

Full Length Research Paper

Classification of chosen orchard pests using the self-organizing feature maps neural network

Krzysztof Pilarski^{1*}, Piotr Boniecki¹, Piotr Slosarz², Jacek Dach¹, Hanna Boniecka-Piekarska³
and Krzysztof Koszela¹

¹Institute of Agricultural Engineering, Poznan University of Life Sciences, Wojska Polskiego 50, 60-625 Poznań, Poland.

²Department of Small Mammal Breeding and Animal Origin Materials, Poznan University of Life Sciences, Złotniki, Słoneczna 1, 62-002 Suchy Las, Poland.

³Department of Entomology and Environmental Protection, Poznan University of Life Sciences, Dąrowskiego 159, 60-594 Poznan, Poland.

Accepted 6 September 2012

The Kohonen neural networks are modelled on the topological properties of the human brain. These networks are also known as self-organizing feature maps (SOFM). One advantage of suggesting a procedure is the ability of the SOFM neural network to determine the degree of similarity occurring between classes. The SOFM network can also be used to detect regularities occurring in the obtained empirical data. 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 SOFM 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 presenting photographs of orchard pests, for example. The resulting model of the Kohonen neural turns to be effective without reference classifier. The average classification error SOFM network during its operation was 0.05532 for the learning set and 0.0762 for the validation set.

Key words: Orchard pests identification, fruit tree pests, Kohonen neural networks, image recognition.

INTRODUCTION

An effective protection of fruit crops against phytophage species requires a precise determination of the pests' quantity and distribution, fixing the terms of fight against pests as well as selection of effective insecticides. Therefore, pest identification is a decisive process in plant's protection. It is extremely important to search for the effective and easy to use methods of optical identification of the pests, which would assist human in the decision-making process. The specialized informatics systems based on the methods of computer image analysis using

modern methods of artificial intelligence are more and more often applied for this purpose (Nowakowski et al., 2009). Recently, the neural modeling has begun to play an important role in this field, and in particular the neural identification of information encoded in the graphic form (Nowakowski et al., 2011). One of the most efficient classifying neural network topologies are the Kohonen networks, also known as self organizing feature maps (SOFM). This type of identification was applied in case of five economically important orchard pests. The larval stages of these pests feed on the surface of the plants (Gagne and Harris, 1998; Hodges and Williams, 2003; Suckling et al., 2007; Japoshvili et al., 2008; Cross and Hall, 2009) or after short feeding on the surface try to bite inside the shoots (Cravedi and Ughini, 1992; Guarino et al., 2002; Kutinkova et al., 2006; Kutinkova, 2008; Ayberk

*Corresponding author. E-mail: pilarski@up.poznan.pl. Tel: +48618466078. Fax: +48618487157.

et al., 2010). These species occur in the Palearctic region and cause a significant annual decline of the fruit yield.

The SOFM networks are most frequently used for a widely understood classification (Slosarz et al., 2011; They perform this task in a relatively untypical way. Because the processing of output values is carried out within the post processing, the result of the network activity is an output variable of nominal character. Each value of this variable represents one definite class. Appropriate neurons occurring in the output layer of the network correspond to particular classes. The connection of the neuron with a given class is indicated by the label with the class name attributed to it. During the action of the network, after the input signal has appeared, the winning neuron is indicated each time (that is the neuron of the lowest activation, which indicates the highest compatibility of weights and the given input pattern). The label of the neuron determines the class to which the input case is ascribed. This untypical structure enables the user to define the output layer of the Kohonen network as a specific two-dimensional "map" of the input multi-dimensional data set (Kohonen, 1982). It enables to place in it an optional number of neurons that are distinguished and established points in this map (Tadeusiewicz and Kohorda, 1997).

