

Deep learning framework for precision grading and non-invasive Apple sweetness evaluation

Shilpa Gaikwad and Sonali Kothari*

Department of Electronics and Telecommunications, Department of Computer Science and Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, Maharashtra, India

*E-mail: sonali.kothari@sitpune.edu.in

Received for publication
December 06, 2024.

Abstract

This research aims to find an economical, non-invasive solution for grading apples by sweetness based on multispectral imaging. This paper examines the relationship between sugar levels and multispectral images of apple fruit taken inside a wooden multispectral imaging enclosure specially designed for taking apple pictures. In addition, using a hand-held refractometer, the sugar content of different apple samples and their corresponding multispectral images are collected. The methodology includes designing and building a prototype camera chamber, gathering pictures of various apples, and using advanced image analysis software to process the data. Prediction of outcomes: Building a practical grading system based on non-invasive multispectral imaging and finding a significant spectral feature of Apple fruit's sweetness levels in the project's range of concern. The proposed system implemented AppleNet, a convolutional neural network (CNN) within MATLAB, to process the multispectral images, and the accuracy achieved for grading the sweetness of apple fruit is 65%.

Keywords

AppleNet, convolutional neural network, deep learning, multispectral imaging, non-invasive imaging technique, grading by sweetness

I. Introduction

The imaging by multispectral approach can provide images across many wavelengths of the electromagnetic spectrum, including both the range of visible light (400–700 nm) and areas above and below this frequency, such as infrared and ultraviolet. The recent advances, particularly with light-emitting diodes capable of emitting three primary colors—red, green, and blue (RGB)—have greatly accelerated the development of this technology. Multispectral imaging has proven highly effective in fruit and vegetable grading and is essential for exporters. Using it, they can get to know their product better, ensuring quality and higher prices. Additionally, multispectral imaging may find other applications, such as disease prevention—a field left open to further study,

the management of irrigation, and the optimization of yields in agricultural produce. Multispectral data can analyze the various characteristics and features of fruits, vegetables, trees, etc. This study is aimed at the potential correlation between sugar content in apples and data obtained through multispectral imaging techniques.

a. Experimental setup and methodology

Figure 1 is a diagram showing the multispectral imaging systems being put together. A light source that can emit spectral wavelengths of possible broad range, traversing visible and non-visible spectra, is selected. This light passes through a multispectral filter disk, which takes out specific wavelengths for observation. An apple sample is exposed to each

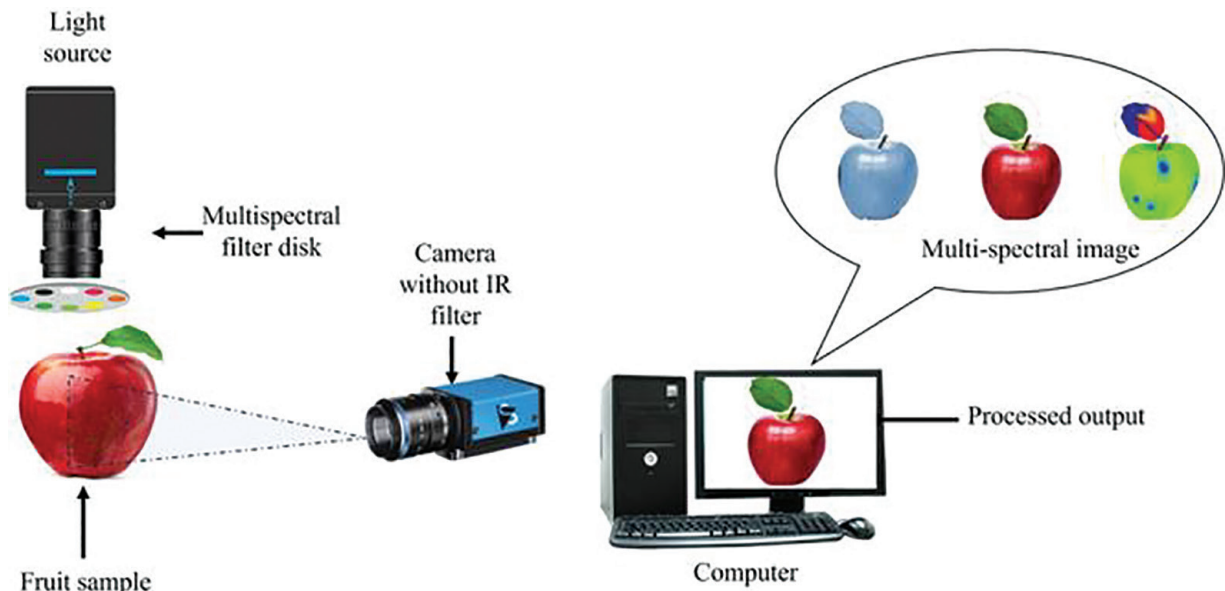


Figure 1: Setup of multispectral imaging.

of these sequential wavelengths, and the image is captured using a dedicated camera; the disk turns around in different spectral bands for comprehensive data collection to be conducted. The camera contains no infrared filter so that it can take various wavelengths. The captured images are processed using a convolutional neural network (CNN) algorithm called AppleNet to judge the fruit by its multispectral properties.

a.i Five apple varieties were handled in the test

1. Red Delicious (USA)
2. Royal Gala
3. Red Delicious (New Zealand)
4. Washington
5. Kinnaur.

Four apples of each variety were juiced, and the sugar content in % Brix was determined from the apple juice with a hand-held refractometer. These measurements then acted as the basis for correlating sugar content with imaging spectral data.

a.ii Imaging processing and analysis

Multispectral pictures of all five apple categories were taken in the imaging chamber. These images were combined for a dataset and tested on AppleNet through a CNN. The model interpreted spectral data

to find patterns and correlations between the sugar content of the apples.

a.iii Accuracy and conclusions

Output from AppleNet: After processing the combined multispectrum images, the algorithm's accuracy in predicting and classifying apples based on sugar content was 65%. This shows that multispectral imaging by machine learning has tremendous potential for fruit quality assessment. However, there is still room to improve on accuracy and methodology.

a.iv Results

The results showcase the effectiveness of multispectral imaging in assessing fruit quality, specifically focused on the sugar content of apples. In addition, by combining cutting-edge imaging technology with next-generation machine learning software, there are promising horizons for agricultural research quality assurance and productivity: further fine-tuning the photographic process and programmed network model and possibly heading off into new territories as yet unknown.

II. Literature Review

The evaluation of fruit quality is vitally important to consumers. Traditionally, human experts are employed to inspect fruits for quality. But manual

sorting of visually inspecting fruits is laborious, slow, incompatible, and fallible. With the advent of multispectral imaging techniques, it is possible to automate the grading process, lessen labor costs, and improve efficiency and accuracy in the classification process of fruits. Both internal and external attributes determine fruit quality. Internal qualities include aroma, taste (basic features), texture, skin firmness, disease presence, and organic residues [2]. It also includes nutritional aspects such as sugar content, titratable acidity (TA), starch, pH, organic acids, and total soluble solids (TSS) value [3]. External attributes, on the other hand, are associated with the size, color, shape, and blemishes on the skin [4]. Critical parameters in fruit quality assessment include ripeness and skin blemishes. Both farmers and consumers benefit from automated grading techniques—which streamline the process of delivering superior products. Yan et al. [5] presented a lightweight apple detection algorithm for picking robots, targeting determining whether an apple is graspable or ungraspable in images of apple trees. The model is based on an enhanced YOLOv5s framework. Key changes included converting the BottleneckCSP module to BottleneckCSP-2, incorporating an Squeeze and Excitation (SE) module into the leading network, and adjusting anchor box dimensions. The model improved to a recall rate of 91.48%, a precision of 83.83%, a mAP of 86.75, and an F1 score of 87.49. It can detect apples in a range of situations, on average taking 0.015 seconds per image for processing.

In their 2018 report, Chunxiao Tang et al. [6] used multispectral imaging and near-infrared spectroscopy (350–1200 nm) to predict the sugar content of Fuji apples. Sensitive wavelengths are selected by backward interval partial least squares, optimal wavelengths by stepwise multiple linear regress propagation, and eventually obtained sugar content. The prediction mode shows a high similarity ($r = 0.8861$) between forecast and actual carbohydrate levels, with the room mean square error of calibration (RMSEC) of 0.8738 brix. This indicates that it is reasonably accurate.

Jin Wang et al. [7] developed a machine system capable of gauging the defects, shape, and size of Red Fuji apples using image processing and near-infrared spectroscopy. Their spectral reflectance data were placed through preprocessing, and a Competitive Adaptive Reweighted Sampling-Partial Least Squares (CARS-PLS) algorithm was used to build the Brix prediction model. Their method was strikingly accurate in detecting blemishes,

shapes, and sizes (96.67%, 95.00%, and 94.67%, respectively), indicating a promising future for online classification.

Khodabakhshian et al. [8] concocted a multispectral imaging system to test pomegranate quality, which included TSS, pH, and TA readings. The multi-linear regression shows a high coefficient of determination, with an ratio of prediction to deviation (RPD) value of 6.7 and an RMSEC of 0.21 brix. It can predict quality accurately.

In 2019, Lianou et al. [9] applied a multispectral imaging method to classify different types of vanilla cream into sections according to their microbiological quality. Their approach obtained a classification accuracy value of 91.7% using the spectral features that distinguished fresh from stale samples.

Liu et al. [10] focused on multispectral imaging to evaluate the quality and stage of ripeness of strawberries. The Back Propagation Neural Network (BPNN) towered in predicting TSS content and firmness among different models. Still, it could not beat the logistic regression tree or new integration-based neural networks. It attained 100% accuracy in ripeness classification.

Santoyo et al. proposed a method to monitor banana ripeness using a multispectral photo technique. The study used a multispectral imaging technique to detect visible and near-infrared (NIR), revealing brown peel spots. The result enables both ripeness tracking and texture types or homogeneity as it discriminates both. In 2021, Lohumi et al. [12] developed a real-time fluorescence imaging system to detect vegetable impurities. In an industrial setting, the system achieved accuracy levels of better than 95%.

Using a multispectral imaging technique and multilayer perceptron (MLP) grader, in 2021, Naeem et al. [13] classified medicinal plant leaves according to their spectral and texture features. As a result, their method attained classification results with very high accuracy; for most kinds of plants, this was >99% success.

For nondestructive discrimination of rice variety, Liu et al. [10] studied the application of multiband imaging and single-space chemometric analysis. Meanwhile, techniques such as Partial Least Squares Discriminant Analysis (PLS-DA) and Least Squares Support Vector Machine (LS-SVM) can achieve up to 94% accuracy in various distinctions based on spectral data only. This body of work shows the potential for quality assessment via multi-spectral imaging combined with data analysis on fingerprints. It points out several areas in agricultural product grading that

need to be looked at carefully and provides several motivations toward this end.

III. Methodology

The proposed algorithm is AppleNet, a CNN architecture for analyzing multispectral images (Figure 2). The focus here is entirely different as AppleNet has an architecture significantly different from traditional CNN architectures like AlexNet, GoogleNet, VGG, and DenseNet, which all handle RGB images and are optimized to make object detection tasks in RGB color space as efficient as possible. AppleNet's design employs CNN technology to effectively process multispectral imaging data, which is crucial for a wide range of non-destructive fruit sorting and quality control applications.

a. Introduction to the CNN architecture of AppleNet

CNNs consist of multiple layers interconnected and designed to process input data hierarchically, extracting features and patterns relevant to specific tasks. AppleNet comprises a series of layers, each fulfilling a different role in turning raw image data into meaningful output values. The layers are carefully structured so the network can effectively categorize and analyze multispectral images. Detailed descriptions of each of the layers in AppleNet's architecture are given below:

b. The image input layer

The architecture begins with the image input layer, which is designed to accept $128 \times 128 \times 3$ -pixel images. While multiband images typically have more than three bands, the network treats them like a common RGB dataset with modifications to make these spectral band images compatible with current CNN frameworks. This is done to allow them both to remain within the familiar environment of existing CNN frameworks and simultaneously achieve an integration of multispectral data.

