Piotr BONIECKI, Krzysztof NOWAKOWSKI, Robert Jacek TOMCZAK
Uniwersytet Przyrodniczy w Poznaniu, Instytut Inżynierii Rolniczej
ul. Wojska Polskiego 50, 60-637 Poznań
e-mail: bonie@up.poznan.pl

NEURAL NETWORKS TYPE *MLP* IN THE PROCESS OF IDENTIFICATION OF CHOSEN VARIETIES OF MAIZE

Summary

During the adaptation process of the weights vector that occurs in the iterative presentation of the teaching vector, the MLP type artificial neural network (MultiLayer Perceptron) attempts to learn the structure of the data. Such a network can learn to recognise aggregates of input data occurring in the input data set regardless of the assumed criteria of similarity and the quantity of the data explored. The MLP type neural network can be also used to detect regularities occurring in the obtained graphic empirical data. The neuronal image analysis is then a new field of digital processing of signals. It is possible to use it to identify chosen objects given in the form of bitmap. If at the network input, a new unknown case appears which the network is unable to recognise, it means that it is different from all the classes known previously. The MLP type artificial neural network taught in this way can serve as a detector signalling the appearance of a widely understood novelty. Such a network can also look for similarities between the known data and the noisy data. In this way, it is able to identify fragments of images presented in photographs of e.g. maize grain. The purpose of the research was to use the MLP neural networks in the process of identification of chosen varieties of maize applying the image analysis method. The neuronal classification shapes of grains was performed with the use of the Johan Gielis super formula.

SIECI NEURONOWE TYPU MLP W PROCESIE IDENTYFIKACJI WYBRANYCH ODMIAN KUKURYDZY

Streszczenie

Podczas iteracyjnej korekcji wektora wag, zachodzącej w trakcie procesu uczenia sieci neuronowej typu MLP (perceptron wielowarstwowy), następuje adaptacja (przez tworzony model neuronowy) wiedzy zawartej w strukturze analizowanych danych. W badaniach prowadzonych w dyscyplinie inżynieria rolnicza, istotne znaczenie ma proces pozyskiwania informacji zakodowanej w postaci graficznej, np. w formie zdjęć cyfrowych. Często zmiennymi reprezentatywnymi, które w sposób wystarczający charakteryzują zobrazowany obiekt, są wybrane współczynniki kształtu. Celem badań było wykorzystanie sieci neuronowych typu MLP w procesie identyfikacji wybranych odmian kukurydzy z wykorzystaniem metod analizy obrazu. Wykorzystana metoda klasyfikacji polegała na rozpoznawaniu różnic kształtów analizowanych obiektów. Neuronowa identyfikacja została wykonana z użyciem super formuły Johana Gielisa.

1. Introduction

Neural models have proven to be highly effective tools serving, inter alia, in data processing to solve problems that cannot be tackled with the use of conventional computers and software. This is because neural networks process data in a parallel and scattered way. Such models are usually considerably faster than sequential processing on conventional digital machines arranged in series. A key advantage of neural networks is their ability to model the problem at hand without developing an algorithm for its solution [8]. The approach closely resembles the operation of the human brain, which has actually inspired and served as a prototype for neural modelling. One area in which humans still continue to defeat computers is identification and recognition in the broad sense of these terms [2].

The purpose of the study is to explore the classification abilities of neural networks in identifying selected agricultural objects. The examples used for such classification were the maize grain. The proposed classification method involved recognizing the shapes of the researched objects presented as images. The caryopsis shapes were identified with the use of a superformula, as it is referred to, proposed by the Belgian engineer *Johan Gielis*. The formula helps map any shape using independent parameters which, for the purposes of this paper, have been identified as distinguishing features. These are used to construct learning sets for the neural models.

2. *Johan Gielis'* superformula

In 1997, *Johan Gielis*, a Belgian engineer, mathematician and biotechnology expert, published a formula (being a generalization of the superellipse formula developed by *Gabriel Lamé*) which describes any shape both in two- and three-dimensional space [4]. The formula has been dubbed the superformula and is occasionally referred to as the *Gielis'* superformula. In a two-dimensional polarized system, the formula is [3]:

$$\frac{1}{r} = \sqrt[n_1]{\left|\frac{1}{a}\cos\left(\frac{m}{4}\phi\right)\right|^{n_2} + \left|\frac{1}{b}\sin\left(\frac{m}{4}\phi\right)\right|^{n_3}} \tag{1}$$

where:

r and ϕ - denote point coordinates in a polarized system,
m - denotes the symmetry coefficient,
n_1, n_2, n_3 - denote the shape coefficients,

a, b - denotes the horizontal and vertical dimensions.

As he developed his notion, *Johan Gielis* proposed the following generalized formula in 2003:

$$r(\theta) = \varphi(\theta)\left[\left|\frac{\cos\left(\frac{1}{4}m\theta\right)}{a}\right|^{n_2} + \left|\frac{\sin\left(\frac{1}{4}m\theta\right)}{b}\right|^{n_3}\right]^{-\frac{1}{n_1}} \tag{2}$$

where:

r, θ - denote point coefficients in a polarized system,

$\varphi(\theta)$ - denotes any mathematical function.

Just as in formula (1), formula (2) can also be applied to manipulate the emerging shape with the use of shape coefficients n_1, n_2 and n_3 and alter the shape of function $\varphi(\theta)$ so as to produce any complex shape.

3. Design and implementation of an artificial neural network

In order to generate caryopsis shapes and then identify the proper superformula coefficients, use has been made of a JAVA applet shape generator (the applet was written by *Holger Hoffmann*) available in an interactive format e.g. at http://www.activeart.de/dim-shops/training/SuperShape/.

Empirical data in the form of chosen maize's caryopsis photographs, were transformed into a set of adequate caryopsis shapes with the use of the *Hoffman's* applet. A record was also made of parameters a, b, m, n_1, n_2 and n_3 which describe the analyzed shapes.

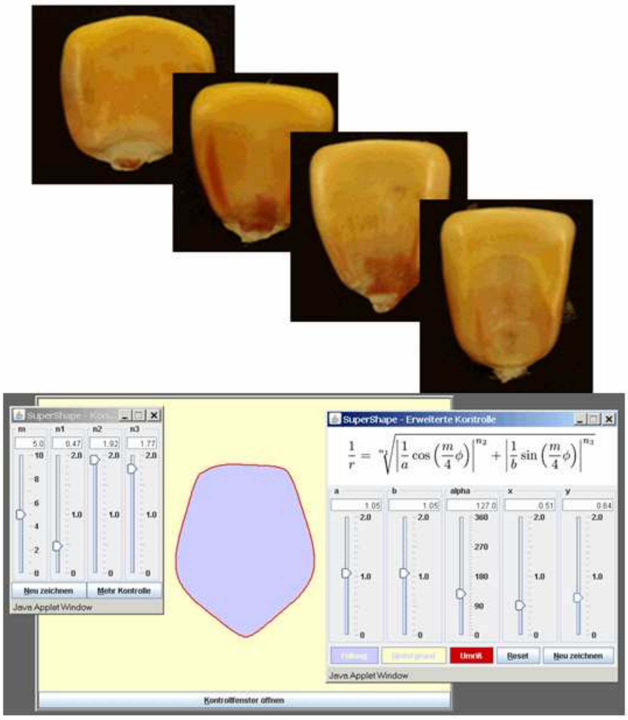

Fig. 1. Pictures of chosen maize grain of the maize [5] and shape generating of grain with usage of *Holger Hoffmann's* applet

Parameters a, b, m, n_1, n_2 were assumed to be the distinguishing features of the caryopsis images tested. These were used to design a learning set for the construction of a neural model capable of classifying wheat, barley and maze caryopses. A fragment of the set is given in Figure 2.

imput variables

output variables

learning vectors

a	b	m	n_1	n_2	n_3	variety
1,03	0,95	2	0,33	1,2	1,23	1
1,13	0,88	2	0,43	0,98	0,11	1
1,01	1,02	2	0,38	1,18	1,28	3
1,12	1,19	2	0,56	0,66	0,55	2
1,07	1.12	2	0,66	1,12	0,44	3
1,16	0,98	2	0,33	1,22	1,31	3
1,17	0,99	2	0,33	0,66	0,31	2
1,12	1,1	2	0,32	1,24	0,44	1
1,03	1,14	2	0,88	1,11	1,78	2
1,13	0,99	2	0,43	0,74	0,41	3
1,10	1,21	2	0,68	0,78	1,45	3
1,12	0,95	2	0,59	1,23	0,19	2
1,16	0,92	2	0,54	0,61	1,22	3

Fig. 2. Fragment of training data set

where:

a, b - denote the horizontal and vertical dimensions,

m - denotes a multiplicity of a symmetry axis,

n_1, n_2, n_3 - denote shape coefficients,

variety - denotes species name (**1**: Clarica FAO 230, **2**: Clarica FAO 2350, **3**: Clarica FAO 280)

The learning set was made up of 210 cases. The cases were placed randomly in a learning set. Designed in such a way, the data set became the basis for learning in the process of developing a neural classification model.

The desired neural network topology was generated with a "Neural networks" module integrated into the commercial software system *Statistica v.7.0*. [7]. At stage one, use was made of the effective "*Automatic network designer*" tool. The authors assessed individual model features such as the rate of learning, validation and test errors, the learning, validation and testing related quality of the neural network, the *ROC* curve (*Receiver Operating Characteristic*)and classification problem statistics [7].

Out of a set of the ten best neural network typologies, the researchers selected the *MLP* neural network characterized by the smallest *RMS* (*Root Mean Square*) error rate. Its structure: *MLP* 6:6-36-1:1 is provided in Figure 3.

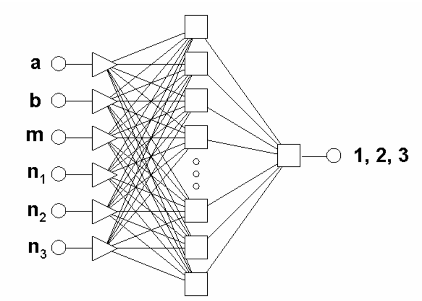

Fig. 3. Structure of developed ***MLP*** type neural network

Statistics of model that best identified varieties of maize:

- learning error - 0,1022,
- validation error - 0,1244,
- test error - 0,1377,
- learning quality - 0,9809,
- validation quality - 0,9666,
- test quality - 0,9811,
- filed under *ROC* curve - 0,9854.

The network then went through a process of "additional learning" that relied on the use of a *conjugate gradient* algorithm (500 epochs) and the *Lovenberg-Marquardt* algorithm (1000 epochs).

To define the weight of individual input parameters (relative to the adopted distinguishing features), the authors analyzed the sensitivity of the neural model generated [7]. The "Sensitivity analysis" module placed in the *Statistica v.*7.0 package was used to demonstrate that the input variables of highest rank are those representing caryopsis shapes, i.e. coefficients n_1, n_2 and n_3. It means that the resulting neural model has a limited sensitivity to changes in the values of features a, b and m, which describe the dimensions and spatial orientation of the modelled caryopses.

4. Conclusions

The studies suggest it is advisable to apply the artificial neural network technology, computer-aided image analysis and *Johan Gielis'* superformula to identify varieties of maize. The neural model developed and verified by the author demonstrates it is advisable to apply it to identify varieties of maize caryopses based on representative features established in caryopsis image analysis. In practical applications, the model has proven superior to the identification capacities of man, in particular where identification speed and duplicability are of the essence. It's possible to make ultimate conclusions as follows:

1. The generated neural model is an effective tool for the classification of varieties of maize and offers a high confidence ratio.
2. As the artificial *MLP* neural network is capable of generalizing knowledge, it easily classifies cases from outside of the learning set.
3. *Johan Gielis'* superformula proves effective in solving such problems. The neural model developed with its use simplifies the selection of distinguishing features to identify the relationship between input data (in the form of shape coefficients) and the output signal.
4. During the learning process, the neural model rejected the equation components describing the size of objects and their spatial orientation. This shows that it is possible to omit image normalization when describing objects with the use of the coefficients generated by the superformula, which accelerates learning data set generation.

5. Literature

[1] Boniecki P., Weres J.: The utilization neural networks technique to expectation of size gatherings chosen agricultural fetuses. Journal of Research and Applications in Agricultural Engineering, 2003, Vol. 48(4), 56-60.

[2] Boniecki P.: Neural networks of *MLP* and *RBF* type as complementary approximate models in the triticale crop prediction process", Journal of Research and Applications in Agricultural Engineering, 2004, Vol. 49(1), 28-33.

[3] Boniecki P., Olszewski T.: Neuronal modeling of selected agricultural objects with usage of Johan Gielis' super shape", Journal of Research and Applications in Agricultural Engineering, 2008, Vol. 53(1), 22-26.

[4] Cieślar K.: Superformula", Paper Foton, Institute of Physics of Cracow University, Cracow, in Polish, (2004).

[5] Nowakowski K., "Neural identification of selected mechanical damage to kernels", Ph.D. dissertation: Nowakowski K., Poznan, in Polish, 2008.

[6] Osowski S., "Neural networks for processing information", Publishing house of Warsaw Technical University, Warsaw, in Polish, 2000.

[7] Tadeusiewicz R., Lula P., "Statistica Neural Networks PL: Introduction to neural networks", StatSoft Poland, Cracow, in Polish, 2001.