Development of a Backpropagation Neural Network to Assist Cell Identification in Herbivore Diet at the Semiarid Chaco of North-West Argentina

Rosanna Costaguta[1] & Ledda Larcher[2]

[1]Departamento de Informática, Facultad de Ciencias Exactas y Tecnologías,
Universidad Nacional de Santiago del Estero, Argentina
rosannac@unse.edu.ar
[2]INDEAS (Instituto para el Desarrollo Agropecuario del Semiárido),
Facultad de Agronomía y Agroindustrias, Universidad Nacional de Santiago del Estero, Argentina
llarcher@unse.edu.ar

Abstract. Raising goats is a strong component of regional economies and a main activity of little farmers in Argentina. Knowing the diet is fundamental to correct nutritional lacks resulting from grazing native species, as it will help to design management strategies both for the native forage resource and for the nutritional state of the animal. This fact has a preponderant relevance when goat production is done at natural areas. Microhistology is one of the more frequently used methodology to determine the botanical composition of herbivorous diet. The specialist working at a microscope, according to his training and experience, can identify the species taken by the animal and to which those cells belong. Neural network approaches are suited for recognition problems, this article presents a backpropagation neural network initially trained to identify the four most significant vegetal species found in goat diet that graze freely in the Semiarid Chaco of North-West Argentina.

1 Introduction

Goats are animals that graze from different groups of species: herbaceous, woody, cactus and shrubs; this browsing herbivore takes leaves and fruits that allows it to reach a nutritional balance. At semiarid regions, forage woody species are important as a feeding complement since they prevail at seasons the grass disappears. Goat production at natural areas becomes easier by the selective habit of this animal, which allows it to keep regular nutritional levels [1].

For long time raising goats has been the main activity of numerous little farmers in Argentina, thus becoming a strong component of regional economies. Generally, this is a subsistence system since farmers belong to the sector with less resource. These are extensive systems, based on free daily raising in areas without determined bounds, with night gathering, without sure water supply neither sanitary or production controls [2].

To make reference to the behaviour of an animal species at a forage offering, concepts like palatability, preference and selectivity must be mentioned, for which is necessary a previous knowing of the diet botanical composition [3]. This way, the knowledge of herbivorous diet helps understanding some of the ecological relationships between the primary producer (vegetal community) and the consumer (herbivorous). Besides, botanical composition of diet and its change through seasons, play an important role when designing management strategies both for the native forage resource, as well as the nutritional state of the animal [4]. From the goat production perspective, knowing the diet is fundamental to correct nutritional lacks that result from grazing native species, especially during winter season, when forage crisis is severe [5].

There are many methodologies to establish the botanical composition of herbivore diet [6, 7, 8, 9, 10, 11], such as total density, relative frequency, ranges, among others; each adapts to a particular ecological situation. Microhistology is one of the techniques used in determining botanical composition of herbivorous diet [3], however involves the knowledge of a great number of species and their histological characters, making the successful identification highly dependant on the technicians training and experience. The problem easily fell into the general area of pattern recognition. There were reasons to believe that the neural network approaches would be remarkably well suited for cell recognition problems. Firstly, the vegetal species cells to work with, form a very specific class of patterns. Thus, the pattern sets seemed well confined and constrained. Secondly, as neural networks are robust, adaptive and trainable from examples, images can have deformations and noise. Thirdly, as efficiency is an important aspect in the consulting application, which could require rapid searches through databases, neural networks can be tailored and trained to fit this response time requirement. Developing an automatic reasoning model, as artificial neural networks are, that could help in pattern recognition was found as a solution to increase accuracy on microscopic technicians when doing their job.