The topology of the SOFM network differs considerably from the structures of other neural networks (Kohonen, 1982; Kohonen, 1990). The structure discussed is basically a one-layer network. It consists of an input layer and an output layer which processes the data presented. The output layer is built of radial neurons. This layer is also defined as a layer forming a topological map. Neurons in the layer forming the topological map are taken into consideration as if they were placed in space according to some predetermined order. Usually for convenience and better perception we imagine them to be nodes of a two-dimensional network (Boniecki et al., 2009, 2011, 2012). The paper presents usage of the SOFM neural network model for classification and identification of five selected larvae, economically important fruit tree pests, made on the basis of information encoded in the form of digital images. The color of classified insects larvae was adopted as a representative criterion (Kurdthongmee, 2008; Mirik et al., 2006).

MATERIALS AND METHODS

Characteristics of selected orchard pests

The research material for recognition of images presenting orchard pests consisted of photographs obtained from the book entitled "Pests on fruit trees" written by Wiech (1999). The subject of recognition concerned selected pests of fruit trees. An orchard is an excellent location for different types of organisms to feed. A quick recognition of the danger (optimally in the larval state) can protect the orchard-owner from heavy losses caused by the appearance of a pest. The presentation was limited to five species of pests occurring in the larval state. The insects belong to different families and feed on different fruit trees. The following species have been

selected:

Pear leaf curling midge - Dasineura piri (Bouché 1847)

Systematic: It belongs to the Order, Diptera, and Family, Cecidomyiidae. It is a small dipteran of a dark, brown-black body, about 2 mm long, in appearance resembling a mosquito, with long legs and antennae. It has white elongated eggs difficult to notice. Its cylindrical larvae is white, legless, jumpy, and grows up to 2 mm length. The host plants are pear trees. The dominant colour of this pest is yellow.

Apple leaf curling midge - Dasineura mali (Kieffer 1904)

Systematics: It belongs to the Order, Diptera, and Family, Cecidomyiidae. It has a small size – about 1.5 to 2 mm dipteran – in appearance resembling a mosquito. It has red elongated eggs. Larvae of the midges are cylindrical; initially they are cream white, later become orange or orange-red; they are legless, and move in a characteristic rolling way. The host plants are apple trees. The dominant colour of this pest is white.

European fruit lecanium - Parthenolecanium corni (Bouché 1844)

Systematics: The Order here is Hemiptera and Family, Coccidae. Females have reduced limbs and do not have wings. Their body is covered in hardened cherry-brown "coat" of cuticle in the shape of a convex 3 to 4 mm long bowl. They spend their lives motionless, attached with the proboscis to the background. Males are winged and able to fly, have reduced oral organs. Their body length is 2.5 mm, and width, 1 mm. They feed most frequently on apple trees and plum trees but they attack other plants as well. The dominant colours of this pest are yellow and brown.

Leopard moth - Zeuzera pyrina (Linneus 1761)

Systematics: It belongs to the Order, Lepidoptera, and the Family, Cossidae. It is a butterfly of white, black speckled wings, with wingspan of 5 to 7 cm. It has oval, pale yellow eggs. Its caterpillar is yellow with two rows of black warts. It feeds on many different species of trees and is dangerous first of all in young orchards. In July and August, females lay (single) eggs on the bark, leaf tails and in other places. Caterpillars feed initially collectively under the bark; in the second year they hollow tunnels in branches. The dominant colour of this pest is brown colour.

European goat moth - Cossus cossus (Linneus 1758)

Systematics: The Order here is Lepidoptera while the Family is Cossidae. It is a very large butterfly with grey wings, and wingspan of 7 cm (male) to 9 cm (female). The body of the butterfly is thick and hairy. The caterpillar is pink-red, up to 10 cm long. They attack different fruit, park and forest trees. Caterpillars hollow corridors under the bark (the first year of development) as well as tunnels in the trunk and thick boughs (second year). The dominant colour of this pest is red colour.

Neural analysis

To create the Kohonen network, the Statistica Neural Networks v.8.0 program was used. The input data consisted of 5 input

Table 1. RMS errors and performance of the best SOFM neural networks.

