RESEARCH ARTICLE

BON VIEW PUBLISHING

A Robust SLIC-Based Approach for Segmentation Using Canny Edge Detector

Srikanta Pal[1], Ayush Roy[2] , **Palaiahnakote Shivakumara[3] and Umapada Pal[2],***

[1]*Maynooth International Engineering College, Maynooth University, Ireland*

[2]*Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, India*

[3]*Faculty of Computer Science and Information Technology, University of Malaya, Malaysia*

Abstract: An accurate image segmentation in noisy environment is complex and challenging. Unlike existing state-of-the-art methods that use superpixels for successful segmentation, we propose a new approach for noise-robust simple linear iterative clustering (SLIC) segmentation that incorporates a Canny edge detector. Generally, methods do not consider the boundaries of various objects of an image, which leads to incorrect assignment of the pixels to a specific cluster which violates the boundaries of the objects. This leads to faulty superpixels. By leveraging Canny edge information, the proposed method modifies the pixel intensity distance measurement to overcome boundary adherence challenge. Furthermore, we adopt a selective approach to updating cluster centers, focusing on pixels that contribute less to the noise using a selection method which focuses on choosing pixels closer to the mean of the pixel distribution. These selected pixels only contribute to the calculation of the centroid of the assigned superpixel. Extensive experiments on synthetic noisy images demonstrate the effectiveness of our approach. It significantly improves SLIC's performance in noisy image segmentation and boundary adherence, making it a promising technique for vision processing tasks.

Keywords: MSER, deep learning, Swin transformer, text detection, license plate number detection

1. Introduction

Computer vision is used in various fields like disease diagnosis [1], text detection [2, 3], etc., to automate tasks for faster and accurate life. Image segmentation is an important aspect for various complex deep learning tasks to spot the region of interests, which are difficult to spot sometimes with human eye. Segmenting region of interest in the images of different situations and application is an integral part of image understanding and image captioning. For an accurate segmentation of vital region that signifies dominant information, superpixel segmentation has been introduced as a preprocessing step for various computer vision applications, including object tracking, object localization, segmentation, and image co-saliency detection. This is because superpixels represent vital information in images. It involves grouping pixels with similar low-level features, such as similarity in color or position, into coherent regions called superpixels. By operating at the superpixel level instead of individual pixels, computations and decisions can be made more efficiently. For instance, simple linear iterative clustering (SLIC) is widely preferred due to its computational efficiency and ability to generate visually meaningful superpixels. When evaluating the quality of superpixel segmentation algorithms, several important criteria are considered. First, it is crucial that the boundaries of resulting superpixels align accurately with the boundaries in the original image, ensuring boundary adherence of

images as shown in Figure 1. The distance, which considers both color and spatial proximity, between a cluster center c_k and a pixel p_i is smaller than the distance between another cluster center c_t and the same pixel p_i. Consequently, the pixel p_i might be wrongly assigned to the cluster associated with c_k instead of c_t. Consequently, the generated superpixels may not accurately preserve the boundaries between the two regions. Second, the resulting superpixels should be compact and have regular shapes. Finally, a good superpixel algorithm must be straightforward and computationally efficient. By fulfilling these criteria, SLIC and other effective superpixel algorithms enable improved performance in various computer vision tasks. Their ability to generate meaningful superpixels enhances the accuracy and efficiency of subsequent image processing and analysis.