This work aims to present the implementation of an artificial neural network (ANN) that takes as input digital images of epidermic pieces and recognises cells of different botanical species in goat diet at the region called Chaco Semiárido of North -West Argentina (NOA). The species considered were: *Zizyphus mistol, Acacia* sp*., Prosopis alba* and *Eupatorium* sp*., Justicia campestris, Clematis montevidense, Celtis palida, Lippia turbinata, Lantana* sp*., Geoffroea decorticans, Wissadula* sp*., Sphaeralcea sp* commonly found in goat diet. Images used in this work are stored in a database. [12].

2 Artificial Neural Networks

ANNs have been developed as generalizations of mathematical models of human cognition or neural biology [14]. ANNs emulate a biological network, and precisely because they are models, there exists a simplification of the system and only the relevant elements are used [13]. ANNs are information-processing systems made up of several simple, highly interconnected processing elements that dynamically process the information, thus giving a response to external inputs [14]. These processing units or nodes -which are equivalent to human neurons-, are generally devices that receive a determined number of input signals over a communication link with an associated weight, apply an activation function to this input to determine the output signal. These nodes are simple devices that must have the ability of storing these signals and process them according to their weights. Nodes or neurons interconnect forming networks, which originates the neural network denomination. The pattern of connections between the neurons, the method used in determining the weights of connections and the activation function used characterizes every ANN [14].

Neurons distribute in the network forming levels or layers, each having a determined number of nodes. Generally, three levels can be distinguished, corresponding to the input layer, hidden layer and output layer. The input layer receives the information directly from sources outside the network. Hidden layers are internal, do not have direct contact with the environment; the number of hidden layers can vary from zero to a high number. The output layers transfer the information from the network to the environment. Nevertheless, the network is not always a multi-layered one, the number of layers and neurons within each of them will change according to the ANN objectives and needs. Generally, three layers are enough, although cases exist when the ANN learns faster if having more than one hidden layer [15].

3 The Developed ANN

Previous to the development there was some research and test of different types of ANN, focusing on the problem to solve, weakness and strengths of different architectures, difficulties found during the development and training of various applications. This study lead to the selection of a backpropagation ANN (BANN). The choice relays on its ability to learn how to relate input patterns (examples) with the corresponding class, and its ability to adapt the weights from intermediate neurons to learn the existing relationship between a pattern set given as an example and the corresponding output (backwards propagation of error). This makes possible that, after the training, the BANN can recognize new input vectors even with noise or incomplete, giving as result an active output if the new input is similar to one of the showed during the learning. This important characteristic, demanded to every learning system [13], is the generalization ability, understood as the ability of giving satisfactory outputs to inputs the system has never seen during the training.

The developed BANN is made up of 3 layers (input – hidden – output), 64 nodes for the input layer, and 8 for the hidden layer as shown in Fig. 1. The input for the net is a gray scale digital photograph resulting from a picture taken from a portion of a slide. The output layer has only 4 nodes since the training and tests performed so far involve the 4 species most commonly found (i.e., patterns) in the goat diet. The output generated in this layer values means the response to the input, not only if the pattern presented whether belongs or not to a category, but to which category.

