J. Electrical Systems 20-11s (2024): 4684-4693

Umesha D. K[1]
Venkata Krishna J[2]
Dhanesh R[3]
Ravithej S V. [4]
Varalaxmi Adimurthy[5]

Segmentation of Arecanut Bunches Using Deep Learning Technique

Journal of Electrical Systems

Abstract: Image segmentation plays a crucial role in computer vision, enabling the extraction of detailed features from images for a wide range of applications. Convolutional Neural Networks (CNNs) have demonstrated exceptional performance in achieving accurate segmentation outcomes across various domains. This study presents a novel method for segmenting unharvested arecanut bunch images using CNN models. An optimized U-Net model was utilized for segmentation, and a comparative analysis was conducted across three different color spaces: RGB, saturation, and grayscale. The results indicate that the proposed model performs best with RGB images, achieving a Dice coefficient of 91.15%, which is notably higher compared to the segmentation of images in the other two color spaces. This research underscores the superior accuracy of using RGB images for the segmentation of arecanut bunches, offering valuable insights for applications in precision agriculture.

***Index Terms*--** Arecanut, U-Net, Segmentation, Convolution Neural Network.

I. INTRODUCTION

Agriculture is a key driver of economic growth in many developing nations, including India, where it plays a significant role in contributing to the country's Gross Domestic Product (GDP) [1]. Among the various commercial crops cultivated in India, arecanut (Areca catechu), commonly known as betel nut, stands out due to its extensive use in a range of products, including medicinal preparations, tea powders, and soaps[2]. This versatility has established arecanut as one of India’s most important commercial crops, crucial for both local economies and broader economic growth.

In agricultural image processing, segmentation is a vital technique that improves the precision and effectiveness of image-based analyses. In arecanut cultivation, the use of Convolutional Neural Networks (CNNs) for segmentation is especially important. This process involves dividing an image into meaningful regions to emphasize key objects or features while separating them from irrelevant background elements [3]**.** CNN models have revolutionized segmentation tasks by providing precise and efficient extraction of target objects from images [4]. In the context of arecanut farming, this capability is particularly beneficial. For instance, segmented images enhance disease prediction a model by offering clearer views of the arecanut bunches, which improves the accuracy of disease detection and management. Additionally, accurate segmentation facilitates more effective maturity analysis, allowing for a detailed examination of the arecanut bunches’ ripeness without interference from background noise [5]**.** Furthermore, segmentation supports the

[1] Research Scholar, Department of CSE, Srinivas University Institute of Engineering and Technology, Srinivas University, Mukka, Managalore,
Associate Professor, Department of CSE, Srinivas University Institute of Engineering and Technology, Srinivas University, Mukka, Managalore.
Department of Studies in Computer Science, Davangere University, Davangere
Lecturer, Department of Computer Science & Engineering, Government Polytechnic, Chamarajnagar
Assistant Professor, Govt. first grade College, Bengaluru.

grading and sorting of arecanut images by isolating individual bunches [6], which is essential for assessing their quality and sorting them according to specific criteria. This process is integral to automated classification systems, where CNN models can classify segmented arecanut bunches into various categories, such as different maturity levels or quality grades. Overall, the application of CNN-based segmentation to remove unwanted backgrounds from arecanut images not only enhances the quality and accuracy of disease prediction, maturity analysis, and grading but also contributes to the efficiency and effectiveness of agricultural practices. This technological advancement paves the way for more sophisticated and automated systems in agriculture, driving progress in precision farming and quality control.

This paper proposes a method that automatically segments the arecanut bunches using an Optimized U-Net model. The primary focus of this work is on accurately segmenting the arecanut bunches. The major contribution of the proposed work is the use of an Optimized U-Net model for segmentation of arecanut bunches and comparative analysis was conducted across three different color spaces: RGB, saturation, and grayscale images.