Figure 1
Illustration of the problem associated with Euclidean distance

Pixel p_i
Cluster centre c_t
Cluster centre c_k
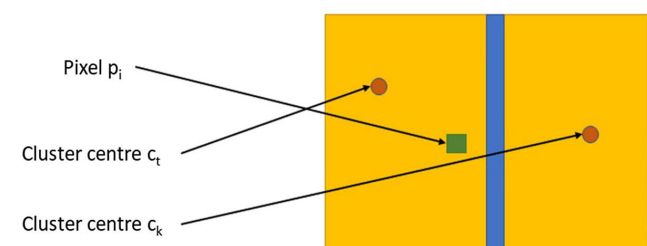
*Corresponding author: Umapada Pal, Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, India. Email: umapada@isical.ac.in

However, most of the existing SLIC-based methods may not be effective for noisy images because the existing SLIC is sensitive to noise and background complexity. Therefore, there is a need for extending the present SLIC to work on noisy images. Thus, this work aims at improving the existing SLIC based on Canny edge features and pixel intensity distance to perform grouping for superpixels segmentation. Hence, the contribution of the proposed works is as follows. (i) Use of Canny edge features to find fine details in the noisy images is new compared to superpixel segmentation. (ii) Proposing distance between pixels intensity values based on Canny edge information to group actual image pixels in noisy images is another contribution of the proposed work.

The paper is structured as follows: Section 2 provides a succinct overview of related works, while Section 3 delves into the proposed method. The experimental findings are presented in Section 4, and the paper concludes with Section 5.

2. Literature Review

Superpixels are groups of pixels that exhibit similarities in terms of color or position. These superpixels have gained popularity in computer vision due to their wide range of applications, particularly in image segmentation as a crucial preprocessing step. Numerous methods have been developed for superpixel segmentation, which can be broadly categorized into graph-based methods and clustering-based methods [4–10].

2.1. Graph-based methods

In certain image processing approaches, an image is represented as a graph $G = (V, E)$, with V representing the nodes and E representing the edges. Each pixel in the image I is represented as a node $v_i \in V$, and neighboring pixels are connected by edges $e_j \in E$. The weight of each edge (say w) between two nodes is determined by the similarities in pixel position and color. The superpixels are generated by minimizing the cost of weighted edges. The normalized cuts algorithm [9], proposed by Shi et al., recursively partitions the graph of all the pixels in an image based on contour and texture cues. It optimizes a criterion that considers the total similarity and dissimilarity between different groups, resulting in superpixels with regular shapes. However, this algorithm fails to preserve the boundaries of image and is computationally expensive, involving a large number of matrix calculations with a time complexity of $O(N^{3/2})$ [5]. To address these challenges, [7] introduced a fast graph-based algorithm utilizing a bottom-up clustering approach. In this method, every pixel is considered as an initial cluster, and pixels are merged with the application of a greedy algorithm. This approach adheres well to image boundaries but produces superpixels that are highly irregular in shape and size. Compactness is not explicitly considered. The time complexity of this algorithm is $O(N\log N)$. The choice of algorithm depends on the specific requirements of the application, balancing boundary adherence, computational efficiency, and the desired characteristics of the resulting superpixels.

2.2. Clustering-based methods

Another approach to generate superpixels is through gradient ascent methods, which generate superpixels based on gradients. These approaches initiate the process with an initial set of K cluster centers placed randomly across various pixels within the image. The updation of the cluster centers is achieved until a convergence criterion is reached, and the image is segmented into K clusters. The mean shift [6] clustering algorithm associates each pixel with a probability density function and clusters the pixels based on the image's color modes and intensity probability density function. This method is a nonparametric clustering method and does not require the number of clusters to be given beforehand. The clustering procedure is based on the variation in the local density values, with each pixel being clustered with other pixels having the same probability density value. This method generates nonuniform-sized superpixels having irregular shapes, and its time complexity is $O(N^2)$. The turbopixels algorithm [5] produces superpixels efficiently by initializing seeds placed regularly on the image and using a level set-based geometric flow evolution process that relies on image gradients in order to establish evenly distributed superpixels across the image plane. These superpixel boundaries emerge at the intersections of two flow patterns. Although the time complexity is $O(N)$, this algorithm is relatively slow in practice and generates superpixels with relatively poor boundary adherence. Superpixels have gained popularity as an efficient method for image processing and find widespread applications in various fields, including medical imaging [11–14], marine Synthetic aperture radar (SAR) imagery [15–17], and Unmanned Aerial Vehicle (UAV)/drone imagery [18, 19]. In Giraud et al. [20] and Zhou et al. [21], generalized decomposition approach has been used for superpixel segmentation in omnidirectional images [20] and color, and gradient information is used for vine spread for superpixel segmentation [21]. The review on different methods of superpixel segmentation indicates that superpixel-based methods are robust to different situations. This observation motivated us to explore superpixel-based methods for segmentation in this work.
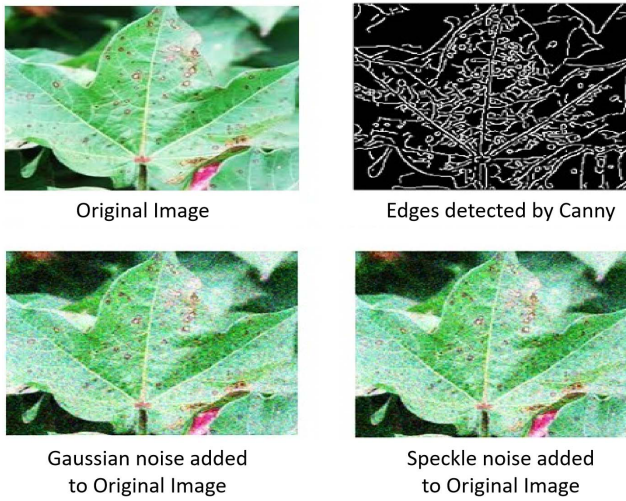
3. Proposed Method

It is noted from the review on existing superpixels segmentation that the methods are not robust to noise images. It is true that, when we look at real-time applications, introducing noise while capturing images is common due to faults in the devices and other external environmental factors. It is also noted from the review that the SLIC is the most popular approach for superpixels segmentation. However, this method does not perform well for noisy images. This is because the method works at pixel level. This observation motivated us to propose modifications to the current SLIC such that it can withstand the challenges of noisy images. Therefore, the proposed model is a modified SLIC for superpixels segmentation. The proposed work leverages Canny edge features (shown in Figure 2) and modifies pixels intensity distance for clustering actual image pixels despite the noisy pixel present in the images. The intuition behind this approach is that Canny edge operator detects the high gradient value, which represents edge pixels in the images for edge detection. Therefore, the Canny edge operator outputs edges though noise pixels are present in the images.

3.1. Overview of SLIC segmentation

This algorithm relies solely on the desired number of clusters, denoted as K. The RGB image is first converted to the LAB color space. In the initialization step, K cluster centers are evenly sampled on a regular grid of S pixel spacing. The grid-spaced S pixels represent the area surrounding each cluster center. The grid interval is represented by $S = \sqrt{N/K}$ (N represents the number of image pixels). The cluster centers are

Figure 2
Some sample images and the edges detected by Canny edge detector

Original Image

Edges detected by Canny

Gaussian noise added
to Original Image

Speckle noise added
to Original Image

$$c_j = \left[l_j a_j b_j x_j y_j\right]^T \tag{1}$$

In Equation (1), the variables x and y are the pixel's coordinate position, while l, a, and b correspond to the respective channels in the CIE Lab* color space. To ensure that the cluster centers do not coincide with image edges, a precaution is taken. The centers are adjusted within a 3 × 3-pixel neighborhood, aiming to position them at locations with the lowest gradient. This prevents the seeding of superpixels with noisy pixels or overlapping with neighboring clusters. Each pixel p_i is then associated with the nearest cluster center c_j. The algorithm focuses on calculating feature-based similarities only between each cluster center and the pixels within a 2S × 2S search region, where the search region encompasses the cluster center. This approach differs from the conventional *K*-means clustering algorithm, where each of the cluster centers is compared against all other pixels in the entire image. By limiting the comparisons to a smaller region, the algorithm significantly reduces the number of feature-based similarity calculations, resulting in a notable speed advantage. The time complexity of the algorithm is $O(N)$, where N represents the number of pixels in the image. This particular process of association depends on the combination of spatial proximity and color proximity calculations between each pixel pi and the nearest cluster center c_j. The color proximity is

$$dc(c_j, p_i) = \sqrt{[\![(l_j - l_i)]\!]^2 + [\![(b_j - b_i)]\!]^2 + [\![(c_j - c_i)]\!]^2} \tag{2}$$

And, the spatial proximity is defined as:

$$ds(c_j, p_i) = \sqrt{[\![(x_j - x_i)]\!]^2 + [\![(y_j - y_i)]\!]^2} \tag{3}$$

The combined (and normalized) proximity is

$$Dc(c_j, p_i) = \sqrt{[\![\frac{dc(c_j, p_i)}{m}]\!]^2 + [\![\frac{ds(c_j, p_i)}{S})^2]\!]} \tag{4}$$

In Equation (4), the variable S represents the anticipated maximum spatial distance within a designated cluster. The

constant m, which lies in the range of [1.40], denotes the maximum color distance and can vary for each cluster or image. After the initial clustering, the mean vector [labxy]^T is used to update the cluster centers, which represent the average values of the LAB color channels and pixel coordinates within each cluster. To measure the convergence of the algorithm, the residual error denoted as *Err* is calculated. *Err* is the difference between the new positions of the cluster centers and their previous positions (*Err* ≤ threshold value incorrectly assigned to cluster c_k instead of with cluster center c_t). The algorithm iterates until *Err* falls below a specified threshold value. It is important to note that in the iterations, some pixels may be incorrectly associated with a cluster c_k instead of their corresponding cluster center c_t. Addressing this issue is crucial as it reduces convergence time and improves the overall quality of the image. Figure 1 provides a visual illustration of this improvement. By refining the association of pixels to the appropriate cluster centers, the algorithm can achieve faster convergence and enhance the segmentation results, resulting in improved image quality.

3.2. SLIC with Canny edge detector

SLIC employs the Euclidean distance metric for pixel similarity, which can lead to missing boundaries of content-sensitive superpixels. To address this limitation, the Canny edge detector, a commonly used edge detection method, is incorporated as an initial step in SLIC. The Canny edge detector locates image edges by identifying local maxima in the image gradient. By integrating Canny edge detection into SLIC, the idea is to artificially decrease the feature-based similarity of a pixel along the path to the closest cluster center if that path consists of edge pixels identified by the Canny edge detector. This approach leverages the intuition that incorporating Canny edge detection can modify the feature-based similarity measurement, reducing the impact of edge pixels on the clustering process. By doing so, SLIC with Canny edge detection aims to enhance the preservation of boundaries and improve the accuracy of content-sensitive superpixel segmentation. This approach addresses the problem of boundary adherence commonly encountered in superpixels with content sensitivity. Equation (4) is modified as:

$$Dc(c_j, p_i) = \sqrt{[\![\frac{dc(c_j, p_i)}{m}]\!]^2 + [\![\frac{ds(c_j, p_i)}{S})^2]\!] + \frac{G(c_j, p_i)}{R}} \tag{5}$$

In Equation (5), G is 0 if the pixel is not an edge pixel and is equivalent to the gradient magnitude calculated by the Canny edge detector if it is an edge pixel. The segmentation results for various values of R (see Equation (5)) show better results than SLIC segmentation, thus solving the boundary adherence problem (also illustrated in Figure 3 where the dotted yellow circle highlights the effectiveness of the solution).

