

FACTA UNIVERSITATIS

Series: **Automatic Control and Robotics** Vol. 12, N° 1, 2013, pp. 19 - 30**INTELLIGENT CLASSIFIERS FOR NON-DESTRUCTIVE
DETERMINATION OF FOOD QUALITY***UDC (543.42), (664.64.016:641)***Tania P. Titova, Veselin G. Nachev,
Chavdar I. Damyanov, Plamen I. Nikovski**Department of Automatics, Information and Control Systems,
University of Food Technologies, Plovdiv, Bulgaria

Abstract. *The paper analyzes the possibilities to non-destructively determine food quality (potatoes, eggs) by means of the spectra of transmission in the visible and near-infrared regions of the electromagnetic spectrum. The research includes the creation and testing of a training sample of representative samples and the evaluation of the possibilities for classification using Neural Classifier and Support Vector Machines method (SVM).*

Key words: *non-destructive quality evaluation, pattern recognition, food quality, classifiers*

1. FOOD QUALITY

The recent developments in the food and flavour industry have been marked by an on-going upsurge in the requirements concerning food product quality. This necessitates the search for new methods for food product quality evaluation and sorting, which turns out to be a difficult task.

Undoubtedly, the most essential and understandable aspect of the quality of foods is their safety related to human health. The consumer at the end of the food chain has the right to demand and expect to consume quality, safe, and absolutely harmless food products.

One of the most promising methods for non-destructive quality evaluation of various quality-related parameters of foods is the usage of spectroscopy in the visible (VIS) and near-infrared (NIR) spectral regions. It is known that most complex organoleptic and functional characteristics for food product quality evaluation correlate with their VIS and NIR spectral ones [5, 13]. Typical are the VIS and NIR spectroscopic methods which have great many applications and have proven to possess the following advantages:

Received March 15, 2013

Corresponding author: Tania P. Titova

Department of Automatics, Information and Control Systems, University of Food Technologies, Plovdiv, Bulgaria

Email: t_titova@abv.bg

- they are proven-in-practice methods for the routine compositional, functional, and sensory analysis of food components, ingredients, and final products;
- they do not require preliminary processing of the product under study or damage its wholeness;
- the analysis is conducted easily and quickly;
- they are also suitable for an on-line control of processes.

Egg quality

Kramer (1951) was the first to define the term "quality" – as the property of some foods to have an impact on consumers [3].

Egg quality is a general term which adheres to standards determining the internal and external state of eggs. External egg quality has to do with the egg shell, its purity, structure, shape, etc., whereas internal egg quality is connected to the egg white – its purity and viscosity, the size of the air cell, the shape and coloring of the egg yolk, etc.

The market and manufacturing criteria for egg quality amount to requirements concerning egg freshness, the presence of internal and external defects, weight, egg density and shape (egg index, asymmetry), the ratio between the size of the albumen substance and the egg yolk weight and the size of the germinal disc [2, 5, 12].

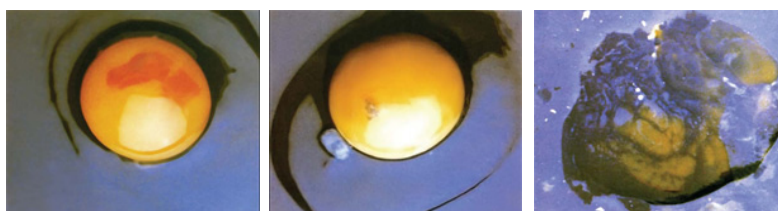


Fig. 1 Internal egg defects – a) Blood spots; b) Meat spots; c) Rotten eggs.

The literature on the problem demonstrates that the optical spectral methods for egg quality evaluation can be thought of as a rival alternative to visual ovoscopy. Back in 1937, Harvey substantiated the possibility for using X rays in order to discover the internal defects of fruits and vegetables, especially the internal cavities of potatoes [14]. It was his idea to use optical methods in the non-destructive quality evaluation and sorting of eggs, which in the 1960s-1980s was successfully applied by Karl Norris [8] to the quality evaluation of fruits, vegetables, and eggs.

