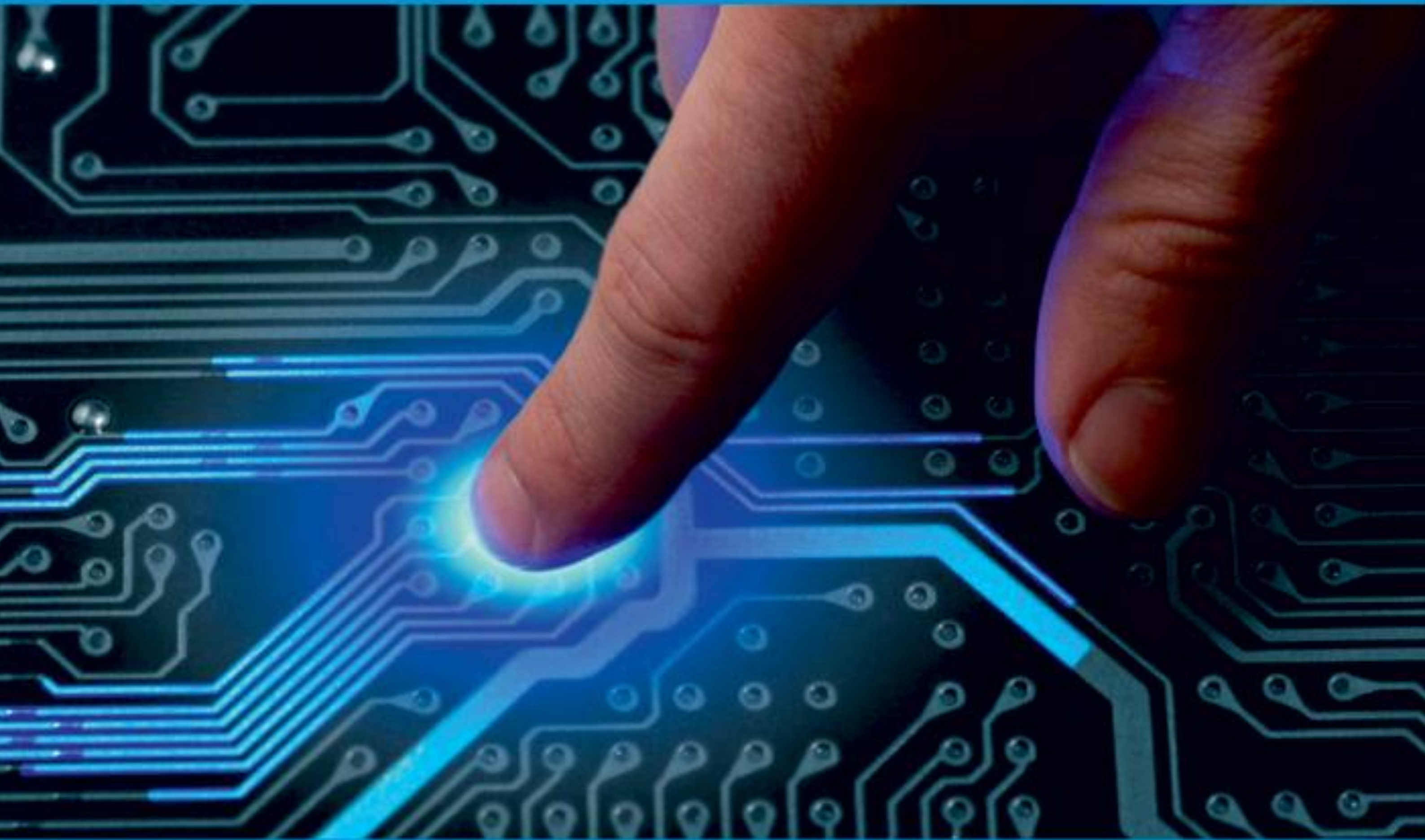
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 2, February 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 10, Issue 2, February 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1002039 |