Type	Inputs	LError	VError	TPerformance	VPerformance	Training
SOFM	12	0.05532	0.0762	0.7638889	0.7266667	KO100724 ^b

LError and VError : RMS error for the learning file and the validation file, LPerformance and VPerformance: SOFM performance for the learning file and the validation file, Training means SOFM learning methods: KO 100724b indicates that the SOFM (Center Assignment) algorithm was used, that the best network discovered during that run was selected (for "best" read "lowest verification error") and that this network was found on the 100724 epoch.

variables and one nominal output variable (that served to label the topological map after completing the teaching process).

In order to obtain the representation of knowledge appropriate for the neural network that is, to process the information included in the image into a sequence of numbers, the previously created original computer system called „PictureSOFM” was used (Kurdthongmee, 2008; Mirik et al., 2006; Penn, 2005; Garcia and Gonzales, 2004). It serves to digitalise the image presented in the file format .bmp (of previously defined dimensions), carrying out the classification based on the model of RGB colours. The result of application operation is then recorded in the text file (Adebayo et al., 2012). The result is presented in the form of a sequence of numbers of a length adequate for the adopted number of pixels in the marked fragment of the image for example, 2 × 2 generates a sequence of 12 numbers changing within the range from 0 to 255.

The text file generated using the „PictureSOFM” program was used to create a teaching set for the designed Kohonen neural network, whose aim was to carry out classification based on the model of RGB colours (Mirik et al., 2006; Kurdthongmee, 2008). Making use of the „PictureSOFM” program, 540 cases were generated per one insect, which (for 5 pests) resulted in a set of 2700 teaching cases. The data file were divided standard randomly into learning file, validating data and testing file. The technique of random mixing of cases was used (Alhoniemi et al., 1999).

In the process of teaching of the SOFM type neural network, the knowledge of the output variable is not necessary since the process of teaching this network develops in an “unsupervised” way. The teaching of the SOFM type neural network was implemented in two separate stages. In the first phase, a high teaching coefficient was applied (~0.7) in combination with a large range of neighbourhood (~1). At this stage, modification of neuron weights was the most extensive. However, in the second phase of teaching a small value of teaching coefficient was used (~0.2 to 0.1) in combination with small neighbourhood (equal to 0). At this stage, the weights of neurons were modified to a minimum extent and the process of the so-called repeated and supplementary teaching of the network was carried out.

RESULTS

The most commonly used measure of quality of Kohonen neural network is an error root mean square error (RMS). RMS error is usually most convenient to interpret a single value that describes the total network error. This is the total error committed by the network on a set of data. It is determined by summing the squares of individual errors and then dividing the resulting sum by the number of included value and the designation of the square root of the quotient obtained.

The results of action of the generated SOFM type neural network are presented separately for the teaching set and the validation set. According to the standard

procedure the teaching set was used in the teaching process of the neural network. To check appropriateness of action of the generated neural network the so-called validation set was used which was randomly isolated from the set of the data possessed. It included cases not presented in the course of teaching the network. It enabled to carry out a verification of quality of the generated and taught neural network. RMS errors and performance of the best SOFM neural networks are shown in Table 1. The error occurs when the network is executed on the training and verification, respectively. This is the root mean square (RMS) of the errors on each individual case, where the error on each individual case is measured by the network's error function. If the network performs classification, the performance measure indicates the proportions of cases which are correctly classified. This takes no account of doubt options, and so a network with conservative Accept and Reject thresholds (confidence limits) may have a low apparent performance, as many cases are not correctly classified. Classification statistics were determined separately for the training set and validation and shown in Table 2. They represent the number of cases classified into the class.

The topological map creates the output layer of the SOFM network. Neurons in this layer are usually placed in two-dimensional space, and every neuron represents one cluster of teaching data. In the case discussed, neurons in the topological map were assigned specific classes representing particular orchard pests. In the generated topological map, the clusters of data are marked with colours assigned as shown in Figure 1. The topological map represents graphically the classification abilities of the generated SOFM network. At the top, the case number (1200) was provided, next the number of the neuron in the topological map (218) and the class to which a given case was classified. After establishing for every neuron the location assigned to it in the topological map, a fill in square is drawn representing the degree of activation of a given neuron, and the bigger squares filled indicate a larger closeness of the model of a given neuron with reference to the tested case. Additionally, the winning neuron is surrounded with a rectangular frame.