The spectrophotometry of eggshell colour, egg yolk colour, blood and meat cells constitutes the first attempt at an automated egg quality control. The initial identification of blood and meat cells is based on the measurement of light transmission at 575 nm where hemoglobin has an absorbing band. The disadvantage of this method lies in the dependence of transmitted light on the size and colour of the egg shell, diseases and contamination. The determination of food product defects, based on spectral analysis, has to do with the fact that fresh white eggs have maximum light transmission at $573\text{-}580\text{ nm}$, brown eggs have maximum light transmission at $590\text{-}596\text{ nm}$ and $636\text{-}638\text{ nm}$, and eggs with blood cells have an absorbing maximum at $538\text{-}540\text{ nm}$ and $573\text{-}576\text{ nm}$.

The research conducted and the analysis of egg permeability in the wavelength range of 400-1100 nm led to the discovery of sections characteristic of the different quality parameters such as eggshell colour, presence of blood and meat cells, putrefaction, etc.

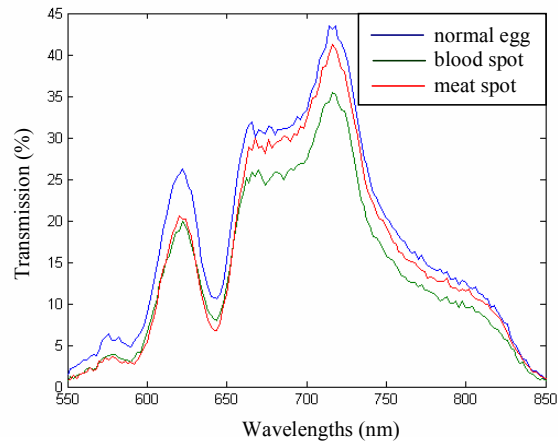


Fig. 2 Spectrograms of a healthy egg and the commonest internal defects – blood spot and meat cell

Figure 2 shows the spectral characteristics of three of the commonest defects – a healthy egg with a blood spot and a meat cell. The characteristics manifest that most of the information concerning quality is carried by the 550-850 nm region where the permeability of the defect-bearing eggs is lower than those of the healthy ones.

Potato quality

Potatoes are products with high nutritional value, so they have an important stake in the daily meals. They are the group of 15 food products of exceptional importance for humanity. The main difficulty in the automatic qualification of potatoes are associated with a wide variety of defects and diseases (external and internal). Greening is one of the most common defects affecting both appearance and taste. Another defect is darkening due to mechanical damage, which leads to rot of potato tubers changes.

Potatoes are classified on the basis of the presence and degree of damage of the tubers (internal and external defects, diseases, genetic anomaly, injuries, etc.). The expert classifies each tuber into one of the following classes - D_1 (1st quality), D_2 (2nd quality) и D_3 (3rd quality). Figure 3 presents examples of the realizations of a signal $U(n)$ which correspond to different product qualities - D_1 , D_2 , D_3 . Each realization $U[n]$ (Fig. 3) of the light transmission coefficient consists of $L=25$ values, registered by the longitudinal scanning of the tuber. The $U[n]$ function possesses almost complete spectrophotometric information about the product and is invariant to disturbances and product thickness.