Semantic Segmentation for Arecanut Classification using Support Vector Machine

Ujwal U.J[1], Smitha M.L[2], Savitha C.K[3]

Professor & Head, Department of Computer Science, KVG College of Engineering, Sullia, Karnataka, India[1]

Professor, Department of Computer Science, KVG College of Engineering, Sullia, Karnataka, India[2,3]

ABSTRACT: Arecanut is the major and popular commercial crop in many parts of India. Varieties of arecanuts are grown by the farmers in different parts of India. Classification of arecanuts based on their variety enhances the quality and allots a grading for the consumers and buyers. In this paper, a supervised classifier Support Vector Machine (SVM) is used for classifying the arecanut. Semantic segmentation is carried out to separate foreground from the background. Experimentation is performed over the image dataset consisting of 1200 images where 98% of the dataset is used for training and 2% of the dataset for testing. Finally, the performance of the proposed system is evaluated through metrics such as precision, recall and F –measure.

KEYWORDS*:* Semantic Segmentation, Feature Extraction, Support Vector Machine (SVM)

I. INTRODUCTION

Arecanut is one of the major cash crops of India. It plays an active role in Indian family and is used for refreshment and also as a part of tradition and culture. Areca nut is also popularly known as betel nut and has a vital role in the preparation of medicines in the field of Ayurveda. The people residing in coastal region grow this in different kinds as a commercial crop. India is one of the largest producer of arecanut in this world and exports it to many other parts of the world. Karnataka, Kerala and Assam are the three states that grow arecanut in India. Nearly 89% of arecanut is cultivated by the three states of Karnataka, Kerala and Assam. Other parts of India such as Meghalaya, Tamil Nadu, and West Bengal, the crop is grown in a very small area even though there are lots of consumers. The arecanut is cultivated in large number especially in places of Uttara Kannada, Dakshina Kannada and Udupi regions located in Karnataka and will occur as shown in Fig. 1.

II. RELATED WORK

Classification of arecanut plays a vital role while grading during buying and sales. The agriculturist can get a good rate while selling if the arecanut is of good quality. The automatic machines developed aren't efficient to completely remove the husk from the nut because of variation with the size of the betel nut, and therefore the arecanuts are damaged due to the non-availability of the optimum process parameters and hence, they are expensive. Proper classification helps in fixing the right price for right crop. If the arecanuts are properly classified into different categories and sorted then the farmer earns a good price. The labour charges incurred while sorting is also saved by using a good classification system. The prior works found in literature are automatic arecanut grading systems that focused on color, texture, and shape as features [2]. These methods are based on segmentation, masking and classification [3]. Classification based on texture features[1], [4] have been reported in literature where segmentation was done using global based threshold and Otsu methods [6] followed by classification using mean and GLCM [9]. Some of the works were found in literature using color features for segmentation with SVM for two grade boiling and Non-boiling classification [5]. Few works also focused on arecanuts which were husk free and were classified by the use of texture features with NN classifier [6],[8].

III. PROPOSED SYSTEM

The proposed model consists of three different stages as depicted in Fig 2. The first stage is segmenting the areca from its background which is mainly done to avoid the shadow effect occurred while capturing the image.

The images are segmented using Semantic segmentation and the foreground object is retrieved after preprocessing. The second stage used here is to extract the features from the segmented image.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 10, Issue 2, February 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1002039 |

In the third stage, the classification of arecanut into five different classes is carried out. For the purpose of classification, well known supervised classifier Support vector machine (SVM) is used. After training, the test images are given in testing phase to match with the trained image. Upon proper matching, the test image will be labelled with the class of the trained image.

