STATISTICAL ASSESSMENT OF BANANA RIPENING USING SMARTPHONE-BASED IMAGES

by

HARITHA R. NAIR

(2019-19-001)

THESIS

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE IN AGRICULTURE

Faculty of Agriculture

Kerala Agricultural University

DEPARTMENT OF AGRICULTURAL STATISTICS

COLLEGE OF AGRICULTURE

VELLAYANI, THIRUVANANTHAPURAM – 695 522

KERALA, INDIA

2022

DECLARATION

I, hereby declare that this thesis entitled **"STATISTICAL ASSESSMENT OF BANANA RIPENING USING SMARTPHONE-BASED IMAGES"** is a bonafide record of research work done by me during the course of research and that the thesis has not previously formed the basis for the award to me of any degree, diploma, associateship, fellowship or other similar title, of any other University or Society.

Place: Vellayani

Date: 05-02-2022

Haritha R. Nair

(2019-19-001)

CERTIFICATE

Certified that this thesis entitled **"STATISTICAL ASSESSMENT OF BANANA RIPENING USING SMARTPHONE-BASED IMAGES"** is a record of research work done independently by **Ms. Haritha R. Nair (2019-19-001)** under my guidance and supervision and that it has not previously formed the basis for the award of any degree, fellowship or associateship to her.

Place: Vellayani
Date: 05-02-2022

Dr. Pratheesh P Gopinath
(Major advisor, Advisory Committee)
Assistant Professor
Department of Agricultural Statistics
College of Agriculture, Vellayani

CERTIFICATE

We, the undersigned members of the advisory committee of **Ms. Haritha R. Nair (2019-19-001)**, a candidate for the degree of **Master of Science in Agriculture** with major in Agricultural Statistics, agree that the thesis entitled **"STATISTICAL ASSESSMENT OF BANANA RIPENING USING SMARTPHONE-BASED IMAGES"** may be submitted by **Ms. Haritha R. Nair (2019-19-001)**, in partial fulfilment of the requirement for the degree.

Dr. Pratheesh P. Gopinath
(Chairman, Advisory Committee)
Assistant Professor
Department of Agricultural Statistics
College of Agriculture, Vellayani,
Thiruvananthapuram 695 522

Dr. Brigit Joseph
(Member, Advisory Committee)
Associate Professor and Head
Department of Agricultural Statistics
College of Agriculture, Vellayani,
Thiruvananthapuram 695 522

Dr. Geetha Lekshmi P. R.
(Member, Advisory Committee)
Assistant Professor,
Department of Post Harvest Technology,
College of Agriculture, Vellayani,
Thiruvananthapuram – 695 522

Dr. Sreekumar J.
(Member, Advisory Committee)
Principal Scientist
Department of Agricultural Statistics
ICAR-Central Tuber Crops Research Institute
Thiruvananthapuram 695 017

Dr. Rafeekher M.
(Member, Advisory Committee)
Assistant Professor and Head,
Department of Floriculture and Landscape architecture,
College of Agriculture, Vellayani,
Thiruvananthapuram – 695 522

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

First and foremost, praises and thanks to the Almighty, for his blessings and kindness at each moment and to complete the research successfully.

I thank each and every one, who contributed in some or other way to the completion of this work. With great pleasure, I would like to express my heartfelt gratitude and indebtedness to my major advisor Dr. Pratheesh P Gopinath, Assistant Professor, Department of Agricultural Statistics, for his constant support, effective guidance, expert practical suggestions, whole hearted cooperation, unreserved help, patient hearing and concern which he extended throughout the course of research work. With his guidance only, I could follow the correct methodology and complete the work in an effective way. I value his knowledge which nurtured this research in the right direction.

I express my sincere gratitude to Dr. Brigit Joseph, Associate Professor and Head, Department of Agricultural Statistics, for her expert advice and guidelines, constructive criticisms, timely suggestions and constant encouragement throughout the course of work.

I am grateful to Dr. Geetha Lekshmi P. R., Assistant Professor, Department of Post Harvest Technology, for her remarkable helps, suggestions and critical evaluation of this work.

I am very thankful to Dr. Sreekumar J., Principal Scientist, Department of Agricultural Statistics-CTCRI, for his support, encouragement, advices and cooperation throughout the research work.

I express my sincere gratitude to Dr. Rafeekher M., Assistant Professor and Head, Department of Floriculture and Landscape architecture, for his remarkable help and support.

I express my gratitude to all the teaching and non- teaching staffs of Department of Agricultural Statistics for all their cooperation and help. I thank Neethu mam for constant support and help in completion of this work.

I am thankful to my classmates Anjana, Ardra and Harinath for their unquestionable help and support throughout the course. Special thanks to Jerin, Harsha, Meera, Krishenendhu, Akhila, Venkat, all my friends Gifty, Arunima and Aparnna for their support, helps and cooperation throughout the study period.

I owe my special thanks to Anusha Bindu, Vikash, Irshad chettan, Neethu miss, Lekshmi and Vaishnavi for their constant support and help in all aspects during the period.

I thank my seniors, Neethu chechi, Nayana chechi, Adarsh chettan and Gokul chettan and juniors Sandra, Keerthana and Malena for their remarkable help, support and cooperation.

I wish to acknowledge with gratitude the award of fellowship by the Kerala Agricultural University during the tenure of the M.Sc. (Ag) programme.

I am extremely obliged to my father Rajasenan, mother Remadevi, brother Harikrishnan grandma Omana amma and my puppy Vicky for their mental support, prayers, blessings and affection to complete this venture successfully.

Once again, I thank everyone and apologize if I ever missed to mention any names.

Haritha R. Nair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AIC	Akaike Criterion
ANN	Artificial neural network
ANOVA	Analysis of variance
B	Blue
CCD	charge-coupled device
CDA	Canonical discriminant analysis
CIE	Commission on illumination
CNN	Convolutional neural network
CT	computed tomography
CVS	Computer vision system
CWAD	Canada Western Amber Durum
DMC	Dry matter content
DOY	Days of year
ET	electrical tomography
et al.,	Co-workers
FAO	Food and Agriculture Organization of the United Nations
Fig.	Figure
G	Green
HRS	Hard Red Spring
HSI	Hue, Saturation, Intensity
i.e.	that is
kHz	kilohertz
MLNN	multi-layer feed-forward neural network
MRI	magnetic resonance imaging
MSNN	multi-structure neural network
NIRS	Near-infrared spectroscopy
nm	nanometer
PLS	Partial least square

R	Red
RMSE	Root mean square error
RBF	Radial basis function
SIG	Sigmoid
SSC	Soluble solids content
SWNIR	Short wave near-infrared
TSS	Total Soluble Solids

Introduction

1. INTRODUCTION

Globally bananas are the most widely consumed fruit. Bananas account for 16 percent of global fruit production. More than 114 million tons of bananas were produced worldwide in 2014, according to the FAO (Food and Agriculture Organization of the United Nations). Bananas are elliptical fruits with a solid, creamy flesh pulp wrapped in inedible skin that is naturally framed. It is one of the world's most popular fruits. Bananas are ranked fourth among the world's food crops. The fruit is rich in vitamins C and B6, potassium, fiber, tryptophan, and amino acid.

For its flavour, texture, and high nutritional content, the banana is one of the most well-known and extensively consumed fruits in the world. It's critical to keep the peel color of same-batch bananas on the shelf uniform to improve market acceptance. Bananas are typically harvested green, transported long distances, and then matured at their destination before being marketed. Bananas are often harvested when they are fully matured green, and they remain firm and green without noticeable changes in peel color, texture, or composition before ripening begins (depending on temperature, humidity, and age at harvest) (Dadzie and Orchard, 1997).

Ripening is irreversible once it starts, involving various chemical changes, changes in fruit texture, and the creation of volatile components (Drury et al., 1999). Furthermore, ripening causes non-homogeneous peel color changes, as well as the emergence of brown blotches on a yellow background. Consumers regard the color of the peel to be the first quality criteria they evaluate, and it is linked to specific tastes or purposes that can influence acceptance or rejection. In reality, peel color is linked to physical and chemical changes in bananas throughout ripening, such as pulp color (Wainwright and Hughes, 1989, 1990), pH, starch conversion to sugar, and flavor development (Ramaswamy and Tung, 1989; Ward and Nussinovitch, 1996; Chen and Ramaswamy, 2002). The synthesis of pigments (such as carotenoids) and the degradation of the green pigment chlorophyll are linked to the removal of the green hue and the resulting yellowing of the peel (Ammawath et al., 2001).

Visual inspection is used in the trade to assess the seven phases of ripening, which are related to pigment changes in the banana peel: Stage 1: green; Stage 2: green

with yellow flecks; Stage 3: more green than yellow; Stage 4: more yellow than green; Stage 5: green tip and yellow; Stage 6: all yellow; and Stage 7: yellow with brown flecks (Li et al., 1997). Visually, ripeness is determined by comparing the color of the peel to established color charts that depict the seven phases of ripening (von Loesecke, 1950; Li et al., 1997), and sometimes by instrumental methods (Wainwright and Hughes, 1990).

The stage of maturity of a fresh banana fruit has a significant impact on the fruit's quality during ripening and marketability after ripening. Farmers will benefit greatly from the capacity to recognize the maturity of fresh banana fruit, as it will assist them to optimize the harvesting phase and prevent the harvesting of either under-matured or over-matured bananas.

Some quality evaluation is still done manually by experienced inspectors in the agriculture sector, which is tedious, laborious, expensive, and intrinsically unreliable due to its subjective nature. The introduction of computer-based image processing techniques was prompted by increased expectations for objectivity, consistency, and efficiency. Computer vision, which uses image processing techniques, has recently advanced to the point that it can now quantitatively evaluate complex size, shape, color, and texture features of food. Image processing technologies are becoming increasingly significant in the assessment of food quality because they retain accuracy and consistency while eliminating the subjectivity of manual inspections. They have a wide range of applications and can be effective alternatives for human visual decision-making.

Because of the continual physiological changes that occur throughout fruit ripening, fruit quality can be assessed as a function of time, contributing to the development of empirical techniques of evaluation (Abbott, 1999). Nonetheless, ethylene has been identified as the primary hormone, in connection to certain genes, that participates in the ripening process of banana fruit until it reaches senescence (Bapat *et al*., 2010). The depolymerization processes of carbohydrates resulted in the breakdown of cell wall structure, which caused the continual and complicated changes

in physiological, biochemical, and mechanical characteristics of fruits during the ripening process (Prasanna *et al.,* 2007).

Color is a prominent descriptor and influential attribute in image analysis for agricultural products, and it typically simplifies object extraction and identification from an image. Color vision has a lot of spatial resolution, which may be used to measure the color distribution of substances. An object's color characteristics can be derived by looking at every pixel inside the object's boundaries. Color has been demonstrated to be effective for objective measurement of a wide range of food products including fruit, grain, meat, and vegetables.

Agricultural products' color contributes more to quality assessment than any other single component (Kays, 1991). Because color has been shown to correspond well with various physical, chemical, and sensory indicators of product quality, it is considered a fundamental physical feature of agricultural goods and foodstuffs. The spectral distributions of three elements determine color: the light source, the reflected light from the sample, and the observer's visual sensitivity. The International Commission on illumination (CIE) defined each of these in 1931 to stimulate the visual system using a set of primaries (*R*, *G*, *B*) and color matching functions. However, the observed light intensity of a specific object is determined by the illuminating light's intensity and spectral distribution, as well as the spectral distribution of the object's reflectivity. CIE later specified two more human-related and less illumination-dependent color measurements in 1976, namely*, L*a*b** or *CIELAB* and *L*u*v** or *CIELUV* (Robertson, 1976).

The majority of fruit quality evaluation procedures, such as determining the pulp to peel ratio and determining fruit firmness, are destructive and rely on rheological features. Furthermore, due to the difficulty of destructively assessing every unit of fruit, these approaches do not adequately monitor the quality of banana fruits during the ripening phase, using only a small number of samples as representative of the entire.

In this context, the present study is an attempt to assess the TSS content of bananas at various ripening stages using smartphone-based images. Since the ripening of banana is characterised by peel color changes, conversion of starch to sugars, fruit

softening and development of characteristic flavour, the study aims to find a useful relationship between TSS and peel color change. For the peel color change evaluation color parameters are measured. From linear and nonlinear models between TSS and color parameters, the model that best fit data are also identified. By identifying the best fitting model for TSS and color parameters, the TSS can be predicted easily.

1.1 SPECIFIC OBJECTIVES OF THE STUDY

1. To develop a suitable model to establish the relationship between Total Soluble Solids (TSS) and *L**(lightness), *a**(green-red ratios), *b**(blue-yellow ratios) values and for prediction of TSS values using *L**, *a**, *b** values.
2. To develop a protocol for accurate data collection to assess TSS content in Banana using smart-phone-based images

1.2 SCOPE OF THE STUDY

Identification of the best-fitting model between *L**, *a**, *b**, and TSS from linear and nonlinear models will help in the prediction of TSS using color parameters. So, the ripening stage of bananas can be assessed from the images captured using smartphones. Developing a protocol for accurate data collection facilitates the elimination of errors in similar studies in the future. Also, deep learning using a Convolutional neural network (CNN) can be developed to categorize images into different TSS range categories automatically.

1.3 LIMITATION OF THE STUDY

The study was conducted for a limited period, therefore the number of samples was limited. The convolutional neural network developed couldn't identify more green than yellow stage and more yellow than green stage (stage 3 and 4 from Von Losecke's color chart) due to a low number of samples and image quality variations.

1.4 PRESENTATION OF THE THESIS

This thesis is presented in five chapters namely introduction, review of literature, materials and methods, results and discussion, and summary. The first

chapter introduction is devoted to the importance, objectives, scope, limitation, and future aspects of the present study. The second chapter attempts a critical review of past work done related to the current study. The third chapter describes various statistical methods and techniques used in the study to analyze the data. The fourth chapter results and discussion presents the inferences drawn from the analysis. The last chapter summary summarizes the entire research followed by references and abstract.

1.5 FUTURE LINE OF THE STUDY

The present study is limited to one variety of bananas. This can be further extended to different varieties to understand the trend in ripening. The study can be continued further with a higher number (at least 10,000) of samples to improve the efficiency of classification using deep learning. The study can be extended to different fruits having a peel color change during ripening.

Review of literature

2. REVIEW OF LITERATURE

For any investigation, the findings of previous studies lay the path for understanding the methodologies that may be used in the current study. In addition, the previous studies fill in the gaps in the existing information and serve as the foundation for new studies. The purpose of this chapter is to critically review the literature of previous research work that is relevant to the present study. The research work of many researchers on various statistical procedures and their application to real-world situations, not just in agriculture but also in other fields, has been critically assessed under the subheadings listed below.

a. Image-based color, morphology, and maturity assessment
b. Biochemical changes during ripening
c. Color parameters and ripeness relationship
d. Physio-phenological modelling
e. Deep learning method for grading and classification

2.1 IMAGE-BASED COLOR, MORPHOLOGY, AND MATURITY ASSESSMENT

For inspecting and grading fresh-market peaches, Miller and Delwiche (1989) developed a color vision system. The red, green, and blue inputs were reduced to two-dimensional chromaticity coordinates using diffused lighting and normalized brightness. For classification, peach color was matched to normal peach maturity colors. Shearer and Payne (1990) successfully sorted bell peppers using a color image analysis system. The intensity levels of RGB pixels were mapped to one of eight possible hues. The color quantitative variables were derived from the relative hue distributions of pixels in six orthogonal viewpoints. The accuracy of evaluating bell peppers by color was up to 96 percent.

Casady *et al*. (1992) developed a trainable algorithm for inspecting soybean seed quality using a color machine vision system. Color chromaticity coordinates and seed sphericity were employed as classifying variables. To improve cracks in the egg picture without unnecessarily boosting other surface characteristics and noise,

Goodrum and Elster (1992) used the 'filter factor,' i.e., a modified unsharp filter transform, which is a Laplace transform of an image added to the same image. This procedure, followed by a contrast stretch, had excellent results. The sensitivity to translucent patches was reduced, but cracks sensitivity was raised. White and Sellers (1994) created a real-time color inspection system for detecting foreign elements on a peanut conveyor belt. The system may be trained to recognize and distinguish distinct color signatures or fingerprints on a variety of foreign materials and foods.

McCarthy *et al.* (1995) developed a method for detecting bruising in magnetic resonance images of apples that were both rapid and computerized. To discriminate between bright pixels representing the vascular system and those depicting bruising, a computationally simple thresholding approach was applied.

So and Wheaton (1996) devised a method for smoothing a binary oyster picture that involves shrinking, expanding, and closing. Small things (such as noise) were eliminated during the shrinking step, whereas the expansion procedure filled holes and concavities in the objects. Initially, the shrink and expand method was employed. The binary image was then smoothed further with a closure procedure to remove any tiny items that remained after the shrink and expand the operation, or to separate the objects from the binary image. The closure technique split a dark convex object that intersected the image into two or more parts by filling the streak at narrow places along with the streak.

Utku and Koksel (1998) applied a filter to remove possible noise from the image data for extracting wheat grain characteristics. The noise was suppressed using a nonlinear filtering approach (median filtering). The median filter approach may be thought of as a subset of the rank statistic filter, which allows the edges to be kept while filtering away the peak noise. As a result, before implementing an edge detection approach, the median filter is frequently utilized. For segmenting flaws in 'Golden Delicious' apples, Leemans *et al*. (1998) employed two types of filters: a '3×3 median filter' and a '3×3 box filter' to segment the major apple defects as much as feasible.

Ahmad *et al*. (1999) developed an RGB color feature-based multivariate decision model that includes six color features such as averages, minimums, and

variances for RGB pixel values to differentiate between asymptomatic and symptomatic soybean seeds for inspection and grading. A linear discriminant function was used to obtain an overall classification accuracy of 88 percent for asymptomatic and symptomatic seeds with the highest chance of occurrence.

Segnini *et al.* (1999) developed a computer-based video image analysis system to measure the color of potato chips in $L^*a^*b^*$ color space. The light reflection impact from the undulating surface of the chips was not significant in the color measurement since this technology is not sensitive to light intensity. The presence of unwanted spots on the chip surface was also measured. The technique's sensitivity to distinguish 'colors' was found to be in line with the human eye's capacity. The video image analysis technique was found to be useful for differentiating potato chip colors as sensitively as the human eye.

Papadakis *et al.* (2000) developed a technique to measure the color of the bottom surface of microwaved pizza by measuring the color profile at various locations on the surface using a photoshop software program. They adopted the $L^* < 60$ criteria. Using Level and Percentile, the percentage of the dark surface was calculated. Lightness, which may be translated to L^*, is the same as Level. For every range of L^*, the value of Percentile can also be found. Microwaving the pizza in contact with a susceptor can overcome the problem with the crispiness and brown color of the bottom. Using the L^* values crispness and color of the bottom surface of the cooked pizza with and without susceptor was illustrated.

Based on color attributes of scabby wheat kernels acquired by machine vision, Ruan *et al.* (2001) developed an automated system to assess the weight percentage of scabby wheat kernels. Sun and Brosnan (2003) used three phases of thresholding to analyze a pizza sauce picture. Using the RGB model, the whole pizza picture was first segmented from the white background. The segmentation of pizza sauce from the pizza base was then done by putting the HSI values in the following ranges: [220, 14], [0, 125], and [0, 200]. Finally, the light zones of pizza sauce were segmented by using the following HSI values: [2, 14], [53, 125], and [106, 200], respectively.

Abdullah *et al.* (2001) demonstrated how imaging technology may be used to provide quality control recommendations for agricultural commodities in both the visible and non-visible spectra. Machine vision technology was examined for use in color grading of oil palms in the visible spectrum. A prototype automated oil palm inspection system was successfully developed and tested, based on a previously established algorithm for color evaluation. The accuracy of reclassification varied from 84 to 92 percent, and the results show that this approach is 20 percent more consistent than human grading. Radar tomography was the equipment under investigation in the non-visible spectrum. It was utilized to map the moisture content of the grain, which is invisible to the naked eye. The presented approach was proved to be fairly effective in reconstructing homogeneous and heterogeneous media throughout a moisture range of 12–39 percent. Outside of this range, the image generated had a poor signal-to-noise ratio due to the inversion algorithm's removal of the diffraction effect.

Subedi *et al*. (2007) tested the use of short wave near-infrared (SWNIR) spectroscopy (400-1100 nm) to assess mango (*Mangifera indica* L.) fruit maturity and as a harvest timing guide to ultimate eating quality. Flesh color, dry matter content, and a visual rating of maturity were used to determine fruit ripeness. The total soluble solids concentration of extracted juice was used to determine the eating quality of fully ripe fruits. In terms of the wavelength range of SWNIR, partial least squares (PLS) regression models based on the second derivative of absorbance spectra for DM, TSS, Hunter b, and visual maturity ranking were optimized. The optimal TSS and DM models employed the same wavelength area and yielded PLS regression coefficient charts that were quite comparable. They suggested that the models were unable to differentiate between soluble and insoluble carbohydrates in the fruit. DM and Hunter b models based on numerous harvest dates were satisfactory when employed in population prediction. Calibration models on TSS of mature fruit helped predict an independent population (R=0.92 with SEP 0.67 and bias= 1.25 percent TSS) created using SWNIR spectra obtained (non-destructively) of hard green mango (R=0.90). They concluded that the SWNIR approach may be utilized to measure fruit maturity (as flesh Hunter b or percent DM) and forecast future TSS of fruit after ripening at the time of harvest.

Due to the tiny viewing area of the equipment, measuring and reporting heterogeneous fruit color changes throughout ripening is difficult with the instruments provided (chromometer and colorimeter). Another way for capturing and quantifying whole fruit color features is to use calibrated computer vision systems (CVS). CVS errors are caused by product curvature, according to studies. It was confirmed in this study that 55 percent and 69 percent of the measured a^* and b^* color values on a curved surface were within the range observed for the same flat surface. The device's use in recording descriptive color data throughout fruit development is illustrated using 'B74' mangoes (Kang *et al.*, 2008).

According to Padda *et al.* (2011), the a^* and b^* values of the flesh were the greatest color qualities to follow changes throughout the early stages of 'Keitt' mango ripening, while the b^* value of the peel was regarded superior during later stages of ripening. The best techniques to measure changes in fruit during ripening were the penetrometer, flesh a^* value, and total soluble solids content, according to a combined Canonical discriminant analysis (CDA) that included the best ways to assess each ripening feature as well as titratable acidity (from filtered juice from whole fruit.

Intaravanne *et al.* (2012) proposed and demonstrated a two-dimensional banana ripeness level sensor utilizing a cell phone for the first time. They developed a portable design consisting of a smart mobile phone, white light sources, and ultra violet light sources. The 2-D spectral analysis for estimating the banana ripeness level was based on the fact that the color of the banana changes from green to yellow during ripening and can be observed under white light illumination. The blue-green luminescent spectrum in a wavelength band of 400-500 nm with an excitation at 350 nm wavelength trends to reduce when the banana goes from yellow to brown-yellow. They chose color ratios from two broad-spectral images under white and ultraviolet illuminations and were utilized to classify the complete banana into immature, ripe, and overripe zones spatially and specifically.

Chaudhary and Prajapati (2014) stated that the categorization of bananas can be improved by digital image processing. All of the algorithms that have been examined and are now available can accurately estimate the quality of a single banana, but when

they applied them to a bunch of bananas, either the picture capture setup was inadequate or the algorithm's performance degraded. Based on their findings, they concluded that picture capture is the most essential phase in the process. The most popular color space employed by fruit categorization is a common color space, and even characteristics do not vary much. When color and texture characteristics were included, the classification procedure yielded the best results. Adding additional factors to the analysis will not necessarily improve the outcome, which is why it is critical to employ a suitable combination of factors.

Watanawan *et al.* (2014) used near-infrared spectroscopy (NIRS) to successfully forecast mango fruit ripeness features. NIRS readings were highly linked with hardness and dry matter content (DMC) at harvest, and they accurately predicted total soluble solids (TSS). Except for fruit weight and L^* values, which exhibited no or poor connection with NIRS values, other quality metrics had lower predictability values or R^2 ranging from 0.70-0.84. The optimal harvest maturity was found to be 105-112 DAFB, indicating the necessity to employ the current harvest maturity recommendation's upper limit (91-105 days from full bloom) to guarantee strong internal qualities as well as a pleasing fresh look. They concluded that the NIRS facility might be integrated into a continuous fruit packaging process as part of a quality assurance system.

Ali *et al.* (2018) tested the principle and use of NIR spectroscopy in the region of 1000 to 2500 nm, for predicting the Brix and pH values of bananas at different maturation stages. With R values of 0.81 and RMSEP values of 3.90, the PLS (Partial least squares) models demonstrated a strong connection between the reference and projected Brix values. The absorbance spectra linked with pH had R^2 values of 0.69 and RMSEP values of 0.36 for the prediction accuracy of pH values. These findings showed that NIR spectroscopy may be used to estimate sugar concentration in bananas without destroying them. The non-destructive method of evaluating fruit quality at the maturity stage may be able to overcome the limitations of traditional analysis and prevent post-harvest losses. Aside from that, the prediction models may be used to create a portable online machine system for sorting and grading diverse agricultural products.