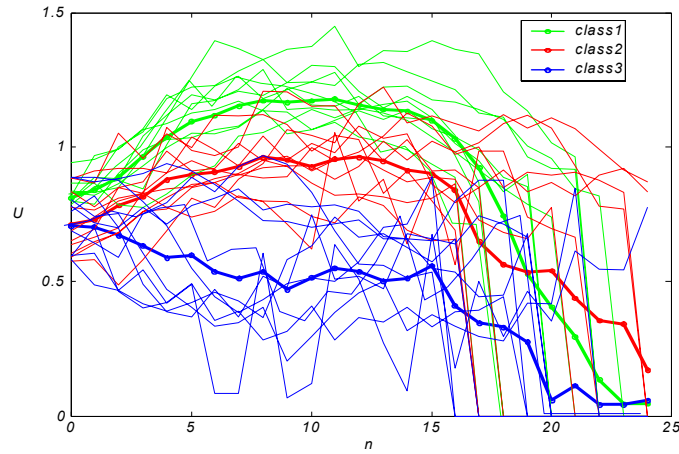


Fig. 3 Images of potato tubers from three quality fractions – D_1 , D_2 and D_3

2. MATERIALS AND METHODS

Support Vector Machines (SVM). SVM is one of the recently established methods in Statistical Learning Theory for regression and classification problems [11, 16]. The definition of the classification problem involving two classes is the following: there are N -class learning images, represented by the vectors (y_i, x_i) , $i=1\dots N$ where: y_i is a class label; $x_i = [a_3^i; d_3^i; d_2^i; d_1^i] \in \mathbb{R}^d$ – a feature vector. The classifier is represented by the function $f(x, \mathbf{w}, b) : x \rightarrow y_i$. SVM is used to determine an optimal separating hyperplane so that the learning examples with notations for two classes +1 and -1 are situated on both sides of the hyperplane at a maximum distance (margin) between the boundary support vectors for the two classes (Fig. 4).

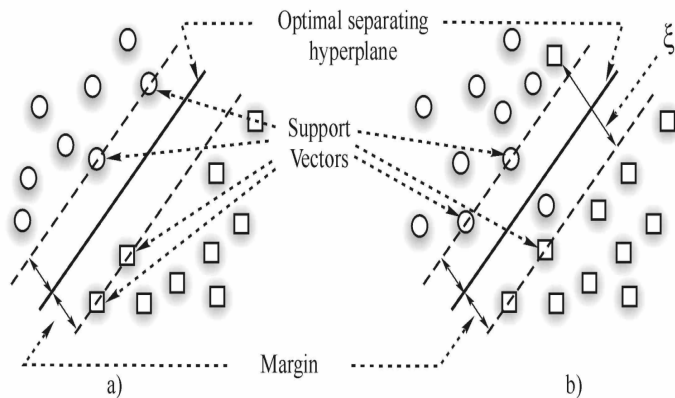


Fig. 4 An optimal hypersurface, linear (a) and non-linear (b) class separation.

The solution is defined by the function $\mathbf{w}\mathbf{x} + b = 0$ with parameters (\mathbf{w}, b) . For vectors which do not cut across the hyperplane, the following conditions are met: $\mathbf{w}\mathbf{x} + b > 0$ or $\mathbf{w}\mathbf{x} + b < 0$. Thus, the classifier is defined as $f(\mathbf{x}; \mathbf{w}, b) = \text{sign}(\mathbf{w}\mathbf{x} + b)$. Support vectors lie on two parallel hypersurfaces $\mathbf{w}\mathbf{x} + b = \pm 1$. The maximization of the margin with these two equations results in an optimization problem with the following additional conditions:

$$\min \left\{ \frac{1}{2} \|\mathbf{w}\|^2 \right\} \text{ for } y_i(\mathbf{w}\mathbf{x} + b) \geq 1, i = 1 \dots N \quad (1)$$

Since the learning vectors are not always linearly separable, in (1) a regularization parameter C is introduced as well as variables which account for errors ξ_i in the decision function. In this case, the problem is expressed in the following way:

$$\begin{cases} \min \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \right\}; \\ y_i(\mathbf{w}\mathbf{x} + b) \geq 1 - \xi_i; \\ \xi_i \geq 0, i = 1 \dots N. \end{cases} \quad (2)$$

To solve the optimization problem (2), a Lagrange form of dual is used to get:

$$\begin{cases} \max \left\{ \sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j y_i y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \right\}; \\ C \geq \lambda_i \geq 0, i = 1 \dots N; \\ \sum_{i=1}^N \lambda_i y_i = 0, i = 1 \dots N. \end{cases} \quad (3)$$

Lagrange variables λ_i have a non-zero value only for the support vectors.

SVM can be generalized in the case of non-linearly separable surfaces. Typical for this method is the projection of data into a new high-dimensional space where the data are linearly separable. The operation in this space is simulated via the so-called Kernel Method. If the vectors $\mathbf{x} \in \mathfrak{R}^d$ map into a new space H , through the function $\Phi: \mathfrak{R}^d \rightarrow H$, then the product $\langle \mathbf{x}_i, \mathbf{x}_j \rangle$ from (3) is transformed with $\langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$. The product $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$ is a kernel which is directly used in (3), without explicit computing of Φ ("kernel trick").

The SVM analytic function is defined only for the separation of two classes. For multi-class separation (M -class), two approaches are used: One-Against-All – M classifiers are applied iteratively to separate the point of each class from all the others; One-Against-One – $M(M-1)/2$ classifiers separate all class pairs, the final solution being the class determined most frequently [4, 18].

SVM makes use of the so-called *kernel*. The most commonly used kernels are linear kernels, polynomial kernels, and radial basis functions:

$$K(\mathbf{x}, \mathbf{z}) = \langle \mathbf{x}, \mathbf{z} \rangle; \quad K(\mathbf{x}, \mathbf{z}) = (1 + \langle \mathbf{x}, \mathbf{z} \rangle)^n; \quad K(\mathbf{x}, \mathbf{z}) = \exp(-\|\mathbf{x} - \mathbf{z}\| / 2\sigma^2);$$

Neural classifier. In recent years, artificial neural networks (ANN) have been widely studied and tested for various applications. Nevertheless, there remain some questions which are still unsolved, which substantiates follow-on research. One of these questions has to do with the determination of a suitable size and topology of the network. The construction of a neural network (NN) is a complicated task because in most cases it is very difficult to find out its optimal configuration. If the network possesses too many free parameters, i.e. weights, neurons, etc., the problems with the local minimum and over-training cannot undoubtedly be avoided. In principle, it is difficult to discover a general quantitative criterion to optimize the overall functioning of the network. In most cases, the optimal synthesis of a feed-forward network is, in fact, to find rational network architecture or approximate the best possible one which will solve the problem. The criteria for evaluation usually belong to some of the characteristics: rate of training, complexity of the network, precision of training, generalization properties, etc. An essential issue concerning the synthesis of a neural network is the complexity of the model and the accuracy of training, which directly influence the generalization properties of the network [5, 7, 17].

Different structures of the neural network (NN) have been studied and the most suitable one has been taken to be the feed-forward NN which constitutes a multilayer perceptron (MLP) (Fig. 5).

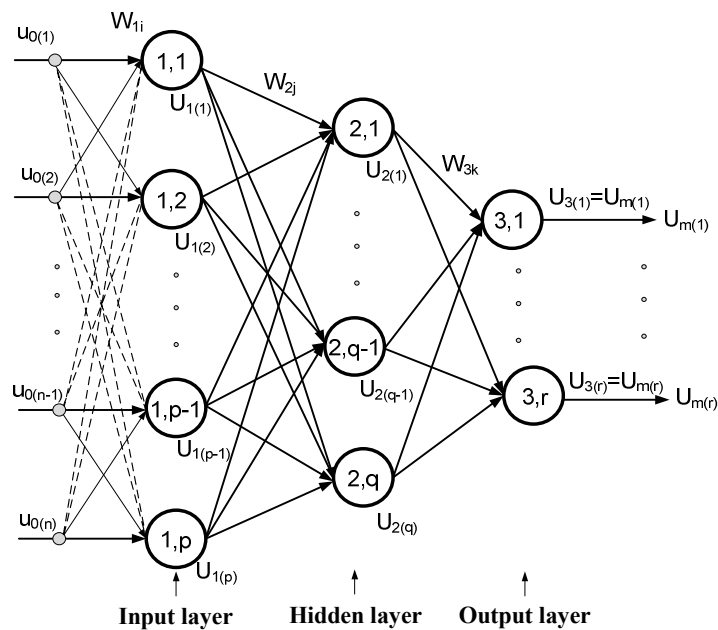


Fig. 5 Structure of Neural Networks

The NN consists of an input layer, an internal layer (or layers) and an output layer. Each layer contains simple process elements, "neurons", connected in a certain manner to the layer following, the connections having specific weights W that determine the knowledge accumulated in the network [10, 15].

3. CALCULATION ERRORS

In the practice of automatic grading and sorting of food products, the classes are more than two, but the evaluations of the credibility of recognition in general include mainly two types of errors: *highway and marginal*.

Highway (major, basic) error is calculated according to the expression:

$$e_i = \frac{\sum_{k=1}^N m_{ik} - m_{ii}}{m_i^r} \times 100, \% , \quad (4)$$

where: $i = 1, 2, \dots, N$, is the number of the respective class; m_i^r – the number of objects (products), assessed on the basis of expert evaluation (EE) in class i ; m_{ik} – the number of objects according EE in class i , which after the algorithmic classification (AC) have ended up in class k by mistake; m_{ii} – the correctly identified objects in the class with the number i , or the number of the objects with EO in class i , left after the algorithmic classification in class i .

The highway error is a constant for the control algorithm and does not depend on the distribution of the objects in the sample. It can also be said that with a large sample the highway error does not depend on the quantity of the objects in the respective class.

Marginal (alternative, field) error is determined by the expression:

$$g_i = \frac{\sum_{k=1}^N m_{ki} - m_{ii}}{\sum_{k=1}^N m_{ki}} \times 100, \% . \quad (5)$$

The marginal error reflects the performance of the algorithm with all the classes simultaneously, which means that the ratio between the quantities of products in separate classes affects the error.

In a synthesized form the error is the difference between the expert evaluation and result of the program realization of the algorithm after training.

Summarizing the bialternative case of error, the highway error shows how many objects have mistakenly gone to other classes, i.e. how many objects are not assigned to the class to which they actually belong, and the marginal – how many objects have mistakenly come from other classes and in this sense are equivalent to the type I and type II errors in mathematical statistics.

If the highway error characterizes the accuracy with which the algorithm classifies a given class, i.e. what percentage of objects wrongly "flow" out of the respective class to other classes, on the other hand the marginal error reflects the "pollution" of a given class with objects from other classes.

The mean values for the highway and marginal errors can be calculated respectively by the expressions:

$$\bar{e} = \frac{1}{N} \sum_{i=1}^N e_i ; \quad \bar{g} = \frac{1}{N} \sum_{i=1}^N g_i . \quad (6)$$

The averaged errors (6) are not equivalent for those who accept the solutions of the classification. In the recognition system they reflect respectively the risk of the developer and the risk of the user.

Total (resultant) error is determined by the formula:

$$e_0 = \frac{100}{m} \sum_{i=1}^N \left(\sum_{k=1}^N m_{ik} - m_{ii} \right), \%$$

$$\text{or} \quad e_0 = 1 - \left(\sum_{i=1}^N m_{ii} \right) / m = 1 - Acc.$$

where: $m = \sum_{k=1}^N m_k^c$ is the total number of objects in the training sample. The total error can also be calculated by the equivalent formula:

$$e_0 = \sum_{i=1}^N e_i P(D_i), \quad (7)$$

where $P(D_i) = m_i^c / m$, ($i = 1, 2, \dots, N$) are the a priori probabilities of distribution of the objects of the respective classes i . According to the rule: $\sum_{i=1}^N P(D_i) = 1$, because the features $S \in D_i$ form a complete group of incompatible events.

After the training the performance of the algorithm, and more precisely its accuracy, is evaluated using a control (test) sample. The resulting error of the test is shown in *table of errors*, which in the particular case of Table 1 provides a general picture of the errors in four classes of recognition ($N = 4$). The structure of the errors corresponding to Table 1 is graphically presented with the graph of Fig. 6.

The formulas for the highway and marginal errors have a universal application and can be used for different recognition algorithms: from neuron type [6], in selecting the complex of features through a genetic algorithm or Karhunen–Loève decomposition [1, 9] and others.

Table 1 Table of errors

EE_j AC_j	m_{i1}	m_{i2}	m_{i3}	m_{i4}	m_i^r	e_i
m_{1j}	m_{11}	m_{12}	m_{13}	m_{14}	m_1^r	e_1
m_{2j}	m_{21}	m_{22}	m_{23}	m_{24}	m_2^r	e_2
m_{3j}	m_{31}	m_{32}	m_{33}	m_{34}	m_3^r	e_3
m_{4j}	m_{41}	m_{42}	m_{43}	m_{44}	m_4^r	e_4
m_j^c	m_1^c	m_2^c	m_3^c	m_4^c	m	\bar{e}
g_j	g_1	g_2	g_3	g_4	\bar{g}	e_0

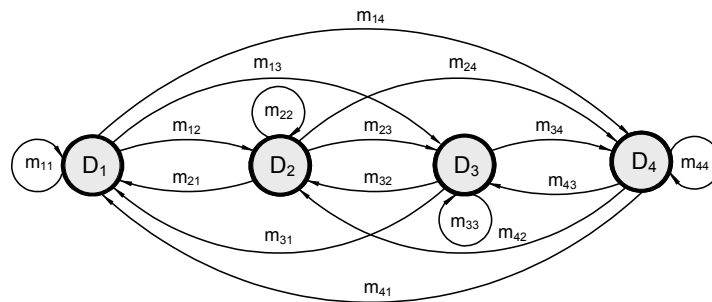


Fig. 6 Graph of the errors ($N = 4$)

4. EXPERIMENTALLY OBTAINED RESULTS

Forming a training sample of eggs. During the experimental studies, the permeability of 95 eggs was measured, after which expert evaluation was performed and the eggs were subdivided into three groups: 28 healthy eggs (with no defects), 37 eggs with blood spots and 30 eggs with putrefaction and mould (Table 2). The measurements were made in the VIS and part of the NIR region in the 450-900 nm range with the USB4000 spectrophotometer of Ocean Optics (USA). In addition, on the basis of the literature on the problem, five intermediate spectrum points – 561, 577, 581, and 640 nm – were included.

On the basis of a metric proximity criterion, the spectral pattern of the objects was preliminarily reduced to 25 frequencies of permeability.

The composition of the testing sample and that of the training sample were selected at random, 60% of the sample as proportional to each class is used for classifier training and 40% - for testing. The distribution of the objects in classes is shown in Table 3.

Table 2 Quality groups (classes) for quality evaluation on the basis of internal egg quality.

Quality groups	Composition/characteristics
Class 1 - D_1	Absence of blood spots or meat spots
Class 2 - D_2	Presence of blood spots and meat spots
Class 3 - D_3	Presence of mould

The samples used in the experiment are shown in Table 3.

Table 3 Composition of the training sample and the testing sample when determining internal egg defects.

Classes	Training sample	Testing sample	Total number of objects
Class 1	10	18	28
Class 2	14	23	37
Class 3	12	18	30
Total	36	59	95

Forming a training sample of potatoes. The reference evaluation and classification of the potato tubers was carried out with the help of the methodology outlined in [5]. It is worth pointing out that the recognition algorithm does not diagnose the disease of the infected potato but identifies only the presence of a defect in the tissue. As a result of scanning, signals are received which are proportional to the optical densities of transmission OD_T and reflection OD_R . The signals have different shapes depending on the respective wavelength, the shape and size of the tubers, the presence and location of the defects.