Fig. 1: Sample of Arecanuts in Dakshina Kannada Region

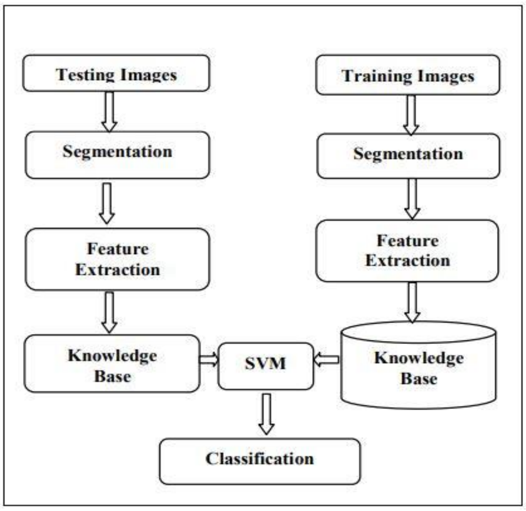


Fig. 2: Proposed Model for Arecanut Classification

- **Input Image**: Both the test image and training images are fetched from the data set stored in computer. To maintain uniformity, the captured images were resized and stored in the dataset.
- **Image Segmentation**: Image segmentation techniques have been applied to split the pixels of the image into two subsets: object area and the background.
- **Feature Extraction:** Feature extraction is the process to retrieve the most important data from the raw data. The major goal of the feature extraction is to extract set of features, which maximizes the recognition rate with least number of elements.
- **Training Classifier for Testing**: This stage is training support vector machine classifier for classification of the test image using features stored in the knowledge base.
- **Trained Model:** The training model consists of classifier which is used to classify the images and storage unit to store the extracted features of the training images.
- **Classification:** This is the final stage of the proposed model which focuses on the labelling the unknown Arecanut image to the known finite classes based on similarity of the images.

In this work, samples of five different classes of arecanut as shown in Fig. 3 are collected for the purpose of classification and the steps involved are discussed in detail below:

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 10, Issue 2, February 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1002039 |

a. Segmentation

Segmentation is the method of partitioning the image into sub parts so as to get the area of interest for further processing. The procedure of partition is carried out based on two major properties of intensity, discontinuity and similarity.

Semantic Segmentation: Semantic Segmentation is the process of assigning a label to every pixel in the image. This is in stark contrast to classification, where a single label is assigned to the entire picture. Semantic segmentation treats multiple objects of the same class as a single entity as shown in Fig. 4. Semantic segmentation algorithm provides a fine-grained, pixel-level approach to develop computer vision applications. It tags every pixel in an image with a class label from a predefined set of classes.

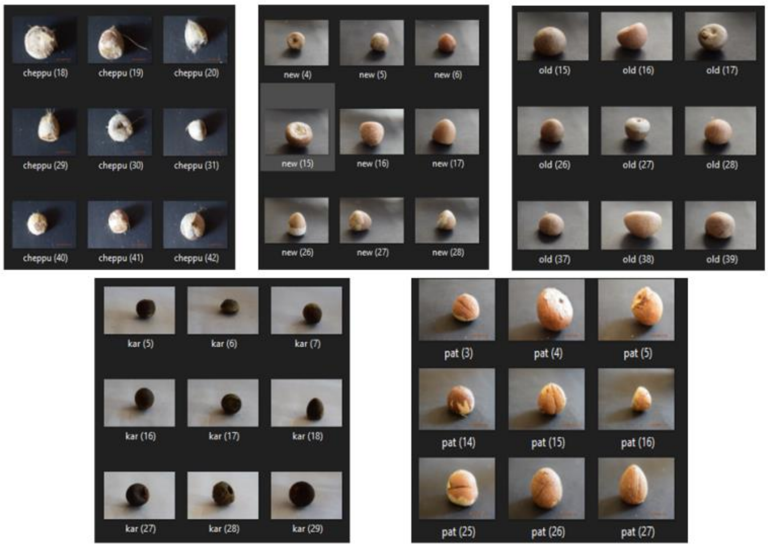


Fig. 3: Five classes of Arecanut

Fig. 4: Results after performing Semantic Segmentation

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 10, Issue 2, February 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1002039 |

b. ***Feature Extraction***

Feature extraction is the process of extracting the information of an image which aids in classification. The colour, shape, and texture features are most commonly used in computer vision applications [1]. Color histograms are most widely used feature extraction method for color image retrieval applications where RGB color space is used. When the segmented image is fed into this extraction method, color histogram is generated for each and every image from the given set of data. These histograms consist of proportion of each color pixels of the image. For a color image, histogram is the collection of bins, where each bin specifies the probability of specific color pixels of the image. While generating the histogram for a given image, each pixel is mapped to a bin and each bin consists of number of pixels having the same color [10].