Hufkens *et al.* (2019) reported that near-surface remote sensing can be used to quantify physiologically important phenological stages in agricultural crops in highly heterogeneous settings, such as smallholder farming, where satellite remote sensing-based vegetation indices fail to capture the start of wheat heading and the end of the tillering phases. They also showed that near-surface imaging can record agronomically significant crop damage occurrences (such as wheat lodging) that are difficult to identify with moderate or coarse spatial resolution satellite data. Information about crop phenology and damage has a lot of potential for improving estimates of heterogeneous agricultural productivity in smallholder systems, as well as for facilitating the delivery of interventions like crop insurance to help farmers cope with the financial risks posed by extreme weather events. Furthermore, by providing high-resolution ground-truth data not currently accessible through national crop cut surveys, these picture recordings potentially cover a vital information vacuum

2.2 BIOCHEMICAL CHANGES DURING RIPENING

Williams *et al.* (1989) conducted a study on peel samples of the banana fruit obtained at 3-week intervals during a 6-month fruit growth cycle using Scanning electron microscopy. They examined the changes in epidermal cells and the epicuticular surface and observed that before the immature bunch emerged from the pseudostem, cell divisions took place. There were no obvious epicuticular structures on the fruit surface before bunch emergence. Unstructured ridges appeared on the epidermal cells within 21 days of bunch emergence and were coated with epicuticular wax. During the development of the fruit, cells on the surface expanded. The ridges on the epicuticular surface of the banana fruit flattened three weeks after fruit harvest (168 days from bunch emergence), and the cells expanded parallel to the circle of the fruit.

The electrical characteristics of banana fruit were investigated to establish a quick and non-destructive evaluation method and to manage the ripening process. To assess the difference in capacitance generated by the introduction of a banana fruit between the plates, a 5 V sine wave AC power source and a rectangular parallel plate capacitor sample were utilized. An equivalent capacitor was created to eliminate the influence of the air gap between the plates. The relationship between dielectric constant

and banana fruit quality indicators was explored. As a result of the ripening procedure, the dielectric constant of banana fruit dropped. Experiments revealed that the optimal sine wave frequency for predicting the amount of ripeness is 100 kHz. At this frequency, the coefficient of determination (R^2) of ripeness level prediction was 0.94. The maturity degree of banana fruit may be accurately predicted using this approach (Alimardani *et al.*, 2011).

Tapre and Jain (2012) investigated the physic, chemical, and mechanical characteristics of bananas (Musa sp var 'Robusta') at three advanced phases of maturation (stages 5, 6, and 7). As the fruit ripened, numerous physical changes were detected in the fruit, including an increase in the pulp to peel ratio, a decrease in the degree of peel greenness, and a drop in polyphenol oxidase activity. From stage 5 to stage 7, mechanical characteristics deteriorated considerably. At different phases of banana ripening, there was a significant variation in firmness. During the various phases of ripening, a similar tendency was found for additional mechanical characteristics such as cohesiveness, chewiness, fracture force, and stiffness. They observed that from stage 5 to stage 7, the moisture content, titratable acidity, pectin content, total sugar, and TSS of the pulp increased, but the starch content dropped.

Jamaludin *et al.* (2014) stated that the dielectric constant of banana fruit decreased as a result of the ripening process. The parameters used to characterize materials were dielectric constant, permittivity, dielectric loss, and impedance. Experiments have indicated that 100 kHz is the best sine wave frequency for forecasting ripeness. The coefficient of determination (R^2) of ripeness level prediction was 0.94 at this frequency. This method may be used to reliably forecast the maturity of banana fruit.

Sankhe and Bhosale (2015) designed a low-cost gadget to forecast the maturity degree of banana fruit. The dielectric constant of a banana is used to calculate its ripeness level. The developed method is capable of accurately predicting the maturity degree of banana fruit. Capacitance was directly affected by sample dimensions. A pair of parallel plate capacitor sensors were used in a capacitance sensing system. The efficacy of this approach to forecasting the quality of bananas during the ripening phase

was investigated in this study. The results revealed that this dielectric property-based approach may detect changes in the quality characteristics of banana fruits during the ripening phase. The maturity of any uniformly sized fruit is precisely proportional to capacitance across the fruit, according to this experiment.

Zulkifli *et al.* (2016) reported the physicochemical features of *Musa Acuminata cv.* Berangan at different ripening stages in the updated form: (1, 2, and 3). Color, for example, is prone to misinterpretation since it might be interpreted differently by various people. As a result, a variety of tests were carried out to identify the association between ripening phases and physicochemical attributes of the fruit, such as color, pH, TSS, and firmness. As the ripening phases progressed, substantial changes in physicochemical qualities and chemical attributes were observed. As the ripening phases progressed, substantial changes in physicochemical qualities and chemical attributes were observed. In comparison to other metrics, the associations between ripening phases and hardness and pH levels were statistically significant. They confirmed that the quality attributes of Berangan bananas can be predicted using these correlations.

Mohapatra *et al.* (2017) designed a dielectric property measuring device and used it in this study to detect changes in red banana dielectric properties during the ripening process at various ripening temperatures. The potential of this approach to predict the ripening assessment of bananas during ripening treatment was investigated in this research. The results demonstrated that this dielectric property-based approach is capable of estimating changes in the quality characteristics of banana fruits during the ripening phase. They confirmed that as the ripening phases of red bananas progress, the values of attributes such as capacitance and relative permittivity grow continually, whilst impedance and admittance gradually decrease.

2.3 COLOR PARAMETERS AND RIPENESS RELATIONSHIP

L^* denotes the lightness factor, while a^* and b^* denote the two-color axes, with a^* denoting the red-green axis and b* denoting the yellow-blue axis. The results for L^* indicate the same pattern as those for the Munsell value, which is to be expected. The findings for the b^* values indicate the same pattern as the Munsell chroma. During

ripening, the a^* value steadily rises, correlating to the degradation of chlorophyll and the loss of the peel's green color which is similar to the Munsell hue. Except for starch and water, the correlation coefficients between Munsell hue and all of the chemical parameters measured were highly significant ($P< 0.01$). Except for water, a^* provided higher correlation coefficients that were highly significant in all cases. The correlation coefficients for value and L^* were similar, but not very high. The correlation coefficients for b^* were greater than the chroma correlation coefficients. There were high and extremely significant associations ($P< 0.01$) between color, a^*, and the visually evaluated ripening stage. The relevance of color as an indication of banana maturity is depicted by correlation coefficients involving hue and a^*, which represent the accuracy of the visual assessment. Because the correlation coefficients between peel gloss measured at 85° and each of the following parameters: color, a^*, pH, °Brix, starch, water, visual stage of ripeness, and roughness were so high, this measurement might be useful for both monitoring and forecasting ripening. The fact that gloss was highly associated with hue and a^* was of great significance (Ward and Nussinovitch, 1996).

Sangwine (2000) reported that the CIELAB space is for reflected light (from surfaces or materials that have been illuminated by a light source), whereas the CIELUV space is for emitted light (such as from light-emitting electronic displays). These have a brightness component (similar to luminance) and two chrominance components. Perceptual uniformity is a key feature of these spaces: each unit distance in the space corresponds to a uniform perceived difference in color. This makes them helpful for exact colorimetry (which is what they were designed for), and these spaces have been used in image processing applications to quantify departure from a particular color (for example, in industrial inspection for color quality control). For applications where the results must match human perception, the CIELAB color space is required.

Vizhanyo and Felfoldi (2000) reported that simple cluster analysis was insufficient to distinguish disease-induced browning from natural browning of the mushroom; they found that converting RGB values to a^* and b^* color components and removing intensity resulted in a significantly better separation for mushroom diseases. In order to accentuate color variations in true-color images of diseased mushrooms,

intensity normalization and image transformation techniques were used. All of the diseased patches were labelled as "diseased," but none of the healthy, senescent mushroom sections were labelled as "diseased."

Pixel pre-processing is a simple but significant image processing method that turns an input picture into an output image by converting each output pixel to the same coordinates as the input pixel. Local pre-processing methods, which are also known as filtration, employ a tiny neighbourhood of a pixel in an input picture to produce a new brightness value in the output image (Du, C.-J. and Sun, D.-W., 2004).

Mendoza *et al.* (2006) conducted a sensitivity analysis of the implemented computer vision system (CVS) utilizing bananas and demonstrated that when the color of the background was changed from black to white, the color measurements of lightness (L^* values) were dramatically modified. Furthermore, the color study of samples with curved surfaces demonstrated that the $L^*a^*b^*$ is better suited to the color representation of surfaces or materials illuminated by a light source. The degree of curvature, shadows, and glossiness of the surfaces were less impacted by these color profiles than the RGB and HSV color systems, making them more acceptable for color measurements of food surfaces. For 125 color sheets, the correlation coefficients for XYZ values derived from calibrated CVS and colorimeter demonstrated acceptable color matching with an $R2 > 0.97$ in all regressions.

Juncai *et al.* (2015) observed color change in L^*, a^*, and b^* values of the stalk, middle, and tip sections of bananas followed a similar pattern during ripening, it was sufficient to explain the ripening procedure using color values of middle parts. Only the a^* values were monotonically growing along with L^*, a^*, and b^* values from ripening stage 1 to ripening stage 7. The a^* value grew clearly from ripening stage 1 to ripening stage 7, while the a^* value increase from ripening stage 5 to ripening stage 6 was not as obvious, making it difficult to determine banana ripening stages based on a^* value alone. They used a support vector machine method to classify the ripening stages by color value L^*, a^*, and b^* as input data.

Meng-han *et al.* (2015) claimed that the three color channel thresholds from B, L^*, and b^* were utilized to create an algorithm for automated picture segmentation of

bananas in a crate. The qualitative validation revealed that the outlines of manually and automatically segmented sections were very comparable. The results of the quantitative assessment revealed that the average area ratio of the 10 evaluated samples was greater than 80 percent, indicating that the automatic segmentation algorithm's performance is satisfactory.

The study conducted by Shamili (2019) revealed that the chemical makeup of mango fruit was influenced by holding temperatures and storage durations. Fruit TSS could be accurately estimated using normalized a^* values produced from converted digital pictures. The results of this investigation demonstrated that image parameters may be used to accurately estimate TSS. According to regression analysis, the polynomial model better described the TSS than other models. Furthermore, when employing Normalized a^* for TSS estimate at 5 °C, the polynomial regression exhibited a sufficient R^2 (0.98). The findings may be utilized to construct a solid method for forecasting the quality and grading mangos stored at low temperatures, as well as simplifying TSS monitoring.

Moreno *et al.* (2020) identified simple, objective, and repeatable assays to characterize the ripening process of bananas and plantains. Non-destructive measures like central diameter and color parameter a^*, in particular, were shown to be highly associated with, and hence predict, quality indicators of the ripening stage like TSS and total starch of the pulp.

2.4 PHYSIO-PHENOLOGICAL MODELLING

Carvalho *et al.* (2008) worked with the evaluation of the maturation stages of Prata bananas; which featured an increase in moisture and soluble content while a decrease in starch content during maturation. The results of the response surface methodology were adjusted using a quadratic model to produce the best total soluble solids extraction. These optimized values were compared to the model's anticipated values, suggesting that the model accurately represented the region analyzed and fit the experimental data. Furthermore, the assays carried out on scaled-up equipment proved the efficiency and practical applicability of the technique investigated in this study.

Richardson *et al.* (2015) concluded that in some instances, spline approximation might be a feasible alternative to adaptive machine learning. Using cubic spline approximation has various advantages. Cubic splines offer a predictable behavior by generating a continuous, twice differentiable approximation between data points. Computing the coefficients for a spline can be time-consuming, but it has the advantage of not requiring any training. Splines also require a small number of data points to create a decent estimate. Splines, of course, have disadvantages. Because the solution is global, finding the coefficients of the complete curve is required before a single point can be approximated. The conventional spline approach used here does not work well with noisy data. Finally, while cubic spline can be assessed quickly, the memory and computational costs become prohibitive when the number of input dimensions is enormous.

Parker *et al.* (2020) demonstrated that for *Vitis vinifera L.*, a temperature-based model may be used to forecast the time to varying target sugar concentrations. Furthermore, two models, the best Sigmoid model (best SIG) and the GSR model, have been created successfully (model efficiencies of 0.5–0.6). These models were used to successfully characterize a wide range of cultivars for the time to various sugar goal concentrations, resulting in the most comprehensive classifications of the grapevine's thermal time to target sugar concentrations to date. The two best-fit models for the Days of year (DOY) to reach the target sugar concentrations were chosen using the Akaike Criterion (AIC) (which evaluates model complexity and goodness of fit in a single criterion) and evaluated for model efficiency (EF) and prediction error (RMSE, root means squared error), followed by a sensitivity analysis and model validation.

2.5 DEEP LEARNING METHOD FOR GRADING AND CLASSIFICATION

Nakano *et al.* (1992) developed a technique for classifying the quality of the exterior look of apples using a neural network. With 27 units in the input layer, 10 units in the hidden layer, and three units in the output layer, the model was a three-layer neural network. The findings of the experiment showed that the classifier could categorize apple quality into three groups based on their external appearance.

A machine vision system was trained to discriminate between green and yellow 'Golden Delicious apples and good and greened potatoes. The HSI (Hue, Saturation, and Intensity) color system was found to be very useful for color evaluation and image processing. By expressing features using hue histograms and using multivariate discriminant techniques, the vision system was able to inspect potatoes and apples with over 90% accuracy. The vision system misclassified more when the number of hue bins was reduced by selecting significant characteristics or summing groups of hue bins. Color classification is a significant way for evaluating quality features that should be included in a larger automated quality inspection and grading system (Tao *et al.*, 1995).

Ghazanfari *et al.* (1996) suggested a multi-structure neural network (MSNN) classifier that comprised of four parallel discriminators to identify four classes of pistachio nuts. Each discriminator used physical features taken from the nuts' photos as input and was a feed-forward neural network with two hidden layers and a single-neuron output layer. When compared to the performance of a multi-layer feed-forward neural network (MLNN) classifier, the MSNN classifier had an average classification accuracy of 95.9 percent, a gain of over 8.9 percent over MLNN.

On the color grading of apples, Nakano (1997) used two neural network models. One is used to categorize pixels, and it has a more than 95 percent accuracy rate. Another neural network was created that can grade an apple's entire surface color into 'superior', 'excellent', 'good', 'poor color', and 'injured'. The grading accuracies for 'superior', 'poor color', and 'injured' were all extremely high, but not so much for 'excellent' and 'good'.

For the categorization of X-ray pistachio nut images, Casasent *et al.* (1998) employed a novel neural network. This neural network produced higher-order decision surfaces with fewer hidden layer neurons than other classifiers. It used new techniques to choose the number of hidden layer neurons, as well as adaptive algorithms that prevent ad hoc parameter selection difficulties, allowing the best classifier parameters to be chosen without having to analyze the test set results. The classification results suggested that serious defects may be reduced to 2 percent, with just 1 percent of excellent nuts being discarded. Luo *et al.* (1999) used a multilayer neural network

classifier to classify cereal grains as well as healthy and six types of damaged Canadian Western Red Spring wheat kernels using chosen morphological and color features derived from the grain sample images.

Kim and Schatzki (2000) used a neural network classifier to categorize apples into three different watercore levels: clean, mild, and severe using eight features retrieved from an X-ray scanned apple picture. The findings revealed that the system successfully classified apples into clean and severe categories with an error rate of 5-8 percent.

Visual inspection and the construction of a machine-reading method, as well as discriminant analysis of the data collected, were used to assess the color of muffins. Light and dark-colored muffins were segregated using a classification system. The accuracy of the system was tested by comparing the color of 4 cm diameter muffins that had been pre-graded and those that had not been pre-graded. When tested on 200 samples, the automated system successfully classified 96 percent of pre-graded muffins and 79 percent of ungraded muffins. In most situations, the algorithm technique was able to categorize muffins with an accuracy of greater than 88 percent, although inspectors' quality choices varied by 20 to 30 percent. When comparing inspectors with machine vision, the results of the vision system revealed that there is a possibility for a considerable discrepancy in quality evaluation across inspectors, resulting in a larger number of misclassifications. Misclassification with machine vision, on the other hand, was consistently low, occurring only in the majority of samples with somewhat poor discriminant scores (Abdullah et al., 2000).

Paliwal *et al.* (2001) studied the classification accuracy of nine different neural network architectures using five different types of cereal grains i.e., Hard Red Spring (HRS) wheat, Canada Western Amber Durum (CWAD) wheat, barley, oats, and rye. Eight morphological parameters were extracted and utilized as input to the neural networks for each kernel: area, perimeter, length of the main axis, length of the minor axis, elongation, roundness, Feret diameter, and compactness. HRS wheat, CWAD wheat, and oats had classification accuracies of over 97 percent, whereas barley and rye had accuracies of around 88 percent.

Using backpropagation neural networks, Kavdir and Guyer (2002) used backpropagation neural networks to classify Empire and Golden Delicious apples depending on their surface quality characteristics. With pixel grey values and texture features derived from the full apple image as input, a 2-class and a 5-class classification were performed. In the 2-class classification, classification success ranged from 89.2 to 100 percent. Empire apples had a classification success rate of 93.8 to 100 percent in the 5-class classification, whereas Golden Delicious apples had a classification success rate of 89.7 to 94.9 percent.

Blasco *et al.* (2003) developed machine vision techniques for online quality estimation of oranges, peaches, and apples, as well as to assess their efficacy in terms of the following quality attributes: size, color, stem location, and detection of external blemishes. Fruits could be identified from the background using a segmentation approach based on Bayesian discriminant analysis. As a result, the problem of determining size was successfully handled. The colorimetric index values that are presently used as standards were well correlated with the colors of the fruits assessed by the technique. Positive findings were obtained in the location of the stem and the identification of blemishes. On-line testing of the classification system with apples yielded good results when categorizing the fruit in batches, as well as 86 and 93 percent repeatability in flaw identification and size estimate, respectively.

According to Du and Sun (2004), the most common sensor techniques utilized in image acquisition for food quality evaluation are charge-coupled device (CCD), magnetic resonance imaging (MRI), Ultrasound, computed tomography (CT), and electrical tomography (ET). To increase the quality of an image for subsequent processing, two image pre-processing approaches may be used: pixel pre-processing and local pre-processing. For pixel pre-processing, color space transformation is the most popular method, and image smoothing techniques are important for local pre-processing. The four basic strategies used to segment food products are thresholding-based, gradient-based, region-based, and classification-based approaches. The four most popular classes used to quantify the object attributes in a food product photograph are size, shape, color, and texture. To do classification, three primary approaches are now used: statistical, fuzzy logic, and neural networks. Statistical methods are

distinguished by the presence of an explicit underlying probability model, whereas fuzzy classification employs fuzzy logic principles and neural networks mimic biological nerve systems.

Adebayo *et al.* (2016) investigated the possibility of exploiting the optical features of bananas to predict quality and categorize bananas into ripening stages 2 to 7. Backscattered pictures of bananas at various stages of ripening were used to extract absorption, decreased scattering, and effective attenuation coefficients using five laser wavelengths of 532, 660, 785, 830, and 1060 nm. For the prediction of quality variables and categorization of bananas, ANN models were utilized. The study discovered a link between the optical qualities of bananas and their ripening phases. The absorption coefficient was found negatively correlated with ripening stages, but the reduced scattering coefficients were positively correlated with ripening stages. They reported that the optical characteristics of bananas, such as absorption, reduced scattering, and effective attenuation coefficients can be a good and effective means of forecasting banana quality features and categorization into distinct ripening phases. They found that when the absorption and reduced scattering coefficients were used, the visible wavelength region of 532, 660, and 785 nm gave the highest correlation coefficient (R) range of 0.9768-0.9807 for chlorophyll prediction and 0.9553-0.9759 for elasticity prediction, while the near-infrared region of 830 and 1060 nm gave an R range of 0.9640-0.9801 for prediction of the soluble solids content (SSC).

Sanaeifar *et al.* (2016) implemented a computer vision system to analyze the color of bananas in *RGB*, *L*a*b**, and HSV color spaces, and variations in color characteristics of bananas over time were used to predict quality indicators quantitatively. The radial basis function (RBF) was used as the kernel function of support vector regression (SVR), and color characteristics from several color spaces were chosen as the model's inputs, with total soluble solids, pH, titratable acidity, and hardness calculated as the output. When compared to data produced using an artificial neural network, the experimental results showed an improvement in predictive accuracy (ANN).

Hamza and Chtourou (2018) designed and applied an artificial neural network classifier to estimate the ripeness of apple fruits based on color. The proposed method consisted of four steps. In the first phase, a region of interest was extracted using a threshold segmentation method and certain morphological processes. Color-based characteristics were extracted from the segmented apple images and separated into training and testing data in the second stage. The final stage was to determine the classifier training parameters. The classification was completed in the final stage using the trained ANN. The simulation data set was collected and exploited for the training and testing phases: 80 percent of the total images were utilized for training and 20 percent of the total images were used to test the classifier. The training dataset is divided into three classes, each representing one of the three stages of apple ripening.

Mazen and Nashat (2019) explored an artificial neural network-based system to classify the maturity stage of banana fruits. To appropriately distinguish between the four banana fruit classes, the proposed model employs Tamura's texture properties as well as a novel feature known as the ripening factor. When compared against other supervised classification algorithms such as the SVM, naive Bayes, KNN, decision tree, and discriminant analysis classifiers, the system performs best. The green and overripen classes have a class recognition accuracy of 100 percent, while the yellowish-green and mid-ripen classes have a class recognition accuracy of 97.75 percent. The classification model's simplicity, high identification rate, and speed (18s for the 89 test bananas) make it ideal for creating a productive and lucrative computer vision machine in the food processing industry.

Kamble *et al.* (2020) proposed a system to classify and then detect the fruit ripening stage using CNN. The ripening stages of three fruits i.e., mango, apple, and banana, were recognized and categorized. Classification accuracy for banana was 96 percent, whereas ripening stage accuracy of banana was 88 percent. This system has the advantage of being generic for three fruits, but it also has the disadvantage of being complex. Other fruits that do not undergo physical changes as they ripen are ineligible for use in this system. The technique also has the drawback of being limited to only three fruits.

Sri *et al.* (2020) used several toolboxes including Statistics and Machine Learning toolbox, Neural Network toolbox, and Image Processing toolbox for constructing an algorithm. They obtained the training data in the form of image categories, image classification using CNN, and prediction. To compensate for the over-fitting effect, the original dataset was augmented using data augmentation techniques such as translations. A positive image and a negative picture were selected from the collected photos to generate a triplet for each image. Eighty percent of the photos acquired were used in the training phase, while the remaining twenty percent were used in the testing process. For perfect accuracy, they used Unsupervised Learning. They focused on Unsupervised Learning and used the CNN algorithm to obtain 92 percent training accuracy and 80 percent testing accuracy.

Materials and Methods

3. MATERIALS AND METHODS

The aim of the present study is to develop a suitable model to establish the relationship between Total Soluble Solids (TSS) and *L**(lightness), *a**(green-red ratios), *b**(blue-yellow ratios) values, and for prediction of TSS values using *L**, *a**, *b** values also to develop a protocol for accurate data collection to assess the TSS content in banana using smart-phone based images. For the model fitting *R, G, B*, *L**, *a**, *b** color parameters obtained from the smartphone images and TSS measured from banana fruit pulp were used. This chapter describes the procedural details in selecting the sample, methods of analysis in the upcoming subheadings.

3.1 Sample collection

3.2 Image collection and pre-processing

3.3 TSS measurement

3.4 Color analysis

3.5 Machine learning

3.6 Statistical analysis

3.7 Deep learning

3.1 SAMPLES OF BANANAS

As the first step of image collection, Nendran fields in the Trivandrum district were identified. Two fields were randomly selected from the sampling frame. Good quality Nendran bananas with only minor shape and peel color flaws were obtained using the following procedures.

- Healthy banana plants with good-quality bananas were identified randomly.
- From each of the identified plants, three hands (Ripening stage 1; green) with 10 fingers by hand were collected.
- The bananas were stored in a normal day/ night cycle.
- Bananas were taken randomly from each hand and their ripening stage changes were identified by observing color changes and the development of brown spots daily during 10-12 days. Plate 1 represents the stages of ripening of bananas.

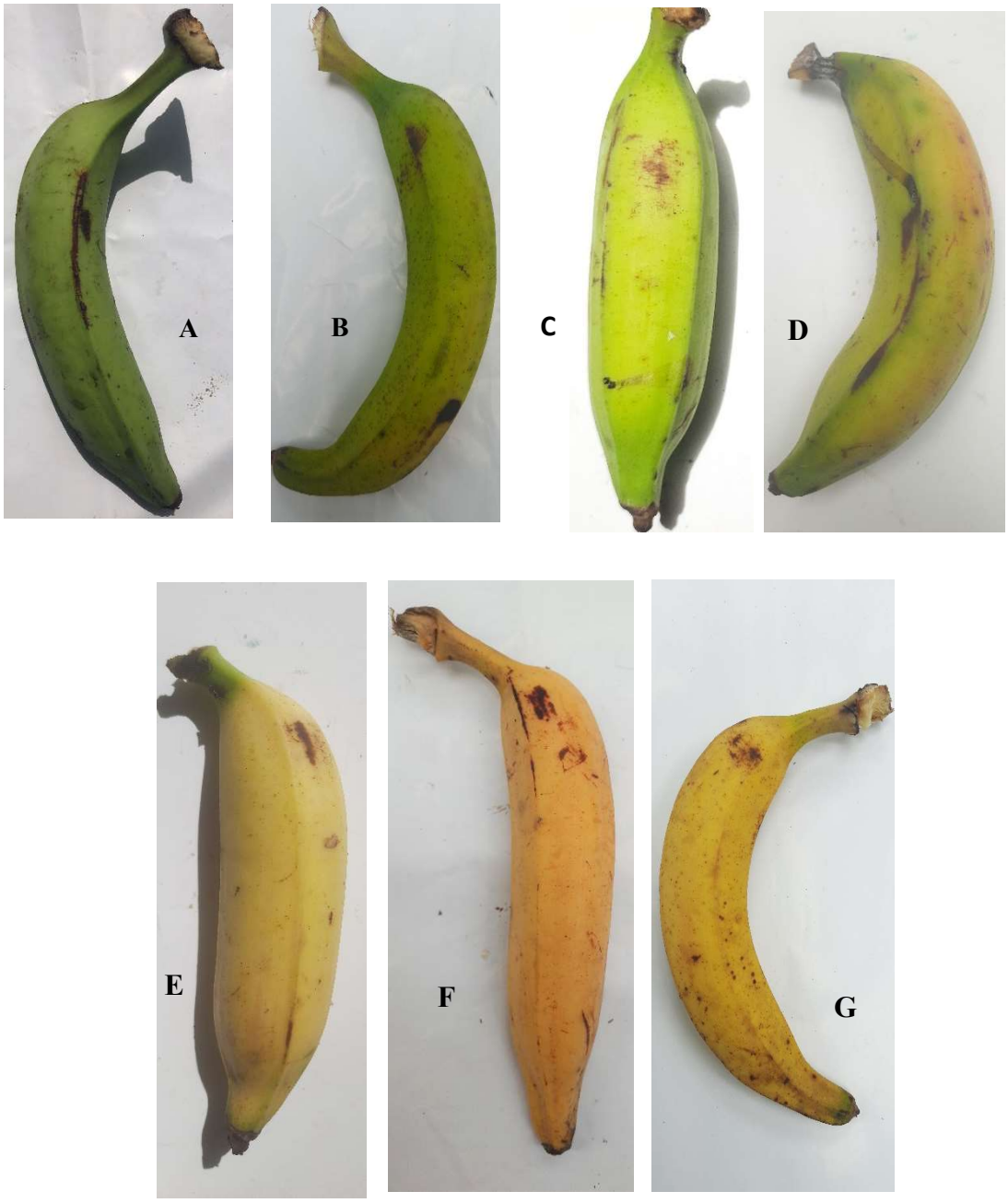

Plate 1. Stages of banana ripening (A) Stage 1-Totally green, (B) Stage 2-Green with yellow traits, (C) Stage 3- More green than yellow, (D) Stage 4- More yellow than green, (E) Stage 5-Yellow with green tip, (F) Stage 6- Yellow, (G) Stage 7-Yellow with brown spots

3.2 IMAGE COLLECTION AND PRE-PROCESSING

3.2.1 Image acquisition

For a good quality image for further processing, illumination is a critical component. The lighting situation can have a significant impact on the image quality obtained. A high-quality image can assist to cut down on the time and complexity of the image processing processes. A proper portable lighting box to keep bananas developed for proper image acquisition as given in plate 2.

Since the basic purpose of the study is to collect the images using smartphones and form a background for developing an application in the future following smartphones with different resolutions and camera quality were selected for the study. Smartphone specifications are given in Table 1.

Table 1. Smartphones used and their specifications

Sl. No.	Devices used	specifications
1	Samsung Galaxy M30s	Processor- Samsung Exynos 9611 4 GB RAM 48 MP rear camera 6000×8000 pixel Android 9 Pie operating system
2	Apple iPhone 8 plus	Processor- Apple A11 Bionic 3 GB RAM 12 MP rear camera 4032×3024 pixel iOS 11 operating system
3	Samsung Galaxy M31	Processor- Samsung Exynos 9611 6 GB RAM 64 MP rear camera 6936×9248 pixel Android 10 operating system

Banana samples were placed on the table covered with a non-reflecting white paper as a background of the image and the bananas were placed in a box model structure (Plate 2) to avoid outside light interference. For white light illumination, two of 36 W fluorescent lamps were fixed at ceiling above the experiment setup. Three smartphones were used for image acquisition. Smart phones were placed at a distance of 20 cm above the banana.

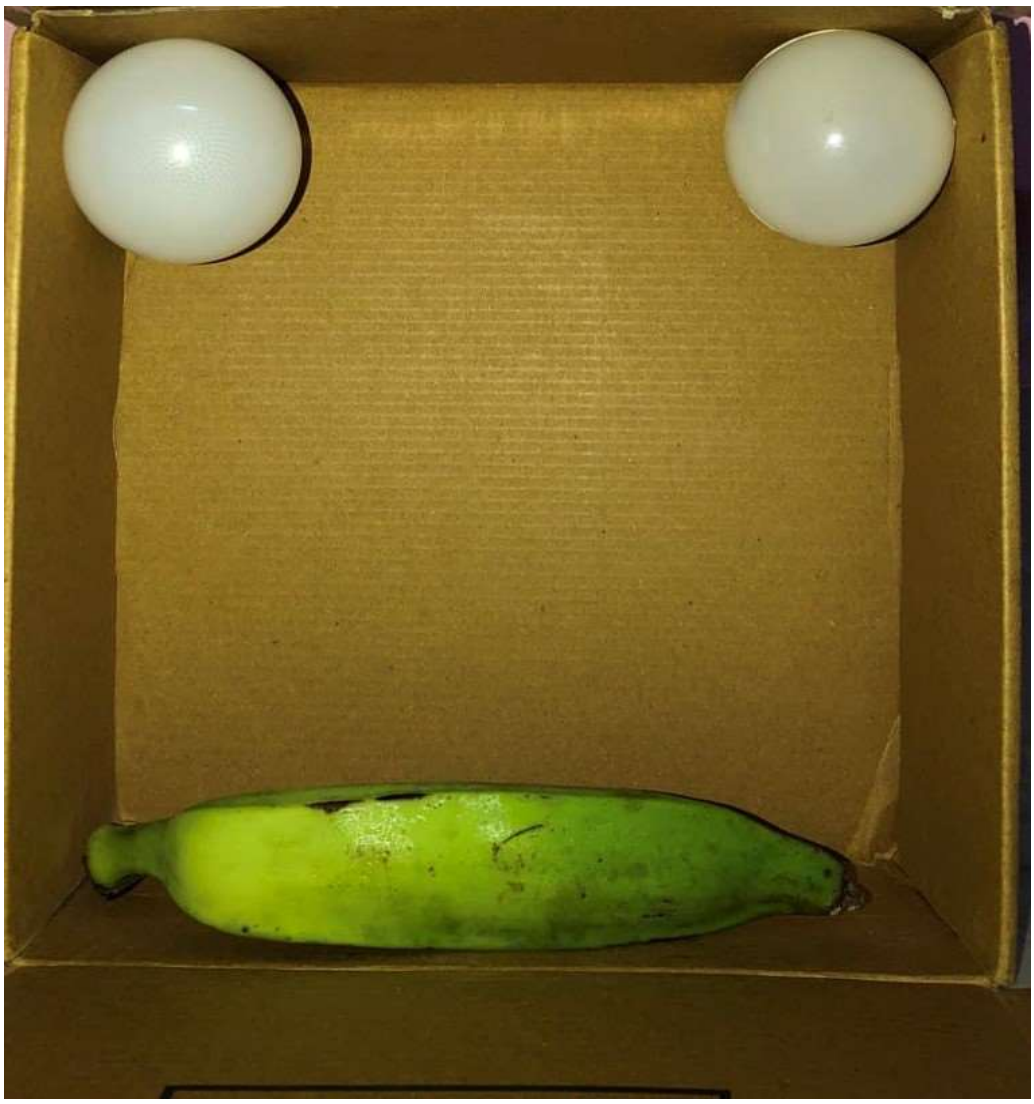

Plate 2. Image capturing setup for banana with two 36W white light source

3.2.2 Image pre-processing

Various forms of noises can be seen in the images collected. These noises can damage the quality of an image, preventing it from providing accurate information for image processing later on. To increase the quality of an image, operations must be done on it to eliminate or reduce degradations that occurred during the image acquisition process. Pre-processing is used to improve image data by suppressing unwanted distortions or enhancing certain image features that are useful for future processing, resulting in a more acceptable image for a given application than the original. According to the size of the pixel neighbourhood that is utilised for the computation of a new pixel, two main types of image pre-processing methodologies for food quality evaluation may be identified: pixel pre-processing and local pre-processing.

3.2.2.1 Pixel pre-processing

Pixel pre-processing is similar to pixel-by-pixel copying, with the exception that the values are changed according to the specified transformation function. The most frequent pixel pre-processing approach for evaluating quality is colour space transformation. The majority of applications where colour space transformations have been used for image pre-processing are likely to be based on HSI (hue, saturation, and intensity) colour space (Li *et al*., 2001; Tao *et al*., 1995). However, image pre-processing has also been done with the *L*a*b** colour space (Vizhanyo and Felfoldi, 2000).

3.2.2.2 Local pre-processing

Local pre-processing methods calculate the new value by averaging the brightness values of some neighbourhood points with attributes comparable to the processed point. Local pre-processing applications can be quite simple to highly complicated. Local pre-processing techniques can be used to blur sharp edges or maintain edges in the image, depending on the situation.

3.2.3 Image segmentation

Image segmentation is a method in which a digital image is broken down into its constituent items, which is a difficult process due to the image's rich visual

information. Thresholding-based, region-based, gradient-based, and classification-based segmentation are four alternative philosophical approaches to image segmentation that have been established for colour quality evaluation. Thresholding-based and region-based approaches have been employed in the majority of applications. Gradient-based and classification-based techniques, on the other hand, are utilised less often.

3.2.3.1 Thresholding-based segmentation

Image thresholding segmentation is a simple form of image segmentation. Threshold-based segmentation is a particularly useful approach for scenes with solid objects lying on a contrasting background, since it effectively differentiates the item from the remainder of the image. Some thresholding-based fruit quality evaluation methods accomplish segmentation directly by thresholding, while others integrate thresholding with additional approaches. Thresholding works best when the items of interest have a consistent inner grey level and are set against a different but consistent grey level background.

3.2.3.2 Region-based segmentation

A region can be classified as a collection of linked pixels with comparable attributes. Pixels might be comparable in terms of intensity, colour, and other factors. In this sort of segmentation, there are some established rules that a pixel must follow in order to be grouped into related pixel areas

There are two types of region-based segmentation methods: region growing-and-merging (GM) and region splitting-and-merging (SM). The first is a bottom-up approach that clusters pixels or sub-regions into bigger regions based on a set of homogeneity criteria, while the second is a top-down method that splits an image into smaller and smaller regions until specific requirements are fulfilled. Region-based algorithms are more computationally expensive than simpler approaches, such as thresholding-based segmentation, but they may use several picture attributes directly and simultaneously to determine the final boundary location. Because strong priori knowledge is not accessible, it shows the most potential in the segmentation of food products.

3.2.3.2.1 Gradient-based and classification-based segmentation

By partitioning the image into sets of interior and exterior points, the thresholding approach achieves segmentation. Gradient-based techniques, on the other hand, try to detect the edges directly based on their large gradient magnitudes. Based on categorization approaches such as statistics, fuzzy logic, and neural networks, classification-based systems attempt to allocate each pixel to different objects. In contrast to the typical technique of segmentation followed by classification, this segmentation is based on the results of classification. This method aims to determine the best way to segment the entire set of input images in order to improve classification accuracy.

3.2.4 Object measurement

Once the images are successfully segmented into discrete objects of interest, the objects can be defined and represented for further processing and analysis by measuring the specific characteristics of each item. A segmented object can be represented in terms of its external or interior characteristics in general. Several characteristics may be utilised to characterise an item, and these traits are compared to data from known objects to categorise it into one of many categories. The greatest features to utilise are those that are the easiest to measure and contribute significantly to the categorization.

3.2.5 Classification

Objects are identified by classifying them into one of the finite sets of classes, which entails comparing the measured properties of a new object to those of a known object or other known criteria to determine if the new object belongs to a certain category of objects. The three primary classification approaches are statistical, fuzzy logic, and neural networks. They share the goal of simulating the behaviour of a human decision-maker, and they have the advantage of consistency and, to a lesser extent, explicitness.

The collected images were pre-processed to eliminate the background interference. Region of interest (ROI) tool was used to select the image with the background omitted (Plate 3).

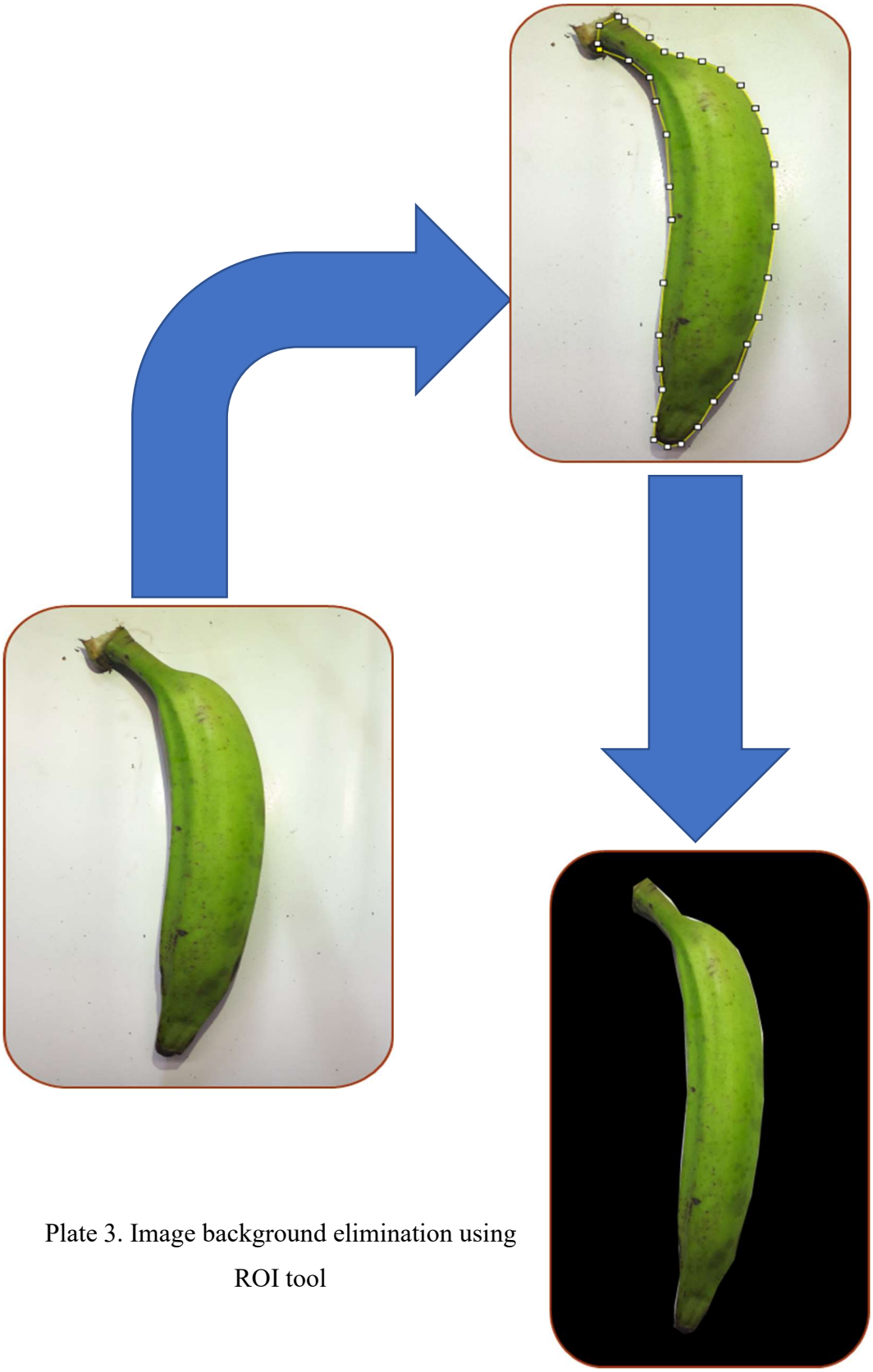

Plate 3. Image background elimination using ROI tool

3.3 TSS MEASUREMENT

Total soluble solids (TSS) were measured for selected fruits at various ripening time points. These time points spanned from harvest through fruit senescence, which was visually determined as the day when the fruits attained complete maturity (peel colour entirely yellow with brown spots) (Wang et al., 2014). Samples of banana were blended using a fruit juicer. The TSS were determined using a Hand-held refractometer (Plate 4).

3.3.1 Hand held refractometer

Hand held refractometer have an illuminator flap which generates diffused light at a grazing angle, which helps to hold the sample in place. Light goes through the sample, through the measuring prism and perhaps other lenses, and then onto the measuring scale, where it may be read. It must be calibrated before collecting readings to account for temperature changes (using calibration screw and distilled water).

To begin the measurement, open the illuminator flap (a little hinge connects it to the instrument) and place the fruit extract sample on the measuring prism surface. A pipette can be used to place the sample on the prism. After closing the flap, gaze through the eyepiece and read the scale's result. It may be required to point the refractometer in the direction of a light source (such as the Sun or a lamp) for better reading, but ambient light is generally sufficient throughout the day. After finishing measuring, wipe the prism and flap dry with a clean, soft cloth.

Plate 4. Hand held Refractometer

3.4 COLOUR ANALYSIS

3.4.1 RGB model

All colours are represented by varying combinations of red, green, and blue in the Red-Green-Blue colour model. This is an additive colour model, which means the three primary colours are blended or 'added' in different quantities to create a new colour. Each of the three additive primaries in the *RGB* colour space can have a value ranging from 0 to 255. As a result, every conceivable colour in the colour space is represented as an *RGB* triplet *(R*, *G*, *B*), each with its own *R*, *G*, and *B* values. When the values of red, green, and blue are all zero, the resultant colour in the colour space is black. Similarly, if all three have a value of 255, the result is white.

3.4.2 CIE *L*a* b**

CIE *L*a*b**, sometimes known as the CIELAB colour model, is one of the colour spaces created from the CIE XYZ space in 1976. The *L*a*b* * colour model uses location to define colour in 3D colour space. It was created to operate as a device-independent guide that specified all viewable colours to the human eye.

It is based on the Opponent - Colors Theory, which states that the cones in the human retina see colours as changes of two opposites:

a) Light and Dark shades of colour

b) Red and Green

c) Yellow and Blue.

The following are the parameters used in the CIE *L*a*b** models:

1. The colour brightness, denoted by the letter *L**, which ranges from 0 to 100. An *L** value of 0 corresponds to black, while a *L** value of 100 corresponds to diffuse white.

2. *a** represents the chroma/position of a colour between red/magenta and green. Green is indicated by negative *a** values, whereas magenta is indicated by positive *a** values.

3. The position of a colour between yellow and blue, denoted by the letter *b**. Blue denotes negative values, whereas yellow denotes positive ones.

Color representation on surfaces or materials with natural curvatures and undulations that are illuminated by a light source is better with the *L*a*b** colour system. Shadows and areas of glossiness on the object surface had less of an impact on the registered light intensities utilising *L*a*b**. Therefore, the *L*a*b** system is suggested as the best space for colour quantification in fruits.

3.4.3 *RGB* and *L*a*b** comparison

Table 2. Comparison of *RGB* and *CIELAB* colour spaces

Model	Parameters	Effective parameters	Advantages	Disadvantages
RGB	Red, Green, Blue	*R*, *G*, *B*	i. it is used as the base colour space in a variety of applications, because no modifications are necessary to display information on the screen. ii. Because of its additive properties, it's commonly used in visual displays iii. Considered as computationally feasible system.	i. Non useful for Color recognition and object specification. ii. Determining a specific colour is difficult in RGB color model iii. It's a hardware-oriented system
CIE *Lab*	*L*(luminance) *a* (red to green ratio) *b* (blue to yellow ratio)	*L*	i. Perceived as uniform ii. Device- independent	i. Suffer from unintuitive

3.4.4 *RGB* to *L*a*b** conversion

Images *RGB* data can be transformed to *L*a*b** values. Briefly, *RGB* values will be converted to *R'G'B'* (by dividing each value by 255). Using gamma curve fitting, the *R'G'B'* values will be transformed to linear values (*sRGB*) using the following equations (Stokes *et al.,*1996):

$$f(x) = \frac{x}{12.92} \quad ; \quad x \leq 0.03928 \tag{3.1}$$

$$f(x) = \left(\frac{x + 0.055}{1.055}\right)^{2.4} \quad ; \quad x > 0.03928\,; \tag{3.2}$$

Where x is a value of R', G', or B' and $f(x)$ is the converted value in sR, sG or sB. Then the *sRGB* values transformed to XYZ parameters using coefficients from the International Telecommunication Union.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} sR \\ sG \\ sB \end{bmatrix} \tag{3.3}$$

where *sR, sG* and *sB* are components of *sRGB* (Stokes *et al.,*1996).

The XYZ values will be used to determine the *L**, *a**, and *b** values (Papadakis *et al.,* 2000):

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) - 16 \tag{3.4}$$

$$a^* = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \tag{3.5}$$

$$b^* = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \tag{3.6}$$

3.4.5 Image Normalization

To make comparisons easier, the sRGB and *L*a*b** scales were normalized between 0 and 1. The normalized values for *sR*, *sG* and *sB* were calculated by dividing each pixel's value by 255. and the following equations for *L**, *a**, and *b** (Papadakis *et al.,* 2000):

$$\text{normalized } L^* = \frac{L^*}{100} \quad 3.7$$

$$\text{normalized } a^* = \frac{a^*+120}{240} \quad 3.8$$

$$\text{normalized } b^* = \frac{b^*+120}{240} \quad 3.9$$

For the images collected, *RGB* and *L*a*b** were extracted using ImageJ software.

3.4.6 ImageJ

ImageJ is a public domain software for processing and analysing scientific images. It is a free Java image processing programme based on the Macintosh's National Institutes of Health and the Laboratory (NIH). It runs on any computer that has a Java 1.4 or later virtual machine, either as an online applet or as a downloadable application. There are downloadable distributions for Windows, Mac OS, Mac OS X, and Linux. 8-bit, 16-bit, and 32-bit images can be displayed, edited, analysed, processed, saved, and printed with it. TIFF, GIF, JPEG, BMP, DICOM, FITS, and "raw" are just a few of the image formats it can read. It can display "stacks," which are a collection of images in a single window. Because it is multithreaded, time-consuming operations like reading image files can be performed in parallel with other tasks.

It can compute statistics on the area and pixel value of user-defined choices. It can calculate angles and distances. It can generate density histograms and line profile plots. Standard image processing functions including contrast modification, sharpening, smoothing, edge detection, and median filtering are all supported. Scaling, rotation, and flips are among the geometric adjustments it does. Zooming up to 32:1 and down to 1:32 is possible. At every magnification factor, all analysis and processing operations are available. The application may run unlimited number of windows (images) at the same time, with the only limitation being available memory.

Real-world dimensions measurements in millimetres may be obtained via spatial calibration. Calibration for density or grey scale is also offered. ImageJ was built with an open architecture that allows it to be extended via Java plugins. ImageJ's built-

in editor and Java compiler may be used to create custom acquisition, analysis, and processing plugins. Plugins built by users may address practically any image processing or analysis difficulty.

Then *RGB* values were measured using *RGB* measure plugin. This plugin menu separately measures the red, green and blue channels of an *RGB* image. Then using converting *RGB* image to *Lab* stack plugin the image is converted to new image-stack containing 3 slices *L*, *a*, *b* (Plate 5) and values were measured for corresponding images for all devices.

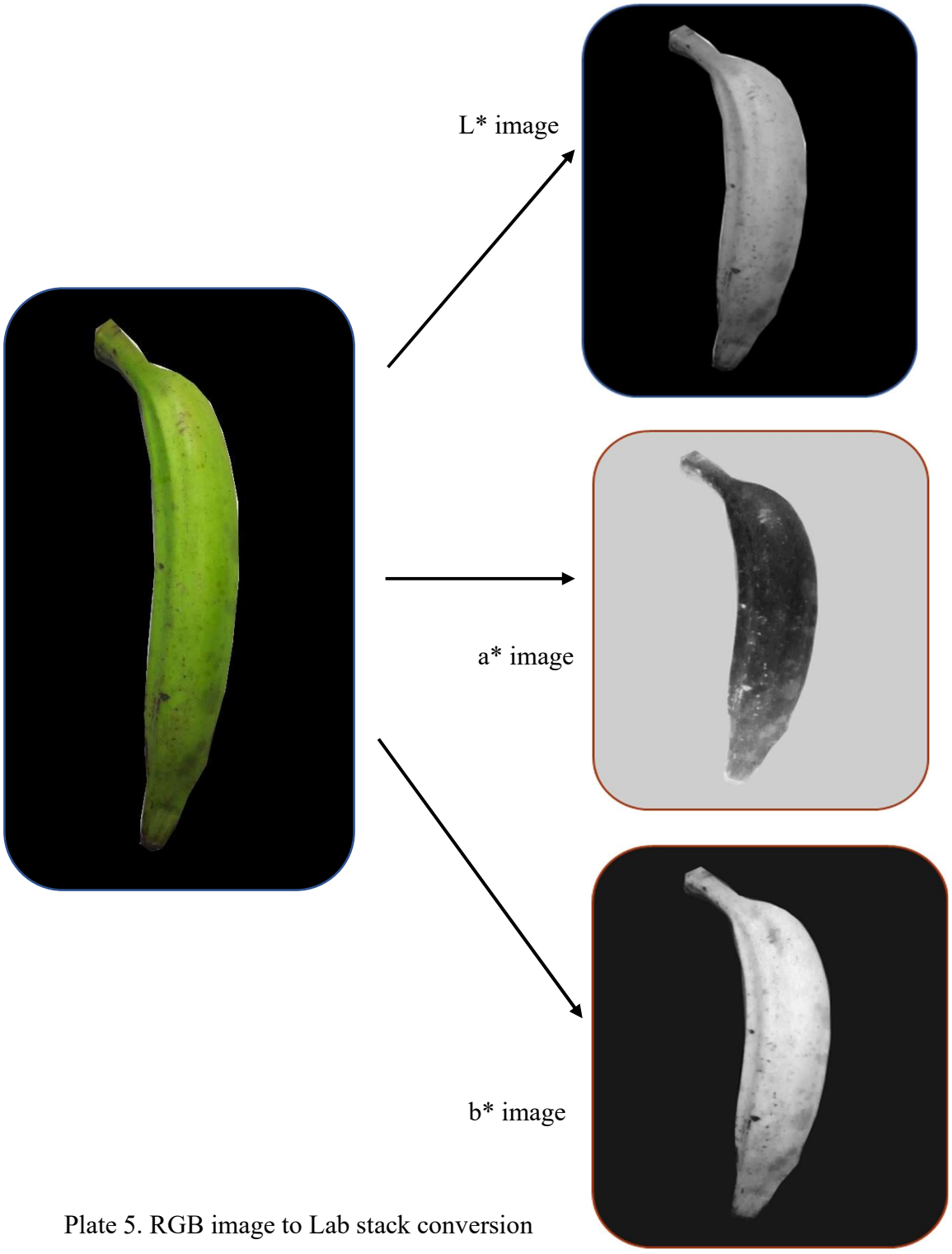

Plate 5. RGB image to Lab stack conversion

3.5 MACHINE LEARNING

Machine learning (ML) is the study of computer algorithms that can learn and develop on their own with experience and data. Machine learning employs programmed algorithms to receive and analyse input data in order to anticipate output values that are within a reasonable range. As fresh data is fed into these algorithms, they learn and optimise their processes in order to increase performance over time, developing 'intelligence' in the process.

