

## **A Hybrid Approach of Preprocessing and Segmentation Techniques in Automatic Rice Variety Identification System**

**Komal<sup>1\*</sup>, G. K. Sethi<sup>2</sup>, R. K. Bawa<sup>1</sup>**

<sup>1</sup>Department of Computer Science, Punjabi University, Patiala, Punjab, India

<sup>2</sup>Department of Computer Science, M. M. Modi College Patiala, Punjab, India

Received 11 July 2021, accepted in final revised form 2 November 2021

### **Abstract**

Image processing techniques play an important role with various images such as rice grain identification, wheat, fruits, medical, vehicle, and digital text images in image acquisition, image preprocessing, clustering, segmentation, and classification techniques. In the application of object detection and classification images, preprocessing and segmentation techniques are used. This paper delves into the specifics of automated segmentation processes, focusing on rice variety identification and classification images. The aim is to talk about the issues that come up when segmenting digital images and the relative merits and drawbacks of the various methods for preprocessing and segmenting images that are currently available. In this paper, we propose a hybrid approach of Preprocessing and Segmentation techniques to develop an automatic rice variety identification system.

*Keywords:* Thresholding; ANN; Image processing; Rice image; Segmentation.

© 2022 JSR Publications. ISSN: 2070-0237 (Print); 2070-0245 (Online). All rights reserved.

doi: <http://dx.doi.org/10.3329/jsr.v14i1.54811>

J. Sci. Res. **14** (1), 205-213 (2022)

### **1. Introduction**

Digital Image processing is a term used in imaging science to describe the mathematical operations used to improve images or derive useful information from them. The study of image processing techniques for different images has been explored in this paper. One of the image processing techniques used in the agriculture industry to recognize the disease in crop plant images and distinguish varieties of crop plants such as rice and wheat can be used to identify different varieties automatically using automation computer vision [1]. New advancements and developments in computer technologies are contributing immensely to rice variety identification, classification, quality inspection, and rice quality grading. Digital images help recognize and classify defects in crop plants, such as rice, which vary in size, color, and form [2]. Another option is to use a BP neural network or a support vector machine to segment the image. Object shape and color features were used to train and integrate a classifier, a hybrid classifier, and a K-means clustering segmentation [3]. Object's region was used to boost edge definition in the image using the

---

\* Corresponding author: [komalsharma00061@gmail.com](mailto:komalsharma00061@gmail.com)

histogram equalization image enhancement, and binary segmentation algorithms are performed using a Bayberry image segmentation approach based on salient object detection homomorphic filtering. It may segment a variety of crop plants, including rice, wheat, and spices. It uses preprocessing to enhance the image for human viewing and to make subsequent processing steps on the resulting images by machines easier. It also provides precise details about the vein pattern in the captured image [4,5]. In order to predict nitrogen deficiency in rice, a convolutional neural network (CNN) model was employed in which raw images were preprocessed to subtract white backgrounds [6]. Transfer learning is a machine learning approach that is redefined as a start to take care of an alternate issue utilizing the knowledge gathered from an established model. The authors of that study improved this established model by utilizing pre-trained CNN models dependent on transfer learning using rice leaf disease images [7]. The degree of severity is given due importance to classify the panicle blast of *Orzya Sativa* using an automatic methodology consisting of a small and robust CNN model [8]. Moisture content is also considered an important feature of rice seed, which determines the water quantity in the rice seed [9]. The rice chalkiness is also evaluated using image processing techniques using the whole area and chalky rice grains [10]. The texture features exhibit similarity, and geometric features prove distinct when deep analysis of rice seed classification performance using different textural and geometric features was performed [11].

The following are some typical image processing steps, but not all image processing systems would need all of them simultaneously [12].

1. *Image Acquisition* is "A sensor, such as a Scanner, camera, captures and digitizes an image."

2. *Image Preprocessing* is "the first phase in preparing an image for higher-level processing." The goal is twofold: to remove undesirable features that will obstruct further processing and extract desirable features that reflect valuable information in the image. In this stage, the computer reduces noise and, in some cases, improves certain object features that are critical to comprehending the picture [13].

3. *Image Segmentation* is "used to minimize the amount of data to make analysis easier. Image Analysis and Image Compression also benefit from segmentation" [14]. "The process of breaking down an image into its constituent regions or objects is known as segmentation," to read the image correctly and accurately identify the image's material, image segmentation must separate the object from the context [15].

4. *Representation and Description* here "Representation is the process of converting raw data into a form that can be processed by a computer [16]. Description, also known as feature extraction, is the process of extracting features that result in quantitative information of interest or features that are important for distinguishing one object type from another".

5. *Recognition and Interpretation* here "Recognition is the process of assigning a mark to an object based on its descriptors' details [17]. The process of assigning meaning to a group of recognized objects is known as interpretation."

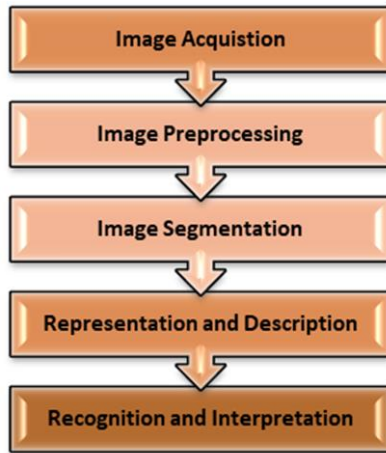


Fig. 1. Phases of image processing.

## 2. Pre-processing

Pre-processing [18] is critical due to noisy, unreliable, and incomplete data. It is one of the preliminary steps that must be completed in order to achieve high step accuracy. "Data in object identification and classification uses different Preprocessing techniques [19]. Although geometric transformations of images (e.g., rotation, scaling, and translation) are listed among preprocessing methods, preprocessing aims to improve the image data by suppressing unwanted distortions or enhancing certain image features necessary for further processing [20]. Image preprocessing methods exploit the significant redundancy of images. Different denoising methods have been suggested.

Table 1. Comparison of different preprocessing techniques.

Sl. No	Preprocessing Technique	Author	Advantages	Disadvantages
1.	Adaptive Median Filter	Ajay <i>et al.</i> [15]	1) Used to smooth non-repulsive noise without causing edge blurring. 2) In the case of high-density impulse noises, keep edge details.	1) When the impulse noise is greater than 0.2, it does not perform well.
2.	Morphological Operations	Anami <i>et al.</i> [16]	1) Allow identifying lesions of different sizes and forms, including those with complex shapes.	1) Based on the principles of infimum and supremum, which necessitates proper color ordering
3.	Mean filter (or) average filter	Chaugule and Mali [17]	1) It is simple to carry out and reduces variance.	1) Averaging operations trigger image blurring, which affects feature localization. 2) Impulse noise is not eliminated. 3)Affects the mean value of all pixels in the vicinity.

4.	Gabor Filter	Devi <i>et al.</i> [18]	1) Individuality 2) It is highly unique to a time and scale. 3) FFT-based Fourier analysis is fast. 4) Applicable to the measurement of stationary signals.	1) For FFT to work, the image must be around the size of a power of two. 2) An issue with the boundary condition 3) A signal's frequency domain and time domain descriptions are inversely related.
5.	Image Normalization	Ghatkamble [19]	1) When photos are normalized before being endorsed, the size and position of the endorsements would be consistent across the data set's pages. 2) If images are printed, using normalized images will avoid printing issues caused by the page size and orientation changes.	1) Image normalization is a time-consuming process that, in large cases, can add considerable time to the e-Discovery export process. 2) Using normalization software that is poorly built will degrade the overall image quality.
6.	Histogram Equalization	Guzman and Peralta [20]	1) Simple and effective in enhancing picture contrasts.	1) This approach fails if the picture contains grey values that are physically separated from one another.
7.	Weighted Median Filter	Hobson <i>et al.</i> [21]	1) Removes salt and pepper noise effectively. 2) Noise is easily eliminated	1) The most significant flaw is the inability to round corners and map texture regions to a uniform hue.
8.	Weiner Filter	Kambo and Yerpude [22]	1) Minimizing mean square error is a time-consuming operation. 2) Ability to deal with both deterioration and noise	1) A fair degradation function estimation is ineffective.

### 3. Segmentation

It divides an image into regions with different properties, such as color, form, brightness, contrast, and grey. A digital grayscale image is used as the process's data. (for example, an image with some object). The process produces anomalies as a result of its output. Segmentation is used to provide more detail than is available in images. Various techniques such as neural networks, decision trees, rule-based algorithms, and Bayesian networks are used." Various methods for segmenting object detection and classification image regions are proposed. "Pixel-based methods, Edge-based methods, Region-based methods, Model-based methods, Texture-based methods, ANN-based methods, Fuzzy Theory-based methods, and Genetic Algorithm-based methods [22]."

Table 2. Comparison of different segmentation techniques.

Sl. No	Segmentation Technique	Author	Advantages	Disadvantages
1.	Pixel Based	Ajay et al. [15]	1) This algorithm works well when the region homogeneity criterion is simple to define.	1) Costly in terms of computational time and memory. 2) There may be under and over-segmentation in the picture, as well as holes in the field.
2.	Edge Based	Anami et al. [16]	1) It does not necessitate any prior awareness of the image's material. 2) It has a low computational time.	1) If there are too many edges, it is a time-consuming operation.
3.	Region Based	Chaugule and Mali [17]	1) It's easier to describe and put into practice.	1) It's difficult to figure out how many clusters there are. 2) It's difficult to make use of spatial data. 3) Feature selection is difficult to comprehend.
4.	Deformable Models Approach	Devi et al. [18]	1) This method is resistant to noise and erroneous edges. 2) Ability to create a surface from a collection of images	1) To position the initial model, a manual approach is needed. 2) Choosing parameters is difficult.
5.	Texture Based	Ghatkamble [19]	1) Using training data, it is simple to solve complex problems. 2) Error detection is a simpler procedure.	1) Training is expensive and subject to human error. 2) It is important to examine different image forms.
6.	Artificial Neural Network Based	Guzman and Peralta [20]	1) It does not necessitate the development of time-consuming services. 2) Could fully take advantage of the parallel existence of neural networks.	1) A longer period of training. 2) The initialization can have an effect on the result.
7.	Fuzzy theory Based	Hobson et al. [21]	1) A single fuzzy rule was used to emphasize the importance of the image's feature-based and spatial details. 2) The membership functions' structure and related parameters were derived automatically.	1) Noise sensitivity. 2) It is computationally costly. 3) Determining fuzzy membership is a difficult task.
8.	Genetic Algorithm Based	Kambo and Yerpude [22]	1) Prove to be good at breaking out of local optima while still providing much versatility. 2) It's good at boosting contrast and delivering natural-looking images. 3) A straightforward method for tackling complex optimization problems.	1) Simple GA can converge extremely slowly or fail in complex designs due to convergence to an 2) Inappropriate local optimum.

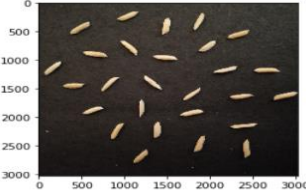
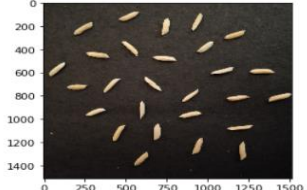
#### 4. Proposed Methodology

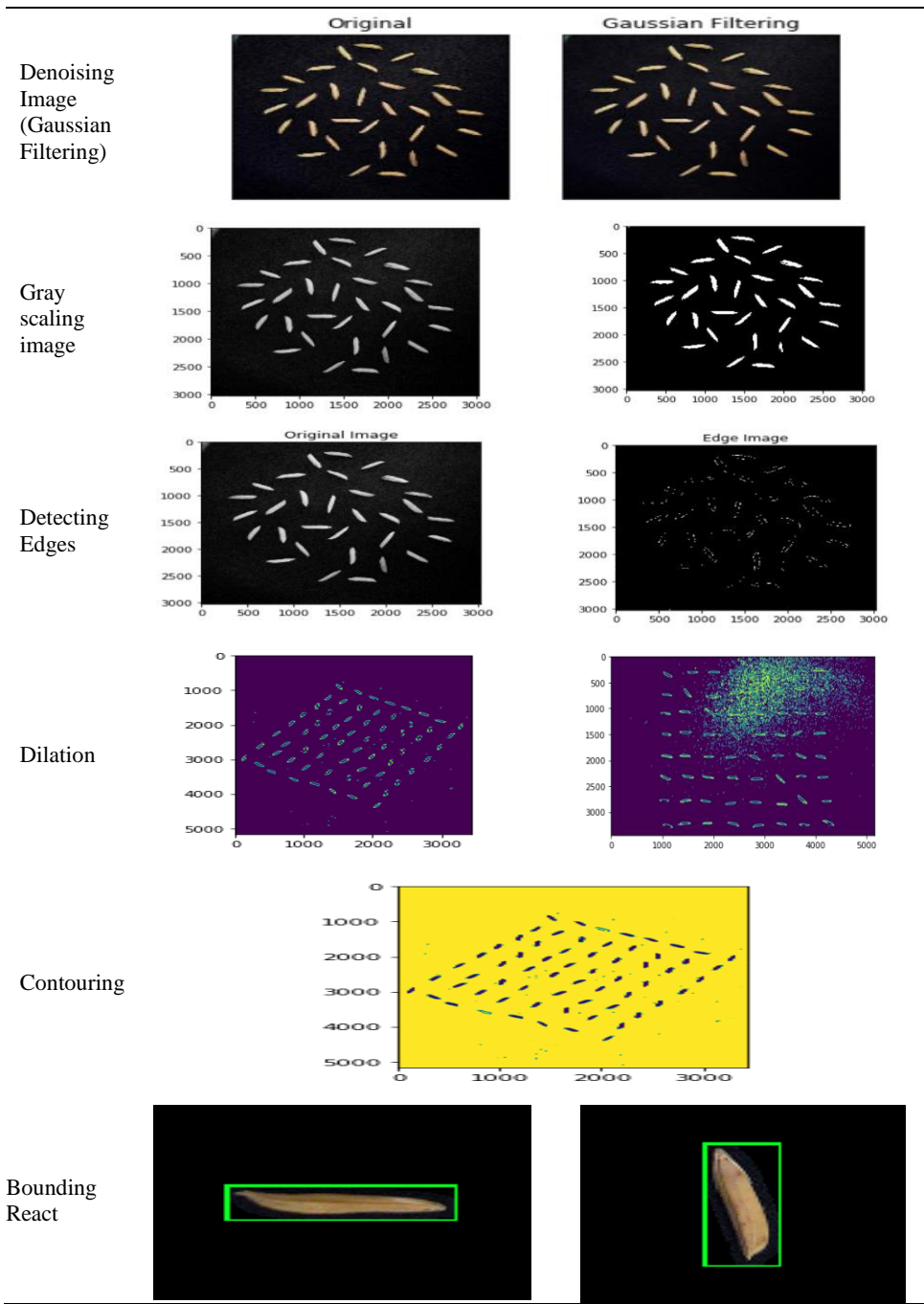
In this research paper, we propose a hybrid approach of Preprocessing and Segmentation techniques to develop an automatic rice variety identification system that is capable of identifying the variety or quality of food grains based on extracted features from the rice images. The proposed system preprocesses the images from datasets, extracts the features, segments these images, and classifies the images based on extracted features. The proposed system is capable of identifying rice grains as well as rice seeds. In the phases of preprocessing and segmentation, we propose the following algorithm to gain better results in the variety identification process of rice images. Using a proposed algorithm, we have succeeded in removing impurities present in the samples of images and obtaining clean images for the classification process.

##### *Proposed hybrid algorithm*

1. Read Image
2. Apply Preprocessing Techniques
  - i. Resize the Image
  - ii. Denoise the Image using 5x5 Gaussian Filter
  - iii. Smooth the Image using low-pass Filter Kernel
  - iv. Apply augmentation techniques for enhancement of Images.
  - v. Convert the Image into grayscale using AdaptiveThresholding Method
3. Apply Segmentation Techniques
  - i. Use the Canny Edge Detection to detect the edges
  - ii. To dilate the images to highlight the edges and remove the dirt particles
  - iii. Use contours to detect objects individually
  - iv. Use Masking to highlight objects with different color schemes
  - v. Use Bounding Rect to bound the individual objects in square boxes and find a region of interest

Some of the results of the proposed hybrid algorithm are given below:

Operations	Original Image	Resulted Image
Resizing Image		



## 5. Conclusion

Various preprocessing and segmentation approaches are examined in this paper. It is well known that all approaches are effective for a variety of purposes. In the table, the benefits and drawbacks of these approaches are discussed. "Preprocessing filters such as Gabor and Histogram Equalization have been used to enhance different imaging modalities and diagnoses. The simplest segmentation methods are based on grey level techniques like thresholding, i.e., pixel-based and region-based methods, but they have limited implementations". However, by combining it with other techniques, the overall performance of all techniques can be increased. Segmentation using a Genetic Algorithm is being used to find approximate solutions to optimization problems. In this paper, we have also proposed an algorithm of preprocessing and segmentation techniques to gain better results in the identification and recognition process of images in automatic object detection and recognition system.

## References

1. A. Asuncion and D. J. Newman, UCI Machine Learning Repository Irvine, CA (University of California, School of Information and Computer Science, 2013).
2. Komal, G. S. Sethi, and R. K. Bawa, *J. Sci. Res.* **64**, 305 (2020).  
<https://doi.org/10.37398/JSR.2020.640242>
3. Handbook of Medical Image Processing and Analysis, ed. I. Bankman (Academic press, 2008).
4. J. C. Bezdek, R. Ehrlich, and W. Full, *Comput. Geosci.* **10**, 191 (1984).  
[https://doi.org/10.1016/0098-3004\(84\)90020-7](https://doi.org/10.1016/0098-3004(84)90020-7)
5. P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, *J. Ambient Intell. Human Comput.* **11**, 5703 (2020). <https://doi.org/10.1007/s12652-020-01938-8>
6. P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, *Comput. Electron. Agricult.* **175**, 105527 (2020). <https://doi.org/10.1016/j.compag.2020.105527>
7. P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. C. Rajpoot, *Arch. Phytopathol. Plant Protection* **54**, 1001 (2020). <https://doi.org/10.1080/03235408.2020.1869386>
8. P. K. Sethy, N. K. Barpanda, S. K. Behera, S. Dash, and A. Pattnaik, *Int. J. Info. Technol. Elect. Eng.* **8**, 26 (2019).
9. P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, *Asian J. Multidimensional Res. (AJMR)* **7**, 74 (2018).
10. P. K. Sethy and A. K. Chatterjee, *Int. J. Appl. Eng. Res.* **13**, 35 (2018).
11. H. P. Ng, S. H. Ong, K. W. C. Foong, P. S. Goh and W. L. Nowinski, *Medical Image Segmentation Using K-Means Clustering and Improved Watershed Algorithm - 2006 IEEE Southwest Symposium on Image Analysis and Interpretation* (2006) pp. 61-65, doi: <https://doi.org/10.1109/SSIAI.2006.1633722>.
12. N. R. Pal and S. K. Pal, *Pattern Recog.* **26**, 1277 (1993).  
[https://doi.org/10.1016/0031-3203\(93\)90135-J](https://doi.org/10.1016/0031-3203(93)90135-J)
13. P. Singh and R. S. Chadha, *Int. J. Adv. Res. Comput. Sci. Software Eng. Res. Paper* **3** (2013).
14. G. Ajay, M. Suneel, K. K. Kumar, and P. S. Prasad, *Int. J. Soft Comput. Eng.* **2**, 35 (2013).
15. B. S. Anami, N. M. Naveen, and N. G. Hanamaratti, *Int. J. Signal Process. Image Process. Pattern Recog.* **8**, 19 (2015). <https://doi.org/10.14257/ijsip.2015.8.4.02>
16. A. Chaugule and S. N. Mali, *J. Eng.* **2014**, ID 617263 (2014).  
<https://doi.org/10.1155/2014/617263>



17. T. G. Devi, P. Neelamegam, and S. Sudha, Machine Vision Based Quality Analysis of Rice Grains - *IEEE Int. Conf. on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017)* pp. 1052-1055. <https://doi.org/10.1109/ICPCSI.2017.8391871>
18. R. Ghatkamble, *Int. J. Comput. Sci. Technol.* **4**, 95 (2016).
19. J. D. Guzman and E. K. Peralta, Classification of Philippine Rice Grains using Machine Vision and Artificial Neural Networks - *World Conf. on Agricultural Information and IT* (Tokyo, Japan, 2008) pp. 24-27.
20. D. M. Hobson, R. M. Carter, and Y. Yan, Characterization and Identification of Rice Grains through Digital Image Analysis – *Conf. Record - IEEE Instrumentation and Measurement Technology Conf.* (2007) pp. 1-7. <https://doi.org/10.1109/imtc.2007.379116>
21. R. Kambo and A. Yerpude, *Int. J. Comput. Trends Technol.* **11**, 80 (2014). <https://doi.org/10.14445/22312803/IJCTT-V11P117>