Krzysztof NOWAKOWSKI, Piotr BONIECKI, Robert Jacek TOMCZAK, Barbara RABA
Poznan University of Life Sciences, Department of Agricultural Engineering, Poland
e-mail: krzysztof.nowakowski@up.poznan.pl; bonie@up.poznan.pl; rjt@up.poznan.pl; braba@up.poznan.pl

IDENTIFICATION PROCESS OF CORN AND BARLEY KERNELS DAMAGES USING NEURAL IMAGE ANALYSIS

Summary

The subject of the study was to develop a neural model for identification of mechanical damage to grain caryopses based on digital photographs. The authors has selected a set of universal features that distinguish damaged and healthy caryopses. As a result of this study it has been performed an artificial neural network of a multilayer perceptron type whose identification capacity is near of the human' s one.

IDENTYFIKACJA USZKODZEŃ ZIARNIAKÓW KUKURYDZY I JĘCZMIENIA BROWARNEGO Z WYKORZYSTANIEM NEURONOWEJ ANALIZY OBRAZU

Streszczenie

Celem projektu badawczego było opracowanie modelu neuronowego do identyfikacji mechanicznych uszkodzeń ziarniaków na podstawie ich cyfrowych fotografii. Wybrany został zestaw cech charakterystycznych na podstawie, których możliwa jest klasyfikacja ziarniaków na zdrowe i uszkodzone. W wyniku badań otrzymano sztuczną sieć neuronową typu perceptron wielowarstwowy charakteryzującą się zdolnościami identyfikacyjnymi zbliżonymi do umiejętności człowieka.

1. Introduction

The primary purpose in assessing the number of mechanically damaged caryopses is to gauge the quality of cereal grains. In order to ensure highest quality, the level of impurities must be minimised. One form of impurities are damaged caryopses. The secondary goal is to evaluate applied agricultural techniques. Classical methods for identification of damages are limited to use, e.g.: a set of mechanical sieves. They allow the separation of healthy from damaged seeds on the basis of their size. Another way is to identify the damage done by man. This method, despite a high accuracy (about 97%), has several defects. The first is the time needed for identification. It extends along with the increase in the number of kernels. Another is lack of reproducibility of the identification for the same cases. The solution to the problem of identification on a similar level of accuracy but without the defects of the above, seems to be the application of neural image analysis. The purpose of the study is to explore the classification abilities of neural networks in identifying mechanically damaged caryopses. 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 like the human brain.

2. Materials and methods

For the implementation of the proposal selected two grains: corn and malting barley. Samples consisted of 300 grains for each grain. After this analysis, the sample was 39 corn kernels mechanically damaged. However, in a sample of 36 damaged barley kernels. Pictures of all the kernels were performed in the chamber to shadow-free photographs. The chamber allows the use of any light source and parameters. At the same time do not see the photographs of shadows and glare objects.

Each caryopsis was placed on a black background and photographed three times with rotation through 120 ° which allows for exposure of the entire surface (Figure 1). The process of image acquisition allowed for collection of 1800 photos of kernels. Then the photos were processed into a learning data for artificial neural networks.

To conduct the identification process was necessary to designate characteristics tags. This should be a set of variables, which allows for unambiguous interpretation state of the grain: healthy or damaged. To identify damage were selected information about color and shape of kernels. Information about the shape were obtained by setting the shape coefficients of kernels:

a) first circuit coefficients

$$R_{C1} = 2\sqrt{\frac{S}{\pi}} \tag{1}$$

b) second circuit coefficients

$$R_{C2} = \frac{L}{\pi} \tag{2}$$

c) Malinowska coefficient

$$R_m = \frac{L}{2\sqrt{\pi \cdot S}} \tag{3}$$

where: *L* – circuit, *S* – field,

d) undimension coefficient

$$R_s = \frac{L^2}{4\pi \cdot S} \tag{4}$$

e) Feret coefficient

$$R_F = \frac{|L_n}{L_v} \quad (5)$$

where: L_N – maximum vertical size, L_V – maximum horizontal size.

Figure 1. Photos of corn and barley cereals

Color information would read from each pixel was describing kernel. Caryopsis color by RGB color space was coded in the manner described below to form a single number (Figure 2). The proposed coding allows to quickly return to representation in RGB. For the feature extraction and creation of learning sets used author's information systems the "Corn V 1.1" and "Hordeum V 1.1".

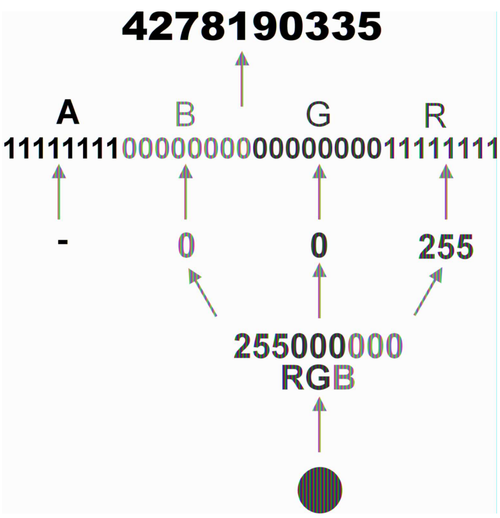

Figure 2. Coding scheme of information about color

Both systems allow to write data to CSV format acceptable by the artificial neural network simulator. To create and learning networks simulator was used with package Statistica.

3. Results and conclusions

During the learning process has been tested a lot of topology and learning methods of artificial neural networks. As the input variable, the information on colors and geometric factors. Output variable take the form of text: HEALTHY or damaged (see Figure 3).

Training set was divided into three subsets: learning, validation and test. Measures determining the quality of the network are determined independently for the training set, validation and testing. Measures determined based on the training set allows an assessment of the ability possessed by the network to approximate. They speak about the precision with which the network will be appointed by the variable output for input vectors presented during learning. Knowing how to network performance for data outside the training set is called the ability to generalize. The evaluation of the possible measures of network properties designated under the test set. Validation set is used to calculate the quality indicators used to monitor the learning process (to include identification of network over learning / decrease the ability to generalize).

During the network learning stage, damage maize and barley caryopses was best identified by the single-hidden-layer MLP network. This may be owed to the dual nature of the identification task. The number of neurons in the hidden layer to 47. Learning relied on the back propagation of errors at the first stage and on the conjugate gradient method in the second stage. For the model to identify damage to corn kernels the network was learned 50 epochs using an error back propagation and 173 epochs conjugate gradients method. The model identifies the damages of malting barley was trained in 50 epochs using an error back propagation method and 153 epochs conjugate gradients method. Statistics of model that best identified corn damage:

- learning error – 0,1069,
- validation error – 0,1371,
- test error – 0,1384,
- learning quality – 0,9907,
- validation quality – 0,9785,
- test quality - 0,9718,
- filed under ROC curve – 0,9713.

Statistics of model that best identified barley damage:

- learning error – 0,1009,
- validation error – 0,1181,
- test error – 0,1144,
- learning quality – 0,9897,
- validation quality – 0,9885,
- test quality - 0,9918,
- filed under ROC curve – 0,9711.

Classification statistics are presented in the tables below.

Table 1. Statistical classification for the model to identify damage to corn kernels

	Learning set	Validation set	Testing set
Total number of cases	450	225	225
Identified correctly	432	216	217
Wrongly identified	18	9	8
Unspecified	0	0	0

Table 2. Statistical classification for the model to identify damage to barley kernels

	Learning set	Validation set	Testing set
Total number of cases	450	225	225
Identified correctly	430	215	216
Wrongly identified	20	10	9
Unspecified	0	0	0

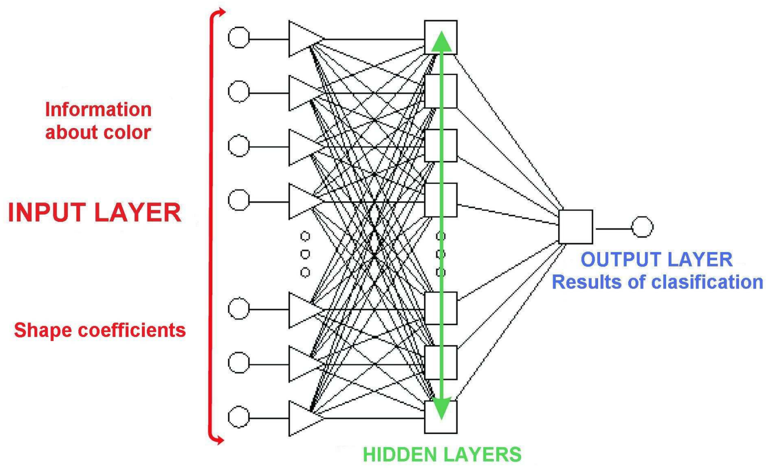

Figure 3. Neural network topology

The studies suggest it is advisable to apply the artificial neural network technology and computer-aided image analysis to identify damage. This conclusion is further supported by satisfactory characteristics of the best performing identification model. The approximate average global error rate of the top model is 4,5%. The approximate average error rate of humans in damage classification is 2% hence, the identification quality of the neural model is inferior to that of a man. On the other hand, neural networks' definite advantage over man is their speed and identification repeatability. The model is capable of identifying a much larger number of caryopses than a human sighter. The only constraint on the number lies in limitations of image acquisition and the processing power of the hardware. Another strong advantage lies in the absence of natural limitations that confine humans. A neural model can work continuously as it e.g. does not succumb to fatigue.

The neural model developed and verified by the authors demonstrates it is advisable to apply it to identify macrodamage in maize and barley caryopses based on representative features established in caryopsis image analysis.

4. Literature

[1] Boniecki P.: Elements of neural modeling in agriculture. Publisher University of Life Sciences in Poznań, 2008. ISBN 978-83-7160-473-7, in Polish.

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

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

[4] Nowakowski K.: Neural image analysis. Chapter in a monograph: Elements of agricultural systems engineering. Publisher University of Life Sciences in Poznań, 2008. Wyd. I, ISBN 978-83-7160-501-7, in Polish.

[5] Osowski S.: Neural networks to processing of information. Publishing house of Warsaw Technical University, Warsaw, 2000, in Polish.

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

Praca finansowana ze środków na naukę w latach 2009-2012 jak projekt badawczy – NN 313 264136