are used as our approach.

The neural network model is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use [8,7].

The information-processing unit of a neural network is called neuron, each neuron is formed by a vector of signals inputs $x = [x_1, x_2, \dots, x_n]^T$, a vector of weights $w = [w_1, w_2, \dots, w_n]^T$ and a nonlinear mapping $f : \mathbb{R} \mapsto [0, 1]$ called transfer function, that is non-decreasing, such that $\lim_{x \rightarrow \infty} f(x) = 1$ and $\lim_{x \rightarrow -\infty} f(x) = 0$ (e.g. Sigmoid Function). A neuron k is described mathematically by the following equations.

$$y_k = f(u_k + b_k), \quad u_k = \sum_{j=1}^m w_{kj}x_j, \quad (1)$$

where y_k is the response of the k -th neuron due an input signal x , f is the transfer function and define the output of the neuron, u_k is a linear combination between input signal and synaptic weights taking m units having connections to j -th unit, b_k is the bias, whose value allows to control the behavior of the layer, and contribute for a successful learning.

Learning is a process by which the free parameters of a neural network are adapted through a process of simulation by the environment in which the network is embedded.

Considering $j = 1, \dots, k$ neurons outputs associated with each input signal, the error signal at the output of j -th neuron at n iteration is defined by $e_j(n) = d_j(n) - y_j(n)$, where $d_j(n)$ is the desired output for j -th neuron and $y_j(n)$ is the output of j -th neuron at n iteration given by equation 1.

The total error energy obtained by the sum of all neurons in the output layer is defined by

$$E(n) = \frac{1}{2} \sum_j (e_j(n))^2. \quad (2)$$

By the training process, the network parameters are adjusted under the combined influence of the training vector and the error signal. The adjustment of the parameters is carried out iteratively until the neural network emulate the training vector. The *backpropagation* algorithm is efficient for training MLP with training vector, the weights are updated according of the minimization of the total error energy (Eq.2). Then, each weight value is softly adjusted in the direction of the negative of the gradient according with the following equation

$$\Delta W_{ki} = -\eta \frac{\partial E}{\partial W_{ki}} \quad \Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}}, \quad (3)$$

where η is a proportionality constant called the learning-rate, the smaller the learning-rate, the smaller the changes to the weight from one iteration to the next. With the aim to accelerate the learning process, a momentum term α is incorporated in updating weight, the momentum term keep the moving in a certain direction, adding a fraction of the previous change to the new one as follows $\alpha \Delta w_{ki}(n) = \alpha(w_{ki}(n) - w_{ki}(n-1))$.

In the case of single hidden layer, W_{ki} denotes the connection of the i -th neuron in input layer to k -th neuron of the hidden layer, and similarly w_{jk} denotes the connection of the k -th neuron to the j -th neuron which belongs to the output layer.

For measuring network efficiency, we consider the Mean Square Error (MSE) criteria, which represents the cost function as a measure of learning performance for a given training set, MSE is updated at each epoch and is defined by the equation (4), where is considered N training patterns:

$$MSE = \frac{1}{N} \sum_{p=1}^N E(n). \quad (4)$$

After the training of the network, a validation is performed with a set of patterns that are not used during the learning process, with the goal to appreciate the quality of the net for pattern recognition [8,22,10].

3 Experiments and Results

In this section, we present our experiments and results of the proposed MFN for obstacle detection by images recognition, built by a single hidden layer and log-sigmoid like transfer function. The *backpropagation* algorithm is used for the training of the network. A single hidden layer network with enough neurons is an universal approximator [10]. If the machine for pattern recognition is designed as a MFN using a supervised learning algorithm, the task of feature extraction is performed by the computational units in the hidden layers of the network [8].

3.1 Data Set

For the training phase and validation of the network, a database consisting of images of highways with obstacles and without obstacles was built. Such images were acquired via the Internet. We consider images with natural sunlight, an average size of 1000×500 pixels and taken from a zenith perspective. We also consider images with obstacles such as cars, mudslides, landslides and lack of road sections, as well as images without the above items. Examples of these images are shown in elements of the Figure 1.

The images of the database are samples of constant dimensions of 71×71 pixels, some with context, i.e. the images have characteristics of the environment in which the highway is located. Contextual influences on object recognition are evident when local features are insufficient because the object is small, occluded or camouflaged [18].

For the database, two subsets of images are created: training and validation. These subsets are formed from a percentage of the total image to consider, for each image 14 samples from seven shifts and seven rotations are obtained. Recognition systems improve efficiency if distortions of the images are included in the training patterns: displacements and rotations [12,1]. This article proposed



Fig. 1. Images database, (a) obstacle image, (b) unobstructed image.

that the rotations are performed with an integer random angle of between 10 and 350 degrees, considering that the route of the UAV will be defined and images taken may have some rotation of the initial reference frame, displacements are considered so that samples are not single context. In Figure 2 can see a set of samples with context, rotation and displacement.



Fig. 2. Sample images.

3.2 Training Patterns

Three types of image pre-processing have been considered to form the training patterns.

1. The images were converted to grayscale and then binarized by the Otsu method ¹.
2. The images were converted to grayscale and normalized according to the mean and standard deviation for each sample.
3. The images are considered in the RGB color space and normalized according to the mean and standard deviation for each sample.

Each sample is a subset of the scanned image, so we consider the samples as a matrix of 71×71 elements, where each element is the pixel value at that point, each matrix corresponding to each sample, is written as a column vector, transposing each row of the matrix in a column vector of 5041×1 for the first two pre-processing and $15,123 \times 1$ for the latter. These vectors contain the information that the net use to classify each pattern.

¹ A description of the algorithm to find global binarization levels can be found in [6]

3.3 ANN Topology

The structure of the proposed ANN is MFN. In Figure 3, a scheme of information processing of an image in the RGB color space is shown.

In this application, the number of entries is associated with each digitized sample. In order to propose the best treatment for the classification of the images, experiments with each of the pre-processed images are performed, considering a training set of 2385 patterns for training and 477 patterns for validation of the recognition obtained from the net. There is only one output, whose value corresponds to the absence (0) or presence (1) of an obstacle.

For the determination of the number of neurons in the hidden layers, there are different heuristic criteria. As the number of inputs increases, the ratio of hidden-layer neurons to inputs decreases [19]. Successive iterations of testing and error were performed to find a efficient and effective topology. The proportion of the training and validation sets corresponds to 80% and 20% respectively of the images in database. The parameters of the ANN are set small for each experiment, the learning-rate $\eta = 0.01$ and momentum factor $\alpha = 0.001$. The hidden and output layer have a log-sigmoid like transfer function.

Table 1. Simulation with different preprocessed images.

| Binarized images | | | | | |
|-----------------------|--------|-------|---------|----------|-------------|
| Topology | MSE | Epoch | Time | Training | Recognition |
| 5041 : 1000 : 1 | 0.0091 | 41 | 6h 3m | 64.31% | 56.21% |
| 5041 : 2000 : 1 | 0.0059 | 57 | 7h 55m | 57.62% | 51.94% |
| 5041 : 2000 : 500 : 1 | 0.013 | 28 | 7h 45m | 54.38% | 51.94% |
| Grayscale images | | | | | |
| 5041 : 2000 : 1 | 0.0051 | 74 | 16h 25m | 55.30% | 50.27% |
| 5041 : 1000 : 500 : 1 | 0.014 | 75 | 5h 57m | 47.33% | 48.05% |
| RGB images | | | | | |
| 15123 : 1 | 0.0239 | 124 | 81.70s | 95.22% | 73.80% |
| 15123 : 500 : 1 | 0.0001 | 124 | 6h 42m | 100% | 82.73% |
| 15123 : 300 : 1 | 0.0005 | 103 | 5h | 99.91% | 81.81% |
| 15123 : 4000 : 1 | 0.0005 | 200 | 60h 48m | 99.91% | 76.50% |
| 15123 : 500 : 100 : 1 | 0.0041 | 75 | 14h 30m | 97.31% | 78.33% |

As shown in Table 1, by using binarized and grayscale images, the training process last more than 6 hours and an incomplete training on the total pattern is achieved. Considering the RGB images, the ANN achieved more than 95% of training at different time given the size of the database. We simulated nets with images of small size of 36×36 pixels but there are slight improvements in the training and validation process, which we associate that with this type of

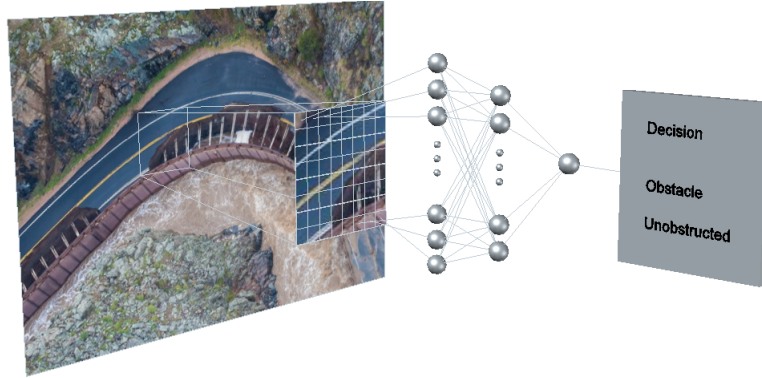


Fig. 3. Information processing. The image is taken from database, then is considered in the RGB color space and normalized, the values resulting from this operation are the entries of the net, then the classifier that is the MFN proposed take a decision, about whether or not there is an obstacle in the image.

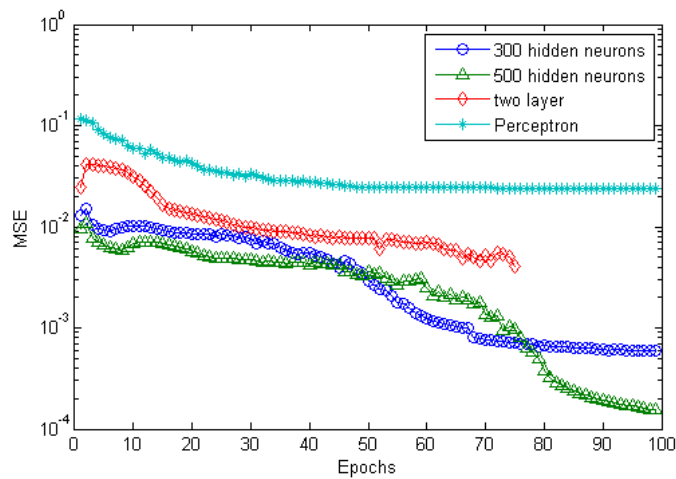


Fig. 4. Comparison of the learning curves associated with different topologies.

network structure is inadequate for treatment with highway images in grayscale and binarized.

In Figure 6, a set of images that the network could not classified correctly is shown. In the first three items, we can see that the obstacle is very small, so the lack of similar patterns in the training set result in misclassification. In the fourth image we can see that much of the sample is context, then the response of ANN result in misclassification.

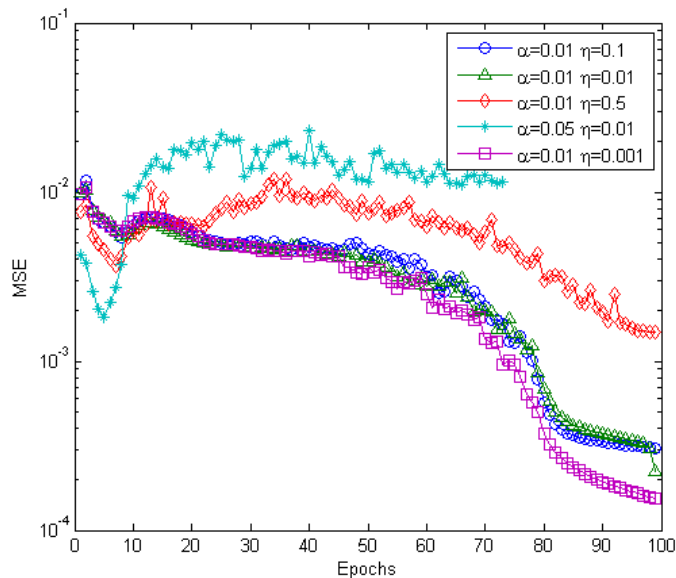


Fig. 5. Comparison of the different parameters.

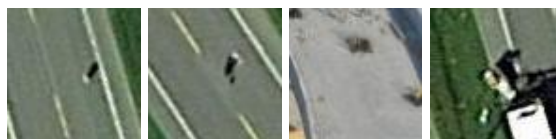


Fig. 6. Misclassified images.

3.4 ANN Parameters

To find the learning rate and momentum factor of the parameters of the RNA, various tests were performed considering the same topology, weights are generated by the implementation in Matlab of the generator of observations of a uniform variable MT19937 with seed 2085268696. Via experimentation, we obtain the following results:

For values of the learning rate $\alpha \in [0.05, 1]$, the network has been unstable manner.

For values of the momentum $\eta \in [0.6, 1]$, the network has been unstable manner, due the factor is too large.

The best values for parameters are $\eta = 0.01$ and $\alpha = 0.001$

In Figure 5 a comparison of different parameters is shown, we plots the best learning curves selected for our experiments, which they consisted of simulating nets with different parameters.

4 Conclusions

In this work, a design of a MFN with a single hidden layer has been proposed. The MFN was trained by using normalized RGB images, as we have seen for treatment of grayscale and binarized images, a different structure of ANN is required, so we conclude that when objects in the images are indistinguishable with the background, the color of the objects is a powerful feature for pattern recognition.

For the optimal values of the learning-rate and momentum, we use the definition: The η and α that, for either the worst-case or on average, yield converge to the global minimum in the error surface with the least number of epochs. With the results obtained, we propose an MFN with a single hidden layer with topology 15123:500:1, learning-rate $\eta = 0.01$ and momentum term $\alpha = 0.001$. The transfer function of the net is log-sigmoid for the hidden and output layer. With this topology and parameters we get the best result of recognition of 82.73% with images not used in training and of several highways.

As future work the MFN will be tested in real world controlled scenarios into the on-board computer of an UAV. The features of the considered obstacles are limited to static objects and cars. Dynamic obstacles are not considered in this work, so it is necessary to design an algorithm to make the recognition of cars in proper circulation and those that are an obstacle as detained cars and overturned cars. Regarding the training process, database should be expanded, because as we have seen the net misclassified small object associated with small obstacles on highways as persons or stones. In terms of image acquisition and processing, filtering processes are necessary for image reduction noise and thus obtain images suitable for classification.

References

1. Baidyk, T., Kussul, E.: *Redes neuronales, visión computacional y micromecánica*. Editorial Itaca (2009)
2. Basu, J.K., Bhattacharyya, D., hoon Kim, T.: Use of artificial neural network in pattern recognition. *International Journal of Software Engineering and Its Applications* (2010)
3. Bertozzi, M., Broggi, A.: A parallel real-time stereo vision system for generic obstacle and lane detection. *IEEE Transactions on Image Processing* 7(1), 62–81 (1998)
4. Bishop, C.M.: *Pattern Recognition*. Clarendon Press, Birmingham, UK (1995)
5. Chen, J., Zhao, P., Liang, H., Mei, T.: Motion planning for autonomous vehicle based on radial basis function neural network in unstructured environment. *Journal Sensors* (2014)
6. Chen, Z., Tu, Y.: Improved image segmentation algorithm based on OTSU algorithm. *International Journal of Advancements in Computing Technology (IJACT)* 4(15) (2012)
7. Cruz, P.P.: *Inteligencia artificial con aplicaciones a la ingeniería*. Alfaomega, México (2010)

8. Haykin, S.: *Neural Networks: a comprehensive foundation*. Prentice-Hall Inc., Ontario, Canada (1998)
9. Hinojosa, C.: *Utilizando sistemas de gestión inteligente para mejorar la seguridad vial*. Versión electrónica presentada en *Actualidad vial*, Zaragoza, España (2013)
10. Hornik, K., Stinchcombe, M.: Multilayer feedforward networks are universal approximators. *Neural Networks* 2(5), 359–366 (1989)
11. Huha, K., Park, J., Hwang, J., Hong, D.: A stereo vision-based obstacle detection system in vehicles. *Optics and Lasers in Engineering* 46(2), 168–178 (2008)
12. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proceeding of the IEEE* 86(11), 2278–2344 (1998)
13. Liu, F., Cao, B.: A neural network enhanced stereo vision obstacle detection and avoidance system for unmanned ground vehicle. *2nd International Conference on Advances in Computer Science and Engineering* (2013)
14. McGee, T., Sengupta, R., Hedrick, K.: Obstacle detection for small autonomous aircraft using sky segmentation. *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA)* pp. 4679–4684 (2005)
15. Nedeveschi, S., Danescu, R., Frentiu, D., Marita, T., Oniga, F., Pocol, C., Schmidt, R., Graf, T.: High accuracy stereo vision system for far distance obstacle detection. In: *IEEE Intelligent Vehicles Symposium*. Parma, Italy (2004)
16. Ochoa, J.A.C., Trinidad, J.F.M.: Reconocimiento de patrones. *Komputer Sapiens III(II)* (2011)
17. Ohe, I., Kawashima, H., Kojima, M., Kaneko, Y.: A method for automatic detection of traffic incidents using neural networks. *Bulletin of Mathematical Biophysics* (1995)
18. Oliva, A., Torralba, A.: The role of context in object recognition. *TRENDS in Cognitive Sciences* (2007)
19. Priddy, K.L., Keller, P.E.: *Artificial neural networks: an introduction*. The International Society for Optical Engineering, Bellingham, USA (2005)
20. Raju, S., Sanjay, K., Kumar, T.S., Madhini, B.: Semi autonomous vehicle to prevent accident. *International Journal of Technology Enhancements and Emerging Engineering Research* 2(5), 42–45 (2014)
21. Ramirez Q., J.A., Chacón M., M.I.: Redes neuronales artificiales para el procesamiento de imágenes, una revisión de la última década. *Revista de ingeniería eléctrica, electrónica y computación* 9(1) (2011)
22. Rosenblatt, F.: The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65(6) (1958)
23. Yu, H., Hong, R., Huang, X., Wang, Z.: Obstacle detection with deep convolutional neural network. *Sixth International Symposium on Computational Intelligence and Design* (2013)