3.3. Noise-robust SLIC with Canny

To make the method noise-robust, we first consider the $[lab]^T$ color intensities and calculate the mean and standard deviation for each superpixel cluster. Let us assume the mean is m_i and the standard deviation is std_i for the ith cluster. Now, we select the pixels p for updating the cluster center as:

Figure 3
Segmentation results of SLIC compared with SLIC+Canny. The highlighted region (yellow encirclement) illustrates the boundary adherence problem overcome using the Canny edge detector

SLIC   R=1.5   R=0.7
                          SLIC+CANNY
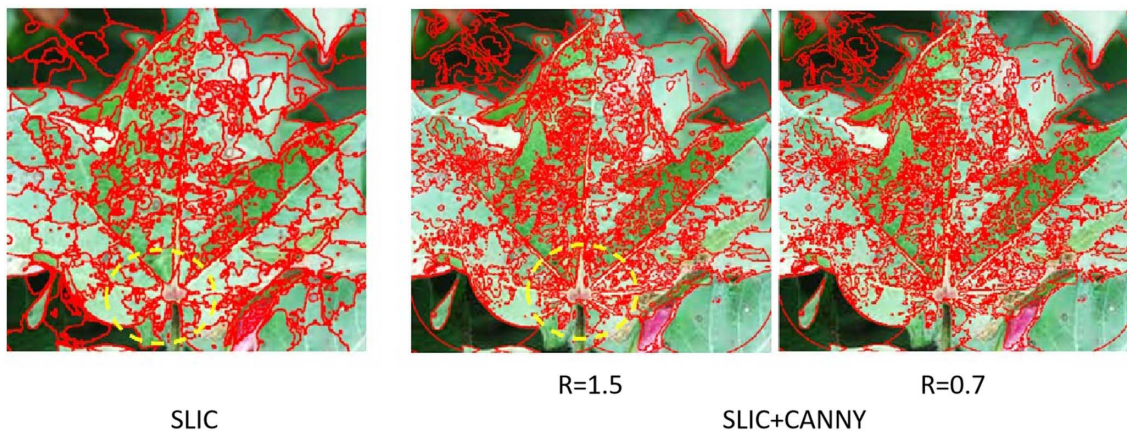
$$p_i \in [m_i - a * std_i, m_i + a * std_i] \qquad (6)$$

In Equation (6), p_i represents the $[lab]^T$ color intensities of pixel p in the ith cluster. It is evident from Figure 4 that noise-robust SLIC+Canny performs better than SLIC + Canny when noise is incorporated into the image. The importance of this step is shown in Figure 4.

4. Experimental Results

Dataset: The Berkeley Segmentation Dataset and Benchmark (BSDS500) is a widely used dataset for superpixel segmentation in computer vision research. It comprises 500 natural images of varying sizes and resolutions, collected from diverse sources. The dataset is divided into three subsets: 200 training images, 100 validation images, and 200 test images. Each image in BSDS500 is accompanied by manually annotated ground truth segmentations, providing pixel-level labeling of superpixels. The ground truth annotations are used for evaluating the performance of superpixel segmentation algorithms. The BSDS500 dataset offers a challenging and comprehensive benchmark for assessing the effectiveness and accuracy of various superpixel segmentation methods across different image types and scenes. Researchers often employ this dataset to develop and compare different algorithms for image segmentation,
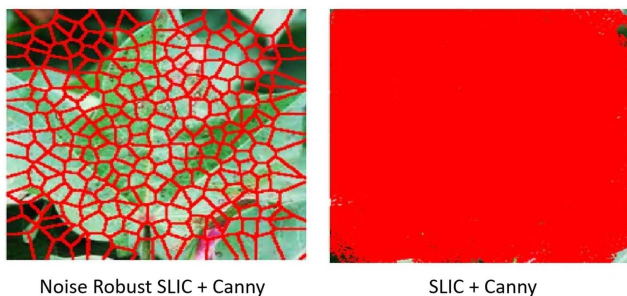
and it has significantly contributed to advancements in superpixel segmentation techniques.

Peak signal-to-noise ratio (PSNR) is a widely used metric in image and video processing to quantify the quality of a reconstructed or compressed signal compared to its original, pristine version. It is expressed in decibels and measures the ratio of the maximum possible power of the original signal (denoted as Pmax) to the mean squared error (MSE) between the original and the reconstructed signals (denoted as MSE). The PSNR formula is given by:

$$PSNR = 10 \left(\log_{10} \frac{P_{max}{}^{\wedge}2}{MSE} \right)$$

The higher the PSNR value, the better the reconstruction quality, as it indicates a smaller difference between the original and reconstructed signals, implying higher fidelity and lower perceptual distortion.

The study utilized the BSDS500 dataset to assess the effectiveness of the proposed model in handling noise during segmentation. The results presented in Figure 5 indicate that the noise-robust SLIC, which incorporated a Canny edge detector, outperformed both the standard SLIC and its enhanced versions in terms of segmentation accuracy. Furthermore, Figure 6 demonstrates that the proposed approach achieves comparable performance with other methods in the absence of noise in normal images. However, when noise is introduced into the image, the proposed approach outperforms other methods. Furthermore, it is seen in Fulkerson et al. [22] that SLIC is computationally extremely fast, way faster than deep learning counterparts and other segmentation methods giving comparative results on standard datasets.

In the same way, Table 1 refers to PSNR values of the proposed and existing methods calculated on the BSDS500 dataset. For the noisy images, the proposed method reports the highest PSNR value compared to the existing methods. However, for the images without noise, the existing methods [20, 21] report better PSNR values than the proposed method. This is due to the existing methods being developed for noise-free images while the proposed method is developed for noisy images. Therefore, one can infer that the proposed method is robust to noisy images but not noise-free images.

Figure 4
Segmentation results of SLIC + Canny and noise-robust SLIC + Canny on a noise-incorporated image

Noise Robust SLIC + Canny   SLIC + Canny

Figure 5
Boundary recall score for the various methods on the BSDS500 dataset for noise added to the images
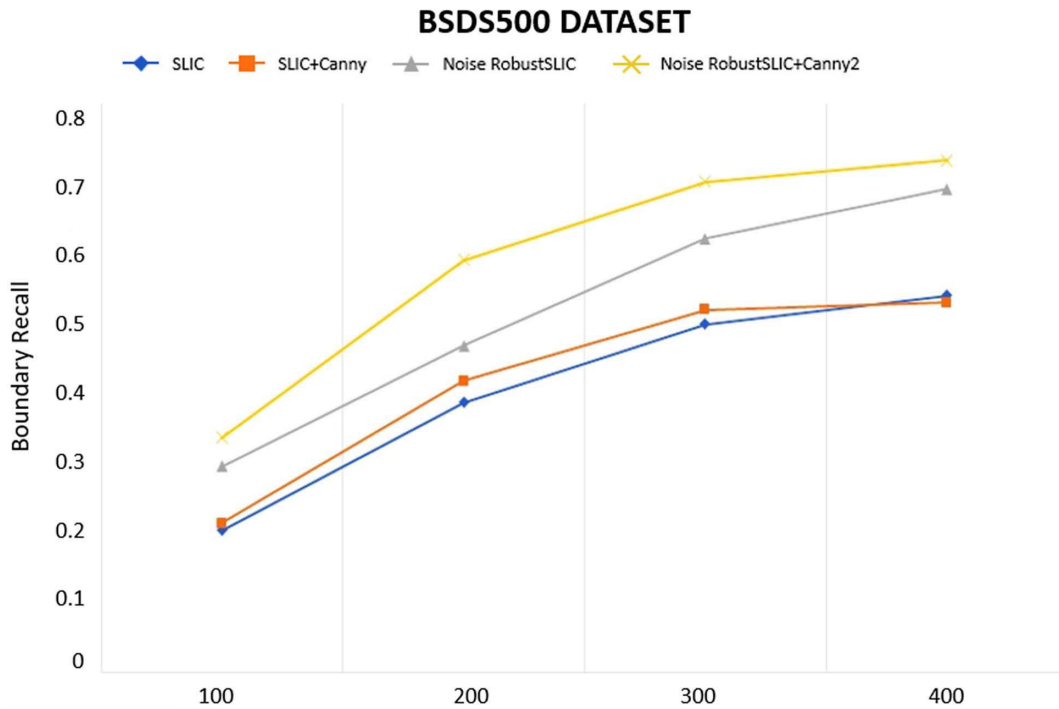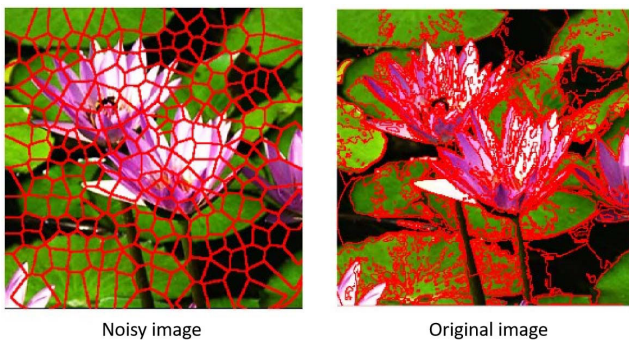

Figure 6
Illustration of the proposed superpixel segmentation technique on original image and noise-incorporated image

Noisy image   Original image

pixel intensity values such that it groups actual image pixels despite the presence of noisy pixels. The results on the standard dataset show the segmentation performance of the proposed method is better than existing methods, especially for noisy images. However, it is noted that the Canny edge operator is sensitive to degradation and poor-quality images. Therefore, for blurry images, the performance of the method degrades. To further enhance this methodology, future endeavors could explore utilizing more sophisticated and robust edge detection algorithms instead of relying solely on the Canny edge operator. Additionally, we plan to work on reducing the time complexity of pixel filtering to make the algorithm more efficient for real-time applications.

Supportive Materials

The codes will be made available at https://github.com/AyushRoy2001/Noise-robust-SLIC-Segmentation-using-Canny-edge-detector.git.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

Palaiahnakote Shivakumara is an editor-in-chief and Umapada Pal is an advisory board member for *Artificial Intelligence and Applications*, and were not involved in the editorial review or the decision to publish this article. The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data supporting the findings of this study is BSDS500, which is openly available at: https://opendatalab.com/OpenDataLab/BSDS500.

Table 1
Performance of the proposed method

Method	PSNR	
Type	No noise	Noise
SLIC alone	29.240	10.163
SLIC + Canny	30.215	12.311
Giraud et al. [20]	31.01	16.97
VSSS [21]	31.36	12.58
Proposed method	30.168	24.906

5. Conclusion and Future Work

We have proposed a modified SLIC for segmenting noisy images. To make the proposed method robust to noise, the proposed work explores Canny edge features, and it modifies the distance between

Author Contribution Statement

Ayush Roy: Conceptualization, Methodology, Software, Investigation, Writing – original draft, Visualization. **Palaiahnakote Shivakumara:** Conceptualization, Validation, Writing – original draft, Supervision, Project administration. **Umapada Pal:** Conceptualization, Writing – review & editing. **Srikanta Pal:** Validation.

References

[1] Roy, A., Acharya, A., Biswas, S., Ray, S., & Ganguly, B. (2022). Identification and classification of human mental stress using physiological data: A low-power hybrid approach. In *IEEE 6th International Conference on Condition Assessment Techniques in Electrical Systems*, 136–139. https://doi.org/10.1109/CATCON56237.2022.10077709.

[2] Nandanwar, L., Shivakumara, P., Ramachandra, R., Lu, T., Pal, U., Antonacopoulos, A., & Lu, Y. (2021). A new deep wavefront based model for text localization in 3D video. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(6), 3375–3389.

[3] Roy, A., Shivakumara, P., Pal, U., Mokayed, H., & Liwicki, M. (2023). Fourier feature-based CBAM and vision transformer for text detection in drone images. In M. Coustaty and A. Fornés (Eds.), *Document analysis and recognition – ICDAR 2023 workshops* (pp. 257–271). Cham: Springer Nature Switzerland.

[4] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(11), 2274–2282.

[5] Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision, 59*, 167–181.

[6] Levinshtein, A., Stere, A., Kutulakos, K. N., Fleet, D. J., Dickinson, S. J., & Siddiqi, K. (2009). Turbopixels: Fast superpixels using geometric flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(12), 2290–2297.

[7] Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(8), 888–905.

[8] Tan, Z., Wan, L., Feng, W., & Pun, C. M. (2013). Image co-saliency detection by propagating superpixel affinities. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2114–2118.

[9] Vedaldi, A., & Soatto, S. (2008). Quick shift and kernel methods for mode seeking. In Forsyth, D., Torr, P., & Zisserman, A. (Eds.), *Computer vision–ECCV 2008: 10th European conference on computer vision* (pp. 705–718). Springer Berlin Heidelberg.

[10] Wang, P., Zeng, G., Gan, R., Wang, J., & Zha, H. (2013). Structure-sensitive superpixels via geodesic distance. *International Journal of Computer Vision*, *103*, 1–21.

[11] Budak, C., & Mençik, V. (2022). Detection of ring cell cancer in histopathological images with region of interest determined by SLIC superpixels method. *Neural Computing and Applications*, *34*(16), 13499–13512.

[12] Faragallah, O. S., El-Hoseny, H. M., & El-Sayed, H. S. (2023). Efficient COVID-19 super pixel segmentation algorithm using MCFO-based SLIC. *Journal of Ambient Intelligence and Humanized Computing*, *14*(7), 9217–9232.

[13] He, F., Mahmud, M. P., Kouzani, A. Z., Anwar, A., Jiang, F., & Ling, S. H. (2022). An improved SLIC algorithm for segmentation of microscopic cell images. *Biomedical Signal Processing and Control*, *73*, 103464.

[14] Massari, F., Dabus, G., Cortez, G. M., Singh, J., Kuhn, A. L., Naragum, V., . . . , & Puri, A. S. (2023). Super large-bore ingestion of clot (SLIC) leads to high first pass effect in thrombectomy for large vessel occlusion. *Journal of NeuroInterventional Surgery*, *15*(7), 664–668.

[15] Sun, Q., Liu, M., Chen, S., Lu, F., & Xing, M. (2023). Ship detection in SAR images based on multi-level superpixel segmentation and fuzzy fusion. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–15.

[16] Wang, X., Li, G., Plaza, A., & He, Y. (2022). Revisiting SLIC: Fast superpixel segmentation of marine SAR images using density features. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–18.

[17] Xu, X., Du, Y., Cui, G., & Yang, X. (2022). Ocean front detection from SAR imagery using SLIC superpixel segmentation method. In *IEEE International Geoscience and Remote Sensing Symposium*, 2578–2581.

[18] Behera, T. K., Bakshi, S., Nappi, M., & Sa, P. K. (2023). Superpixel-based multiscale CNN approach toward multiclass object segmentation from UAV-captured aerial images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 1771–1784.

[19] Sheela, J. E. D., Jansi Rani, P. A., & Paul, M. A. (2023). Super pixels transmission map-based object detection using deep neural network in UAV video. *The Imaging Science Journal*, *71*(8), 767–775.

[20] Giraud, R., Pinheiro, R. B., & Berthoumieu, Y. (2023). Generalization of the shortest path approach for superpixel segmentation of omnidirectional images. *Pattern Recognition*, *142*, 109673.

[21] Zhou, P., Kang, X., & Ming, A. (2023). Vine spread for superpixel segmentation. *IEEE Transactions on Image Processing*, *32*, 878–891.

[22] Fulkerson, B., Vedaldi, A., & Soatto, S. (2009). Class segmentation and object localization with superpixel neighborhoods. In *2009 IEEE 12th International Conference on Computer Vision*, 670–677.