Machine learning algorithms are divided into four categories: supervised, semi-supervised, unsupervised, and reinforcement.

3.5.1 Categories of machine learning algorithms

3.5.1.1 Supervised learning

The machine is taught by example in supervised learning. The operator gives the machine learning algorithm a known dataset containing desired inputs and outputs, and the algorithm must figure out how to get to those inputs and outputs. The algorithm analyses patterns in data, learns from observations, and generates predictions while the operator knows the proper answers to the issue. The operator corrects the algorithm's predictions, and the process repeats until the algorithm achieves a high degree of accuracy/performance. Regression falls under supervised learning.

3.5.1.2 Semi-supervised learning

Semi-supervised learning is similar to supervised learning, except it incorporates both labelled and unlabelled data. Labelled data is information with relevant tags so that the algorithm can interpret it, but unlabelled data does not have those tags.

3.5.1.3 Unsupervised learning

Here, the machine learning programme looks for patterns in the data. There is no answer key or human operator available to give assistance. Instead, the machine analyses available data to discover correlations and relationships. The machine learning algorithm is left to evaluate enormous data sets and respond to them in an unsupervised

learning process. The algorithm tries to organise the data in some way such that its structure can be described.

3.5.1.4 Reinforcement learning

Reinforcement learning is concerned with regimented learning processes in which a machine learning algorithm is given a set of actions, parameters, and end values to work with. The machine learning algorithm attempts to explore several alternatives and possibilities after setting the rules, monitoring and assessing each output to decide which is the best. Reinforcement learning instructs the system to learn via trial and error. It takes what it has learned in the past and adapts its approach to the situation in order to produce the best possible result.

3.5.2 Training and Test sets

3.5.2.1 Training data

This type of data builds up the machine learning algorithm. The model evaluates the data repeatedly to learn more about the data's behaviour and then adjusts itself to serve its intended purpose.

3.5.2.2 Test data

After the model is built, testing data once again validates that it can make accurate predictions. Test data provides a final, real-world check of an unseen dataset to confirm that the ML algorithm was trained effectively.

In the study the data was split into 80 percent training and 20 percent test data for the model fitting and prediction of TSS.

3.6 STATISTICAL ANALYSIS

3.6.1 Linear regression model

Linear regression models are used to show or predict the relationship between a dependent and an independent variable.

3.6.2 Model accuracy assessment

3.6.2.1 Residual standard error (RSE)

Even if the correct regression line is known, it is difficult to perfectly predict the dependent variable from the independent variable due to the inclusion of error components. The RSE estimate is a measure of prediction error. The lower the RSE, the more accurate the model is. The RSE is an estimate of the standard deviation of error. It is, roughly speaking, the average deviation of the response from the true regression line. It is computed using the following formula:

$$RSE = \sqrt{\frac{1}{n-2}RSS} = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n-2}} \quad 3.10$$

Residual sum of squares (RSS) is calculated by using the formula

$$RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad 3.11$$

By dividing the RSE by the mean outcome variable, the error rate can be calculated.

3.6.2.2 R^2 statistic

The RSE is an absolute measure of the model's lack of fit to the data. However, because it is expressed in *Y* units, it is not always clear what defines a good RSE. The R^2 statistic is another way to measure fit. It is expressed as a proportion — the proportion of variance explained — and hence always has a value between 0 and 1, regardless of the scale of *Y*.

R^2 is computed by the formula

$$R^2 = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS} \quad 3.12$$

where $TSS = \sum(y_i - \overline{y})^2$ is the total sum of squares, and RSS is the residual sum of squares.

An R^2 value close to 1 indicates that the model explains a large percentage of the variance in the outcome variable. The R^2 represents the correlation coefficient between the observed values of the outcome variable (y) and the fitted (i.e., predicted) values of y in multiple linear regression.

The R^2 has the drawback of increasing always as more variables are included to the model, even if those variables are very weakly associated with the response. Adjusting the R^2 to account for the number of predictor variables is one solution.

The adjustment in the summary output's "Adjusted R Square" value is a correction for the amount of x variables in the prediction model.

3.6.2.3 Root mean squared error

The root mean squared error, or RMSE prediction error, is the most used metric for evaluating linear regression model performance. The primary concept is to compare the model's predictions against actual observed values to see how bad/erroneous they are. It is the average difference between the observed known outcome values in the test data and the predicted outcome values by the model. The model is better if the RMSE is low.

3.6.2.4 F-Statistic

The F-statistic determines the model's overall significance. It determines whether there is a non-zero coefficient for at least one predictor variable. A large F-statistic corresponds to a statistically significant p-value ($p < 0.05$).

3.6.3 Multiple linear regression

The relationship between two or more independent variables and one dependent variable is estimated using multiple linear regression.

Initially Multiple linear regression was fitted between *R*, *G*, *B* and TSS using R software. The data was split into 80 percent training and 20 percent test data before fitting the model. *R*, *G*, *B* color values were plotted against TSS using linear models separately for all devices. Then multiple linear regression model was fitted for *L**, *a**, *b** and TSS. The linear model plots were obtained separately for

*L**, *a**, *b** against TSS. Fitted models between color parameters and TSS were evaluated.

*L*a*b** is the best space for colour quantification in fruits therefore, the data was visualised between *L**, *a**, *b** and TSS for all the devices.

Since a non-linear trend appeared on visualization, nonlinear approaches were considered further for modelling. Splines, an often-superior approach to modelling nonlinear relationships was used to model *L**, *a**, *b** and TSS.

3.6.4 Spline regression

Spline regression is a type of non-linear and non-parametric regression that attempts to overcome the difficulties that linear and polynomial regression techniques have. In linear regression, the full dataset is taken into account at the same time. The dataset is partitioned into bins in spline regression. After that, each data bin is made to fit into its own model. Each function is referred to as a piecewise step function since there are independent functions that suit the bins.

3.6.4.1 Knots

Knots are the spots where the data is partitioned into bins for fitting separate model. Number and placement of knots may have dramatic effect on fit and smoothness. 3 knots (0.25, 0.5, 0.75) were used in the study

3.6.4.2 Generalized cross-validation

For spline smoothing, generalised cross-validation (GCV) is a popular parameter selection criterion, but it can produce poor results if the sample size is insufficient.

3.6.5 Generalized Additive Models (GAMs)

A Generalized Additive Model (GAM) uses smooth functions (like splines) for the predictors in a regression model. It helps to flexibly model non-linear relationships and to fit spline models with automated selection of knots. Data may be fitted using smooths, or splines, which are functions that can take on a variety of forms, using a GAM. The gam () function in R from the mgcv package may be used to fit GAM via

an approach known as backfitting. This method fits a model with many predictors by updating the fit for each predictor individually while keeping the others constant. When GAM is fitted, the independent variable, x, is wrapped in the s (), that is the smooth function, which specifies that the relationship is flexible. GAMs' flexible smooths are really made up of a number of smaller functions. These are referred to as "basis functions." Each smooth is the product of a number of basis functions, each of which is multiplied by a coefficient, which is a model parameter.

The reference number of degrees of freedom used for hypothesis testing is the Ref.df value (on the basis of the associated F-value). The edf value is the number of effective degrees of freedom, which can be interpreted as an estimate of how many parameters are required to represent the smooth. The smooth's quantity of non-linearity is indicated by the edf value. The edf value indicates the amount of non-linearity of the smooth. If the edf value for a smooth is (close to) 1, the pattern is (close to) linear. A value greater than one denotes a more complex pattern (i.e., non-linear). The maximum edf value is k minus one (as the intercept is part of the parametric coefficients). If the edf value is close to its maximum, it may be required to increase the base dimension to avoid over smoothing.

Spline regression was fitted using R software using splines package and mgcv package. TSS and L*, a*, b* values were modelled using spline regression and was compared with the linear model fitted earlier. After comparing both the models best fitting model was identified for the prediction of TSS.

Possibility of deep learning using Convolutional neural network (CNN) was also explored in this study.

3.7 DEEP LEARNING

Deep learning is a subset of machine learning that is essentially a three or more layers neural network. These neural networks are designed to mimic the human brain's behaviour by allowing it to "learn" from large amounts of data. While a single-layer neural network may produce approximate predictions, more hidden layers can assist to optimise and improve for accuracy.

3.7.1 Convolutional neural network (CNN)

CNN is a type of artificial neural network used in image recognition and processing that is specifically designed to process pixel data. It can recognise features and patterns within an image, allowing tasks such as object detection and recognition. It is largely utilised in computer vision and image classification applications. The layers of a CNN consist of an input layer, an output layer and a hidden layer that includes multiple convolutional layers, pooling layers, fully connected layers and normalization layers (Fig. 1)

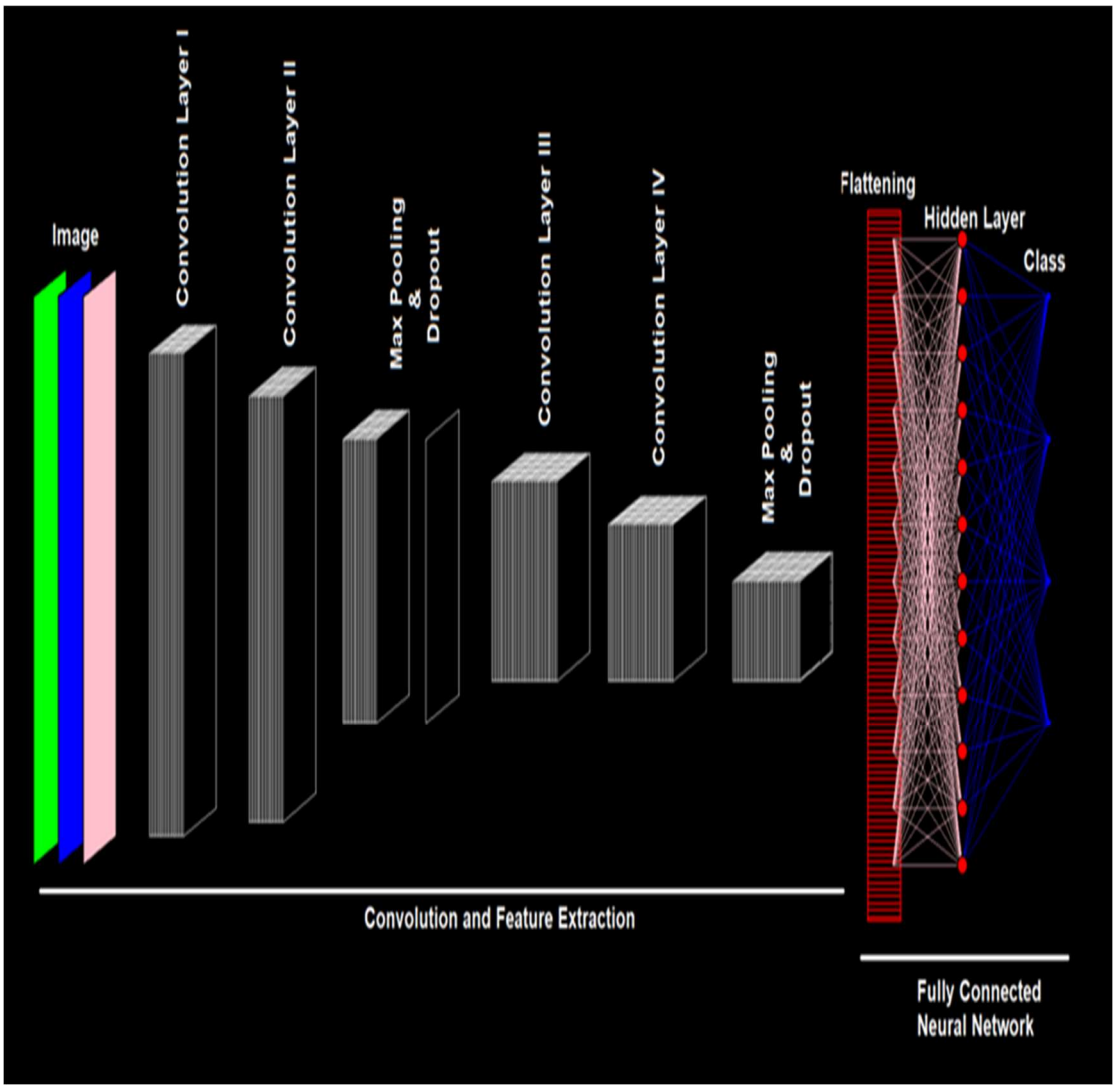

Fig. 1. Architecture of the Convolution neural network (online source: https://towardsdatascience.com/a-laymans-guide-to-building-your-first-image-classification-model-in-r-using-keras-b285deac6572)

3.7.1.1 Convolution layers: Different features of the input are extracted using the convolution operation. Low-level features such as edges, lines, and corners are extracted by the first convolution layer. Higher-level layers extract features at a higher level. Convolution is performed by sliding a smaller frame over the entire image, which acts as a receptive field for the temporal-spatial features. This frame is called as a Kernel/Filter. The weights of the neural network are equivalent to the filter values. At each given time, a single feature output is generated by a linear combination of these filter values and the image's foregrounding pixel values. A feature map is created when a set of these values is formed.

3.7.1.2 Rectified Linear Unit (ReLU): An activation function is applied over each of the feature maps once they are generated. Because a ReLU implements the function $y = \max(x,0)$, this layer's input and output sizes are the same. It improves the decision function's and overall network's nonlinear qualities without impacting the convolution layer's receptive fields.

3.7.1.3 Pooling/subsampling layers: The resolution of the features is reduced by the pooling/subsampling layer. It strengthens the features against noise and distortion.

3.7.1.4 Fully connected layers: The last layers of a CNN often used are fully connected layers. These layers add up a weighted total of the previous layer's characteristics, showing the exact mix of "ingredients" needed to get a given intended output result. In the case of a fully connected layer, each element of each output feature is calculated using all of the elements of the preceding layer's features.

From the images obtained using all the three devices three categories with 30 samples each were categorised based on TSS. The categories were Raw (TSS 4-10), Medium (TSS 11-17), Ripe (TSS 18-32). 25 images from each category were selected as training set and 5 images were selected as test set. The images were uploaded in R and image classification model was developed using Tensorflow framework.

3.7.2 Learning curves to diagnose machine learning model performance

- Learning curves are plots that depict how learning performance changes over time as a function of experience.

- Model performance learning curves on the train and validation datasets can be used to diagnose if a model is underfit, overfit, or well-fit.
- Learning curves of model performance can be used to determine if the train or validation datasets are representative of the problem domain

Learning curves are commonly used in machine learning for algorithms like deep learning neural networks that learn (optimise their internal parameters) incrementally over time. The present state of a machine learning model can be examined at each step of the training algorithm during training. On the training dataset, it can be tested to see how effectively the model is "learning." It can also be tested on a hold-out validation dataset that isn't included in the training set. The validation dataset is used to determine how well the model is "generalising."

- Train Learning Curve: This is a learning curve derived from the training dataset that indicates how effectively the model is learning.
- Validation Learning Curve: This is a learning curve derived from a hold-out validation dataset that indicates how effectively the model generalises.

3.7.3 Epoch

When an entire dataset is passed forward and backward through the neural network only once. While the epochs are running graphics of juxtaposed curves of loss and accuracy for training and validation sets can be obtained.

In the study, 100 epochs were performed to avoid overfitting and to increase the generalization capacity of neural network.

Results and Discussion

4. RESULTS AND DISCUSSION

The study entitled "Statistical Assessment of Banana ripening using smartphone-based images" has been carried out at the Department of Agricultural Statistics, College of Agriculture, Vellayani, Thiruvananthapuram during the year 2019-2020. In order to develop of suitable model to establish the relationship between Total Soluble Solids (TSS) and L^*(lightness), a^*(green-red ratios), b^*(blue-yellow ratios) values of bananas were collected from randomly selected fields in Trivandrum district. Images were collected on a daily basis and TSS for the same is measured. Best fitting model was identified for prediction of TSS values using L^*, a^*, b^* values. A protocol for accurate data collection to assess TSS content in Banana using smart-phone-based images was developed. Along with this, the possibility of deep learning using Convolutional neural network (CNN) was also explored. Various statistical methods were used to fit the model between TSS and color parameters. The results of the study are given under the following subsections.

4.1 Color parameters extraction

4.2 Linear model fitting between R, G, B, L^*, a^*, b^* and TSS

4.3 Data visualisation between L^*, a^*, b^* and TSS

4.4 Spline model fitting between L^*, a^*, b^* and TSS

4.5 Comparison of fitted models and identification of suitable model

4.6 Protocol for accurate data collection using smartphone

4.7 Convolutional neural network for identification and classification of banana images into different categories

4.1 COLOR PARAMETERS

For the evaluation of color parameters, 377 images of Nendran variety were collected using three different smartphones having different resolution properties. Sampling distribution for each device is given in Table 6. For each set of bananas, the images were captured from the day of collection to a ripened stage with brown spots. Table 3, 4 and 5 represents the color parameters along with TSS for day 1-6 for devices

1, 2 and 3 respectively. It was found from the tables that, the *L** and *b** values varied non uniformly throughout the cycle. The TSS increased over storage time. During ripening, the a* value steadily rises, indicating the degradation of chlorophyll and the loss of the peel's green colour.

Table 3. *L**, *a**, *b** values of images using device 1 and their TSS for a set of bananas for initial six days after collection

	*L**	*a**	*b**	TSS
Day 1	42.19	-11.89	34.97	5
Day 2	30.02	-6.65	20.91	7.2
Day 3	29.93	-3.22	28.27	10.9
Day 4	56.25	3.41	55.72	17.8
Day 5	42.45	9.95	40.86	25.6
Day 6	34.86	11.31	19.15	30.8

Table 4. *L**, *a**, *b** values of images using device 2 and their TSS for a set of bananas for initial six days after collection

	*L**	*a**	*b**	TSS
Day 1	41.85	-12.81	29.06	4.2
Day 2	38.04	-7.63	21.00	7.2
Day 3	43.60	-3.65	30.98	13.8
Day 4	59.12	1.57	41.64	17.8
Day 5	46.43	7.68	30.67	28.2
Day 6	53.20	7.04	30.77	31.8

Table 5. *L**, *a**, *b** values of images using device 3 and their TSS for a set of bananas for initial six days after collection

	*L**	*a**	*b**	TSS
Day 1	40.18	-15.56	26.33	4.6
Day 2	38.31	-6.07	35.17	8
Day 3	43.34	-3.79	43.01	13.8
Day 4	60.57	2.67	50.66	17.8
Day 5	55.81	6.07	51.13	26.2
Day 6	50.29	14.56	42.86	30.6

Table 6. Number of images captured using three smartphones

Devices	No. of samples
Device 1	129
Device 2	126
Device 3	122

4.2 LINEAR MODEL FITTING BETWEEN *R*, *G*, *B*, *L**, *a**, *b** AND TSS

Linear model was fitted between *R*, *G*, *B* and TSS for all the images captured using three smartphones. Machine learning model was implemented for fitting linear models. Total sample was divided into 80 percent training and 20 percent test set.

Table 7. *RGB* and TSS linear model coefficients for device 1

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	12.42	1.65	7.52	9.05e-12 **
R	0.31	0.01	18.00	< 2e-16 **
G	-0.32	0.02	-16.13	< 2e-16 **
B	0.07	0.03	2.45	0.0154 *
Multiple R^2				0.80
Adjusted R^2				0.80
RMSE for predicted TSS				3.83
R^2 value for prediction				0.84

For device 1, out of the total 129 samples, 104 samples were taken as training and 25 samples were taken as test set. It is evident from Table 7 that *R*, *G* and *B* have significant relationships with TSS for images captured using device 1. Since p-value is less than 0.05, R and G found to be significant at 1 percent level of significance. B found to be significant at 5 percent level of significance.

Therefore, the model will be

$$TSS = 12.42 + 0.31\,R - 0.32\,G + 0.07\,B \qquad 4.1$$

Regression coefficients of *R* and *B* were positive indicating a positive correlation between TSS and *R*, *B* color values whereas, negative coefficient of *G* suggests that there is a negative correlation between *G* and TSS. Adjusted R^2 obtained was 0.80, meaning that 80 percent of the variance in the measure of TSS can be predicted by *R*, *G* and *B*. The RSE was 4.04 corresponding to a 29.73 percent error rate. The RMSE and R^2 values for predictions were 3.83 and 0.84 respectively.

Table 8. *RGB* and TSS linear model coefficients for device 2

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	18.21	2.03	8.93	5.27e-15 **
R	0.56	0.02	22.73	< 2e-16 **
G	-0.69	0.05	13.49	< 2e-16 **
B	0.12	0.06	2.03	0.0439 *
Multiple R^2				0.81
Adjusted R^2				0.80
RMSE for predicted TSS				2.98
R^2 value for prediction				0.90

For device 2, out of the total 126 samples, 102 samples were taken as training and 24 samples were taken as test set. The results presented in table 8 shows that for the images captured using device 2, *R*, *G* and *B* values have significant relationship with TSS. *R* and *G* found to be significant at 1 percent level of significance. B found to be significant at 5 percent level of significance. For *R* and *B* estimates were positive

indicating a positive correlation with TSS. In the case of *G*, there is a negative correlation with TSS.

Therefore, the model will be

$$TSS = 18.21 + 0.56\,R - 0.69\,G + 0.12\,B \qquad 4.2$$

Adjusted R^2 obtained was 0.80, meaning 80 percent of the variance in the measure of TSS can be predicted by *R*, *G* and *B*. The RSE obtained was 4.02 corresponding to a 23.70 percent error rate. The RMSE and R^2 values for predictions were 2.98 and 0.90 respectively.

Table 9. *RGB* and TSS linear model coefficients for device 3

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8.71	1.56	5.57	1.56e-07 ***
R	0.36	0.01	23.68	< 2e-16 ***
G	-0.33	0.01	-16.81	<2e-16***
B	0.01	0.02	0.59	0.551
Multiple R^2				0.84
Adjusted R^2				0.84
RMSE for predicted TSS				2.22
R^2 value for prediction				0.95

The result presented in table 9 shows that for the images captured using device 3, *R* and *G* values have significant relationship with TSS. *R* and *G* found to be significant at 1 percent level of significance. *B* value found to be non-significant. The regression coefficient of R colour value was positive while for *G* it was negative.

Therefore, the model will be

$$TSS = 8.71 + 0.36\,R - 0.33\,G \qquad 4.3$$

Adjusted R^2 was 0.84, meaning 84 percent of the variance in the measure of TSS can be predicted by *R* and *G*. The RSE for the model was 3.57 corresponding to 18.67 percent error rate. The RMSE and R^2 values for predictions were 2.22 and 0.95 respectively.

Fig. 2, 3 and 4 shows the linear regression lines fitted for *R*, *G*, *B* and TSS using R software for device 1. Fig. 5, 6 and 7 shows the linear regression lines fitted for *R*, *G*, *B* and TSS for device 2. For device 1, 2 and 3, linear regression plots were prepared by taking one colour parameter against TSS by keeping the remaining colour parameters under control. From Fig. 2, 3 and 4, it was seen that the regression line for *R*, *G*, *B* and TSS for device 1 does not pass through most of the points. Similarly for device 2 and 3, the data points were farther from the linear regression line in Fig. 5, 6, 7 and Fig. 8, 9, 10.

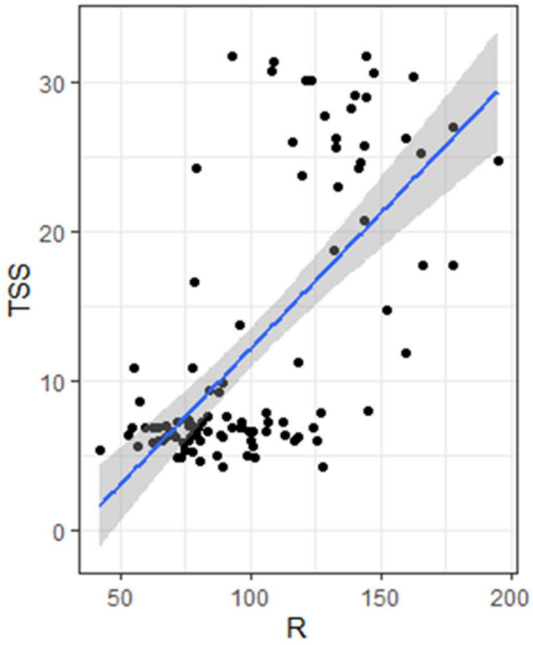

Fig.2. Linear regression plots between *R* and TSS for device 1

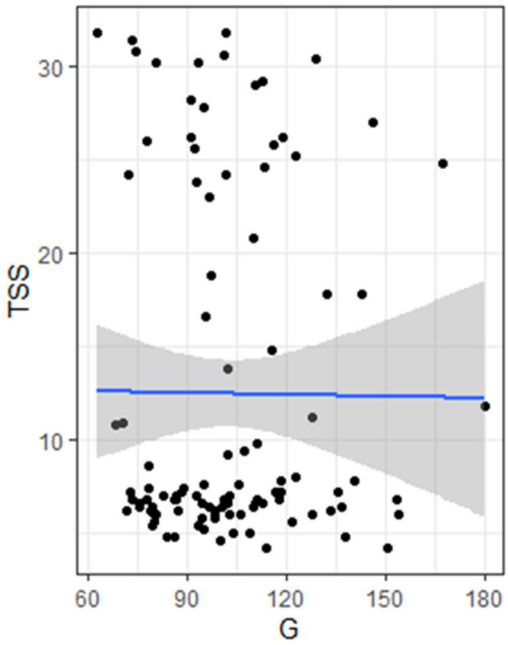

Fig.3. Linear regression plots between *G* and TSS for device 1

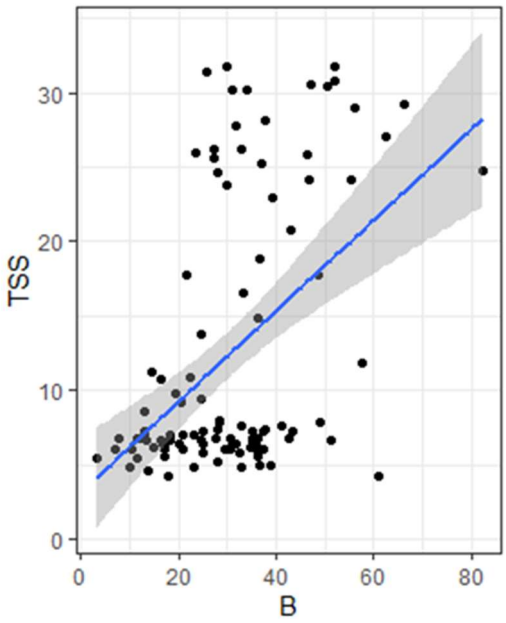

Fig.4. Linear regression plots between *B* and TSS for device 1

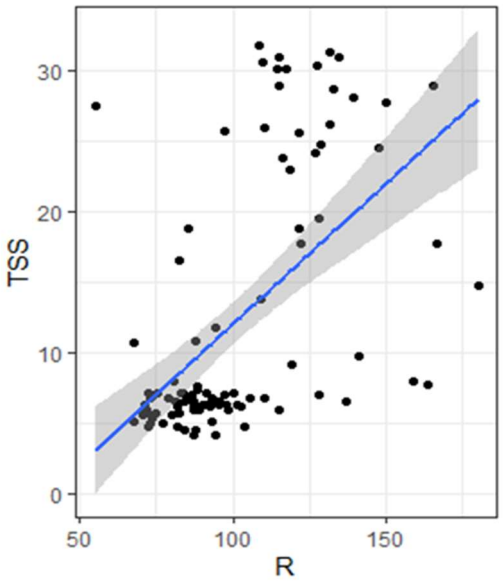

Fig.5. Linear regression plots between *R* and TSS for device 2

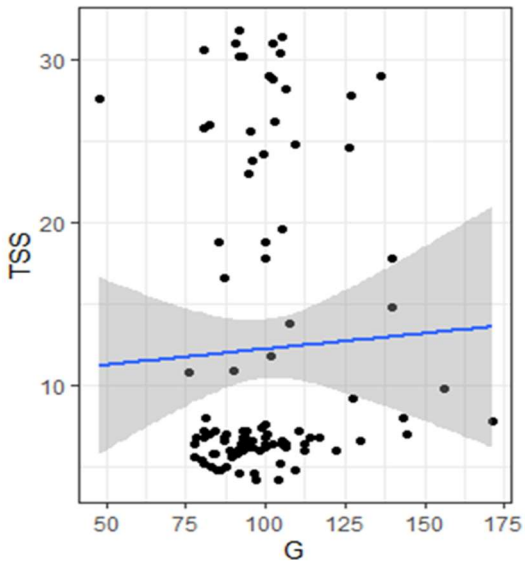

Fig.6. Linear regression plots between *G* and TSS for device 2

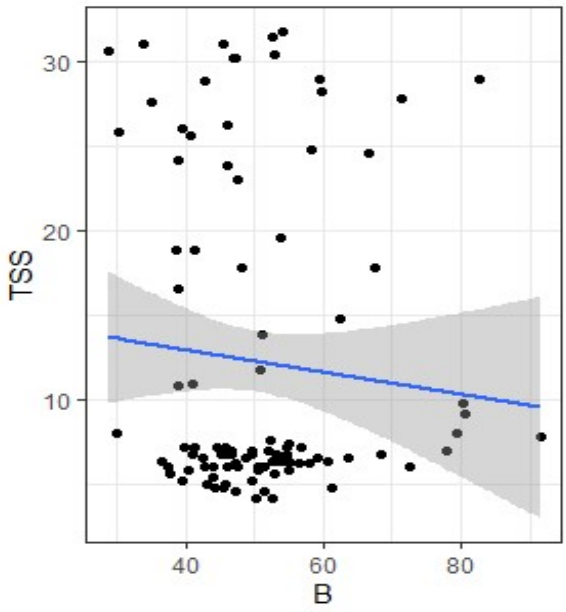

Fig.7. Linear regression plots between *B* and TSS for device 2

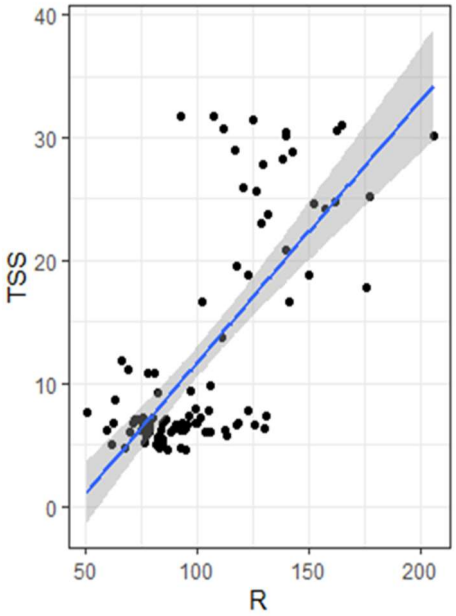

Fig.8. Linear regression plots between *R* and TSS for device 3

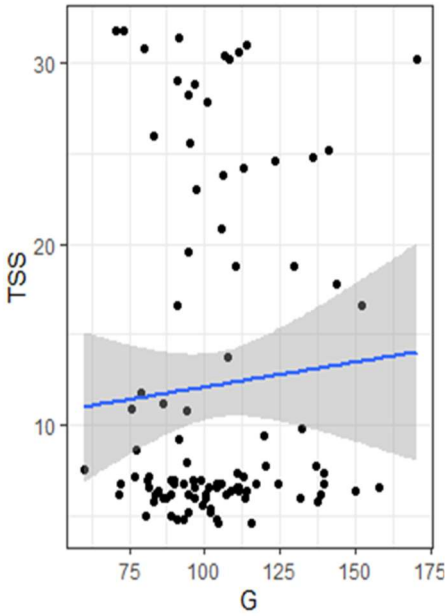

Fig.9. Linear regression plots between *G* and TSS for device 3

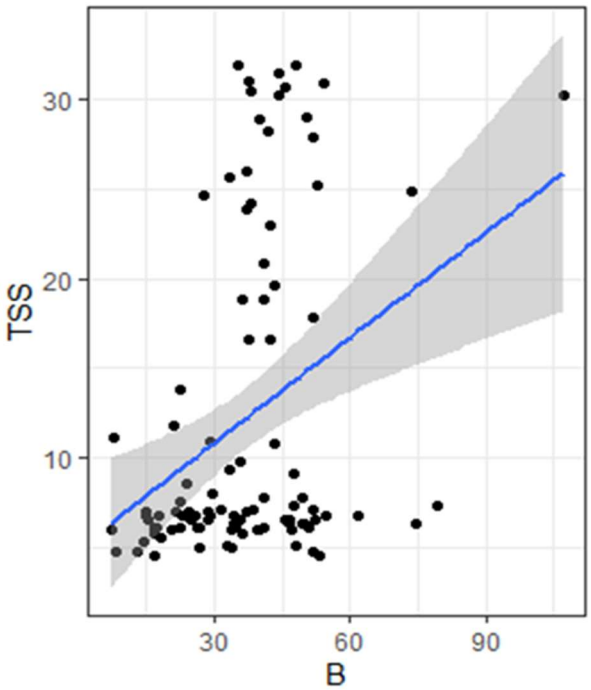

Fig.10. Linear regression plots between *B* and TSS for device 3

But in the case of the *RGB* colour model it is a device dependent model. This implies that a certain *RGB* triplet may differ from device to device. This is because the colour components (phosphors or dyes) used by various manufacturers may react differently to varying *R*, *G*, and *B* values. As a result, an *RGB* value may not always represent the same colour on different devices. Because it specifies colours in terms of additive primaries, *RGB* is not an intuitive colour model.

Therefore, the extracted *L*a*b** color values were used for fitting linear models to predict TSS. *L* a* b** is device-independent. As a result, unlike the *RGB* system, there will be no differences in *CIELAB* models on various devices. The Luminosity/Brightness component is quite similar to how people perceive lightness. It may be used to properly modify the colour balance. The *CIELAB* colour gamut is much larger than the *RGB* colour gamut, and it also includes the gamut of the latter model. The colour range of this model is so wide that it contains colours that the human eye cannot see. In other words, it comprises imaginary colours as well. By altering chroma, brightness, and occasionally hue, colour management software, such as built-in image processing tools, would select the closest colour present in the spectrum of colours visible to humans.

Table 10. *L*a*b** and TSS linear model coefficients for device 1

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	11.82	1.80	6.54	1.37e-09 **
L	0.01	0.06	0.19	0.84854
a	0.67	0.03	19.90	< 2e-16 **
b	0.17	0.06	2.86	0.00488 *
Multiple R^2				0.78
Adjusted R^2				0.78
RMSE for predicted TSS				4.04
R^2 value for prediction				0.84

The results presented in Table 10 revealed that the parameters a^* and b^* had a significant relationship with TSS for images captured using device 1. a* found to be significant at 1 percent level of significance and b* found to be significant at 5 percent level of significance. Regression coefficients of a^* and b^* were positive indicating a positive correlation with TSS.

Therefore, the model will be

$$TSS = 11.82 + 0.67\, a^* + 0.17\, b^* \qquad 4.4$$

Adjusted R^2 obtained was 0.78, meaning 78 percent of the variance in the measure of TSS can be predicted by a^* and b^* color values. The RSE was 4.25 corresponding to 31.39 percent error rate. The RMSE and R^2 values for predictions were 4.04 and 0.84 respectively.

Table 11. $L^*a^*b^*$ and TSS linear model coefficients for device 2

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	19.60	2.08	9.38	4.36e-16 ***
L	0.01	0.07	0.21	0.827
a	1.20	0.05	21.38	< 2e-16 **
b	0.01	0.10	0.10	0.914
Multiple R^2				0.82
Adjusted R^2				0.81
RMSE for predicted TSS				2.86
R^2 value for prediction				0.91

The result presented in table 11 indicated that only a^* and TSS had a significant relationship. a* found to be significant at 1 percent level of significance. L^* and b^*

found to be non-significant. For a^*, the estimate was positive indicating a positive correlation with TSS.

Therefore, the model will be

$$TSS = 19.60 + 1.2\ a^* \qquad 4.5$$

Adjusted R^2 obtained was 0.81, meaning 81 percent of the variance in the measure of TSS can be predicted by a^*. The RSE obtained was 3.90 corresponding to 22.78 percent error rate. The RMSE and R^2 values for predictions were 2.86 and 0.91 respectively.

Table 12. $L^*a^*b^*$ and TSS linear model coefficients for device 3

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	9.64	1.62	5.95	2.74e-08 **
L	0.16	0.05	3.16	0.002 **
a	0.74	0.02	25.23	< 2e-16 **
b	0.07	0.05	1.40	0.162
Multiple R^2				0.85
Adjusted R^2				0.85
RMSE for predicted TSS				2.16
R^2 value for prediction				0.95

It was found from Table 12 that, L^* and a^* had a significant relationship with TSS. L^* and a^* were significant at 5 percent level of significance. Color value b* found to be non-significant.

Therefore, the model will be

$$TSS = 9.64 + 0.16\ L^* + 0.74\ a^* \qquad 4.6$$

Adjusted R^2 was 0.85 and RSE was 3.48 with 18.17 error rate. The RMSE and R^2 values for predictions were 2.16 and 0.95 respectively.

The fitted line plots in Fig. 11-19 displayed that the model didn't fit properly on the data. Most of the points were lying farther from the regression line showing a nonlinear relationship with TSS for all the devices. Linear models of RGB have shown a significant and high adjusted R-Square than linear models of L^* a^* b^* and TSS.

From the linear models between L^*, a^*, b^* and TSS, it was found that when the camera specification varied, the significant ones also varied. Hence it will be difficult to find a generalized model using L*, a*, b* colour parameters. The colour value a* found to be significant for all the three devices.

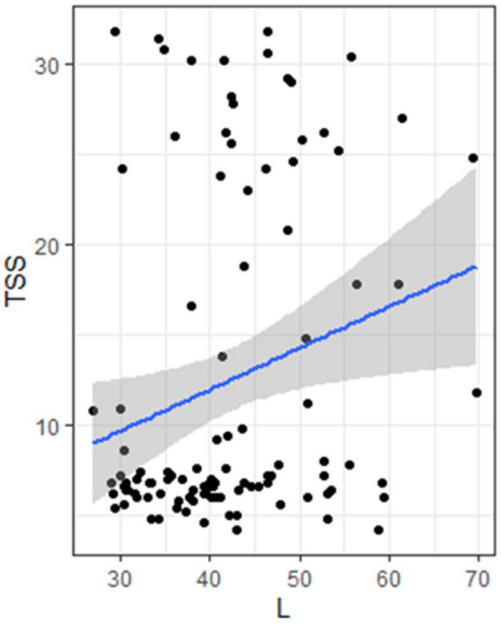

Fig. 11. Linear regression plots between *L**and TSS for device 1

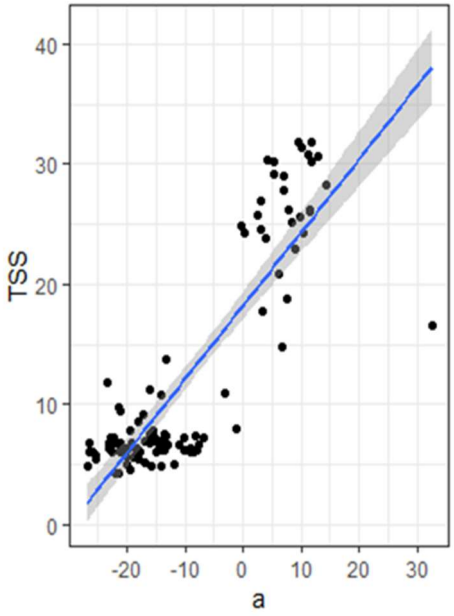

Fig. 12. Linear regression plots between *a**and TSS for device 1

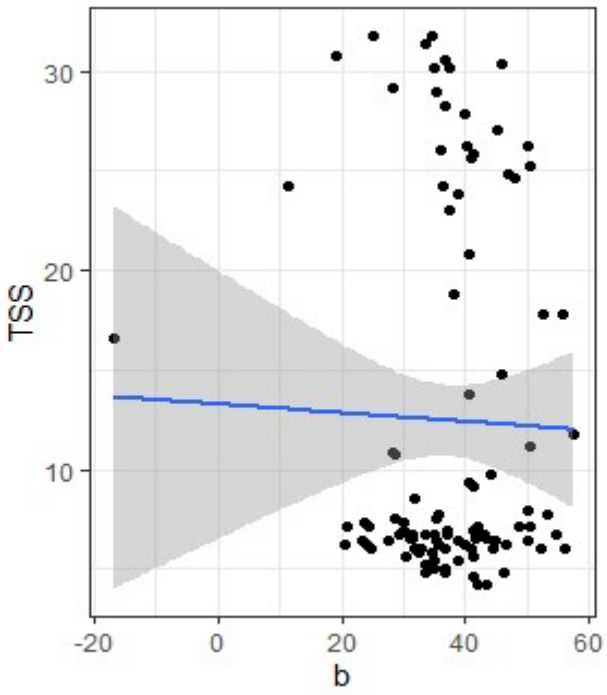

Fig. 13. Linear regression plots between *b**and TSS for device 1

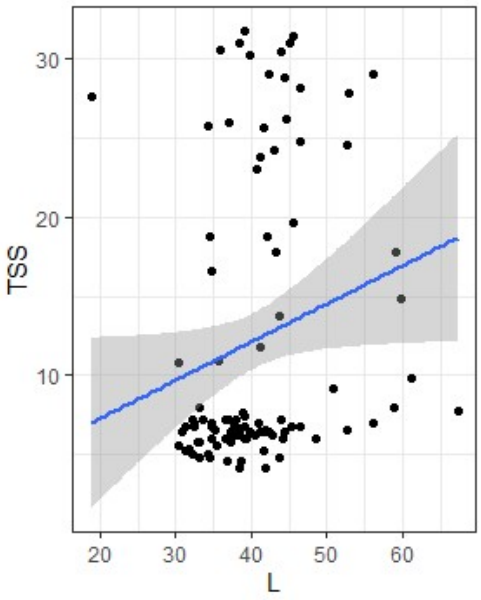

Fig. 14. Linear regression plots between *L**and TSS for device 2

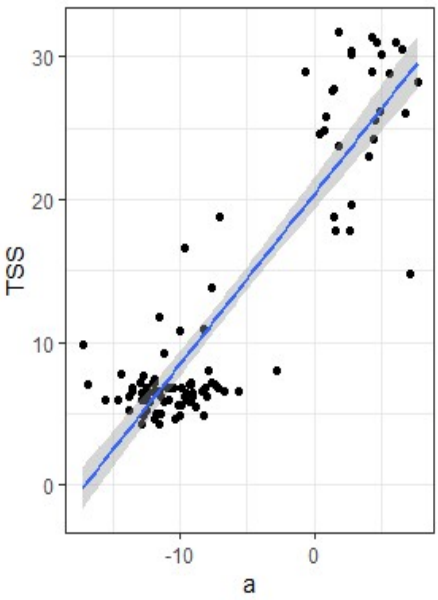

Fig. 15. Linear regression plots between *a**and TSS for device 2

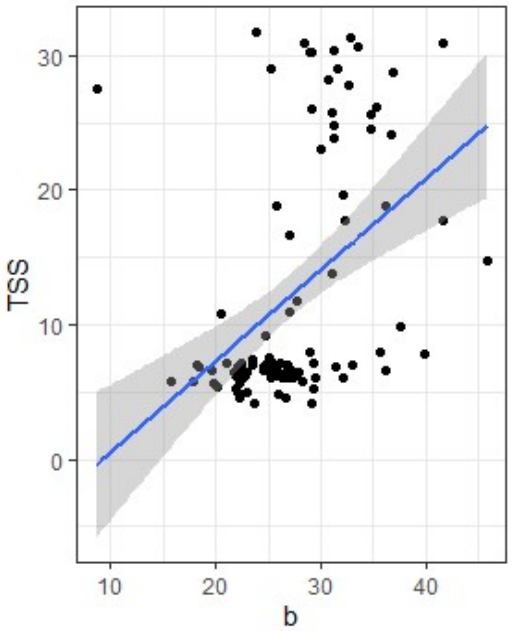

Fig. 16. Linear regression plots between *b**and TSS for device 2

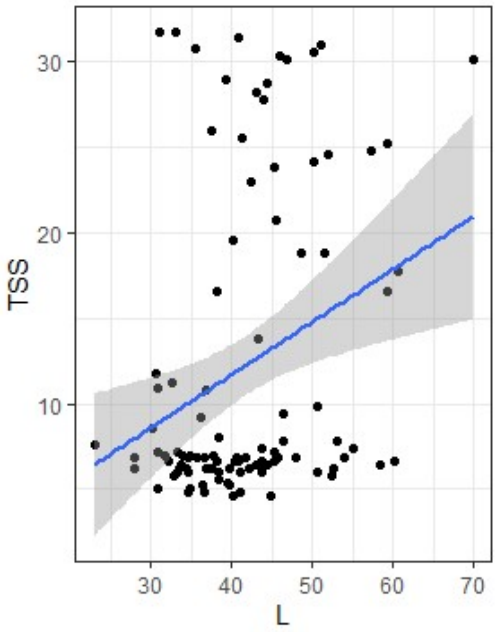

Fig. 17. Linear regression plots between *L**and TSS for device 3

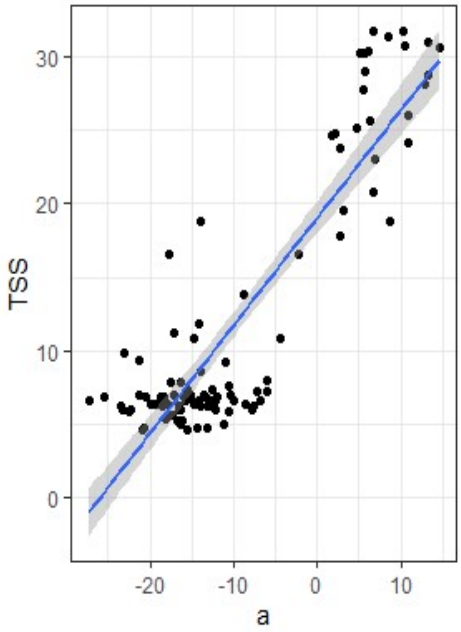

Fig. 18. Linear regression plots between *a**and TSS for device 3

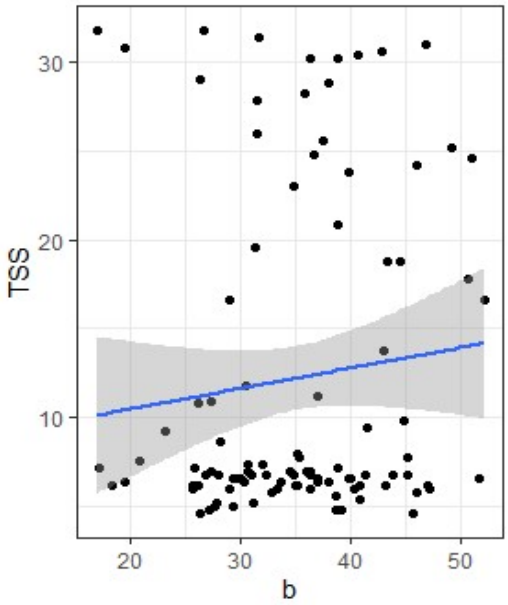

Fig. 19. Linear regression plots between *b**and TSS for device 3

4.3 DATA VISUALISATION BETWEEN *L**, *a**, *b** AND TSS

Data has been visualised to see the trend of data points. Scatter plots were created using R software and the 'ggplot2' package. The function 'geom_point ()' is used. The function 'stat_smooth ()' is used to add regression lines to the scatter plot. Fig. 20, 21, 22 picturise the trend between *L**, *a**, *b** and TSS for device 1. It is clear from the plot that the relationship is non-linear. Similarly for device 2 and device 3 scatter plots were prepared (Fig.23-28) and a non-linear relationship was found between colour parameters and TSS. Therefore, the possibility of non-linear regression was explored in the study. Since all the scatterplots showed a non-linear relationship but can't say it is strictly nonlinear hence spline regression was tried to fit in the data points.

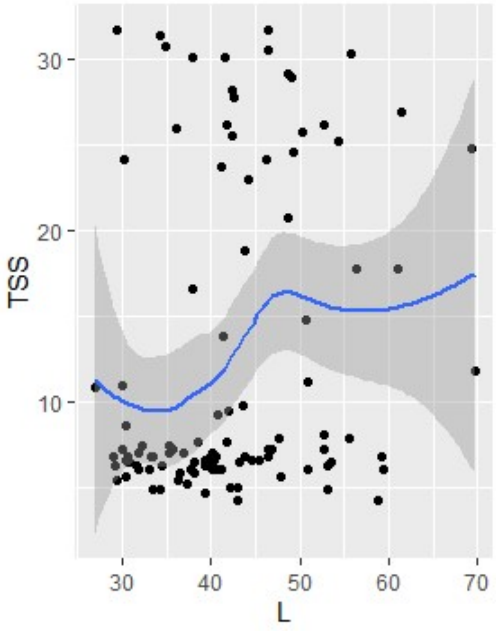

Fig. 20. Scatterplot of *L** and TSS for device 1

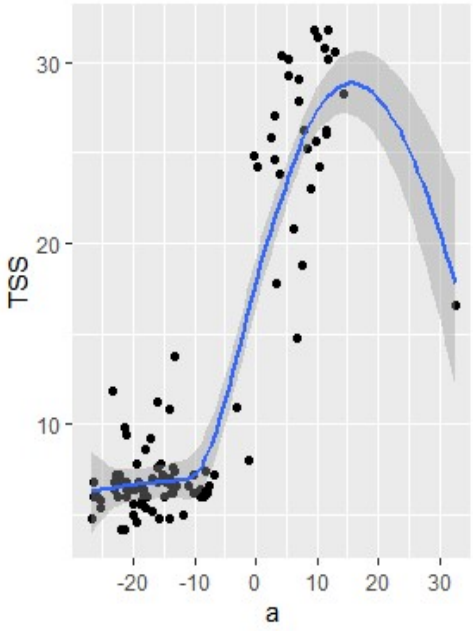

Fig. 21. Scatterplot of *a** and TSS for device 1

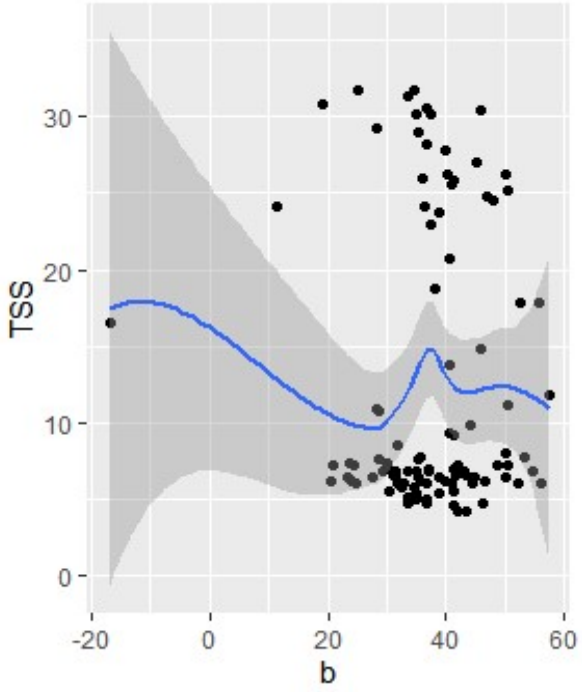

Fig. 22. Scatterplot of *b** and TSS for device 1

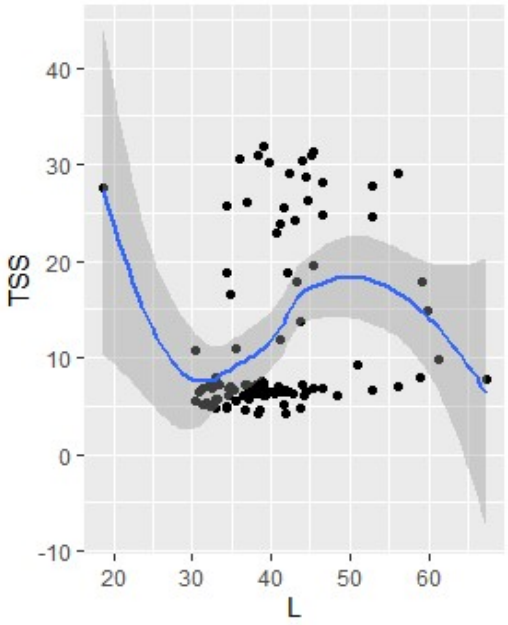

Fig. 23. Scatterplot of L^* and TSS for device 2

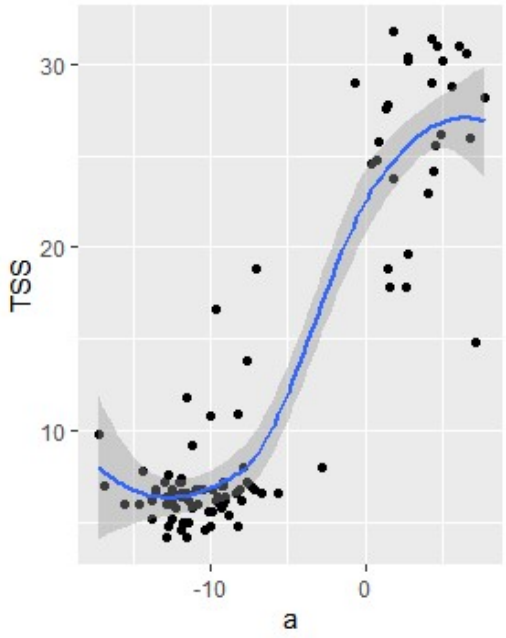

Fig. 24. Scatterplot of a^* and TSS for device 2

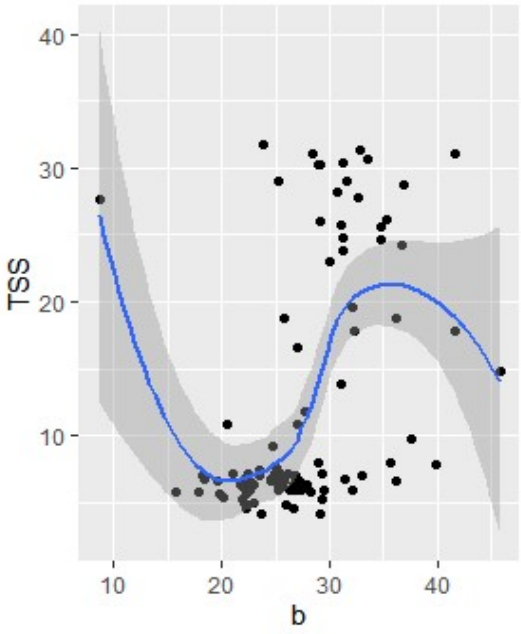

Fig. 25. Scatterplot of b^* and TSS for device 2

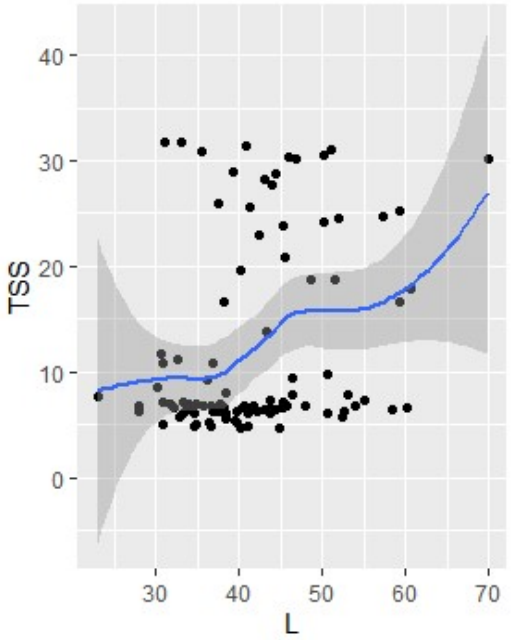

Fig. 26. Scatterplot of L^* and TSS for device 3

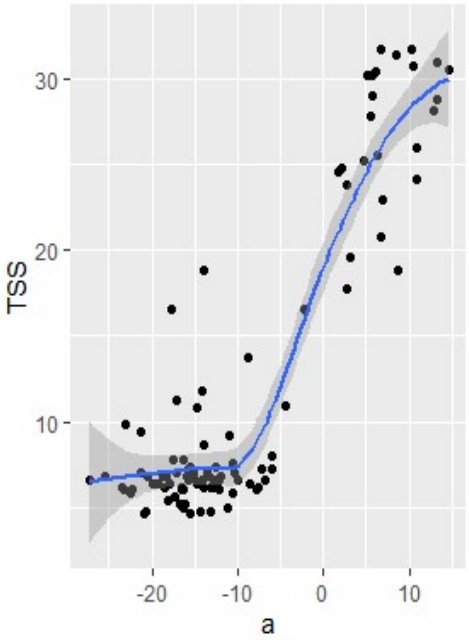

Fig. 27. Scatterplot of a^* and TSS for device 3

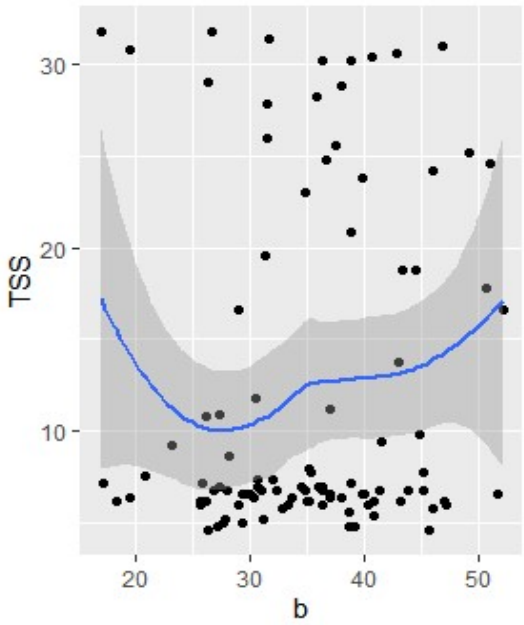

Fig. 28. Scatterplot of b^* and TSS for device 3

4.4 SPLINE MODEL FITTING BETWEEN *L**, *a**, *b** AND TSS

Once the linear regression model found to be unfitted for specifying the relationship between *L**, *a**, *b** and TSS, spline regression model was tried for fitting data points. The s function is a symbolic wrapper that is used in the model to signify a smooth term.

Table 13. Parametric coefficients of spline regression model for *L*a*b** and TSS; device 1

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	12.45	0.26	46.52	<2e-16 ***

Table 14. Approximate significance of smooth terms for device 1

	edf	Ref.df	F	p-value
s(L)	2.92	3.66	2.74	0.02832 *
s(a)	7.85	8.63	102.88	< 2e-16 ***
s(b)	1.00	1.00	6.99	0.00962 **
Adjusted R^2				0.91
Deviance explained				92%
GCV				8.50

For device 1, it is clear from the Table 14 that the smoothing terms of *L**, *a** and *b** were significant. It is evident that *L**, *a**, *b** were significant covariates and *a** was the most significant covariate since the corresponding p-value is the lowest (<2e-16). The p-value associated with each smooth indicates whether it is significantly different from 0 or not. The adjusted R^2 value obtained was 0.91 which is close to 1 and is good. The proportion of the total deviance explained by the current model was 92 percent.

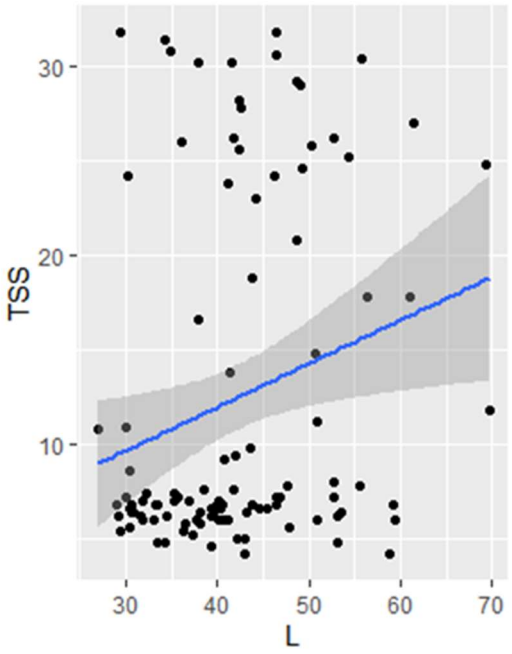

Fig. 29. Spline fitting between *L** and TSS for device 1

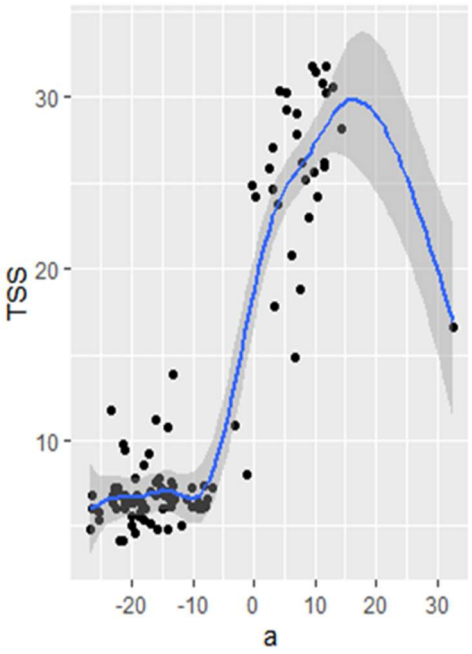

Fig. 30. Spline fitting between *a** and TSS for device 1

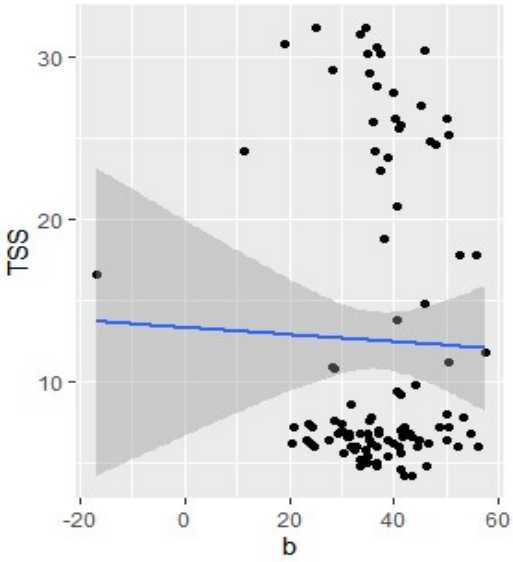

Fig. 31. Spline fitting between *b** and TSS for device 1

Table 15. Actual and predicted values of TSS using spline model using device 1 data

SL NO	ACTUAL TSS	PREDICTED TSS
1	4.6	5.72
2	6.2	12.52
3	6.5	7.31
4	29.0	27.34
5	27.6	26.57
6	5.2	5.82
7	18.8	6.79
8	31.0	28.49
9	28.8	27.69
10	31.0	27.37
11	8.0	8.96
12	7.2	9.27
13	6.2	6.35
14	6.0	6.43
15	16.6	12.55
16	7.4	7.30
17	6.8	6.69
18	7.0	6.88
19	6.8	6.54
20	4.6	5.64
21	17.6	8.27
22	7.2	5.99
23	6.0	6.51
24	19.6	24.83
25	6.6	7.03

Table 16. Parametric coefficients of spline regression model for L*a*b* and TSS; device 2

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	12.21	0.27	43.67	<2e-16 ***

Table 17. Approximate significance of smooth terms for device 2

	edf	Ref.df	F	p-value
s(L)	6.18	7.23	1.30	0.2458
s(a)	8.26	8.82	54.59	<2e-16 ***
s(b)	5.29	6.37	2.68	0.0217 *
Adjusted R^2				0.90
Deviance explained				92.3%
GCV				10.01

It is evident from Table 17 that the smoothing terms of a^* and b^* were significant. a^* found to be most significant with p-value <2e-16 for device 2 as well. The adjusted R^2 value was 0.90 and 92.3 percent deviance was explained by the model.

Table 18. Actual and predicted values of TSS using spline model using device 2 data

SL NO	ACTUAL	PREDICTED
1	29.2	28.6
2	4.6	7.29
3	6.2	7.45
4	6.5	7.52
5	6.8	7.24
6	31.8	29.67
7	30.8	26.04
8	7.4	6.48
9	9.4	7.38
10	6.0	7.18
11	25.2	22.21
12	26.2	21.63
13	5.8	7.40
14	6.4	7.34
15	8.6	6.91
16	6.8	6.64
17	11.2	9.03
18	6.2	7.18
19	7.0	10.24
20	5.4	7.00
21	24.2	17.78
22	17.6	12.62
23	6.0	7.65
24	6.6	9.24

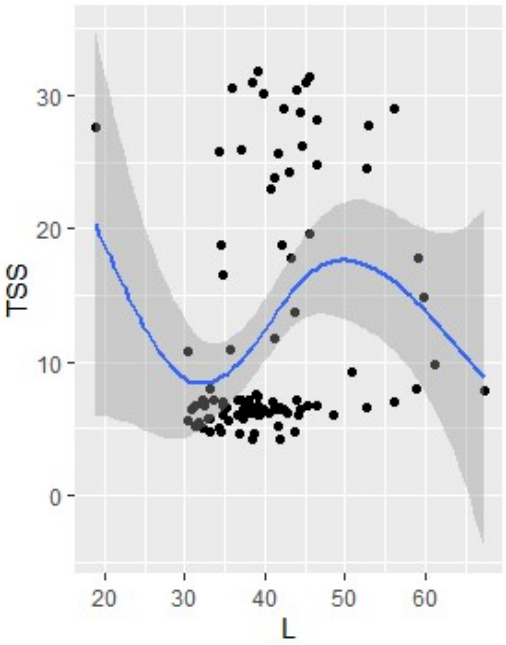

Fig. 32. Spline fitting between *L** and TSS for device 2

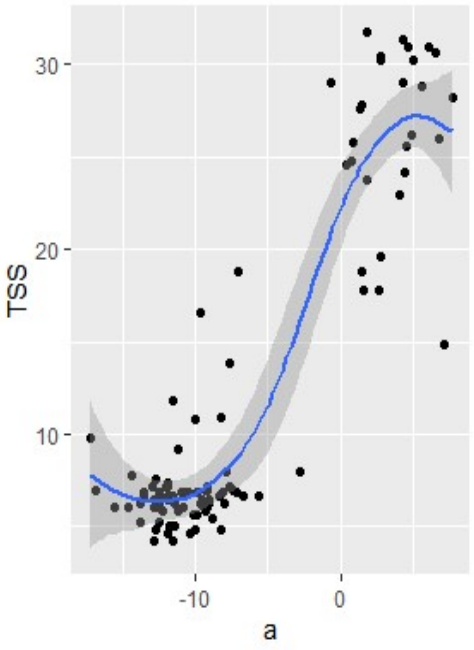

Fig. 33. Spline fitting between *a** and TSS for device 2

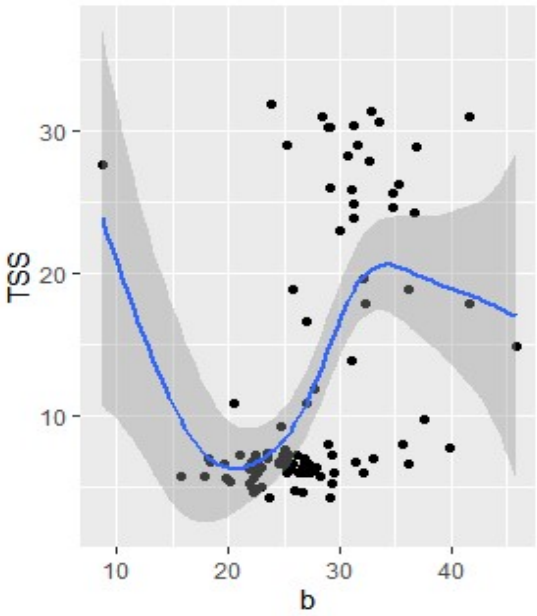

Fig. 34. Spline fitting between *b** and TSS for device 2

Table 19. Parametric coefficients of spline regression model for L*a*b* and TSS; device 3

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	12.20	0.28	42.33	<2e-16 ***

Table 20. Approximate significance of smooth terms for device 3

	edf	Ref.df	F	p-value
s(L)	1.71	2.17	2.09	0.133
s(a)	5.26	6.40	119.51	<2e-16 ***
s(b)	1.00	1.00	0.16	0.685
Adjusted R^2				0.89
Deviance explained				90.7%
GCV				9.04

From Table 20, a* was the most significant with corresponding p-value <2e-16. L* and b* were found to be non-significant. Adjusted R^2 was 0.89 and 90.7 percent variance was explained by the model.

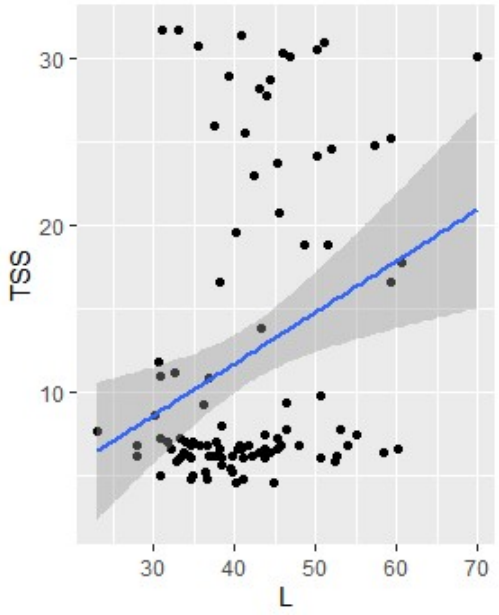

Fig. 35. Spline fitting between *L** and TSS for device 3

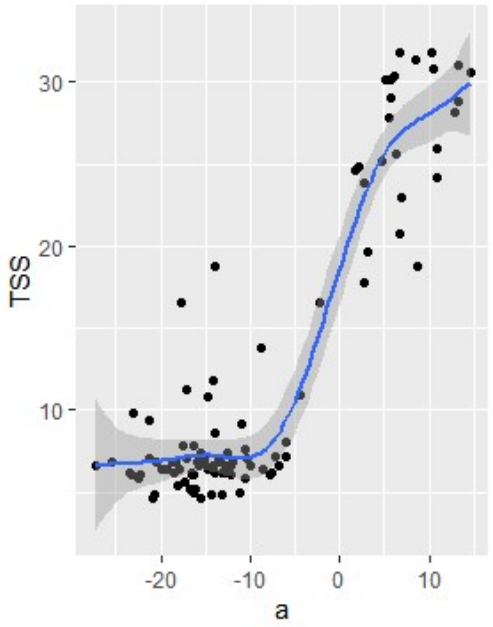

Fig. 36. Spline fitting between *a** and TSS for device 3

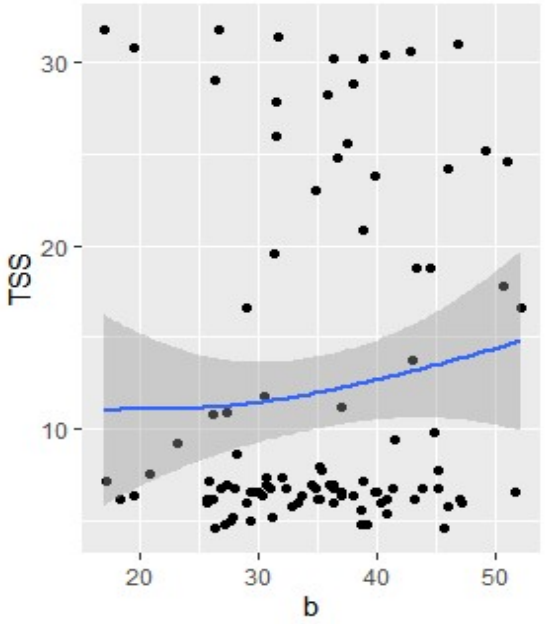

Fig. 37. Spline fitting between *b** and TSS for device 3

Table 21. Actual and predicted values of TSS using spline model using device 3 data

SL NO	ACTUAL	PREDICTED
1	6.0	7.14
2	25.8	20.65
3	4.2	7.12
4	6.5	11.32
5	6.8	6.81
6	29.0	27.02
7	27.6	27.15
8	6.0	7.35
9	26.2	29.25
10	7.2	6.21
11	6.0	6.95
12	31.0	27.98
13	6.4	9.65
14	5.8	6.60
15	6.4	6.50
16	6.6	8.89
17	7.6	6.88
18	7.4	6.72
19	7.2	7.08
20	6.8	5.63
21	7.2	6.01
22	4.6	5.92
23	26.2	28.65

4.5 COMPARISON OF FITTED MODELS AND IDENTIFICATION OF SUITABLE MODEL

When the linear model was fitted between TSS and *RGB* values for all the devices, each of the independent variables were found to be significant. For device 1, when the linear model was fitted between *R*, *G*, *B* and TSS, all the three parameters were found to be significant for calculating TSS. Adjusted R^2 for the model was 0.80. In case of device 2, *R*, *G* and *B* were significantly related to TSS when the linear model was fitted. 0.80 was the adjusted R^2 for corresponding. For device 3, *R* and *G* found to be significant with an adjusted R^2 of 0.84 for linear model fitting.

For device 1, L*, a*, b* and TSS linear model showed that a^* and b^* were the significant parameters with adjusted R^2 value 0.78. Similarly linear model for device 2 revealed that only a^* is significant with adjusted R^2 value of 0.81. For device 3, L^* and a^* were found to be significant with adjusted R^2 value of 0.85.

Even though the adjusted R^2 for three models were much closer, a^* is found to be significant in all the cases. When prediction accuracy was compared for linear model between TSS vs *RGB* and TSS vs $L^*a^*b^*$, *RGB* model found to predict TSS much accurately than $L^*a^*b^*$ color space. Therefore, *RGB* is found to be a better fit in linear models.

Spline model between L^*, a^*, b^* and TSS for device 1 showed that the smoothing terms of L^*, a^* and b^* were significant with an adjusted R^2 of 0.91. For device 2, smoothing terms of a^* and b^* were significant with adjusted R^2 of 0.90. In case of device 3, smoothing term of a^* was significant with adjusted R^2 of 0.89. R-square values obtained were much higher for spline model fitting of TSS and $L^*a^*b^*$ values than linear models with a good percentage of explained variation. Spline regression model was found to be the best fit for L^*, a^*, b^* and TSS values. R-square values were much higher with a good percentage of variation explained.

4.6 PROTOCOL FOR ACCURATE COLOUR PARAMETERS MEASUREMENT USING SMARTPHONE

This study was conducted to pave the way for easy maturity assessment of bananas without involving human perception during the process.

4.5.1 Sample collection

Choose good quality bananas at the green stage with minor shape defects randomly. Identify healthy banana plants from randomly selected fields and collect three hands from each plant with minimum 10 fingers by hand.

4.5.2 Image collection

- Capture images inside a box model structure (plate 6) to avoid light interference and use black as background.
- Use white light illumination inside the box to complete the experimental setup. Ensure that this illumination is as evenly distributed as possible.
- While using a smartphone for taking images keep it in manual mode to avoid resetting and do not mark over the banana while taking photos. Capture images from the same distance. A distance of 20 cm was used in the study.

4.6.3 TSS measurement

Measure TSS using handheld refractometer at various ripening time points from harvest to fruit senescence.

4.6.4 Image calibration and pre-processing

- Calibrate the image to bring the *RGB* values of color patches in the input image to the standard *RGB* values. Color calibration eliminates variation in lighting conditions to avoid the variation in pixel intensities.
- Pre-processing is used to improve image data by suppressing unwanted distortions or enhancing certain image features that are useful for future processing.

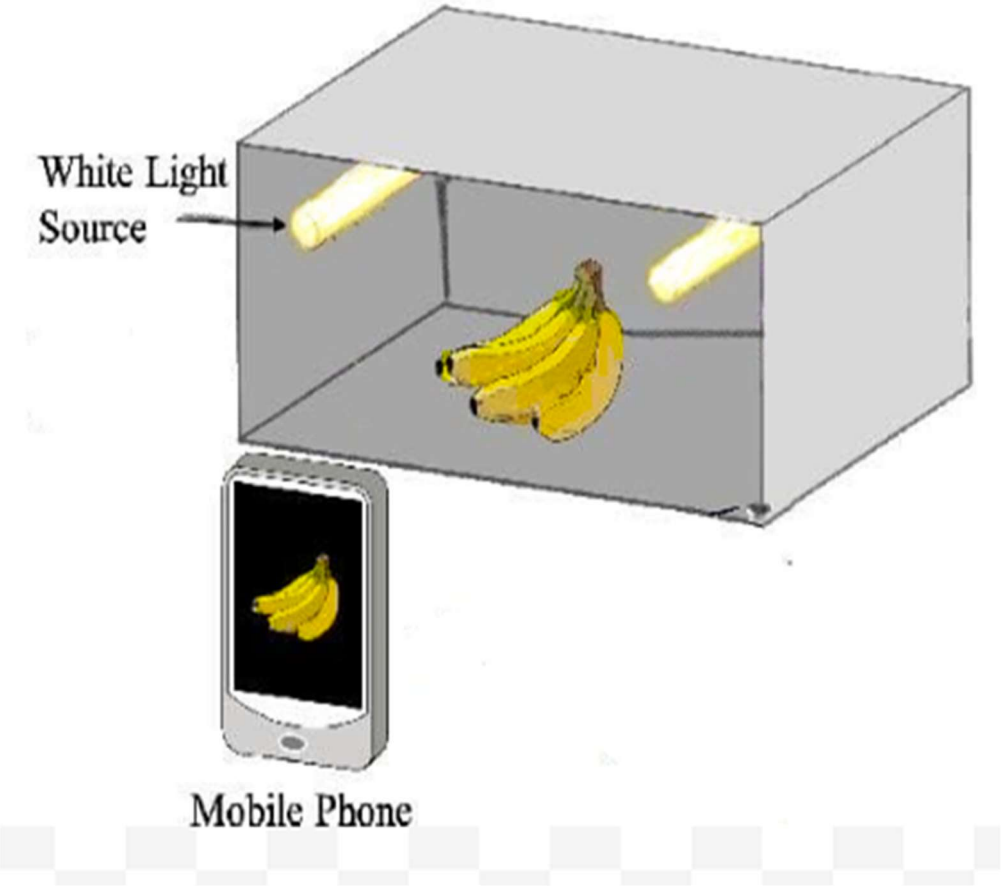

Plate 6. Box model for image capturing

4.6.5 Image analysis

ImageJ is an open-source java-based software for processing and analyzing scientific images developed at the National Institute of Health and the Laboratory. Upload the collected images in imageJ for colour measurements.

4.6.6 Background elimination and color space conversion

Select Region of interest (ROI) from the image and remove background. RGB values can be directly obtained using imageJ. Then convert the image to image-stack containing 3 slices *L**, *a**, *b** and measure the values.

4.6.6 Model fitting

Tabulate *R*, *G*, *B*, *L**, *a**, *b** and TSS values. Then visualize the data and find the relationship among them. Fit linear or nonlinear regression based on visualization and find the best fitting model.

4.7 CONVOLUTIONAL NEURAL NETWORK

A novel CNN architecture was used for the classification and detection of banana ripening stages by analyzing the collected images of it. Tensor-flow was used as the main framework to implement the proposed Convolutional Neural Network architecture.

A 14- layer convolutional neural network was developed. There were 6 convolutional layers and 2 max pooling-2d layers in the structure. After the flattening layer, two dense layers were implemented.

4.7.1 Samples

Three categories with 30 images each were selected as sample. The categories were shown in Table 21.

Table 22. Identified categories of banana and TSS for the respective categories

Category	TSS
Raw	4-10
Medium	11-17
Ripe	18-32

From the sampling frame 30 images with suitable TSS in the range were chosen and separated. For the training set 25 images from each category were selected. And for the test set 5 images from each category were selected.

The R script for building the CNN model is given in Appendix 1. Model detail is shown below:

Model: "sequential_2"

Layer (type)	Output Shape	Param #
conv2d_11 (Conv2D)	(None, 98, 98, 100)	2800
conv2d_10 (Conv2D)	(None, 96, 96, 80)	72080
max_pooling2d_3 (MaxPo	(None, 32, 32, 80)	0
dropout_5 (Dropout)	(None, 32, 32, 80)	0
conv2d_9 (Conv2D)	(None, 30, 30, 80)	57680
conv2d_8 (Conv2D)	(None, 28, 28, 80)	57680
conv2d_7 (Conv2D)	(None, 26, 26, 80)	57680
conv2d_6 (Conv2D)	(None, 24, 24, 80)	57680
max_pooling2d_2 (MaxPo	(None, 8, 8, 80)	0
dropout_4 (Dropout)	(None, 8, 8, 80)	0
flatten_1 (Flatten)	(None, 5120)	0

```
_________________________________________________________________
dense_3 (Dense)        (None, 500)         2560500
_________________________________________________________________
dropout_3 (Dropout)    (None, 500)         0
_________________________________________________________________
dense_2 (Dense)        (None, 3)           1503
=================================================================
Total params: 2,867,603
Trainable params: 2,867,603
Non-trainable params: 0
_________________________________
```

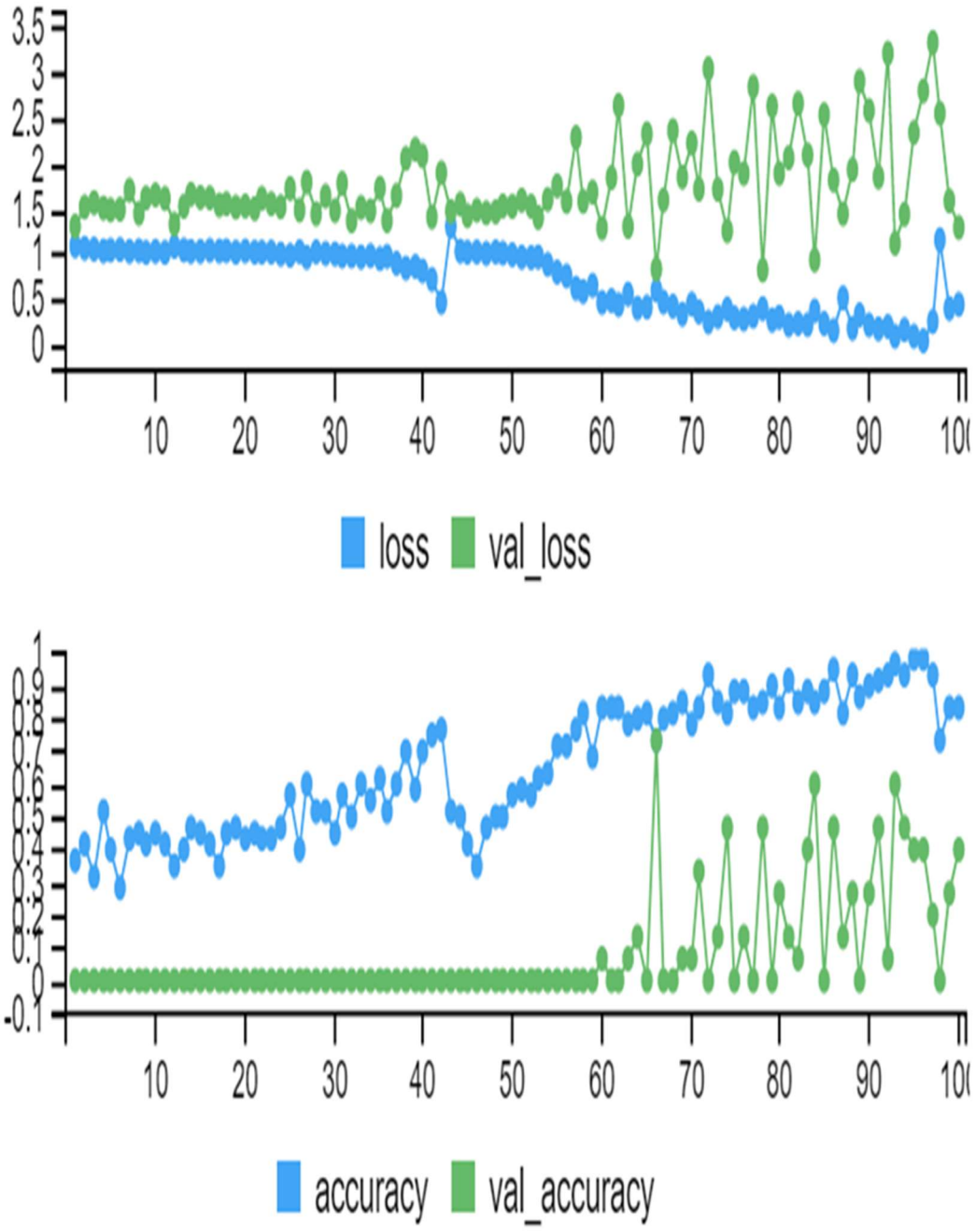

Fig. 38. Graphic of loss and accuracy for training and validation sets

On fitting the training data, each epoch appeared in the console area. While the epochs were running, a graphic was obtained in the RStudio viewer (Fig. 38). 100 epochs were used.

These are the juxtaposed curves of loss and accuracy for training and validation sets. Once the training was completed, a freshly trained model was evaluated.

Table 23. Model Evaluation Training set

loss	accuracy
0.4586261	0.8400000

Table 24. Evaluation Result (Training Set)

	Ripe	Raw	Medium
Ripe	25	0	3
Raw	0	24	8
Medium	0	1	14

Table 25. Model Evaluation Test set

loss	accuracy
0.81	0.73333

Table 26. Evaluation Result (Test Set)

	Ripe	Raw	Medium
Ripe	5	0	0
Raw	0	4	3
Medium	0	1	2

The model evaluation of the trained set revealed 84 percent accuracy with loss percent of 45 (Table 23). The evaluation result of the trained set is given in table 22. It is evident from the table 24 that the model successfully identified and categorized all 25 ripe fruits. From the 25 raw images, the model was able to categorize 24 raw and one identified as medium category. From the 25 medium ripe images, 14 images were categorized as medium ripe, 8 were identified as raw and 3 were identified as ripe.

Then the trained model was tested and evaluated over a test set. On evaluation of the model for the test set, it was found from table 25 that the model classified the test images of bananas into raw, ripe and medium with 73 percent accuracy and loss of 81 percent.

The developed model successfully identified all 5 ripe images from test set. From the raw image set, the model identified 4 raw and 1 image was identified as medium ripe category. In the case of 5 medium ripe images, 3 were identified as raw and 2 were identified as medium category (Table 26). It was found that the test dataset's accuracy is slightly lower than the training dataset's accuracy. Overfitting is the reason for this gap between training accuracy and test accuracy. When a machine learning model overfits, it performs worse on new data than it did on training data.

The test set images of bananas classified by the model is given in Fig. 39. Since the sample size was much smaller for developing a neural network, more categorization based on TSS and color parameters was difficult. In practice, at least 10,000 images are needed for building a good CNN model.

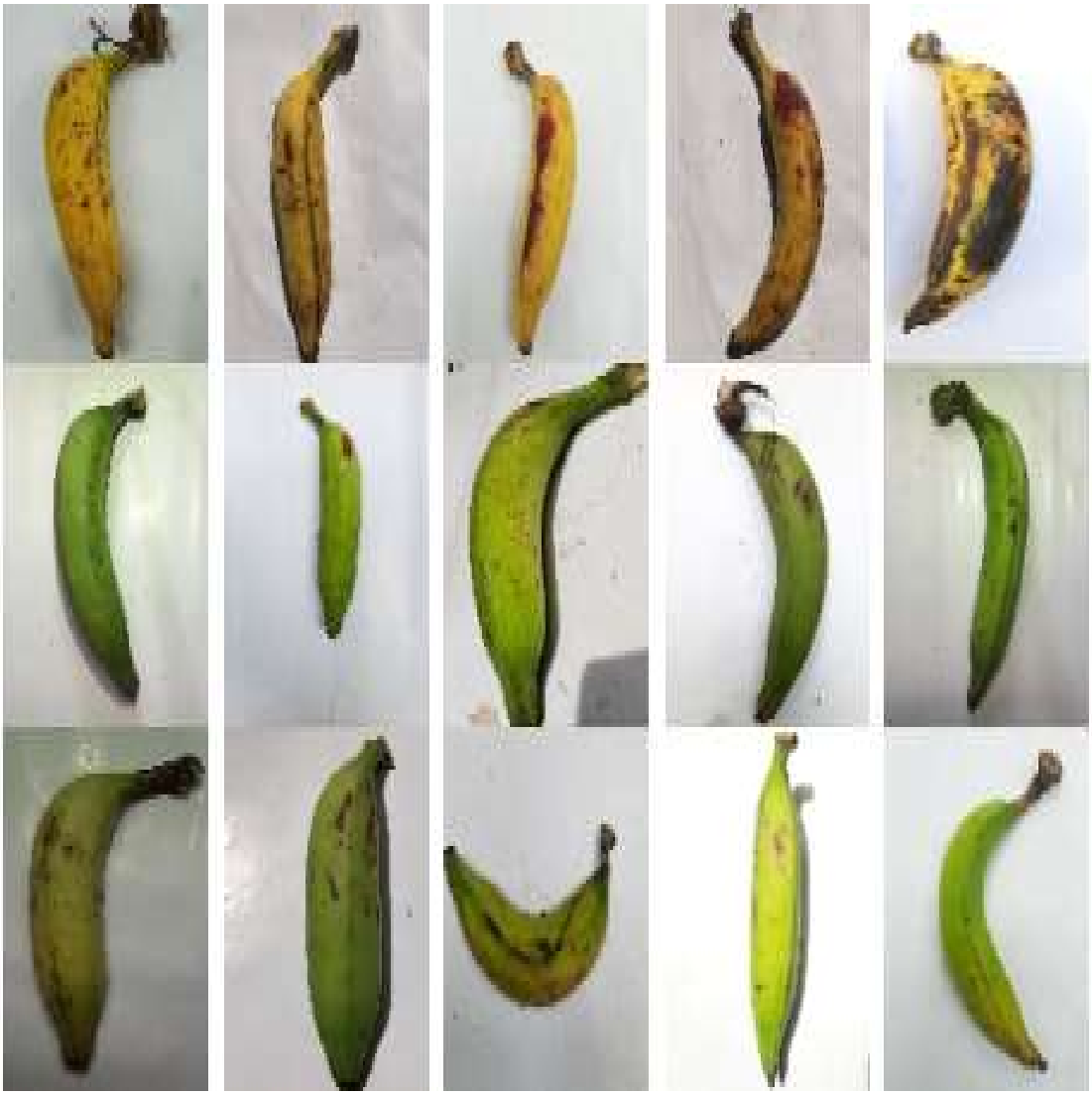

Fig. 39. Diagram of test images identified and classified by model

Summary

5. SUMMARY

An optimum ripening stage for bananas is required for quality control and consumer acceptance. In image analysis, color is a major descriptor and influential attribute. Computer-based image processing approaches can provide an automated and non-destructive tool for banana ripening classification. Recent advances in deep learning have led to breakthroughs in long-standing tasks such as feature extraction, image segmentation, and image classification in the field of artificial intelligence. Convolutional neural network (CNN) is one of the most successful approaches among these, with a wide range of applications in image classification.

The study entitled "Statistical assessment of banana ripening using smartphone-based images" conducted in the Department of Agricultural Statistics, College of Agriculture, Vellayani during the year 2020-21 was an attempt to assess the ripening stages of bananas using the images of bananas captured with smartphone by predicting the TSS. In addition to this, a CNN model was developed to classify the images into different stages of ripening by analyzing the color values from the images. The color parameters *R*, *G*, *B*, *L**, *a**, and *b** were extracted from the images. The TSS of bananas were measured using a handheld refractometer.

Initially, the bananas were collected from randomly identified Nendran fields in the Trivandrum district. Images of bananas were taken using three different smartphones having different resolution characteristics. Banana samples were placed in a box model structure to avoid outside light interference with a non-reflecting white paper as a background of the image. For white light illumination, two of 36 W fluorescent lamps were fixed at the ceiling above the experiment setup. Smartphones were placed at a distance of 20 cm above the banana. For each set of bananas, the images were captured from the day of collection to ripened stage. TSS was measured daily during 10-12 days from the day of collection.

ImageJ software was used for extracting the color values from the image. The collected images were pre-processed to eliminate the background interference. The *RGB* values were measured using the *RGB* measure plugin. Then using converting an *RGB* image to *Lab* stack plugin the image is converted to a new image-stack containing *L*, *a*, *b* stacks, and values were measured for corresponding images for all devices.

For all of the images acquired with three smartphones, a linear model was fitted between *R*, *G*, *B*, and TSS. For fitting linear models, a machine learning model was used. A total of 80 percent of the sample was used for training and 20 percent for testing.

For device 1, out of the total 129 samples, 104 samples were taken as training and 25 samples were taken as the test set. *R*, *G,* and *B* were found to have a significant relationship with TSS. Since p-values for *R* and *G* were less than 0.01, they were found to be significant at 1 percent level of significance. *B* was found to be significant at 5 percent level of significance. Adjusted R^2 obtained was 0.80 which was closer to 1 indicating a good fit of the model. The RSE was 4.04 corresponding to a 29.73 percent error rate. The RMSE and R^2 value for predictions were 3.83 and 0.84 respectively.

For device 2, out of the total 126 samples, 102 samples were taken as training and 24 samples were taken as the test set. *R*, *G,* and *B* values had a significant relationship with TSS. *R* and *G* were found to be significant at 1 percent level of significance. *B* was found to be significant at 5 percent level of significance. Adjusted R^2 obtained was 0.80. The RSE obtained was 4.02 corresponding to a 23.70 percent error rate. The RMSE and R^2 value for predictions were 2.98 and 0.90 respectively.

For device 3, out of the total 122 samples, 99 samples were taken as training and 23 samples were taken as the test set. *R* and *G* values were found to have a significant relationship with TSS. *R* and *G* were found to be significant at 1 percent level of significance. *B* value was found to be non-significant. Adjusted $R^2$was 0.84. The RSE for the model was 3.57 corresponding to 18.67 percent error rate. The RMSE and R^2 value for predictions were 2.22 and 0.95 respectively.

Since the *RGB* color model is a device-dependent model, an *RGB* value may not always represent the same color on different devices. Therefore, the extracted *L*a*b** color values were used for fitting the linear model.

The parameters *a** and *b** had a significant relationship with TSS for the images captured using device 1. *a** was found to be significant at 1 percent level of significance and *b** was found to be significant at 5 percent level of significance. Adjusted R^2 obtained was 0.78. The RSE was 4.25 corresponding to a 31.39 percent error rate. The RMSE and R^2 value for predictions were 4.04 and 0.84 respectively.

For device 2, only a^* and TSS were found to have a significant relationship. a^* found to be significant at 1 percent level of significance. L^* and b^* were found to be non-significant. Adjusted R^2 obtained was 0.81. The RSE obtained was 3.90 corresponding to a 22.78 percent error rate. The RMSE and R^2 value for predictions were 2.86 and 0.91 respectively.

For device 3, L^* and a^* were found to have a significant relationship with TSS. L^* and a^* were significant at 5 percent level of significance. Adjusted R^2 was 0.85 and RSE was 3.48 with an 18.17 error rate. The RMSE and R^2 value for predictions were 2.16 and 0.95 respectively.

When data has been visualized to see the trend of data points, a non-linear relationship was found between color parameters and TSS for all three devices. Then, the spline regression model was tried for fitting data points.

For device 1, the smoothing terms L^*, a^*, and b^* were significant. The adjusted R^2 value obtained was 0.91 which is close to 1 and is good. The proportion of the total deviance explained by the current model was 92 percent. The smoothing terms of a^* and b^* were significant. a^* found to be most significant with p-value <2e-16 for device 2 as well. The adjusted R^2 value was 0.90 and 92.3 percent deviance was explained by the model. For device 3, a* was the most significant with a corresponding p-value <2e-16. L* and b* were found to be non-significant. Adjusted R^2 was 0.89 and 90.7 percent variance was explained by the model.

When a linear model was fitted between TSS and *RGB* values for all the devices, each of the independent variables was found to be significant. When the spline model was fitted, the adjusted R^2 for the three models were much closer, a^* is found to be significant in all the cases. When prediction accuracy was compared for linear model between TSS vs *RGB* and TSS vs $L^*a^*b^*$, the *RGB* model was found to predict TSS much accurately than $L^*a^*b^*$ color space. Therefore, *RGB* is found to be a better fit in the linear model. The spline regression model was found to be the best fit for L^*, a^*, b^*, and TSS values. R-square values were much higher with a good percentage of variation explained.

Also, a protocol for accurate data collection was developed. A CNN model was developed for the classification and detection of banana ripening stages by analyzing the collected images of it. The model classified the test images of bananas into raw, ripe, and medium with 73 percent accuracy and loss of 81 percent. When the trained model was tested and evaluated over the test set, it was found that the model classified the test images of bananas into raw, ripe, and medium with 73 percent accuracy and loss of 81 percent.

5.1 SUGGESTIONS

- The present study is limited to one variety of bananas. This can be further extended to different varieties to understand the trend in ripening.
- The study can be continued further with a higher number (at least 10,000) of samples to improve the efficiency of classification using deep learning.
- The study can be extended to different fruits having a peel color change during ripening.

References

6. REFERENCES

Abbott, J. A. 1999. Quality measurement of fruits and vegetables. Postharvest Biol. Technol. 15(3): 201-225.

Abdullah, M. Z., Aziz, S. A., and Mohamed, A. M. D. 2000. Quality inspection of bakery products using a color-based machine vision system. *J. Food Qual*. 23(1): 39–50.

Abdullah, M.Z., Guan, L.C., Lim, K.C., and Karim, A.A. 2004. The applications of computer vision system and tomographic radar imaging for assessing physical properties of food. *J. Food Eng*. 61: 125–135.

Adebayo, S. E., Hashim, N., Abdan, K., Hanafi, M., and Mollazade, K. 2016. Prediction of quality attributes and ripeness classification of bananas using optical properties. *Sci. Hortic.* 212, pp. 171-182.

Ahmad, I. S., Reid, J. F., Paulsen, M. R., and Sinclair, J. B. 1999. Color classifier for symptomatic soybean seeds using image processing. *Plant Dis*. 83, 320–327.

Ali, M. M., Janius, R. B., Nawi, N. M., and Hashim, N. 2018. Prediction of total soluble solids and pH in banana using near infrared spectroscopy. *J. Eng. Sci. Technol*. Vol 13, pp. 254-264.

Alimardani, R., Soltani, M., and Omid, M. 2011. Evaluating banana ripening status from measuring dielectric properties. *J. Food Eng*. 105(4): 625-631.

Ammawath, W., Che Man, Y.B., Yusof, S., and Rahman, R.A. 2001. Effects of variety and stage of fruit ripeness on the physicochemical and sensory characteristics of deep-fat fried banana chips. *J. Sci. Food Agric.* 81: pp. 1166-1171.

Bapat, V. A., Trivedi, P. K., Ghosh, A., Sane, V. A., Ganapati, T. R., and Nath, P. 2010. Ripening of fleshy fruit: Molecular insight and the role of ethylene. Biotechnol. Adv. 28(1): 94-107.

Bhosale, S. and Sankhe, D. 2015. Ripeness inspection system for Banana. *Int. J. Comput. Appl.* vol 6, pp. 0975-8887.

Blasco, J., Alexios, N., and Molto, E. 2003. Machine vision system for automatic quality grading of fruit. *Biosystems Eng*. 85(4): 415-423.

Carvalho, G. B. M., Silva, D. P., Santos, J. C., Filho, H. J. I., Vicente A. A., Teixeira, J. A., Felipe, M. G. A., and Silva, J. B. A. 2008. Total soluble solids from banana: evaluation and optimization of extraction parameters. *Appl. Biochem. Biotechnol*. 153: pp. 34-43.

Casady, W. W., Paulsen, M. R., Reid, J. F., and Sinclair, J. B. 1992. A trainable algorithm for inspection of soybean seed quality. *Trans. ASAE*, 35: 2027–2034.

Casasent, D. A., Sipe, M. A., Schatzki, T. F., Keagy, P. M., and Lee, L. C. 1998. Neural net classification of X-ray pistachio nut data. *Food Sci. Technol.-Lebensmitte/Wissenschaft & Technologie*, 31: 122–128.

Chaudhary, S. and Prajapati, B. 2014. Quality analysis and classification of Bananas. *Int. J. Advanced Res. Comp. Sci. Software Eng.* vol. 4, pp. 869-874.

Chen, C.R., and Ramaswamy, H.S. 2002. Color and texture change kinetics in ripening bananas. *Lebensm-Wiss u-Technol*. 35: pp. 415-419.

Dadzie, B.K., and Orchard, J.E. 1997. Routine Post-Harvest Screening of Banana/Plantain Hybrids: Criteria and Methods. INIBAP Technical Guidelines 2. International Plant Genetic Resources Institute, Rome, Italy, 61p.

Drury, R., Hörtensteiner, S., Donnison, I., Bird, C.R., and Seymour, G.B. 1999. Chlorophyll catabolism and gene expression in the peel of ripening banana fruits. *Physiol. Plant*. 107, pp. 32-38.

Du, C.-J. and Sun, D.-W., 2004. Recent developments in the applications of image processing techniques for food quality evaluation. *Trends Food Sci. Technol*. 15: 230–249.

Ghazanfari, A., Irudayaraj, J., and Kusalik, A. 1996. Grading pistachio nuts using a neural network approach. *Trans. ASAE*, 39: 2319–2324.

Goodrum, J. W. and Elster, R. T. 1992. Machine vision for crack detection in rotating eggs. *Trans. ASAE*, 35: 1323– 1328.

R. Hamza, R. and Chtourou, M. 2018. Apple ripeness estimation using Artificial Neural Network. *2018 Int. Conf. High Perform. Computing Simul.* pp. 229-234.

Hufkens, K., Melaas, E. K., Mann, M. L., Foster, T., Ceballos, F., Robles, M., and Kramer, B. 2019. Monitoring crop phenology using a smartphone based near-surface remote sensing approach. *Agric. For. Meterolo.* 265: 327-337.

Intaravanne, Y., Sumriddetchkajorn, S., & Nukeaw, J. 2012. Cell phone-based two-dimensional spectral analysis for banana ripeness estimation. *Sensors and Actuators B: Chemical*, *168*, 390–394.

J Hou, Y Hu, L Hou, K Guo, T Satake, 2015. Classification of ripening stages of bananas based on support vector machine. *Int. J. Agric. Biol. Eng*. 8(6), pp. 99–103.

Jamaludin, D., Abd Aziz, S., and Ibrahim, N. U. A. 2014. Dielectric based sensing system for banana ripeness assessment. *Int. J. Environ. Sci. Dev.* 5: 286-289.

Kamble, P. R., Marathe, R. S., Jha, S. K., Ranvare, S. S., and Katti, J. V. 2020. Development of an effective system to identify fruit ripening stage for Apple, Banana and Mango. *Int. J. Advanced Sci. Technol*. 29: 2766-2772.

Kang, S.P., East, A.R., and Trujillo, F.J. 2008. Color vision system evaluation of bicolor fruit: a case study with "B74" mango. *Postharvest Biol. Technol*. 49 (1), pp. 77-85.

Kavdir, I. and Guyer, D. E. 2002. Apple sorting using artificial neural networks and spectral imaging. *Trans. ASAE*, 45: 1995–2005.

Kays, S.J. 1991. *Postharvest Physiology of Perishable Plant Products*. Van Nostrand Reinhold, New York, 312p.

Kim, S. and Schatzki, T. F. 2000. Apple watercore sorting system using X-ray imagery: I. Algorithm development. *Trans. ASAE*, 43: 1695–1702.

Leemans, V., Magein, H., and Destain, M. F. 1998. Defects segmentation on 'Golden Delicious' apples by using color machine vision. *Comput. Electr. Agric*. 20: 117–130.

Li, J., Tan, J. and Shatadal, P. 2001. Classification of tough and tender beef by image texture analysis. *Meat Sci*. 57: 341–346.

Li, M., Slaughter, D.C., and Thompson, J.E. 1997. Optical chlorophyll sensing system for banana ripening. *Postharvest Biol. Technol.* 12(3), pp. 273-283.

Luo, X., Jayas, D. S., and Symons, S. J. 1999. Comparison of statistical and neural network methods for classifying cereal grains using machine vision. *Trans. ASAE*. 42: 413–419.

Mazen, F. M. A. and Nashat, A. A. 2019. Ripeness classification of Bananas using an Artificial Neural Network. *Arabian J. Sci. Eng*. 44, pp. 6901-6910.

McCarthy, M. J., Zion, B., Chen, P., Ablett, S., Darke, A.H., and Lillford, P.J. 1995. Diamagnetic Susceptibility changes in apple tissue after bruising. *J. Sci. Food Agric.* 67(1): 13-20.

Mendoza, F., Dejmek, P., and Aguilera, J. M. 2006. Calibrated color measurements of agricultural foods using image analysis. *Postharvest Biol. Technol*. 41, 285-295.

Meng-Han, H., Qing-Li, D., Bao-Lin, L., Lei-Qing, P., and Walshaw, J. 2015. Image segmentation of bananas in a crate using a multiple threshold method. *J. Food Process Eng.* pp. 1745-4530.

Miller, B. K. and Delwiche, M. J. 1989. A color vision system for peach grading. *Trans. ASAE*, 32: 1484–1490.

Mohapatra, A., Shanmugasundaram, S., and Malmathanraj, R. 2017. Grading of ripening stages of red banana using dielectric properties changes and image processing approach. *Comput. Electr. Agric*. 143: 100-110.

Moreno, J. L., Tran, T., Cantero-Tubilla, B., Lopez-Lopez, K., Lavalle, L. A. B. L., and Dufour, D. 2020. Physiochemical and physiological changes during the ripening of Banana (Musaceae) fruit grown in Colombia. *Int. J. Food Sci. Technol.* Vol. 56, pp. 1171-1183.

Nakano, K. 1997. Application of neural networks to the color grading of apples. *Comput. Electr. Agric*. 18: 105–116.

Nakano, K., Kurata, K., and Kaneko, M. 1992. Studies on sorting systems for fruits and vegetables. *J. Soc. of Agric. Structures*, *Japan*, 23: 81-86.

Padda, M.S., do Amarante, C.V.T., Garcia, R. M., Slaughter, D.C., and Mitcham, E.J. 2011. Methods to analyze physico-chemical changes during mango ripening: a multivariate approach. *Postharvest Biol. Technol*. 62: 267-274.

Paliwal, J., Visen, N. S., and Jayas, D. S. 2001. Evaluation of neural network architectures for cereal grain classification using morphological features. *J. Agric. Eng. Res*. 79: 361–370.

Papadakis, S., Abdul-Malek, S., Kamdem, R.E., and Yam, K.L. 2000. A versatile and inexpensive technique for measuring color of foods. *Food Technol*. 54: 48–51.

Parker, A. K., Cortazar-Atauri, G., Geny. L., Spring, J., Irvine, J. D., Schultz, A., Molitor, H., Lacombe, D., Graca, T., Monamy, A., Stoll, C., Storchi, M., Trought, P., Hofmann, M., Van Leeuwen, R., and Amber, C. 2020. Temperature-based grapevine sugar ripeness modelling for a wide range of Vitis vinifera L. cultivars. *Agric. For. Meterolo*. 285-286.

Prasanna, V., Prabha, T. N., and Tharanathan, R. N. 2007. Fruit ripening phenomena-an overview. *Crit. Rev. Food Sci. Nutr.* 47(1): 1-19.

Ramaswamy, H.S., and Tung, M.A. 1989. Technical note: Textural changes as related to color of ripening bananas. *Int. J. Food Sci. Technol.* 24: pp. 217-221.

Richardson, J., Reiner, P., and Wilamowski, B. M. 2015. Cubic spline as an alternative to methods of machine learning. IEEE 13th International Conference on Industrial Informatics (INDIN), pp. 110-115.

Robertson, A. L. 1976. The CIE 1976 color difference formulae. Color Research Application, 2(1), pp. 7-11.

Ruan, R., Shu, N., Luo, L. Q., Xia, C., Chen, P., Jones, R., Wilcke, W., and Morey, R. V. 2001. Estimation of weight percentage of scabby wheat kernels using an automatic machine vision and neural network based system. *Trans. ASAE*, 44: 983–988.

Sanaeifar, A., Bakhshipour, A., and de la Guardia, M. 2016. Prediction of banana quality indices from color features using support vector regression. Talanta 148, pp. 54-61.

Sangwine, S.J. 2000. Color in image processing. *Electron. Commun. Eng. J*. 12(5): 211–219.

Segnini, S., Dejmek, P., and Oste, R. 1999. A low cost video technique for color measurement of potato chips. *Lebensm. -Wiss. U.-Technol*. 32: 216–222.

Shamili, M. 2019. The estimation of mango fruit total soluble solids using image processing technique. *Sci. Hortic.*249: 383-389.

Shearer, S. A. and Payne, F. A. 1990. Color and defect sorting of bell peppers using machine vision. *Trans. ASAE*, 33: 2045–2050.

So, J. D. and Wheaton, F. W. 1996. Computer vision applied to detection of oyster hinge lines. *Trans. ASAE*, 39: 1557–1566.

Sri, M. K., Saikrishna, K., and Kumar, V. V. 2020. Classification of ripening of banana fruit using convolutional neural networks. *Int. Conf. Innovative Advmt. Eng. Technol.* pp. 1-6.

Stokes, M., Anderson, M., Chandrasekar, S. and Motta, R. 1996. A standard default color space for the internet-sRGB, Version 1.10. In: *International Color Consortium (ICC)*, 1899 Preston White Drive, Reston, VA, November 5.

Sun, D. W. and Brosnan, T. 2003. Pizza quality evaluation using computer vision-part 1 pizza base and sauce spread. *J. Food Eng.* 57: 81–89.

Tao, Y., Heinemann, P. H., Varghese, Z., Morrow, C. T., and Sommer, H. J. 1995. Machine vision for color inspection of potatoes and apples. *Trans. ASAE*, 38(5): 1555–1561.

Tao, Y., Heinemann, P. H., Varghese, Z., Morrow, C. T., and Sommer, H. J. III. 1995. Machine vision for color inspection of potatoes and apples. *Trans. ASAE*, 38: 1555–1561.

Tapre, A. R. and Jain, R. K. 2012. Study of advanced maturity stages of Banana. *Int. J. Advanced Eng. Res. Stud.* Vol 1, pp. 272-274.

Utku, H. and Koksel, H. 1998. Use of statistical filters in the classification of wheats by image analysis. *J. Food Eng*. 36: 385–394.

Vizhanyo, T. and Felfoldi, J. 2000. Enhancing color differences in images of diseased mushrooms. *Comput. Electr. Agric*. 26: 187–198.

Von Loesecke, H.W. 1950. Bananas: Chemistry, Physiology, Technology. Interscience, New York, pp. 108-109.

Wainwright, H., and Hughes, P. 1989. Objective measurement of banana pulp color. *Int. J. Food Sci. Technol.* 24: pp. 553-558.

Wang, J., Tang, X. J., Chen, P. S., and Huanh, H. H. (2014). Changes in resistant starch from two banana cultivars during postharvest storage. Food Chemistry, 156, pp. 67-118.

Ward, G. and Nussinovitch, A. 1996. Peel gloss as a potential indicator of banana ripeness. *LWT- Food Sci. and Technol*. 29(3): 289-294.

Ward, G., and Nussinovitch, A. 1996. Peel gloss as a potential indicator of banana ripeness. *Lebensm-Wiss u-Technol*. 29: pp. 289-294.

Watanawan, C., Wasusri, T., Srilaong, V., Wongs-Aree, C., and Kanlayanarat, S. 2014. Near infrared spectroscopic evaluation of fruit maturity and quality of export Thai mango (*Mangifera indica* L. var. Namdokmai). *Int. Food Res. J*. 21(3): 1109-1114.

White, K.W. and Sellers, R. J. 1994. *Foreign materials sorting by innovative real time color signatures in Food Processing Automation III*. Proceedings of the FPAC Conference, ASAE, St. Joseph, Michigan, p. 29.

Williams, M. H., Vesk, M., and Mullins, M. G. 1989. A scanning electron microscope study of the formation and surface characteristics of the peel of the banana fruit during its development. *Bot. Gaz*. 150(1): 30–40.

Zulkifli, N., Hashim, N., Abdan, K., and Hanafi, M. 2016. Evaluation of physiochemical properties of *Musa acuminate cv.* Berangan at different ripening stages. *Int. Food Res. J. 23(suppl)*: 97-100.

Abstract

STATISTICAL ASSESSMENT OF BANANA RIPENING USING SMARTPHONE-BASED IMAGES

by

HARITHA R. NAIR

(2019-19-001)

ABSTRACT OF THE THESIS

Submitted in partial fulfilment of the

requirements for the degree of

MASTER OF SCIENCE IN AGRICULTURE

Faculty of Agriculture

Kerala Agricultural University

DEPARTMENT OF AGRICULTURAL STATISTICS

COLLEGE OF AGRICULTURE

VELLAYANI, THIRUVANANTHAPURAM – 695 522

KERALA, INDIA

2022

ABSTRACT

The research work entitled "Statistical assessment of banana ripening using smartphone-based images" was carried out at College of Agriculture, Vellayani during the period 2019 to 2021. The objectives were the development of suitable model to establish the relationship between Total Soluble Solids (TSS) and *L**(lightness), *a**(green-red ratios), *b**(blue-yellow ratios) values and for prediction of TSS values using *L**, *a**, *b** values. Development of a protocol for accurate data collection to assess TSS content in Banana using smart-phone-based images. Good quality Nendran variety with only minor shape and peel colour flaws were obtained from a nearest field randomly chosen for the study. Each time 3 hands at the ripening stage 1 (green) with 10 fingers by hand were collected. The fruits were stored in a normal day/ night cycle. Bananas were taken randomly from each hand and their color changes and development of brown spots were measured daily during 10-12 days. Banana samples were placed on the table covered with a non-reflecting white paper as a background of the image. For white light illumination, two of 36 W fluorescent lamps were fixed at ceiling above the experiment setup. Three smartphones were used for image acquisition. Smart phones were placed at a distance of 20 cm above the banana. Samples of banana were blended using a fruit juicer. The TSS were determined using a digital refractometer. For the images obtained, RGB and L*a*b* were extracted using ImageJ software. The observations on TSS, R, G, B, L*, a*, b* were used for fitting regression models after splitting the data into train (80%) and test (20%) sets.

When linear model was fitted between TSS and R, G, B values for all the three devices, each of the independent variables were found to be significant. Adjusted R-squared values obtained were 0.80, 0.80, and 0.84 for the three devices. It means about 80% of the variation in the TSS was explained by R, G, B values. For the predicted values of TSS R-squared values were 0.84, 0.90, and 0.95. Hence linear model was found to be better fit for predicting TSS. Since RGB color model is device dependent model, it may not always represent the same colour on different devices. But in case of CIE L*a*b*, it is device independent and shadows and areas of glossiness on the object surface had less impact. Therefore, linear model was fitted between TSS and L*, a*, b* values. Adjusted R-squared values obtained were 0.78, 0.81, and 0.85 for the three

devices. For the predicted TSS values R-squared values were 0.84, 0.76, and 0.95. Therefore, linear model between TSS and RGB model found to predict TSS much accurately than L*a*b* color space when prediction accuracy was compared.

On visualization of data, TSS and L*a*b* found to have non-linear relationship for all the devices. When spline regression was fitted between TSS and L*, a*, b* values R-Squared obtained were 0.91, 0.90, and 0.89, which was higher compared to R-squared values for linear model. Also, deviance explained by the models were 92%, 92.3%, and 90.7% for corresponding device 1,2 and 3. Therefore, spline regression found to be better model for TSS and L*, a*, b* data and for prediction of TSS values.

Protocol for accurate data collection was developed with modification in the procedure performed. Possibility of Deep learning was explored in the study using CNN. Convolutional neural network (CNN) was developed using 3 categories Raw (TSS 4-10), Medium (TSS 11-17) and Ripe (TSS 18-32) with 30 samples each. 25 images from each category were taken as training set and 5 were taken as test set. 100 epochs were performed to mitigate overfitting and to increase the generalization capacity of the neural network. Model evaluation of training set gave an accuracy of 84% with loss value 0.45. For the training set, all 25 from ripe category were able to identify into that particular category. In case of raw 24 were identified as raw with 1 identified as medium. For medium 14 were identified as medium,3 identified as ripe and 8 identified as raw. Model evaluation of test set provided 73% accuracy with 0.81 loss. The model successfully classified 5 ripe bananas, 4 raw bananas (1 classified as medium) and 2 medium bananas (3 classified as raw).

The results of the research work to identify the best fitting model concluded that RGB model found to predict TSS much accurately than L*a*b* color space when linear regression model was fitted and spline regression model was found to be the best fit for L*, a*, b* and TSS values, R-squared values were much higher with a good percentage of variation explained. The CNN developed classified images into raw, medium, and ripe with approximate accuracy of 74%. Therefore, CNN can be used to predict range of TSS in no time, if a large number of images are uploaded into this model. The CNN can be optimized further with higher number (atleast 10,000 samples) of samples to improve the efficiency of classification.

Appendices

APPENDIX 1

R code for CNN

```
library(reticulate)
install_miniconda(force = TRUE)
library(keras)
library(dplyr)
library(EBImage)
library(BiocManager)
install_keras()
BiocManager::install("EBImage")
#set working directory
setwd("D:/STUDENTS/Haritha/Pratheesh/raw")
card<-readImage("june_3.jpg")
print(card)
getFrames(card, type = "total")
display(card)
### access all ripe
setwd("D:/STUDENTS/Haritha/Pratheesh1/ripe")
img.card<- sample(dir()); #-------shuffle the order
cards<-list(NULL);
for(i in 1:length(img.card))
{ cards[[i]]<- readImage(img.card[i])
```

```
cards[[i]]<- resize(cards[[i]], 100, 100)} #resizing to 100x100
ripe<- cards # Storing stack of images in # matrix form in a list
#-------------------------------------------------------------
### access all raw
setwd("D:/STUDENTS/Haritha/Pratheesh1/raw")
img.card<- sample(dir()); #-------shuffle the order
cards<-list(NULL);
for(i in 1:length(img.card))
{ cards[[i]]<- readImage(img.card[i])
cards[[i]]<- resize(cards[[i]], 100, 100)} #resizing to 100x100
raw<- cards # Storing stack of images in # matrix form in a list
#-------------------------------------------------------------
### access all medium
setwd("D:/STUDENTS/Haritha/Pratheesh1/medium")
img.card<- sample(dir()); #-------shuffle the order
cards<-list(NULL);
for(i in 1:length(img.card))
{ cards[[i]]<- readImage(img.card[i])
cards[[i]]<- resize(cards[[i]], 100, 100)} #resizing to 100x100
medium<- cards # Storing stack of images in # matrix form in a list
#-------------------------------------------------------------
###Training Set
train_pool<-c(ripe[1:25],
```

```
 raw[1:25],
 medium[1:25]) # The first 25 images from each
# are included in the train set
train<-aperm(combine(train_pool), c(4,1,2,3)) # Combine and stacked
#### Test set
test_pool<-c(ripe[26:30],
 raw[26:30],
 medium[26:30]
 )
test<-aperm(combine(test_pool), c(4,1,2,3)) # Combined and stacked
##See images in test set
par(mfrow=c(3,5)) # To contain all images in single frame
for(i in 1:15){
 plot(test_pool[[i]])
}
par(mfrow=c(1,1)) # Reset the default
###### hot encoding for categorical data
#one hot encoding
train_y<-c(rep(0,25),rep(1,25),rep(2,25))
test_y<-c(rep(0,5),rep(1,5),rep(2,5))
#####
train_lab<-to_categorical(train_y) #Catagorical vector for training
#classes
```

```
test_lab<-to_categorical(test_y)#Catagorical vector for test classes
# Model Building
model.card<- keras_model_sequential() #-Keras Model composed of a
#-----linear stack of layers
model.card %>%
 #---------Initiate and connect to #----------------------------(A)-------------------------------
----#
 layer_conv_2d(filters = 100, #----------First convoluted layer
 kernel_size = c(3,3), #---40 Filters with dimension 4x4
 activation='relu', #-with a ReLu activation function
 input_shape = c(100,100,3)) %>%
 #----------------------------(B)---------------------------------#
 layer_conv_2d(filters = 80, #---------Second convoluted layer
 kernel_size = c(3,3), #---40 Filters with dimension 4x4
 activation='relu') %>%
 #----------------------------(C)---------------------------------#
 layer_max_pooling_2d(pool_size = c(3,3) )%>% #--------Max Pooling
 #-----------------------------------------------------------------#
 layer_dropout(rate = 0.25) %>% #-------------------Drop out layer
 #----------------------------(D)---------------------------------#
 layer_conv_2d(filters = 80, #-----------Third convoluted layer
 kernel_size = c(3,3), #----80 Filters with dimension 4x4
 activation='relu')%>% #--with a ReLu activation function
```

```
#--with a ReLu activation function
#----------------------------(E)--------------------------------#
layer_conv_2d(filters = 80, #----------Fourth convoluted layer
kernel_size = c(3,3), #----80 Filters with dimension 4x4
activation='relu') %>% #--with a ReLu activation function
#----------------------------(E.1)--------------------------------#
layer_conv_2d(filters = 80, #----------Fourth convoluted layer
kernel_size = c(3,3), #----80 Filters with dimension 4x4
activation='relu') %>% #--with a ReLu activation function
#----------------------------(E.2)--------------------------------#
layer_conv_2d(filters = 80, #----------Fourth convoluted layer
kernel_size = c(3,3), #----80 Filters with dimension 4x4
activation='relu') %>% #--with a ReLu activation function
#----------------------------(F)--------------------------------#
layer_max_pooling_2d(pool_size = c(3,3)) %>% #---------Max Pooling
#-----------------------------------------------------------------#
layer_dropout(rate = 0.35) %>% #-------------------Drop out layer
#----------------------------(G)--------------------------------#
layer_flatten()%>% #---Flattening the final stack of feature maps
#----------------------------(H)--------------------------------#
layer_dense(units = 500, activation='relu')%>% #-----Hidden layer
#--------------------------(I)--------------------------------#
layer_dropout(rate= 0.25)%>% #-------------------Drop-out layer
```

```
#-----------------------------------------------------------------#
layer_dense(units = 3, activation='softmax')%>% #-----Final Layer
#---------------------------(J)---------------------------------#
compile(loss = 'categorical_crossentropy',
optimizer = 'sgd',
metrics = c("accuracy")) # Compiling the architecture
model.card %>% summary()
#fit model
history<- model.card %>%
fit(train,
train_lab,
epochs = 100,
batch_size = 4,
validation_split = 0.2
)
model.card %>% evaluate(train,train_lab) #Evaluation of training set pred<-
model.card %>%
predict_classes(train) #-----Classification
pred<-model.card %>% predict(train) %>% k_argmax()
pred<-as.matrix(pred)
Train_Result<-table(Predicted = pred, Actual = train_y)
Train_Result
model.card %>% evaluate(test, test_lab) #-----Evaluation of test set
```

```
pred1<-model.card %>% predict(test) %>% k_argmax()

pred1<- pred<-as.matrix(pred1)

Test_Result<-table(Predicted = pred1, Actual = test_y) #-----Results

Test_Result

rownames(Train_Result)<-

 rownames(Test_Result)<-

 colnames(Train_Result)<-

 colnames(Test_Result)<-c("ripe", "raw", "Medium")

print(Train_Result)

print(Test_Result)
```

APPENDIX 2

The actual values and predicted values of TSS of testing data using linear model fitting between R, G, B and TSS for devices 1, 2 and 3 respectively

SL NO	ACTUAL TSS	PREDICTED TSS
1	4.6	5.06
2	6.2	6.34
3	6.5	5.82
4	29.0	23.46
5	27.6	26.02
6	5.2	5.48
7	18.8	10.19
8	31.0	23.33
9	28.8	27.84
10	31.0	24.98
11	8.0	14.11
12	7.2	13.47
13	6.2	6.73
14	6.0	5.90
15	16.6	18.24
16	7.4	10.64
17	6.8	6.76
18	7.0	6.03
19	6.8	2.46
20	4.6	3.95
21	17.6	13.00
22	7.2	3.86
23	6.0	6.54
24	19.6	22.46
25	6.6	10.13

SL NO	ACTUAL TSS	PREDICTED TSS
1	29.2	24.95
2	4.6	5.36
3	6.2	6.76
4	6.5	7.55
5	6.8	7.31
6	31.8	30.06
7	30.8	23.14
8	7.4	8.52
9	9.4	6.12
10	6.0	5.35
11	25.2	25.40
12	26.2	25.09
13	5.8	6.34
14	6.4	7.15
15	8.6	3.71
16	6.8	0.74
17	11.2	11.98
18	6.2	9.57
19	7.0	9.63
20	5.4	8.29
21	24.2	22.03
22	17.6	17.20
23	6.0	7.58
24	6.6	11.26

SL NO	ACTUAL TSS	PREDICTED TSS
1	6.0	7.85
2	25.8	23.23
3	4.2	6.35
4	6.5	8.45
5	6.8	6.75
6	29.0	23.51
7	27.6	26.07
8	6.0	5.11
9	26.2	27.67
10	7.2	6.23
11	6.0	9.88
12	31.0	25.88
13	6.4	5.54
14	5.8	5.95
15	6.4	5.80
16	6.6	6.64
17	7.6	6.59
18	7.4	6.58
19	7.2	7.90
20	6.8	3.55
21	7.2	5.74
22	4.6	3.51
23	26.2	25.19

APPENDIX 3

The actual values and predicted values of TSS of testing data using linear model fitting between L*, a*,b* and TSS for devices 1, 2 and 3 respectively

SL NO	ACTUAL TSS	PREDICTED TSS
1	4.6	5.59
2	6.2	6.03
3	6.5	6.23
4	29.0	22.21
5	27.6	24.06
6	5.2	5.90
7	18.8	10.53
8	31.0	22.49
9	28.8	26.52
10	31.0	25.60
11	8.0	14.22
12	7.2	12.39
13	6.2	6.52
14	6.0	5.95
15	16.6	15.82
16	7.4	8.80
17	6.8	5.59
18	7.0	5.68
19	6.8	4.14
20	4.6	5.48
21	17.6	10.71
22	7.2	3.16
23	6.0	7.91
24	19.6	22.25
25	6.6	11.14

SL NO	ACTUAL TSS	PREDICTED TSS
1	29.2	25.52
2	4.6	5.10
3	6.2	6.58
4	6.5	7.37
5	6.8	7.19
6	31.8	29.36
7	30.8	24.29
8	7.4	8.37
9	9.4	5.83
10	6.0	5.24
11	25.2	25.70
12	26.2	25.34
13	5.8	6.15
14	6.4	6.95
15	8.6	3.70
16	6.8	0.60
17	11.2	11.91
18	6.2	9.88
19	7.0	9.84
20	5.4	8.27
21	24.2	22.42
22	17.6	17.28
23	6.0	7.43
24	6.6	11.09

SL NO	ACTUAL TSS	PREDICTED TSS
1	6.0	7.93
2	25.8	23.19
3	4.2	5.88
4	6.5	9.12
5	6.8	6.78
6	29.0	23.79
7	27.6	26.95
8	6.0	5.29
9	26.2	27.39
10	7.2	6.17
11	6.0	9.74
12	31.0	26.08
13	6.4	5.59
14	5.8	5.80
15	6.4	5.63
16	6.6	6.72
17	7.6	6.27
18	7.4	6.50
19	7.2	7.66
20	6.8	3.38
21	7.2	5.68
22	4.6	3.69
23	26.2	25.51