Table 4 shows the distribution by class quality sample of potatoes.

Table 4 Quality groups (classes) for quality evaluation potatoes

Quality groups	Composition/characteristics
Class 1 - D_1	without external and internal defects
Class 2 - D_2	light mechanical damage, impaired surface, atypical coloration
Class 3 - D_3	waste, various decay and others disease states, greening

Table 5 shows the distribution of objects in the class, and their division in the training and testing samples.

Table 5 Composition of the training sample and the testing sample when determining potatoes defects

Classes	Training sample	Testing sample	Total number of objects
Class 1	754	379	1133
Class 2	813	426	1239
Class 3	562	260	822
Total	2129	1065	3194

Classifiers. These two methods are used in order to guarantee the acquisition of the best solution for the given problem in view of its description and the sample available.

The choice of classifiers to determine internal egg quality and potato quality are preceded by a considerable number of experiments and analyses. The preliminary selection is necessitated by the presence of many parameters which in some cases determine the accuracy of the model. Table 6 and Table 7 generalizes the possibilities for the realization of this application.

The architecture of the NN used has a fixed 25/15/3 structure, which means that the NN has a 15-dimensional input vector, 25 neurons in the first layer and 3 neurons at the output. The NN has been trained via applying the "backpropagation" algorithm in a Matlab environment. The structures of the networks and the parameters have been selected by validation with a certain number of experiments. Classification errors have been evaluated after training the network for a given number of epochs.

In order to realize the algorithm for classification with a SVM, a MATLAB environment has been used once more, by applying the multi-alternative synthesis scheme "One against One".

Table 6 Results from testing the classifiers using spectral features and PCA of eggs data.

Model: Classifiers using spectral features							
Classes	1		2		3		$e_o, \%$
Errors	$g_b, \%$	$e_i, \%$	$g_b, \%$	$e_i, \%$	$g_b, \%$	$e_i, \%$	
SVM, Gaussian kernel function	3.2	11.0	7.2	8.0	6.2	0.9	6.1
Neural classifier	2.3	1.8	6.0	7.5	8.1	4.0	4.9
Model: Classifiers using PCA							
Classes	1		2		3		$e_o, \%$
Errors	$g_b, \%$	$e_i, \%$	$g_b, \%$	$e_i, \%$	$g_b, \%$	$e_i, \%$	
SVM, Gaussian kernel function	0.7	5.2	6.0	8.6	8.3	4.0	5.5
Neural classifier	0.2	0.0	3.2	1.6	6.3	0.0	1.9

Table 7 Results from testing the classifiers using spectral features and PCA of potato data.

Model: Classifiers using spectral features							
Classes	1		2		3		$e_o, \%$
	$g_i, \%$	$e_i, \%$	$g_i, \%$	$e_i, \%$	$g_i, \%$	$e_i, \%$	
Errors							
SVM, Gaussian kernel function	4.7	6.8	8.3	7.3	5.7	8.4	6.8
Neural classifier	4.3	7.8	10.9	8.5	7.6	10.2	8.2
Model: Classifiers using PCA							
Classes	1		2		3		$e_o, \%$
	$g_i, \%$	$e_i, \%$	$g_i, \%$	$e_i, \%$	$g_i, \%$	$e_i, \%$	
Errors							
SVM, Gaussian kernel function	4.1	5.9	7.2	6.4	4.9	7.3	5.9
Neural classifier	3.8	7.1	9.8	7.6	6.8	9.2	7.4

5. CONCLUSION

The results from the study substantiate the following conclusions:

1. Light permeability may be used as a characteristic which helps the objective determination of internal and external food quality.
2. The problem of non-destructive quality assessment of potatoes the best results in the evaluation of classifiers is obtained using the SVM method, while the determination of internal defects eggs best results were obtained when shared use of neural classifier and PCA.
3. The evaluations of the classification errors in testing deviate to some extent from the testing and validation errors. This is due to the relatively small samples. When the samples are small, the differences in the partial values for each error influence the total value in the case of a disbalance of samples.

REFERENCES

- [1] A. Forbes, "Classification-algorithm evaluation: Five performance measures based on confusion matrices," *Journal of Clinical Monitoring and Computing*, vol. 11, pp. 189–206, 1995. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/7623060>
- [2] A. J. Coutts, Gr. C. Wilson, *Optimum egg quality - A practical approach*, Sheffield, Great Britain, 2007.
- [3] A. Kramer, What is quality and how it can be measured: From a food technology point of view. In *Market Demand and Product Quality. Mktg. Res. Workshop Rept.*, Michigan State College, 1951.
- [4] C. Bahlmann, B. Haasdonk, H. Burkhardt, "Online handwriting recognition with support vector machines - A kernel approach," in *Proc. of the 8th International Workshop on Frontiers in Handwriting Recognition (IWFHR)*, Ontario, Canada, 2002, pp. 49–54. [Online]. Available: <http://dx.doi.org/10.1109/IWFHR.2002.1030883>
- [5] Ch. Damyanov, *Non-destructive quality evaluation in automatic sorting systems*, UFT Academic Publishing House, Plovdiv, p. 360, 2006. (in Bulgarian)
- [6] Ch. Damyanov, "Analysis and evaluation of errors in training and testing of recognition systems," *Information, Communication and Control Systems and Technologies ICCST*, vol. 1, no. 1, pp. 13–19, 2012. [Online]. Available: <http://icgst.uni-ruse.bg/>
- [7] Chr. M. Bishop, *Neural networks of pattern recognition*, Oxford Clarendon Press, Oxford, 1995.

- [8] K. N. Norris, J. D. Rowan, "Automatic detection of blood in eggs," *Agricultural Engendering*, pp. 154–159, 1962.
- [9] M. Mladenov, *Analysis and evaluation of grain quality*, University publishing center of RU "A. Kanchev", Ruse, 2011. (in Bulgarian)
- [10] J.A. Fomin, G.R. Tarlovskiy, "Statistical theory of pattern recognition," Radio and Communication, Moscow, 1986. (in Russian)
- [11] J. P. Marques de Sa, *Pattern Recognition: Concepts, Methods, and Applications*, Springer, 2001.
- [12] P. Chen, "Quality evaluation technology for agricultural products," in *Proc. of International Conference of Agricultural Machinery Engineering*, Seoul, Korea, 1996, pp. 171–204.
- [13] P. C. William, K. H. Norris, *Near infrared technology in the agricultural and food industries*, American Association of Cereal Chemists Inc., St. Paul, Minnesota, USA, 1987, (2nd ed. 2003).
- [14] R. B. Harvey, "The X-ray inspection of internal defects of fruit and vegetables," *Journal of the American Society for Horticultural Science*, vol. 35, pp. 156–157, 1938.
- [15] R. O. Duda, P. E. Hart, D. Stork, *Pattern Classification*, 2nd eds., New York: John Wiley, 2001.
- [16] V. N. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.
- [17] W. J. Wang, X. C. Tang., W. C.Li, "A variable structure neural networks model and its applications," in *Proc. of IEEE Region 10 Conference on Computer, Communication, Control and Power Engineering – TENCN '93*, 1993, vol. 2, pp. 799–802. [Online]. Available: <http://dx.doi.org/10.1109/TENCN.1993.320134>
- [18] Y. Lin, Y. Lee, G. Wahba, *Support vector machines for classification in nonstandard situations*, Technical Report 1016. Department of Statistics, University of Wisconsin, Madison 2000.