DISCUSSION AND CONCLUSION

The acceptance of the colour of the analysed image as

Table 2. Classification statistics for the learning and the validation files.

Parameter	Learning file					Validation file				
	<i>Dasyneura mali</i> Kieff	<i>Dasyneura piri</i> Bonche	<i>Zeuzera pyrina</i> L.	<i>Cossus cossus</i> L.	<i>Parthenolecanium corni</i> Bouche	<i>Dasyneura mali</i> Kieff	<i>Dasyneura piri</i> Bonche	<i>Zeuzera pyrina</i> L.	<i>Cossus cossus</i> L.	<i>Parthenolecanium corni</i> Bouche
	p2	p1	t1	t2	m1	p2	p1	t1	t2	m1
Total	366	354	373	345	362	174	186	167	195	178
Correct	362	306	340	318	339	172	160	151	165	159
Wrong	4	48	33	27	23	2	26	16	30	19
Unknown	0	0	0	0	0	0	0	0	0	0

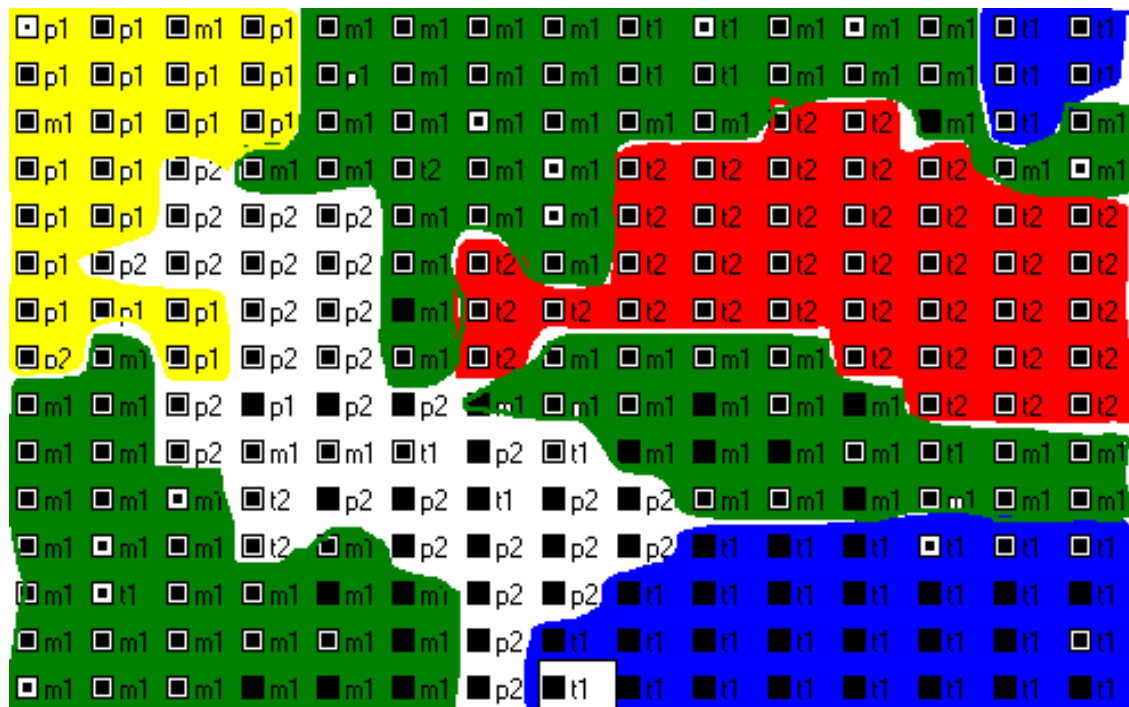


Figure 1. Classifying topological map type Kohonen (for output matrix 15x15). Yellow, *Dasyneura piri* (Bonché) (indicated as p1); Blue, *Zeuzera pyrina* (L.) (indicated as t1); Red, *Cossus cossus* (L.) (indicated as t2); Green, *Parthenolecanium corni* (Bouché) (indicated as m1).

