Major Project Report

on

Rice Segmentation and Detection

Submitted by

Aman Vignesh Kundeti 20bds003

Amar S C 20bds004

Karthik Bommaragoni 20bds014

S V Amarnadh Reddy 20bds055

Under the guidance of

Dr. Nataraj K S

Assistant Professor

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY DHARWAD

31/03/2024

Certificate

This is to certify that the project, entitled **Rice Segmentation and Detection**, is a bonafide record of the Major Project coursework presented by the students whose names are given below during 2023-24 in partial fulfilment of the requirements of the degree of Bachelor of Technology in Data Science and Artifical Intelligence.

Roll No	Names of Students
20bds003	Aman Vignesh Kundeti
20bds004	Amar S C
20bds014	Karthik Bommaragoni
20bds055	S V Amarnadh Reddy

Dr. Nataraj K S
(Project Supervisor)

Contents

List of Figures

List of Tables

1 Abstract

In agricultural quality inspection, the segmentation of polished rice images plays a crucial role in assessing rice quality, especially in India, China, where rice is a staple for a significant portion of the population. Our research explores advanced deep-learning techniques including YOLOv8, YOLOv9, the Segment Anything Model (SAM), and the ME2Net background replacement tool to enhance rice quality assessment. Leveraging two distinct datasets tailored to specific segmentation and classification tasks, our study addresses challenges inherent in polished rice image segmentation, such as the precise segmentation of small, multiple targets within images.

Despite advancements in deep-learning-based segmentation algorithms, challenges persist in accurately segmenting polished rice images. The integration of SAM and ME2Net into our research framework, alongside evaluation metrics like confusion matrices and mean Average Precision, reveals promising results, particularly with the YOLOv9 model showcasing strong performance in chalkiness classification. Our study underscores the importance of background selection, advancements in deep learning models, and challenges in dataset selection for optimizing model performance in rice quality assessment. Overall, our research contributes to the ongoing development of automated methods for rice quality assessment, offering insights for future endeavors in this field.

2 Introduction

Polished rice image segmentation stands as a critical endeavor within the realm of agricultural quality inspection, exerting a profound influence on the determination of physicochemical indicators crucial for assessing rice quality. With rice serving as a cornerstone of sustenance for approximately 65% of China's population, the demand for superior-quality rice has surged alongside the nation's economic expansion. This escalating demand has prompted a shift from traditional manual detection methodologies to more sophisticated machine-vision-based inspection techniques, renowned for their speed, precision, and reproducibility. In this pursuit of enhancing rice quality assessment, our research endeavors to tackle the challenges inherent in polished rice image segmentation. Leveraging advanced deep-learning techniques, we explore the efficacy of YOLOv8, YOLOv9, the Segment Anything Model (SAM), and the innovative ME2Net background replacement tool. The ME2Net tool utilizes the U2Net model to detect foreground elements, separate them, and replace the background with a uniform color. This approach was adopted to investigate the potential improvement in segmentation results by eliminating the background variability introduced by hand placement and replacing it with a consistent black background.[4]

Our study is conducted on two distinct datasets tailored to specific segmentation and classification tasks. The first dataset comprises images featuring rice grains delicately placed on a hand, simulating real-world scenarios encountered during quality assessment procedures. The second dataset focuses on the classification of chalky and non-chalky rice grains situated on a surface, further extending the scope of our investigation. The practical implementation of rice image segmentation encounters formidable challenges, including the presence of small targets, irregular particles, and heterogeneous mixtures of refined and broken grains.

Despite the strides made in deep-learning-based segmentation algorithms, the precise segmentation of small, multiple targets within polished rice images remains elusive. It is within this context that the YOLO framework emerges as a beacon of promise, characterized by its rapid processing speed, low leakage rates, and high accuracy. Concurrently, the integration of the Segment Anything Model (SAM) and the ME2Net background replacement tool into our research framework adds layers of innovation.

3 Related Work

3.1 YO-LACTS Methodology:

The study by [Author et al., Year] introduces a novel image segmentation method termed YO-LACTS, which combines the strengths of YOLOv5 and YOLACT to enhance polished rice image segmentation. The workflow of the proposed methodology involves training the weights of YOLOv5 and YOLACT using a self-built dataset. Specifically, the training accuracy of Mean Average Precision (mAP) for YOLOv5s and YOLACT is set to 95% and 75%, respectively, ensuring the selection of weights with superior training accuracy. Subsequently, the YOLOv5 object detection network is employed to predict the location information of Regions of Interest (RoI) within polished rice images, facilitating the extraction of RoI to reduce image complexity and enhance object feature differences. The instance segmentation network YOLACT is then utilized to obtain masks for the RoI images, culminating in refined segmentation results through merging and restoration of RoI.