c. ***Classification***

This is the final stage of the proposed model which focuses on labelling the unknown arecanut image to the known finite classes based on similarity of features. Classification is a supervised learning technique which has training and testing stage. In training stage, a sample is labelled and kept in a specific class. While testing, a unknown sample image will be compared with trained sample and if it is found similar to the training image, it is labelled with that of a training image. In this work, Support Vector Machine (SVM) is used for classification [7].

SVM is mainly devised for classifying a two class problem by separating the classes using creating of hyperplane. A hyperplane is a process of boundary fitting to a region of points which are similar. This makes the similar points to classify into a specific class. Once a hyperplane is created for training samples, for testing samples, we need to check whether they are within the same boundary or not. If it lies within the boundary it will be classified to a known created class. For a simple two attribute problem the hyperplane is either a straight line or a curve for separating them.

The direct approach SVM for multi class problem is not well suited, so it will be avoided. For multi-class problem different approaches will be employed. A multi category classification classifies the pattern by computing the c linear discriminant functions given by,

$$g_i(x) = W^{'}(x) + W_{i0}$$

for i = 1,2,......,c and assign x to the category corresponding to the largest discriminant. A One-Versus-One classifier is constructed by dividing n-classes into n / (n-1)/2 binary classes. Each case will have a binary classifier responsible for dividing different pairs of classes.

Learning is done by training the data with any of the two labels corresponding to the classes. If it encounters any class labels which are not similar, it will be ignored. During testing, a pattern is given to each one of the binary classifiers.

The output is given by S_{xy}. This provides confidence score of the binary classifier for dividing class x and y as per the previous class. The class having highest confidence scope will be the output class for the classifier. The scope matrix of the classifier is given by,

$$S = \begin{bmatrix} S_{11} & .... & S_{1x} \\ S_{21} & .... & S_{2x} \\ . & .... & \\ S_{x1} S_{x2} & .... & S_{xy} \end{bmatrix}$$

In SVM, training is easy and operates nicely with dimensional data. Both continuous and categorical data can be easily handled. It predicts the nonlinear relations within the data and is a non-parametric technique. It deals well with erroneous data. Henceforth, the SVM classifier has been utilized to classify the arecanut into five different classes.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 10, Issue 2, February 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1002039 |

IV. DATASET AND EXPERIMENTATION

The dataset for experimentation is obtained by capturing the samples as the standard dataset was not available for arecanuts. The dataset comprising of 1200 images were used to classify into five classes as shown in Fig. 5.

Fig. 5: Sample Images in Arecanut Dataset

The five classes considered in this work are 1: Cheppugotu 2: Karigotu 3: Patora 4: New Areca and 5: Old Areca. Real time data collection has been done for construction of dataset. These images for the dataset were captured using Nikon digital camera at arecanut processing centers in Sullia Taluk, Dakshina Kannada district of Karnataka state in India.

Out of the known metrics for validating machine learning models, we choose *Accuracy* and *F1 as* they are the most used in supervised machine learning. For the accuracy score, it shows the percentage of the true positive and true negative to all data points. So, it's useful when the data set is balanced. For the F1 score, it calculates the harmonic mean between precision and recall, and both depend on the false positive and false negative. So, it's useful to calculate the F1 score when the data set isn't balanced.