Section 2 provides an overview of related works on arecanut segmentation. Section 3 outlines the materials and methods utilized in this study. Section 4 discusses the experiments conducted, while section 5 presents the study's results. Finally, section 6 draws conclusions and offers recommendations for future research directions.

II. Related Work

In recent years, there has been increasing interest in utilizing image processing and deep learning techniques to enhance agricultural practices, particularly in automating tasks such as crop grading, sorting, and segmentation. Image-based systems offer promising solutions for agricultural challenges, such as arecanut maturity identification, by facilitating precise segmentation in complex environments. Arecanut (*Areca catechu*), a key commercial crop in India, requires accurate methods for assessing maturity to enhance yield quality and optimize market value.

Several studies have explored various image processing methods for segmenting arecanut bunches. One of the earliest works focused on segmentation and classification using statistical methods [7]. In this research, three-sigma control limits are applied to segment raw arecanut images based on their size and shape characteristics. This statistical approach was effective in identifying and categorizing arecanuts, providing a quantitative framework for segmenting images with high precision, and laying the groundwork for future image processing techniques in this domain. In the work on segmenting defective regions in fruits and vegetables, [8] developed an image processing technique that effectively isolates regions of interest based on defects. This method, which utilizes color, texture, and shape analysis, was designed to enhance postharvest quality control through automated systems. Their approach is particularly relevant to agricultural applications requiring accurate segmentation in dynamic environments, such as arecanut maturity identification.

Building on these earlier methods, HSV (Hue, Saturation, Value) color model is used to improve segmentation outcomes for arecanut bunches [9]. This research demonstrated how the HSV model’s separation of chromatic content from intensity allowed for more effective segmentation, particularly when dealing with complex agricultural backgrounds. The ability of the HSV model to distinguish arecanut bunches from surrounding noise significantly enhanced image clarity, making it a useful tool in agricultural image processing. In further developments, [10] investigated the YCgCr color model, emphasizing its robustness in varying lighting conditions. The study revealed that separating luminance (Y) from chrominance (Cg and Cr) improved segmentation accuracy by reducing background interference. This model demonstrated superior performance in comparison to traditional RGB methods, providing a more reliable solution for arecanut segmentation in diverse environmental conditions. The comparative study in [11] systematically evaluated the performance of RGB, HSV, and YCgCr color models. Their findings showed that while the commonly used RGB model struggled with background interference, both HSV and YCgCr models produced better results by separating color components more effectively. The YCgCr model, in particular, was noted for its ability to isolate arecanut bunches with higher accuracy, reducing background clutter and improving overall segmentation outcomes.

Recently, deep learning methods have gained significant attention for their effectiveness in addressing complex segmentation tasks. Convolutional Neural Networks (CNNs), in particular, have been employed for segmenting arecanut bunches, demonstrating their potential in agricultural image processing applications [12]. This study demonstrated the effectiveness of CNN-based models like U-Net in accurately segmenting arecanut bunches, even in challenging lighting and background conditions. These models significantly outperformed traditional image processing methods, offering enhanced segmentation precision and robustness. [13] Developed a deep learning model for detecting and segmenting obscured green fruits, addressing challenges like occlusion and color similarity with the background. The model showed high segmentation accuracy, making it relevant for agricultural tasks, such as arecanut bunch segmentation, where complex environments and lighting variations hinder accuracy. This model holds promise for broader applications in automated fruit harvesting and yield estimation.

Despite these advancements, there remains a need for lightweight models that can perform accurate segmentation in real-time while managing the complexities of agricultural environments. Although deep learning techniques have shown promise, optimizing these models for real-time application in dynamic field conditions remains an ongoing challenge. Further research is required to develop models that balance accuracy, efficiency, and computational cost, offering farmers practical solutions for arecanut maturity identification and other agricultural applications.

In summary, although substantial progress has been made in arecanut segmentation, there is still potential for further advancements. Traditional color models like HSV and YCgCr have contributed valuable insights, but deep learning models present a promising solution for overcoming the challenges encountered in real-world agricultural settings. Future efforts should focus on optimizing these models for enhanced accuracy and real-time performance, ultimately enabling automated systems that can assist farmers and improve crop management techniques.

III. MATERIALS AND METHODS

In the proposed work, Arecanut images are collected and pre-processed as per the proposed models requirements. The pre-processed images are then segmented using proposed models. Details of the proposed methodology are described below.

2.1. Dataset Collection

The arecanut dataset available in [14] is utilized in this study to assess and validate the proposed method. This dataset consists of 1,017 images of arecanut bunches including 629 unripened and 388 ripened bunches, providing a comprehensive basis for analysis. These images were captured using an OPPO F3 mobile phone, equipped with a 16-megapixel rear camera, Octa-core processor, 4GB RAM, and the Android operating system. The phone was mounted on a selfie stick that extended from 5.31 inches to 28.7 inches. Photographs were taken between 9 AM and 1 PM, which coincides with the typical harvesting period, maintaining a distance of approximately 50 cm between the camera and the arecanut bunches. The phone was positioned at a 45-degree angle using the stick. The plants chosen for image capture were between 7 and 9 years old and stood approximately 12 to 14 feet tall. To assess the maturity level of the arecanut bunches, an expert's evaluation was sought. All images were resized to 256x256 pixels and saved in JPEG format.

2.2. Image Pre-processing

*E*ach images in the dataset are resized in to 224 × 224 pixels for the efficient computations to meet memory requirements. The arecanut maturity levels are associated with the color of the fruit. In this work, the resized RGB images are converted into HSV images and the saturation component channel is considered for the input to the deep learning models. The saturation component of an image indicates its color intensity relative to brightness [15]. This component can be derived from an RGB image using Equation 1, as referenced in.

$$Sat = 1 - \frac{3}{(P1 + P2 + P3)}[min(P1, P2, P3)] \tag{1}$$

where P1, P2 and P3 are primary color components like blue, red and green respectively.

To compare the effect of color space on segmentation grayscale arecanut images are obtained from the RGB images and stored separately for further analysis. Figure 1 show sample images of 3 different datasets used in this proposed work.

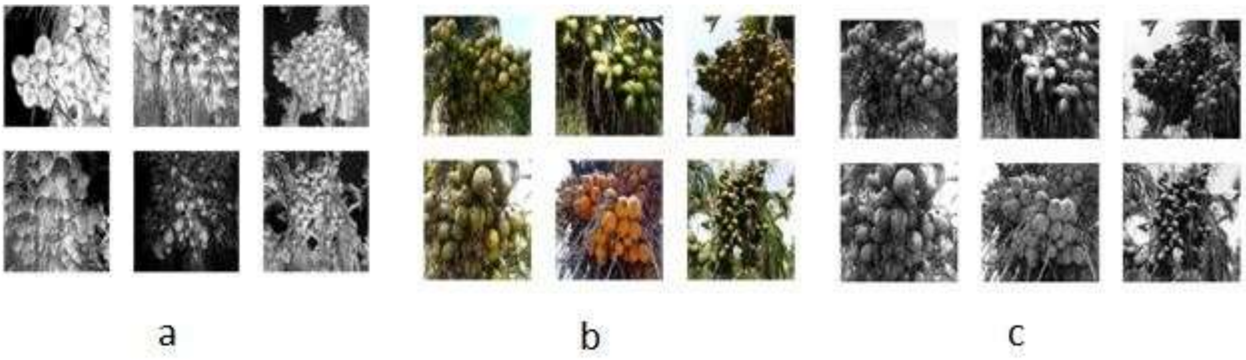

Figure 1: Sample images in the used dataset.

(a) Saturation component images b) RGB images (c) Grayscale images

These pre-processed arecanut images contains unwanted background which makes it very difficult for other image analysis like maturity detection, disease prediction, grading, sorting and classification of the arecanut bunches. To overcome this issue in the proposed work the arecanut bunch images are segmented to remove unwanted background in the images.

2.3. Segmentation: Optimized U-Net Model

In this proposed work original U-Net architecture [16] is customized to segment the arecanut bunch images which remove the undesired backdrop in the image. Figure 2 depicts the proposed segmentation model. The architecture of the model features a "U" shape, with the contracting path on the left and the expansive path on the right. Open boxes in the figure represent feature maps, with the number of channels in each map indicated at the top. The shaded box on the contracting side denotes the input layer, while the shaded box on the expansive side indicates the copied feature map from the corresponding contracting stage. Arrows between the boxes illustrate the various operations performed.

The contracting path functions as a conventional convolutional network, comprising three stages. Each stage on the left side of the U-Net (contracting path) includes two 3×3 convolutions activated by the rectified linear unit (ReLU) and a 2×2 max pooling layer with a stride of 2 for down-sampling. After each stage in the contracting path, the number of feature maps is doubled.

Conversely, the expansive path on the right side consists of stages that involve an up-sampling operation using a 2×2 convolution, followed by a concatenation with the corresponding features from the contracting path, and two 3×3 convolutions with ReLU activation. The number of stages in the expansive path matches that of the contracting path. A 1×1 convolution is applied in the final layer to produce the desired number of classifications.

The proposed model accepts input images sized $224 \times 224 \times 3$ and outputs a segmented binary mask of size

$224 \times 224 \times 1$. The segmented images corresponding to the input can be generated by multiplying the original unsegmented image by the binary mask produced by the model. Table 1 outlines the primary differences between the proposed optimized U-Net model and original U-Net model [16].

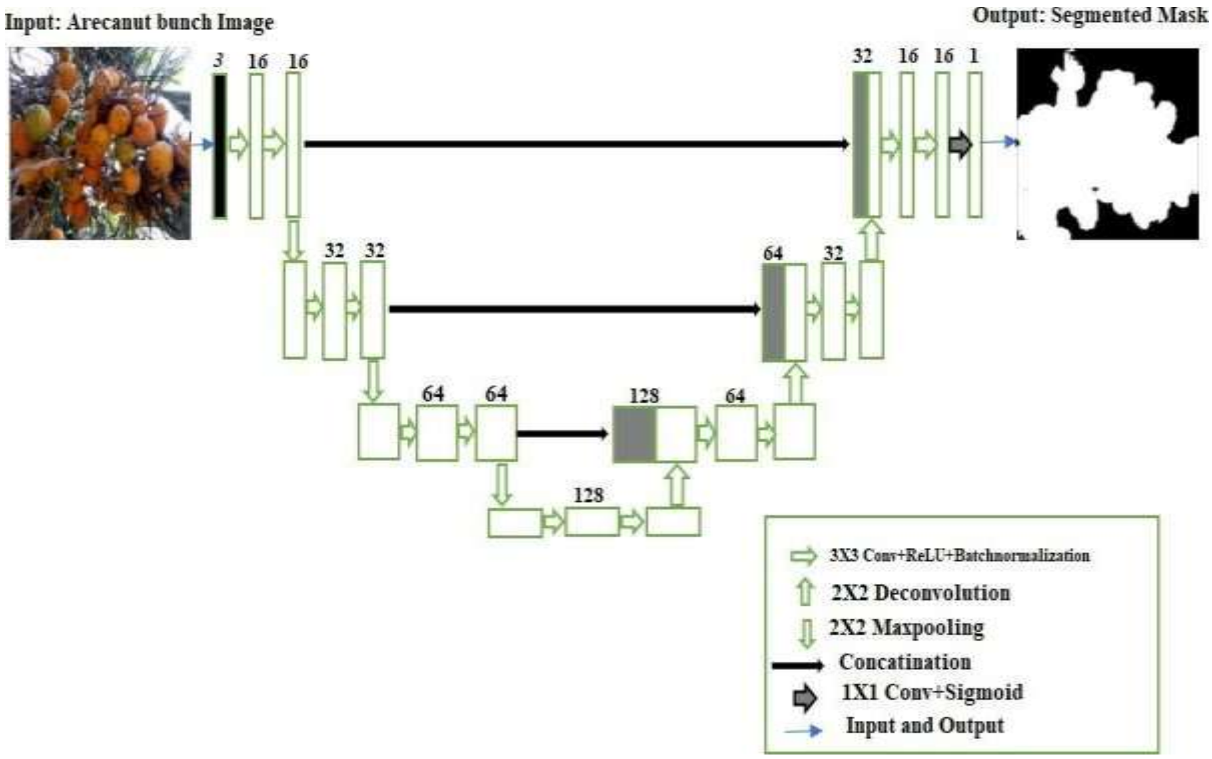

Figure 2: Proposed Optimized U-Net Architecture

Characteristic	**Original U-Net**	**Proposed Optimized U-net**
Depth levels within both the contracting and expansive path	4	3
Each Depth level in contracting path consists of	Two 3×3 convolutions followed by ReLU activation and a 2×2 max pooling layer.	Two 3×3 convolutions with ReLU activation, Batch Normalization and a 2×2 max pooling layer."
Number of feature channels in each depth level	64,128,256,512	16,32,64
Padding in convolution operation	Padding is not applied, which leads to a reduction in the size of the feature maps after each convolution operation.	Padding is applied to maintain the same size of the feature maps before and after each convolution operation.
Input and Output size	Input: 572*572*1 Output: 388*338*2	Input: 224*224*3 Output 224*224*1
Data augmentation during the model training	Used	Not Used

Table:1 Difference between proposed U-Net and original U-Net .

IV. EXPERIMENTAL SETUP

The proposed model was implemented using Keras (version-2.4.0) [17] framework with TensorFlow (version-2.4.1) [18]. Python is the language used for implementation because it not only comes with a large set of libraries that can be utilized for Machine Learning, but it is also easily accessible.

U-Net Model Training:

To enhance the performance of the model, it is essential to configure hyper-parameters effectively. The Random Search algorithm [19] is employed to identify the optimal hyper-parameter settings, which are then applied to the optimized U-Net model for segmenting arecanut bunch images. The hyper-parameters utilized during the training of the optimized U-Net model are presented in Table 2.

The model accepts input images sized 224 × 224 × 3 and generates a segmented output mask of size 224 × 224 × 1. To accelerate the training process, Batch Normalization [20] is implemented, and a dropout rate of 0.05 is applied during both down-sampling and up-sampling to mitigate overfitting. The model is compiled using the Dice loss function, which evaluates the overlap between the predicted and actual samples, calculated via Equation 2. The Adam optimizer [21] is utilized with a learning rate of 0.006 for updating the model's weights.

In this study, RGB images, saturation images and grayscale images are segmented using the proposed U-Net model. Each dataset, consisting of 1,017 images, is divided into training, validation, and test datasets for model evaluation, as detailed in Table 3. The training dataset serves as input, while the corresponding ground truth masks act as labels. Figure 3 illustrates the training curves of the proposed U-Net model applied to RGB images over 40 epochs. The graph indicates that the model is effectively trained with minimal overfitting. The segmented output images generated by the proposed model are stored separately for future reference.

Hyper-parameter	Number of Depth levels	Learning Rate	Dropout	Optimizer	Batch size
Values	3	0.006	0.05	Adam	16

Table 2. Hyper-parameter setting for the proposed U-Net model.

Table 3: Description of training, validation and test datasets. (The numbers represents the number of images in the different dataset).

	Training dataset	Validation dataset	Test dataset	Total
Number of images	732	183	102	1017

Table 3: Number of images in training, validation and test datasets.

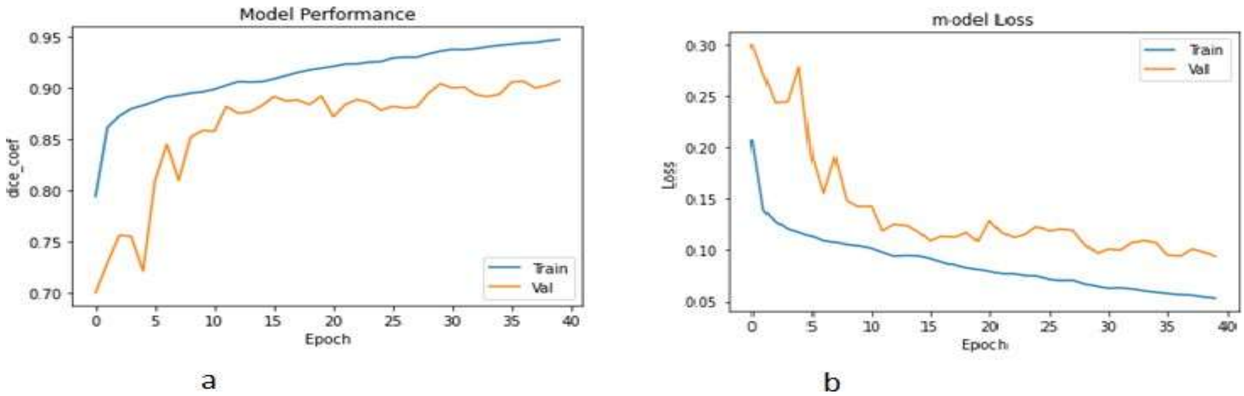

Figure 3: Training graphs for proposed *U− Net* model on RGB Images (a) Dice Coefficient versus Epoch (b) Dice loss versus Epoch

V. RESULTS AND DISCUSSIONS

Following the training phase, the model's performance is assessed by comparing the outputs generated from the test image dataset with the expert-graded ground truth images. In this study, the evaluation metrics used include the Dice Coefficient (DC), Recall, and Precision. The Dice Coefficient quantifies the degree of correspondence between the ground truth and the predicted output. In contrast, Recall indicates the model's capability to recognize positive samples effectively. Precision, on the other hand, calculates the proportion of true positive samples relative to all predicted positive samples. Together, these metrics provide a robust evaluation of the model's performance. The calculations for these metrics are outlined as follows:

$$\text{Dice coefficient} = \frac{2*TP}{(2*TP)+FP+FN} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

Where TP, TN, FP, and FN represent True Positive, True Negative, False Positive, and False Negative, respectively.

Comparison of different metrics of the proposed segmentation model on 3 different test dataset and other previous work has been tabulated i n Table 4. It is observed from the results that segmentation is more accurate on RGB images using proposed method and segmentation using proposed work is better than other previous work on arecanut images.

Authors	**Method used**	**Arecanut Image Used**	**Metrics**			
			DC (%)	IOC (%)	Precision (%)	Recall (%)
Dhanesha R, et al [11]	YCgCr Colour threshold	Immature	79.93	----	----	----
		Mature	83.24	----	----	----
		Over Mature	87.69	----	----	----

Anitha A.C., et al. [21]	U-Net	Unripe areca	----	58.07	74.71	77.15
		Ripe Areca	----	54.61	61.53	87.07
Anitha A.C. et al. [21]	Mask R-CNN	Unripe areca	----	65.98	89.86	73.14
		Ripe Areca	----	61.01	73.57	81.84
Proposed Work	Proposed UNet Model (Considered both Ripe and Unripe areca images for segmentation)	**RGB images**	**91.46**	**-----**	**93.30**	**76.50**
		Saturation images	90.30	-----	91.85	77.72
		Grayscale Images	89.64	-----	92.21	76.40

Table 4: Comparison of other previous work with proposed segmentation work.

Qualitative results of the proposed segmentation method on three type of images are shown in Figure 5 , first column shows the original image taken from Areca field, second column shows manually created Ground Truth image, third column shows segmented binary mask obtained from segmentation model and the fourth column shows Segmented RGB image which is produced by multiplying the binary mask with original images.

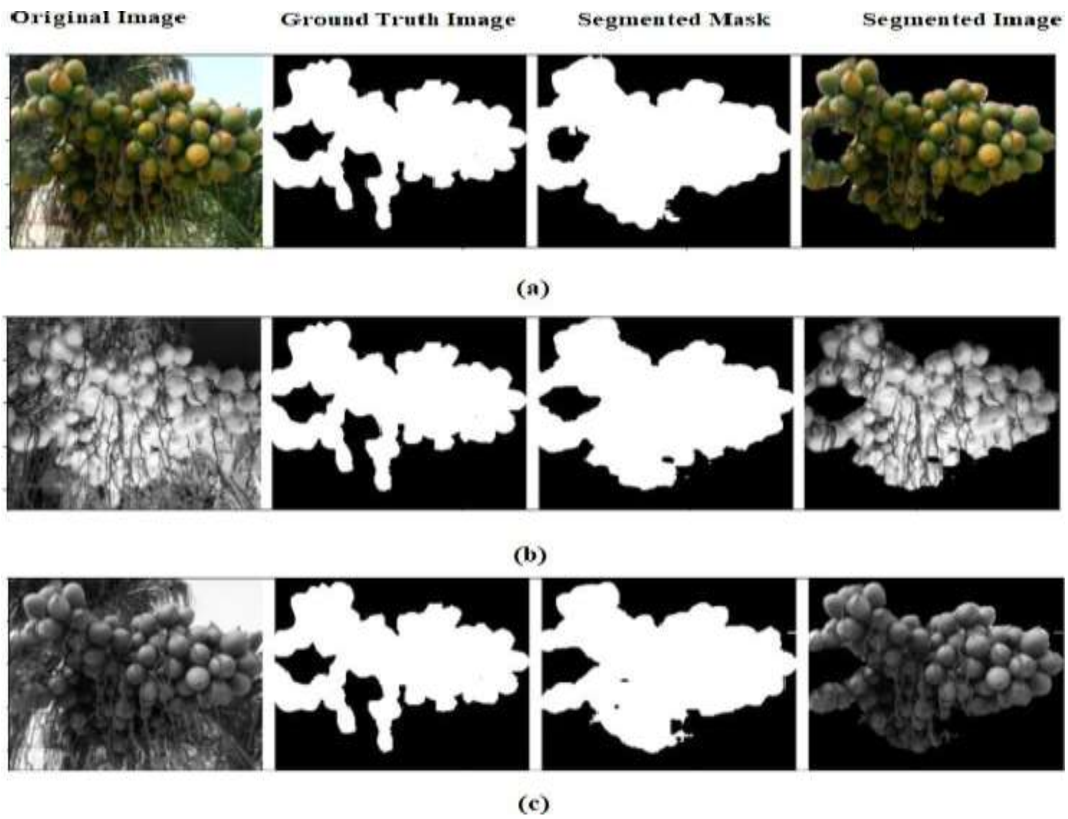

Figure 5: Results of proposed segmentation method (a)RGB Images, (b)Saturation Image and (c) Grayscale Image.

VI. CONCLUSION

Arecanut bunch segmentation is a pivotal task in precision agriculture, crucial for accurately analyzing and managing arecanut crops. In this study, we introduced a novel approach for segmenting unharvested arecanut bunch images using an optimized U-Net model. We evaluated the performance of the segmentation across three different color spaces: RGB, saturation, and grayscale.

Results of this work indicates that the optimized U-Net model performs best with RGB images, achieving a Dice coefficient of 91.15%, which is significantly higher than the results obtained with saturation, grayscale images and also segmentation results of the other previous work. This demonstrates that RGB images provide the most accurate segmentation of arecanut bunches using proposed model. The superior performance of the RGB-based model underscores its effectiveness and potential for improving precision agriculture practices. This research offers valuable insights for more accurate arecanut bunch analysis and better decision-making in crop management.

REFERENCES

[1] Seth, D., & Shah, M. (2016). The role of agriculture in economic development: Indian context. *Journal of Commerce and Management Thought, 7*(2), 201-215. https://doi.org/10.5958/0976-478X.2016.00012.0

[2] Gupta, S. K., et al. (2023). Areca catechu L. (Betel nut): A review on its nutritional, medicinal, and toxicological aspects. *Journal of Food Science and Technology, 50*(1), 123-135.

[3] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2022). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44*(7), 3523-3542. https://doi.org/10.1109/TPAMI.2021.3059968

[4] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3431-3440).

[5] Dhanesha, R., & Shrinivasa Naika, C. L. (2019). A novel approach for segmentation of arecanut bunches using active contouring. In *Integrated Intelligent Computing, Communication and Security* (pp. 677-682).

[6] Zhang, X., Zhang, Y., & Li, Y. (2022). Deep learning-based grading and sorting of agricultural products: A review. *Computers and Electronics in Agriculture, 196*, 107015.

[7] Danti, A., & S. M. (2012). Segmentation and classification of raw arecanuts based on three-sigma control limits. *Elsevier Journal of C3IT-2012, 4*, 215-219. https://doi.org/10.1016/j.protcy.2012.05.032

[8] Rozario, L. J., Rahman, T., & Uddin, M. S. (2016). Segmentation of the region of defects in fruits and vegetables. *International Journal of Computer Science and Information Security, 14*(5), 399.

[9] Dhanesha, R., & Shrinivasa, N. C. L. (2018). Segmentation of arecanut bunches using HSV color model. In *Proceedings of the 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)* (pp. 37–41). https://doi.org/10.1109/ICEECCOT43722.2018.9001632

[10] Dhanesha, R., Shrinivasa Naika, C. L., & Kantharaj, Y. (2019). Segmentation of arecanut bunches using YCgCr color model. In *2019 1st International Conference on Advances in Information Technology (ICAIT)* (pp. 50-53). https://doi.org/10.1109/ICAIT47043.2019.8987431

[11] Dhanesha, R., Umesha, D. K., Shrinivasa Naika, C. L., & Girish, G. N. (2021). Segmentation of arecanut bunches: A comparative study of different color models. In *2021 IEEE Mysore Sub Section International Conference (MysuruCon)* (pp. 752-758). https://doi.org/10.1109/MysuruCon52639.2021.9641680

[12] Anitha, A. C., Dhanesha, R., Shrinivasa, C. L., Krishna, A. N., Kumar, P. S., & Sharma, P. P. (2022). Arecanut bunch segmentation using deep learning techniques. *International Journal of Circuits, Systems and Signal Processing, 16*, 1064-1073.

[13] Liu, M., Jia, W., Wang, Z., Niu, Y., Yang, X., & Ruan, C. (2022). An accurate detection and segmentation model of obscured green fruits. *Computers and Electronics in Agriculture, 197*, 106984. https://doi.org/10.1016/j.compag.2022.106984

[14] Davangere University. (n.d.). Arecanut database. http://davangereuniversity.ac.in/arecanut-database/

[15] Fairchild, M. (1998). *Color appearance models*. Reading.

[16] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_2823

[17] Chollet, F. (2015). Keras: A deep learning library for Theano and TensorFlow. *arXiv preprint arXiv:1502.06975*.

[18] Abadi, M., Barham, P., Brevdo, E., et al. (2016). TensorFlow: Large-scale machine learning on heterogeneous systems. *arXiv preprint arXiv:1605.07603*.

[19] Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research, 13*(2).

[20] Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift.

[21] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization.