Real-Time Visual Inspection System for Grading Fruits Using Computer Vision and DEP Learning Techniques

R. Sowmiya[1], K. Praveena[2]

[1, 2]M.C.A

[1, 2]A.R.J College of Engineering and Technology, Mannargudi

Abstract

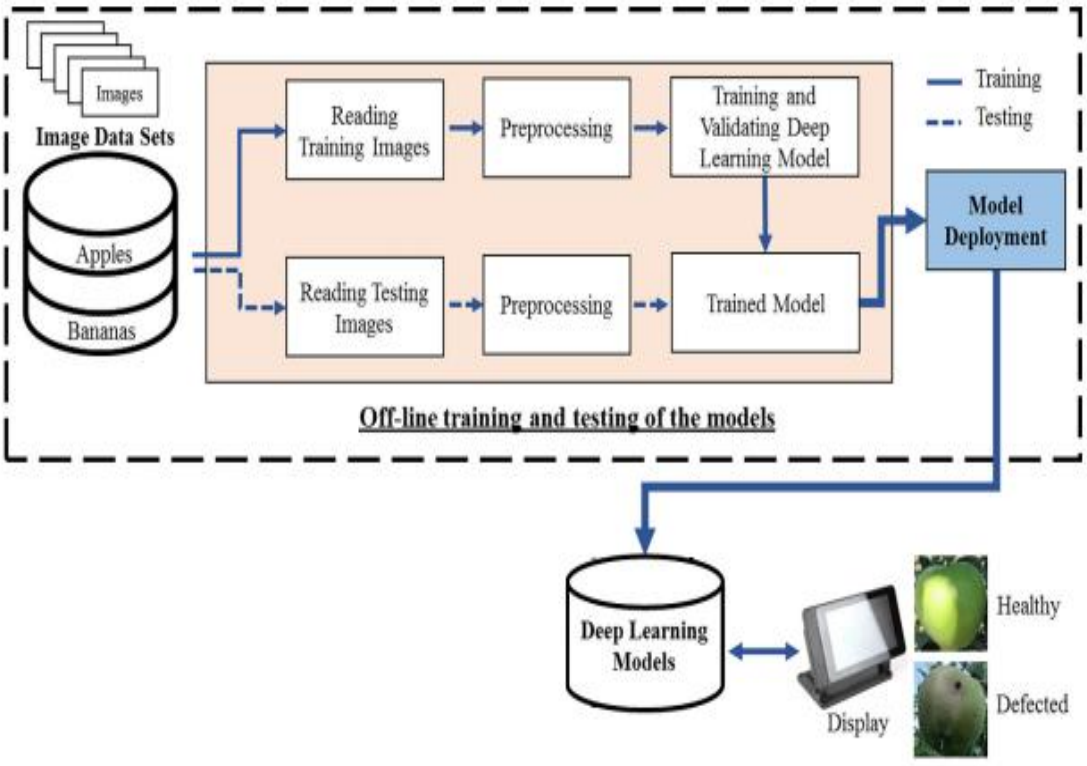
Traditional manual visual grading of fruits has been one of the important challenges faced by the agricultural industry due to its laborious nature as well as inconsistency in inspection and classification process. Automated defects detection using computer vision and machine learning has become a promising area of research with a high and direct impact on the domain of visual inspection. In this study, we propose an efficient and effective machine vision system based on the state-of-the-art deep learning techniques and stacking ensemble methods to offer a non-destructive and cost-effective solution for automating the visual inspection of fruits' freshness and appearance. We trained, tested and compared the performance of various deep learning models to find the best model for the grading of fruits.

I. INTRODUCTION

Agriculture industry has a vital role in the economy of many countries in the world. An important sector of agriculture business is the production and supply of the fresh fruits and vegetables to the vendors and markets. The growing demand for effective food production and quick and safe supply to the market has led to the development and use of various innovative technologies in this industry. The technologies such as Internet of Things (IoT) based smart farming has been found useful in improving the quality of fruit and vegetable yields. Moreover, the use of intelligent logistic by medium and large scale enterprises has reduced the time to sort, package and deliver the production to the market. However, at the level of small scale agricultural enterprises and farmers, limited embracing of the new technologies has been noticed.Two major challenges in adoption of these technologies by them are the overall increased cost and the requirement for learning specialized skills. Hence, there is a rising need to develop low-cost and easy-to-use solutions for these enterprises and farmers so that they can take more advantage of the new technologies. This study focuses on the automation of grading and sorting process which is one of the important phase of the fresh fruits' supply chain system. Since outer appearance reflects the freshness of a fruit and marks its selling point, it is one of the major criterion in grading the fruits. The tasks of grading and sorting of fruits according to their outer appearance and freshness are still laborious and challenging in nature at the level of small scale enterprises. The manual quality control highly depends on the trained humans to inspect the outer appearance of the fruits and then make a decision about their grading. Automating the grading and sorting of fruits is not only an essential step to deal with the inconsistency in classification but it also helps in reducing the labor cost and time spent in packaging, pricing and supplying these fruits to the markets or vendors. Thus, an alternative solution for low budget enterprises is to opt for a low-cost intelligent fruits grading system.

II. SYSTEM ARCHITECTURE

SVM works to solve classification problems for classes that are linearly separable or nonlinearly separable. The algorithm creates a hyperplane that separates the data into two classes. An ideal hyperplane is considered to have a maximum margin between classes. The data above the hyperplane belong to the first class, and the data below the hyperplane belongs to the second class. When the data are non-separable due to their overlap, the solution is to transform the data from a nonlinear data space to a separable data space by finding the hyperlevel that causes the fewest possible errors.This mechanism is called the lazy learning model because it is characterized by not mining data from knowledge. The algorithm trains and stores the training dataset for future use. When classifying a new condition (test point), the algorithm compares the test point with the stored data and finds the similarity between the test points and the stored dataset (training dataset). According to the principle of Euclidean distance, KNN measures the distance between the test data point and the stored data. Then, a new point is mapped to the nearest neighbour. A decision Tree is a machine learning algorithm called symbolic learning, wherein decision rules are closely related to trees and control the flow of operations. This algorithm consists of the root node (represents the entire dataset), the branch node (represents the features) and the root node (represents the final decision). In each decision, two or more branches are generated from each child node. The process continues until the process reaches the last decision, and there is no new decision (leaf node), meaning that each case depends on the values of the features involved in making a decision. Thus, the process starts from the root node and moves each state to the next branch until it reaches the leaf node. The training dataset is divided into several subsets of data according to the selected feature values. The features remaining in each subset are evaluated, and logical decision rules are formulated to the next level of the tree.



II.A System Architectures

Input design is one of the most important phases of the system design. Input design is the process where the input received in the system are planned and designed, so as to get necessary information from the user,

eliminating the information that is not required. The aim of the input design is to ensure the maximum possible levels of accuracy and also ensures that the input is accessible that understood by the user. The input design is the part of overall system design, which requires very careful attention. If the data going into the system is incorrect then the processing and output will magnify the errors.A synchronization packet (commonly known as the timing reference signal) occurs immediately before the first active sample on every line, and immediately after the last active sample (and before the start of the horizontal blanking region).A systems flowchart specifies master files, transaction files and computer programs. Input Data are collected and organized into groups of similar data. Once identified, appropriate input media are selected for processing. The output devices to consider depend on factors such as compatibility of the device with the system, response time requirements, expected print quality and number of copies needed. . All nodes in the network may depart or fail unpredictably.

III. SYSTEM TESTING

The philosophy behind testing is to find the errors. A good test is one that has a high probability of finding an undiscovered error. A successful test is one that uncovers the undiscovered error. Test cases are devised with this purpose in mind. A test case is a set of data that the system will process as an input. However the data are created with the intent of determining whether the system will process them correctly without any errors to produce the required output. The user interface testing is important since the user has to declare that the arrangements made in frames are convenient and it is satisfied. When the frames where given for the test, the end user gave suggestion. Based on their suggestions the frames where modified and put into practice.Testing is a series of different tests that whose primary purpose is to fully exercise the computer based system. Although each test has a different purpose, all work should verify that all system element have been properly integrated and performed allocated function. Testing is the process of checking whether the developed system works according to the actual requirement and objectives of the system.After performing the validation testing the next step is output testing of the proposed system. Since the system cannot be useful if it does not produce the required output. Asking the user about the format in which the system is required tests the output displayed or generated by the system under consideration. Here the output format is considered in two ways. One is on screen and another one is printed format. The output format on the screen is found to be corrected as the format was designed in the system phase according to the user needs. And for the hardcopy the output comes according to the specifications requested by the user.This project is application based project, and the modules are interdependent with the other modules, so the testing cannot be done module by module. So the unit testing is not possible in the case of this driver. So this system is checked only with their performance to check their quality.

IV. SOFTWARE DESCRIPTION

Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. It was created by Guido van Rossum during 1985- 1990. Like Perl, Python source code is also available under the GNU General Public License (GPL). This tutorial gives enough understanding on Python programming language.Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages. Python is a MUST for students and working professionals to become a great Software Engineer especially when they are working in Web Development Domain.Pandas is mainly used for data analysis and associated manipulation of tabular data in Data frames. Pandas allows importing data from various file formats such as comma-separated values, JSON, Parquet, SQL database tables or queries, and Microsoft Excel. Pandas allows various data manipulation operations such as merging, reshaping, selecting, as well as data cleaning,

and data wrangling features. The development of pandas introduced into Python many comparable features of working with Data frames that were established in the R programming language. The panda's library is built upon another library NumPy, which is oriented to efficiently working with arrays instead of the features of working on Data frames.NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays.Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK.

V. CONCLUSION

A deep learning based low-cost machine vision system for grading the fruits based on their outer appearance or freshness. Various state-of-the-art deep learning models and stacking ensemble deep learning methods were applied to two data sets of fruits. The results of this study show that models and their stacked combinations have the highest accuracy in grading the test set and real samples as compared to the other deep learning models. Moreover, the application of deep learning models has been found more accurate in classification of fruits as compared to the results previously reported while applying traditional machine learning techniques using the feature extraction methods. However, there are certain limitations of the current study which require further investigations. In our proposed system, classification is solely based on the outer appearance of the fruit and it only uses single view of the fruits' images. In future, the system will be trained and tested using multi-view vision system for capturing image data sets. In addition, for real-time testing of the system, the lighting box will be setup over a conveyor belt with an environment where multiple camera systems will be installed to capture the images of the fruits from multi-dimensions.

VI. REFERENCE

[1] Tripathi MK, Maktedar DD. A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: a survey. Inf Process Agric 2020;7:183–203.
[2] Yan B, Wu XH, Ye B, Zhang YW. Three-level supply chain coordination of fresh agricultural products in the Internet of Things. IndManag Data Syst 2017;117:1842–65.
[3] Nukala R, Panduru K, Shields A, Riordan D, Doody P, Walsh J. Internet of Things: a review from Farm to Fork. In: Proc. Proceedings of 27th Irish Signals and Systems Conference. Londonderry, UK; 2016. p. 1–6.
[4] Dubey SR, Jalal AS. Apple disease classification using color, texture and shape features from images. Signal, Image Video Process 2016;10:819–26.
[5] Moallem P, Serajoddin A, Pourghassem H. Computer visionbased apple grading for golden delicious apples based on surface features. Inf Process Agric 2017;4:33–40.
[6] Mazen FMA, Nashat AA. Ripeness classification of bananas using an artificial neural network. Arab J SciEng 2019;44:6901–10.
[7] Tan K, Lee WS, Gan H, Wang S. Recognising blueberry fruit of different maturity using histogram oriented gradients and colour features in outdoor scenes. BiosystEng 2018;176:59–72.
[8] Silalahi DD, Rean˜ o CE, Lansigan FP, Panopio RG, Bantayan NC. Using genetic algorithm neural network on near infrared spectral data for ripeness grading of oil palm (Elaeisguineensis Jacq.) fresh fruit. Inf Process Agric 2016;3:252–61.
[9] Sabzi S, Abbaspour-Gilandeh Y, Garcı´a-Mateos G. A new approach for visual identification of orange varieties using neural networks and metaheuristic algorithms. Inf Process Agric 2018;5:162–72.