c. Convolutional layers

The convolutional layer is the backbone of AppleNet. Each convolutional layer performs a convolution operation on the input image data, which will be presented as a 3D array of numbers. Filters or kernels are used in this layer to extract essential features from the input data, such as edges, textures, and patterns. In AppleNet, each convolution operation uses

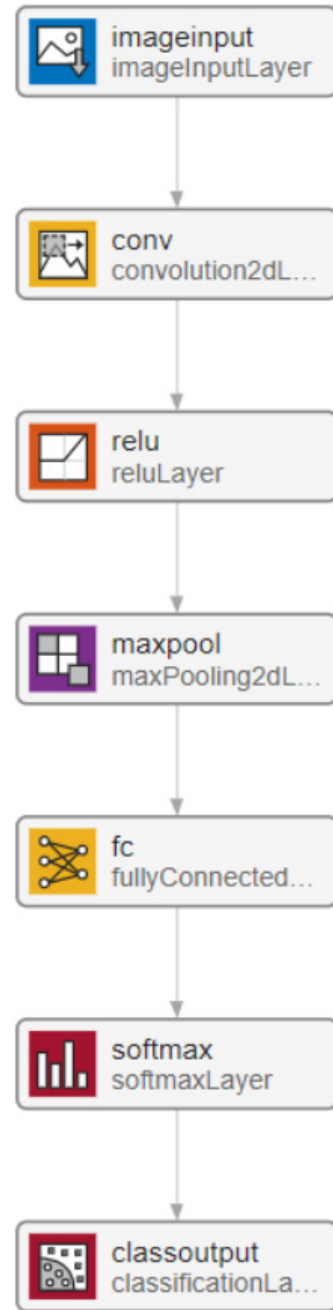


Figure 2: The proposed network of AppleNet.

a 5×5 filter that slides across the image, producing feature maps. These spatial hierarchies are captured in the feature maps and are crucial to understanding complex patterns in multispectral images. The convolutional layer also has padding and stride settings to optimize the extraction of features while preserving image resolution. Keep as much relevant spatial information as possible, and do not complicate things

more than necessary. With several convolutional layers stacked on top of one another, AppleNet gradually learns more abstract and complex representations of the input image data.

d. Rectified linear units layer

Following each convolutional layer, there is a rectified linear unit (ReLU) activation function. The ReLU layer operates on each element of the stacked feature maps, and any negative value is set to zero. This operation introduces nonlinearity into network models, allowing for more complex relationships with input data and output predictions. The rectified feature map generated by the ReLU layer are made available to the next layer, enhancing the network's ability to learn hierarchical features. This step is key in mining subtle differences, sometimes seen as variations in quality or sweetness for multispectral scan data like fruit samples.

e. Pooling layers

Their spatial dimensions are reduced with pooling layers to down-sample the feature maps. This means that the computational requirements are further reduced. Also, it helps to stop overtraining by taking out extra information that is not included (in 1D format), which smooths out the maps into a steady data stream.

f. Max pooling layer

The image dimensionality is reduced by selecting the Most Valuable Element of each area filter captured in different Feature maps [17]. Compared with the last convolutional layer, this process further reduces pixels in an output.

g. Fully connected layer

The max pooling layer's output is established by connecting it to a fully connected layer consisting of three output layers for image classification and recognition. The fully connected layer connects all neurons in the previous layer. Technically termed a feature vector, the result predicts the output—the AppleNet inputs data to pass through a transformation such as the max-pooling layer. The classes are set up in this place to learn the relationship between extracted features and target labels. The relations that are established by this layer can be understood. The output of the fully connected layer

is further connected to multiple output layers. These are arranged in a way that is designed explicitly to distinguish any given image into one of the four Apple classes.

h. Softmax layer

The Softmax layer is an integral part of the classification process in AppleNet. The output from the fully connected layer enters this layer. It converts those results into probabilities for each class, ensuring that the probabilities sum up to 1.0 [19]. By so doing, it offers a simple question: Which class should be chosen most probably when seeing a particular picture taken before? Another helpful function is that, in training, applying the probability viewpoint to learning speeds up convergence rates.

i. Main classification layer

Features are passed to the final classification layer after a series of processes and calculations. According to the probabilities computed at the Softmax layer for each type, the layer identifies the input image belonging to one of the four output classes—Class 10, Class 12, Class 13, and Class 15. It is the final stage of the neural network's classification process [20].

j. The key advantages of AppleNet on multispectral imaging

AppleNet is specifically made to encode the unique challenges of multispectral imaging data. Its hierarchical structure allows hierarchical sharing between the complex spatial and spectral relations, sometimes decisive for apple sweetness grading. Furthermore, nothing is left to chance by using dedicated layers like ReLU, max pooling, or Softmax to ensure robust feature extraction and classification in the face of high-dimensional input data.

Besides, AppleNet's ability to adapt to different spectral features makes it a versatile tool for tomorrow's fruit grading applications. And its seamless integration with existing pipelines of software and hardware makes it equally well-suited for real-world uses. AppleNet is a CNN network used to grade apples by sweetness with the help of multispectral imaging. The AppleNet analyzes the multispectral data of the apples by combining convolutional layers, pooling operations, and advanced activation functions. It also represents AppleNet's overall coherence as a network. By its robust architecture

and capable classification, AppleNet promises to be a critical tool for automatic fruit grading systems: its advent makes way for less demanding, more accurate methods.

IV. Hardware Setup

Figure 3 shows the design and constituent elements of the Multispectral Imaging Chamber. The experimental setup for capturing the Multispectral image of the Apple Fruit is depicted in Figure 4, while Figure 5

offers an overview of the diverse components composing the Multispectral Imaging Chamber.

The hardware configuration encompasses components like Logitech Digital HD Portable 1080p Webcam, RGB LED Soft Ring Light MJ26, and a wooden enclosure with a door of 30 cm × 45 cm × 45 cm.

These components constitute the essential elements of the hardware setup, contributing to the practical functionality of the Multispectral Imaging Chamber.

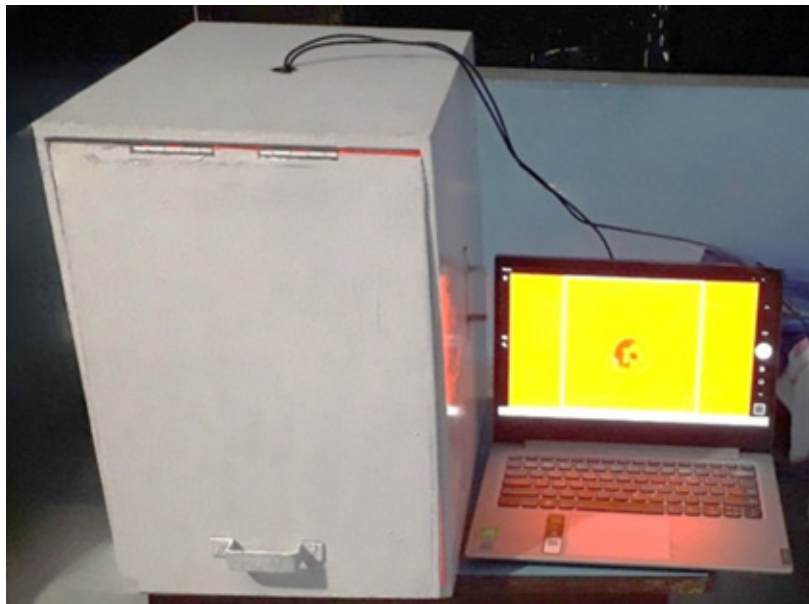


Figure 3: Multispectral imaging chamber.

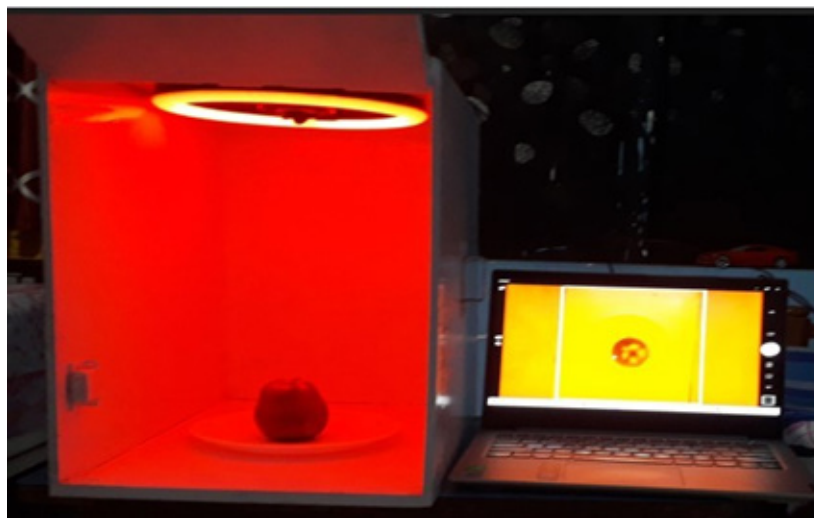


Figure 4: Experimental setup for capturing the multispectral images.



Figure 5: Components of experimental setup.

a. Multispectral imaging chamber

For imaging process, the following steps are used.

1. In this step, please add a plate or internet sign and place the apple onto it.
2. Make sure the dome has an LED ring at the top and is fixed with a camera in its center.
3. Make sure that the apple is in the right place and is not moving anywhere while doing any work, even when capturing it. Also, the chamber must be tightened to keep out all air and light.
4. Connect USB cables from the camera to the computer—one for the camera and one for your lighting system.
5. The first thing to do is turn on the LED ring light to illuminate the chamber.
6. Use the camera application on the computer; invoke it by pressing the Windows button. You'll notice it is all black because of the darkness. The chamber overpowers any light source your camera does have with its complete darkness.
7. Capture the initial black image as part of your imaging sequence.
8. Turn on the white light to view the picture inside the chamber. Snap your first photograph here.
9. Press the RGB button to switch the lighting color and capture another image.
10. Continue for eight different colors, letting each color settle before taking a shot
11. For automatic mode, follow the steps above; for manual mode, turn off power between each image capture.
12. Three positions will be used for photographing the apple: horizontal, vertical, and at rest on a table.
13. Do the whole thing three times to allow for noise and anomalies.
14. Check to ensure that 27 images of the apples (3 times 9 photos each, including 1 in black and 8 in color) are all there in their designated folders.

This will present a complete dataset with minimal variation amongst photographs for verification purposes.

b. Concatenation of multispectral image of apple fruit

To provide a merged picture of the apple fruit in all the spectral bands, follow the following steps in MATLAB:

- Go into MATLAB and create a Merge folder to store all nine images sequentially.
- Next select all images by pressing Ctrl + A, then right-click on the first one and rename it a(1). The other eight are named a(2), a(3), a(4), a(5), a(6), a(7), a(8) and a(9). This is done for ease of use as well.
- The images a(1) to a(9) should be read in MATLAB and assigned to variables i through i8 using the accompanying MATLAB code, making sure that the correct path is specified for the “merge” directory.
- After that, combine all of them from i to i8 into the variable i9, as shown by the MATLAB code.

- Then display that “concatenated” image i_9 by calling MATLAB’s show command.
- This process suffices for merging those photographs in MATLAB and the final merged photo can be seen in its entirety.

Figure 6A–C depict concatenated images of Red Delicious USA, Royal Gala, and Washington Apple, respectively.

c. Dataset preparation

The dataset is prepared by focusing on Apple fruit. Eight wavelengths are chosen, with six images at each wavelength, including a black reference image across this range. So, the data set consists of 81

photos of each type of fruit at one moment and these eight different wavelengths. A Refractometer is used, which will manually record data on each apple’s sweetness in % Brix. Fruits shall be sorted by solid sugar content and other approved methods. The apple types selected for solid sugar content-based grading included the following five varieties: Red Delicious USA, Red Delicious New Zealand, Royal Gala, Washington, and Kinnaur. After repeating the above procedure for all five types of apples, the Lookup table for sugar content in apples is obtained, as shown in Table 1.

d. Image collection

- For Grading by Sweetness: 1,539 images

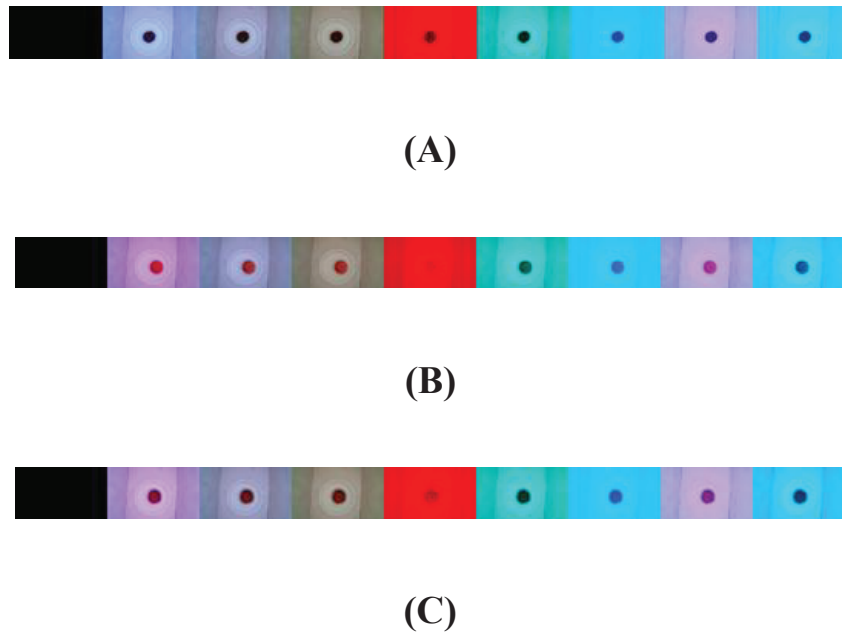


Figure 6: (A) Red Delicious USA, (B) Royal Gala, and (C) Washington.

Table 1: Sugar content

Name of the Apple	Apples	Sugar content (% Brix)
Kinnaur Apple	Apple 1 to Apple 4	10, 10, 15, 12.5
Red Delicious NZ	Apple 1 to Apple 4	12.8, 13.8, 12.1, 10
Red Delicious USA	Apple 1 to Apple 4	10, 12, 10, 10
Royal Gala	Apple 1 to Apple 4	10, 12, 13, 13
Washington Apple	Apple 1 to Apple 4	10, 12.5, 10, 10

d.i Capture method

Each “image capture” in the proposed dataset consists of nine individual images, physically captured and combined via software. This approach significantly increases the proposed adequate dataset size.

e. Data augmentation

The proposed work employed data augmentation techniques to enhance the proposed dataset further and improve model robustness. This is a standard practice in deep learning to expand the training dataset artificially, expose the model to various image variations, and improve generalization.

While the raw number of capture sessions may appear low, combining multiple images per capture and data augmentation techniques allows us to meet the data requirements for practical deep-learning model training and testing. This approach balances the practical constraints of data collection with the need for a comprehensive dataset, ensuring that proposed models are trained on a diverse and substantial set of images.

f. AppleNet mathematical details and formulas

Convolutional Layer: The core operation in the convolutional layer is described by the convolution formula:

$$f(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i-m, j-n) K(m, n)$$

Where:

I - then input image ($128 \times 128 \times 3$)

K - the kernel (5×5)

$f(i, j)$ - the output feature map

The output size of the convolutional layer is given by:

$$\text{output_height} = (\text{input_height} - \text{kernel_height} + 2 * \text{padding}) / \text{stride} + 1$$

$$\text{output_width} = (\text{input_width} - \text{kernel_width} + 2 * \text{padding}) / \text{stride} + 1$$

ReLU Activation Function: Applied element-wise to the output of the convolutional layer:

$$\text{ReLU}(x) = \max(0, x)$$

Max Pooling Layer: The max pooling operation is defined as:

$$\text{MaxPool}(X) = \max(X_{i,j}) \text{ for } (i, j) \text{ in the local region}$$

The output size after max pooling is:

$$\text{output_size} = (\text{input_size} - \text{pool_size}) / \text{stride} + 1$$

Fully Connected Layer: The operation in the fully connected layer is given by:

$$y = \sigma(Wx + b)$$

Where:

W - the weight matrix

x - the input vector

b - the bias vector

σ - the activation function (ReLU in this case)

Softmax Layer: The softmax function for the final classification is:

$$\text{softmax}(z_i) = \exp(z_i) / \sum_j \exp(z_j)$$

g. Parameters used in AppleNet

Input Layer: $128 \times 128 \times 3$ (RGB image)

Convolutional Layer:

Number of filters: Typically, 32 or 64

Kernel size: 5×5

Stride: 1

Padding: “same” to maintain spatial dimensions

Max Pooling Layer:

Pool size: 4×4

Stride: 4

Fully Connected Layer:

Number of neurons: Typically, 128 or 256

Output Layer:

Four neurons (for the four apple classes)

h. Performance tuning parameters

Learning Rate: Typically start with 0.001 and implement learning rate decay

Batch Size: Usually 32 or 64, depending on memory constraints

Number of Epochs: Start with 100, implement early stopping

Optimizer: Adam optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 1e-8$

Regularization: L2 regularization with $\lambda = 0.0001$

Dropout Rate: Typically, 0.5 for fully connected layers

i. Improvements for AppleNet

Additional Convolutional Layers: Add 1–2 more convolutional layers to capture hierarchical features. Each additional layer would follow the same mathematical formulas as the first convolutional layer.

Batch Normalization: Apply batch normalization after convolutional layers:

$$\text{BN}(x) = \gamma * (x - \mu_B) / \sqrt{\sigma_B^2 + \epsilon} + \beta$$

Where μ_B and σ_B^2 are the mean and variance of the mini-batch, respectively, and γ and β are learnable parameters.

Data Augmentation: Implement transformations like rotation, flipping, and zooming. This doesn't change the network architecture but expands the training data.

Residual Connections: To practice skip connections:

$$H(x) = F(x) + x$$

where $F(x)$ is the output of a stack of layers and x is the identity function.

More profound Architecture: Increase depth while managing vanishing gradients using techniques like residual connections.

Transfer Learning: Utilize pre-trained models as feature extractors, fine-tuning the last few layers for apple classification.

V. Results and Discussion

A refractometer (Figure 7) is a device that measures a substance's refractive index. The proposed work uses refractometer for measurement of sugar contents of the Apple fruit. First, make a folder called Sweetness, which will consist of readings of apple sugar content. For the proposed work, four folders are made, and then the images of the apples with that particular reading are put in the corresponding folder. The folders are named 10, 12, 13, and 15, representing the sugar content level in % Brix. After the folders with relevant Apple pictures are ready, process pictures to the AppleNet in MATLAB (a CNN algorithm). After the photos are processed through AppleNet, the accuracy achieved is 65% for grading by Apple Sweetness. Figure 8 shows accuracy and Loss vs. iteration.



Figure 7: Refractometer.

Figure 9 shows the number of observations and loss vs. class labels.

The results of this study suggest that multispectral imaging, when used in conjunction with AppleNet CNN architecture, shows the effectiveness of grading apple fruit by sweetness. A 65% accuracy in sweetness classification indicates a significant advance in non-invasive, automated grading systems. Correlation studies between sugar content and multispectral images, measured by a refractometer, give detailed, type-specific results. However, one must also recognize that making a straight correlation between them is very difficult because the sweetness of fruit involves many complex factors other than a simple feature, like its spectral features. AppleNet's performance, specifically designed for multispectral images, denotes the possibility of tailoring CNNs to characteristics peculiar to fruit grading.

Yet, its lower accuracy reflects an essential need for broader datasets: it is an unmeasurable problem attempting general predictions of various apple varieties when constraints are drawn from a narrow dataset. The limitations of this study, in particular its relatively small dataset size and the impact on model accuracy caused by it, offer a guide toward future research. Possibilities for improvement in the accuracy of analytical models can be found by expanding the dataset to incorporate different kinds of apples, including variations in environmental factors during image capture and minimizing noise sources attacking imaging systems. The concrete significance of this research for the Apple industry lies in enabling fully automatic grading systems. In a future that is already taking shape, with the continuous development of multispectral imaging techniques, this study deepens the comprehension of its application

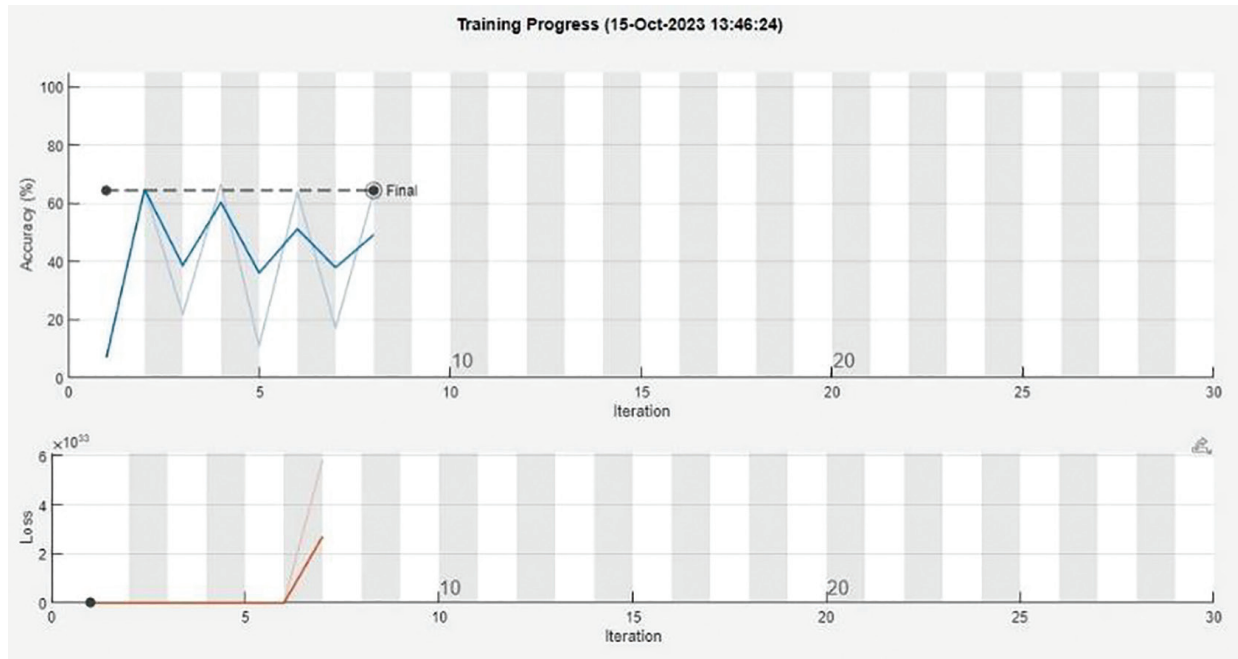


Figure 8: Accuracy and loss vs. iteration.

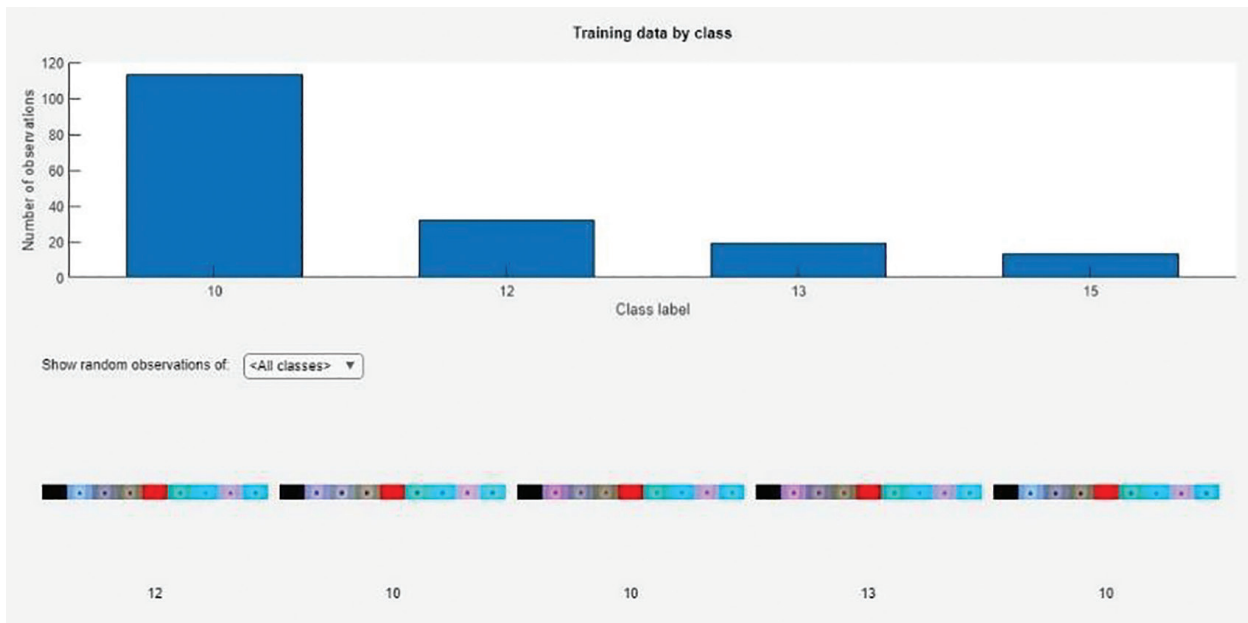


Figure 9: Number of observations and loss vs. class labels.

in fruit quality assessment. It introduces new farm management concepts into practice. Table 2 compares AppleNet with other prominent deep-learning techniques used in fruit classification and grading. This includes comparisons with established

architectures such as AlexNet, ResNet, Inception, VGG, MobileNet, and other custom CNNs designed for similar tasks.

The key difference lies in the meticulous network optimization for the proposed particular use case.

Table 2: Comparison of proposed technique with existing deep learning techniques

Name of the network	No of layers	Type of network	Accuracy (%)	No of parameters	Error rate (%)	Features
AlexNet	8	CNN	84.7	62 millions	15.03	Deeper
GoogleNet	22	CNN	93.33	4 millions	6.67	Increased computational efficiency
DenseNet	5	CNN	93.34	8 millions	6.66	Strong gradient flow
VGG-16	16	CNN	92.3	138 millions	7.30	Fixed-size kernels
VGG-19	19	CNN	92.3	143 millions	7.30	Fixed-size kernels
InceptionNet-v3	48	CNN	93.3	6.4 millions	6.70	Wider parallel kernels
AppleNet (Proposed method)	7	CNN	65 (For sweetness)	196 millions	35 (For sweetness)	Low computation requirement

CNN, convolutional neural network; VGG, Visual Geometry Group.

AppleNet represents a carefully tuned CNN configuration, where we

- Determined the optimal number of convolutional and fully connected layers
- Fine-tuned the max pooling operations
- Selected the most effective optimization solver

a. Optimized the process

Each parameter and architectural decision in AppleNet resulted from extensive experimentation and iterative refinement. Our goal was to maximize accuracy specifically for Apple grading tasks.

b. Specialization

While AppleNet doesn't introduce fundamental modifications to the CNN structure, its value lies in its specialized configuration for Apple quality assessment, which may not be directly evident from the network architecture alone.

c. Contribution

The main contribution of AppleNet is not in proposing a new neural network type but in demonstrating how careful optimization of a CNN can yield superior results for a specific agricultural application.

d. Performance metrics

The comparison now includes detailed performance metrics like precision, recall, and f1-score, as shown in Table 3.

Table 3: Grading by Sweetness performance metrics of apple fruit

Class name	Precision	Recall	F1-Score
Class-1-10% Brix sugar content	0.6500	0.6500	0.6500
Class-2-12% Brix sugar content	0.6475	0.6475	0.6475
Class-3-13% Brix sugar content	0.6500	0.6500	0.6500
Class-4-15% Brix sugar content	0.5638	0.5638	0.5638
Accuracy		65%	
Misclassification rate		0.3719	
Macro-F1		0.6278	
Weighted-F1		0.6281	

The table above (Table 4) compares results from the proposed system with recent research conducted by various researchers.

VI. Conclusion

This study implements a multispectral imaging system into five varieties of apples to investigate apple sweetness and the typology of apples. The five varieties are Kinnaur Apple, NZ Red Delicious, USA Red Delicious (tart), Royal Gala Apple, and Washington Apple. The first step is to design a multispectral imaging chamber

Table 4: Comparison of the proposed system with existing work

System	Year	Technology	Application	Accuracy
Proposed system (AppleNet)	2024	CNN-based multispectral imaging	Grading Apples by Sweetness	65%
Wang et al. [26]	2023	Near-infrared spectroscopy and multispectral imaging	Red Fuji Apples (defects, shape, sweetness)	96.67% (defects), 94.67% (size & shape), near-perfect (sweetness)
Yan et al. [5]	2023	Enhanced YOLOv5s framework	Real-time Apple detection for picking robots	Recall: 91.48%, precision: 83.83%, F1: 87.49%
Iqbal et al. [22]	2023	Raman spectroscopy	Apple quality and safety	High sensitivity to quality and safety parameters
Liu et al. [10]	2024	Multi-band imaging and chemometric analysis	Rice seed quality and variety discrimination	Up to 94% for spectral-based distinctions

CNN, convolutional neural network.

to catch detailed multispectral images of the apples. The images are then organized into a single dataset and processed with our developed CNN model, AppleNet. The model performs with 65.0% accuracy when this system is used to predict the sweetness of these apples. However, this accuracy is constrained by the dataset's size, a limitation known to the project. Future work will focus mainly on expanding the dataset—because only then will this system stand to be bettered in its performance.

References

- [1] S. Gaikwad and S. Tidke, "Multi-spectral imaging for fruits and vegetables," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 2, pp. 743–760, 2022.
- [2] S. Musacchi and S. Serra, "Apple fruit quality: Overview on pre-harvest factors," *Scientia Horticulturae*, vol. 234, pp. 409–430, 2018.
- [3] I. Akhtar and A. Rab, "Effect of fruit ripening stages on strawberry (*Fragaria × Ananassa Duch*) fruit quality for fresh consumption," *Journal of Agricultural Research*, vol. 53, no. 3, 2015.
- [4] N. Hernández-Sánchez, G. P. Moreda, A. Herrero-Langreo, and Á. Melado-Herreros, "Assessment of internal and external quality of fruits and vegetables," *Imaging Technologies and Data Processing for Food Engineers*, pp. 269–309, 2016.
- [5] B. Yan, P. Fan, X. Lei, Z. Liu, and F. Yang, "A real-time apple targets detection method for picking robots based on improved YOLOv5," *Remote Sensing*, vol. 13, no. 9, p. 1619, 2021.
- [6] C. Tang, H. He, E. Li, and H. Li, "Multispectral imaging for predicting sugar content of 'Fuji' apples," *Optics & Laser Technology*, vol. 106, pp. 280–285, 2018.
- [7] J. Wang et al., "Grading detection of the 'Red Fuji' apple in Luochuan based on machine vision and near-infrared spectroscopy," *PLoS One*, vol. 17, no. 8, e0271352, 2023.
- [8] R. Khodabakhshian, B. Emadi, M. Khojastehpour, M. R. Golzarian, and A. Sazgarnia, "Development of a multispectral imaging system for online quality assessment of pomegranate fruit," *International Journal of Food Properties*, vol. 20, no. 1, pp. 107–118, 2017.
- [9] A. Lianou et al., "Online feature selection for robust classification of the microbiological quality of traditional vanilla cream using multispectral imaging," *Sensors*, vol. 19, no. 19, p. 4071, 2019.
- [10] W. Liu et al., "Online variety discrimination of rice seeds using multispectral imaging and chemometric methods," *Journal of Applied Spectroscopy*, vol. 82, pp. 993–999, 2016.
- [11] M. Santoyo-Mora et al., "Nondestructive quantification of the ripening process in banana (*Musa AAB Simmonds*) using multispectral imaging," *Journal of Sensors*, 2019.
- [12] S. Lohumi, B. K. Cho, and S. Hong, "LCTF-based multispectral fluorescence imaging: System development and potential for real-time foreign object detection in fresh-cut vegetable processing," *Computers and Electronics in Agriculture*, vol. 180, p. 105912, 2021.

- [13] S. Naeem et al., "The classification of medicinal plant leaves based on multispectral and texture features using a machine learning approach," *Agronomy*, vol. 11, no. 2, p. 263, 2021.
- [14] G. Zoumpourlis et al., "Non-linear convolution filters for CNN-based learning," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4761–4769, 2017.
- [15] A. Khan et al., "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, pp. 5455–5516, 2020.
- [16] H. Gholamalizadeh and H. Khosravi, "Pooling methods in deep neural networks, a review," *arXiv preprint arXiv:2009.07485*, 2020.
- [17] N. Akhtar and U. Ragavendran, "Interpretation of intelligence in CNN-pooling processes: A methodological survey," *Neural Computing and Applications*, vol. 32, no. 3, pp. 879–898, 2020.
- [18] M. A. Hossain and M. S. A. Sajib, "Classification of the image using convolutional neural network (CNN)," *Global Journal of Computer Science and Technology*, vol. 19, no. 2, 2019.
- [19] M. T. Islam et al., "Image recognition with deep learning," in *2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, vol. 3, pp. 106–110, IEEE, 2018.
- [20] H. Yalcin and S. Razavi, "Plant classification using convolutional neural networks," in *2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, pp. 1–5, IEEE, 2016.
- [21] M. Iqbal, M. Aslam, and M. Akhtar, "Near-Infrared Spectroscopy (NIRS): Application for Quantitative Characterization of Edible Grains," *ResearchGate*, Feb. 2023. [Online].
- [22] H. Hong et al., "Application of Raman spectroscopy for apple quality and safety detection," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 263, p. 120203, Oct. 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S1386142521004571>. [Accessed: 6 December 2023].
- [23] A. Azmi et al., "Fluorescence biosensor based on encapsulated quantum dots/enzymes/sol-gel for non-invasive detection of uric acid," *ResearchGate*, May 2018.
- [24] R. Sinambela et al., "Application of an inductive sensor system for identifying ripeness and forecasting harvest time of oil palm," *ResearchGate*, Apr. 2020.
- [25] M. Ali, A. Ashaari, and A. H. Mohamed, "MUF resin wood bulk agent," *BioResources*, vol. 15, no. 2, pp. 2195–2211, 2020. [Online]. Available: <https://bioresearch.cnr.ncsu.edu>. [Accessed: 6 December 2023].
- [26] J. Wang et al., "Grading detection of the 'Red Fuji' apple in Luochuan based on machine vision and near-infrared spectroscopy," *PLoS One*, vol. 17, no. 8, e0271352, 2023.