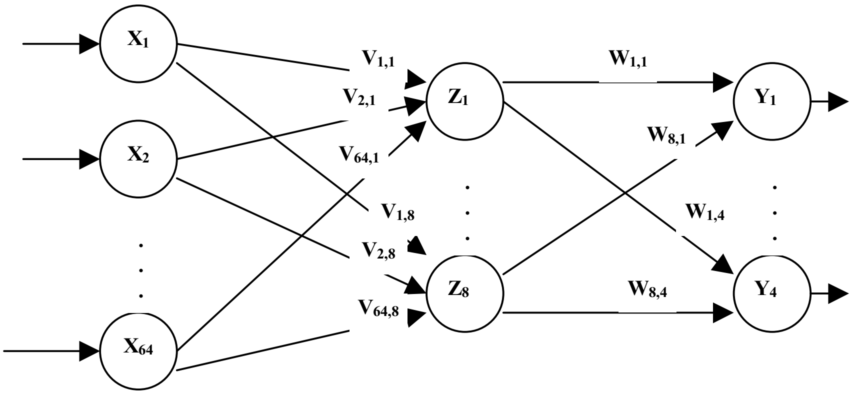


Fig. 1. Architecture of the developed BANN

The learning algorithm was centered in the Delta rule generalization [13, 14, 15] using a bipolar sigmoid activation function. In an earlier implementation, the initial weights were set with random values between (-0.5; +0.5) as is commonly the case; then some modifications were made looking forward to improve the ability of hidden units to learn. This was accomplished by distributing the initial weights and biases, so that, for each input pattern, it would be likely that the net input to one of the hidden units would be in the range in which that hidden neuron will learn most readily (Nguyen-Widrow initialization method) [14].

To train the BANN in recognizing the epidermic cells that let identify the species intaken by goats, two disjoint set of test patterns were created: a set of training patterns and a set of test-training patterns.

The sets of patterns were built using images (photographs taken at 10x). Each species characteristics have, at that magnitude, enough details to be recognized. The training set was established with the assistance of a specialist, making a visual selection among them and choosing those that presented minimal overlap. Then some of the cells were identified and outlined by the specialist. Initially the network was trained to recognize *Zizyphus mistol, Acacia* sp.*, Prosopis alba* and *Eupatorium* sp., this selection was made taking account of differences between cells shape (Fig. 2) and their significance in goat diet.

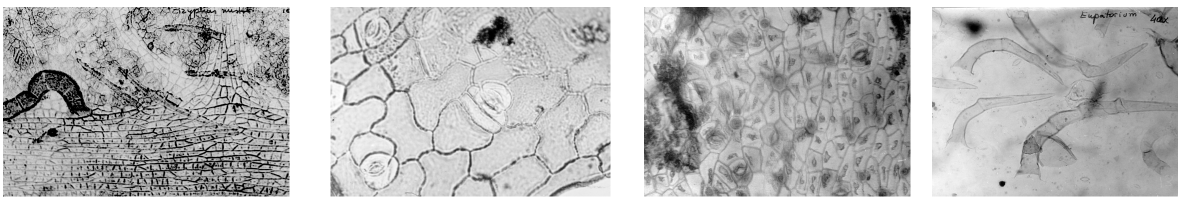
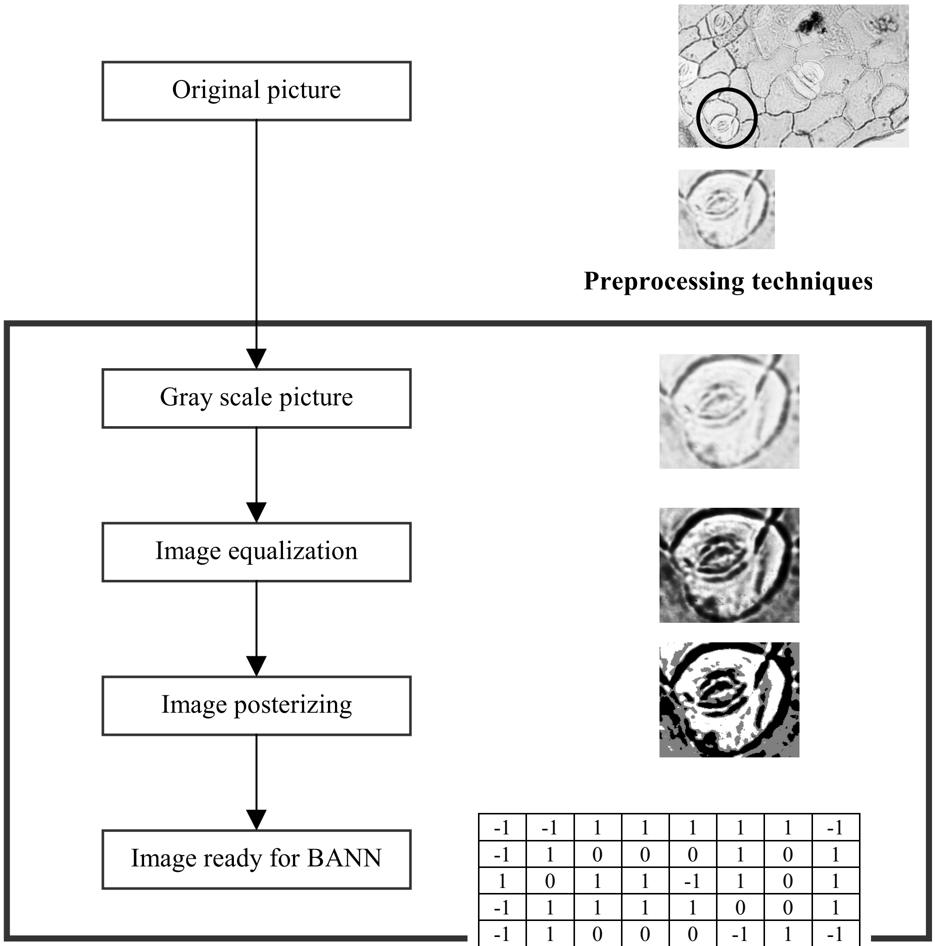


