NON-DESTRUCTIVE ASSESSMENT OF EXTERNAL AND INTERNAL QUALITY ATTRIBUTES OF FRUITS USING COMPUTER VISION AND VIS–NIR SPECTROSCOPY

Qobilov H.Kh.
PhD, associate professor of Bukhara state technical university
Ogomurodov U.H.
PhD student of Bukhara state technical university
E-mail: ogomurodov@gmail.com

Annotation: This article covers the issue of increasing efficiency through the integration of computer vision and near-infrared (NIR) technologies in the automatic sorting of fruits and vegetables by quality. This approach compared to traditional visual evaluation methods allows quick determination external and internal quality indicators of a product, in a non-contact and reliable manner. It also looks at ways to further optimize and apply these systems using machine learning (ML) techniques.

Keywords: computer vision, near infrared (NIR), automatic sorting, fruit and vegetable quality, machine learning, food safety, spectral analysis, PCA.

Аннотация. В данной статье рассматривается вопрос повышения эффективности за счёт интеграции технологий компьютерного зрения и ближней инфракрасной (NIR) спектроскопии в автоматическую сортировку фруктов и овощей по качеству. По сравнению с традиционными методами визуальной оценки, данный подход позволяет быстро, бесконтактно и надёжно определять как внешние, так и внутренние показатели качества продукции. Также в статье рассматриваются пути дальнейшей оптимизации и применения подобных систем с использованием методов машинного обучения.

Ключевые слова: компьютерное зрение, ближний инфракрасный диапазон (NIR), автоматическая сортировка, качество фруктов и овощей, машинное обучение, безопасность пищевых продуктов, спектральный анализ, PCA.

Annotatsiya. Ushbu maqolada meva va sabzavotlarni sifat bo‘yicha avtomatik saralashda kompyuter ko‘rishi va yaqin infraqizil (NIR) texnologiyalarini integratsiya qilish orqali samaradorlikni oshirish masalasi ko‘rib chiqilgan. An’anaviy vizual baholash usullariga nisbatan ushbu yondashuv mahsulotning tashqi va ichki sifat ko‘rsatkichlarini tez, kontaktlarsiz va ishonchli tarzda aniqlash imkonini beradi. Shuningdek, maqolada ushbu tizimlarni yanada optimallashtirish va qo‘llash usullari, shu jumladan mashina o‘rganish (ML) texnikalaridan foydalanish yo‘llari ham ko‘rib chiqilgan.

ISSN 3030-3907

Kalit so‘zlar: kompyuter ko‘rishi, yaqin infraqizil (NIR), avtomatik saralash, meva va sabzavot sifati, mashina o‘rganish, oziq-ovqat xavfsizligi, spektral tahlil, PCA.

Introduction.

Today, with increasing production of agricultural products, he need for advanced technologies that meet all requirements of storage, processing and quality delivery to consumers is increasing. In particular, non-contact, automated approaches instead of traditional methods for export-grade and delivery of fruits and vegetables are becoming more important. This article analyzes the possibilities of creating an effective sorting system by combining computer vision and NIR technologies.

Determination methods for assessing the quality of the fruit. There are several sorting methods available to evaluate fruit quality, such as machine imaging, infrared and near-infrared (VIS-NIR) spectroscopy, hyperspectral imaging, impact analysis, ultrasound, X-ray, and magnetic resonance imaging [1].

Machine vision techniques aim to automatically assess the external quality characteristics of fruits, such as color, size, shape, and skin defects, by analyzing images of the fruit surface by a digital color camera and scattered light sources. Figure 1 illustrates a typical view of machine vision techniques.

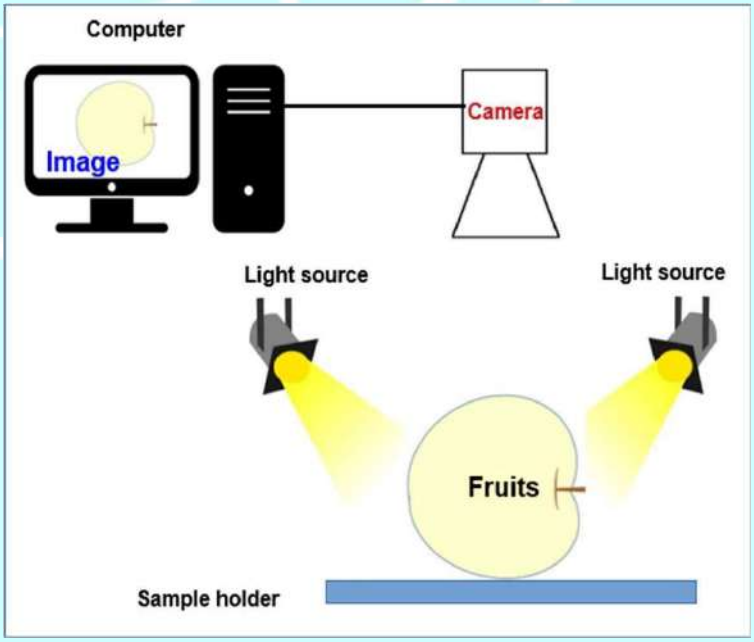


Figure 1. Typical installation of machine vision equipment Images captured by a color camera using machine vision

It is usually sent to a computer for processing and analysis to obtain and evaluate the appropriate properties of the fruit. Machine vision techniques are suitable for real-time application of fruit exterior quality assessment, as it is a fast, non-perishable, reliable, and cost-effective method [3]. Nowadays, machine vision techniques have been developed and in practice, fruit sorting is widely

ISSN 3030-3907

applied for commercial purposes. However, this method is limited only to an external quality assessment because it uses a color camera to photograph the task.

The VIS-NIR spectroscopy method uses a spectrometer that covers a wide wavelength range in the infrared and near-infrared ranges to collect spectral data of fruits. When a light source strikes the surface of the fruits, it can be reflected, absorbed or transmitted by the fruit at each wavelength depending on the chemical and physical properties of the fruits.

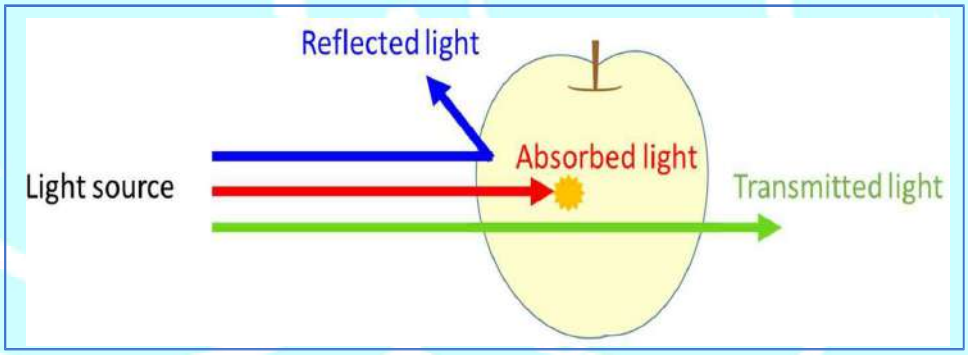


Figure 2. Interaction of light source and fruits

Due to the complex interaction between the light source and the fruit, several statistical and other analyses, sometimes called chemometry, are performed to extract information related to the intrinsic quality characteristics of the fruit from the collected spectral data [1,2]. A demonstration of a typical VIS-NIR or NIR spectroscopy setup for reflected light capture is illustrated in Figure 3. The VIS-NIR spectroscopy technique has great advantages in terms of rapid evaluation of internal properties due to the fact that no chemical reactions have occurred on fruit samples. In addition, this method can evaluate several internal characteristics at the same time. However, VIS-NIR spectroscopy is a point-sensing method, and spectral data are collected only in a small fraction of fruits [2].

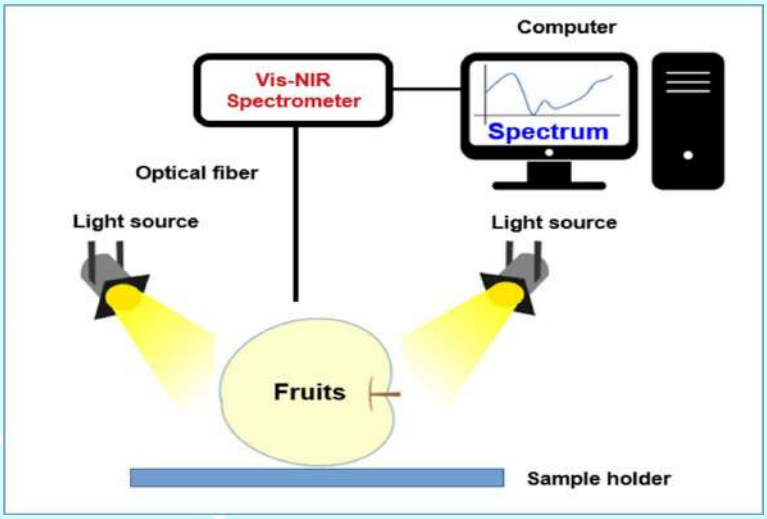


Figure 3. Typical Setup of VIS-NIR Spectroscopy Technique

Hyperspectral imaging is a spectroscopy technique to collect images of fruits at several to several hundred different wavelengths at a time. In this method, a

ISSN 3030-3907

hyperspectral camera is used to collect spatial and spectral data called a hypercube of two spatial dimensions and one spectral dimension. Therefore, the chemical components and exterior properties of the fruit can be easily extracted from the hypercube [3]. Figure 4 shows the hyperspectral imaging technique and a sample setup of hypercube data. Hyperspectral imaging, which has a large amount of spatial and spectral data, is a powerful tool for evaluating the external and internal quality of fruits. Pathmanaban et al. v. [4] Considers some applications of hyperspectral vision in fruit quality evaluation. However, low-speed data acquisition and high cost are the main drawbacks of hyperspectral imaging techniques for applications.

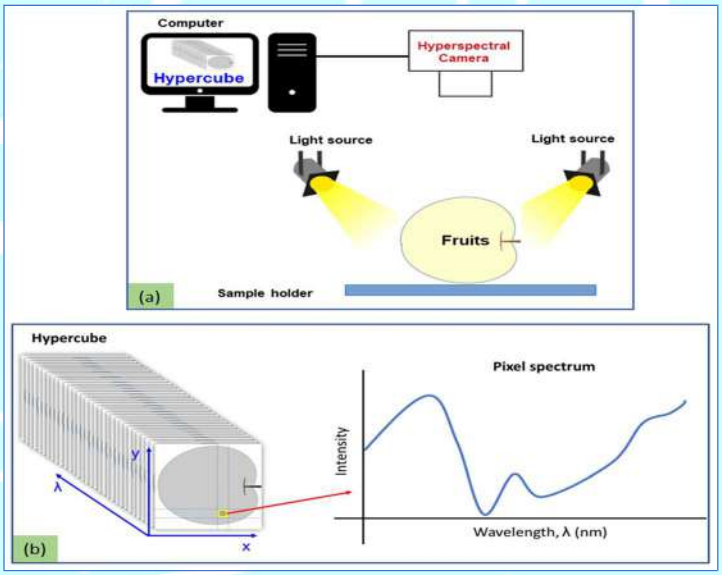


Figure 4. Typical hyperspectral imaging is the setting (a) and the hypercube data (b).

Impact analysis, as described in Figure 5, is also used to assess the consistency or intrinsic damage of fruits based on Hertz's communication theory [1]. A spherical impactor is made of a material that has permanent physical properties. A small vibration or acceleration sensor is mounted on the impactor to measure maximum deformation, maximum force, and contact time between the fruit and the impactor. This data corresponds to the mechanical properties of the fruits. This method is being applied practically for the evaluation of consistency in fruits as diverse as peaches, tomatoes, apples, kiwifruit and mangoes [1,5]. However, this method is not suitable for fruits with low hardness, since the fruits can be permanently deformed by external forces from the action.

ISSN 3030-3907

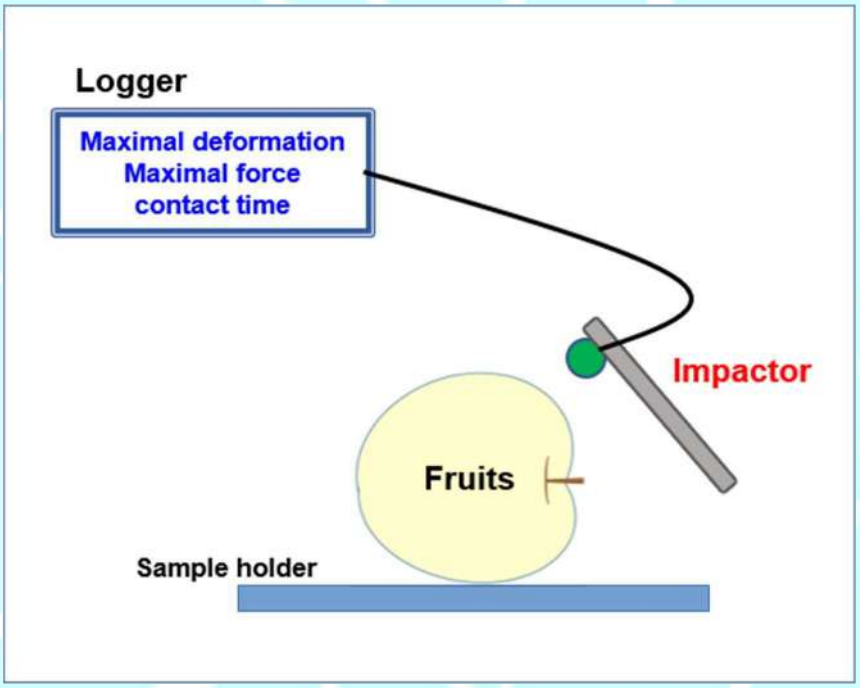


Figure 5. Impact analysis method

Ultrasound is a schematic diagram of a sound wave with a frequency of 20 kHz considered as a quality determination technique without degradation for evaluating fruit quality[1]. Therefore, this method was used to measure the internal quality of fruits, mainly their consistency. Application of ultrasonic techniques to evaluate the intrinsic properties of fresh fruits and vegetables is shown by Mizrach [6] and Awad et al. (2015) comprehensively presented.

X-ray technique is widely used in medical diagnosis. The basic principle of X-ray applications is based on the absorption of X-ray photons by the material, that is, the X-rays depend on local density, atomic number, and energy [1]. Therefore, if an object has multiple characteristics with different absorption properties, these characteristics can be distinguished in X-ray images. As for the evaluation of the quality of the fruit, this method is mainly used to detect high contrast unknown substances in the fruits or to assess the extent of damage and moisture content on the inside. Some applications and potentials of X-ray techniques are detailed in Haff et al. [7].

Magnetic resonance imaging (MRI) techniques show electromagnetic radiation in the radio frequency range with an object based on interaction. By the application of a strong magnetic field of this technique, the transformation of the protons (^{1H} nuclei) present in the fruit juice is observed and information is obtained regarding the spatial distribution of proton density, relaxation parameters, and self-diffusion parameters within the fruit [1]. This information can be used to provide detailed information about the internal structure of the fruit [8]. Thus, previous studies have successfully applied this method to measure the intrinsic damage of several types of fruits, such as watermelons,

ISSN 3030-3907

apples, peaches, pears, and citrus fruits [1,4,8].

Table 1 shows some of the key characteristics of the main non-perishable methods for evaluating fruit quality [1]. These non-disruptive methods include machine vision, VIS-NIR spectroscopy, impact analysis, and hyperspectral imaging techniques that are commercialized and applied in practical applications. On the contrary, ultrasound, X-ray, magnetic resonance imaging techniques are still at the investigation stage in the laboratory, because they are very expensive and have a low frequency.

Table-1.

Comparison of the main destructive methods of evaluating the quality of the fruit.

Element		Machine View	VIS-NIR Spectroscopy	Impact analysis	yperspectral View	Ultrasonography	X-ray	MRI
Quality attributes	Tint	O			O			
	Size, shape	O			O		O	O
	Surface defect	O			O			
	ɔntent of soluble solids (Sugar content)		O		O			
	Maurity		O		O			
	Hardness		O	O	O	O		
	Internal damage		O	O	O		O	O
	Starch		O		O	O		
	Moisture content		O		O		O	O
Usage	[Pg 100]	O	O	O	O	O	O	O
	The industrial system	O	O	O				
	explosive equipment	O	O					
Speed (*fruit/sec*)		10	6	10	0.5	0.2	0.1	0.1
Price (*US$ 1000*)		1	5	1	10	100	200	200

In this study, the sweetness of the fruits (soluble solid composition) is

ISSN 3030-3907

mainly a characteristic. Among the non-disruptive methods summarized in Table 1 are VIS-NIR spectroscopy and hyperspectral imaging potential

While hyperspectral imaging is the most powerful method to capture all the external and internal characteristics of fruits, this method is a high cost and low speed compared to spectroscopy techniques. It is not suitable for applications in this area. Therefore, the VIS-NIR spectroscopy technique was chosen to evaluate the intrinsic quality of the fruits in this study.

Evaluation of the level of ripeness (sweetness) of the fruits with VIS-NIR

Spectral data in the VIS-NIR range include information on C-H, O-H, N-H, C-S, and S-S vibrational absorption, which are mainly associated with major compounds in fruits, such as water, sugar, chlorophyll, carotenoids, and anthocyanins [1,9,10]. Absorption bands of fruit pigments such as chlorophyll, carotenoids, and anthocyanins in the visible region occurred between 460 nm and 680 nm, respectively, as shown in Figure 6 [11]. Figure 7 shows the most visible absorption bands in the near-infrared region [10]. Therefore, the VIS-NIR spectroscopy method can be used to measure many intrinsic properties of fruits.

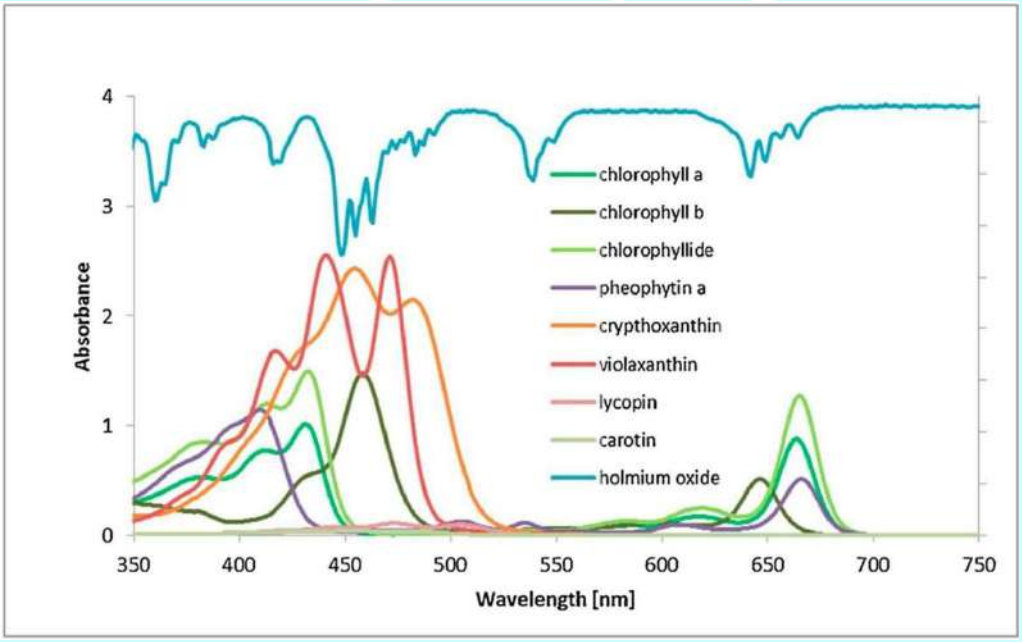


Figure 6. Absorption Spectra of Pigments in Acetone

ISSN 3030-3907

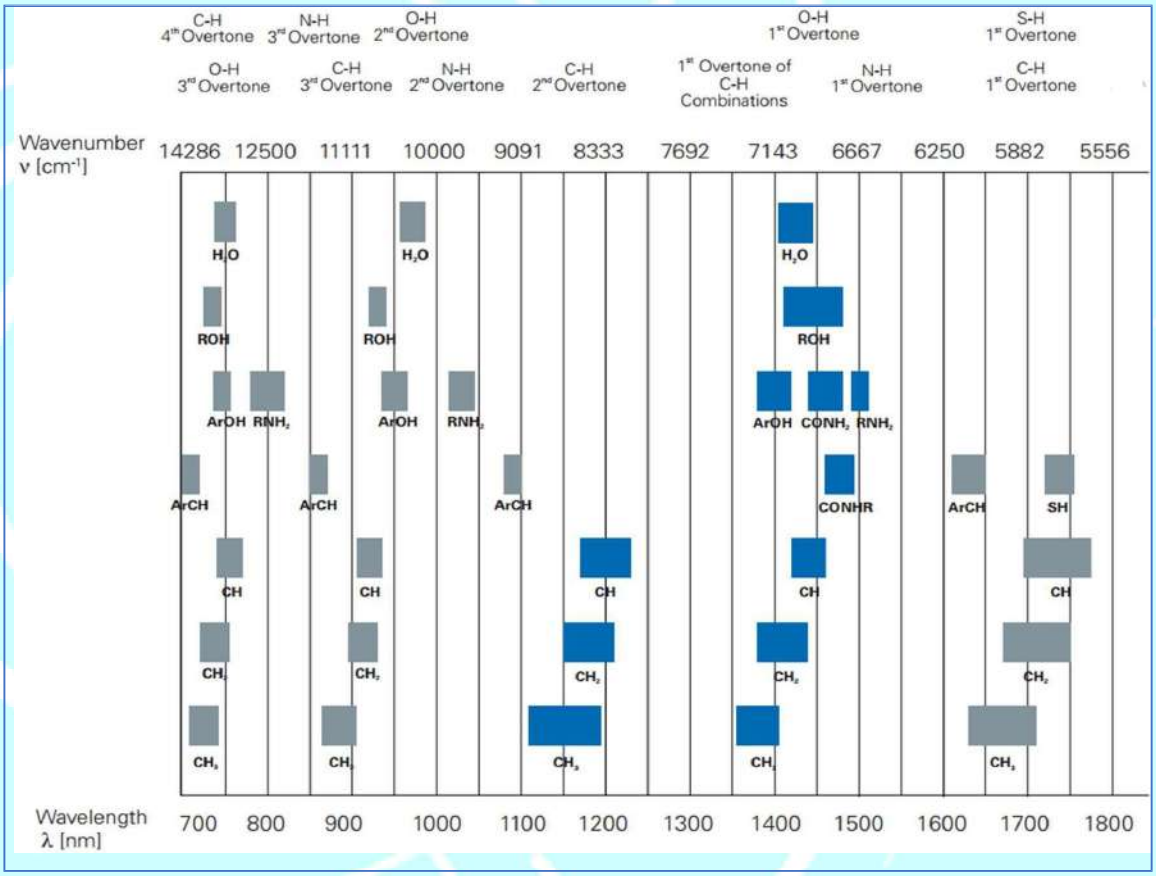


Figure 7. NIR Absorption Networks (Brooker Optics Inc., MA, USA)

In previous literature, many previous studies have proven that VIS- NIR spectroscopy has been used in both on-line and off-line applications to assess the quality of many fruits such as mangoes, apples, pears, and melons. [3] and Wang et al. (2006). [2] listed numerous previous studies and commercial systems using this technique. However, the practical introduction of this method in the field has not been widely used due to the lack of consistency with the change in working conditions. In addition, this method is still not convenient for farmers or small producers, since a skilled operator is required to process spectral data and create prediction models, and the cost of portable devices is relatively high for them.

In order to support farmers or small-scale producers and encourage the widespread use of this technique in the field, simplified spectrometric systems have been developed by several previous studies using filters or light-emitting diodes (LEDs) instead of traditional spectrometers. [5] They created a multi-band sensor system to assess the maturity of the oil palm orchard.

Four sets of discrete search-detector work in parallel to collect reflection spectra with an operating distance of 60 cm. [6] tested the potential of an LED-based prototype to quickly evaluate the ripeness of white grapes. The four channels consisted of four LEDs, four photodiodes, and optical fibers. Jam and Chia [8] have developed a five-band portable sensor to classify pineapple quality

ISSN 3030-3907

based on sweetness. A light source with a photodiode and five near-infrared LEDs is connected to a special pentagonal fixture.

These studies have shown the possibility of filter or LED-based devices for the evaluation of fruit quality in a simple and cost-effective way. However, the production and reproduction of this research system is still limited, since individual devices are required for such customized fittings or optical parts and all discrete components. In addition, these studies focused mainly on classification tasks, with quantitative prediction tasks not fully mentioned.

Regression Models.

For quantitative prediction of fruit sweetness, a correlation was established between sweetness (Brix values) and spectral data of fruits with regression models.

Partial least squared regression (PLSR) is the most common analysis method used in spectral analysis to generate a regression model. In the PLSR analysis, the actual variables are translated into new orthogonal variables (hidden variables). These new variables are independent of each other. Therefore, this method is particularly appropriate when the number of wave networks is greater than the number of samples and can work efficiently with data that include multicollicity [12].

Multilinear regression (MLR) is another common method in which original variables are used to construct a regression model. This method is simple and easy to understand the contribution of individual variables to the predictive outcome. However, multilinear regression (MLR) is not appropriate if the number of wave networks is greater than the number of samples and tends to multicolliarity [13]. Therefore, effective wave plots should be selected before constructing a multicollinearity regression (MLR) model.

In this study, PLSR was used for the first time to evaluate and compare the performance of regression models with AS7265x data and USB2000+ data, and to identify potential wave networks to create a simple model with MLR. Several efficient wave networks were then selected to create an optimal MLR model with AS7265x data.

Compared in terms of performance of the PLSR and MLR models, both the calibration model (R2, RMSEC) and the prediction model (R2, RMSEP) were determined using the detection coefficient (R2) and the root mean square error (RMSE).

Equations (1), (2), (3), and (4).

$$R^2c = 1 - (\Sigma^{nc} i_{=1} (yi - \hat{y}i)^2) / (\Sigma c^n i_{=1} (yi - \bar{y}i)^2) \quad (1)$$

$$RMSEC = (2)\sqrt{\frac{1}{n}\sum_{i=1}^{n_c}(y_i - \hat{y}_i)^2}$$

ISSN 3030-3907

$$R^2p = 1 - (\Sigma np_{k=1}(yk - \hat{y}_k)^2) / (\Sigma^{np}_{k=1}(yk - \bar{y}_k)^2) \quad (3)$$

$$RMSEC = \sqrt{\frac{1}{n}\sum_{k=1}^{n_c}(yk - \hat{y}k)^2} \quad (4)$$

Where ŷi, ŷk denotes the predicted Brix value of the i sample in the calibration set and the k-th sample in the prediction set; yi, y, is the measured Brix values of the i sample in the k calibration set and the k-th sample in the prediction set; n_c, the number of samples in the np calibration set and the prediction set. In general, a better model should have higher values of R^2_c and higher values of R_{p2} and lower values of RMSEC and RMSEP.

Classification Models

To evaluate apples based on their sweetness, several well-known classification algorithms, namely k-nearest neighbors (k-NN), discriminant analysis (DA), support vector machine (SVM), and random forest (RF), were applied with different hyperparameters to find a suitable algorithm for this purpose. The basic idea behind these algorithms is summarized briefly as follows [14]:

- k-NN is a supervised learning algorithm that classifies a test data point based on the similarity of k to the closest points in the training data. In this algorithm, there is no need to build the model or adjust any parameters. For the classification of a test point, the distance between this point and all training points is measured. Then k training points with the smallest distance (the closest) are taken. Finally, among these k points, the most common label is given to the test point.
- In the DA algorithm, it is assumed that the instructional data within each classroom is derived from the Gaussian distribution. The DA makes predictions by estimating the probability that new test data will belong to each class using Bayes' theorem. The new test data is then assigned to the class with the highest probability.
- It is assumed that it is possible to construct a hyperplane in SVM that perfectly distributes instructional data into different classes. The distance between the hyperplane and the nearest data points is called the boundary. These points are called support vectors. SVM tries to find the best hyperplane that can distinguish classes by the largest margin. After that, new test data is assigned to the class, depending on which side of the hyperplane it is on.

The performance of the classification models was evaluated by cross-validation method. Based on the results of the cross-examination classification, the predictive effectiveness was evaluated in terms of accuracy and uncertainty calculated based on the confusion matrix. The confusion matrix for the two-class classification models has four states: True Positive (TP), False Positive (FP),

ISSN 3030-3907

False Negative (FN), and True Negative (TN), as shown in Figure 8.

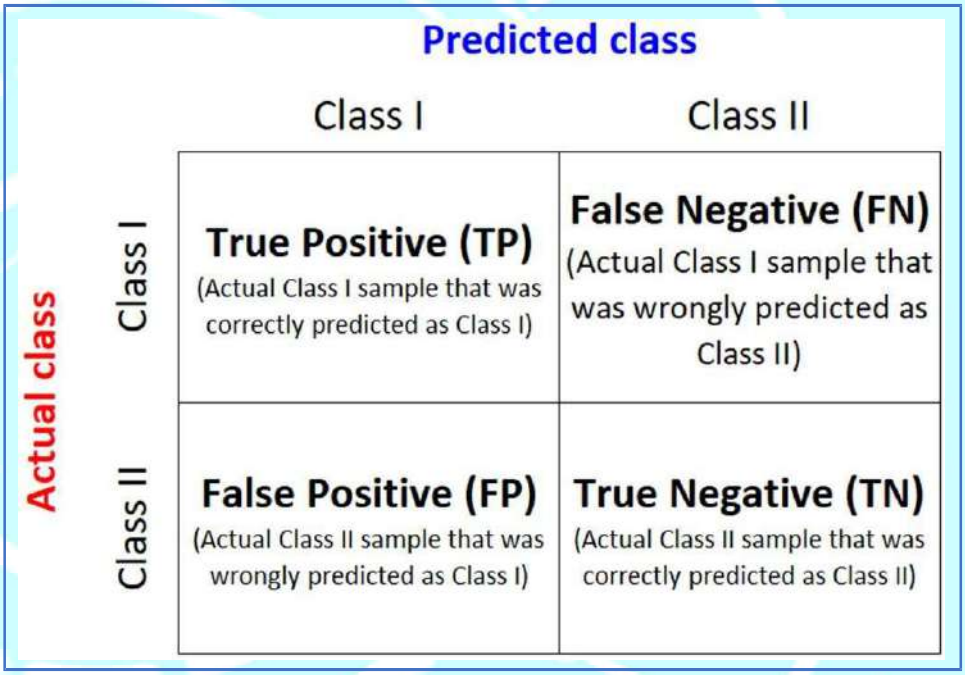


Figure 8. Confusion matrix of classification models

Based on the results of the cross-validation classification, the prediction performance of the models was evaluated in terms of accuracy and accuracy according to equations (5) and (6).

$$\text{Precision} = \frac{TP}{TP+FP+TNF}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

Accuracy refers to the total number of correctly classified samples for all samples. This metric is often helpful when the importance of all four cases is equal. However, in the case of fruit classification, the accuracy criterion has been used in conjunction with the accuracy criterion to select the best classifier, because the "real" class (high sweetness) is more important than "fake" to satisfy customer need.

Conclusion. This study proposes a method for estimating the external (color, shape, defect) and internal (sweetness, moisture, hardness) quality parameters of fruits and vegetables **in an accurate and non-contact manner by integrating** computer vision (CV) **and** near-infrared (NIR) technologies.**computer imaging quickly** evaluates external properties (with 90% accuracy), while NIR quickly assesses internal quality indicators (soluble solids, moisture). This significantly improves the sorting speed.In the prediction of fruit sweetness (Brix) using PLSR and MLR recomposition models, an RMSE $\leq$ of 0.5 and an $R^2 \geq 0.92$ help to predict the outcome with high accuracy.

Proposals: Integrating the system with IoT in the future and analyzing the data on a cloud platform; By mounting the NIR sensor on drones or robots, it will

ISSN 3030-3907

be possible to establish quality control in the field, and this case study will create the scientific basis for automated quality control in the fruit and vegetable industry.

REFERENCES

1. Nicolaï, B.M.; Defraeye, T.; De Ketelaere, B.; Herremans, E.; Hertog, M.L.A.T.M.; Saeys, W.; Torricelli, A.; Vandendriessche, T.; Verboven, P.; Ketelaere, B. De; et al. Nondestructive Measurement of Fruit and Vegetable Quality. *Annu. Rev. Food Sci. Technol.* **2014**, *5*, 285–312, doi:10.1146/annurev-food-030713-092410.
2. Wang, H.; Peng, J.; Xie, C.; Bao, Y.; He, Y. Fruit Quality Evaluation Using Spectroscopy Technology: A Review. *Sensors* **2015**, *15*, 11889–11927, doi:10.3390/s150511889.
3. Lu, Y.; Lu, R. Quality Evaluation of Apples. In *Computer Vision Technology for Food Quality Evaluation*; Da-Wen Sun, Ed.; Elsevier, 2016;
4. Pathmanaban, P.; Gnanavel, B.K.; Anandan, S.S. Recent application of imaging techniques for fruit quality assessment. *Trends Food Sci. Technol.* **2019**, *94*, 32–42, doi:10.1016/j.tifs.2019.10.004.
5. Homer, I.; García-Ramos, F.J.; Ortiz-Cañavate, J.; Ruiz-Altisent, M. Evaluation of a Non-Destructive Impact Sensor to Determine On-Line Fruit Firmness. *Chil. J. Agric. Res.* **2010**, *70*, 67–74, doi:10.4067/S0718-58392010000100007.
6. Mizrach, A. Ultrasonic technology for quality evaluation of fresh fruit and vegetables in pre- and post-harvest processes. *Postharvest Biol. Technol.* 2008, *48*, 315–330.
7. Haff, R.P.; Toyofuku, N. X-ray detection of defects and contaminants in the food industry. *Sens. Instrum. Food Qual. Pure.* 2008, *2*, 262–273.
8. Srivastava, R.K.; Talluri, S.; Beebi, S.K.; Rajesh Kumar, B. Magnetic Resonance Imaging for Quality Evaluation of Fruits: a Review. *Food Anal. Methods* **2018**, *11*, 2943–2960, doi:10.1007/s12161-018-1262-6.
9. Walsh, K.; Golic, M.; Lawson, P. Short-Wavelength Near-Infrared Spectra of Sucrose, Glucose, and Fructose with Respect to Sugar Concentration and Temperature. *Appl. Spectrosc. Vol. 57, Issue 2, pp. 139-145* **2003**, *57*, 139–145.
10. BrukerOptics *Guide for Infrared Spectroscopy*; 2009;
11. Walsh, K.B.; Blasco, J.; Zude-Sasse, M.; Sun, X. Visible-NIR 'point' spectroscopy in post-harvest fruit and vegetable assessment: The science behind three decades of commercial use. *Postharvest Biol. Technol.* **2020**, *168*, 111246, doi:10.1016/j.postharvbio.2020.111246.
12. Irfan, M.; Javed, M.; Raza, M.A. Comparison of shrinkage regression methods for remedy of multicollinearity problem. *Middle East J. Sci. Res.* **2013**, *14*, 570–

ISSN 3030-3907

579, doi:10.5829/idosi.mejsr.2013.14.4.488.
13. Naes, T.; Mevik, B.-H. Understanding the collinearity problem in regression and discriminant analysis. *J. Chemom.* **2001**, *15*, 413–426, doi:10.1002/cem.676.
14. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer Texts in Statistics; Springer New York: New York, NY, 2013; Vol. 103; ISBN 978-1-4614-7137-0.

ISSN 3030-3907