In SVM, we used two different kernel functions polynomial and Radial Basis Function (RBF). SVM with RBF kernel function is outperforming SVM with Polynomial kernel function. It has achieved greater accuracy during training of the image dataset. The proposed system calculates the accuracy and F1 scores to show the performance difference between the two selected kernel functions (Poly and RBF) on the same data set as shown in Fig. 6.

```
The predicted Data is :
[4 2 1 4 4 1 4 2 0 4 0 1 4 1 2 0 3 4 2 0 3 0 0 1 4 1 2 4 4 1 2 1 1 0 3 0 0
 1 3 3 0 1 1 4 1 0 4 4 4 2 0 3 3 0 4 4 1 0 0 1 1 1 4 2 4 1 0 2 2 1 4 4 0 1
 2 1 1 1 2 1 4 4 4 3 1 2 4 2 1 1 4 3 4 4 3 4 0 1 3 0 1 2 0 0 1 4 4 0 1 4 0
 0 4 0 1 3 3 0 4 1 4 3 1 1 0 3 4 4 4 0 4 3 4 1 3 3 0 2 1 1 3 0 1 1 1 3 1 4
 4 0 0 0 0 2 0 4 0 0 0 2 4 2 4 2 0 0 3 4 0 0 4 2 0 4 0 1 1 0 4 4 1 2 0 1 4
 1 3 1 1 4 3 4 4 4 4 1 0 0 1 0 3 4 4 4 1 4 4 0 3 1 4 3 4 4 1 1 2 0 0 1 4 1
 1 0 2 3 2 4 1 0 3 2 4 1 4 2 4 4 1 3 2 1 0 4 1 4 1 0 1 4 1 1 4 4 4 0 4 0 2
 2 4 3 0 0 3 3 2 1 0 4 0 4 1 0 4 0 0 4 4 2 0 1 4 1 4 1 0 2 3 0 0 0 0 0 1 4
 3 4 4 2 0 1 1 1 0 1 4 4 1 0 0 2 3 2 3 2 3 3 4 3 1 4 4 2 0 4 4 3 4]
The actual data is:
[0 2 1 0 0 1 4 0 0 0 0 1 0 1 3 0 4 0 2 0 2 0 0 1 4 1 0 4 4 1 2 1 1 0 3 0 0
 1 3 3 4 1 1 4 1 2 4 4 2 4 3 3 3 0 4 4 1 0 4 1 1 1 4 2 4 1 0 2 3 1 4 4 0 1
 2 1 1 1 0 1 4 4 4 3 1 2 4 0 1 1 4 3 0 4 3 0 0 1 3 2 1 4 0 0 1 4 4 0 1 2 0
 0 4 4 1 3 3 0 4 1 2 3 1 1 0 3 4 4 4 0 4 3 4 1 3 3 0 4 1 1 3 0 1 1 1 2 1 0
 3 4 0 0 0 4 0 4 0 0 0 4 3 2 4 3 0 4 3 0 0 0 0 2 0 4 0 1 1 0 4 0 2 2 0 1 4
 1 2 1 1 0 2 4 4 4 4 1 0 0 1 0 3 4 0 4 1 0 4 0 3 1 4 3 2 4 1 1 2 0 0 1 4 1
 1 4 2 3 2 4 1 0 3 2 4 1 4 3 4 0 1 3 2 1 0 4 1 4 1 0 1 0 1 1 4 4 4 0 4 0 0
 0 4 3 0 0 0 2 2 1 4 4 0 2 1 0 4 4 0 4 2 4 0 1 2 1 4 1 2 3 3 0 0 2 0 0 1 4
 3 0 4 2 0 1 1 1 0 1 0 4 1 0 3 2 2 2 3 0 3 3 4 2 1 4 2 2 0 4 4 3 4]
The model is 78.419452887538% accurate
```

Fig. 6: Performance of the proposed system

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 10, Issue 2, February 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1002039 |

The proposed work can help the farmers/business persons to automatically classify the Arecanut much faster without any man power. It also helps to obtain higher accuracy than a human performing the same task.

The classification results for different categories of arecanut obtained by the proposed system are shown below in figures (Fig. 7 to Fig. 11):

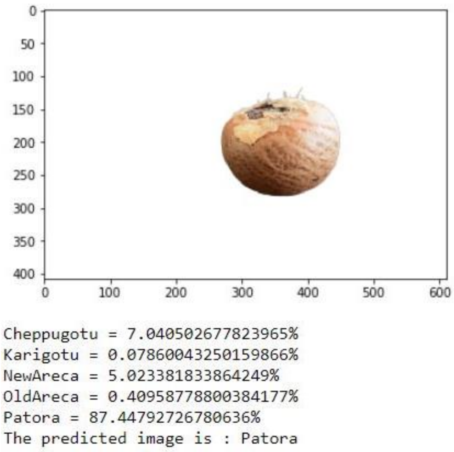


Fig. 7: Predicted image after classification as Patora

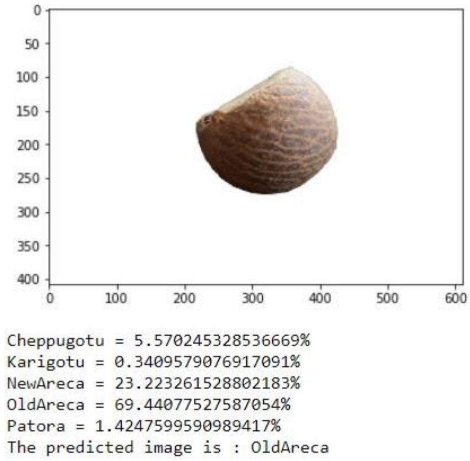


Fig. 8: Predicted image after classification as Old Areca

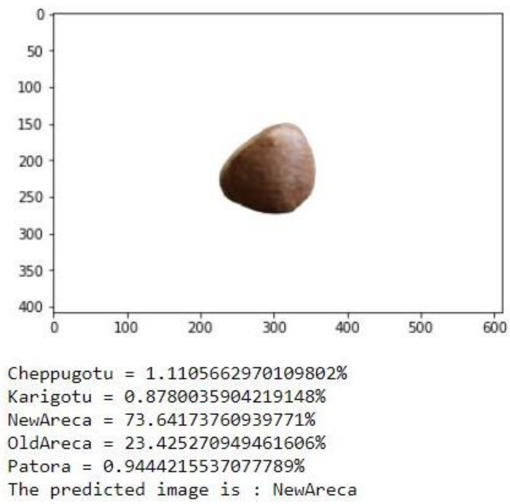


Fig. 9: Predicted image after classification as New Areca

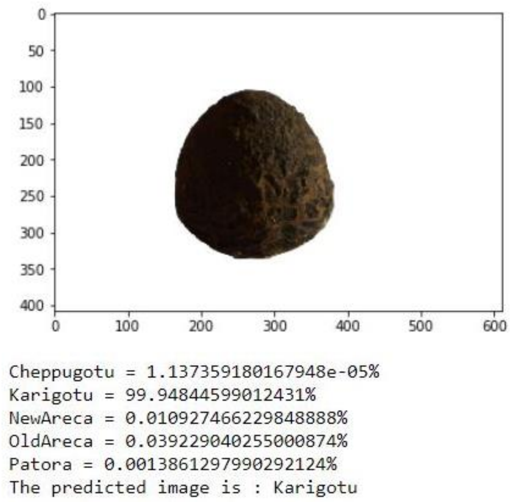


Fig. 10: Predicted image after classification as Karigotu

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 10, Issue 2, February 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1002039 |