Fig. 2. Pictures showing *Zizyphus mistol, Acacia* sp.*, Prosopis alba* and *Eupatorium* sp cells

A digitalization process using a scanner follows the selection, when the photograph is not stored in the database mentioned above [12]. All images go through a digital image process wherein the outlined characteristics are more elaborately separated from other objects (other cells or portions of them) within the image. Quality of input images is improved using different image enhancement techniques (changing it to 8 bit gray scale – equalizing – posterizing into 3 levels) [16], making the necessary transformations so as to enable the extraction of suitable features for processing in the BANN. The resulting image keeping only their main characteristics is then translated into a 64 element bidimensional arrays (8 rows, 8 columns) that will hold the values –1; +1 or 0, (which stand for white, gray or black) (Fig. 3).



-1	-1	1	1	1	1	1	-1
-1	1	0	0	0	1	0	1
1	0	1	1	-1	1	0	1
-1	1	1	1	1	0	0	1
-1	1	0	0	0	-1	1	-1
-1	-1	1	0	0	1	-1	-1
-1	-1	-1	1	1	-1	-1	-1
-1	-1	-1	1	1	-1	-1	-1

Fig. 3. Block diagram of image processing using *Acacia* sp.

Obtaining the training and test-training pattern sets was a complex task as cells are not bidimensional figures, but vary in shape, location, standpoint of sight, etc. and the BANN had to be trained to recognize main features on each of these variations.

The BANN was trained as follows: weight adjustments were made based on the training patterns; at intervals during training, the error was computed using the training-testing patterns. As long as the error for the training-testing patterns continued decreasing, training continued. When the error began to increase, it was interpreted as the net starting to memorize the training patterns too specifically, (which meant, losing its ability to generalize). At this point, training was terminated. After the training phase ended, and in order to verify that the BANN was able to make the desired association between the input vectors and patterns never presented before, it was provided different patterns presuming they could be noisy, incomplete, distorted, and new ones [15].

The network shows at this point a good performance in the recognition of the four species used through its learning state. Some experimental results are shown in Table 1.

Table 1. Training and test results

	Zizyphus mistol	*Acacia* sp.	*Prosopis alba*	*Eupatorium* sp
Patterns used for training	24	14	18	12
Patterns used for test	8	8	8	8
Correct answers	5	7	6	5
Incorrect answers	3	1	2	3
Reliability (%)	62.5	87.5	75	62.5

The BANN was programmed using C language on a Pentium 100 desktop computer. This choice is based in the main features of this language: simple creation and use of libraries, code reutilization, ability to handle recursive callings, easy of creation and function use, compact and fast execution of final code.

4 Conclusions

This article shows the successful application of ANN in a field such as microhistology. In particular, the BANN will be extremely useful for the technician who will have his efficiency and efficacy increased by having an automatic reasoning model, trained in determining the species found in diets of goats grazing freely in the Semiarid Chaco of NOA. At the same time, knowing the nutritional level of the animal diet would let the farmer to incorporate, when necessary, elements those insure maintaining the quality of the animals.

As future lines of work, the following aspects are being taken in consideration:

(a) incorporating the whole set of species and train the BANN in their recognition;

(b) changing the number of the hidden layer nodes and/or the number of these layers, looking forward to find a better convergence to the solution;

(c) developing a software module to automatically transform the photographs into files having the required input format to the BANN.

5 Acknowledgments

To the CICYT – UNSE for the given support to the project “Composición botánica, amplitud y calidad forrajera de la dieta de caprinos pastoreando en el monte chaqueño”.

References

[1] Agraz-García, A., *Caprinotecnia 3*, Limusa: México, 1989.

[2] Salinas, A. & Dellepiane, G., *Situación del Sector Caprino Nacional*, Ministerio de Economía y Obras y Servicios Públicos, Secretaría de Agricultura, Ganadería y Pesca: Argentina, 1994.

[3] Pelliza de Sbriller, A., Acerca de la microhistología. *Comunicación Técnica. Recursos Naturales, Dieta, INTA*, 32, 1993.

[4] Mellado, M., Foote, M., Rodriguez, A. & Zarate, P., Botanical composition and nutrient content of diet selected by goats grazing on desert grassland in northern Mexico, *Small Ruminant Research*,. 6, pp. 141-150, 1991.

[5] Renolfi, R., *Control de arbustos mediante pastoreo con caprinos en un arbustal-bosque abierto de la región chaqueña semiárida*, Universidad Nacional de Mar del Plata, Escuela para Graduados en Ciencias Agropecuarias: Argentina, 1994.

[6] Havstad, K. & Donart, G., The microhistological technique: testing two central asssumptions in South-Central

New Mexico. *J. Range Manage*, 31(6), pp. 469-470, 1978.

[7] Holecheck, J. & Gross, B., Training needed for quantifying simulated diets from fragmented range plants. *J. Range Manage*, 35(5), pp. 644-647, 1982.

[8] Norbury, G., Microscopic analysis of herbivore diets - a problem and a solution. *Aust. Widl*, 15, pp. 51-57, 1998.

[9] Sparks, D. & Malechek, J., Estimating percentage dry weigth in diets using microscopic technique. *J. Range Manage*, 21 (4), pp. 264-265, 1968.

[10] Westoby, M., Rost, G.& Weis, J., Problems with estimating herbivore diets by microscopically identifying plant fragments from stomachs. *J. Mammals*, 57, pp. 167-172, 1976.

[11] Barker, R., A technique to simplify herbivore diet analysis. *Aust. Widl.,* 13, pp. 569-573, 1986.

[12] Degano C.; Larcher, L.; Renolfi, R. & Martiarena R., A database to help in the determination of species that conform the diet of herbivorous, when microhistology is used. *Pakistan Journal of Botany*, 31(2), pp. 295-300, 1998.

[13] Hilera, J. & Martínez, V., *Redes neuronales artificiales. Fundamentos, modelos y aplicaciones*, Addison-Wesley Iberoamericana: México, 1995.

[14] Fausett, L., *Fundamentals of Neural Networks. Architectures, algorithms and applications*, Prentice Hall: U.S.A., 1994.

[15] Freeman, J., & Skapura, D., *Redes neuronales. Algoritmos, aplicaciones y técnicas de programación*, Addison-Wesley/Díaz de Santos: España, 1993.

[16] Quek, C., Tan K., & Sagar, V., Pseudo-outer product based fuzzy neural network fingerprint verification system. *Neural Networks*, 14, pp. 305-323, 2001.