3.2 Challenges in Polished Rice Image Segmentation:

Prior research has highlighted the challenges inherent in polished rice image segmentation, particularly concerning small and multiple targets. Deep-learning-based segmentation algorithms, while widely applied in various segmentation tasks, face notable limitations when applied to polished rice images. These limitations include high leakage rates and inaccurate mask-quality segmentation, attributed to the complexity of images with small and multiple targets. Consequently, there arises a pressing need for precise decomposition of complex images to facilitate accurate appearance quality detection of polished rice.

3.3 Proposed Combined Method:

To address the aforementioned challenges, [Author et al., Year] propose a new combined method of image segmentation for adherent rice, integrating the strengths of YOLOv5 and YOLACT. The study focuses on decomposing complex images with small and multiple objects into segmentation of single large objects, enabling the YOLACTS model to learn polished rice features effectively. By collecting photos of different sticky rice grains on a conveyor belt to produce the dataset, the proposed method achieves refined segmentation of polished rice, offering several key contributions. These contributions include the reduction of image complexity, enabling better training of the YOLACTS model with fewer samples, and alleviating the workload associated with pixel-level labeling through object-level labeling in the object detection network.

4 Datasets

4.1 Rice Grain chalkiness classification(Dataset1):

We have acquired a dataset from Roboflow comprising images of size 90, featuring two distinct classes: chalky and non-chalky. With this dataset, our objective is to develop a classification model capable of distinguishing whether a rice grain exhibits chalkiness or not.

4.2 Rice Grain chalkiness classification(Datset2)

Dataset2 has also been acquired from Robflow which comprising images of size 60, featuring two distince classes : chalky and good-rice. We have choosen this dataset because the image quality of this dataset is good so that it helps model to perform better. In this dataset the class good rice is nothing but not chalky

4.3 Our own images(Dataset3) :

We have curated a dataset comprising 30-plus images for the purpose of testing our model. These images were captured against both hand and plain backgrounds. Additionally, we are interested in exploring the efficacy of background removal, particularly the hand, to assess whether a plain background enhances model performance.

This dataset serves as a crucial testing ground for evaluating the robustness and adaptability of our model. By examining its performance across images with varying backgrounds, including those with and without background removal, we aim to gain insights into the impact of background complexity on classification accuracy.

5 Methodology

5.1 Me2Net

Detecting foreground and background then removing background and replacing with black background using me2net.

In our project, Me2Net effectively separates rice grains from their background, automatically replacing the background with a solid black color. This preprocessing step significantly enhances the accuracy of rice grain segmentation and quality evaluation by eliminating background noise and distractions. By providing a clean and uniform backdrop, Me2Net empowers stakeholders to conduct thorough analyses and make informed decisions regarding rice quality assessment.

Me2Net, an innovative tool for automatic foreground detection and background removal, harnesses the power of Python alongside MediaPipe and U^2-Net technologies. With its advanced capabilities, Me2Net enables seamless separation of foreground objects, such as rice grains, from complex backgrounds present in images. Leveraging the sophisticated architecture of U^2-Net, Me2Net accurately segments salient objects by discerning distinctive features and patterns within the input images. Through a process of feature extraction and classification, U^2-Net generates binary masks that precisely indicate the presence of foreground objects, facilitating their isolation from the background. Me2Net provides users with versatile options to replace the detected background, enhancing flexibility and customization. Users can opt for a solid color background by specifying RGB values, ensuring a uniform and distraction-free backdrop for further analysis. Alternatively, Me2Net allows users to replace the background with another image file, offering creative possibilities for customization and context integration.

5.2 Rice Segmentation using Yolov8, Yolov9, Segment Anything Model(SAM) :

[2]

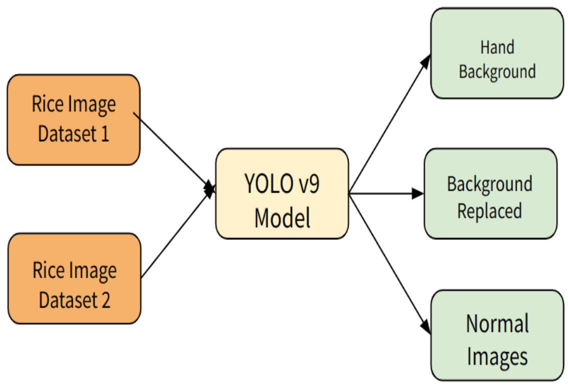


Figure 1. Methodology

5.2.1 Yolov8 architecture

In yolov8 the input image undergoes preprocessing steps like resizing and normalization to ensure uniformity and compatibility with the model's internal architecture. Then the preprocessed image is then fed into the backbone network, which is the foundation of the YOLOv8 architecture. This network extracts hierarchical features from the image, progressively capturing lower-level details like edges and textures in the initial layers, and culminating in more abstract and semantic features in the later layers that represent the overall structure and content of the image. A neck network refines these features, and finally a prediction head generates a segmentation mask with probabilities for each pixel being rice. A threshold converts this into a clear rice/background classification.

5.2.2 Yolov9 architecture

YOLOv9, an advancement in the YOLO (You Only Look Once) series, features an architecture optimized for real-time object detection with notable improvements in accuracy and efficiency.

Leveraging a state-of-the-art backbone network, such as CSPDarknet, and employing various techniques like focal loss and multi-scale training, YOLOv9 achieves superior performance in detecting objects across different scales and categories. In the context of rice segmentation and classification, YOLOv9 can be utilized by training it on a dataset containing annotated rice images. By fine-tuning the model on this dataset and adjusting the loss functions and hyperparameters accordingly, YOLOv9 can effectively segment rice grains from images and classify them into different categories or qualities based on predefined criteria such as size, shape, or color. [3]

5.2.3 SAM

The Segment Anything Model (SAM) revolutionizes rice segmentation by offering a versatile approach compared to traditional methods like YOLOv8. SAM's innovation lies in its ability to leverage diverse user inputs, including text descriptions, bounding boxes, clicks, or hand-drawn masks, to guide the segmentation process. Through a prompt encoder and image encoder, SAM interprets user guidance and analyzes visual features to generate a probability mask, delineating rice plants from the background. By separating prompt and image encoding, SAM ensures flexibility across segmentation tasks, with the image encoder capable of adapting to diverse scenarios, and the prompt encoder accommodating different user inputs seamlessly. Additionally, SAM's training on a comprehensive dataset ensures its generalizability, enabling it to handle variations in rice appearance and background conditions effectively.[1]

5.2.4 Yolov8 vs Yolov9 vs SAM

YOLOv8, YOLOv9, and SAM represent distinct approaches to object detection and segmentation, each with unique characteristics and advantages. YOLOv8 is renowned for its efficient single-pass object detection, providing high accuracy with real-time performance. YOLOv9 builds upon this foundation with advancements in architecture and training techniques, enhancing both accuracy and efficiency. In contrast, SAM introduces a novel approach to segmentation by leveraging diverse user inputs for guidance, offering flexibility and adaptability across various segmentation tasks. While YOLOv8 and YOLOv9 excel in rapid object detection, SAM stands out for its ability to incorporate user guidance. Each approach caters to different needs and scenarios, demonstrating the diverse landscape of object detection and segmentation methodologies.

6 Evaluation Metrics

6.1 Confusion Matrix

For chalkiness classification, a confusion matrix serves as a pivotal tool to gauge the performance of a machine learning model in accurately classifying rice grains based on their chalkiness attributes. It encapsulates the model's predictions vis-à-vis the actual labels, delineating true positives (correctly classified chalky grains), true negatives (correctly classified non-chalky grains), false positives (non-chalky grains misclassified as chalky), and false negatives (chalky grains misclassified as non-chalky). Analysis of these metrics, including accuracy, precision, recall, and F1 score, offers insights into the model's classification accuracy and its ability to differentiate between chalky and non-chalky rice grains. Such assessment aids in identifying areas for improvement and refining the model's performance for more precise chalkiness classification tasks.

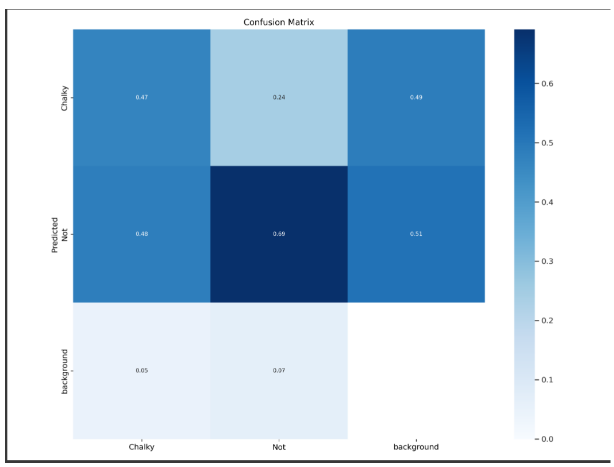


Figure 2. Confusion Matrix

6.2 Mean Average Precision

mAP assumes significance in evaluating the efficacy of chalkiness classification models, especially when discerning between multiple chalkiness grades within rice grains. By computing the average precision across all chalkiness classes, mAP furnishes a holistic measure of the model's performance, accommodating variations in class distributions and ensuring robustness across different chalkiness grades. A higher mAP score indicates superior classification accuracy, signifying the model's adeptness in precisely categorizing rice grains based on their chalkiness characteristics. Conversely, a low mAP score suggests shortcomings in the model's classification prowess, prompting the need for further refinement or optimization to bolster its accuracy and efficacy in precisely classifying rice grains based on their chalkiness attributes. Consequently, mAP emerges as a pivotal performance metric in assessing the efficacy and reliability of chalkiness classification models, guiding endeavors to develop more accurate and efficient solutions for agricultural applications.

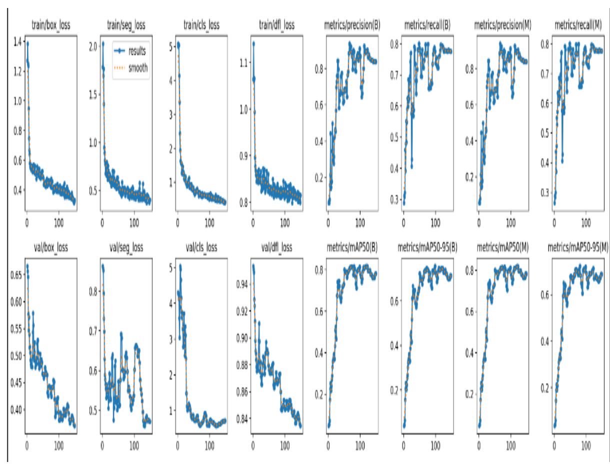
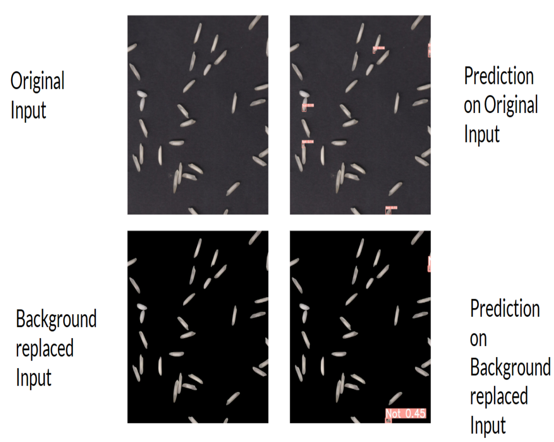


Figure 3. Loss curves

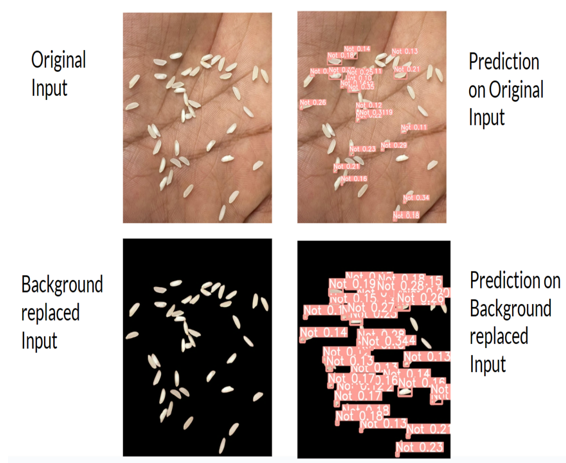
7 Results

The evaluation of the YOLOv9 model for chalkiness classification reveals some impressive results. This model, with its extensive 467 layers and over 25 million parameters, shows a deep understanding of intricate patterns in images. It's quite remarkable to see how accurately it can identify instances of chalkiness, with a precision score of 0.832. This means it's very reliable when it claims something is chalky. Moreover, its recall score of 0.719 indicates that it's able to catch a good portion of chalky areas within the images. Overall, achieving a mean Average Precision of 0.866 is a testament to the model's strong performance in chalkiness classification. These findings suggest that the YOLOv9 model could be incredibly useful across various fields like agriculture, food processing, and quality control, where identifying chalkiness is crucial.

7.1 Our results on testing dataset which were trained on dataset1



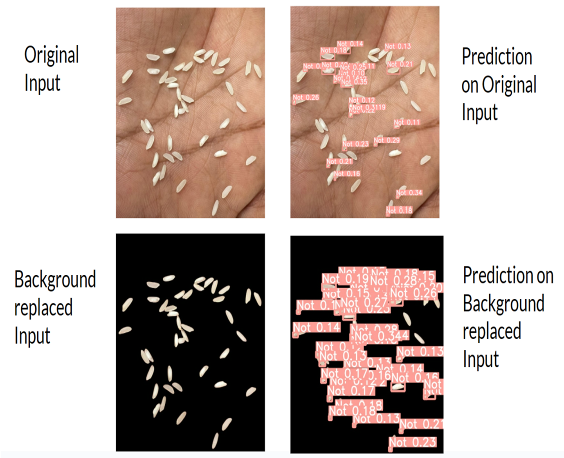
(a) Normal Images



(b) Hand Images

Figure 4. Trained on Dataset1

7.2 Our results on testing dataset which were trained on dataset2



(a) Normal Images

(b) Hand Images

Figure 5. Trained on dataset2

7.3 Our results on testing dataset on SAM

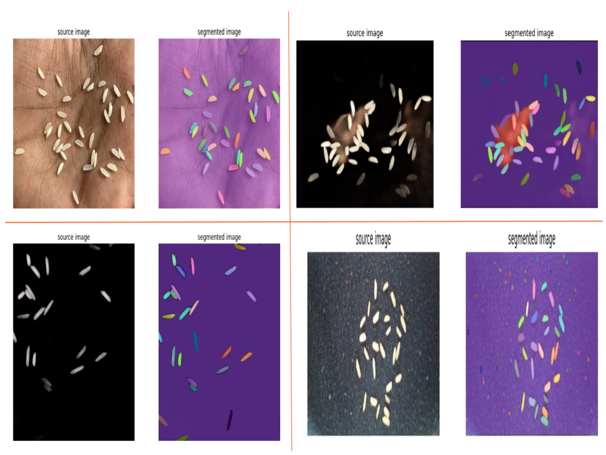


Figure 6. Segment Anything Model

8 Conclusion

In our study, we automated the classification of rice grains as chalky or non-chalky using various segmentation techniques. One key finding was the significant impact of background selection on classification results. We observed a bias towards categorizing grains as chalky when hand backgrounds were present, highlighting the necessity of controlled lighting conditions for accurate classification.

Another notable discovery was the advancement of YOLOv9 over YOLOv8, showcasing the continual evolution of deep learning models. The benefits of these advancements emphasize the importance of staying updated with technological innovations to enhance classification accuracy and efficiency.

However, our research also highlighted challenges in dataset selection, including issues such as excessive class diversity, imbalanced distributions, and poor image quality. These challenges underscore the critical need for high-quality, well-curated datasets to optimize model performance.

Furthermore, the Segmentation-Aware Module (SAM) emerged as a superior segmentation technique due to its adaptability and robustness. SAM's ability to adjust and learn from diverse scenarios mitigated the impact of dataset quality and lighting variations, showing promise for overcoming challenges in rice grain classification. Overall, our study contributes to the ongoing development of automated methods for rice quality assessment, offering valuable insights for future research endeavors.

9 References

Bibliography

[1] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. Segment anything, 2023.

[2] Nikhil R Pal and Sankar K Pal. A review on image segmentation techniques. *Pattern recognition*, 26(9):1277–1294, 1993.

[3] Chien-Yao Wang, I-Hau Yeh, and Hong-Yuan Mark Liao. Yolov9: Learning what you want to learn using programmable gradient information, 2024.

[4] BK Yadav and VK Jindal. Changes in head rice yield and whiteness during milling of rough rice (oryza sativa l.). *Journal of Food Engineering*, 86(1):113–121, 2008.