the main attribute has an effect on the quality of action of the taught neural network as a tool for identification of the presented object in the form of an image. It can be supposed that taking into consideration other attributes of the image (texture, exposure to light or other physical features), accuracy of reactions of the taught neural network could be improved. It should be noted that to teach the SOFM network it is sufficient to possess only a set of teaching data – known models are not necessary. The knowledge of names of pests was used only for labelling the SOFM topological map. It facilitated the visualisation of the functioning of the network during its exploitation. It should be emphasized that the SOFM neural network demonstrated also a certain ability to generalize, in such a way that, the cases “similar” to those known to is tried to classify to the classes. The cases unknown to the network (representing other pests) were classified as foreign.

The desirable feature enables the Kohonen neural network to identify the pests correctly based on the presentation of images not originating from the teaching set, that is, noisy photographs taken under different light exposure conditions and using different quality of the equipment. The following conclusions can be drawn:

1. The good quality of SOFM neural network is confirmed by a small errors RMS and good performance. It was generated during the model's operation with the use of learning and validation data file.
2. Similar numerical value LError and VError shows good approximate properties of the generated SOFM neural network.
3. The smallest error network is committed to identifying *Dasyneura mali* Kieffer. This orchard pest has a clear, uniform white colour on the entire surface. Therefore, it is easily identifiable on the digital images, both man-made as well as by the neural model.

ACKNOWLEDGEMENT

The part of the study was supported by grant no. N N312 212136 from the Polish Ministry of Science and Higher Education.

REFERENCES

- Adebayo JA, Rabee R, Ibrahim DK (2012). Neural computing modeling of the reference crop evapotranspiration. *Environ. Model. Softw.* 29:61-73.
- Alhoniemi E, Hollmen J, Simula O, Vesanto J (1999). Process monitoring and modeling using the self-organizing map. *Int. Computer-Aided. Eng.* 6(1):3-14.
- Ayberk H, Hakyemez A, Cebeci H (2010). Light trap surveys for moths in Sile region of Istanbul, Turkey. *Afr. J. Biotechnol.* 9 (29):4624-4630.
- Boniecki P, Dach J, Nowakowski K, Jakubek A (2009). Neural image analysis of maturity stage during composting of sewage sludge. *International Conference on Digital Image Processing (ICDIP 2009)*, 7-9 March 2009, BANGKOK, PROCEEDINGS: 200-203 DOI: 10.1109/ICDIP.2009.85.
- Boniecki P, Nowakowski K, Tomczak RJ (2011). Neural networks type MLP in the process of identification chosen varieties of maize. 3rd International Conference on Digital Image Processing (ICDIP 2011), April 15-17, 2011, CHENGDU, PROCEEDINGS of SPIE Volume: 8009 Article Number: 800908 DOI: 10.1117/12.896184.
- Boniecki P, Dach J, Pilarski K, Piekarska-Boniecka H, (2012). Artificial neural networks for modeling ammonia emissions released from sewage sludge composting. *Atmospheric Environment* 57: 49-54.
- Cross JV, Hall DR (2009). Exploitation of the sex pheromone of apple leaf midge *Dasineura mali* Kieffer (Diptera: Cecidomyiidae) for pest monitoring: Part 1. Development of lure and trap. *Crop Prot.* 28:139-144.
- Cravedi P, Ughini V (1992). Monitoring of some injurious Lepidoptera in quince orchards. *Acta Phytopathologica Et Entomol. Hungarica* 27 (1-4):185-188.
- Gagne RJ, Harris MO (1998). The distinction between *Dasineura* spp. (Diptera: Cecidomyiidae) from apple and pear. *Proceeding of the Entomological Society of Washington.* 100(3):445-448.
- Garcia HL, Gonzalez IM (2004). Self-organizing map and clustering for wastewater treatment monitoring. *Eng. Appl. Arti. Intell.* 17(3):215-225.
- Guario A, Marinuzzi V, Bari G (2002). Preliminary results of field control of *Zeuzera pyrina* L. (Lepidoptera, Cossidae) in Apulia. *Acta Horticulturae* 586:815-817.
- Hodges GS, Williams MC (2003). Description and illustration of the first instars of *Parthenolecanium corni* (Bouche) and *Parthenolecanium quercifex* (Fitch) (Hemiptera: Coccidae). *J. Entomol. Sci.* 38(4):489-493.
- Japoshvili G, Gabroshvili N, Japoshvili B (2008). The parasitoid complex of *Parthenolecanium corni* Bouché (Hemiptera: Coccoidea) in the city of Tbilisi and its surroundings and comparison with some other European countries. *Bull. Entomol. Res.* 98(1):53-56.
- Kohonen T (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics.* 43: 59-69.
- Kohonen T (1990). Improved versions of learning vector quantization. *International Joint Conference on Neural Networks San Diego, CA.* 1:545-550.
- Kutinkova H (2008). Seasonal monitoring of two xylophagous orchard pests by pheromone traps in Bulgaria. *Acta Hortic.* 767:411-415.
- Kutinkova H, Adreev R, Arnaoudov V (2006). The leopard moth borer, *Zeuzera pyrina* L. (Lepidoptera: Cossidae) – important pest in Bulgaria. *J. Plant Protect. Res.* 46(2):111-115.
- Kurdthongmee W (2008). Colour classification of rubberwood boards for fingerjoint manufacturing using a SOM neural network and image processing. *Comput. Elect. Agric.* 64:85-92.
- Mirik M, Michels G, Kassymzhanova-Mirik S, Elliott NC, Catana V, Jones DB, Bowling R (2006). Using digital image analysis and spectral reflectance data to quantify damage by greenbug (Hemiptera: Aphididae) in winter wheat. *Comput. Electr. Agric.* 51:86-98.
- Nowakowski K, Boniecki P, Dach J (2009). The identification of mechanical corn kernels damages basis on neural image analysis. *International Conference on Digital Image Processing (ICDIP 2009)*, 7-9 March 2009, BANGKOK, PROCEEDINGS: 412-415 DOI: 10.1109/ICDIP.2009.78.
- Nowakowski K, Boniecki P, Tomczak RJ (2011). Identification process of corn and barley kernels damages using neural image analysis. 3rd International Conference on Digital Image Processing (ICDIP 2011), April 15-17, 2011, CHENGDU, PROCEEDINGS of SPIE Vol. 8009 Article Number: 80090C DOI: 10.1117/12.896664.
- Penn BS (2005). Using self-organizing maps to visualize high dimensional data. *Comput. Geosci.* 31(5):531-544.
- Suckling DM, Walker JTS, Shaw P, Manning LA, Lo P, Wallis R, Bell V, Sandanayaka RM, Hall DR, Cross JV, E-Sayed AM (2007). Trapping *Dasineura mali* (Diptera: Cecidomyiidae) in Apples. *J. Econ. Entomol.* 100(3):745-751.
- Slosarz P, Stanisz M, Boniecki P, Przybylak A, Lisiak D., Ludwiczak A (2011). Artificial neural network analysis of ultrasound image for the estimation of intramuscular fat content in lamb muscle. *Afr. J. Biotechnol.* 10 (55):11792-11796.
- Tadeusiewicz R, Korohoda P (1997). *Computer Analysis and Image*

Processing. Progress of Telecommunication Foundation Publishing
House, Krakow.
Wiech K (1999). Pests of fruit trees. Plantpress, Krakow.