International Journal of Trend in Research and Development, Volume 9(3), ISSN: 2394-9333
www.ijtrd.com

Segmentation of Images Using the K-Mean Technique after Pre-Processing Using the Nonparametric Confidence Limits Algorithm

[1]Aseel Muslim Eesa, [2] Bushra Saad Jasim and [3]Hayder Raaid Talib
[1]Faculty of Administration and Economics/ Sumer University, Iraq Department of Statistics
[2]University of Baghdad - College of Administration and Economics - Department of Statistics
[3]Faculty of Administration and Economics/ Sumer University, Iraq Department of Statistics

Abstract: The remediation and improvement of digital images are regardedas one of the important fields that are used in many branches of science such as mathematics, statistics, computation, and others that are specially used for remediation. When dealingwith images or sending them through a specific channel, they exposeto disruptionand noisiness. Therefore, it should be used ways and methods to extinguish these images from the noisinessto clarity the features of the image. In this research, a nonparametric statistical confidence boundary algorithm was proposed to process the noisy images Pre-ly, and it gave good and highly efficient results, as the noisiness was removed from the images during and after that the images were cut and the external features of the images were determined by separating the regions of the homogeneity. The method was applied to four. The images taken in the photo captured the noise, the photo was taken to watch the noise.

Keywords: *Image processing, image pre-processing, Segmentation Image, Noise, k-mean, nonparametric Statistical Confidence*

I. INTRODUCTION

The process of division images (Segmentation Image) is one of the most important special processes for remediation and analysis of images that are used for dividing a digital image into multiple parts. that is, the set of image points in an area are similar according to some criteria of homogeneity such as color, density, or texture, to determine Objects and borders in the image. Image segmentation is used in applications medical images (identification of tumors and other diseases, measurement of tissue volumes, computer-guided surgery, diagnosis, treatment planning, study of anatomical structure), locating objects in satellite images (roads, forests, etc.), face recognition, Fingerprint recognition, etc.[2] [7]

Image segmentation can be implemented by many methods and techniques belonging to different image processing techniques such as methods based on edges and others based on regions as well as clustering and thresholding. In this paper, it is dependent on the technique of (Clustering) in the segmentation of the images. It means a collection of similar parts with each other and the different parts from other groups within one cluster so that it is easy to understand and deal with. Subsequently, it is possible to use this type of analysis in understanding the complex nature of vocabulary.[7] [9]

There are statistical methods capable of implementing the cluster-based sham task if the classification techniques are based on the formation of cluster regions consisting of sham elements based on the criteria of similarity or proximity. The k-mean algorithm was taken in the segmentation process. The sham method, which is considered one of the most widespread [8] cluster methods, aims to divide the studied data set into (K) areas of homogeneous and separate clusters. The centers of the cluster areas as a measure of these common characteristics.[6]

Sometimes, the researcher faces some problemsbefore the segmentation process as the problem of noisiness. the natural noisiness (Noise Gaussian) is a type of noisiness. It results from the Image Acquisition system, in which the continuous electrical signal is converted into the digital form accepted by the calculator. The second type is speckle noisiness. It is a double noise that has a granular pattern and is an inherent characteristic of the ultrasound image and the SAR image.[3]

The existence of natural noisiness and spot noisiness significantly deteriorates the quality of the acquired image.

Therefore, the image must be processed before the segmentation process, and this is what we will do as the image is pre-processing.

Preprocessing techniques are used to improve the quality of an image before it is processed in the application. These preprocessing techniques are also called filtration and resolution enhancement, and preprocessing reduces disruption and improves accuracy[10].

Using nonparametric confidence limits by removing noise from the image and then cutting the image using the (k-mean) technique. There are some studies on this topic, (2018) researchers (Buenestado and Acho)[6] developed a new method for image segmentation based on the statistical confidence limits tool along with the well-known Otsu algorithm.

II. IMAGE SEGMENTATION BASED ON NONPARAMETRIC STATISTICAL CONFIDENCE INTERVAL APPROACH

In this paper, a new method for preprocessing the image is proposed, then the images are cut, and the nonparametric confidence limits were used for the Pre- processing.

The nonparametric confidence limits algorithm is one of the modern algorithms for image treatment before the segmentation process. Confidence intervals are based on aligned normal distribution and estimates[1]

It gives narrow confidence intervals with the variance and square bias remaining at the same degree. so the amount of bias is a function of each of the core functions and the derivatives of both the core function and the probability density function, while the amount of variance represents a function of the conditional variance and for each of the core and probability density functions [5][6]

,the nonparametric confidence limits technique depends on the estimation of the kernel function:

$$\hat{f}_n(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h_n}\right)$$

It is a real finite continuous symmetric function and its integral is equal to one $\int K(u)\,du = 1$, and h refers to the bandwidth parameter.

Assuming the existence of the second derivative of the core function K(u) is found, the nonparametric certainty limits are found through the following algorithm:

1- Calculate the variance of the entire image population $\sigma_n^2(x)$ by considering each pixel intensity as a point-wise datum.

2-dissociate the original image into r sub-images.

3-Through the Cartesian coordinate $I_k(i,j), k = 1, \dots, n_p$, which represents the density of the gray-band image unit for each given sub-image, the bandwidth parameter (4) is calculated.

4-Calculate the core estimator as well as the probability density function $\hat{f}_n(x)$.

5-Calculate the variance of the function $\hat{f}_n(x)$ according to the following formula $\sigma_n^2(x) \simeq \frac{\hat{f}_n(x)\int K(u)^2\ du}{nh_n}$

6-Calculate the nonparametric confidence limits for each sub-image $I_k(i,j), k = 1, \dots, r$

$$p\left(\hat{f}_n(x) - Z_{\frac{\alpha}{2}} * \sigma_n(x) < I_k(i,j) < \hat{f}_n(x) + Z_{\frac{\alpha}{2}} * \sigma_n(x)\right) = 1 - \alpha$$

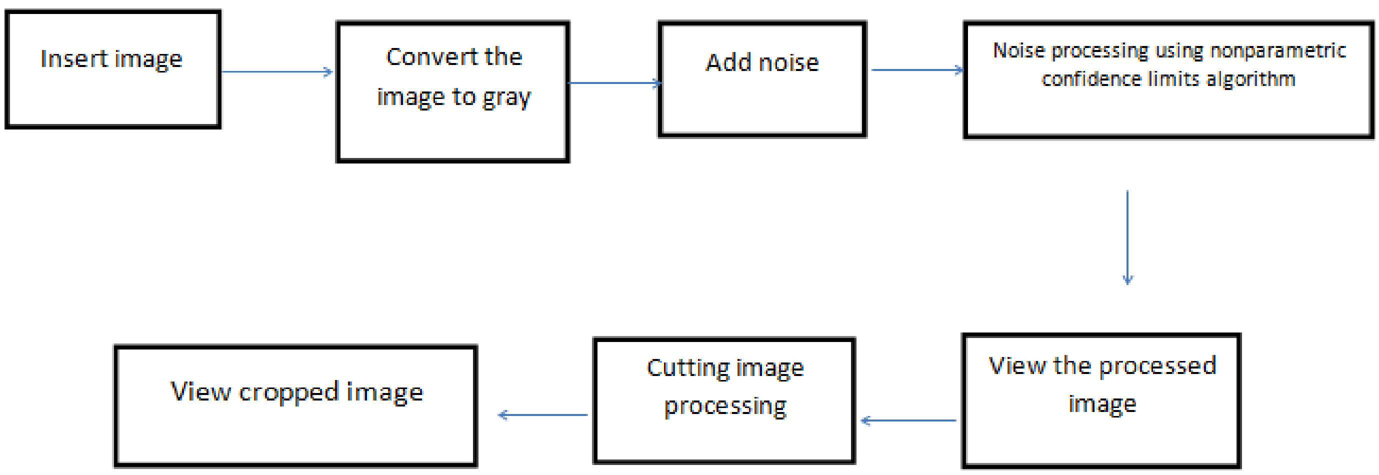
For each sub-image $I_k(i,j)$,if $I_k(i,j) < \hat{f}_n(x) + Z_{\frac{\alpha}{2}} * \sigma_n(x)$ and $I_k(i,j) > \hat{f}_n(x) - Z_{\frac{\alpha}{2}} * \sigma_n(x)$ then $I_k(i,j) = I_k(i,j)$,otherwise

$$I_k(i,j) = h\hat{f}_n(x)$$

7-Reconstruct the preprocessed image by configuring the resulting sub-images from above.

8-Slice the resulting image from step (5) using K-Mean to get the final image.

A figure is a tool that shows the stages of work, as shown as a tool:



Since the Pre-remediation process using the nonparametric statistical confidence limits, we note that the parameter of the bandwidth affects the work of the algorithm, and if it is very small, it gives an opaque picture.

Therefore, it is preferred when extracting the package parameter to use the rule of thumb (Plug-In PI).[4][11]

III. PRACTICAL SIDE

To demonstrate the importance of using the nonparametric statistical confidence limits algorithm in the image processing mentioned in the theoretical side, the algorithm was implemented on four different images using the MATLAB program.

As shown in the results listed below:

Figure 1: (a) Original Images, (b) The original noisy images, (c) The pre-processed images by using our IC method, (d) The images segmentations by invoking the reference method, (e) The images segmentation by using our design.

In Figure (1), the four original images (a) were converted to grayscale images, then natural noisiness (Noise Gaussian was added to the four images with a value of 0.02 as shown in images(b).after applying the nonparametric confidence limits algorithm used, I was able to remove noisiness from the images while showing the most important features as well as highlighting the boundaries of the features as shown in the pictures (c).and by cutting the images (b) in a distorted (K-Mean) way, it has produced image d)), as we note that the cut images have shown the most prominent features while showing the noise that may appear in the images (b),but after cutting the images (c) by the (K-Mean) method, we notice that the resulting images have cut images that contain the most important features while removing the unimportant ones.

CONCLUSION

Through the results of the practical side, it is clear that the nonparametric statistical confidence limits algorithm has given satisfactory results in the Pre- treatment in removing the spots and natural noises from the noisy images, as it proved its efficiency in extracting all the features of the images while removing the noisiness that may appear in the images.

We note the effectiveness of this method when cutting images using the (K-Mean) method, as it has given the most important features.

References

[1] Buenestado, Pablo Iand Acho, Leonardo, (2018)." Image Segmentation Based on Statistical Confidence Intervals ", Journals Entropy, Vol. 20, November 46,Issue 1.
[2] Carlo V. Fiorio, (2004)," Confidence intervals for kernel density estimation ", The Stata Journal, vol. 4, Number 2, pp. 168–179.
[3] Dhanachandra, Nameirakpam & Chanu, Yambem Jina, (2017)," A New Approach of Image Segmentation Method Using K-Means and Kernel-Based Subtractive Clustering Methods", International Journal of Applied Engineering Research, ISSN 0973-4562 Volume 12, Number 20 (2017) pp. 10458-10464.
[4] Gonzalez, R. C., Woods Richard E., 2002, "Digital Image Processing", Prentice – Hall, Inc, London.
[5] Hmood, Munaf Yousif, (2011), "estimate the nonparametric regression function using canonical kernel", journal of economics and administrative sciences, Vol. 17, Issue 61, pp.212-225.
[6] Ilea Dana Elena and Whelan Paul F. (2006), "Color image segmentation using a spatial k-means clustering algorithm",
[7] Liang Qu, Xinghui Dong, and Fading Guo, (2009), "Automatic K-Means for Color Enteromorpha Image Segmentation", Third International Symposium On Intelligent Information Technology Application, IEEE Computer Society, 224 - 227.
[8] Prabhishek Singh & Raj Shree, (2016)," Speckle Noise: Modelling and Implementation", International Science Press,9(17), pp. 8717-8727.
[9] Rajeshwar Dass, Priyanka, Swapna Devi, (2012)," Image Segmentation Techniques ", IJECT Vol. 3, Issue 1, pp.66-70.
[10] Rajeshwari. S.& Sree.T. Sharmila,(2013)," Efficient quality analysis of MRI image using preprocessing techniques"
[11] Wasserman, Larry, (2006),"All of Nonparametric Statistics" A Concise Course in Nonparametric Statistical Inference.