UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Computer Science
Data Science Curriculum

Karina Rao

Produce Quality and Pesticide Residue Estimation Using Light Sensing

Master's Thesis (15 ECTS)

Supervisor: Huber Flores, PhD

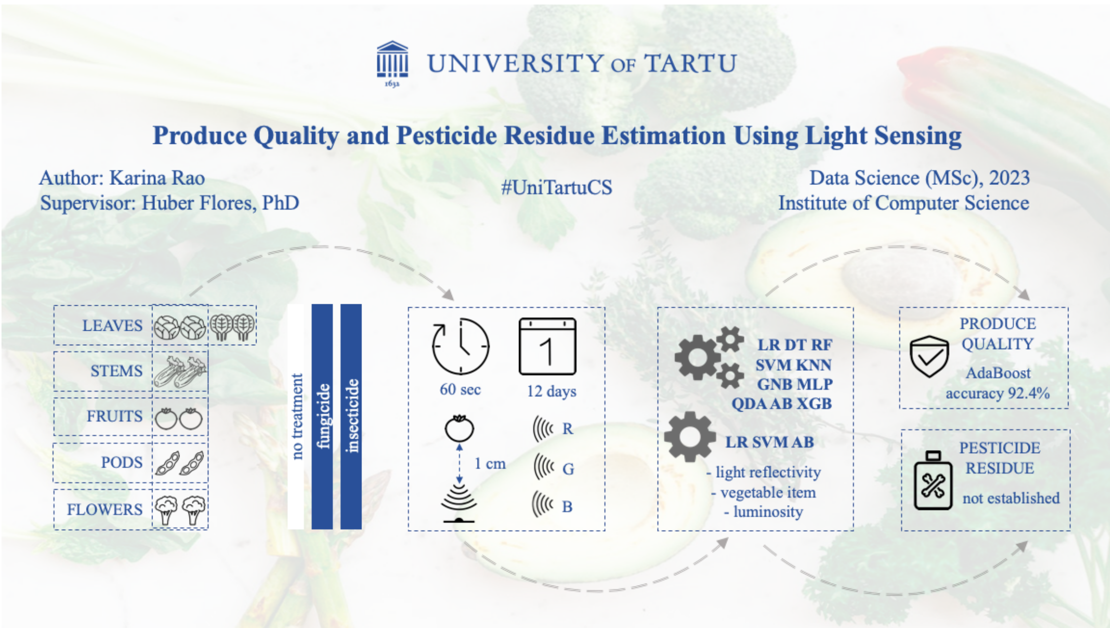
Tartu 2023

Produce Quality and Pesticide Residue Estimation Using Light Sensing

Abstract:

While produce quality estimation across various stages in the value chain is essential to tackle food loss and waste, determining pesticide residue in fresh produce can alleviate the threat to human health and the environment. Light sensing offers a non-invasive and cost-effective method to establish unique fingerprints for fresh produce. During a 12-day produce decomposition period, it was established that light reflectivity is effective for the quality estimation of vegetables. The AdaBoost classification model with blue light reflectivity value, vegetable items and luminosity as input features achieved a performance accuracy of 92.4%. While measuring reflectivity intensity, it is important to account for varying lighting conditions (luminosity). Notwithstanding the success of predicting the quality of fresh produce, light sensing failed in pesticide residue estimation.

Graphical abstract:



Keywords:

Light sensing, produce quality, pesticide residue, machine learning

CERCS:

P170 Computer science, numerical analysis, systems, control

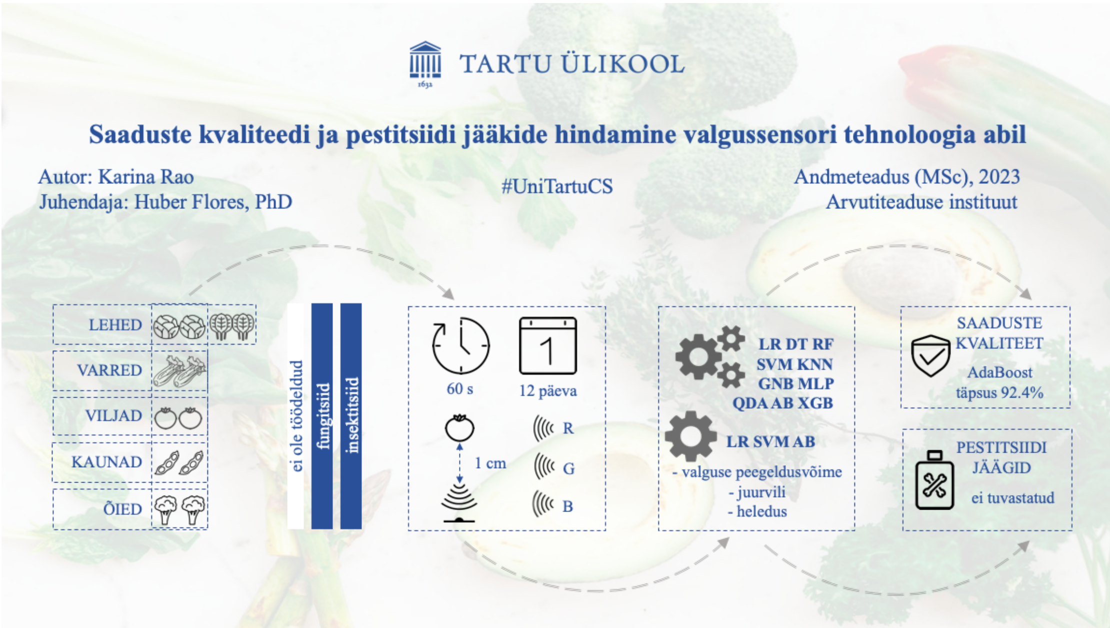
P176 Artificial intelligence

Saaduste kvaliteedi ja pestitsiidi jääkide hindamine valgussensori tehnoloogia abil

Lühikokkuvõte:

Kui toodangu kvaliteedi hindamine väärtusahela eri etappides on oluline toidukao ja raiskamise vähendamiseks, siis pestitsiidide jääkide kindlaks määramine saadustes leevendab ohtu inimeste tervisele ja keskkonnale. Valgussensori tehnoloogia võimaldab mitteinvasiivselt ja kulutõhusalt tuvastada värskete saaduste ainulaadseid sõrmejälgi. Uurimistöö 12-päevase saaduste lagunemisperioodi jooksul tehti kindlaks, et valguse peegeldusvõime on tõhus viis juurviljade kvaliteedi hindamisel. AdaBoost klassifikatsioonimudel, mille sisendparameetrid olid sinise valguse peegeldusvõime väärtus, juurviljade tüüp ja heleduse tase, saavutas täpsuse 92.4%. Peegeldusvõime intensiivsuse mõõtmisel on oluline arvestada erinevate valgustingimustega (heleduse tase). Hoolimata edust ennustada värskete saaduste kvaliteeti, valgussensori tehnoloogia abil ei õnnestunud hinnata pestitsiidide jääke.

Visuaalne kokkuvõte:



Võtmesõnad:

Valgussensori tehnoloogia, toodangu kvaliteet, pestitsiidide jäägid, masinõpe

CERCS:

P170 Arvutiteadus, arvutusmeetodid, süsteemid, juhtimine

P176 Tehisintellekt

Contents

1 Introduction

Food quality and safety are central issues in food economics [20] that get attention from different stakeholders throughout the agriculture and food production supply chain. The advances in computing technologies are benefiting food quality and safety detection and monitoring. These emerging technologies are helping farmers, producers and retailers to improve efficiency and respond to consumer demands. Besides direct economic profitability, indirect and second-order benefits are valuable to the market players and society at large. Eradication of hunger and poverty, clean water, sustainable land use, responsible production and consumption, mitigating climate change, and sustainable life on land and water are the United Nations sustainable development goals (SDG) [31] that can be linked to food quality and safety.

Computer vision technology has been extensively deployed for quality estimation of fresh produce, such as fruits and vegetables. The first computer vision systems for produce quality assessments examined surface characteristics to detect visual defects or estimate ripeness [10]. Machine learning techniques are often used in conjunction with computer vision. Other computing technologies for quality estimation make use of thermal imaging and light sensing [19, 49]. The safety of fresh produce has been investigated less by advanced computational methods. This thesis examines the application of light sensing for both produce quality and pesticide residue estimation.

Quality estimation encompasses various characteristics (e.g. colour, texture, shape and sugar content) and refers to the ripeness stage of fresh produce. The importance of quality estimation consists in food loss and waste prevention. Every year around one-third of all food produced for human consumption is lost and wasted across the entire supply chain [14]. This has serious implications for food security, especially in the context of a growing population and limited resources. Nearly every third person does not have access to adequate food and about 12% of the global population is severely food insecure [15]. Quality assessment solutions can limit food loss and waste throughout the supply chain, consequently causing less environmental harm as fresh produce often involves long transportation and cold storage.

Pesticide residue estimation is concerned with determining the levels of substances in produce resulting from the use of pesticides. Pesticides constitute the second largest group of man-made chemicals after fertilizers [16] and pose a threat to human health and the environment. The sustainable use of pesticides is advocated by the European Commission in its communication on the Chemicals Strategy for Sustainability Towards a Toxic-Free Environment that is part of the European Green Deal. The communication underlines societal concern over the use of pesticides: “84% of Europeans are worried about the impact of chemicals present in everyday products on their health, and 90% are worried about their impact on the environment” [4]. New techniques and methodologies for detecting pesticide residue can invoke stronger control over pesticide usage.

Zuniga *el al* demonstrate the use of reflective green light sensing as an innovative low-cost approach for quality estimation of fresh produce [49]. The authors validate their method through empirical benchmarks showing it can establish unique fingerprints for different produce and estimate the quality or ripeness. This is established via correlation between the changes in the green light values and the so-called transpiration coefficients of produce. Although the study covers a diverse range of fruit categories, it is essential to examine the same effect on vegetables to generalize the method. Moreover, a wider spectrum of light sources other than red light is worth considering.

Therefore first, this thesis aims to corroborate the work of light sensing for quality estimation of produce. Red, green and blue light reflectivity is used to establish fingerprints instead of merely green light. All three light sources are evaluated for their fitness. Several categories of vegetables are examined to demonstrate the generalisability of light sensing for produce quality estimation. Second, since there is non-existing research about light sensing for pesticide residue estimation, the goal is to explore the possibility of light reflectivity for assessing pesticide residues in fresh produce. Fungicide and insecticide categories of pesticides are investigated. Both objectives are achieved via a single experimental setup.

2 State of the Art

This chapter reviews the literature with respect to the core topics of the thesis, presenting the current state to date for produce quality and pesticide residue estimation. Produce quality estimation covers visual inspection, odor and magnetic sensing, soft X-ray and thermal imaging, light sensing and a number of other methods. For pesticide residue estimation, the two main routines are chromatography and biosensors, that is supplemented by a few other methods.

2.1 Produce quality estimation

2.1.1 Visual inspection

The most common method for quality estimation is manual or automated visual inspection. Manual inspection by human graders is time-consuming and labor-intensive. Moreover, it lacks accuracy due to human error and subjective evaluations [48].

In an attempt to overcome the shortcomings of manual examination, computer vision technology was introduced to inspect and grade food products in the late 1980s. Computer vision systems provide rapid, economic, hygienic, consistent and objective quality assessment [18]. The first automated systems for produce quality assessments were detecting visual defects or estimating ripeness by examining surface characteristics, e.g. texture coarseness [3]. Color, texture, size, shape and defects are common features inspected by traditional computer vision systems (TCVS) where RGB color cameras acquire images.

Multispectral and hyperspectral computer vision systems surpass TCVS in a few defects that are challenging to detect with TCVS due to the dominance of spectral images [3]. The shortcoming of capturing the images of fruits and vegetables from one direction can be overcome by modeling the fruit shape as a 3D spheroid to obtain a global score of the fruit. This method matches the defects between adjacent views to prevent counting them more than once and examines the whole surface [1]. While severely injured fruits and vegetables are easily identified visually, hidden internal physical damage caused by mechanical injury is difficult to detect. Therefore, visual methods are limited to external quality factors.

2.1.2 Odor sensing

E-nose technologies are sensors that respond reversibly to volatile compounds and generate electric signals that define odor fingerprints. These are widely used in agriculture for quality evaluation, process monitoring and detection of crop diseases. The e-nose has been used in monitoring aroma changes during ripening and shelf life assessment of fruits and vegetables [27]. The fruits' shelf life has been evaluated at room temperature and during cold storage. Humidity, ambient temperature and atmosphere, and the presence of other gases can be a challenge for electronic sensor systems.

2.1.3 Magnetic sensing

Nuclear magnetic resonance (NMR) instruments work by application of an external magnetic field and enable the detection of variations in the concentration or state of water and fats in fruits and vegetables. This can be used for assessing ripeness, defects and decay, primarily suited for produce with high water content [34]. NMR systems require high magnetic fields and sophisticated electronics; hence the limitation of this method is bulky and expensive equipment.

2.1.4 Soft X-ray imaging

Soft X-rays are electromagnetic waves that have wavelengths ranging from 1 to 100 nm. It only takes a few seconds to produce an X-ray image. These are appropriate for the evaluation of agricultural products, since they have low penetration power and can reveal internal density changes [32]. The method is used in the seed industry to detect internal voids, defects, insect infestation and insect damage [24, 33, 29].

2.1.5 Thermal imaging

Thermal imaging measures the reflection of light in terms of heat. It has been widely adopted in agriculture for monitoring the growth quality of fruits and vegetables, e.g. estimation of seasonal diameter growth of fruits [7]. Other applications include nursery produce monitoring [25], pathogens detection in crops [41], maturity evaluation [8] and bruise detection [9].

Thermal dissipation from organic produce can be used as a sensing modality. A thermal imaging based approach proposes to assess the quality of produce through human touch interactions. Thermal radiation is transferred from users to objects as they touch them and the dissipation of this heat is examined to establish a thermal dissipation fingerprint [10].

2.1.6 Light sensing

Light sensing has been examined by using red spectrum to capture the decomposition of produce over time. The quality state of fruits has unique fingerprints as reflected light values correlate with the transpiration coefficients of produce [49]. This method is low cost and easy to implement throughout the supply chain, however it requires specialized devices and new sensor designs.

2.1.7 Other methods

Spectroscopy measures the absorption of different light wavelengths instead of using light reflection. Regulatory food safety inspections rely on spectroscopy as a highly accurate method that is able to estimate internal quality factors. However, the equipment is bulky and expensive, and measurement taking and analysis are time-consuming.

Other proposed methods of produce quality estimation include bio-inspired soft tactile sensor [38] and texture sensor based on highly sensitive hair-like cilia receptors [28]. The limitation of surface analysis is the need for physical interaction that can damage the fruit. Internal quality factors can be assessed by absorption of wireless signals to estimate the water content of produce [22], which in turn correlates with the state of fruits and vegetables. The wireless technique is highly sensitive to the measurement setup though.

Furthermore, maturity level can be evaluated by changes in volatiles emitted by fruit during ripening. This can be investigated by gas chromatography (GC) with headspace sampling and GC combined with mass spectrometry (GC-MS) [2]. Expensive devices and sophisticated analysis limit the use of these methods. Besides, these methods can only be applied to measure volatiles in unsealed packages of food. To overcome this limitation, a tunable diode laser spectroscopy method has been developed to detect volatiles that are sealed in containers or packages [45].

2.2 Pesticide residue estimation

Unlike quality estimation, visual inspection of fruits and vegetables is usually of no use to determine pesticide residues as there may be no visible traces. Therefore, computer vision systems are not a fit for this purpose. Chromatography and biosenors are the main methods for pesticide residue estimation.

2.2.1 Chromatography

Gas and liquid chromatography or chromatographic methods coupled with mass spectrometry are classical techniques for the detection of pesticides [6, 26]. Their pros are sensitivity, separation and identification abilities. However, the cons are laboriousness, highly skilled manpower and costly instruments. Moreover, these require pre-treatment and extraction processes as detection is not done in real samples. This makes real-time on-site detection of pesticide residues unfeasible.

2.2.2 Biosensors

Biosensors overcome several limitations of conventional methods, e.g. sample preparation. These are analytical devices that detect changes in biological processes and convert them into electrical signals. The main detection methods in biosensors are optical, electrochemical, piezoelectric and molecular imprinted polymer [30]. Biosensors are frequently used in medicine, environmental monitoring, the food industry and agriculture. In agriculture, biosensors are generally used for the detection of pesticides. Sensor-based techniques offer several advantages such as low-cost, simple, rapid operation, highly sensitive and selectivity on-site detection with detection limits lower compared to the classical chromatographic methods [35].

2.2.2.1 Enzyme-based biosensors

Enzyme-based biosensors are based on the inhibition reaction or catalytic activity of several enzymes in the presence of pesticides. Since some pesticides have a similar mode of action affecting the activity of the same enzyme, enzyme-based biosensors are mostly used for screening purposes and unspecific for individual pesticides [23]. Thus, they can determine the total pesticide content and not a particular pesticide.

2.2.2.2 Immunosensors

Immunosensors are biosensors that use antibodies or antigens as the sensing element and provide concentration-dependent signals. Electrochemical immunosensors are used more frequently, optical sensors less and piezoelectric sensors least actively for the detection of pesticides in fruits and vegetables [11]. Amperometric, potentiometric, conductometric and impediometric for electrochemical; and fluorescence, colorimetric, chemiluminescence, electrochemiluminescence, surface plasmon resonance and surface-enhanced Raman spectroscopy for optical sensors are the proposed immunosensors for pesticide detection.

High sensitivity, convenience, simplicity and broad linear range are the advantages of the immunosensors. However, a few of them are widely applied due to concerns around cost, usability and speed of analysis [13]. Advances in nanotechnology promise further improvement and miniaturization of biosensor devices. Nanomaterials have improved the concept of flexibility, stability, optical transparency and compatibility [46]. The drawback of using nanomaterials such as metal and metal oxide nanoparticles is their toxic effect [37].

2.2.3 Other methods

Hyperspectral and near infrared (NIR) imaging harness the dual advantage of image and spectrum. The detection process is non-destructive, non-polluting and does not require pre-treatment of sample. NIR sensing is proposed for the detection of pesticides in food [43, 17] and can provide preliminary screening for pesticides and other chemical residues in produce [40]. NIR microscopic image technique has demonstrated the ability to detect pesticide concentration in a vegetable [42]. Hyperspectral imaging combined with machine learning algorithms has shown potential in pesticide residue estimation in vegetables [47, 21, 44].

Interferometric sensing uses laser light that changes the speed when passing through contaminated produce. The change is measured and determined against a set of existing values to spot bacteria and pesticides. The plasmo-photonic bimodal multiplexing sensor is proposed to detect pesticide residue without the use of chemicals or dyes as a marker [5]. The photonic sensor is expected to cut the time for pesticide residue detection in fruits and vegetables from a few days to minutes.

3 Methodology

The methodology chapter introduces the key methods and techniques used in the study. First, the concept and benefits of light reflectivity as a sensing modality are presented. Then, the core elements of machine learning techniques and approaches applied for data analysis are outlined.

3.1 Light reflectivity

Reflection is the process of electromagnetic radiation that the light wave returns. The incoming light wave is called an incident wave and the wave that bounces away from the object is termed the reflected wave. When the light wave reflects from the object, it returns in a certain manner that mirrors the original wave. The law of reflection states that upon reflection from a smooth surface, the angle of the reflected wave (r) from the normal is equal to the angle of the incident wave (i), i.e. $r = i$ (see Figure 1).

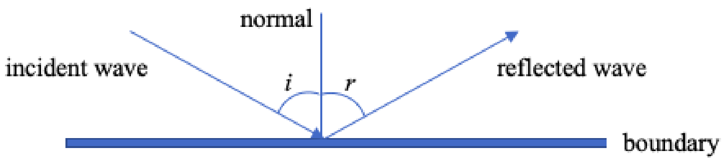


Figure 1. Return of reflected light wave from the incident light wave at an equal angle.

The sensing modality of light reflectivity measures how much light is reflected by the object instead of how much it absorbs or transmits. Light reflection can either occur at the boundary between two media (surface reflection) or at the interior of a medium (volume reflection). The reflectivity of the object is determined by its physical properties, such as color, texture and composition.

3.1.1 Reflectivity by color

The color of an object affects its reflection properties. When the light wave strikes the surface of an object, the light can be absorbed, transmitted or reflected. The object's color determines which wavelengths of light are absorbed and which are reflected. For example, a red surface of an object absorbs the green and blue and reflects red wavelengths. Therefore, given separate red, green and blue incident light wave sources, the reflected light amount from a red object is the highest for the red light source and less for green and blue light sources.

3.1.2 Reflectivity by texture

The amount of light reflected by an object and how it is reflected is dependent upon the smoothness or texture of the surface. The reflection of light can be generally categorized into two types: specular and diffuse reflection. Specular reflection is defined as light reflected from a smooth surface at a definite angle. Regarding the law of reflection, for specular reflection, the reflected light follows a trajectory of the same angle as the incident light. On the contrary, diffuse reflection is produced by a rough surface that reflects light in all directions.

Therefore, the texture of an object determines the amount of light reflected back and captured at the point of the light source. For specular reflection, the light reflectivity values are higher as all waves travel back. Whereas in the case of diffuse reflection, only some waves come back to the source point since the rest of them scatter in different directions.

3.1.3 Reflectivity by composition

The composition of an object determines its physical and chemical properties, including its reflectivity. For example, metals are highly reflective due to the presence of free electrons that interact with incident light waves and reflect it back. On the contrary, non-metallic materials such as ceramics, glass and plastics have lower reflectivity due to their chemical composition and lack of free electrons. Fresh produce has even lower reflectivity due to its organic composition, while a high index of refraction determines higher reflectivity for materials.

3.1.4 Measure of light reflectivity

A photoresistor can be used to capture the value of light reflectivity. A photoresistor is a type of resistor whose resistance varies with the amount of light falling on its surface. As the amount of light falling on the photoresistor changes, its resistance also changes and this change in resistance can be measured and used to determine the level of light reflectivity. The output of the photoresistor can be calibrated using a known light source or a reference standard to provide accurate measurements of light reflectivity.

3.1.5 Benefits of light sensing

There is a wide range of benefits of light sensing compared to other sensing modalities. These advantages make it a versatile and useful method for various applications, including imaging, sensing, communication and monitoring. In the context of light reflectivity, the benefits of light sensing are the following: non-invasive, high sensitivity, cost-effective, fast response and low power consumption.

3.1.5.1 Non-invasive

Light sensing is a non-invasive sensing modality that does not require physical contact with the object. This makes it an ideal solution for applications where physical contact is not possible or undesirable such as monitoring. Besides, non-invasive methods are typically faster and less labor-intensive than invasive ones, which require physical sampling or alteration of the product.

3.1.5.2 High sensitivity

Light sensors are highly sensitive and detect even small changes in light reflectivity. This makes them perfect for applications such as biomedical sensing, where even small changes in light reflection can provide valuable information.

3.1.5.3 Cost-effective

Generally light sensors are low-cost and widely available. This makes them accessible to a wide range of applications and industries. It is paramount for farmers in developing countries and small businesses to have access to low-cost devices. Moreover, it can increase the efficiency of the production process, since expensive and time-consuming laboratory tests can be skipped.

3.1.5.4 Fast response

Light sensors can rapidly detect changes in light reflectivity, making them perfect for applications requiring real-time or near-real-time sensing, such as industrial automation and robotics. The need for fast-response portable equipment to use in the field and packinghouse is recognized [39]. Furthermore, quick results for the state of fresh produce contribute to consumer protection. Light sensing devices allow consumers to exercise their power while shopping instead of relying on regulatory food inspections that are rather sporadic.

3.1.5.5 Low power consumption

Many light sensors have low power consumption and can be easily integrated into low-power electronic devices. This makes them ideal for battery-operated devices and IoT applications. Besides, low-power devices reduce energy consumption and therefore contribute to reduced cost and environmental sustainability.

3.2 Machine learning

Machine learning is a subfield of artificial intelligence that encompasses algorithms and statistical models for making predictions or decisions based on data without being explicitly programmed to do so. This is achieved by training models to improve performance on a specific task by learning from data. Overall, machine learning algorithms can be divided into supervised, unsupervised and reinforcement learning.

3.2.1 Supervised learning

Supervised learning involves training a model on labeled data, where the target output is known for each input example in order to predict outputs for new, unseen instances. Regression and classification are the two main types of supervised learning problems. In regression, the output feature is a continuous variable and the objective is to predict a numerical value. On the other hand, for classification, the output feature is a categorical variable and the objective is to predict a label or class for a given input.

3.2.2 Input feature selection

Feature selection involves the identification of the most relevant and informative input features for predicting the output variable. Thus, irrelevant or redundant features are discarded.

Input feature selection is important in the context of overfitting. The selection of too many features or wrong features can lead to overfitting. Overfitting occurs when the model is trained too well on the training data to the point where it begins to memorize the noise or random fluctuations in the data rather than learning the underlying patterns and relationships. This can lead to poor performance on new, unseen data, as the model has not learned to generalize beyond the training data.

3.2.3 Model training

Model training comprises data preparation, model initialization, training, prediction and evaluation. Both regression and classification follow the same steps in model training. The difference lies in the used algorithms and evaluation metrics.

3.2.4 One-hot encoding

One-hot encoding, a technique that converts categorical data into numerical data to be used as input for the machine learning algorithm, is deployed in the data preparation step. For example, the types of vegetables constitute categorical variables that require transformation. The one-hot encoder creates a binary vector representation for each category in the data. This ensures the model does not interpret the numerical values as continuous data and mistakenly attempts to draw relationships between the values.

3.2.5 Train-test split

Data preparation also involves the train-test split, which divides the available data into the following two sets: a training and a testing set. The training set is used to train the machine learning model, while the test set allows for evaluating model performance on unseen data. The train-test split function from the scikit-learn Python machine learning library is used with a ratio of 80% for training and 20% for testing. The original dataset is split into two subsets by means of random selection.

3.2.6 Separate and single model training

Two main approaches to building models can be distinguished – separate and single model training. Separate model training refers to separate machine learning models for each specific task or problem. For example, in the case of fresh produce, separate models for each vegetable type are trained. This approach can lead to more accurate predictions for each specific task as the model is optimized specifically for that task. Single model training, on the other hand, trains a single machine learning model to perform multiple tasks or solve multiple problems. This approach can be more efficient and easier to manage since only a single model needs to be trained rather than multiple separate models.

3.2.7 Scikit-learn

The scikit-learn library provides a wide range of tools and algorithms for the implementation of machine learning models in Python. It is a user-friendly and powerful library that is built on top of other scientific Python libraries such as NumPy, SciPy and matplotlib.

The regression and classification models from the scikit-learn Python machine learning library are deployed for model training. There are about 40 regressors and 30 classifiers available in the scikit-learn library. These include both traditional machine learning models as well as some deep learning models. The latest stable release at the time of the data analysis is scikit-learn version 1.2.2 [36].

3.2.8 Performance metrics

Performance metrics are used to evaluate the performance of a machine learning model. The evaluation metric choice depends on the problem being solved and the type of model used. For regression, the common metrics include mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE). The classification problem is often evaluated by accuracy, F1 score and area under the curve (AUC).

In this thesis, model accuracy is used to compare various models in their prediction of the target class. The accuracy measures the proportion of correct predictions made by the model. The MAE score, that is the average absolute difference between the predicted and true values, is given for the regression problem. However, the evaluation of the regression model is made by plotting the true value and predicted value.

4 Experimental design

In this chapter, the experimental design is outlined for the use of light sensing to estimate produce quality and pesticide residue. These two distinct objectives are pursued via a single experiment. First, the selection of fresh produce obtained from a retail outlet is described. Then, the experimental setup and measurements are discussed in detail. Further, the application of pesticides is specified. Finally, an account of the overall procedure is given for a 12-day experimental study.

4.1 Objects

Fresh produce commonly available at retail stores is considered for the experiment. This covers the following five categories of vegetables: leaves (cabbage and spinach), stem (celery), fruits (tomato), pods (bean) and flowers (cauliflower). Roots, bulbs, tubers and seeds are omitted from the study since pesticide application while these grow does not take place directly on the edible plant part, e.g. potato tubers are growing underground and therefore do not get directly sprayed. Moreover, tubers and bulbs can be stored at room temperature for a considerable period of time without losing quality. Hence produce decomposition from ripe to decayed would not be captured for these vegetable categories over a 12-day period.

The fresh produce is bought at one time from the same retail outlet. All the vegetables are initially in a fresh state without signs of decomposition. Two samples from each vegetable category are selected to have similar appearance, shape and size to ensure that samples' differences only come from inherent properties. Each produce sample is marked with a unique identifier, e.g. spinach-1 and spinach-2.

4.2 Testbed

A testbed is set up to measure the vegetable items for light reflectivity. The setup comprises a white table where samples of vegetable items are placed. The table is positioned in the room in such a way that each side of the table receives the same amount of light. Since touching the produce can affect its surface, each item is kept on its spot and therefore, the measurements are taken by moving the instrument around the items.

4.3 Apparatus

The apparatus (see Figure 2) to measure light reflectivity consists of three laser sensors – red, green and blue light – and a photoresistor to receive the reflected light. The photoresistor is connected to the analog input pin of an M5StickC PLUS ESP32 development board, which integrates Wi-Fi capabilities. The board controls the sampling frequency (5Hz) and records the light values to a CSV file along with a timestamp for each record.

Also the light reflectivity values can be seen on the device screen in real time. The light sensor is easy to deploy and its components are low-cost. This makes light sensing an affordable solution that can be easily scaled.

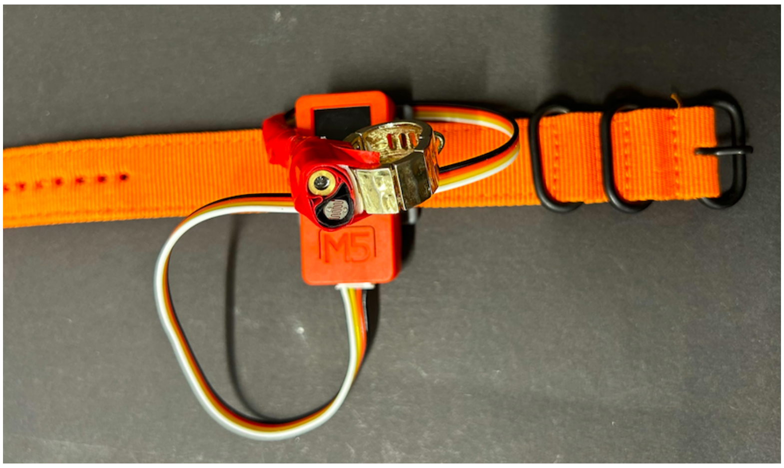


Figure 2. The device with three laser sensors (red, green and blue) and photoresistor to measure light reflectivity.

Additionally, three more devices are deployed to obtain reference values. First, a non-invasive durometer depicts the firmness of the vegetable items. Second, a room thermometer records the daily temperature. Third, a Lux meter application running in the smartphone Xiaomi M11 measures the experimental testbed's light intensity (lux).

4.4 Pesticide application

The most common pesticides are selected from the two main categories - fungicide (Switch) and insecticide (KarateZeon) - that are freely available at various retail stores. First, Switch by Baltic Agro is a fungicide where the active substances are cyprodinil and fludioxonil. Cyprodinil poses a risk to human health by being liver toxic as well as impacting reproductive and developmental processes. The human health concerns of fludioxonil include liver and kidney damage. Second, KarateZeon by Baltic Agro is an insecticide where the active substance is lambda cyhalothrin. Lambda cyhalothrin is highly neurotoxic and can cause short-term effects or tremors. These pesticide substances are widely found in the contamination of European vegetables [12].

The pesticide solutions are prepared according to the application instructions and sprayed evenly over the vegetable. The solutions are mixed in two identical new spray bottles so no other residue mixes with the pesticide dilution. Two samples of produce for each pesticide treatment are taken. Additionally, one two-samples produce set is not treated with anything and constitutes the baseline (no treatment group).

4.5 Procedure

Measurements are collected from each set of fungicide-, insecticide- and untreated vegetable items once a day over a 12-day long period. Variations in light reflectivity capture changes in surface and characteristics of produce.

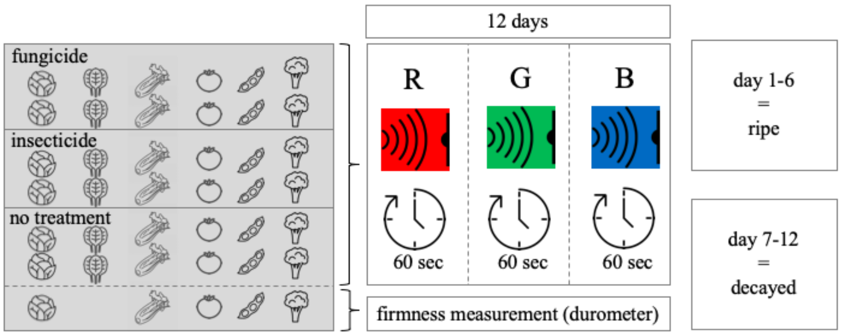


Figure 3. The procedure of light sensing for produce quality and pesticide residue estimation.

Pesticides are only applied on the first day and the surface of produce is left to dry completely before the first measurement is taken. Every day measurements are taken in the second half of the day. Each sample is measured over a 60-second period by moving the sensor around the item from a distance of one cm. Every time the two samples in each set are chosen randomly.

First, measurements are taken with red light, then the same procedure is repeated with green and finally with blue light. Sample-1 and sample-2 of no pesticide group is measured for firmness with durometer, while sample-3 is left untouched by durometer. The daily temperature and light intensity (lux) values are recorded. The produce is kept at room temperature (≈ 24°C) throughout the experiment.

5 Results

This chapter presents the results of the experimental study conducted in August 2022. First, red, green and blue light reflectivity is assessed as a sensing modality. This includes an account of light characterization and generalization. From comparing the three light sources, blue light is chosen as a basis for further analysis. Then, produce quality and pesticide residue estimation are examined based on the light reflectivity values.

5.1 Light characterization

Light reflectivity values from red, green and blue light demonstrate that fresh produce can be characterized by variations in light intensity across different vegetable items. The Kruskal-Wallis test confirms the differences to be significant between various vegetables for red ($x^2 = 67493, \eta^2 = 0.75, p < 0.05$), green ($x^2 = 76754, \eta^2 = 0.85, p < 0.05$) and blue light ($x^2 = 73306, \eta^2 = 0.82, p < 0.05$). These variations constitute unique fingerprints that allow to determine the vegetable type irrespective of pesticide treatment.

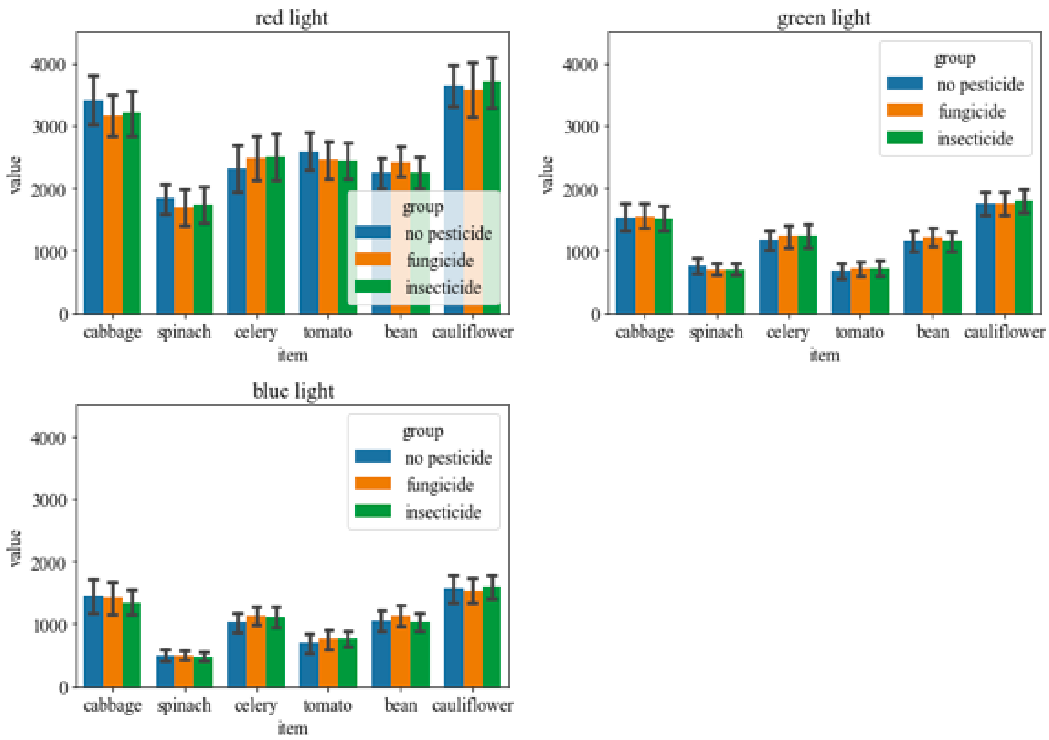


Figure 4. Item-wise median light values and standard deviation of treatment groups (no pesticide, fungicide and insecticide) for red, green and blue light sensing.

Figure 4 shows the item-wise light reflectivity median values and standard deviations for no pesticide, fungicide and insecticide groups. While cabbage, spinach, tomato and cauliflower tend to have more distinctive values across three lights, celery and bean exhibit more similar fingerprints throughout all three lights. Posthoc comparisons (Dunn-Bonferroni) prove that differences are statistically significant ($p < 0.05$) for all vegetables given red, green and blue light with the exceptions for the green light: spinach-tomato ($x^2 = -0.11, p > 0.05$) and bean-celery ($x^2 = 1.05, p > 0.05$); for blue light: bean-celery ($x^2 = 0.52, p > 0.05$); for red light: bean-celery ($x^2 = 1.68, p > 0.05$) and celery-tomato ($x^2 = 1.31, p > 0.05$). Thus it appears that all three lights grapple with differentiating bean and celery, and blue light outperforms red and green light sources when it comes to capturing vegetable items by light reflectivity values.

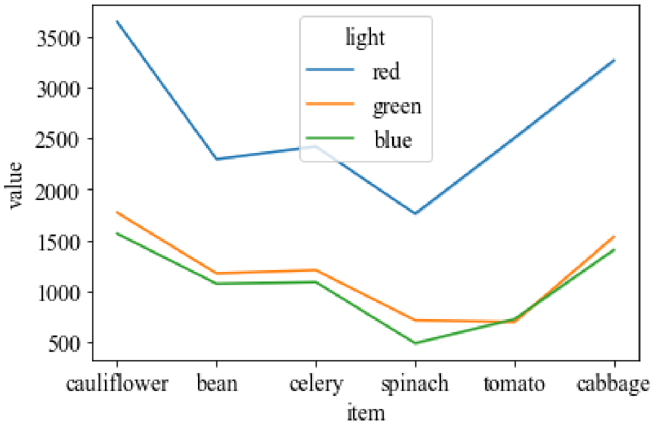


Figure 5. Item-wise median light values for red, green and blue light sensing.

Red laser results in the highest variance in light reflectivity, whereas green and blue lights have a lower variance. The reflectivity from green and blue light is on par ranging from 500 to 1500, whereas red light values fall in a higher range from 1500 to 4000. Figure 5 shows that overall reflectivity patterns for different items are very similar, e.g., cauliflower and cabbage have the highest and spinach have the lowest values. Kolmogorov–Smirnov test also shows no significant difference ($p > 0.05$) for all the light sources, i.e. for blue-green: KS=0.17, for blue-red: KS=0.33, and for green-red: KS=0.33.

5.2 Light generalization

For red light, no matter whether the pesticide is on the vegetable or not, the Wilcoxon test using light values as a condition shows that differences in measurements between the two samples (sample-1 and sample-2) are not statistically significant ($p > 0.05$), except the cabbage ($x^2 = 174, p < 0.05$) and bean ($x^2 = 108, p < 0.05$). For green light, the exception of significance between the two samples is bean ($x^2 = 192, p < 0.05$) and for blue light is cabbage ($x^2 = 202, p < 0.05$). This result indicates that light sensing can be generalized to identify different vegetable samples of the same item type (see Figure 6). The lack of generalizability of cabbage for red and blue light can be attributed to the uneven color of the surface.

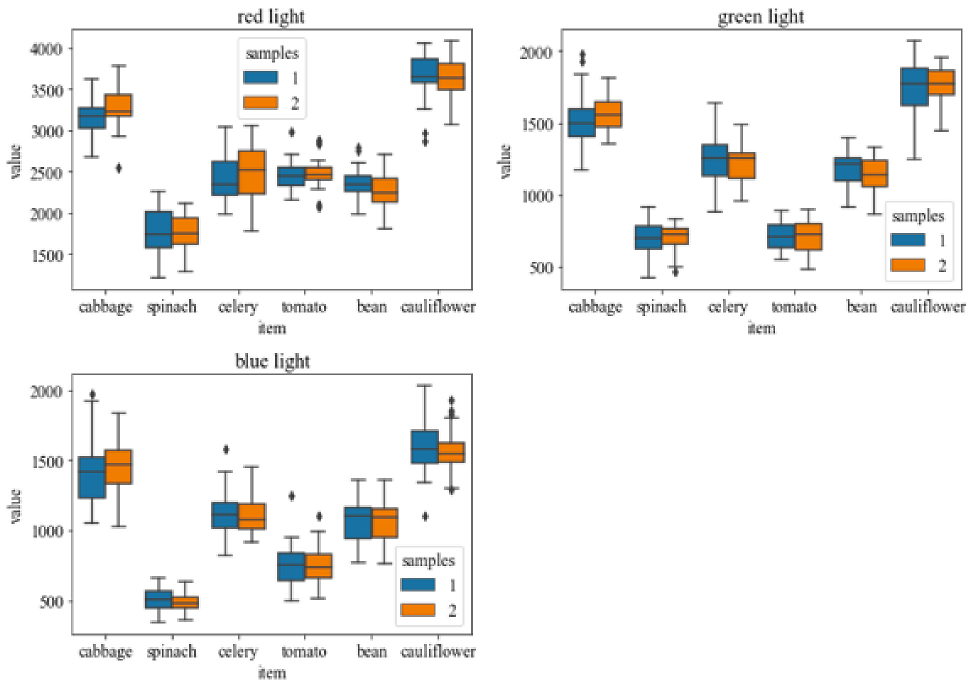


Figure 6. Item-wise median light values and standard deviation for samples 1 and 2 of vegetables given red, green and blue light.

As the durometer applies pressure on the surface, it can cause damage, speeding up the decomposition process and affecting the light sensing. To assess this, the light values of the vegetables for which durometer measurements are available are separately compared to a third sample of the same produce that is not measured with the durometer (no pesticide group). Figure 7 depicts the three samples for different light sources.

For blue light, Friedman tests show no significant difference across different samples for all the vegetables ($p > 0.05$), verifying that the durometer does not damage the vegetable during the experiment and that the decomposition characteristics remain similar across the produce of the same type. However, for the red light, it is cabbage ($x^2 = 18.2, p < 0.05$), tomato ($x^2 = 11.2, p < 0.05$) and bean ($x^2 = 9.5, p < 0.05$), and

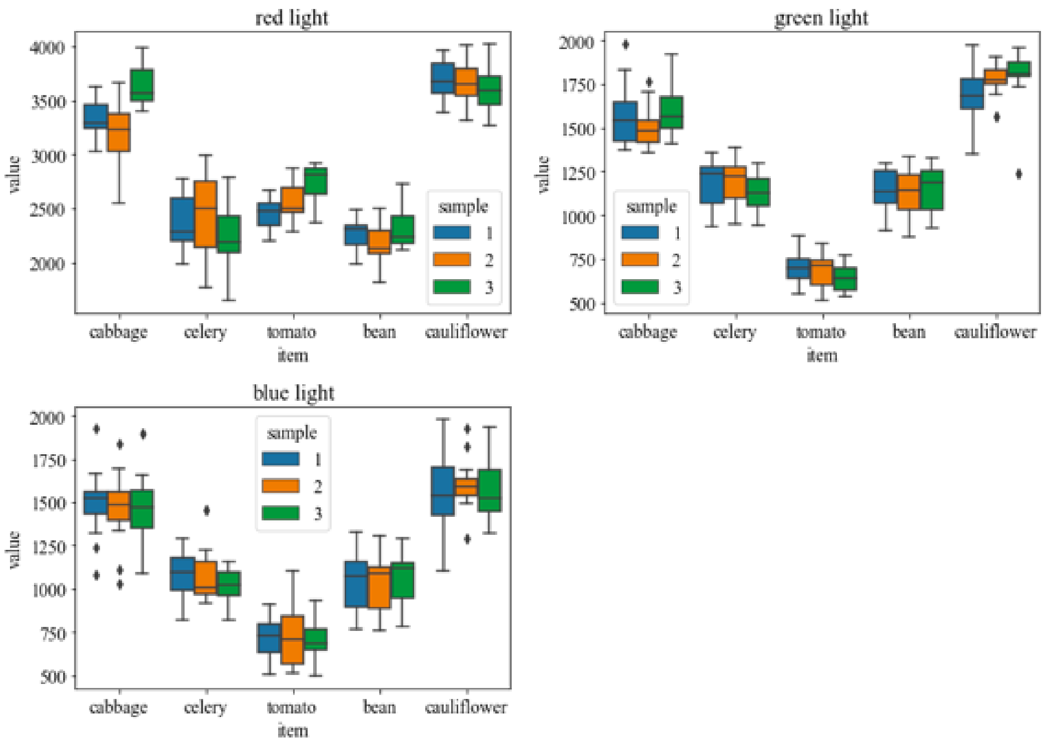


Figure 7. Item-wise median light values and standard deviation for samples 1, 2 and 3 of no pesticide treatment group given red, green and blue light.

for the green light, it is cabbage ($x^2 = 6.5, p < 0.05$) and tomato ($x^2 = 8.7, p < 0.05$) that show disparities across the samples. This suggests that unlike blue, red and green sensors might be more sensitive to capturing any damage to produce.

5.3 Light source comparison

Figure 8 showcases the daily differences in reflectivity of vegetable items for red, green and blue light sensors. On the left is given the mean with the 95% confidence interval (CI); on the right is the median value and the standard deviation (SD). There are negligible differences in mean and median values, hence both are suitable for analysis. Again, the red light exhibits high variance and celery and bean overlap values for all three lights. Moreover, tomato values blend in celery and bean for the red light and in spinach for the green light. Out of the three lights, green appears to be the most stable over time, i.e. the light values for particular items do not follow any trend. Whereas red and blue lights are somewhat dynamic, i.e. changes in light reflectivity in time can be noticed.

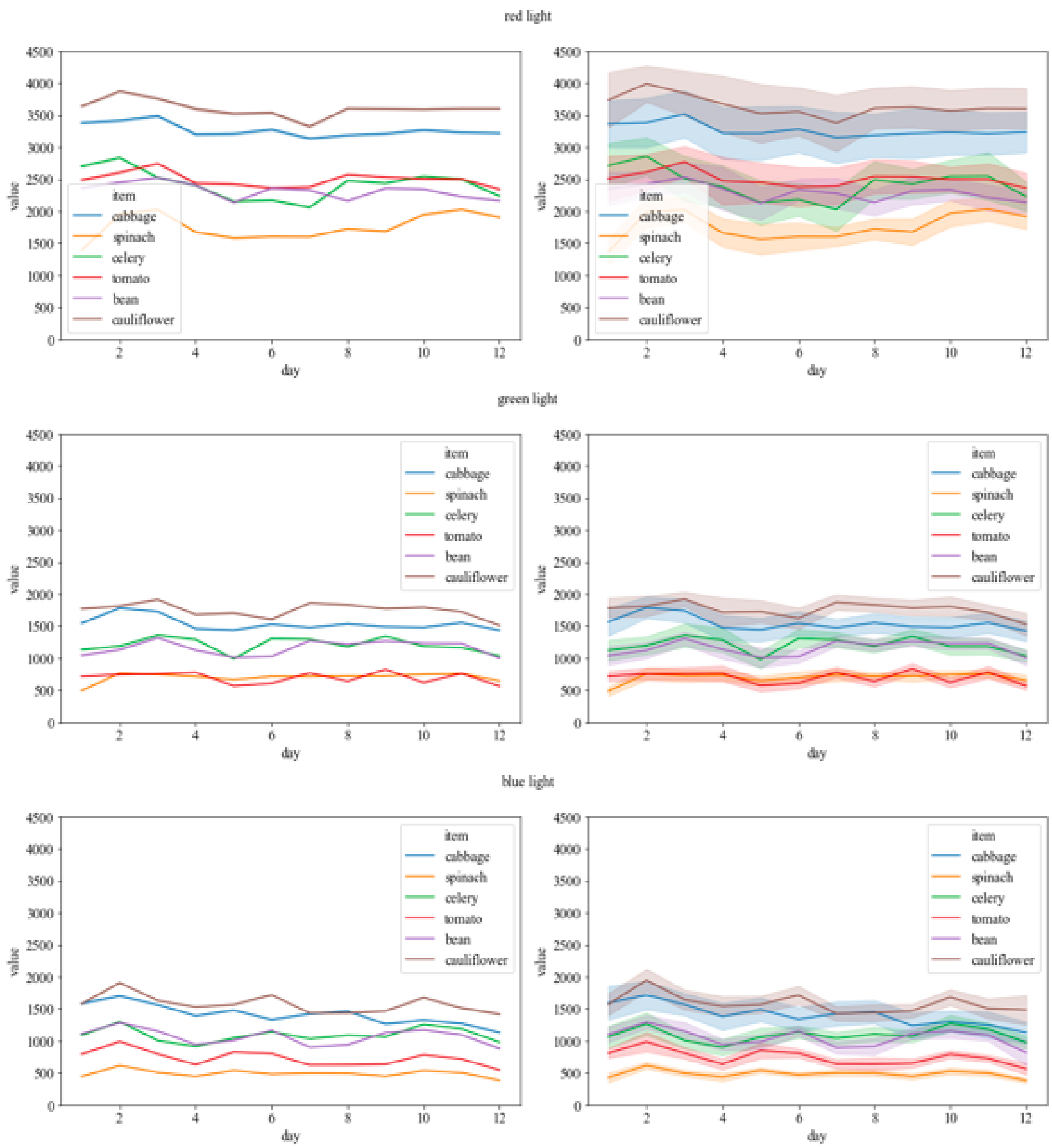


Figure 8. Day-wise mean + 95% CI (left) and median + SD (right) light values of vegetable items for red, green and blue light sensing.

Given the observations of vegetable items from red, green and blue light sensors for different treatment groups (Figure 4), light reflectivity pattern (Figure 5) and 12-day long time span (Figure 8), it appears that blue light is most appropriate for conducting further analysis. Unlike red light, there is a lower variance and unlike green light, there can be observed a trend in light reflectivity over time for the blue sensor. Blue light appears most promising for capturing differences in treatment groups since the median values are not as even as in the case of green light and the standard deviation is not as much overlapping as in the case of red light. Therefore, from here onwards only the reflectivity values from the blue light sensor are considered for produce quality and pesticide residue estimation.

5.4 Produce quality estimation

5.4.1 Capturing quality of fresh produce

The ability of blue light reflectivity to capture produce quality is assessed. For this purpose, light values from only no pesticide treatment group are evaluated. Figure 9 demonstrates the median values with a 95% confidence interval (CI) of sample-1 and sample-2 for each day over the 12-day decomposition period.

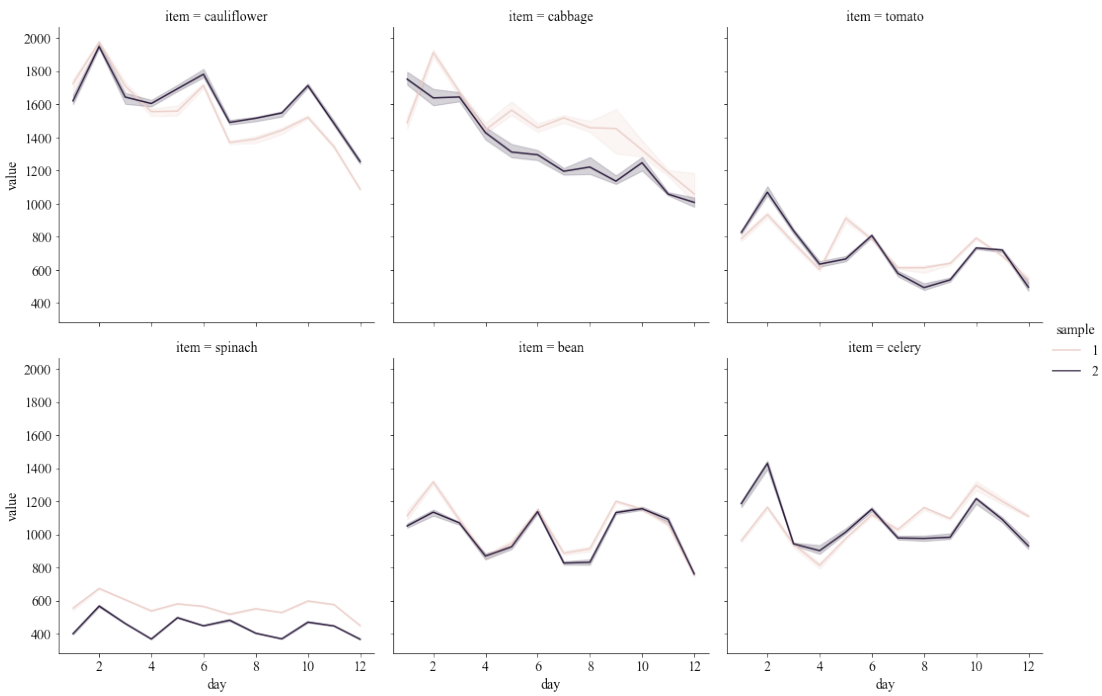


Figure 9. Day-wise median + 95% CI light values of samples 1 and 2 for vegetable items.

Cauliflower, cabbage and tomato have a clearly established trend of decreasing reflectivity intensity due to produce decomposition. There are minor or non-existent changes in time for spinach, bean and celery. During the experiment, the spinach leaves dried up quickly and remained dry, so these items obtained a stable state and reflected in light values. Bean and cabbage first slightly softened and then began to dry and harden, other vegetable items followed a softening trend (see Figure 10 for firmness values from durometer). Figure 11 displays the produce appearance at four stages (day 1, day 5, day 9 and day 12) over a period of 12 consecutive days.

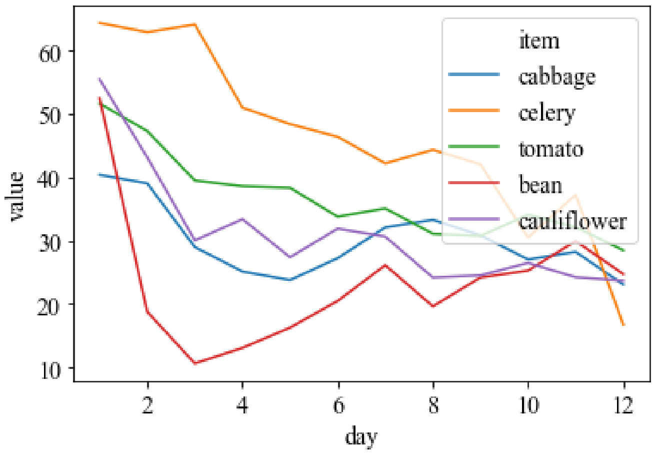


Figure 10. Day-wise average durometer values for vegetable items.

Figure 11. Vegetable items at different stages during 12 consecutive days.

Figure 12 depicts each vegetable item's daily median absolute deviation (MAD) value compared to the corresponding fingerprint on day 1. While light values for cabbage display a linear function of produce decomposition, the rest of the vegetable items are more inconsistent. However, similarly to the earlier figure, other vegetable items too have a certain tendency to change values as a response to produce decomposition.

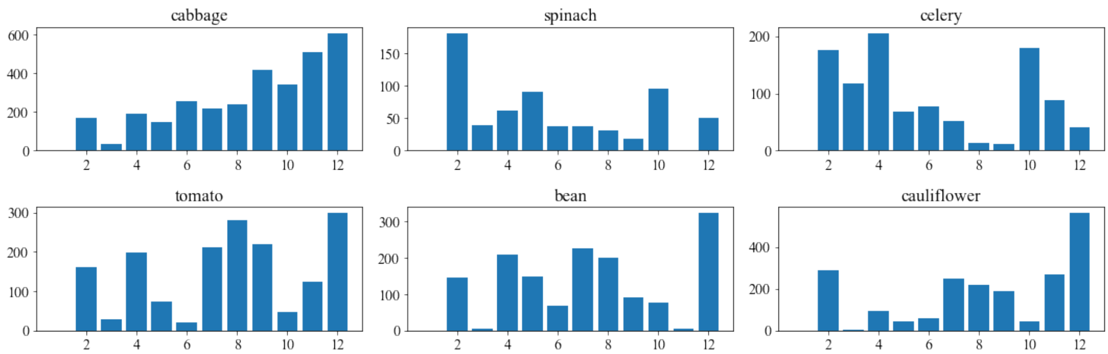


Figure 12. Day-wise median absolute deviation (MAD) of light values for vegetable items.

Next, the daily light values are grouped into two categories: days 1-6 as ripe and days 7-12 as decayed. Figure 13 displays the median and standard deviation (SD) of ripe and decay stages for each vegetable item, distinguishing between sample-1 and sample-2.

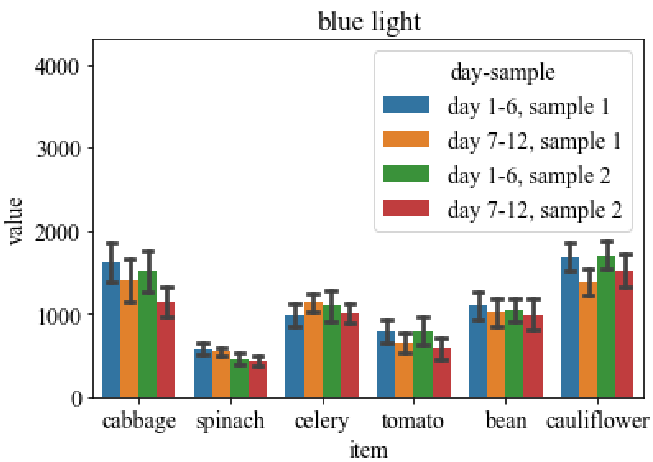


Figure 13. Two-stage median + SD light value of samples 1 and 2 for vegetable items.

The overall tendency for a decline in reflectivity intensity can be observed, which is apparent for both samples of cabbage, cauliflower and tomato. Spinach, celery and bean are less distinguishable from ripe to decayed stages. The Wilcoxon test confirms there are statistically significant differences ($p - values < 0.05$) for all vegetable items of ripe and decayed groups (see Table 1 for test statistic t and p-value).

Table 1. Wilcoxon test statistic t and p-value of ripe vs decayed vegetable items for samples 1 and 2.

	Sample 1		**Sample 2**	
Day 1-6 vs day 7-12	**t**	**p**	**t**	**p**
Cabbage	93116	0.000	19619	0.000
Spinach	105014	0.000	208005	0.000
Celery	63866	0.000	187403	0.000
Tomato	105006	0.000	53333	0.000
Bean	204598	0.000	211655	0.000
Cauliflower	16788	0.000	86644	0.000

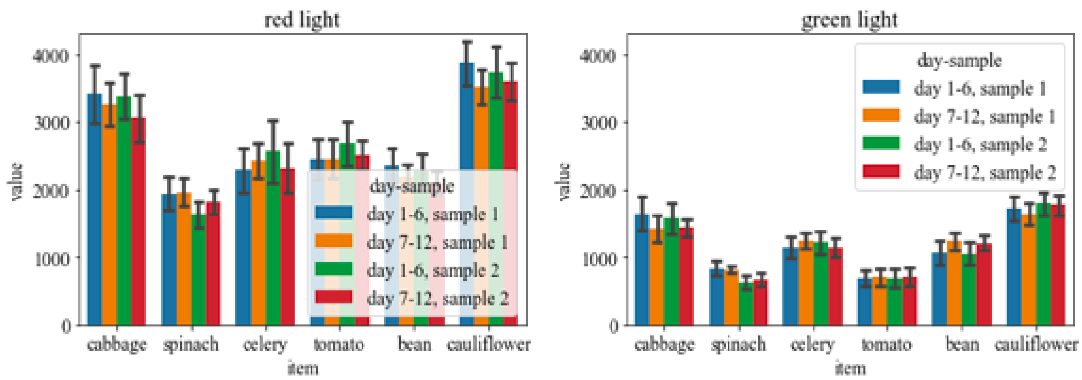


Figure 14. Red (left) and green sensor (right) two-stage median + SD light values.

For the purpose of comparison, the two-stage-wise median and SD light values are also given for red and green light (see Figure 14). These other two lights are evidently underperforming as a sole sensing modality. However, given the combination of all three light sensors, changes in produce quality get well captured for all vegetable items.

5.4.2 Predicting quality of fresh produce

Machine learning models for classification are deployed to predict produce quality based on the light reflectivity values. The following two stages of ripeness are considered to characterize the quality of vegetable items: ripe (day 1-6) and decayed (day 7-12). First, separate models are trained for each vegetable item considering merely light values as a single input feature. Second, the approach of a single machine learning model is explored. Finally, for baseline comparison a regression model is trained to predict firmness values from light reflectivity along with other input features.

5.4.2.1 Separate model classification

The following ten classifiers are explored for model performance: Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gaussian Naïve Bayes (GNB), Multilevel Perceptron (MLP), Quadratic Discriminant Analysis (QDA), AdaBoost (AB) and XGBoost (XGB).

As seen in Table 2, separate model accuracy is highest for cauliflower (78.1%), followed by tomato (76.5%) and cabbage (73.5%). Classifiers for bean (64.3%), spinach (61.6%) and celery (55.9%) are rather poor – these three vegetable items are known for their obscurity from prior analysis.

Table 2. Model accuracy of ten classifiers for individual vegetable items.

	LR	**DT**	**RF**	**SVM**	**KNN**	**GNB**	**MLP**	**QDA**	**AB**	**XGB**
Cabbage	**73.5**	70.2	69.8	72.7	70.7	73.5	72.6	73.5	72.9	72.3
Spinach	59.1	57.2	56.3	61.5	57.0	61.5	53.9	61.5	**61.6**	59.4
Celery	49.1	55.0	55.5	56.5	56.5	55.3	50.9	55.3	**57.5**	55.9
Tomato	74.4	72.7	73.0	**76.5**	71.8	75.2	71.1	75.2	76.1	74.7
Bean	57.4	58.5	58.5	**64.3**	60.6	59.3	52.2	59.3	61.9	59.5
Cauliflower	**78.1**	75.1	74.6	78.0	75.9	78.0	50.5	78.0	77.0	77.0
Average:	65.3	64.8	64.6	**68.3**	65.4	67.1	58.5	67.1	67.8	66.5
Average (-spinach):	66.5	66.3	66.3	**69.6**	67.1	68.3	59.5	68.3	69.1	67.9

Evidently, there is no particular classification model that works best for all vegetable items. However, certain classifiers outperform others: LR for cauliflower and cabbage, SVM for tomato and bean, and AB for spinach and celery. GNB and QDA are as good as LR for cabbage and very close to AB for spinach. The SVM is the most universal classifier, i.e. it scores well across all vegetables with an average accuracy of 68.3%. In the case where spinach is excluded (since no firmness values are recorded for spinach), the average accuracy of SVM stands at 69.6%.

5.4.2.2 Single model classification

Next, a single model is trained for all vegetables except spinach to predict the quality of fresh produce. Performance accuracies of LR, SVM and AB are measured as these classifiers did the best for separate models. Table 3 shows the selected models' accuracy for four different input feature combinations. Across all scenarios, the AB algorithm outperforms the other two models.

Table 3. Single model accuracy of top three classifiers for vegetable items given various input feature combinations.

	LR	**SVM**	**AB**	**Average**
Light reflectivity value (R)	54.6	63.0	**63.1**	60.2
R + vegetable item (V)	66.3	60.6	**67.3**	64.7
R + luminosity (L)	77.4	75.0	**88.1**	80.2
R + V + L	87.2	76.0	**92.4**	85.2

When light reflectivity alone is considered, the accuracy for a single model (AB) is 63.1%. In this case, no information about the type of produce is given. Supplementing the dataset with vegetable type improves the accuracy to 67.3%. However, light reflectivity along with luminosity gives a leap for model accuracy amounting to 88.1%. Further, combining all three input features - light reflectivity, vegetable type and luminosity - takes the accuracy score to 92.4%.

These outcomes highlight the importance of considering lightning conditions as it impacts light reflectivity. If light reflectivity is not corrected with luminosity, then accuracy performance suffers. Therefore, it is essential to include luminosity as an input parameter in the model to account for varying lighting conditions.

5.4.2.3 Regression for firmness prediction

The firmness of fresh produce is often associated with produce quality. Therefore, the daily firmness values measured by the durometer can be considered as a baseline for produce quality. In other words, light reflectivity values should be able to predict the firmness of the vegetable items. For this purpose, a machine learning model for regression is deployed for all vegetables except spinach. Only values of sample-1 and sample-2 are included since sample-3 was not measured for firmness, as it was kept as a reference point to evaluate the effect of the durometer.

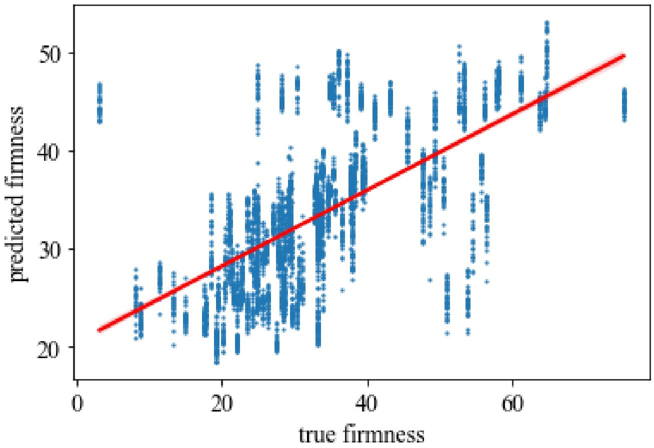


Figure 15. True firmness (x-axis) and predicted firmness (y-axis) of the regression model.

A linear regression model is trained given the light value, lux and vegetable items as input features. The output feature or prediction is the firmness value. As a result, the mean absolute error (MAE) is 7.39. The model performance is illustrated in Figure 15, where the true firmness is given on the x-axis and the predicted firmness on the y-axis. Overall, the regression model is able to predict the firmness value from input features, as depicted by the red line in the figure. Pearson's correlation coefficient between the true and predicted firmness stands at 63%.

5.5 Pesticide residue estimation

5.5.1 Capturing pesticide residue in fresh produce

Light reflectivity values across no pesticide, fungicide and insecticide treatment groups are compared in order to estimate pesticide residue in fresh produce. Wilcoxon test shows that differences in light values for the treatment groups are not statistically significant ($p > 0.05$). Figure 16 exhibits the disparities in treatment groups for each vegetable item. Kruskal-Wallis test finds no statistically significant difference among treatment groups for most vegetable items ($p > 0.05$), as seen in Table 4.

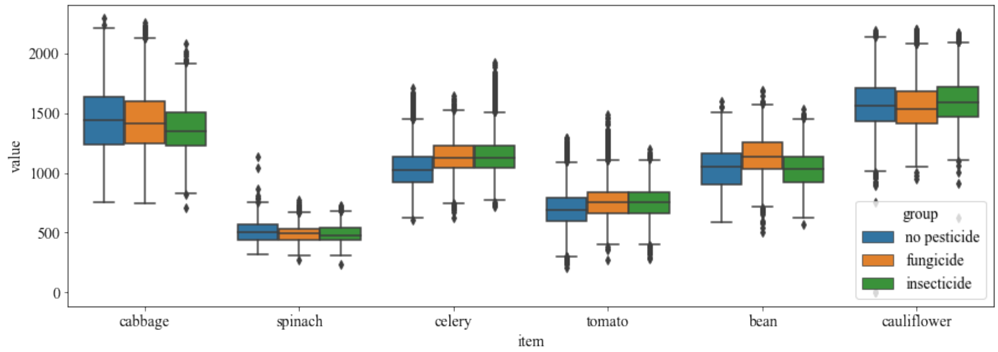


Figure 16. Item-wise light reflectivity values for no pesticide, fungicide and insecticide groups.

Table 4. Kruskal-Wallis test statistic x^2 and p-value of treatment groups for vegetable items.

	x^2	**p**
Cabbage	3.9	0.142
Spinach	1.1	0.571
Celery	7.6	0.023
Tomato	3.0	0.218
Bean	3.0	0.218
Cauliflower	0.3	0.856

Celery and bean are the only vegetables that exhibit statistically significant differences among treatment groups. Posthoc comparisons (Dunn-Bonferroni) reveal only bean's disparities for fungicide-insecticide and insecticide-no pesticide pairs that are statistically significant ($p < 0.05$). Consequently, light reflectivity values lack the ability to capture pesticide residue in fresh produce.

5.5.2 Pesticide residue over time

Figure 17 showcases the median values with a 95% confidence interval (CI) of three treatment groups for each day over the 12-day period. It highlights how the no pesticide, fungicide and insecticide light values go hand in hand for bean, spinach, tomato and celery, and somewhat less for cauliflower and cabbage.

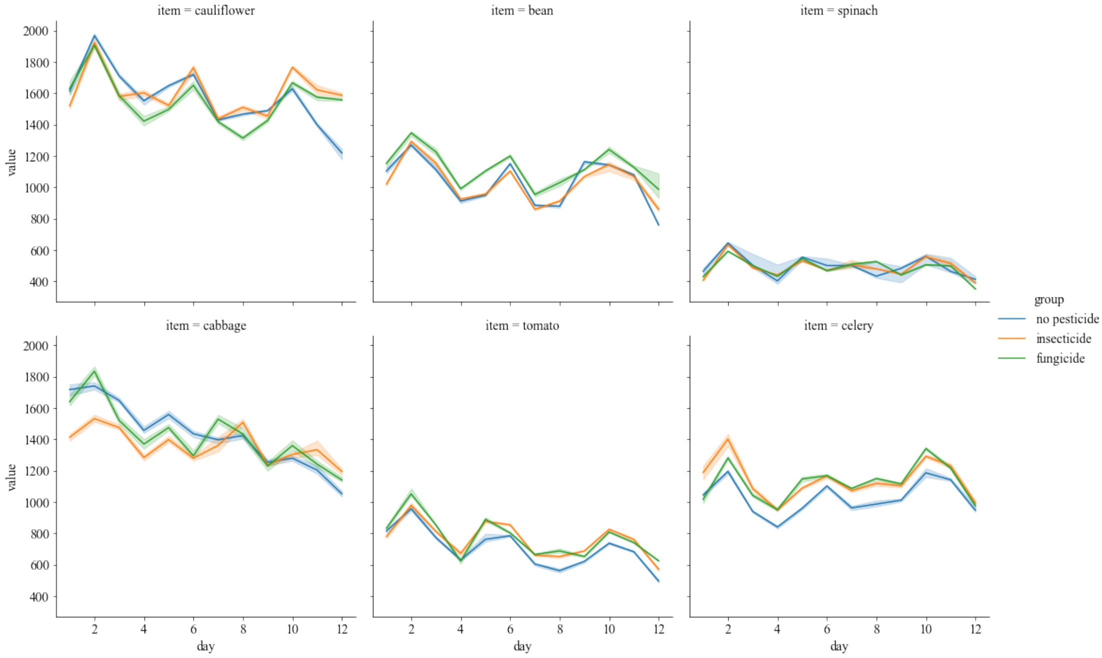


Figure 17. Day-wise median + 95% CI light values of treatment groups for vegetable items.

5.5.3 Predicting pesticide residue

Following, machine learning models are trained to predict pesticide residue in fresh produce. Light reflectivity value, vegetable item and luminosity value are taken as input features to perform classification into no pesticide, fungicide and insecticide groups. The same ten classifiers are deployed as done in machine learning for produce quality estimation. The highest accuracy is scored by XGB (53%), followed by DT (47.9%) and RF (47.7%). These accuracy scores are very low and do not make for a good classification model.

Further, the target class is divided into no pesticide and pesticide (fungicide and insecticide) for a binary classification problem. Performance accuracy sees some improvement for XGB (68.2%), RF (64.6%) and KNN (63.4%), however still remains quite low.

6 Discussion

In this chapter, the implications and limitations of the study are discussed. The areas for potential future work are deliberated.

6.1 On light sensing

6.1.1 Light source color

In this experiment, three different light sources are deployed and a single light (blue) is chosen for further analysis. The advantage of blue light is that none of the examined vegetables (cabbage, spinach, celery, tomato, bean and cauliflower) are blue in color. Since the vegetables are either green, red or white, blue wavelengths are absorbed instead of being reflected from the surface. Hence, the blue light beam is more apt to capture the inner qualities of the produce given the vegetables used in this study. However, for example to estimate the quality of eggplant, the use of red or green light could be worthier. It is, therefore, critical to consider the use of variable light based on the color of the vegetable type so that fewer wavelengths are reflected and more get absorbed.

6.1.2 Single vs multi point measurement

Another aspect to consider while measuring the light reflectivity is to prescribe how the measurements are taken. As done by Zuniga *el al* [49], one way is to place the sensor at a single point and collect the light reflectivity values from that particular spot. The drawback of this method is that produce decomposition is not even. Sporadic decayed dark spots on the surface of produce can affect the measurement, depending on whether the records are taken from that spot or not.

Another approach is to move the sensor from a fixed distance along the object, as was performed in this experiment. In such a scenario, the recorded values reflect better the overall condition of the vegetable item rather than a specific area of the object.

6.1.3 Non-invasive and contactless

The non-invasive nature of light sensing is beneficial where it is not desirable to cause damage to the produce or the method does not work with certain items. For example, given the fragile nature of the spinach leaves, it is impossible to measure the firmness of spinach by applying pressure with the durometer. Here, light sensing becomes particularly useful in estimating the level of ripeness for gentle leafy vegetables, such as spinach.

Another facet where non-invasive and contactless sensing has an advantage is hygiene. Given the recent COVID-19 pandemic, touching fresh produce at supermarkets was highly undesirable or forbidden. Hence, unlike thermal imaging through human touch interactions, light sensing does not threaten the spread of disease.

6.2 Approaches in machine learning

Two approaches to training the machine learning models are explored in this study. In the separate model classification problem, where only light reflectivity value is taken as an input feature, the top model accuracy ranges from 57.7% (celery) to 78.1% (cauliflower). The average performance accuracy of separate classifiers stands at 68.3%. Logistic regression, support vector machine and AdaBoost are the outstanding algorithms. However, the single model classification accuracy reaches 63.1% with a single input parameter and 67.3% with light reflectivity and vegetable items as input features. AdaBoost is found to be a top-performing model for a single model classification problem. Consequently, it is worth considering developing a light sensing unit that can, based on the produce type, adjust the classification model that works best for that particular item.

6.3 Limitations

Unlike the case of cauliflower, cabbage and tomato, the analysis of fresh produce decomposition reveals that spinach, bean and celery do not follow a clear trend of decreasing light reflectivity value. These are less distinguishable from ripe to decayed stages and their separate classification model accuracy is lower. In fact, these vegetable items are often kept on refrigerated shelves in the supermarket. However, in this study these are kept at room temperature, which does not resemble the real case scenario. It is therefore recommended to design a different testbed for refrigerated vegetables to imitate a real-life situation.

6.4 Future work

The work of light sensing for produce quality estimation has been so far focused on the retail stage of the produce value chain. This stage has two main implications for different stakeholders. Namely, quality estimation when buying fresh produce (customer perspective) and dynamic pricing at the retail store (retailer perspective).

After all, the same method can benefit earlier in the value chain. For example, ripeness can be estimated from raw to ripe, aiding the farmers in precise timing decisions for automated harvesting solutions. This is a new potential area for work that can be performed with raw and ripe produce, preferably without detaching the fruit or vegetable from the plant.

7 Conclusion

This thesis explored light reflectivity as a sensing modality for produce quality and pesticide residue estimation in fresh produce. Three different light sources (red, green and blue) were investigated for various vegetable items from five categories (leaves, stem, fruits, pods and flowers). All three lights were able to establish unique fingerprints for different vegetable types. Comparing two samples of the same item indicated the ability of light sensing to generalize. Moreover, with the third sample, it was established that firmness measured with a durometer did not affect produce decomposition. Based on the juxtaposition of the three light sources, the blue light was chosen for further analysis to estimate produce quality and pesticide residue.

The 12-day experiment of produce decomposition confirmed the ability of light reflectivity to capture the quality of fresh produce. This corroborates with the work on green light sensing for fruit categories and therefore endorses the generalizability of light reflectivity for produce quality estimation.

Light sensing offers a non-invasive and cost-effective method with a wide range of applications across the value chain of fresh produce. Besides, it can work for vegetable items, where taking firmness measurements are not viable, such as in the case of spinach. This non-invasive sensing technique can be contrasted with traditional methods, which are time-consuming, expensive, use harmful reagents, need expert laboratory staff and are strongly dependent on rigorously following a standardized protocol to obtain accuracy.

While predicting the quality of fresh produce, it is essential to account for varying lighting conditions, as reflectivity intensity depends on luminosity. With respect to machine learning to classify produce into ripe and decayed, one size does not fit all. Different models work best for various vegetable items for a separate model classification. AdaBoost classifier, with the input features of light reflectivity value, vegetable item and luminosity, reached an accuracy of 92.4%. For a baseline comparison, a linear regression model validated the ability to predict the firmness of vegetable items from light reflectivity value, vegetable items and luminosity as input features.

The estimation of pesticide residue using light reflectivity did not yield anticipated results. The disparities among the no pesticide, fungicide and insecticide treatment groups across vegetable items were not established. Overall, it was observed that changes in light reflectivity due to produce decomposition are major; hence estimation of pesticide residue becomes troublesome.

The future work emancipating from this thesis was proposed. This can fall in the scope of research to be expanded from retail to other stages of the produce value chain, e.g. harvesting. Furthermore, various aspects of light sensing and training of the machine learning model were highlighted.

References

[1] Antonio Albiol, Alberto Albiol, and Carlos J. Sánchez de Merás. Fast 3d rotation estimation of fruits using spheroid models. *Sensors*, 21(6):2232, 2021.

[2] Roberto Beghi, Susanna Buratti, Valentina Giovenzana, Simona Benedetti, and Riccardo Guidetti. Electronic nose and visible-near infrared spectroscopy in fruit and vegetable monitoring. *Reviews in Analytical Chemistry*, 36, 2017.

[3] Anuja Bhargava and Atul Bansal. Fruits and vegetables quality evaluation using computer vision: A review. *Journal of King Saud University-Computer and Information Sciences*, 33(3):243–257, 2021.

[4] European Commission. Annex to the communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions chemicals strategy for sustainability towards a toxic-free environment. *Brussels, European Commission*, 2020.

[5] European Commission. Ultra-compact, low-cost plasmo-photonic bimodal multiplexing sensor platforms as part of a holistic solution for food quality monitoring, 2021. Accessed on 20.02.2022.

[6] Szogyi M Cserhati T. Chromatographic determination of pesticides in foods and food products. *Journal of Nutrition & Food Sciences*, 2(2):1–10, 2021.

[7] M. Lakota D. Stajnko and M. Hočevar. Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. *Computers and Electronics in Agriculture*, 42(1):31–42, 2004.

[8] Jalal Dehghannya, Michael Ngadi, , and Clement Vigneault. Mathematical modeling procedures for airflow, heat and mass transfer during forced convection cooling of produce: a review. *Food Engineering Reviews*, 2(4):227–243, 2010.

[9] Zhenjiao Du, Xiangquan Zeng, Xiangxin Li, Xiaomeng Ding, Jiankang Cao, and Weibo Jiang. Recent advances in imaging techniques for bruise detection in fruits and vegetables. *Trends in Food Science & Technology*, 99:133–141, 2020.

[10] Hilary Emenike, Farooq Dar, Mohan Liyanage, Rajesh Sharma, Agustin Zuniga, Mohammad A. Hoque, Marko Radeta, Petteri Nurmi, and Huber Flores. Characterizing everyday objects using human touch: Thermal dissipation as a sensing modality. *2021 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 1–8, 2021.

[11] Tatyana Ermolaeva, Olga Farafonova, and Nadezhda Karaseva. Possibilities and prospects of immunosensors for a highly sensitive pesticide detection in vegetables and fruits: a review. *Food Analytical Methods*, 12(1):2785–2801, 2019.

[12] Pesticide Action Network Europe. Forbidden fruit: The dramatic rise in the most toxic pesticides found on fruits and vegetables sold in europe and evidence that governments are failing their legal obligations. 2022. Accessed on 20.07.2022.

[13] Ling Fang, Xiaofang Liao, Boyu Jia, Linchun Shi, Linzhi Kang, Lidong Zhou, and Weijun Kong. Recent progress in immunosensors for pesticides. *Biosensors and Bioelectronics*, 164:112255, 2020.

[14] FAO. The state of food and agriculture 2019. moving forward on food loss and waste reduction. *Rome, FAO*, 2019.

[15] FAO, IFAD, UNICEF, WFP, and WHO. The state of food security and nutrition in the world 2021. transforming food systems for food security, improved nutrition and affordable healthy diets for all. *Rome, FAO*, 2021.

[16] American Association for the Advancement of Science. Infographic: Pesticide planet, 2013.

[17] X. Fu and Y. Ying. Application of nir and raman spectroscopy for quality and safety inspection of fruits and vegetables: A review. *Nongye Jixie Xuebao/Transactions of the Chinese Society for Agricultural Machinery*, 44:148–164, 2013.

[18] Narendra V G and Hareesh K S. Quality inspection and grading of agricultural and food products by computer vision-a review. *International Journal of Computer Applications*, 2(1):43–65, 2010.

[19] A. A. Gowen, B. K. Tiwari, P. J. Cullen, and K. McDonnell. Applications of thermal imaging in food quality and safety assessment. *Trends in Food Science Technology*, 21(4):190–200, 2010.

[20] Klaus G. Grunert. Food quality and safety: consumer perception and demand. *European Review of Agricultural Economics*, 32(3):369–391, 2005.

[21] Weiwen He, Hongyuan He, Fanglin Wang, Shuyue Wang, and Rulin Lyu. Non-destructive detection and recognition of pesticide residues on garlic chive (allium tuberosum) leaves based on short wave infrared hyperspectral imaging and one-dimensional convolutional neural network. *Journal of Food Measurement and Characterization*, 15, 10 2021.

[22] Hui Huang, Li Liu, and Michael O. Ngadi. Recent developments in hyperspectral imaging for assessment of food quality and safety. *Sensors (Basel, Switzerland)*, 14:7248 – 7276, 2014.

[23] Xuesong Jiang, Dongyang Li, Xia Xu, Yibin Ying, Yanbin Li, Zunzhong Ye, and Jianping Wang. Immunosensors for detection of pesticide residues. *Biosensors and Bioelectronics*, 23(11):1577–1587, 2008.

[24] C. Karunakaran, D. Jayas, and N.D White. Detection of internal wheat seed infestation by rhyzopertha dominica using x-ray imaging. *ournal of Stored Products Research*, 40:507–516, 2004.

[25] Ilse Kranner, Gerald Kastberger, Manfred Hartbauer, and Hugh W. Pritchard. Non-invasive diagnosis of seed viability using infrared thermography. *Proceedings of the National Academy of Sciences*, 107(8):3912–3917, 2010.

[26] Alsharif AM Lawal A, Tan GH. Recent advances in analysis ofpesticides in food and drink samples using lpme techniquescoupled to gc-ms and lc-ms: a review. *Journal of AOAC INTERNATIONAL*, 99(6):1383–1394, 2016.

[27] Changying Li, Gerard W. Krewer, Pingsheng Ji, Harald Scherm, and Stanley J. Kays. Gas sensor array for blueberry fruit disease detection and classification. *Postharvest Biology and Technology*, 55:144–149, 2010.

[28] Pedro Canotilho Ribeiro Maria Otília Carvalho, Verônica Maria Costa Romão, and Susana Cardoso. Smart fingertip sensor for food quality control: fruit maturity assessment with a magnetic device. *Journal of Magnetism and Magnetic Materials*, 536(2):168116, 2021.

[29] S.K. Mathanker, P.R. Weckler, and T.J. Bowser. X-ray applications in food and agriculture: a review. *Transactions of the ASABE*, 56:1227–1239, 2013.

[30] S.T. Narenderan, Subramania Nainar Meyyanathan, and B. Babu. Review of pesticide residue analysis in fruits and vegetables. pre-treatment, extraction and detection techniques. *Food research international*, 133:109141, 2020.

[31] United Nations. Transforming our world: The 2030 agenda for sustainable development. *New York: UN Publishing*, 2015.

[32] S. Neethirajan, D.S. Jayas, and N.D.G. White. Detection of sprouted wheat kernels using soft x-ray image analysis. *Journal of Food Engineering*, 81:509–513, 2007.

[33] S. Neethirajan, C. Karunakaran, S. Symons, and D.S. Jayas. Classification of vitreousness in durum wheat using soft x-rays and transmitted light images. *Computers and Electronics in Agriculture*, 53:71–78, 2006.

[34] E. E. Okere, E. Arendse, H. Nieuwoudt, O. A. Fawole, W. J. Perold, and U. L. Opara. Non-invasive methods for predicting the quality of processed horticultural food products, with emphasis on dried powders, juices and oils: A review. *Foods (Basel, Switzerland)*, 10(12):3061, 2021.

[35] Hitika B. Patel, Deepak Rawtani, and Yadvendra Kumar Agrawal. A newly emerging trend of chitosan-based sensing platform for the organophosphate pesticide detection using acetylcholinesterase- a review. *Trends in Food Science & Technology*, 2019.

[36] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[37] Deepak Rawtani, Nitasha Khatri, Sanjiv Tyagi, and Gaurav Pandey. Nanotechnology-based recent approaches for sensing and remediation of pesticides. *Journal of environmental management*, 206:749–762, 2018.

[38] Pedro Canotilho Ribeiro, Susana Cardoso, Alexandre Bernardino, and Lorenzo Jamone. Fruit quality control by surface analysis using a bio-inspired soft tactile sensor. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8875–8881, 2020.

[39] M. Ruiz-Altisent, L. Ruiz-Garcia, G.P. Moreda, Renfu Lu, N. Hernandez-Sanchez, E.C. Correa, B. Diezma, B. Nicolaï, and J. García-Ramos. Sensors for product characterization and quality of specialty crops—a review. *Computers and Electronics in Agriculture*, 74(2):176–194, 2010.

[40] Lourdes Salguero-Chaparro, Antonio Jesús Gaitán-Jurado, Víctor Ortiz Somovilla, and Francisco Peña-Rodríguez. Feasibility of using nir spectroscopy to detect herbicide residues in intact olives. *Food Control*, 30:504–509, 2013.

[41] Manfred Stoll, Hans R Schultz, and Beate Berkelmann-Loehnertz. Exploring the sensitivity of thermal imaging for plasmopara viticola pathogen detection in grapevines under different water status. *Functional Plant Biology*, 35(4):281–288, 2008.

[42] Haoran Sun, Liguo Zhang, Lijun Ni, Zijun Zhu, Shaorong Luan, and Ping Hu. Study on rapid detection of pesticide residues in shanghaiqing based on analyzing near-infrared microscopic images. *Sensors*, 23(2):983, 2023.

[43] Ernest Teye, Xingyi Huang, and Newlove Afoakwa. Review on the potential use of near infrared spectroscopy (nirs) for the measurement of chemical residues in food. *American Journal of Food Science and Technology*, 1:1–8, 2013.

[44] Ye W, Yan T, Zhang C, Duan L, Chen W, Song H, Zhang Y, Xu W, and Gao P. Detection of pesticide residue level in grape using hyperspectral imaging with machine learning. *Foods*, 11(11):1609, 2022.

[45] Gao Y, Jiao L, Jiao F, and Dong D. Non-intrusive prediction of fruit spoilage and storage time via detecting volatiles in sealed packaging using laser spectroscopy. *LWT*, 155:112930, 2022.

[46] Rojas-Carillo O. Zamora-Sequeira R., Starbird-Pérez R. and Vargas-Villalobos S. What are the main sensor methods for quantifying pesticides in agricultural activities? a review. *Molecules (Basel, Switzerland)*, 24(14):2659, 2019.

[47] REN Zhan-qi, RAO Zhen-hong, and JI Hai-yan. Identification of different concentrations pesticide residues of dimethoate on spinach leaves by hyperspectral image technology. *IFAC-PapersOnLine*, 51(17):758–763, 2018.

[48] Liyong Zhou, Vikram Chalana, and Yongmin Kim. Pc-based machine vision system for real-time computer-aided potato inspection. *International Journal of Imaging Systems and Technology*, 9(6):423–433, 1998.

[49] Agustin Zuniga, Petteri Nurmi, and Huber Flores. Ripe or rotten? low-cost produce quality estimation using reflective green light sensing. *IEEE Pervasive Computing*, 20:60–67, 2021.

Appendix

I. Licence

Non-exclusive licence to reproduce thesis and make thesis public

I, **Karin Rao**,
(author's name)

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

   reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

   Produce Quality and Pesticide Residue Estimation Using Light Sensing,
   (title of thesis)

   supervised by Huber Flores.
   (supervisor's name)

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Karina Rao
09/05/2023