**SEGREGATION OF DRY ARECA NUT USING DIGITAL IMAGE PROCESSING AND MACHINE LEARNING TECHNIQUES**

Meghana R1, Ramya D Shetty\*

Shri Madhwa Vadiraja Institute of Technology and Management

## *Abstract*---Areca Nut production is dominant in India, especially in the southern parts of India, in the region of Karnataka. Day by day, the labor charges are also increasing with less productivity. This project brings a contribution to the overall sustainability of the areca nut. The method that we propose for feature extraction and classification of the image is Inception version 3 to extract the image features and classify the image accordingly. The second approach is that we use the feature extraction in the image processing and use K-nearest neighbor algorithm for the classification of arecanut images. We have done comparative study based on the above-mentioned approaches.

## *Index Terms*---Inception v3, Recognition Model, Image Processing, Feature Extraction

**1. Introduction**

# In the area of Karnataka state of India Areca's nut production is dominant. Malenadu is about 100 km wide and covers the western east slopes of the Sahyadri Western Ghats. It is situated between Karnataka Coast and the Karnataka tropical Bayaluseeme area, which is popularly called betel nut, as it is commonly used for chewing with betel leaves in the country. It is a palm tree that belongs to the Arecaceae family. It is relevant not only in India, but also in China and Southeast Asia for business and business. [1].Farmers currently face problems in the separation of dry areca nuts, the labor cost is rising every day and reduces productivity. The current problems faced by the farmers are given Figure 1.

# The results on the limitations of the areca nut growers suggest that the organizations and experts concerned should give their attention to help the areca nut growers overcome the constraints of agriculture in general and areca nut growing in particular

# In general, the project brings a contribution to the overall sustainability of the area: improvement of overall on sorting of dry areca nut using digital image processing and machine learning techniques [3].

# Figure 1: Problems faced by Farmers [2]

**2. Literature Review**

**2.1 Normal Image Classification Techniques:**

Edoardo et al. [4], shows a perspective on making the classification process for remote sensing images a more efficient SVM (Support Vector Machine) is introduced as a new active learning strategy. It basically employs the reformulation of original classification into a new problem and also the construction of corresponding significance space to suitably guide the selection of the samples used to better deal with the problem at first. In conclusion, a new active learning strategy has been proposed by the paper developed for SVM classification [4].

Bing et al. [5], focuses on Horror content sharing on the web. It basically speaks interference caused due to in our daily life. Therefore, this paper talks about the dangers of it caused on the mental health of its users. Thereby, a CMIL (Context-aware multi instance learning) algorithm has been presented by this paper. It helps in identifying and picking out horror images. It also employs FSVM (Fuzzy Support Vector Machine) for its strength [5]. We can understand the challenges in horror image recognition. A CMIL has been proposed in this paper taking into account that emotion of horror is not evoked by isolated image regions but by interaction among them. In conclusion, Experimental results have shown that the algorithm is superior to traditional CMIL methods and effective image classification in horror image reorganization [5].

**2.2 Classification of Image Using Neural Networks and Other Advance Techniques:**

Qian Wang et al., In this paper [6] the author is reviewing the image classification techniques that are already in use in the medical field. The concept of “Low level Processing to high level understanding” is mentioned, The author discusses about two steps involved mainly localization and segmentation, in localization region of interest after projecting common space by graph-based groupwise image recognition method and segmentation is a voxel-wise label is acquired with a knowledge aided convolution neural network with deep learning and Resnet 50.This method can overcome the lacking in classifying the small organs in the body can be overcome, By using CNN and Resnet 50 the author says the classification of melanoma disease can get up to the accuracy of 0.94 by using deep learning only the accuracy can go up to 0.87. While describing ophthalmic disease [6].

**2.3** **Summary of Literature Review with Research Gaps:**

With the aid of various algorithms, the recent approaches are segmented. As mentioned in previous articles, different models are used in the segmentation of the arecanut such as the support vector machine (SVM), which makes the input not so reliable. While the system automatically extracts features at a lower level from Deep Learning, such as the confolutionary neural network (CNN), to higher levels and gives good performance. The defined concept is that the process of low levels to a high level of understanding takes a lot less time to identify the subject but the quality of the image decreases so that we can either lose something or gain speed. The high degree of comprehension thus includes utilizing neural networks and fundamental learning concepts.

**3. Implementation**

Seed can be used in a cold, dried, fried, baked, roasted or cured way from the outer layer of fruit. Betel nut is most commonly cut and roll into thin bands in a betel leaf with slacked lime (powder) or broken seashells. The betel quid, betel noodle chew, betel cherry, betel pan or betel paan are known as the box. We seek to define the image using two different methods in this project design. Use the Extraction Technical for digital image processing the first method listed below for image classification.

The k-nearest neighbors (KNN) algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems.

****

Figure a: Bette Figure b: Gorabulu

****

Figure c: Hasa Figure d: Ide

Figure 2: Images of Dry Arecanut

Betel nut's medicinal use is limited, and long-term negative reactions to betel quid chewing are well known.

We are mainly focusing on sorting of boiled arecanut alias dry arecanut, the farmers basically sort this dry arecanut in four different types namely hasa, bette, ide and gorabalu as shown in Figure 2. Mainly farmers are getting delayed in sending their goods to the market because of this delay in the sorting of this arecanut. It takes more time and manpower to complete the task in the current scenario. The main con of designing these kinds of projects is reducing the time constraint and object detection. There is no machine that can detect the object every time with 100 percent accuracy. The advantage of this project is that once we are able to achieve the main goal then this same thing can be used to sort different kinds of nuts or seeds.

The automatic classification of areca nut is the primary classification system for four types of areca nut. The aim of the Layout is to identify the arecanut and to reduce the time and work of farmers to separate the arecanuts and to increase productivity and efficiency of segregation.

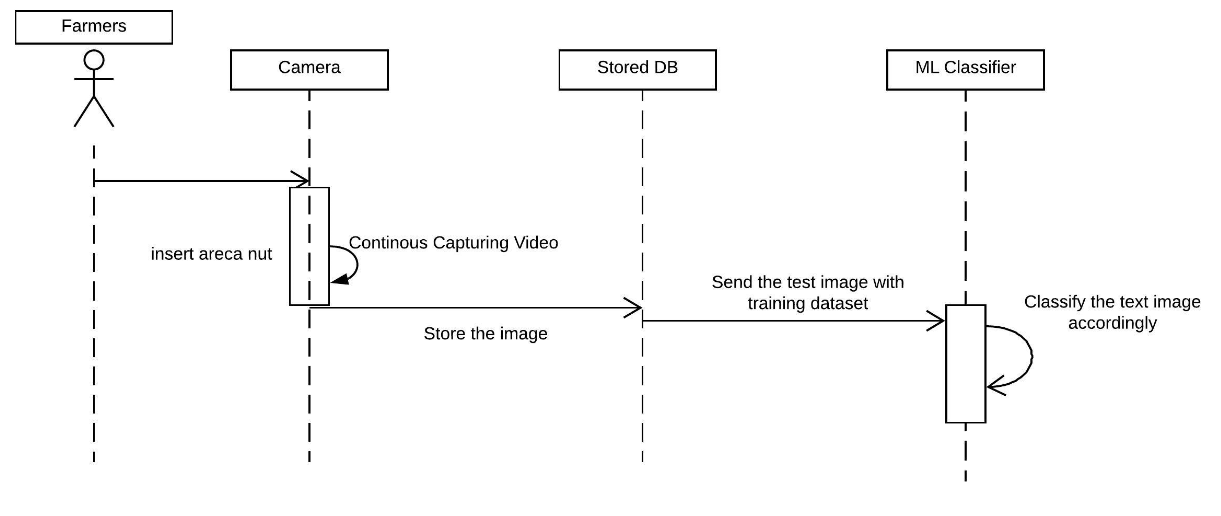
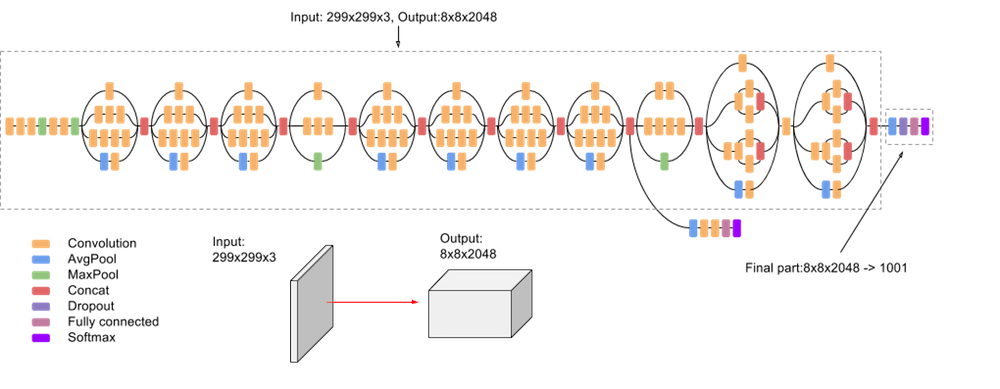


Figure 3: Sequence diagram for arecanut sorting software side

The above Figure 3 is the the sequence diagram for the project which shows the workflow on the software side of our project. First the actor (the farmer) needs to put the arecanut into the machine for the sorting process. The areca nut is taken into the machine by conveyor belts inside the machine a camera is placed in an angle where all the areca nut coming through the conveyor belt can be captured. The capturing of the image is a continuous process a live video loop is much necessary for the smoother transition for camera to capture the raw images. these video part of the frame where areca nut is visible is taken for the processing.

The frame of the image taken is stored in the database for the future training of the machine learning algorithm for gaining more knowledge and higher accuracy in classifying the test data whenever it appears. The second method used for feature extraction and classification of the image we used Inception version 3 to extract the image features and classify the image accordingly.

Figure 4: Model of Inception V3 by google [7]



Inception v3 is a widely-used image recognition model that has been shown to attain greater than 78.1% accuracy on the ImageNet dataset. The model is the culmination of many ideas developed by multiple researchers over the years. The idea as shown in the Figure 4 and working simultaneously.

After the classification output from the classifier based on the result whether the classifier is able to classify the image or not. The output of the classifier is sent to sorting machine for giving the hardware signals to do the task. When the message enters the sorting machine decodes and analyzes the data and transfers the signals to the hardware device for the operation. The sorting machine waits for the message. In accordance with the signals produced to move them to their final destinations, the arecanut is transferred into a specific gunny bag/container. The overall state of the sorted nozzles, cycle time, etc., is to be sent back to the actor (farmer). [8]

**Image Acquisition**: Picture Processing consists of producing a digitally coded image of the visual features of the object, for example a physical scene or an object's interior design.

**Image Pre-processing:** For image operations at the lowest level of abstraction, preprocessing is the common name–both input and output are intensity images. The purpose of pre-processing is to enhance image data to remove unwanted defects or to improve certain image features that are necessary for further processing

**Feature Extraction:** As the data for the classifier, the extracted features. Colour, texture, shape and size were the various features extracted for subsequent classification. As color analyzes do not suggest RGB, HSV color moments including mean, standard deviation and skew are omitted as most of the color information is included in those 3 moments. Of these 14 characteristics, five are considered most important to this proposed method. Energy, comparison, correlation and homogeneity and entropy are these characteristics.[9]

**Algorithm for Feature Extraction:**

Step 1: Decompose input image using 2-D Wavelet or Contourlet transform after converting RGB image into grayscale.

Step 2: Derive Co-occurrence matrices for high frequency sub bands of wavelet or contourlet with 1 for distance and 0; 45; 90 and 135 degrees for θ and averaged.

Step 3: From these co-occurrence matrices, five Haralick texture features called Co-occurrence Texture features are extracted. The size feature was calculated by counting the number of non-zero pixels in the segmented image and then normalizing it with the total number of pixels. In order to extract the shape features, Fourier coefficient method has been used. This method involves the following steps to estimate the shape feature.

1. Estimate the outermost boundary points of the areca nut, let N be the total number of pixels.

2. Determine the centroid (xc, yc) of the areca nut.

3. Find Euclidean distance R (k) from each boundary point (xk, yk) to the centroid.

4. Discrete Fourier Transform is applied to R (k), resulting one dimensional feature vector of the areca nut. Only the first few coefficients are distinct and can be used to distinguish the difference between areca nut shapes. Classification is the final stage in the areca nut grading process. Support Vector Machine (SVM) is a powerful binary supervised classifier and accurate learning technique. SVM can be used for classification or regression challenges. It is very suitable for nonlinear classification. Here the basic idea is to map feature vectors nonlinearly to another space and learn a linear classifier. The linear classifier in new space would be an appropriate nonlinear in the original space. Kernel functions effectively map the original feature vectors into higher dimensional space without explicit calculation.[10]

# 4.Conclusion

# With the support of various algorithms, the recent approaches segment the arecanut. In the segmentation of the Arecanut, as stated in previous articles, different models are implemented, such as Support Vector Machine (SVM), where the input is handmade and the output is not so exact. In deep learning, such as the Convolutionary Neural Network (CNN), lower level features are automatically extracted to higher levels which yield good results. The definition defined in the article says that low-level processing to a high level of understanding dictates that the object takes very little time but the quality of the image is decreased so that we either lose something or have to gain a speed. This high level of knowledge involves the use of neural networks and principles of deep learning to solve this problem.

# References

1. Pramank, A. K. “ Accounting and Management: In Theory and Practice ” Year of Publication:2004. [Online].Available: https://books.google.co.in/books?id=P7S1dbrbfwUC&pg=PA45&redir\_esc=y#v=onepage&q&f=false (Last accessed:05-03-2020)
2. "Commodity: Areca Nut". crnindia.com. Retrieved 31 May2016.[Online].Available:http://crnindia.com/commodity/arecanut.html (Last accessed: 05-03-2020)
3. Young Jin Heo, Se Jin Kim, Dayeon Kim, Keondo Lee, and Wan Kyun Chung “Super-High-Purity Seed Sorter using Low-latency Image-Recognition based on Deep Learning” *IEEE Robotics and Automation Letters*. Vol3,Issue4,October2018.[Online].Available:https://doi.org/10.1109/LRA.2018.284951
4. Edoardo Pasolli, Farid Melgani, Yakoub Bazi “Support Vector Machine Active Learning Through Signiﬁcance Space Construction” *IEEE Geoscience and Remote Sensing Letters*, Vol. 8, Issue 3, May 2011. [Online].Available:https://doi.org/10.1109/LGRS.2010.2083630
5. Bing Li, Weihua Xiong, Ou Wu, Weiming Hu, Stephen Maybank, Shuicheng Yan “Horror Image Recognition Based on Context-Aware Multi-Instance Learning” *IEEE Transactions on Image Processing*. Vol24,Issue12,December2015.[Online].Available:https://doi.org/10.1109/TIP.2015.2479400
6. Qian Wang, Yinghuan Shi, Dinggang Shen “Machine Learning in Medical Imaging” *IEEE Journal of Biomedical and Health Informatics*, Vol 23, No 4, July-2019.[Online].Available:https://doi.org/10.1109/JBHI.2019.2920801
7. Annegreet Van Opbroek, M Arfan Ikram, Meike W Vernooij, Marleen de Bruijne “Transfer Learning Improves Supervised Image Segmentation Across Imaging Protocols” *IEEE Transactions on Medical Imaging*. Vol 34, Issue: 5, May 2015. [Online]. Available:https://doi.org/10.1109/TMI.2014.2366792
8. Qian Wang, Yinghuan Shi, Dinggang Shen “Machine Learning in Medical Imaging” *IEEE Journal of Biomedical and Health Informatics*, Vol 23, No 4, July2019.[Online].Available:https://doi.org/10.1109/JBHI.23
9. Young Jin Heo, Se Jin Kim, Dayeon Kim, Keondo Lee, and Wan Kyun Chung “Super-High-Purity Seed Sorter using Low-latency Image-Recognition based on Deep Learning” *IEEE Robotics and Automation Letters*.Vol3,Issue4,October2018.[Online].Available:https://doi.org/10.1109/LRA.2018.2849513
10. Annegreet Van Opbroek, M Arfan Ikram, Meike W Vernooij, Marleen de Bruijne “Transfer Learning Improves Supervised Image Segmentation Across Imaging Protocols” *IEEE Transactions on Medical Imaging*.Vol34, Issue:5,May2015.[Online].Available:https://doi.org/10.1109/TMI.2014.2366792