sciendo

RURAL SUSTAINABILITY RESEARCH
51(346), 2024
ISSN – 2256-0939

Received: 8 June 2024 Revised: 3 July 2024 Accepted: 31 July 2024

Algorithm for Determination of Pepper Maturity Classes by Combination of Color and Spectral Indices

Miroslav Vasilev, Galya Shivacheva, Vanya Stoykova, *Zlatin Zlatev
Trakia University, Faculty of Technics and Technologies, 38 Graf Ignatiev str., 8602, Yambol, Bulgaria

Abstract. The aim of the present work is to propose methods and tools for classifying sweet pepper into groups according to their degree of maturity based on color and spectral characteristics extracted from color images on the surface of the vegetables. The investigated pepper is two varieties of sweet - red *Banji* and yellow *Liri*. Five groups were formed, depending on the degree of maturity, and 16 color and 11 spectral indices were calculated for each of the groups. By successively using the ReliefF and PLSR methods, a selection of informative features and subsequent reduction of the vector formed by them was carried out, thereby aiming to increase the predictive results and minimize the time for data processing. The obtained classification errors between the individual stages of ripening vary according to the type of pepper and depending on which of the two types of maturity the fruits are in - technical or biological. For red sweet pepper, the separation inaccuracy obtained using a discriminant classifier with a quadratic separation function is in the range of 8 - 19%, while for yellow it is from 5 to 23%. The results obtained in the present work for the classification of pepper into groups according to their degree of maturity would support decision-making in selective harvesting and overall more accurate and efficient management of the harvesting process from the point of view of precision agriculture. The work will continue with studies related to the prediction of various compounds indicating changes in the color of peppers, including chlorophylls, carotenes and xanthophylls. In this way, it is possible to increase the accuracy in determining the degree of ripeness, since in pepper the color does not always follow the same pattern of change from green to yellow to orange to red.
Key words: maturity classification, ReliefF method, discriminant analysis, error rates.

Introduction

The favorable natural conditions and established traditions in Bulgaria in the field of vegetable production are not sufficient in the cultivation of sweet pepper to achieve optimal yields corresponding to those obtained in neighboring countries, where in most cases the production is high-tech. In order to develop a highly efficient and economically justified production of pepper in our country as well, organizational, technological, and resource factors occupy a leading place. One such factor directly related to the need to develop technological innovations and determine the economic feasibility of growing sweet pepper is the stage at which the fruits will be harvested from the field. In the case of pepper, this is because sweet pepper peppers can be consumed both at technical ripeness (peppers are unripe and have a green color) and biologically (when the fruits have the characteristic red or yellow color of a ripe pepper). The economic importance of sweet pepper maturity can be justified by the influence of this indicator on the marketability, quality, and value of the crop. By understanding the different stages of ripeness and harvesting the fruit at the right time, farmers can optimize their yields and profitability while meeting the needs and preferences of their customers (Lehnert et al., 2020).

Climatic advantages and established traditions in the field of vegetable production are not sufficient in the cultivation of sweet pepper to achieve competitiveness and optimal yields in the Republic of Bulgaria. In order to develop highly efficient and economically justified production in this vegetable, organizational, technological, and resource factors occupy a leading place. One such factor directly related to the need to develop technological innovations and determining the economic feasibility of growing sweet pepper is

* Corresponding Author's email:
zlatin.zlatev@trakia-uni.bg

DOI: 10.2478/plua-2024-0010

the stage at which the fruits will be harvested from the field.

As already emphasized, two types of maturity are observed in pepper. The color is of different saturation. At full or biological maturity, a rich range of colors is observed. The colors are purple, scarlet, orange, red, and yellow (Description, 2023).

The evaluation of pepper ripeness, for the purpose of harvesting and sorting the crop, is done visually, relying above all on the experience of the evaluating expert. Modern technical means such as robots and sorting machines, using a variety of sensors, control devices, actuators, and regulating bodies, require that data on pepper maturity be obtained in an automatic manner. For this purpose, video cameras, spectrophotometers, hyperspectral and multispectral cameras, magnetic resonance, and sound sensors are used as sources of data on pepper maturity. To process these data, classification and regression methods are used, as well as significantly more complex computational procedures, such as neural networks (Kootstra et al., 2021).

Postharvest Center of the University of California (Bell Pepper, 2023) offers flyers for various fruits, vegetables, and ornamental flowers. In the pepper section, the team has developed a brochure that briefly describes the development of the plant and what diseases and pests are more common on it. These brochures are extremely useful to producers, traders, and agronomists in the cultivation and distribution of agricultural produce. The scarce information on the color characteristics of the pepper during its ripening can be pointed out as a disadvantage.

In recent years, there has been a trend towards extracting knowledge from databases with visual information about pepper. Harel et al. (2020a) analyze an available image dataset with 296 images. The authors extracted 24 features from pepper images, divided into 4 classes. As a next step, they grade the pepper by degree of ripeness. Their training samples are 20–50% and test samples are 50–80% of the data. Two classification algorithms, logistic regression and random forest, were used. For yellow and red pepper, the authors achieved an accuracy of 90–98%, depending on the classifier and pepper variety used. A major contribution of the team of scientists is the proposed variant for training classifiers with a small sample size of 6-48 images, which achieves the specified classification accuracy. Harel et al. (2020b) continued this work, determining the influence of camera angle and fruit position as factors affecting pepper maturity recognition accuracy. The authors found that the orientation of the fruit had a significant effect on the classification accuracy. When photographing the fetus from below, they get the highest classification accuracy compared to the other observation points. Accuracy increases by 15-20%.

Van Essen et al. (2022), propose a point selection algorithm for obtaining images of pepper when determining its ripeness. Applying this method, the authors increase the classification accuracy by 6% and reduce the harvesting time by 3-21%. A disadvantage of this development is that the algorithm is not efficient enough in all possible cases of photographing pepper fruits. According to the scientists, it is necessary to search for more effective methods for optimal determination of the shooting point of pepper, when determining its maturity.

Cong et al. (2023), propose classification procedures for determining pepper maturity from color digital image data. The authors use a variant of a neural network. Recognition accuracy is increased to 99%.

Some limitations of known solutions in the subject area can be summarized. The pepper images used for classification and regression are at relatively the same degree of maturity. Different classification procedures are used in different publications, which makes comparative analyses and the selection of appropriate algorithms and procedures for a given pepper variety difficult. Due to the different varietal characteristics of pepper, it is not possible to fully generalize the task of recognizing and classifying the degree of ripeness of the fruits of the plant. The techniques of hyperspectral analysis (Lorente et al., 2012) and nuclear and magnetic resonance have been successfully applied in laboratory conditions for the classification of pepper according to degree of maturity. A major limitation of these methods is their direct application in the field. The problem arises from the fact that they are associated with relatively large volumes of data that are obtained from measuring devices, as well as the lower distribution of these devices due to the complexity of use and their higher cost, compared to video cameras and spectrophotometers. In terms of data processing and analysis, complex procedures such as neural networks have been used. These methods increase the accuracy of pepper maturity determination compared to classical ones such as discriminant analysis, decision trees, and regression methods. On the other hand, the use of complex computational procedures is difficult to implement in embedded robot control systems for harvesting and sorting the pepper crop. In addition, there are no clearly defined grades of pepper maturity. There are two main classes: technical and biological maturity, which in various publications the authors divide into several subclasses (Cong et al., 2023; Van Essen et al., 2022; Villaseñor-Aguilar et al., 2020).

The improvement and refinement of the technologies for harvesting and sorting pepper, depending on its degree of maturity, requires producers to introduce such technical and informational means that lead to an effective technological and economic effect. In most cases, when using automatic sorting algorithms, there is a discrepancy between the results obtained in laboratory conditions and directly in the field. This necessitates constant improvement of automatic collecting and sorting machines (Liu et al., 2024).

The aim of this work is to propose methods and tools for classifying sweet peppers into groups according to their degree of maturity based on color and spectral characteristics extracted from color images of the surface of the vegetables.

Materials and Methods

Two varieties of sweet pepper were studied: red *Banji* (*Capsicum chinense*) and yellow *Liri* (*Capsicum annuum*). Data on the color change on the surface of peppers were taken from available literature sources (Harel et al., 2017; Harel et al., 2020; van Essen et al., 2022). Images of 70 yellow and 70 red peppers are divided into five groups (classes), depending on the degree of maturity (classes from 1 to 5). 60% of the images were used to create the classification procedures and 40% to check their accuracy.

This work uses the results obtained by Harel et al. (2017) on knowledge extraction from visual information databases of peppers. It uses the result of Logistic Regression, Random Forest classification algorithms, and nonlinear discriminant analysis with a quadratic partition function. Furthermore, in this research, results from Harel et al. (2020) are incorporated; the work did include the pepper fruit's orientation to add accuracy to the classification.

A total of 16 color indices were calculated according to the known equations, presented in Pathare et al. (2013). For both types of pepper: L, a, b, C, h, YI, WI, BI, SI, CIRG, COL, CI, ECB, FCI, WL, PACI.

The component L (Lab) is related to changes in illuminance levels and ranges from 0 to 100. The next component of the Lab model a(Lab) shows the proportions between red and green color and has a value of -128 to +127, while the changing is within the same limits. b (Lab) indicates changes from blue to yellow hue of color have occurred. The component C (LCh) determines the distance at which the color is located from the center of the color circle. The component h (LCh) indicates at what angle the color is located relative to the origin of the coordinate system of the color circle. White Index (WI) represents how white a color is in terms of human perception. Yellow Index (YI) indicates color changes from white to yellow. The brown index (BI) is related to the oxidation of the phenolic compounds in the products and changes with their variations. SI determines how light or dark an object's color is based on a comparison to white and black. The CIRG index shows the changes in anthocyanin levels in the product. COL is an index reflecting variations in the carotenoid level. CI corresponds to the ratio of the red to yellow-blue component, indicating the dominant color direction. Changes in the ECB index show the extent to which the color shifts from yellow to pink. The FCI index represents the difference between the lightness and the yellow level of the object, while the WL shows the ratio between the latter two. The red to yellow color value is represented by the PACI index.

The presented indices are calculated by the L, a, b, C, and h color components (which are also used as indices) using the following formulas:

$$C = \sqrt{a^2 + b^2} \tag{1}$$

$$h^o = tan^{-1}\left(\frac{b}{a}\right) \tag{2}$$

$$WI = \frac{142,86b}{L} \tag{3}$$

$$YI = 100 - \sqrt{(100 - L)^2 + a^2 + b^2} \tag{4}$$

$$BI = \frac{x - 0,31}{0,17} \quad x = \frac{a + 1,75L}{5,645L + a - 0,012b} \tag{5}$$

$$SI = \sqrt{a^2 + b^2} \tag{6}$$

$$CIRG = \frac{180 - h}{L + C} \tag{7}$$

$$COL = \frac{2000a}{LC} \tag{8}$$

$$CI = \frac{a}{b} \tag{9}$$

$$ECB = \frac{a}{b} + \frac{a}{L} \tag{10}$$

$$FCI = L - b \tag{11}$$

$$WL = \frac{L}{b} \tag{12}$$

$$PACI = \frac{1000a}{L + h} \tag{13}$$

Using functions to convert RGB data to spectral features, spectral data for 5 classes of pepper were obtained (Smits, 2000; Wyman et al., 2013).

The calculated spectral indices are 11 in number: REI, PTI, CTI, TVI, G, NexG, NGRDI, RGBVI, GLI, VARI, and ExG (Cermakova et al., 2019; Atanassova et al., 2019).

Redness Index (REI). It provides information on the boundary VIS and NIR spectral regions. The photochemical reflectance index (PRI) provides information on changes occurring in the amounts of carotenoids. High carotenoid index (CRI) levels suggest an increase in carotenoid values. Triangulation Index (TVI) describes the radiant energy absorbed by the pigment as a function of the relative difference between the red and near-infrared ranges. Green Index (G) indicates variations in green color. The greenness index with normalized kurtosis (NExG) gives information about the extent to which the green color covers the plant under study. Normalized Green-Red Difference Index (NGRDI) presents information about what area of the plant is green and what is red. The red-green-blue vegetation index (RGBVI) signals deviations in the appearance of yellow spots on the investigated vegetation. Green Leaf Index (GLI) values range from -1 to +1. Negative values indicate the presence of dark areas in the captured objects, while positive values refer to green-colored areas. The Visible Atmospheric Resistance Index (VARI) aims to reveal color in the visible part of the spectrum while mitigating differences in illumination and atmospheric pressure. Using the excess green index (ExG), it is possible to obtain information about which areas of the studied object are green and which are of another color. Illumination has a significant influence on this index.

The vegetative indices used in the work were obtained by spectral reflection in the wavelengths 420, 510, 520, 530, 550, 570, 620, 670, 680, 720, 740, and 750 nm. These wavelengths are relevant for maturity analysis in red and yellow peppers because they capture pigment concentration and variation of chlorophyll content (Rizzo et al., 2023). The wavelengths include critical ranges for absorption and reflection by chlorophyll, carotenoids, and anthocyanins that take part in plant-ripening processes. These formulas are, therefore, very different in maturity classes by applying these wavelengths to the vegetative indices. When peppers ripen, this differentiation is based on changes in color and composition.

The indices are calculated by the formulas:

$$REI = \frac{R_{740}}{R_{720}} \tag{14}$$

$$PRI = \frac{R_{530} - R_{570}}{R_{530} + R_{570}} \tag{15}$$

$$CRI = \frac{1}{R_{510}} - \frac{1}{R_{550}} \tag{16}$$

$$TVI = 0{,}5(120(R_{750} - R_{550}) - \\ - 200(R_{670} - R_{550})) \tag{17}$$

$$G = \frac{R_{550}}{R_{680}} \tag{18}$$

$$NExG = \frac{2R_{520} - R_{620} - R_{420}}{R_{520} + R_{620} + R_{420}} \tag{19}$$

$$NGRDI = \frac{R_{520} - R_{620}}{R_{520} + R_{620}} \tag{20}$$

$$RGBVI = \frac{{R_{520}}^2 - R_{620}R_{420}}{{R_{520}}^2 + R_{620}R_{420}} \tag{21}$$

$$GLI = \frac{2R_{520} - R_{620} - R_{420}}{2R_{520} + R_{620} + R_{420}} \tag{22}$$

$$VARI = \frac{R_{520} - R_{620}}{R_{520} + R_{620} - R_{420}} \tag{23}$$

$$ExG = 2R_{520} - R_{620} - R_{420} \tag{24}$$

Using the ReliefF method (Georgieva et al., 2020), a vector of the most informative features for pepper classification, depending on its degree of maturity, was selected. According to Farrell et al. (2019) the most significant features are those with weight coefficients with value above 0.6.

Data volume reduction of the selected feature vector was done with the Partial Least Squares Regression (PLSR) method (Yotova et al., 2016).

The data were classified with discriminant analysis (Kirilova, 2012), by means of a non-linear quadratic dividing function, which has the form:

$$\delta(x) = K + v.L + v'.Q.v \tag{25}$$

where: K is a constant;
L – linear coefficient;
Q – quadratic coefficient;
v=[x;y] – data matrix;
x and y are the first two latent variables;
v' – transpose matrix of v.

The efficiency of using the classification procedure was evaluated by the overall classification error.

$$e = \frac{\sum_{i=1}^{m} FN_i}{\sum_{i=1}^{m} TP_i + \sum_{i=1}^{m} FN_i}.100,\% \tag{26}$$

where: FN is the number of data from class i incorrectly assigned to other classes;
TP is the number of correctly classified data of class i.

Matlab 2017b software system (The Mathworks Inc., Natick, MA, USA.) was used to process the experimental data.

Literature sources regarding color changes at the pepper surface are very similar to those obtained from research data. In this work, an overall analysis was made with a representative population of 140 samples of peppers, split into two parts: 70 samples of red sweet peppers and 70 samples of yellow peppers. The color RGB images were acquired with an DFK 41AU02 (The Imaging Source Europe GmbH., Bremen, Germany) camera mounted on a fixed stage. Since 360-degree characterization of a pepper is not possible in a real greenhouse, the viewpoints are limited to one viewpoint from one side of the pepper. The camera has a distance of 25 cm from the pepper which corresponds to the minimum depth distance of the camera. The pepper samples were evaluated regarding their degree of maturity. The peppers could be rated into five different classes or categories, including an in-depth rating for their ripeness. To ensure a strong and reliable classification in this analysis, a leave-one-image-out cross-validation method was employed. For this, classifiers were built using 60% of the images, which were trained as models. Then, validation was performed on the remaining images, consisting of 40%, to check classifier performance on unseen data. This choice of methodology avoids overfitting and ascertains the checking of model accuracy against a very different set of samples. Level of significance All the statistical analyses and data processing were done assuming a significance level of $\alpha = 0.05$. This means that if the probability of seeing the data assuming there is no true effect is less than 5%, it is considered significant. It is a strict statistical criterion, which states that the deductions made from the data are valid and the changes observed in color or the maturity classifications are not a matter of chance.

Results

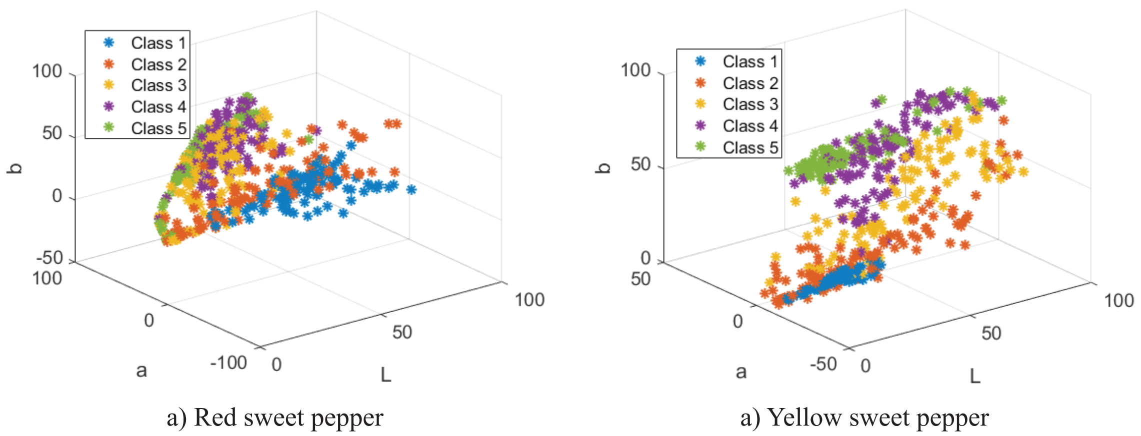
Figure 1 shows the Lab characteristics of red and yellow sweet pepper. For the red peppers, there exists a huge overlap of their data by the first two and the last three classes; this means that there are two possibilities of groupings for the peppers – namely, one class for 1-2 and another for 3-5. This just corresponds to the two stages of ripening – technical and biological. For yellow peppers, this means that classes 1-3 are well separated from one another, though classes 4 and 5 overlap strongly with one another. Again, one has two groups: classes 1-3 represent technical maturity, and classes 4-5 indicate biological maturity.

Table 1 shows data on the variation of color indices, depending on the ripeness class of red sweet pepper. The SI index shows a change proportional to the color change of the pepper during ripening. With the ripening process of the peppers, there is a general reduction of “L” and h while “a” increases, along with “C”, the accumulation of pigments, and the PACI increase. The YI increases a lot, but the WI decreases constantly. The color saturation increases towards the middle stages to reach its peak value, and there is a huge variation in other indices like CIRG, COL, and ECB, among others. All these changes reflect the transition from technical to biological maturity, with a marked shift in color and pigment concentration.

Table 2 shows data on the variation of color indices, depending on the maturity class of yellow sweet pepper. The CI index shows a change proportional to the color change of the pepper during ripening. The rest of the signs show an increase and then a gradual decrease in their values. This is a prerequisite for weaker separability between classes if each feature is used separately.

Figure 2 shows averaged spectral characteristics of red and yellow sweet pepper. The characteristics of the red pepper overlap, and those of the yellow pepper show a smooth change in the ripening process. This is an indication that the direct use of spectral characteristics to determine the degree of ripeness for pepper is not appropriate. It is necessary to reduce the data from these features to an appropriate set of informative features as commented by Rizzo et al. (2023).

Table 3 shows data on the variation of spectral indices, depending on the ripeness class of red sweet pepper. The spectral indices show variation depending on the class of pepper to which they refer. With a



a) Red sweet pepper

a) Yellow sweet pepper

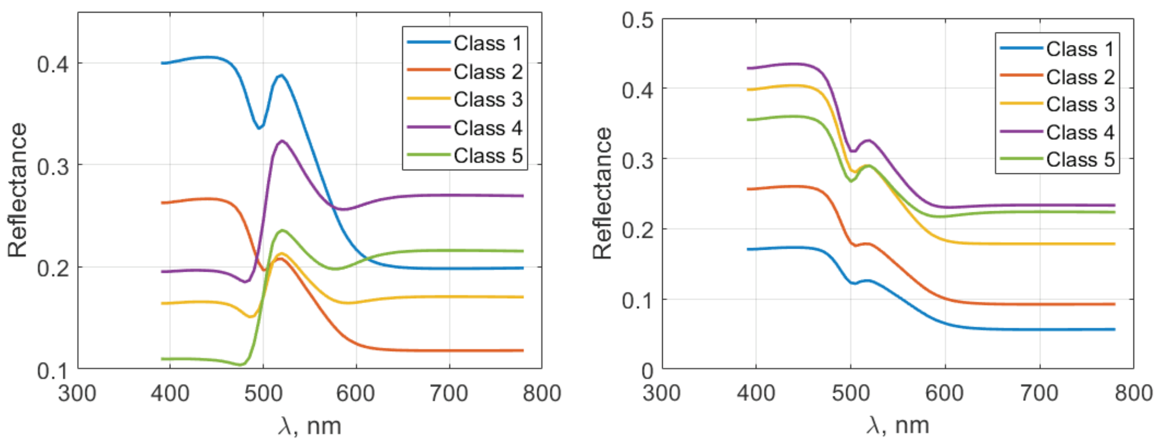
Figure 1. Lab values of sweet pepper.

Table 1

Sweet red pepper color indices values

Index \ Class	Class 1	Class 2	Class 3	Class 4	Class 5
L	46.60±0.00	32.41±16.87	29.13±20.40	41.02±14.53	30.75±11.69
a	25.54±0.00	19.79±60	27.62±13.52	46.60±16.39	45.85±13.82
b	26.57±0.00	29.82±12.45	30.81±15.75	40.76±14.69	33.68±10.61
C	2.62±0.00	3.67±1.05	6.99±2.10	9.26±2.59	8.73±1.28
h	0.76±0.00	0.98±0.23	0.85±0.26	0.73±0.28	0.62±0.15
YI	95.30±0.00	149.31±50.09	160.37±41.92	148.16±37.98	168.28±34.21
WI	33.63±0.00	18.83±16.67	14.33±9.99	12.60±9.10	7.56±8.75
BI	22.29±0.00	39.14±27.11	23.86±84.13	11.94±81.61	9.18±2.68
SI	37.57±0.00	36.77±11.72	42.54±18.95	62.36±19.66	57.15±15.71
CIRG	4.28±0.00	10.57±2.06	8.51±17.81	3.89±12.34	6.89±1.61
COL	635.75±0.00	910.09±514.44	536.21±1548.15	261.74±843.61	484.56±141.14
CI	1.48±0.00	0.76±2.19	1.00±0.48	1.17±0.59	1.48±0.33
ECB	2.11±0.00	1.56±2.08	2.19±1.14	2.36±1.46	3.23±0.70
FCI	21.21±0.00	6.82±21.95	5.94±7.42	7.61±5.16	6.96±7.22
WL	3.78±0.00	1.05±7.85	0.96±0.35	1.04±0.32	0.90±0.34
PACI	648.67±0.00	789.74±294.38	1190.55±560.18	1167.63±902.68	1681.95±429.33

All data are statistically significant at $p<0.05$



a) Red sweet pepper

a) Yellow sweet pepper

Figure 2. Averaged spectral characteristics of sweet pepper.

Table 2

Sweet yellow pepper color indices values

Index \ Class	Class 1	Class 2	Class 3	Class 4	Class 5
L	19.22±6.06	30.14±21.98	50.07±20.76	58.12±16.78	51.68±13.14
a	20.1±3.69	20.97±11.15	15.42±9.88	8.84±5.96	16.4±7.11
b	17.01±4.25	28.17±17.44	48.36±18.45	60.37±12.89	56.2±8.98
C	1.69±0.52	2.73±1.45	5.5±1.94	7.92±1.34	8.49±0.41
h	0.69±0.06	0.90±0.26	1.23±0.21	1.41±0.13	1.28±0.14
YI	134.03±34.80	148.9±41.38	142.81±25.17	151.88±13.81	158.89±16.36
WI	14.78±4.97	16.73±9.09	23.24±6.13	22.89±3.08	22.27±4.68
BI	7.43±4.46	16.99±34.01	277±1316.92	95.24±106.08	86.51±447.86
SI	26.37±5.45	35.87±19.34	51.8±18.18	61.43±12.28	59.13±7.81
CIRG	9.6±3.69	9.32±7.66	4.08±2.50	2.94±0.89	3.09±0.55
COL	1630.93±1252.92	942.14±796.98	236.45±413.70	54.7±86.01	84.05±44.15
CI	1.21±0.14	0.88±0.48	0.37±0.27	0.16±0.14	0.31±0.15
ECB	2.34±0.34	1.75±0.86	0.73±0.51	0.34±0.30	0.67±0.34
FCI	4.15±3.69	6.57±7.79	4.95±5.61	4.67±2.96	6.72±5.14
WL	1.16±0.38	1.06±0.41	1.03±0.19	0.95±0.09	0.91±0.15
PACI	1177.5±334.51	916.63±503.08	370.49±258.98	178.32±163.93	348.27±184.31

All data are statistically significant at $p<0.05$

Table 3

Sweet red pepper spectral indices values

Index \ Class	Class 1	Class 2	Class 3	Class 4	Class 5
REI	0.91±0.01	0.93±0.02	0.96±0.03	0.96±0.02	0.98±0.02
PRI	0.04±0.01	0.04±0.01	0.03±0.01	0.02±0.01	0.02±0.01
CRI	0.09±0.05	1.95±14.43	0.89±2.19	0.38±1.05	1.8±4.81
TVI	6.45±6.93	2.49±2.43	4.15±2.93	8.12±3.98	6.88±3.66
G	1.32±0.47	1.54±0.68	0.88±0.68	0.61±0.18	0.42±0.16
NExG	0.08±0.06	0.12±0.08	0.13±0.13	0.15±0.11	0.29±0.12
NGRDI	0.13±0.11	0.19±0.12	0.17±0.16	0.20±0.13	0.36±0.13
RGBVI	0.13±0.11	0.19±0.12	0.18±0.17	0.21±0.14	0.38±0.15
GLI	0.13±0.11	0.19±0.12	0.18±0.17	0.21±0.14	0.38±0.15
VARI	0.33±0.28	0.48±0.34	0.31±0.29	0.32±0.17	0.53±0.16
ExG	0.08±0.04	0.07±0.06	0.04±0.03	0.09±0.05	0.10±0.04

All data are statistically significant at $p<0.05$

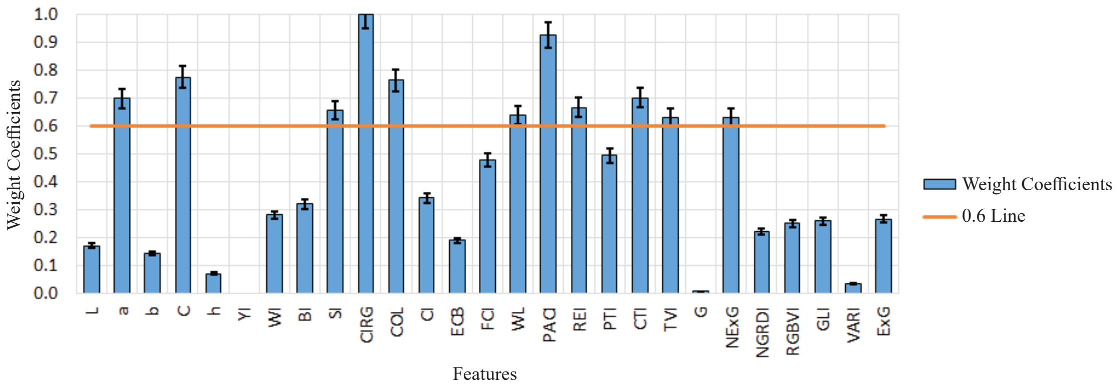


Figure 3. Feature selection.

All data are statistically significant at $p<0.05$

significant change, the G, NExG, and GLI indexes can be distinguished.

Table 4 shows data on the variation of spectral indices, depending on the maturity class of yellow sweet pepper. The spectral indices show variation depending on the class of pepper to which they refer. With a significant change, the G and GLI indexes can be distinguished.

Figure 3 shows the results of a selection of informative features for classification. The features are the same for both types of pepper. More color indices than components are selected. The smallest number are the spectral indices. This is due to their correlation with color characteristics. The most informative features are those with weight coefficients above 0.6 according to Farrell et al. (2019).

A vector of 11 informative features is formed. In this way, the volume of data is reduced by half. It contains 2 color components, 5 colors, and 4 spectral indices. The vector has the following form:

$$FV=[a\ C\ SI\ CIRG\ COL\ WL\ PACI\ REI\ CTI\ TVI\ NExG] \quad (27)$$

The classification results of red sweet pepper are visualized in figure 4. For the classes covering the technical maturity of the pepper, lower classification error values were observed than for the biological maturity deviation.

Between classes 1 and 2, the separation inaccuracy is 9%, and between 2 and 3, also covering technical maturity, it does not exceed 8%. Significantly higher

Table 4

Sweet yellow pepper spectral indices values

Index \ Class	Class 1	Class 2	Class 3	Class 4	Class 5
REI	0.91±0.00	0.91±0.01	0.93±0.01	0.94±0.01	0.95±0.01
PRI	0.04±0.00	0.04±0.00	0.04±0.00	0.04±0.00	0.04±0.00
CRI	0.19±0.07	0.19±0.15	0.10±0.07	0.08±0.03	0.10±0.03
TVI	1.32±0.95	2.71±2.67	2.68±2.28	1.53±1.12	1.24±2.17
G	1.63±0.49	1.89±0.79	1.52±0.33	1.43±0.18	1.29±0.12
NExG	0.12±0.05	0.14±0.06	0.12±0.04	0.11±0.03	0.09±0.02
NGRDI	0.20±0.09	0.23±0.11	0.19±0.07	0.17±0.04	0.14±0.04
RGBVI	0.20±0.09	0.23±0.11	0.19±0.07	0.18±0.04	0.14±0.04
GLI	0.20±0.09	0.23±0.11	0.19±0.07	0.18±0.04	0.14±0.04
VARI	0.52±0.29	0.64±0.37	0.47±0.22	0.42±0.13	0.33±0.10
ExG	0.05±0.02	0.09±0.06	0.12±0.07	0.12±0.03	0.09±0.04

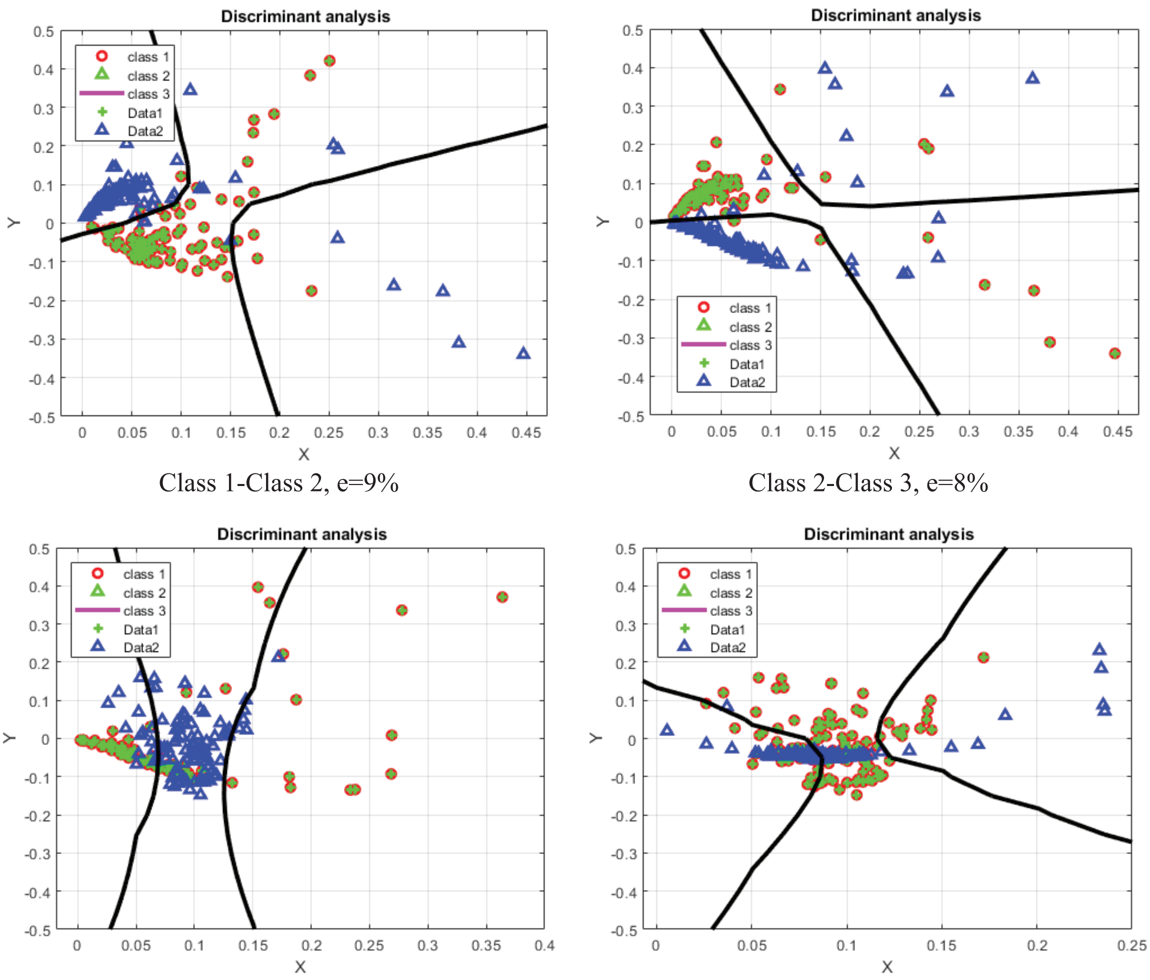
All data are statistically significant at $p<0.05$

The following separation functions are defined for sweet red pepper classes:

Class (C)	Function	
C1-C2	$\delta_{1,2} = 0{,}01 + 31.75x - 34.19y - 164x^2 + 244xy + 33.26y^2$	(28)
C2-C3	$\delta_{2,3} = -0.09 - 3.48x + 26.71y + 29.35x^2 - 170.4xy - 42.88y^2$	(29)
C3-C4	$\delta_{3,4} = 4.51 - 107.16x + 1.15y + 5.36x^2 - 27.38xy - 8.34y^2$	(30)
C4-C5	$\delta_{4,5} = -1.99 + 46.31x + 12.01y - 2.41x^2 - 105.42xy + 21.88y^2$	(31)

The following separation functions are defined for classes of sweet yellow pepper:

Class (C)	Function	
C1-C2	$\delta_{1,2} = -3.88 + 77.17x - 43.22y - 334x^2 + 376xy + 103y^2$	(32)
C2-C3	$\delta_{2,3} = -1.01 + 36.98x - 31.8y - 164x^2 + 169.02xy + 20.25y^2$	(33)
C3-C4	$\delta_{3,4} = 0.41 - 19.26x + 31.04y + 75.15x^2 - 139.42xy - 9.92y^2$	(34)
C4-C5	$\delta_{4,5} = -1.01 + 36.98x - 31.8y - 164x^2 + 169.02xy + 20.25y^2$	(35)



Class 1-Class 2, e=9%

Class 2-Class 3, e=8%

Class 3-Class 4, e=12%

Class 4-Class 5, e=19%

Figure 4. Classification of red sweet pepper with discriminant analysis – general view.

The following separating functions can be defined for the two classes of technical and biological maturity of pepper:

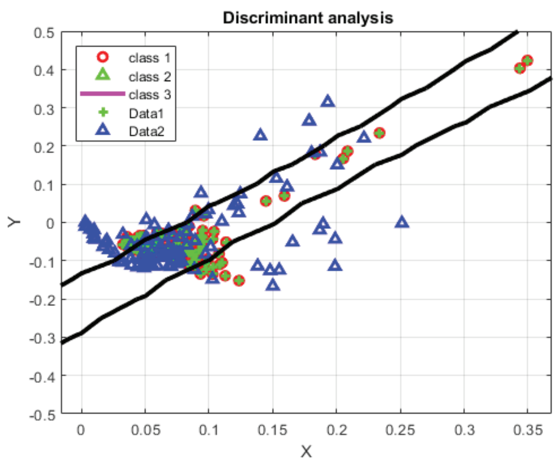
Class	Function	
Red	$\delta_{R\,(T,\ B)} = -0.56 + 33.95x - 45.86y - 1.61x^2 + 198.62xy + 86.97y^2$	(36)
Yellow	$\delta_{Y\,(T,\ B)} = 1.86 - 50.6x + 41.47y + 2.52x^2 - 312xy + 20.69y^2$	(37)

classification errors were observed between biological maturity stages. Between 3 and 4, it is 12%, and at the last two stages of ripening, it reaches 19%.

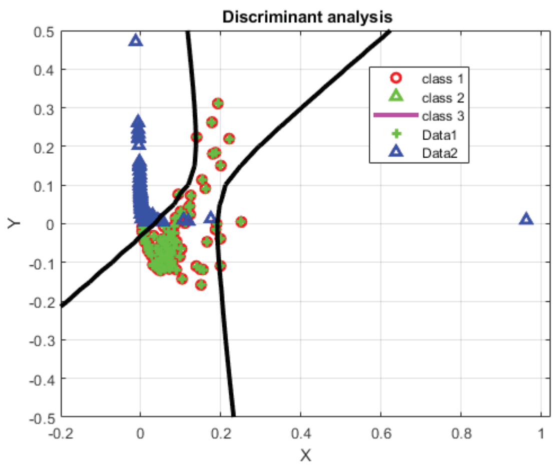
The classification results of yellow sweet pepper are visualized in figure 5. The highest values of the classification error were observed in the classes covering the technical maturity and were smaller between the classes covering the biological maturity of the pepper fruits. Between classes 1 and 2, the most significant separation error is observed (23%), and between the other compared classes it does not exceed 8%.

The differences in the classification errors of the two pepper cultivars in technical and biological maturity can be explained by the different signs (color and spectral) that are greatly influenced by the change of pigments in the pepper fruit. In red sweet pepper, the change from green to red color is due to the carotenoids capsanthin and capsorubin, while in yellow pepper, the change from green to yellow color is mainly due to violaxanthin, and to a lesser extent to lutein and beta-carotene.

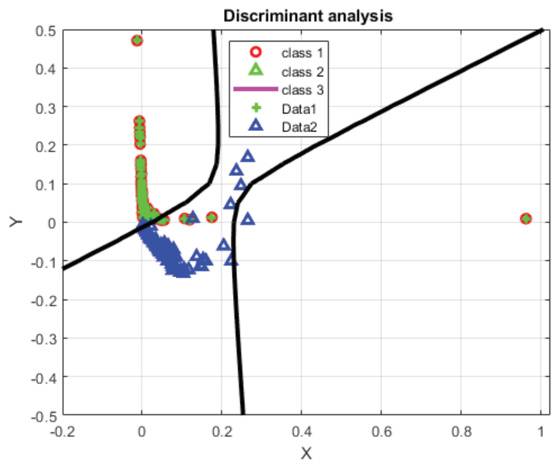
A check was made for the separability of sweet pepper at technical (classes 1 and 2) and biological



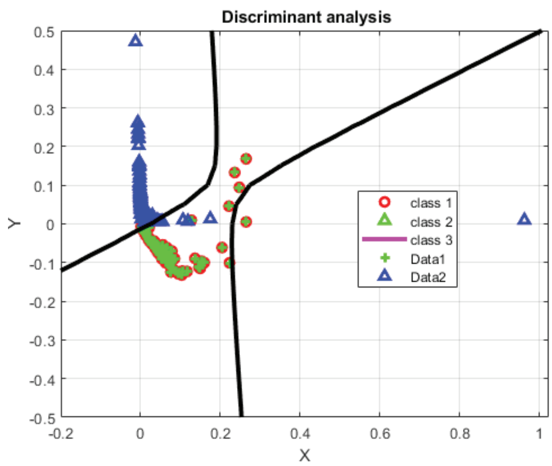
Class 1-Class 2, e=23%



Class 2-Class 3, e=8%

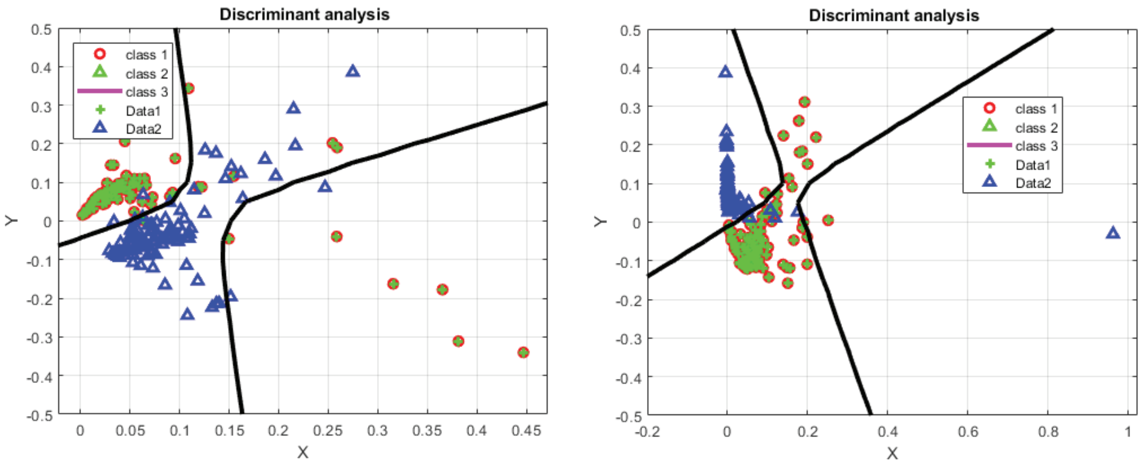


Class 3-Class 4, e=5%



Class 4-Class 5, e=5%

Figure 5. Classification of yellow sweet pepper with discriminant analysis – general view.



a) Red pepper, e=6% a) Yellow pepper, e=4%

Figure 6. Separation between technical and biological maturity of sweet pepper.

maturity (classes 3, 4, and 5). Figure 6 visualizes the results of this check.

Separability, with a classification error not exceeding 6%, was observed between the classes representing technical maturity in red sweet pepper (1 and 2) and classes 3, 4, and 5 covering biological maturity. In the case of yellow sweet pepper, the inaccuracy in separation between the object areas of the two types of maturity—technical and biological—is 4%.

Figure 7 shows a block diagram of the algorithm for determining the degree of ripeness of pepper. The algorithm is similar for both varieties – red and yellow. Once the vector containing color and spectral indices and features is loaded, it is reduced to two latent variables. The resulting latent variables (LVs) are processed with a discriminant classifier using the quadratic partition function presented above. A check is made to see if the pepper data belongs to Class 1. If it belongs to Class 1, a final decision is made - to fall into this group. Otherwise, a check is made to see if it belongs to Class 2. If so, it is assigned to that class. Otherwise, pepper data is assigned to the next class. The procedure continues until the pepper falls into the last class.

Figure 8 shows a block diagram of the algorithm for determining the degree of ripeness of pepper—technical and biological. The algorithm is similar for both varieties—red and yellow. Once the vector containing color and spectral indices and features is loaded, it is reduced to two latent variables. The resulting latent variables (LVs) are processed with a discriminant classifier using a quadratic partition function. A check is made to see if the pepper data belongs to the Technical Maturity Class. If he belongs to this class, a final decision is made—to fall into this group. Otherwise, it is assigned to Class "biological maturity."

Table 5 shows data for a test of pepper classification algorithms, depending on their maturity. The algorithms were applied to evaluate the ripeness of 30 peppers each of the two types of red and yellow pepper, which were not used in its preparation. Grades 1 and 2 fall into the "technical maturity" class, and grades 3, 4, and 5 fall into the biological maturity class for both types of pepper.

The resulting error is e=13-18% for red sweet pepper and e=10-16% for yellow sweet pepper. This shows that in a test, the proposed algorithm can be used to determine the degree of ripeness of red sweet pepper, with an accuracy of 82-87%. For yellow sweet pepper, this accuracy is 84-90%.

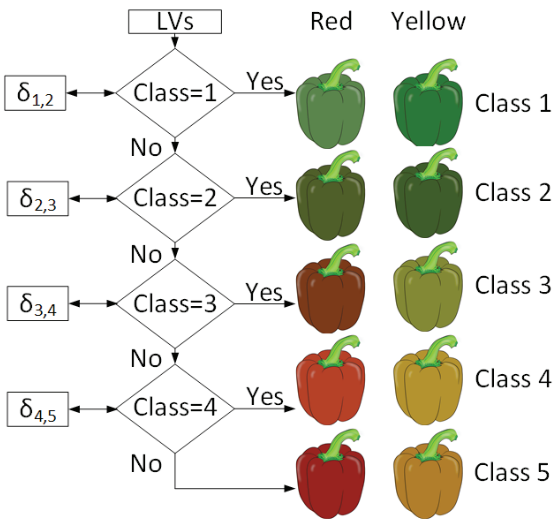


Figure 7. Algorithm for determination of sweet pepper maturity for 5 classes.

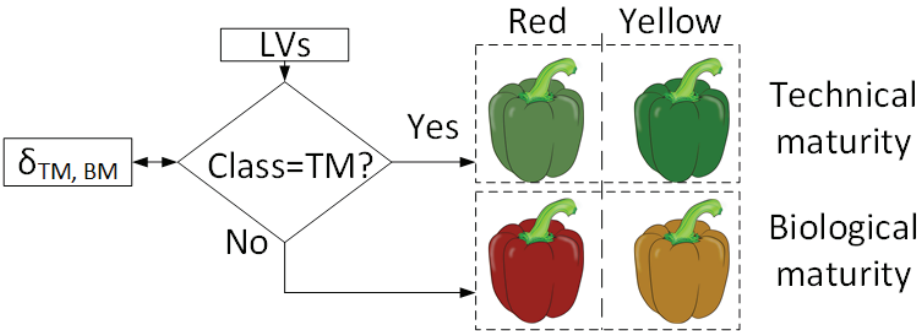


Figure 8. Algorithm for determination of sweet pepper maturity for two classes – technical and biological maturity (TM-technical maturity; BM-biological maturity).

Table 5

Results from algorithm tests (e, %)

Class \ Pepper	Red	Yellow
1-2	18%	16%
2-3	18%	13%
4-5	13%	10%
TM[1]-BM[2]	17%	14%

[1]TM-technical maturity; [2]BM-biological maturity

Discussion

In this paper, the results of Harel et al. (2020a) related to knowledge extraction from pepper visual information databases. To the Logistic regression and Random forest classification algorithms, results for nonlinear discriminant analysis using a quadratic partition function have been added. The results obtained are close. Classification accuracy reaches 90%.

Added the results of Harel et al. (2020b), who used pepper fruit orientation as a factor to improve classification accuracy. In this work, an average color of pepper was used. An accuracy of 82-90% was achieved.

Using the averaged pepper fruit color, this work improved the results of van Essen et al. (2022), whose algorithm is not efficient enough in all possible cases of photographing the pepper fruit, because it is necessary to select certain measurement points on the pepper fruit to increase the accuracy of classification by degree of ripeness.

In recognizing pepper ripeness, Cong et al. (2023), achieve an accuracy of 99% using complex computational procedures. The accuracy achieved in the present work is 10% lower, but it uses established computational methods that are suitable for implementation in embedded control systems for sorting machines and pepper harvesting robots.

Algorithms and procedures used in the present development have an advantage in terms of speed and ease of application in embedded control systems of sorting machines and pepper harvesting robots. The main limitation that can be pointed out is the lower classification accuracy compared to using data from hyperspectral images and processing them with neural networks. More research is needed to improve the classification accuracy of pepper depending on its degree of maturity. These analyzes should be aimed at using procedures and algorithms that can be used in computing devices that do not have high performance, but provide fast operation and sufficient classification accuracy. Also, further research should be aimed at generalizing the pepper classification task according to its degree of ripeness as mentioned by Lapegna et al. (2021).

Conclusion

Harvesting sweet pepper at the desired stage of maturity based on customer preference is essential for optimal economic returns. Overripe peppers are more likely to be damaged during transport and storage, resulting in a loss of quality and quantity. On the other hand, underripe peppers may not have developed their full nutritional potential, which can affect their marketability and price. Therefore, it is important to harvest sweet peppers at the correct stage of maturity to ensure their quality and maximize their economic value. In the present work, algorithms and procedures for determining the ripeness of red and yellow sweet peppers by jointly using their color and spectral characteristics have been compiled and tested. An advantage of these algorithmic procedures is that they do not require relatively large volumes of data and high-performance computing systems. This makes them suitable for application in embedded control systems for sorting machines and pepper harvesting robots. It was found that by a total of 11 informative color and spectral characteristics, the degree of ripeness can be determined for red and yellow sweet pepper. It has been proven that the feature vector can be reduced to two latent variables, which significantly reduces the amount of data used in classification. It was also found that the classification errors vary significantly, according to the type of pepper and the studied maturity—technical or biological. This is due to the difference in the absorption spectra of the different carotenoid groups involved in the color change of pepper fruits from green to red and yellow, respectively. Algorithms for pepper classification were adapted and studied, depending on five degrees of maturity as well as between the two main classes of technical and biological maturity.

The proposed tools are suitable for sorting sweet pepper into groups corresponding to its maturity. Complementing the results presented in the available literature, the methods and procedures proposed in this work can be used in the development of modern embedded control systems for sorting machines and robots for pepper harvesting. More research is needed to improve classification accuracy and generalization of the pepper classification task, depending on its degree of maturity.

References

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, 1-5.

Bell Pepper [online]. (n.d.). Retrieved March 11, 2024, from https://postharvest.ucdavis.edu

Cermakova, I., Komarkova, J., & Sedlak, P. (2019). Calculation of visible spectral indices from UAV-based data: Small water bodies monitoring. In 2019 *14th Iberian Conference on Information Systems and Technologies* (CISTI) (pp. 1-5).

Cong, P., Li, S., Zhou, J., Lv, K., & Feng, H. (2023). Research on instance segmentation algorithm of greenhouse sweet pepper detection based on improved Mask RCNN. *Agronomy*, 13, 196, 1-24.

Description of undersized pepper varieties [online]. (n.d.). Retrieved March 10, 2024, from https://tomathouse.com/4/nizkoroslye-sorta-perca.html (in Bulgarian)

Farrell, A., Wan, G., Rush, S., Martin, J., Belant, J., Butler, A., & Godwin, D. (2019). Machine learning of large-scale spatial distributions of wild turkeys with high-dimensional environmental data. *Ecology and Evolution*, 1-12.

Georgieva, Ts., Mihaylova, A., & Daskalov, P. (2020). Research of the possibilities for determination of some basic soil properties using image processing. In Proceedings of the *7th International Conference on Energy Efficiency and Agricultural Engineering* (EE&AE) (pp. 1-4).

Harel, B., Kurtser, P., Parmet, Y., & Edan, Y. (2017). Sweet pepper maturity evaluation. Advances in Animal Biosciences: *Precision Agriculture*, 8(2), 167-171.

Harel, B., Parmet, Y., & Edan, Y. (2020a). Maturity classification of sweet peppers using image datasets acquired in different times. *Computers in Industry*, 121, 103274, 1-10.

Harel, B., van Essen, R., Parmet, Y., & Edan, Y. (2020b). Viewpoint analysis for maturity classification of sweet peppers. *Sensors*, 20, 3783, 1-22.

Kirilova, E. (2012). Classifier design for identification of corn kernels, damaged by Fusarium moniliforme, using color features. In Proceedings of *University of Ruse, Ruse, Bulgaria*, 2012, 51(3.1), 168-175. (in Bulgarian)

Kootstra, G., Wang, X., Blok, P. M., Hemming, J., & van Henten, E. (2021). Selective harvesting robotics: Current research, trends, and future directions. Current Robotics Reports, 2(2), 95–104.

Lapegna, M., Balzano, W., Meyer, N., & Romano, D. (2021). Clustering algorithms on low-power and high-performance devices for edge computing environments. *Sensors*, 21(16), 5395.

Lehnert, C., McCool, C., Sa, I., & Perez, T. (2020). Performance improvements of a sweet pepper harvesting robot in protected cropping environments. *Journal of Field Robotics*, 37(7), 1197–1223.

Liu, C., & Niu, S. (2024). Automated fruit sorting in smart agriculture system: Analysis of deep learning-based algorithms. *International Journal of Advanced Computer Science and Applications* (IJACSA), 15(1).

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O., & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food Bioprocess Technology*, 5, 1121-1142.

Pathare, P., Opara, U., & Al-Said, F. (2013). Colour measurement and analysis in fresh and processed foods: A review. *Food Bioprocess Technology*, 6, 36-60.

Rizzo, M., Marcuzzo, M., Zangari, A., Gasparetto, A., & Albarelli, A. (2023). Fruit ripeness classification: A survey. *Artificial Intelligence in Agriculture*, 7, 44-57.

Smits, B. (2000). An RGB to spectrum conversion for reflectances. University of Utah, 21 January 2000, 1-10.

van Essen, R., Harel, B., Kootstra, G., & Edan, Y. (2022). Dynamic viewpoint selection for sweet pepper maturity classification using online economic decisions. *Applied Sciences*, 12, 4414, 1-15.

Villaseñor-Aguilar, M.-J., Bravo-Sánchez, M.-G., Padilla-Medina, J.-A., Vázquez-Vera, J. L., Guevara-González, R.-G., García-Rodríguez, F.-J., & Barranco-Gutiérrez, A.-I. (2020). A maturity estimation of bell pepper (*Capsicum annuum* L.) by artificial vision system for quality control. *Applied Sciences*, 10(15), 5097.

Wyman, C., Sloan, P.-P., & Shirley, P. (2013). Simple analytic approximations to the CIE XYZ color matching functions. *Journal of Computer Graphics Techniques* (JCGT), 2(2), 1-11.

Yotova, G., Tsitouridou, R., Tsakovski, S., & Simeonov, V. (2016). Urban air quality assessment using monitoring data of fractionized aerosol samples, chemometrics and meteorological conditions. *Journal of Environmental Science and Health*, Part A, 51(7), 544-552.

Acknowledgement

This work was administrative and technically supported by the Bulgarian national program “Development of scientific research and innovation at Trakia University in the service of health and sustainable well-being”—BG-RRP-2.004-006-C02.