

## APPLICATION OF CLASSIFICATION TECHNIQUES IN THE ANALYSIS OF EXTERNAL CHARACTERISTICS OF EGGS

Mariya Georgieva-Nikolova, Zlatin Zlatev  
Faculty of Technics and Technologies, Trakia University  
38 Graf Ignatiev str., 8602, Yambol, Bulgaria  
e-mail: maria\_bogomilova@abv.bg

**Abstract:** *In this article a comparative analysis is made to determine the influence of vectors of selected features derived from geometric, optical and dielectric characteristics of eggs on the accuracy of classification, depending on their weight. Suitable for classification are the principal components and latent variables that reduce feature vectors containing shape indices (D, A, V), spectral indices (TVI, GLI), dielectric characteristics (C, k), selected by four methods (CORR, SFCPP, RELIEFF, FSRNCA). By comparative studies it is found that the use of classification methods (DT, DA, SVM) are more effective in predicting weight of hen eggs than in quail eggs. The proposed egg analysis methods take precedence over the known solutions in this field as it takes into account changes in the internal properties of quail and hen eggs when stored.*

**Keywords:** *Eggs parameters and models, Image analysis, Shape features, Spectral indices, Dielectric properties, Feature selection, Classification.*

### 1. INTRODUCTION

The processing hens and quail eggs for consumption consists of basic steps, including collection, cleaning, sorting, packing. In egg production, these processes are fully or partially automated [17,19]. When sorting eggs, part of the processes are automated and others are done manually and above all rely on the expert's visual judgment [18,20]. This type of assessment also determines the nature of the assessments. They are subjective, relatively slow, and of high cost. Therefore, when grading eggs, these methods are inappropriate. The quality requirements of eggs, achieved with the most accurate sorting according to their weight, necessitate the continuous development of new methods or the improvement of existing ones.

In recent years, various methods for determining the quality of eggs have been developed and studied, including optical [2,3,6,9,21,30], gas [36], capacitive, ultrasonic [1] sensors. Methods for sharing different types of sensors are less affected. According to Mladenov [22] and Wang et al. [34], by using combined sensor data, better results can be obtained than using each sensor type individually. A number of methods that are efficient in their classification accuracy, such as the Support vector machines Method, Neural Networks, Genetic Algorithms, have been used to process data from these sensors.

The prediction and classification of eggs, depending on their weight, whether using regression methods or complex computational procedures, is of close accuracy [28,33]. There is an obvious need to look for effective, rapid and simplified classification approaches, methods and tools for an automated, objective and sufficiently accurate assessment of key egg indicators [16]. These methods are suitable for direct use of the production line because they offer an accelerated calculation process and reduce the time for sorting eggs.

The purpose of the article is to determine, by means of a comparative analysis, the effect of vectors of selected features derived from the shape, optical and dielectric characteristics of eggs on the accuracy of classification, depending on their weight.

## 2. MATERIAL AND METHODS

A total of 108 quail eggs from three producers were used. 140 hen eggs from three producers. The eggs were purchased from a commercial network in Yambol, Bulgaria. They are selected without internal and external defects. They were stored at  $4\pm 2^{\circ}\text{C}$  and a relative humidity of  $46\pm 3\text{RH}\%$  for a period of 21 days. Measurements were taken on days 1, 7, 14, and 21 of storage.

The measurement system presented in [12] was used to obtain digital images, spectral and dielectric characteristics. The system is expanded with a lighting module and a single-board microcomputer with a capacitor cell included. Capacitance measurement library [24] was used to determine the capacity. The measurement frequency is 5kHz.

Spectral characteristics of the quail and hen eggs were obtained, which was done by converting the values of the LMS model into transmittance spectra in the VIS region, in the range 390-730nm, according to the mathematical dependences presented in [35].

Shape indexes  $d$ ,  $D$ ,  $A$ ,  $P$ ,  $A_i$ ,  $V$ ,  $A_{mr}$ ,  $K_v$ ,  $K_f$ ,  $K_1$ ,  $c_1$ ,  $r_1$ ,  $KA$ ,  $KAM$  according to [12] have been determined.

Spectral indices  $REI$ ,  $PTI$ ,  $CTI$ ,  $TVI$ ,  $G$ ,  $NExG$ ,  $NGRDI$ ,  $RGBVI$ ,  $GLI$ ,  $VARI$ ,  $ExG$  according to [5,7] are determined.

From the dielectric characteristics were measured and calculated: capacity ( $C$ , pF); electrical resistance ( $R$ ,  $G\Omega$ ); electrical conductivity ( $EC$ ,  $\mu\text{S}$ ); dielectric constant ( $k$ ).

The following methods have been used to select features describing the characteristics of eggs:

- Correlation method, *CORR* [32]. Correlation dependence allows the search for unknown relationships between the characteristics describing eggs. In the selection of features, those that have the least possible correlation with each other are sought;
- Sub-feature method of comparable prediction ability *SFCPP* [27]. Through this method, an optimal set of features is found that are mutually and maximally different and can effectively represent the compared objects. The algorithm minimizes the feature set and maximizes the appropriate feature set describing the objects. This algorithm is suitable for the selection of classification features.
- Character selection method by neighbor component analysis *FSNCA* [10]. Weight the characteristics using a diagonal adaptation of the Neighborhood Component Analysis (NCA) method. This algorithm works well when evaluating the significance of characteristics for distance-based models. The algorithm is suitable for selecting classification features;
- Method for ranking significant forecasting parameters *RELIEFF* [26]. The method is a selection function using the ReliefF algorithm for classification and the RReliefF algorithm for regression. This algorithm works well when evaluating the significance of characteristics for distance-based models;
- Method for selecting regression features by analysis of adjacent components *FSRNCA* [11]. Weight the characteristics using a diagonal adaptation of the Neighborhood Component Analysis (NCA) method. This algorithm works well when evaluating the significance of characteristics for distance-based models. The algorithm is suitable for selecting features for regression analysis;

The presented feature selection methods describing the shape, spectral and dielectric characteristics of eggs have selected feature vectors. Tables 1, 2, 3 and 4 present the selected (HSFV) for hen eggs and (QSFV) for quail eggs containing shape indices and spectral indices. Feature vectors (HCFV) for hen eggs and (QCFV) for quail eggs are also presented.

Table 1. Selected shape indices and spectral indices for hen eggs

Feature vector	Method for selection	Shape indices	Spectral indices
HSFV1	CORR	A, P, Ai, V, Amr	TVI, G, GLI, VARI
HSFV2	SFCPP	d, Ai, Amr, Ka	REI, PTI, G, NGRDI
HSFV3	FSNCA	A, P, V, Kf, K1	PTI, NExG, RGBVI, ExG
HSFV4	RELIEFF	d, P, V, Kf	PTI, CTI, TVI, ExG
HSFV5	FSRNCA	A, P, Amr, c1	CTI, TVI, RGBVI, ExG

Table 2. Selected shape indices and dielectric characteristics for hen eggs

Feature vector	Method for selection	Shape indices	Dielectric characteristics
HCFV1	CORR	A, P, Ai, V, Amr	R, k
HCFV2	SFCPP	D, A, Ai, Amr	R, K
HCFV3	FSNCA	d, V, Amr, Kf	C, R
HCFV4	RELIEFF	d, P, V, Kf	R
HCFV5	FSRNCA	A, P, V, Kf, K1	C

Table 3. Selected shape indices and spectral indices for quail eggs

Feature vector	Method for selection	Shape indices	Spectral indices
QSFV1	CORR	D, Ai, Amr, K1	CTI, NExG, GLI, VARI
QSFV2	SFCPP	A, P, Ai, K1	REI, PTI, TVI, G
QSFV3	FSNCA	V, Amr, Ai, d	CTI, TVI, GLI, VARI
QSFV4	RELIEFF	A, Ai, Amr, c1	REI, G, NExG, RGBVI
QSFV5	FSRNCA	A, V, Amr, K1	CTI, TVI, RGBVI, GLI

Table 4. Selected shape indices and dielectric characteristics of quail eggs

Feature vector	Method for selection	Shape indices	Dielectric characteristics
QCFV1	CORR	D, Ai, Amr, K1	C, R, k
QCFV2	SFCPP	D, P, Amr, K1	R, k
QCFV3	FSNCA	d, P, Amr, K1	k
QCFV4	RELIEFF	A, Ai, Amr	C, R, k
QCFV5	FSRNCA	A, V, Amr, K1	C, k

Principal components (*PC*) obtained by the principal component analysis and latent variables (*LV*) obtained by the method of partial least squares regression were used to reduce the data volume of the spectral characteristics [14,23]. A kernel variant of the principal components (*kPC*) was also applied using a second-order polynomial kernel.

The accuracy of classification with the found feature sets, in combination with methods to reduce the amount of their data, was evaluated using four types of classifiers: Naïve Bayes classifier [4]; Decision Tree Method [15]; Discriminant analysis [22]; Support vector machines method [8,31].

The Naïve Bayes Classifier (*NBC*) was used as a benchmark. It is one of the classic algorithms in machine learning and is based on the Bayes theorem for determining the posterior probability of an event occurring. Accepting the "naïve" assumption of conditional independence between each pair of parameters.

The Decision Tree (*DT*) method classifies data into inverted tree branches with internal nodes and leaf nodes. The algorithm can be effectively used for training when building a decision tree model and validation data set in order to decide on the appropriate tree size required to achieve the optimal final model. The following separation functions are used in the Decision Tree method:

- Gdi (*G*) – an algorithm that uses Gini data heterogeneity index;
- Twoing (*T*) – split algorithm generating a balanced decision tree slower than Gdi;
- Deviance (*D*) – algorithm for finding deviation between data.

Discriminant analysis (*DA*) is a multidimensional data analysis that is used when it is necessary to predict the values of a grouping variable. This is known as classification or pattern recognition. The discriminant analysis uses the following separation functions:

- Linear (*L*) - linear separator function, suitable for multivariate normal density data of each group, with an overall covariance estimate;
- Diagonal (*DL*) - is similar to the linear separating function but uses the calculation of the diagonal of a covariance matrix (diagonal linear separating function);
- Quadratic (*Q*) - quadratic separating function (second degree), distributes data with multivariate normal density by calculating covariance and grouping them;
- Diagonal quadratic (*DQ*) - is similar to the quadratic separating function but uses the calculation of the diagonal of a covariance matrix (diagonal nonlinear separating function);
- Mahalanobis (*M*) - splits data into groups by Mahalanobis distance by determining covariance in the data.

The Support Vector Machines Method (*SVM*) is a model of teacher training and related data analysis algorithms used for classification. Each element of the training sample is associated with one of two categories, and the training algorithm builds a model in which the data is transformed into a new space so that there is a separation between them. SVM analysis uses the following separation functions:

- Linear (*L*) - linear separator function, suitable for multivariate normal density data of each group;
- Quadratic (*Q*) - quadratic separating function (second degree), distributes data with multivariate normal density by calculating covariance and grouping them;
- Polynomial (*P*) - polynomial separation function;
- *RBF* - separating function defined by radial basis elements.

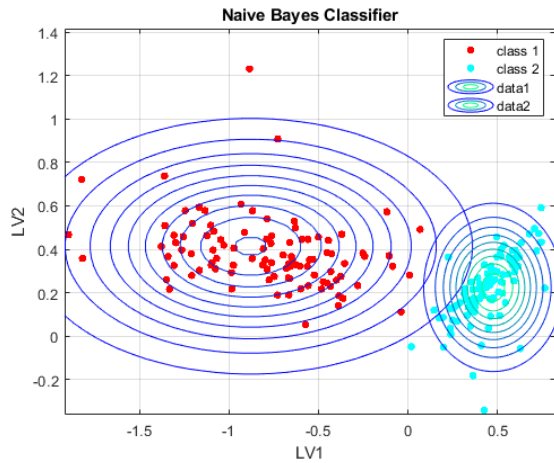
Figure 1 presents examples of classification with the indicated classifiers.

The performance of the classifiers used is estimated by a total classification error [8], which is described by the formula:

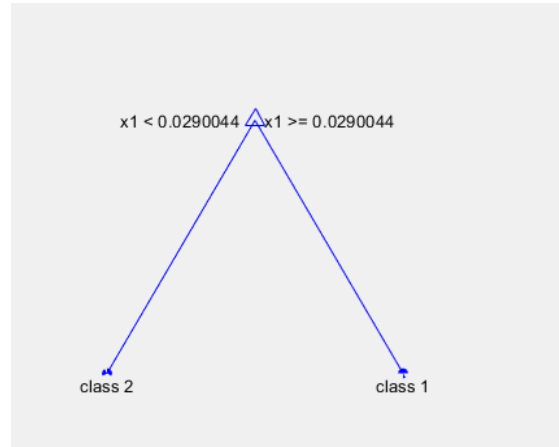
$$e = \frac{\sum_{i=1}^n (\sum_{k=1}^n y_{ik} - y_{ii})}{\sum_{i=1}^n \sum_{k=1}^n y_{ik}} \cdot 100, \% \quad (1)$$

where  $y_{ik}$  is the number of samples in class  $i$  classified by the classifier in class  $k$ ;  $y_{ii}$  number of correctly identified samples;  $k=1\dots n$  - number incorrectly assigned to a class  $i$  relative to the total number of samples;  $n$  - number of classes.

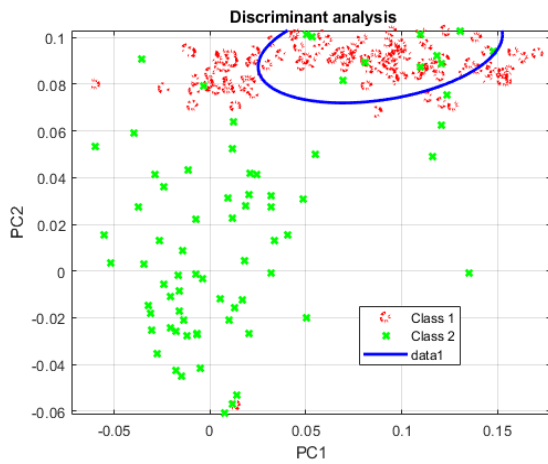
All data are processed at a level of significance  $\alpha=0,05$ .



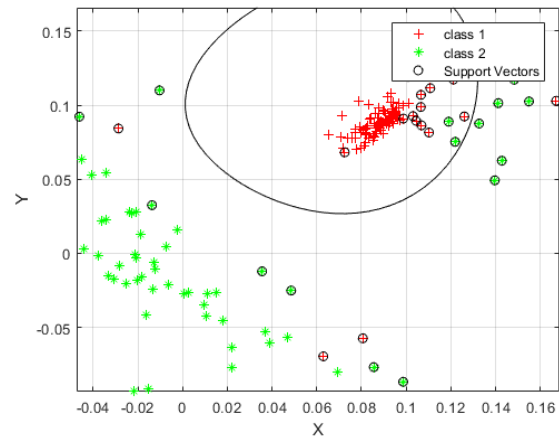
a) Naïve Bayes (NBC)



b) Decision Tree (DT)



c) Discriminant Analysis (DA)



d) Support Vector Machines (SVM)

Figure 1. Examples of the classifiers used

### 3. RESULTS AND DISCUSSION

The results of a classification with four classifiers are presented. A summary analysis of the classification errors obtained is made, depending on the feature vector used and the method for reducing the volume of data. The results obtained are summarized and compared with those in the available literature.

#### 3.1. Classification by Naïve Bayes Classifier (NBC)

Table 5 shows the results of the classification of hen eggs with a Naïve Bayes classifier by feature vectors containing shape indices and spectral indices. The table shows that the feature vectors HSFV1, HSFV2 and HSFV4 have the lowest total classification error, compared to the other feature vectors,  $e=9-11\%$ . Compared to the kernel variant of the principal components used to reduce the amount of data, low values of the total classification error were obtained using latent variables. Significantly larger are the values of this error when using a linear variant of the principal components. Highest error rates were obtained between 1-7 days and 7-14 days.

Table 5. A total classification error  $e_m, \%$  for hen eggs by feature vectors containing shape indices and spectral indices

FV	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
	DRM									
HSFV1	LV		4%	8%	7%	4%	5%	8%	<b>6%</b>	<b>9%</b>
	PC		19%	15%	17%	14%	15%	23%	<b>17%</b>	
	kPC		10%	2%	3%	1%	2%	3%	<b>3%</b>	
HSFV2	LV		4%	3%	4%	3%	5%	4%	<b>4%</b>	<b>11%</b>
	PC		28%	23%	26%	23%	26%	26%	<b>25%</b>	
	kPC		2%	2%	2%	2%	2%	9%	<b>3%</b>	
HSFV3	LV		19%	20%	20%	20%	21%	28%	<b>21%</b>	<b>23%</b>
	PC		52%	53%	8%	52%	8%	7%	<b>30%</b>	
	kPC		10%	19%	21%	13%	17%	19%	<b>17%</b>	
HSFV4	LV		5%	10%	8%	7%	6%	10%	<b>8%</b>	<b>10%</b>
	PC		16%	18%	19%	16%	16%	17%	<b>17%</b>	
	kPC		8%	2%	2%	1%	2%	15%	<b>5%</b>	
HSFV5	LV		15%	16%	10%	20%	9%	12%	<b>14%</b>	<b>13%</b>
	PC		15%	15%	3%	49%	28%	30%	<b>23%</b>	
	kPC		9%	1%	2%	1%	4%	5%	<b>4%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day

Table 6 shows the results of the classification of hen eggs with a Naïve Bayses classifier by shape indices and dielectric characteristics. The feature vectors HCFV1, HCFV4 and HCFV5 have the lowest total classification error, compared to the other feature vectors,  $e=4-8\%$ . These errors are smaller compared to the use of feature vectors containing spectral indices and shape indices. Compared to the kernel variant of the principal components used to reduce the amount of data, low values of the total classification error were obtained using latent variables. Significantly larger are the values of this error when using a linear variant of the principal components. The highest error values were obtained between 1-21 days.

Table 6. A total classification error  $e_m, \%$  for hen eggs by feature vectors containing shape indices and dielectric characteristics

FV	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
	DRM									
HCFV1	LV		6%	9%	21%	1%	2%	2%	<b>7%</b>	<b>8%</b>
	PC		9%	22%	23%	11%	3%	17%	<b>14%</b>	
	kPC		14%	2%	3%	2%	4%	2%	<b>4%</b>	
HCFV2	LV		8%	11%	23%	3%	4%	4%	<b>9%</b>	<b>10%</b>
	PC		11%	24%	25%	13%	5%	19%	<b>16%</b>	
	kPC		16%	4%	5%	4%	6%	4%	<b>6%</b>	
HCFV3	LV		6%	17%	10%	6%	6%	15%	<b>10%</b>	<b>9%</b>
	PC		8%	10%	19%	20%	8%	6%	<b>12%</b>	
	kPC		6%	7%	6%	6%	6%	6%	<b>6%</b>	
HCFV4	LV		1%	12%	5%	1%	1%	10%	<b>5%</b>	<b>4%</b>
	PC		3%	5%	14%	15%	3%	1%	<b>7%</b>	
	kPC		1%	2%	1%	1%	1%	1%	<b>1%</b>	
HCFV5	LV		1%	5%	3%	4%	1%	21%	<b>6%</b>	<b>6%</b>
	PC		9%	17%	13%	5%	4%	12%	<b>10%</b>	
	kPC		0%	1%	1%	0%	2%	1%	<b>1%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day

Table 7 shows the results of the classification by NBC of quail eggs with feature vectors (QSFV) containing shape indices and spectral indices. The feature vectors QSFV1, QSFV2 and QSFV4 have the smallest total classification error, compared to the other feature vectors  $e=8-10\%$ . The QSFV2 feature vector was chosen over QSFV3 because, when using latent variables, it shows a lower value of the total classification error. Compared to the kernel

variant of the principal components used to reduce the amount of data, low values of the total classification error were obtained using latent variables. Significantly larger are the values of this error when using a linear variant of the principal components. Highest error rates were obtained between 1-7, 1-14, and 7-21 days.

Table 7. A total classification error  $e_m, \%$  for quail eggs by feature vectors containing shape indices and spectral indices

FV	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
	DRM									
QSFV1	LV		18%	5%	6%	5%	5%	8%	<b>8%</b>	<b>10%</b>
	PC		13%	12%	15%	11%	12%	14%	<b>13%</b>	
	kPC		10%	2%	6%	3%	29%	1%	<b>8%</b>	
QSFV2	LV		48%	7%	1%	8%	1%	1%	<b>11%</b>	<b>8%</b>
	PC		6%	6%	2%	5%	2%	5%	<b>4%</b>	
	kPC		9%	12%	1%	21%	1%	1%	<b>8%</b>	
QSFV3	LV		37%	23%	1%	20%	1%	1%	<b>14%</b>	<b>8%</b>
	PC		17%	13%	4%	10%	4%	3%	<b>9%</b>	
	kPC		10%	3%	1%	3%	1%	0%	<b>3%</b>	
QSFV4	LV		8%	9%	1%	9%	1%	0%	<b>5%</b>	<b>10%</b>
	PC		44%	41%	8%	41%	7%	6%	<b>24%</b>	
	kPC		1%	1%	0%	1%	0%	0%	<b>1%</b>	
QSFV5	LV		17%	14%	6%	13%	6%	4%	<b>10%</b>	<b>15%</b>
	PC		51%	45%	21%	44%	20%	20%	<b>34%</b>	
	kPC		4%	1%	2%	2%	2%	1%	<b>2%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day

Table 8 shows the results of the classification of quail eggs with a Naïve Bayes classifier (NBC) by shape indices and dielectric characteristics. The feature vectors QCFV1, QCFV4 and QCFV5 have the lowest total classification error, compared to the other feature vectors,  $e=7-9\%$ . These errors are smaller compared to the use of feature vectors containing spectral indices and shape indices. Compared to the kernel variant of the principal components used to reduce the amount of data, low values of the total classification error were obtained using latent variables. Significantly larger are the values of this error when using a linear variant of the principal components. The highest error values were obtained between 1-14 and 1-21 days.

Table 8. A total classification error  $e_m, \%$  for quail eggs by feature vectors containing shape indices and dielectric characteristics

FV	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
	DRM									
QCFV1	LV		6%	16%	16%	4%	6%	15%	<b>10%</b>	<b>8%</b>
	PC		6%	19%	19%	17%	10%	18%	<b>15%</b>	
	kPC		0%	0%	0%	0%	0%	0%	<b>0%</b>	
QCFV2	LV		8%	18%	18%	6%	8%	17%	<b>12%</b>	<b>10%</b>
	PC		8%	21%	21%	19%	12%	20%	<b>17%</b>	
	kPC		2%	2%	2%	2%	2%	2%	<b>2%</b>	
QCFV3	LV		6%	12%	3%	7%	3%	3%	<b>6%</b>	<b>10%</b>
	PC		22%	22%	23%	16%	15%	21%	<b>21%</b>	
	kPC		3%	4%	3%	4%	3%	3%	<b>4%</b>	
QCFV4	LV		3%	9%	0%	4%	0%	0%	<b>3%</b>	<b>7%</b>
	PC		22%	19%	22%	13%	12%	18%	<b>18%</b>	
	kPC		0%	1%	0%	1%	0%	0%	<b>1%</b>	
QCFV5	LV		8%	13%	15%	6%	7%	13%	<b>10%</b>	<b>9%</b>
	PC		13%	19%	19%	13%	13%	20%	<b>16%</b>	
	kPC		0%	0%	0%	0%	0%	0%	<b>0%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day

Based on the classification with the Naïve Bayes Classifier (NBC), it was found that suitable for distinguishing hens eggs by their weight, depending on the day of storage, HSFV1, HSFV2 and HSFV4 feature vectors derived from spectral characteristics and shape indices are suitable. By these indices can be used to classify quail eggs, feature vectors QSFV1, QSFV2 and QSFV4.

Distinguishing hen eggs by their weight, depending on the day of storage, is possible by the feature vectors HCFV1, HCFV4 and HCFV5 derived from the shape indices and dielectric characteristics. According to these indices for the classification of quail eggs, feature vectors QSFV1, QSFV4 and QSFV5 can be used.

Appropriate methods for reducing the amount of data of feature vectors are principal components (PC) and latent variables (LV), because their use results in low values of the total classification error.

The results show that the CORR, SFCPP and RELIEFF methods are suitable for selection of features from shape indices and spectral indices of eggs according to the results obtained with the Naïve Basyes classifier (NBC). The ranking of features by shape indices and dielectric characteristics of eggs is successful by the CORR, RELIEFF and FSRNCA methods.

The selected feature vectors were used to classify eggs by the Decision Tree (DT) method, Discriminant analysis (DA), and Support Vector Machines (SVM) methods.

### 3.2. Classification by Decision Tree (DT)

Table 9 shows the results of the classification of hen eggs by the Decision Tree (DT) method, by feature vectors containing shape indices and spectral indices. The table shows that the HSFV2 feature vector has the lowest total classification error, compared to the other feature vectors,  $e=9\%$  using latent variables (LV). Significantly larger are the values of this error when using the linear variant of the principal components (PC),  $e=28\%$ . The highest error values for this feature vector were obtained between 14-21 days.

Table 9. A total classification error  $e, \%$  for hen eggs by feature vectors containing shape indices and spectral indices

FV	DRM	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
		SF									
HSFV1	LV	G		5%	8%	7%	5%	5%	8%	6%	6%
		T		5%	8%	7%	7%	5%	8%	7%	
		D		5%	8%	7%	5%	5%	9%	6%	
	PC	G		20%	25%	26%	20%	20%	28%	23%	24%
		T		20%	24%	28%	21%	20%	26%	23%	
		D		23%	24%	27%	24%	23%	28%	25%	
HSFV2	LV	G		4%	5%	5%	5%	5%	6%	5%	5%
		T		4%	4%	4%	5%	6%	6%	5%	
		D		5%	4%	4%	5%	6%	7%	5%	
	PC	G		30%	28%	29%	24%	30%	27%	28%	28%
		T		30%	31%	29%	24%	30%	24%	28%	
		D		29%	31%	29%	30%	30%	29%	30%	
HSFV4	LV	G		5%	14%	15%	5%	5%	10%	9%	9%
		T		5%	10%	12%	5%	5%	12%	8%	
		D		6%	12%	12%	5%	5%	12%	9%	
	PC	G		20%	22%	18%	19%	18%	18%	19%	19%
		T		19%	20%	18%	19%	18%	18%	19%	
		D		19%	20%	22%	20%	22%	20%	20%	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; G-Gdi; T-Twoing; D-Deviance



Table 10 shows the results of the classification of hen eggs by the DT method, by feature vectors containing shape indices and dielectric characteristics. The HCFV5 feature vector has the lowest total classification error, compared to the other feature vectors,  $e=3\%$  using LV. Significantly larger are the values of this error when using the PC,  $e=15\%$ . The highest error values for this feature vector were obtained between 1-14 days.

Table 10.

A total classification error  $e, \%$  for hen eggs by feature vectors containing shape indices and dielectric characteristics

FV	DRM	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
		SF									
HCFV1	LV	G		3%	25%	15%	3%	3%	15%	<b>11%</b>	<b>11%</b>
		T		3%	25%	15%	3%	3%	15%	<b>10%</b>	
		D		3%	25%	15%	3%	3%	15%	<b>10%</b>	
	PC	G		14%	23%	25%	14%	14%	24%	<b>19%</b>	<b>19%</b>
		T		14%	23%	24%	14%	14%	23%	<b>19%</b>	
		D		14%	23%	25%	14%	14%	25%	<b>19%</b>	
HCFV4	LV	G		1%	7%	7%	1%	1%	7%	<b>4%</b>	<b>4%</b>
		T		1%	7%	8%	1%	1%	8%	<b>4%</b>	
		D		1%	7%	7%	1%	1%	6%	<b>4%</b>	
	PC	G		12%	26%	18%	12%	12%	18%	<b>16%</b>	<b>17%</b>
		T		12%	28%	18%	12%	12%	18%	<b>17%</b>	
		D		12%	29%	18%	12%	12%	18%	<b>17%</b>	
HCFV5	LV	G		2%	8%	2%	2%	2%	2%	<b>3%</b>	<b>3%</b>
		T		2%	8%	2%	2%	2%	2%	<b>3%</b>	
		D		2%	8%	2%	2%	2%	2%	<b>3%</b>	
	PC	G		13%	22%	14%	13%	13%	14%	<b>15%</b>	<b>15%</b>
		T		13%	22%	14%	13%	13%	14%	<b>15%</b>	
		D		13%	22%	14%	13%	13%	14%	<b>15%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; G-Gdi; T-Twoing; D-Deviance

Table 11 shows the results of the classification of quail eggs with DT, by feature vectors containing shape indices and spectral indices. The QSFV4 feature vector has the lowest total classification error, compared to the other feature vectors,  $e=9\%$  using LV. For the QSFV2 feature vector, lower error values were obtained using a PC. The highest error values for these feature vectors were obtained between 1-7 and 7-14 days.

Table 12 shows the results of the classification of quail eggs with DT, by feature vectors containing shape indices and dielectric characteristics. The QCFV5 feature vector has the lowest total classification error, compared to the other feature vectors,  $e=8\%$  using LV and  $e=18\%$  using PC. The highest error values for this feature vector were obtained between 1-21 and 7-21 days.

Table 11.

A total classification error  $e, %$  for quail eggs by feature vectors containing shape indices and spectral indices

FV	DRM	SF, D	D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, %$	FV $e_m, %$
QSFV1	LV	G	19%	8%	5%	8%	4%	4%	<b>8%</b>	<b>8%</b>
		T	19%	8%	4%	8%	5%	5%	<b>8%</b>	
		D	18%	8%	5%	8%	5%	5%	<b>8%</b>	
	PC	G	12%	10%	19%	10%	13%	10%	<b>12%</b>	<b>12%</b>
		T	12%	10%	19%	10%	12%	10%	<b>12%</b>	
		D	12%	10%	19%	10%	12%	10%	<b>12%</b>	
QSFV2	LV	G	44%	7%	1%	7%	1%	1%	<b>10%</b>	<b>10%</b>
		T	44%	8%	1%	7%	1%	1%	<b>10%</b>	
		D	44%	7%	1%	7%	1%	1%	<b>10%</b>	
	PC	G	9%	7%	7%	7%	7%	10%	<b>8%</b>	<b>8%</b>
		T	9%	7%	7%	7%	7%	11%	<b>8%</b>	
		D	9%	7%	7%	7%	7%	7%	<b>8%</b>	
QSFV4	LV	G	11%	9%	4%	10%	4%	4%	<b>7%</b>	<b>7%</b>
		T	10%	9%	4%	9%	4%	4%	<b>7%</b>	
		D	10%	10%	4%	10%	4%	4%	<b>7%</b>	
	PC	G	48%	41%	13%	42%	14%	13%	<b>29%</b>	<b>29%</b>
		T	46%	42%	13%	41%	14%	13%	<b>28%</b>	
		D	47%	45%	15%	44%	14%	15%	<b>30%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; G-Gdi; T-Twoing; D-Deviance

Table 12.

A total classification error  $e, %$  for quail eggs by feature vectors containing shape indices and dielectric characteristics

FV	DRM	SF, D	D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, %$	FV $e_m, %$
QCFV1	LV	G	22%	22%	7%	22%	7%	7%	<b>15%</b>	<b>15%</b>
		T	23%	22%	7%	22%	7%	7%	<b>15%</b>	
		D	23%	23%	7%	26%	7%	7%	<b>16%</b>	
	PC	G	28%	31%	11%	30%	10%	10%	<b>20%</b>	<b>20%</b>
		T	31%	31%	10%	29%	11%	11%	<b>21%</b>	
		D	28%	31%	10%	28%	11%	11%	<b>20%</b>	
QCFV4	LV	G	30%	19%	7%	19%	7%	8%	<b>15%</b>	<b>15%</b>
		T	29%	19%	8%	19%	7%	7%	<b>15%</b>	
		D	30%	19%	7%	19%	7%	7%	<b>15%</b>	
	PC	G	16%	9%	13%	10%	16%	9%	<b>12%</b>	<b>12%</b>
		T	16%	9%	17%	9%	16%	9%	<b>12%</b>	
		D	17%	9%	16%	9%	17%	9%	<b>13%</b>	
QCFV5	LV	G	9%	7%	9%	7%	11%	7%	<b>9%</b>	<b>8%</b>
		T	9%	7%	10%	7%	10%	7%	<b>9%</b>	
		D	9%	7%	9%	7%	9%	7%	<b>8%</b>	
	PC	G	19%	18%	19%	17%	15%	19%	<b>18%</b>	<b>18%</b>
		T	15%	19%	19%	17%	17%	19%	<b>18%</b>	
		D	21%	20%	20%	20%	21%	19%	<b>20%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; G-Gdi; T-Twoing; D-Deviance

### 3.3. Classification by Discriminant analysis (DA)

Table 13 shows the results of the classification of hen eggs with a DA Classifier by vectors of features containing shape indices and spectral indices. The HSFV2 feature vector has the smallest total classification error, compared to the other feature vectors,  $e=4\%$ , using LV. Using a PC,  $e=22\%$ . Error values keep their levels throughout the measurement period. It can be seen that the highest values of the total classification error are obtained using linear separating functions, regardless of the method used to reduce the amount of data of feature vectors.

Table 13.

A total classification error  $e, \%$  for hen eggs by feature vectors containing shape indices and spectral indices

FV	DRM	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
		SF	L								
HSFV1	LV	L		5%	8%	8%	5%	5%	8%	<b>6%</b>	<b>6%</b>
		DL		5%	8%	8%	4%	4%	8%	<b>6%</b>	
		Q		4%	7%	6%	4%	4%	8%	<b>6%</b>	
		DQ		4%	8%	7%	3%	5%	8%	<b>6%</b>	
		M		4%	6%	5%	4%	4%	8%	<b>5%</b>	
	PC	L		20%	22%	20%	20%	18%	25%	<b>21%</b>	<b>18%</b>
		DL		23%	21%	20%	20%	18%	26%	<b>22%</b>	
		Q		15%	8%	14%	7%	10%	9%	<b>10%</b>	
		DQ		19%	15%	17%	14%	15%	22%	<b>17%</b>	
		M		17%	25%	22%	21%	14%	28%	<b>21%</b>	
HSFV2	LV	L		3%	4%	3%	4%	5%	5%	<b>4%</b>	<b>4%</b>
		DL		4%	4%	5%	5%	5%	5%	<b>4%</b>	
		Q		3%	4%	5%	4%	5%	5%	<b>4%</b>	
		DQ		4%	3%	4%	3%	4%	4%	<b>4%</b>	
		M		3%	3%	5%	3%	4%	5%	<b>4%</b>	
	PC	L		27%	25%	25%	25%	28%	25%	<b>26%</b>	<b>22%</b>
		DL		28%	25%	26%	25%	28%	27%	<b>26%</b>	
		Q		24%	12%	19%	8%	17%	12%	<b>15%</b>	
		DQ		28%	23%	25%	23%	25%	26%	<b>25%</b>	
		M		30%	24%	25%	29%	26%	28%	<b>27%</b>	
HSFV4	LV	L		3%	10%	8%	3%	4%	9%	<b>6%</b>	<b>6%</b>
		DL		6%	9%	8%	6%	5%	10%	<b>8%</b>	
		Q		3%	8%	6%	3%	3%	10%	<b>5%</b>	
		DQ		5%	10%	8%	6%	6%	10%	<b>8%</b>	
		M		3%	7%	6%	3%	3%	9%	<b>5%</b>	
	PC	L		21%	19%	17%	20%	18%	17%	<b>19%</b>	<b>17%</b>
		DL		20%	20%	19%	20%	18%	19%	<b>19%</b>	
		Q		13%	10%	11%	13%	13%	18%	<b>13%</b>	
		DQ		16%	18%	18%	16%	16%	17%	<b>17%</b>	
		M		19%	8%	11%	13%	13%	17%	<b>13%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; L-Linear; DL-DiagLinear; Q-Quadratic; DQ-DiagQuadratic; M-Mahalanobis

Table 14 shows the results of the classification of hen eggs with a Discriminant (DA) Classifier by feature vectors containing shape indices and dielectric characteristics. It can be seen that the feature vector HCFV5 has the lowest total classification error, compared to the other feature vectors,  $e= 2\%$ , using latent variables (LV). Using a linear variant of the principal components (PC),  $e=13\%$ . The error values for this feature vector are highest at

days 1-14 and 1-21. It can be seen that the values of the total classification error are maintained throughout the measurement period, regardless of the separating function used and the method for reducing the amount of data for the presented feature vector.

Table 14.

A total classification error  $e_m$ , % for hen eggs by feature vectors containing shape indices and dielectric characteristics

FV	DRM	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m$ , %	FV $e_m$ , %
		SF									
HCFV1	LV	L		6%	11%	17%	7%	3%	1%	<b>7%</b>	<b>8%</b>
		DL		6%	19%	24%	8%	5%	9%	<b>12%</b>	
		Q		1%	8%	18%	1%	1%	1%	<b>5%</b>	
		DQ		6%	9%	20%	1%	2%	2%	<b>6%</b>	
		M		1%	15%	19%	2%	1%	1%	<b>6%</b>	
	PC	L		18%	23%	24%	21%	3%	20%	<b>18%</b>	<b>16%</b>
		DL		14%	23%	21%	21%	4%	17%	<b>17%</b>	
		Q		8%	19%	20%	10%	3%	14%	<b>12%</b>	
		DQ		9%	22%	23%	10%	4%	17%	<b>14%</b>	
		M		10%	21%	20%	13%	3%	13%	<b>14%</b>	
HCFV4	LV	L		1%	8%	5%	2%	1%	5%	<b>4%</b>	<b>4%</b>
		DL		2%	10%	5%	1%	1%	8%	<b>4%</b>	
		Q		1%	8%	6%	1%	1%	5%	<b>4%</b>	
		DQ		1%	11%	5%	1%	0%	9%	<b>5%</b>	
		M		1%	7%	5%	1%	1%	5%	<b>3%</b>	
	PC	L		2%	4%	20%	17%	8%	1%	<b>9%</b>	<b>8%</b>
		DL		2%	4%	20%	17%	8%	1%	<b>8%</b>	
		Q		1%	2%	13%	13%	3%	1%	<b>6%</b>	
		DQ		3%	5%	14%	15%	3%	1%	<b>7%</b>	
		M		1%	12%	21%	23%	2%	1%	<b>10%</b>	
HCFV5	LV	L		1%	5%	2%	1%	1%	2%	<b>2%</b>	<b>2%</b>
		DL		2%	5%	2%	0%	1%	2%	<b>2%</b>	
		Q		1%	5%	3%	0%	1%	2%	<b>2%</b>	
		DQ		2%	5%	2%	0%	1%	2%	<b>2%</b>	
		M		1%	6%	3%	0%	1%	2%	<b>2%</b>	
	PC	L		14%	20%	16%	13%	9%	16%	<b>15%</b>	<b>13%</b>
		DL		14%	20%	16%	13%	9%	15%	<b>14%</b>	
		Q		7%	16%	14%	5%	4%	13%	<b>10%</b>	
		DQ		9%	16%	14%	5%	4%	13%	<b>10%</b>	
		M		18%	22%	27%	13%	6%	16%	<b>17%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; L-Linear; DL-DiagLinear; Q-Quadratic; DQ-DiagQuadratic; M-Mahalanobis

Table 15 shows the results of the classification of quail eggs with DA by feature vectors containing shape indices and spectral indices. The QSFV4 feature vector has the smallest total classification error, compared to the other feature vectors,  $e=5\%$  using LV. For the QSFV2 feature vector, lower error values were obtained using PC,  $e=8\%$ . The highest error values for these feature vectors were obtained between 1-7, 1-14 and 7-14 days.

Table 15.

A total classification error  $e_m$ , % for quail eggs by feature vectors containing shape indices and spectral indices

FV	DRM	D SF	D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m$ , %	FV $e_m$ , %
QSFV1	LV	L	18%	5%	6%	6%	5%	7%	8%	8%
		DL	19%	5%	6%	5%	5%	7%	8%	
		Q	20%	4%	6%	4%	5%	7%	8%	
		DQ	18%	4%	6%	5%	5%	7%	7%	
		M	19%	4%	6%	4%	5%	7%	8%	
	PC	L	12%	11%	12%	12%	11%	10%	11%	11%
		DL	13%	13%	11%	12%	10%	11%	11%	
		Q	13%	12%	12%	10%	11%	9%	11%	
		DQ	13%	12%	13%	11%	11%	11%	12%	
		M	14%	14%	15%	14%	14%	16%	15%	
QSFV2	LV	L	38%	6%	1%	6%	6%	1%	10%	10%
		DL	56%	6%	1%	6%	6%	1%	13%	
		Q	33%	2%	1%	2%	2%	1%	7%	
		DQ	48%	6%	1%	8%	8%	1%	12%	
		M	32%	1%	1%	2%	2%	1%	7%	
	PC	L	12%	7%	12%	7%	13%	6%	9%	8%
		DL	12%	7%	13%	7%	13%	6%	10%	
		Q	7%	6%	2%	3%	2%	5%	4%	
		DQ	7%	6%	2%	5%	2%	5%	4%	
		M	11%	9%	2%	7%	2%	6%	6%	
QSFV4	LV	L	10%	8%	2%	8%	2%	1%	5%	5%
		DL	8%	8%	2%	8%	2%	1%	5%	
		Q	9%	9%	1%	10%	1%	0%	5%	
		DQ	9%	9%	1%	9%	1%	0%	5%	
		M	12%	10%	1%	11%	0%	0%	6%	
	PC	L	48%	44%	12%	42%	11%	11%	28%	26%
		DL	45%	47%	13%	46%	12%	9%	29%	
		Q	37%	41%	6%	41%	6%	4%	23%	
		DQ	44%	41%	7%	41%	6%	6%	24%	
		M	46%	50%	18%	51%	17%	12%	32%	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; L-Linear; DL-DiagLinear; Q-Quadratic; DQ-DiagQuadratic; M-Mahalanobis

Table 16 shows the results of the classification of quail eggs with a Discriminant (DA) Classifier by feature vectors containing shape indices and dielectric characteristics. It can be seen that the feature vector QCFV1 has the lowest total classification error, compared to the other feature vectors,  $e=10\%$  using latent variables (LV). Using a linear variant of the principal components (PC),  $e=13\%$ . The error values for this feature vector are highest at days 1-14 and 14-21. It can be seen that the values of the total classification error are maintained throughout the measurement period, regardless of the separating function used and the method for reducing amount of data for the presented feature vector.

Table 16.

A total classification error  $e_m, \%$  for quail eggs by feature vectors containing shape indices and dielectric characteristics

FV	DRM	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
		SF									
QCFV1	LV	L		6%	11%	16%	5%	6%	12%	9%	10%
		DL		6%	17%	17%	6%	7%	15%	11%	
		Q		3%	13%	18%	2%	3%	11%	8%	
		DQ		5%	15%	16%	4%	5%	15%	10%	
		M		4%	14%	17%	4%	6%	13%	10%	
	PC	L		7%	18%	18%	9%	9%	19%	13%	13%
		DL		8%	18%	19%	9%	9%	17%	13%	
		Q		9%	13%	18%	11%	10%	17%	13%	
		DQ		6%	19%	19%	12%	9%	18%	14%	
		M		9%	17%	18%	12%	10%	20%	14%	
QCFV4	LV	L		8%	10%	17%	7%	9%	10%	10%	10%
		DL		10%	17%	19%	9%	11%	16%	14%	
		Q		5%	9%	17%	3%	5%	7%	8%	
		DQ		6%	17%	19%	5%	6%	16%	11%	
		M		7%	11%	19%	5%	7%	11%	10%	
	PC	L		25%	20%	20%	9%	9%	19%	17%	15%
		DL		24%	19%	20%	9%	9%	19%	17%	
		Q		20%	4%	6%	10%	10%	18%	11%	
		DQ		22%	19%	21%	11%	12%	19%	17%	
		M		24%	6%	6%	10%	10%	19%	12%	
QCFV5	LV	L		10%	9%	15%	7%	8%	11%	10%	11%
		DL		13%	14%	16%	11%	13%	13%	13%	
		Q		6%	9%	16%	5%	6%	8%	8%	
		DQ		7%	12%	15%	6%	7%	12%	10%	
		M		9%	12%	16%	6%	8%	11%	10%	
	PC	L		12%	18%	18%	12%	11%	18%	15%	14%
		DL		12%	18%	17%	12%	11%	18%	15%	
		Q		11%	16%	18%	10%	11%	19%	14%	
		DQ		13%	19%	18%	12%	13%	19%	16%	
		M		11%	18%	19%	10%	11%	18%	14%	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; L-Linear; DL-DiagLinear; Q-Quadratic; DQ-DiagQuadratic; M-Mahalanobis

### 3.4. Classification with Support Vector Machines (SVM)

Table 17 shows the results of the classification of hen eggs by Support vector machines (SVM) method by feature vectors containing shape indices and spectral indices. The HSFV4 feature vector has the smallest total classification error, compared to the other feature vectors,  $e=1\%$ , using latent variables (LV). Using a linear version of the principal components (PC),  $e=5\%$ . Error values keep their levels throughout the measurement period. It can be seen that the highest values of the total classification error are obtained using linear separating functions, regardless of the method used to reduce the volume of data of feature vectors.

Table 17.

A total classification error  $e_m, \%$  for hen eggs by feature vectors containing shape indices and spectral indices

FV	DRM	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
		SF	L								
HSFV1	LV	L	2%	3%	3%	1%	1%	3%	<b>2%</b>	<b>1%</b>	
		Q	0%	0%	0%	0%	0%	0%	<b>0%</b>		
		P	1%	0%	0%	0%	0%	0%	<b>0%</b>		
		RBF	0%	1%	1%	0%	1%	1%	<b>1%</b>		
	PC	L	19%	17%	15%	19%	14%	19%	<b>17%</b>	<b>6%</b>	
		Q	2%	8%	3%	13%	4%	9%	<b>6%</b>		
		P	0%	0%	0%	1%	0%	0%	<b>0%</b>		
		RBF	1%	0%	0%	1%	1%	2%	<b>1%</b>		
HSFV2	LV	L	1%	1%	2%	1%	3%	2%	<b>2%</b>	<b>1%</b>	
		Q	0%	0%	1%	0%	1%	0%	<b>0%</b>		
		P	0%	0%	0%	1%	0%	0%	<b>0%</b>		
		RBF	1%	0%	1%	0%	1%	0%	<b>1%</b>		
	PC	L	24%	19%	22%	19%	24%	19%	<b>21%</b>	<b>8%</b>	
		Q	0%	6%	4%	4%	4%	4%	<b>4%</b>		
		P	2%	3%	1%	5%	3%	3%	<b>3%</b>		
		RBF	2%	5%	1%	5%	3%	7%	<b>4%</b>		
HSFV4	LV	L	2%	7%	5%	2%	2%	5%	<b>4%</b>	<b>1%</b>	
		Q	0%	0%	0%	0%	0%	2%	<b>0%</b>		
		P	0%	0%	2%	0%	1%	1%	<b>1%</b>		
		RBF	1%	1%	0%	0%	2%	1%	<b>1%</b>		
	PC	L	19%	15%	15%	19%	15%	15%	<b>17%</b>	<b>5%</b>	
		Q	1%	0%	1%	2%	1%	0%	<b>1%</b>		
		P	3%	1%	2%	1%	2%	2%	<b>2%</b>		
		RBF	1%	0%	0%	1%	1%	2%	<b>1%</b>		

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; L-Linear; Q-Quadratic; P-Polynomial; RBF-Radial Basis Function

Table 18 shows the results of the classification of hen eggs with SVM by feature vectors containing shape indices and dielectric characteristics. The HCFV4 feature vector has the smallest total classification error, compared to the other feature vectors,  $e=1\%$  using LV. Using PC,  $e=3\%$ . Error values keep their levels throughout the measurement period.

Table 19 shows the results of the classification of quail eggs with SVM by feature vectors containing shape indices and spectral indices. The QSFV1 feature vector has the lowest total classification error, compared to the other feature vectors,  $e=1\%$ , using LV. Using PC,  $e=3\%$ . It can be seen that the highest values of the total classification error are obtained using linear separating functions.

Table 20 shows the results of the classification of quail eggs by the method Support vector machines (SVM) by feature vectors containing shape indices and dielectric characteristics. The QCFV1 feature vector has the smallest total classification error, compared to the other feature vectors,  $e=2\%$ , using latent variables (LV). Using a linear variant of the principal components (PC),  $e=4\%$ . Error values keep their levels throughout the measurement period. It can be seen that the highest values of the total classification error are obtained using linear separating functions, regardless of the method used to reduce the amount of data of feature vectors.

Table 18. A total classification error  $e_m, \%$  for hen eggs by feature vectors containing shape indices and dielectric characteristics

FV	DRM	SF, D	D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV, $e_m, \%$
HCFV1	LV	L	1%	8%	1%	3%	1%	1%	<b>3%</b>	1%
		Q	0%	4%	0%	2%	0%	0%	<b>1%</b>	
		P	0%	1%	1%	0%	0%	0%	<b>1%</b>	
		RBF	1%	2%	0%	0%	0%	1%	<b>1%</b>	
	PC	L	15%	3%	9%	14%	1%	4%	<b>8%</b>	3%
		Q	6%	2%	6%	5%	1%	3%	<b>4%</b>	
		P	1%	2%	0%	0%	0%	1%	<b>1%</b>	
HCFV4	LV	RBF	0%	0%	1%	1%	0%	0%	<b>0%</b>	1%
		L	1%	3%	3%	1%	0%	4%	<b>2%</b>	
		Q	0%	0%	0%	0%	0%	0%	<b>0%</b>	
		P	0%	0%	0%	0%	0%	0%	<b>0%</b>	
	PC	RBF	0%	1%	1%	0%	0%	1%	<b>1%</b>	3%
		L	1%	2%	18%	17%	6%	0%	<b>7%</b>	
		Q	0%	0%	10%	2%	5%	0%	<b>3%</b>	
HCFV5	LV	P	0%	0%	1%	0%	0%	0%	<b>0%</b>	1%
		RBF	0%	0%	0%	1%	0%	1%	<b>1%</b>	
		L	1%	1%	1%	0%	0%	2%	<b>1%</b>	
		Q	0%	1%	0%	0%	0%	1%	<b>0%</b>	
	PC	P	0%	0%	0%	0%	0%	1%	<b>0%</b>	5%
		RBF	0%	0%	0%	0%	0%	0%	<b>0%</b>	
		L	13%	14%	17%	13%	8%	14%	<b>13%</b>	
QSFV1	LV	Q	0%	0%	0%	0%	0%	1%	<b>0%</b>	1%
		P	0%	0%	0%	0%	0%	1%	<b>0%</b>	
		RBF	0%	0%	0%	1%	0%	1%	<b>1%</b>	
		L	1%	1%	1%	0%	0%	2%	<b>1%</b>	
	PC	Q	0%	0%	0%	0%	0%	1%	<b>0%</b>	3%
		P	2%	0%	0%	0%	0%	0%	<b>1%</b>	
		RBF	0%	1%	1%	1%	0%	1%	<b>1%</b>	
QSFV2	LV	L	4%	4%	2%	4%	3%	3%	<b>3%</b>	5%
		Q	1%	0%	1%	0%	0%	1%	<b>1%</b>	
		P	1%	0%	0%	0%	0%	2%	<b>1%</b>	
		RBF	1%	1%	0%	1%	0%	1%	<b>1%</b>	
	PC	L	8%	8%	8%	8%	7%	9%	<b>8%</b>	2%
		Q	2%	0%	0%	0%	0%	2%	<b>1%</b>	
		P	2%	0%	0%	0%	0%	0%	<b>1%</b>	
QSFV4	LV	RBF	0%	1%	1%	1%	0%	1%	<b>1%</b>	1%
		L	42%	6%	0%	5%	0%	0%	<b>9%</b>	
		Q	39%	2%	0%	2%	0%	0%	<b>7%</b>	
		P	4%	1%	0%	0%	0%	0%	<b>1%</b>	
	PC	RBF	4%	2%	0%	1%	0%	0%	<b>1%</b>	6%
		L	2%	1%	11%	1%	7%	0%	<b>4%</b>	
		Q	0%	0%	6%	0%	6%	0%	<b>2%</b>	
QSFV4	LV	P	0%	0%	0%	0%	0%	0%	<b>0%</b>	1%
		RBF	0%	0%	1%	0%	0%	0%	<b>1%</b>	
		L	5%	5%	0%	5%	0%	1%	<b>3%</b>	
		Q	0%	0%	0%	0%	0%	0%	<b>0%</b>	
	PC	P	0%	1%	0%	1%	0%	0%	<b>1%</b>	6%
		RBF	1%	0%	1%	0%	0%	0%	<b>0%</b>	
		L	30%	33%	4%	34%	4%	4%	<b>18%</b>	
QSFV4	LV	Q	6%	2%	1%	2%	1%	2%	<b>3%</b>	6%
		P	4%	0%	0%	0%	0%	1%	<b>1%</b>	
		RBF	1%	1%	0%	1%	0%	0%	<b>1%</b>	
		L	5%	5%	0%	5%	0%	1%	<b>3%</b>	
	PC	Q	0%	0%	0%	0%	0%	0%	<b>0%</b>	1%
		P	0%	1%	0%	1%	0%	0%	<b>1%</b>	
		RBF	1%	0%	1%	0%	0%	0%	<b>0%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; L-Linear; Q-Quadratic; P-Polynomial; RBF-Radial Basis Function

Table 19. A total classification error  $e_m, \%$  for quail eggs by vectors of traits containing shape indices and spectral indices

FV	DRM	SF, D	D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV, $e_m, \%$
QSFV1	LV	L	4%	4%	2%	4%	3%	3%	<b>3%</b>	1%
		Q	1%	0%	1%	0%	0%	1%	<b>1%</b>	
		P	1%	0%	0%	0%	0%	2%	<b>1%</b>	
		RBF	1%	1%	0%	1%	0%	1%	<b>1%</b>	
	PC	L	8%	8%	8%	8%	7%	9%	<b>8%</b>	3%
		Q	2%	0%	0%	0%	0%	2%	<b>1%</b>	
		P	2%	0%	0%	0%	0%	0%	<b>1%</b>	
QSFV2	LV	RBF	0%	1%	1%	1%	0%	1%	<b>1%</b>	5%
		L	42%	6%	0%	5%	0%	0%	<b>9%</b>	
		Q	39%	2%	0%	2%	0%	0%	<b>7%</b>	
		P	4%	1%	0%	0%	0%	0%	<b>1%</b>	
	PC	RBF	4%	2%	0%	1%	0%	0%	<b>1%</b>	2%
		L	2%	1%	11%	1%	7%	0%	<b>4%</b>	
		Q	0%	0%	6%	0%	6%	0%	<b>2%</b>	
QSFV4	LV	P	0%	0%	0%	0%	0%	0%	<b>0%</b>	1%
		RBF	0%	0%	1%	0%	0%	0%	<b>1%</b>	
		L	5%	5%	0%	5%	0%	1%	<b>3%</b>	
		Q	0%	0%	0%	0%	0%	0%	<b>0%</b>	
	PC	P	0%	1%	0%	1%	0%	0%	<b>1%</b>	6%
		RBF	1%	0%	1%	0%	0%	0%	<b>0%</b>	
		L	30%	33%	4%	34%	4%	4%	<b>18%</b>	
QSFV4	LV	Q	6%	2%	1%	2%	1%	2%	<b>3%</b>	6%
		P	4%	0%	0%	0%	0%	1%	<b>1%</b>	
		RBF	1%	1%	0%	1%	0%	0%	<b>1%</b>	
		L	5%	5%	0%	5%	0%	1%	<b>3%</b>	
	PC	Q	0%	0%	0%	0%	0%	0%	<b>0%</b>	1%
		P	0%	1%	0%	1%	0%	0%	<b>1%</b>	
		RBF	1%	0%	1%	0%	0%	0%	<b>0%</b>	

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; L-Linear; Q-Quadratic; P-Polynomial; RBF-Radial Basis Function



Table 20. A total classification error  $e_m, \%$  for quail eggs by feature vectors containing shape indices and dielectric characteristics

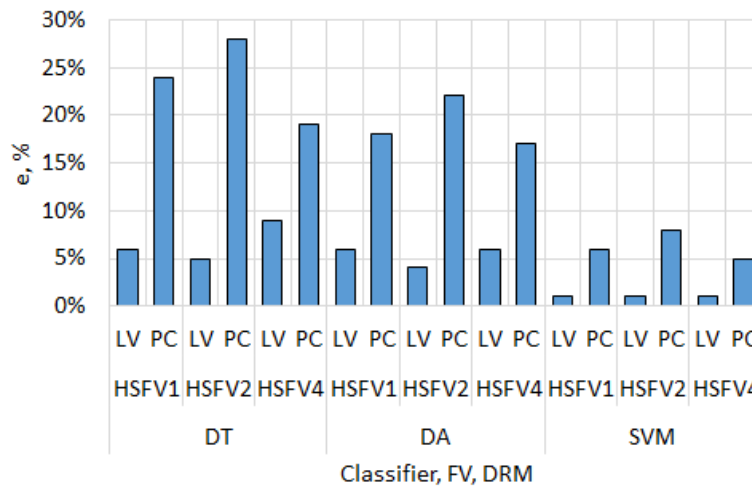
FV	DRM	D		D1-D7	D1-D14	D1-D21	D7-D14	D7-D21	D14-D21	$e_m, \%$	FV $e_m, \%$
		SF	L								
QCFV1	LV	L	4%	6%	7%	3%	4%	6%	<b>5%</b>	<b>2%</b>	
		Q	3%	2%	1%	1%	2%	2%	<b>2%</b>		
		P	0%	1%	0%	0%	0%	2%	<b>1%</b>		
		RBF	0%	1%	0%	1%	1%	3%	<b>1%</b>		
	PC	L	5%	13%	13%	7%	7%	13%	<b>10%</b>	<b>4%</b>	
		Q	2%	0%	2%	1%	1%	1%	<b>1%</b>		
		P	2%	2%	1%	2%	2%	0%	<b>2%</b>		
		RBF	1%	0%	1%	3%	2%	0%	<b>1%</b>		
QCFV4	LV	L	5%	4%	7%	6%	7%	7%	<b>6%</b>	<b>3%</b>	
		Q	4%	1%	1%	4%	6%	1%	<b>3%</b>		
		P	0%	1%	0%	0%	0%	2%	<b>1%</b>		
		RBF	1%	1%	2%	1%	1%	0%	<b>1%</b>		
	PC	L	5%	8%	8%	8%	8%	14%	<b>9%</b>	<b>4%</b>	
		Q	2%	2%	1%	0%	0%	0%	<b>1%</b>		
		P	1%	0%	1%	5%	5%	0%	<b>2%</b>		
		RBF	0%	2%	3%	4%	4%	2%	<b>3%</b>		
QCFV5	LV	L	7%	3%	7%	6%	6%	6%	<b>6%</b>	<b>3%</b>	
		Q	2%	3%	1%	2%	1%	1%	<b>2%</b>		
		P	0%	3%	0%	0%	2%	2%	<b>1%</b>		
		RBF	1%	2%	1%	1%	3%	0%	<b>2%</b>		
	PC	L	11%	16%	15%	11%	11%	15%	<b>13%</b>	<b>5%</b>	
		Q	3%	0%	2%	4%	5%	0%	<b>2%</b>		
		P	3%	1%	1%	2%	0%	1%	<b>1%</b>		
		RBF	4%	3%	0%	4%	2%	0%	<b>2%</b>		

FV-feature vector; DRM-method for reducing the amount of data; D-day; SF-separation function; L-Linear; Q-Quadratic; P-Polynomial; RBF-Radial Basis Function

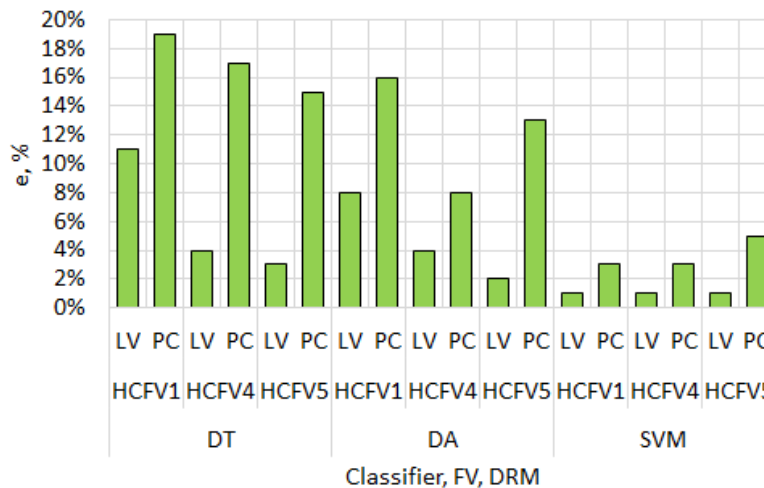
### 3.5. A summary analysis of the results obtained

The analysis of the results obtained was made in terms of the classifier used, the method for reducing the amount of data of selected feature vectors containing shape indices, spectral indices and dielectric characteristics of hens and quail eggs. A comparative analysis was made with the results presented in the available literature.

Figure 2 shows a summarized results obtained from the classification of hen eggs. When using feature vectors containing shape indices and spectral indices and reducing the amount of data by latent variables, the total classification error is in the range  $e=1-5\%$ , and when using principal components  $e=5-28\%$ . The highest values of the total classification error ( $e=6-28\%$ ) were obtained using the Decision Tree method, followed by the Discriminant Analysis method ( $e=6-22\%$ ). The lowest values of total classification error ( $e=1-8\%$ ) were obtained using the SVM Method. When using feature vectors containing shape indices and dielectric characteristics and reducing their amount of data through latent variables (LV), the total classification error is in the range  $e=1-11\%$ , and when using principal components (PC)  $e=11-19\%$ . The highest values of total classification error ( $e=3-19\%$ ) were obtained using the Decision Tree (DT) method, followed by the Discriminant Analysis (DA) method ( $e=2-16\%$ ). The lowest values of total classification error ( $e=1-5\%$ ) were obtained using the Support vector machines (SVM) Method. The results show that lower values of total classification error for classification of hen eggs are obtained using feature vectors containing shape indices and dielectric characteristics.



a) by shape indices and spectral indices

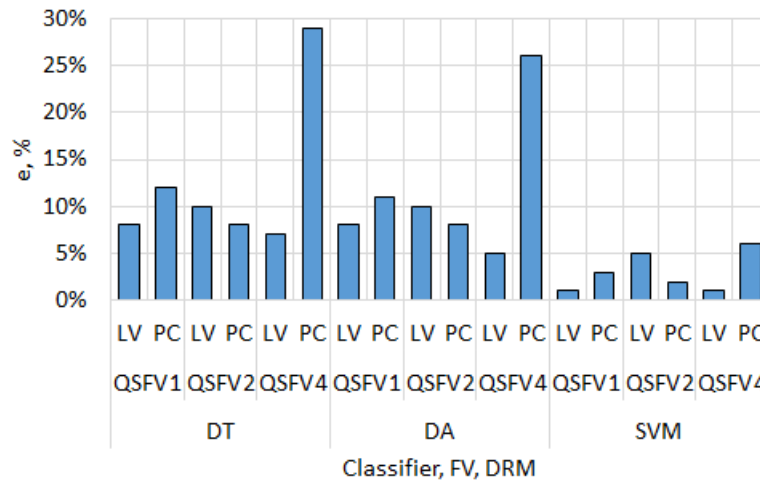


b) by shape indices and dielectric characteristics

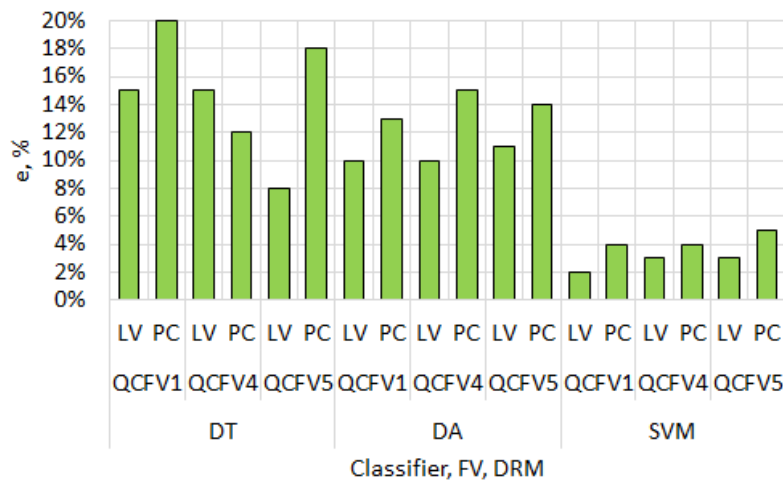
Figure 2. A summary analysis of the results for the classification of hen eggs FV-feature vector, DRM-amount of data reduction method

Figure 3 shows a summarized analysis of the results obtained from the classification of quail eggs. When using feature vectors containing shape indices and spectral indices and reducing their volume of data by latent variables (LV), the total classification error is in the range  $e=1-10\%$ , and when using principal components (PC),  $e=2-29\%$ . The highest values of the total classification error ( $e=8-29\%$ ) were obtained using the Decision Tree (DT) method, followed by the Discriminant Analysis (DA) method ( $e = 5-26\%$ ). The lowest values of total classification error ( $e=1-6\%$ ) were obtained using the SVM Method. When using feature vectors containing shape indices and dielectric characteristics and reducing their amount of data, reduced by latent variables (LV), the total classification error is in the range  $e=2-15\%$ , and when using principal components (PC)  $e=4-20\%$ . The highest values of the total classification error ( $e=8-20\%$ ) were obtained using the Decision Tree (DT) method, followed by the Discriminant Analysis (DA) method ( $e=10-15\%$ ). The lowest values of total classification error ( $e=2-5\%$ ) were obtained using the Support vector machines (SVM)

Method. The results show that lower values of total classification error are obtained using feature vectors containing shape indices and dielectric characteristics.



a) by shape indices and spectral indices



b) by shape indices and dielectric characteristics

Figure 3. A summary analysis of the results for the classification of quail eggs  
 FV-feature vector, DRM-amount of data reduction method

The summarized analysis of the results obtained shows that the vectors of features containing shape indices, spectral indices and dielectric characteristics are suitable for classification depending on the change in weight of both hen and quail eggs. For both types of eggs, the results are comparable. The total classification error values obtained are similar for the two egg types. With the lowest values of this error the SVM Method ( $e=1-5\%$ ) works regardless of the method used to reduce the amount of data. Separation of eggs during storage, depending on their weight, is possible using non-linear separation functions. Consideration of changes in the internal characteristics of eggs is important since the external characteristics of the eggs such as shape index, long and short axis do not change significantly during storage, whereas the change in eggs weight may reach more than 3%. [25].

The results obtained in this paper confirm and supplement those of the available literature. When eggs are stored, they change their optical transmittance characteristics [21,30]. Using a genetic algorithm and a neural network method, a total classification error of  $e=0-18\%$  was obtained. The disadvantage of these methods is that they use complex computational procedures and the processing time is longer than the Decision Tree, Discriminant Analysis and Support vector machines as proposed here [15].

Spectral indices derived from transition spectra combined in vectors of shape index features derived from digital images of eggs and an appropriate classifier yielded a total classification error ( $e=1-5\%$ ) close to that using of complicated computing procedures.

The method presented here is inferior to the prediction accuracy over those using neural networks and classifiers [13,28,29] in the analysis of eggs during storage using visual images data and dielectric characteristics. The disadvantage of these methods is that they use sophisticated computational procedures and require a long time to process data, making them unusable directly on the production line.

#### 4. CONCLUSION

Approaches are adapted to track the change in eggs weight, depending on the stage of storage, are fully geared to the use of methods that would be sufficiently effective with respect to rapid and simplified classification and at the same time providing satisfactory accuracy according to the technological requirements for eggs evaluation.

The CORR, SFCPP and RELIEFF methods were found to be suitable for selection of features from shape indices and spectral indices of eggs according to the results obtained with the Naïve Bayes classifier. The ranking of features from shape indices and dielectric characteristics of eggs is successful by the CORR, RELIEFF and FSRNCA methods.

Appropriate methods for reducing the amount of feature vector data have been shown to be principal components and latent variables, because using them low values of the total classification error are obtained with Naïve Bayes classifier ( $e=4-23\%$ ).

From a comparative study to evaluate the impact of the methods used, on the accuracy of classification, depending on the eggs weight, it is found that the use of the three methods tested for classification of DT, DA and SVM for hen eggs gave better results at predicting their mass than for quail eggs. The comparative analysis shows that high classification accuracy is achieved by the SVM method. In this method, the total classification error is  $e=1-5\%$ . The use of the Decision Tree (DT) and the Discriminant analysis (DA) methods show good results ( $e=5-29\%$ ), which are much lower than those obtained with the Support vector machines (SVM) method.

The proposed egg analysis methods take precedence over the known solutions in this field because it takes into account the internal properties of quail and hen eggs by spectral indices derived from transmittance spectra and their dielectric characteristics, and not only by shape indices that do not change significantly when the product is stored.

#### 5. REFERENCES

- [1] Aboonajmi M., Akram A., Nishizu T., et al. (2010). An ultrasound based technique for the determination of poultry egg quality. *Research in Agricultural Engineering*, Vol.56, pp. 26-32.
- [2] Alikhanov D., Penchev S., Georgieva Ts., Moldajanov A., Shynybaj Z., Daskalov P. (2015). Indirect Method for Egg Weight Measurement Using Image Processing, *International Journal of Emerging Technology and Advanced Engineering IJETAE*, Vol. 5, Iss. 11, ISSN 2250-2459, pp. 30-34.

- [3] Alikhanov J., Penchev S., Georgieva Ts., Moldazhanov A., Shynybay Z., Daskalov P. (2017). An indirect approach for egg weight sorting using image processing. *Food Measurement and Characterization*, Vol. 12, No.1, pp.87-93.
- [4] Arabadzieva-Kalcheva N., N. Nikolov. (2017). Comparative analysis of the naive bayes classifier and sequential minimal optimization for classifying text in bulgarian in machine learning. *Computer science and technologies*, Year XV, No.1, ISSN 1312-3335, pp. 97-105. (in Bulgarian).
- [5] Atanassova S., Nikolov P., Valchev N., Masheva S., Yorgov D. (2019). Early detection of powdery mildew (*Podosphaera xanthii*) on cucumber leaves based on visible and near-infrared spectroscopy. *AIP Conference Proceedings*, 2075, 160014-1-160014-5.
- [6] Baycheva S., Dobрева K., Mihaylova G. (2019). Study on the ethanol extracts of bulgarian white oregano (*Origanum heracleoticum* L.). *Applied Researches in Technics, Technologies and Education (ARTTE)*, Vol. 7, No. 3, pp. 206-214.
- [7] Cermakova I., Komarkova J., Sedlak P. (2019). Calculation of Visible Spectral Indices from UAV-Based Data: Small Water Bodies Monitoring. *2019 14th Iberian Conference on Information Systems and Technologies CISTI*, 1-5.
- [8] Damyanov Ch., Nachev V., Titova T. (2010). Bayes classifiers. *Journal Automatics and informatics*, Vol. 2, pp.54-61.
- [9] Dimov M. (2012). Contemporary method for utilization of rubber wastes. *Management and education*, Vol. 8, No. 4, pp. 173-176.
- [10] FSNCA, Feature selection using neighborhood component analysis for classification. [online]. Available at: <https://www.mathworks.com/help/stats/fscnca.html>. (Accessed: 27.08.2019).
- [11] FSRNCA, Feature selection using neighborhood component analysis for regression. [online]. Available at: <https://www.mathworks.com/help/stats/fsrnca.html>. (Accessed: 27.08.2019).
- [12] Georgieva-Nikolova M., Zlatev Z. (2019). Prediction of the weight of quail eggs through shape features and spectral indices. *International Conference on Technics, Technologies and Education ICTTE 2019*, pp.125-132.
- [13] Javadikia P., Dehrouyeh M., Naderloo L., Rabbani H., Lorestani A. (2011). Measuring the weight of egg with image processing and ANFIS model. *SEMCCO, Part I, LNCS 7076*, (2011), pp. 407-416.
- [14] Kazlacheva Z., Ilieva J. (2018). An investigation of Design of Twist Knot Drapes. *The Aegean International Textile and Advanced Engineering Conference AITAE 2018*, 5-7 September 2018, Lesvos, Greece, IOP Conf. Ser.: Mater. Sci. Eng. 459 (2018) 012079.
- [15] Kirilova E., Daskalov P., Tsonev R., Draganova Ts. (2010), Selection of colour features for recognition of Fusarium damaged corn seeds. *Annual Conference RU & SU*, Vol. 49, book 3.1, ISSN 1311-3321, pp. 125-130. (in Bulgarian).
- [16] Kumar, E., Annamalai, K. (2014). Investigation into optimal fixturing cost of an assembly using genetic algorithm. *Engineering Review*, Vol. 34, No. 2, pp. 85-91.
- [17] Lin H., Zhao J., Sun L., Chen Q., Zhou F. (2011). Freshness measurement of eggs using near infrared (NIR) spectroscopy and multivariate data analysis. *Innovative Food Science and Emerging Technologies*, Vol. 12, pp.182-186.
- [18] Lukanov H. (2014). Egg quality traits in layers from different production types. *Agricultural Science And Technology*, Vol. 6, No. 2, pp.148-151.
- [19] Lukanov H. (2019). Meat colour characteristics of different productive types domestic quails. *Trakia Journal of Sciences*, Vol. 17, No. 1, pp. 42-48.
- [20] Lukanov H., Genchev A., Kolev P. (2019). Egg quality traits in wg, gg and gl japanese quail populations. *Trakia Journal of Sciences*, Vol. 17, No. 1, pp. 49-55.

- [21] Mehdizadeh S., Minaei S., Hancock N., Torshizi M. (2014). An intelligent system for egg quality classification based on visible-infrared transmittance spectroscopy. *Information processing in agriculture*, Vol. 1, pp.105-114.
- [22] Mladenov M. (2015). Complex assessment of the quality of food products through visual images, spectrophotometric and image analysis hyper-spectral characteristics. Monograph. University publishing Center of Rousse University "A. Kanchev", Rousse, 2015. (in Bulgarian).
- [23] Mladenov M., S. Penchev, M. Deyanov, M. Mustafa. (2015). Automatic classification of grain sample elements based on color and shape properties. *University Politehnika of Bucharest, Scientific Bulletin, Series C*, Vol. 73, No. 4, pp. 39-54.
- [24] Nethercott J.: Capacitance measurement library for Arduino, <https://github.com/codewrite/arduino-capacitor>, (accessed: 15.09.2019).
- [25] Ondrušíková S., Nedomová Š., Pytel R., Cwíková O. and Kumbár V. (2018). Effect of different storage times on Japanese quail egg quality characteristics. *Potravinárstvo Slovak Journal of Food Sciences*, Vol. 12, No. 1, pp. 560-565.
- [26] RELIEFF, Rank importance of predictors using ReliefF or RReliefF algorithm. [online]. Available at: <https://www.mathworks.com/help/stats/relieff.html> (Accessed: 27.08.2019).
- [27] SFCPP, Select Subset of Features with Comparative Predictive Power. [online]. Available at: <https://www.mathworks.com/help/stats/feature-selection.html>, (Accessed: 26.08.2019).
- [28] Soltani M., Omid M. (2015). Detection of poultry egg freshness by dielectric spectroscopy and machine learning techniques. *LWT – Food Science and Technology*, Vol. 62, No. 2, pp.1034-1042.
- [29] Soltani M., Omid M., Alimardani R. (2014). Egg quality prediction using dielectric and visual properties based on artificial neural network. *Food analytical methods*, Vol. 8, No. 3, pp. 710-717.
- [30] Titova T., V. Nachev, Ch. Damyanov, N. Bozukov, H. Dinkov. (2012). Determining the freshness of eggs using spectral data and wavelet network model. *Zbornik radova 56. Konferencije za ETRAN, Zlatibor, 11-14. juna 2012*, pp. VI1.5-1-4.
- [31] Titova T., V. Nachev, Ch. Damyanov, P. Nikovski. (2013). Intelligent classifiers for non-destructive determination of food quality. *Facta Universitatis, Series: Automatic Control and Robotics*, Vol. 12, No. 1, pp.19-30.
- [32] Trendafilov K., Delchev N. (2018). Headland turns using the tractor's "fi fth wheel" steering device instead of front steering wheels. *Bulgarian Journal of Agricultural Science*, Vol. 24, No. 6, pp.1137-1147.
- [33] Vasileva V., Gorbunova E., Vasilev A., Peretyagin V., Chertov A., Korotaev V. (2018). Assessing exterior egg quality indicators using machine vision. *British Poultry Science*, Vol. 59, No. 6, pp. 636-645.
- [34] Wang Y., Wang C. (2017). High resolution remote sensing image segmentation based on multi-features fusion. *Engineering Review*, Vol. 37, No. 3, pp. 289-297.
- [35] Wyman C., Sloan P.-P. and Shirley P. (2013). Simple Analytic Approximations to the CIE XYZ Color Matching Functions. *Journal of Computer Graphics Techniques JCGT*, 2 (2), 1-11.
- [36] Yongwei W., Wang J., Zhou B., Lu Q. (2009). Monitoring storage time and quality attribute of egg based on electronic nose. *Analytica Chimica Acta*, Vol. 650, pp.183-188.