International Journal of Scientific Research in Science, Engineering and Technology

Print ISSN - 2395-1990
Online ISSN : 2394-4099

Available Online at : www.ijsrset.com
doi : https://doi.org/10.32628/IJSRSET23103177

Automatic Fruit Quality Inspection System Using Image Processing

Mr. Kannan A[1], Mohammed satham A[2], Udhayanithi D[2], Venkataraman A[2]

[1] Assistant Professor, Department of Electronics and Communication, Rajiv Gandhi College of Engineering and Technology, Kirumampakkam, Puducherry, India

[2] UG - Electronics and Communication, Rajiv Gandhi College of Engineering and Technology, Kirumampakkam, Puducherry, India

ARTICLEINFO

Article History :

Accepted: 10 June 2023

Published: 04 July 2023

Publication Issue :

Volume 10, Issue 4

July-August-2023

Page Number :

01-08

ABSTRACT

The demand for high-quality fruits has increased significantly, driven by consumers' growing emphasis on health and nutrition. To ensure consistent quality control and efficient fruit grading processes, an automatic fruit quality inspection system using image processing techniques has been developed. This system leverages Convolutional Neural Network (CNN) algorithms to achieve an impressive accuracy rate of 99% in fruit quality assessment. The proposed system involves a multi-step process starting with the acquisition of high-resolution fruit images. These images undergo pre-processing to enhance clarity and eliminate noise or artifacts. Subsequently, the pre-processed images are fed into a trained CNN model for feature extraction and classification based on learned patterns. The CNN model has been trained on a large dataset of labeled fruit images, enabling it to recognize quality attributes such as colour, size, shape, and defects.

The system's evaluation involved a diverse range of fruits, encompassing various species, varieties, and maturity stages. The output demonstrate the system's exceptional accuracy, with a 99% success rate in correctly identifying and grading fruit quality attributes. This automatic fruit quality inspection system offers several advantages, including real-time processing, high efficiency, and reduced labor costs, making it suitable for integration into fruit processing facilities and supply chains. [1]

Keywords : CNN, Hyperspectral data, PLSR, SVR

I. INTRODUCTION

Sorting fresh fruits is crucial to ensure high quality, and the current visual inspection process based on color, size, and shape parameters has inherent limitations. It is subjective, variable, tedious, laborious, inconsistent, and easily influenced by environmental factors. Traditionally, fruits have been sorted through manual visual inspection, where trained individuals assess the fruits based on their appearance. This method, though widely used, has its drawbacks. It is time-consuming, prone to human error, and can be influenced by subjective factors such as fatigue and individual judgment.

To address these challenges, there is a growing interest in utilizing innovative and non-contact measurements provided by artificial vision systems, which can measure the entire surface of a fruit sample. These systems offer improved representativeness compared to colorimeters, which rely on point-to-point measurements. Computer vision systems (CVSs) have emerged as effective tools for automating the sorting process and detecting defects in horticultural products.

In the realm of fruit sorting, Convolutional Neural Networks (CNNs) have demonstrated remarkable performance. Unlike traditional sorting methods, such as manual inspection, CNNs can analyze fruit images with high precision and efficiency. They can automatically learn and recognize intricate patterns and features from the images, enabling accurate classification based on color, texture, and shape attributes. This automated approach reduces reliance on human judgment and minimizes subjective biases.

Compared to other traditional sorting algorithms like K-Nearest Neighbors (KNN) and X-ray imaging, CNNs have distinct advantages. They can handle large and high-dimensional datasets, making them suitable for analyzing diverse fruit images. CNNs excel at extracting meaningful features from raw image data, eliminating the need for manual feature engineering. Moreover, CNNs can capture hierarchical representations, allowing them to extract both low-level and high-level features, resulting in a comprehensive analysis of fruit attributes.

In comparison to manual inspection, where human error and fatigue can affect the consistency and accuracy of the sorting process, CNNs provide a reliable and efficient solution. They can be trained on a large dataset of labeled fruit images, enabling them to generalize well and accurately classify unseen fruit samples. This automated approach significantly reduces the subjectivity, variability, and labor-intensive nature of traditional sorting methods

II. RELATED WORK

The study by Shirin Nasr-esfahani et al. introduces a pit detection system for olives using hyperspectral imaging and a Convolutional Neural Network (CNN) classifier. The system aims to address the presence of pits in olives, which can impact product quality and pose risks to consumers.

Hyperspectral data capturing specific wavelengths for detecting pits in olives was collected. A 1D CNN classifier was developed and achieved high accuracies of 99.5% and 97.69% for pitted and whole olives in the training and test sets, respectively. The addition of a dropout layer further improved the accuracy to 98.27% on the test dataset. A comparison with other classifiers revealed the superior performance of the CNN classifier.

The results demonstrate the effectiveness of the proposed hyperspectral-based pit detection system using CNN. The high classification accuracies highlight its potential to enhance product quality and safety in various applications. The CNN classifier outperformed traditional machine learning methods, showcasing its suitability for accurate and efficient pit detection in olives[1].

The work by Xinting Yang et al. introduces a CNN model that combines spatial and channel attention mechanisms for recognizing pests in field-based images. The model incorporates Spatial Transformer Networks (STN) and Improved Split-Attention Networks, achieving high classification accuracies on three datasets. Compared to traditional CNN models and state-of-the-art deep learning methods, this approach outperforms them. Additionally, the researchers constructed six datasets with varying image resolutions, all achieving accuracies above 92%.

Agricultural insect pests can cause significant crop damage, impacting both quantity and quality. Accurate and timely pest identification is crucial for effective monitoring and early warning systems.

Traditional machine learning methods involve multiple sequential stages, while deep learning models, like CNNs, learn features and relationships directly from data. The proposed cascaded architecture, combining STN and ResNet networks, aims to enhance the accuracy and efficiency of pest recognition in field images.[2]

This research, conducted by Wilson Castro, Jimy Oblitas, Miguel De-la-torre, Carlos Cotrina, Karen Bazán, and Himer Avila-George, focuses on the classification of Cape gooseberry fruit based on its ripeness using machine learning techniques and various color spaces. The study involved analyzing 925 fruit samples, which were manually categorized into seven ripeness classes. The results highlighted the significance of both the chosen color space and the classification technique in accurately classifying the fruit based on ripeness. Notably, the combination of the Lab* color space and the SVM classifier demonstrated the highest f-measure among the models. Additionally, the integration of different color spaces through principal component analysis improved the performance of the models, albeit with increased complexity. These findings contribute to the quality assessment of fruit and provide valuable insights for selecting appropriate color spaces in order to achieve precise ripeness classification.[3]

The authors of this paper, Alberto Lucas Pascual, Antonio Madueño Luna, Manuel De Jódar Lázaro, and José Miguel Molina Martínez, focus on analyzing the functionality of olive pitting, slicing, and stuffing machines (DRR) using an IoT system, computer vision, and neural network diagnosis. Their objective is to optimize the performance of these machines by remotely analyzing data obtained through an IoT-based system and a physical chip with neural networks for classification. By integrating hardware-based classification, the machines can operate at their optimal speed, surpassing the limitations of software-based methods. The authors also reference relevant studies that showcase the effectiveness of neural networks in real-time detection, computer vision applications in industries, and agricultural product classification. Through the analysis of the feed chain in olive processing machines, this research contributes to the advancement of efficient and automated techniques in the industry. [4].

The study conducted by Baohua Yang, Yue Zhu, Mengxuan Wang, and Jingming Ning proposes a method to improve the accuracy of estimating yellow tea polyphenols content through multi-feature fusion. The method combines multi-scale wavelet decomposition and feature fusion, incorporating wavelet coefficient features, GLCM texture features, and wavelet texture features. Experimental evaluations were performed on five different types of yellow tea, comparing models that used various features, including partial least squares regression (PLSR) and support vector regression (SVR). The results showed that the model based on multi-feature fusion outperformed individual indicators, with SVR-based models demonstrating superior performance compared to PLSR. The main objective of the study is to enhance the accuracy of tea polyphenols content estimation by extracting multi-scale features from hyperspectral images while preserving the original information within the images.[5]

The authors of the mentioned content are Moallem, P., Serajoddin, A., and Pourghassem, H. The content discusses the importance of visual inspection in food quality assurance programs, particularly in fresh produce sorting. It highlights the limitations of manual inspection, such as subjective judgments, varying human acuity, reduced sorting speeds, and increased production costs. To overcome these challenges, the content suggests the use of computer vision systems, which offer flexibility and can be reasonable substitutes for human visual decision-making. These systems can measure physical dimensions, flag missing or deformed parts, detect discolorations and surface blemishes, and even examine internal defects using transmitted light, X-rays, or ultrasonic images. Digital image analysis and pattern recognition techniques are mentioned as approaches for quality detection in fresh produce[6].

III. EXISTING WORK

The need for high-quality fruits that meet consumer and market standards has led to the demand for an accurate and reliable grading system during the postharvest process. An existing system has been developed to address this need, achieving a high average accuracy of 0.9515 for detecting defects such as calyx and stalk scars on both defective and healthy fruits. This detection is performed using histogram thresholding based on the mean g-r value of these regions of interest. Defective regions are further identified using an RBF-SVM classifier based on LAB color-space pixel values, resulting in an overall accuracy of 0.989 upon validation.

To enable fruit grading, four recognition models were developed based on color and texture features. The algorithms explored include RBF-SVM, linear SVM, quadratic SVM, cubic SVM, random forest, and ANN. The RBF-SVM algorithm outperformed all other models, achieving the highest accuracy of 0.9709 for differentiating healthy and defective fruits. However, as the number of grading categories increased, the accuracy of the grading decreased. The combination of color and texture features yielded the highest accuracy across all grading categories during the evaluation of image features. The developed system, with its accurate grading capabilities, can be effectively used as an inline fruit sorting tool to ensure adherence to and maintenance of quality standards. The system follows a specific architecture where the image is preprocessed, features are extracted, and grading is performed using SVM, random forest, and ANN algorithms. Background removal is achieved through a simple image subtraction technique, and defects such as calyx and stalk scars are detected based on gray values and longitudinal directions of the fruit image. Various features such as color, texture, and shape are extracted to enable accurate grading. The system has been validated and proven to provide reliable results in fruit grading.

In the existing system, one notable observation was that the accuracy of the different models decreased as the number of grading categories increased. This suggests that the complexity of distinguishing and categorizing fruits accurately becomes more challenging when there are multiple grading categories involved.

Furthermore, it is important to note that the existing system is specifically designed to detect defects in Cherry and Heirloom tomatoes. While it demonstrates effectiveness for these particular types of tomatoes, it may not be applicable or as accurate for other fruit varieties. Therefore, the system's scope is limited to the specific types of fruits mentioned.

One of the drawbacks of manual defect identification in fruits is the inherent difficulty involved. Human experts may encounter challenges in precisely identifying and categorizing defects, which can be subjective and prone to errors. This emphasizes the need for automated systems, such as the proposed grading system, to provide objective and consistent defect detection and grading capabilities.

IV. PROPOSED WORK

In the field of image processing for fruit sorting, this study aims to advance the classification of fruit quality using machine learning techniques. The proposed system utilizes transfer learning in conjunction with the EfficientNet model, which is known for its efficiency and effectiveness in analyzing images. By leveraging the model's ability to extract intricate features from high-dimensional fruit images, the objective is to create an approach that excels in both efficiency and accuracy when categorizing various levels of fruit quality.

To commence, a comprehensive dataset is collected through collaborations with fruit farms and suppliers. This dataset encompasses a diverse range of fruit varieties and quality categories, including fruits of excellent quality as well as those displaying defects such as bruises, rot, or damage. This dataset forms the foundation for training the model and ensures its ability to accurately classify a wide range of fruit quality conditions.

The collected fruit images undergo meticulous preprocessing, including the standardization of image resolutions and the normalization of pixel values. Additionally, data augmentation techniques such as cropping, scaling, flipping, shifting, and rotation are applied to enhance the dataset's diversity. These techniques contribute to improving the model's capacity to learn from various fruit variations and enhance its ability to generalize.

The proposed system capitalizes on the strength of transfer learning. Specifically, the upper layers of the EfficientNet model are fine-tuned to adapt to the unique requirements of fruit quality classification. By combining the model's inherent visual pattern recognition capabilities with its customization to the curated dataset, the system becomes adept at accurately discerning the complex characteristics associated with different fruit quality categories.

Training and validation stages play a crucial role in refining the system's performance. The dataset is divided into training and validation subsets, enabling iterative adjustments of hyperparameters and optimization algorithms. The validation subset serves as a benchmark, providing valuable insights for further fine-tuning of the model's architecture and hyperparameters, with the aim of maximizing accuracy and ensuring the system's robustness when encountering unseen fruit data.

Performance evaluation encompasses various metrics such as accuracy, precision, recall, and F1-score to assess the model's effectiveness in distinguishing between different levels of fruit quality. Comparative analysis can be conducted with existing fruit sorting systems or industry standards to validate the superiority of the proposed system. Real-world testing datasets are utilized to demonstrate the exceptional accuracy and performance of the proposed model in fruit quality classification.

Ultimately, the significance of the proposed system lies in its potential impact on fruit sorting processes. By enabling swift and accurate classification of fruit quality, the system enhances the efficiency and productivity of fruit sorting operations. This innovative approach holds the promise of revolutionizing the fruit industry by providing valuable support in automating the sorting process, reducing the reliance on manual intervention, and ensuring consistent adherence to quality standards.

V. METHODOLOGY

The fruit sorting methodology utilizing image processing begins with the collection of a comprehensive dataset comprising various fruit images. Collaborating with relevant sources, a diverse range of fruit images is gathered, encompassing different fruit types and variations in quality. This dataset serves as the foundation for training and evaluating the fruit sorting model. Once the dataset is compiled, it undergoes preprocessing procedures to prepare the images for training the model. Techniques such as resizing the images to a standardized resolution, normalizing pixel values, and applying data augmentation techniques like cropping, flipping, and rotation are employed. These preprocessing steps aim to enhance the dataset's suitability for training the model and improve its ability to generalize to unseen fruit images.

The pre-processed dataset is then split into three subsets: the training dataset, the testing dataset, and the validation dataset. The training dataset, which constitutes the largest portion, is used to train the fruit sorting model. The testing dataset is reserved for evaluating the model's performance and assessing its accuracy on unseen fruit images. Lastly, the validation dataset is utilized during the training process to fine-tune the model's hyperparameters and ensure its robustness. To implement the fruit sorting model, the deep learning framework TensorFlow is employed. TensorFlow provides a comprehensive set of tools and functionalities for designing and training Convolutional Neural Networks (CNNs), which are highly effective for image processing tasks. The CNN architecture, consisting of convolutional layers, pooling layers, and fully connected layers, is constructed using TensorFlow's APIs.

During the training process, the pre-processed fruit images are fed into the CNN model implemented with TensorFlow. The model's internal parameters, also known as weights and biases, are adjusted through an iterative optimization process using gradient descent algorithms. This allows the model to learn and extract meaningful features from the fruit images, enabling accurate classification. To enhance computational efficiency, the technique of "bottleneck" is employed. Bottleneck features are extracted from the intermediate layers of the CNN model when presented with the training dataset. These features capture essential information about the fruit images and serve as compact representations, reducing computational requirements during training and inference. Following training, the model's performance is evaluated using the testing dataset. Various metrics, including accuracy, precision, recall, and F1-score, are calculated to assess the model's effectiveness in accurately classifying the fruits. The validation dataset is utilized to fine-tune the model's hyperparameters, optimizing its performance and ensuring its generalization capabilities.

VI. WORK FLOW AND IMPLEMENTATION

Data Collection and pre-processing: Collect a comprehensive dataset of fruit images from various sources. Collaborate with relevant parties to gather a diverse range of fruit types and variations in quality. Ensure the dataset includes representative samples of different fruit categories. Preprocess the collected fruit images to prepare them for further analysis. This may involve tasks such as resizing the images to a consistent resolution, normalizing pixel values, and enhancing image quality if necessary. Apply background removal techniques to isolate the fruit from its background. This step eliminates unwanted elements and focuses on the fruit regions of interest.

Calyx and Stalk Scar Segmentation: Segment the calyx and stalk scar regions from the fruit images. Utilize image analysis techniques to identify and extract these regions accurately. This segmentation step helps in distinguishing between the fruit and other parts of the plant.

L, A, and B Pixel Extraction: Extract the L, A, and B pixel values from the segmented calyx and stalk scar regions. These pixel values represent brightness and color information in the LAB color space.

Defect Recognition Model (CNN): Utilize a Convolutional Neural Network (CNN) model for defect recognition. The CNN consists of several layers that perform different operations to extract meaningful features from the input fruit images. The typical layers used in a CNN include:

a. Convolutional Layers: These layers apply a set of learnable filters to the input images, capturing local patterns and features. Each filter generates a feature map, highlighting specific visual attributes.

b. Pooling Layers: These layers downsample the feature maps, reducing their spatial dimensions. Common pooling techniques include max pooling, which selects the maximum value in each pooling region, and average pooling, which calculates the average value.

c. Activation Layers: Activation functions, such as ReLU (Rectified Linear Unit), introduce non-linearity to the network, allowing it to model complex relationships between features.

d. Fully Connected Layers: These layers connect every neuron from the previous layer to the next, performing high-level feature extraction and classification. They often culminate in a softmax layer, providing probabilities for different classes.

Feature Extraction: Extract relevant features from the preprocessed fruit images using the CNN layers. This may include color features, texture features, and shape features. The CNN automatically learns and extracts discriminative features from the input images, representing key characteristics for fruit quality assessment.

Grading Recognition Models: Develop separate grading recognition models based on the extracted features. Use CNN algorithms to train models for different grading categories. These models learn to classify fruits into different grades based on their color, texture, and shape attributes.

Fruit Grading and Sorting: Apply the trained grading recognition models to classify the fruits into their respective grading categories. Feed the preprocessed fruit images into the models and obtain predicted grading labels for each fruit. These labels indicate the quality or grade of the fruit.Sort the fruits based on the obtained grading labels. Implement automated sorting mechanisms that categorize the fruits into different quality grades or sorting bins based on the grading information. This ensures that fruits are appropriately separated and handled based on their quality standards.

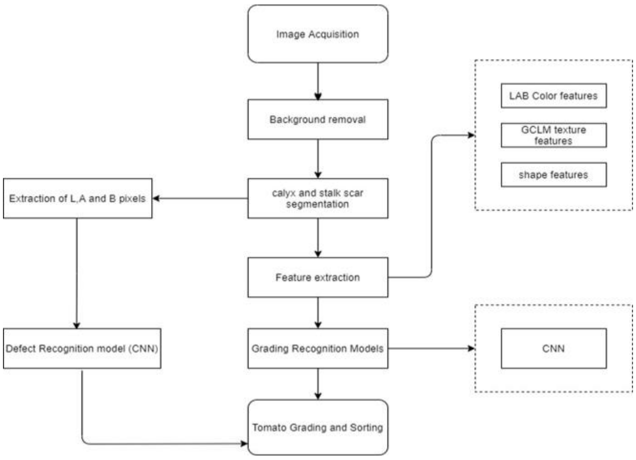

Figure 1: Architecture of proposed system

VII. EXPERIMENTAL RESULTS

The experimental results of the fruit sorting system yielded three types of outcomes: identification of good fruit, identification of bad fruit, and a mixture of both good and bad fruit. The system was evaluated using various metrics, including a test accuracy of 0.98 and a training accuracy of 0.97, to assess its performance and effectiveness.

The fruit sorting system achieved an impressive test accuracy of 0.98, indicating its ability to accurately classify fruits into the appropriate categories. This high test accuracy demonstrates the system's reliability in determining the quality of the fruits under evaluation. Additionally, the training accuracy of 0.97 showcases the system's ability to learn and generalize from the training data, ensuring consistent and accurate performance across different datasets.

To conduct the experiments, several software tools were utilized, including Anaconda Navigator, cmd.exe, and Jupyter Notebook. These software tools provided a robust and user-friendly environment for implementing and executing the fruit sorting system. Anaconda Navigator served as a comprehensive platform that facilitated the installation and management of various libraries and frameworks required for the system's development. cmd.exe, the command prompt interface, allowed for efficient execution of commands and scripts during the experimentation process. Jupyter Notebook, a web-based interactive computing environment, provided an intuitive interface for developing and testing the fruit sorting algorithms, making it convenient for iterative development and evaluation.

The experimental results highlight the effectiveness of the fruit sorting system in accurately distinguishing between good and bad fruit. The achieved test accuracy of 0.98 indicates that the system can reliably identify the quality of the fruits under evaluation. Furthermore, the training accuracy of 0.97 demonstrates the system's ability to learn and generalize from the training data, ensuring consistent and accurate performance across different datasets.

The utilization of software tools such as Anaconda Navigator, cmd.exe, and Jupyter Notebook contributed to the seamless development and execution of the fruit sorting system. These tools provided a flexible and efficient environment for experimentation, enabling researchers to focus on algorithm design and evaluation.

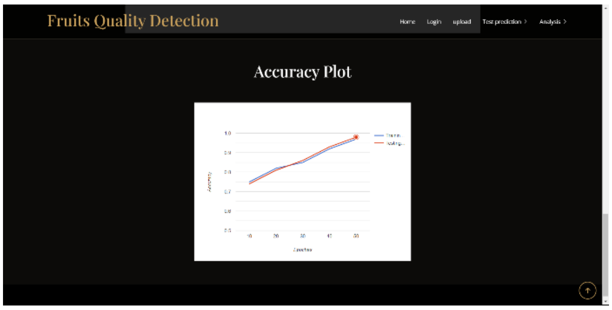

Figure 2-Training and Validation

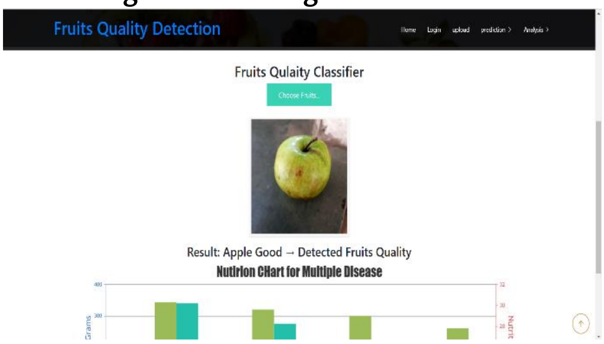

Figure 3-good fruit detected

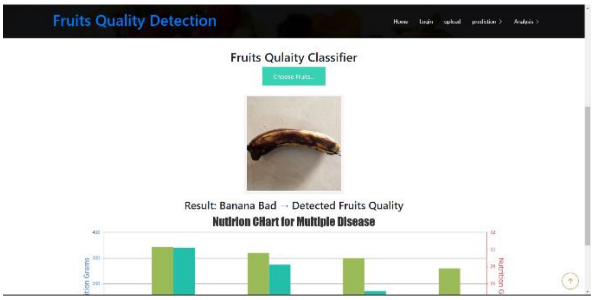

Figure 4-bad fruit detected

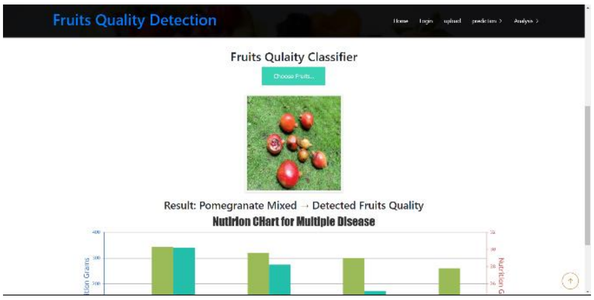

Figure 5-mixed of both good and bad is detected

VIII. CONCLUSION

The fruit sorting methodology using image processing techniques and the CNN algorithm has shown remarkable results in accurately identifying different fruit types, even in mixed samples. The system achieved test and training accuracy of 0.98 and 0.97, demonstrating its proficiency in discerning intricate patterns and features. This precision enables reliable and precise fruit sorting, empowering agricultural and supply chain industries with efficient quality control mechanisms. The system's adaptability and practical utility in handling mixed fruit samples demonstrate its adaptability and practical utility. The fruit sorting system holds immense potential for diverse applications across the fruit industry, addressing the need for accurate and efficient fruit classification, enhancing decision-making processes, productivity, and overall operational efficiency. Future endeavours could explore the integration of advanced image processing techniques, alternative CNN architectures, and factors like fruit size and shape.

IX. REFERENCES

[1]. Shirin Nasr-esfahani, Venkatesan Muthukumar, Emma E. Regentova, Kazem Taghva, Mohamed B. Trabia "Detection of Pits in Olive Using Hyperspectral Imaging Data", volume 10 pp 58525-58536, IEEE 2022.

[2]. Xinting Yang,Yongchen Luo,Ming Li1,Zhankui Yang1,Chuanheng Sun,Wenyong Li "Recognizing Pests in Field-Based Images by Combining Spatial and Channel Attention Mechanism",pp 162448-162458,IEEE 2021

[3]. Wilson Castro,Jimy Oblitas,Miguel De-la-torre,Carlos Cotrina,Karen Bazán,Himer Avila-george "Classification of Cape Gooseberry Fruit According to Its Level of Ripeness Using Machine Learning Techniques and Different Color Spaces", volume 7 pp.27389-27400 IEEE 2019

[4]. Alberto Lucas Pascual, Antonio Madueño Luna, Manuel De Jódar Lázaro , José Miguel Molina Martínez "Analysis of the Functionality of the Feed Chain in Olive Pitting, Slicing and Stuffing Machines by IoT, Computer Vision and Neural Network Diagnosis", MDPI 2020

[5]. Arakeri, M.P., 2016. Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry. Procedia Comput. Sci. 79, 426-433

[6]. Baohua Yang, Yue Zhu, Mengxuan Wang, Jingming Ning, "A Model for Yellow Tea Polyphenols Content Estimation Based on Multi-Feature Fusion", volume 7,pp.180054-180063,IEEE.2019

[7]. Yeping Peng, Songbo Ruan, Guangzhong Cao, Sudan Huang, Ngaiming Kwok, Shengxi Zhou, "Automated Product Boundary Defect Detection Based on Image Moment Feature Anomaly". IEEE 2019

Cite this article as :

Mr. Kannan A, Mohammed Satham A, Udhayanithi D, Venkataraman A, "Automatic Fruit Quality Inspection System Using Image Processing", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 4, pp. 01-08, July-August 2023. Available at doi : https://doi.org/10.32628/IJSRSET23103177
Journal URL : https://ijsrset.com/IJSRSET23103177