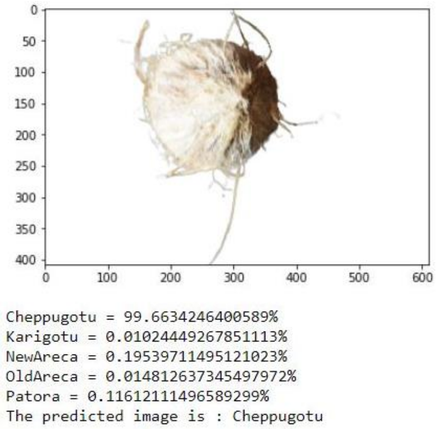


Fig. 11: Predicted image after classification as Cheppugotu

V. CONCLUSION

Arecanut classification is a critical task in computer vision application. By selling the classified arecanuts at their right prices the farmer would be able to earn more profit compared to selling it unclassified. The manual cost and sorting time always impacts the income of the farmers. In this work, classification of areca nut is carried out using SVM classifier by obtaining good accuracy for prediction of arecanut classes.

REFERENCES

[1] Ajit Danti, Suresha M, S. K Narasimhamurthy, "Classification of Diseased Arecanut based on Texture Features", International Journal of Computer Applications (0975 – 8887) Recent Advances in Information Technology, 2014.

[2] S Siddesha, S K Niranjan," Different Classifiers in Classification of Raw Arecanut", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-6, August 2019.

[3] H. Chandrashekhara, M. Suresha, "Classification of Healthy and Diseased Arecanuts using SVM Classifier", International Journal of Computer Sciences and Engineering Vol.7 (2), Feb 2019, E-ISSN: 2347-2693.

[4] Dhanuja K C, Mohan Kumar H P, "Arecanut Disease Detection using Image Processing Technology", International Journal of Engineering Research & Technology (IJERT) Vol. 9 Issue 08, August-2020.

[5] Pushparani M.K, Dr. D Vinod Kumar, Dr. Abdulla Gubbi, "Arecanut Grade Analysis using Image Processing Techniques", International Journal of Engineering Research & Technology (IJERT) NCRACES - 2019.

[6] S Siddesha, S K Niranjan, "Detection Of Affected Regions Of Disease Arecanut Using K-Means And Otsu Method", International Journal of Scienctific and Technological Research,Volume 9, Issue 02, ISSN 2277-8616, February 2020.

[7] Kuo-Yi Huang, "Detection and classification of Arecanuts with machine vision", Published by Elsevier Ltd. Selection and/or peer-review under responsibility of C3IT, 2011.

[8] Suresha M, Chandrashekhara H ,"Classification of Arecanut using Neural Networks with Feed-forward Techniques ", International Journal of Research in Advent Technology, Vol.7, No.3, March 2019 E-ISSN: 2321-9637.

[9] Suresh M, Chandrashekhara H, "Classification of algorithms using soft computing algorithms", International Journal of Computer Engineering and Applications, Volume XI, Issue X, October 17, www.ijcea.com ISSN 2321-3469.

[10] Mr. Ganesh Shetty, Jyotsna U Bailoor, Mathias Christen Sunny, Karthik R Anchan, Melston Joyel Tauro, "Segregation of Cashew Kernel and Arecanut by Using Advanced Color Sorting Mechanism", International Journal of Scientific Development and Research (IJSDR) May 2018.

BIOGRAPHY

Ujwal U.J is a professor & head, Department of Computer Science Engineering, KVG College of Engineering, Sullia, Karnataka, India. He pursued his doctorate from Visvesvaraya Technological University, Belagavi in the year 2020. His research interests are Cloud Computing, Web mining, Data Mining.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 10, Issue 2, February 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1002039 |

Smitha M.L is a professor, Department of Computer Science Engineering, KVG College of Engineering, Sullia, Karnataka, India. She pursued her doctorate from Mangalore University, Mangalore in the year 2016. Her research interests are Image/Video Processing, Pattern Recognition and Machine Learning.

Savitha C.K is a professor, Department of Computer Science Engineering, KVG College of Engineering, Sullia, Karnataka, India. She pursued her doctorate from Visvesvaraya Technological University, Belagavi in the year 2021. Her research interests are Artificial Intelligence and Machine Learning.

SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING