Segmentation Method for Nuclei Cell Image Using Hybrid Techniques

Suhad Fakhri Hussein
Ababil high school-Al-resaffa 1 education
Hai Ur - Baghdad , Iraq
suhad7242@gmail.com

ABSTRACT

In the medical applications, the nuclei images are the most important for classifying the tumors into normal or abnormal (benign and malignant). The important stride in image histometric and cytometry is an automatic segmentation and clustering of cell nuclei. Despite substantial progress, there is a requirement to enhance speed, precision, automation level, and adaptation to modern implementations. A cancer diagnosis for early, cell nuclei automated segmentation is decisive such as characteristics of the nucleus of the cell are fundamentally related to an evaluation of maliciousness. Very little research work was done implemented in cell nuclei segmentation of automated on pleural effusion images cytology, that is seedy treated through previous techniques. Furthermore, cytology pleural effusion image is still defying because of an assortment of cells, the poor contrast of images, and overlapping cells. The recent cancer of cervical diagnosis, Pap smear examination rolls an essential role in that cells of the human obtained of the patient of the cervix are analyzed of precancerous modification. The cells expert is performing the manual analysis of these cells and this work takes time-consuming and labor intensive. In this paper, we present a novel approach of obtaining a Nuclei cell segmentation method using a hybrid technique and evaluate this technique using the Mean Square Error (MSE) between the centroid from image segmentation and its original label. Two types of image data sets, benign and malignant images, are implemented in this work. The result obtained in this work shows a confident approach of segmentation for nuclei cell image. The average rational error of the MSE obtained in this work is about 11.361 %.

KEYWORDS: Nuclei Cell, Segmentation, Confocal microscopy, Clustering, image two-dimensional analysis, Dynamic Programming**.**

تجزئة صورة خلية النواة باستخدام تقنية هجينة

الملخص

صور الخلايا النووية (nuclear images) من الصور الحديثة والمهمة في الاستخدامات الطبية. ومن الخطوات المهمة امكانية التعامل مع هذه الصور وتجزئتها لدراسة كل جزء منها. وهناك حاجة لقياس جودة صور الخلايا النووية ولعملية التجزئة الأوتوماتيكي لهذه الصور. هناك دائماً الحاجة لتعزيز السرعة والدقة ومستوى الأتمتة والتكيف مع التطبيقات الحديثة لهذه الصور ومهم كذلك بالنسبة للتشخيص المبكر للسرطان ،اذ ان تجزئة الخلايا الآلية لنواة الخلية تعتبر حاسمة لأن خصائص نواة الخلية ترتبط بشكل أساسي بتقييم الورم الخبيث. في هذا المجال كان هناك عدد ضئيل من الأعمال البحثية حول التجزئة الالية للنواة الخلوية على صور الانصباب الجنبي الخلوي والتي يتم معالجتها بشكل غير طبيعي من خلال هذه التقنيات. اضافة الى ذلك لا تزال صورة الانصباب الجنبي الخلوية تواجه مشاكل بسبب مجموعة متنوعة من الخلايا والتباين السيئ للصور وتداخل الخلايا. في تشخيص سرطان عنق الرحم ، يعتبر دور فحص بطانة عنق الرحم يؤدي دورًا أساسيًا في تحليل الخلايا البشرية التي يتم الحصول عليها من عنق الرحم وان عملية التحليل اليدوي لهذه الخلايا من قبل خبير الخلايا لهذه الصور يستغرق وقتا طويلا. في هذا البحث، نقدم طريقة جديدة للحصول على طريقة تجزئة خلية النواة باستخدام تقنية هجينة ويتم تقييم هذه التقنية باستخدام الخطأ التربيعي المتوسط (MSE) بين النقطه الوسطى للصورة المقسمة والتقسيمة الأصلية. يتم تنفيذ الطريقة المقترحة على نوعين من مجموعات بيانات الصور ، الصور الحميدة والخبيثة (benign and malignant)، وان النتائج التي تم الحصول عليها في هذا البحث نتائج جيدة من حيث تجزئة صورة الخلية النووية. اذ يبلغ متوسط مربع الخطأ التي تم الحصول عليه في هذا العمل حوالي ١١,٣٦١٪ والذي يعتبر ذو نسبة قليلة مقارنةً بالاعمال المقاربة لهذا العمل.

I. INTRODUCTION

In multicellular organisms, cells communication is a main to the evolution and the balance of living organisms and tissues within an organism. additionally, Changes communication from one cell to another several processes of related diseases, like tumorigenesis. Efforts for understanding the quality these contacts are downstream impacts in cells attitude is common and is predominately the optical microscope was performed from fluorescence samples classified to the interest of specified molecules [1]. the delineation from the overlapped cells or cells nuclei is the one of the maximum enjoyable and challenging issues in the automated analysis of microscopic images. Especially in a case for well-known Pap, the overlapping regions of the cell in

microscopic slides are an extremely popular phenomenon Pap smear. The automated detection of the nuclei positions in nuclei images has been successfully pointed out by [2] and [3].

In last decade, different automated microscopic cellular image analysis algorithms have been presented [4]-[5]. Especially, different methods to segment the nucleus exist in literature most of them were utilized a semi-automatic method or manually into get more precise results of segmentation [6]. Yet, interaction of the user is prevented analysis of the image of the automated cell [6], [7]. Thus, promoting an advanced and the nucleus is uncensored. segmentation techniques are needful into to emphasize the successful automatic cell of analysis of the image. In spite for different nuclei cell segmentation methods in the market, but still, only a few works have been implemented for segmentation and clustering techniques together.

In this work, we present a novel algorithm using a novel segmentation method using Otsu thresholding and Poisson equation. Also, five different clustering techniques are used to cluster segmented nuclei images then to evaluate these techniques and select the optimum one. The rest of this paper is organized as follows: a literature survey corresponding to this work is presented in Section II. Section III represents the Methodology. Section IV expresses the MSE evaluation method. In Section V, result and discussion are presented. Finally, a conclusion is offered in section VI.

II. LITERATURE SURVEY

The authors of [8] presented an analysis algorithm of automatic quantitative image of breast cell histopathology (BCH) images. They applied top-bottom hat transform for the nuclei segmentation to improve image quality. To obtain regions of interest (ROIs) thereby realizing precise, multi-scale region-growing (WDMR) and location wavelet decomposition are combined. Double-Strategy Splitting Model (DSSM) including morphology of adaptive mathematical and Curvature Scale Space (CSS) detection of the corner approaches are implemented into separate nested cells of best precision and strongest. Based on color spaces and for the cell nuclei classification, four features based on the shape and 138 features based on textual are extracted. Then, they obtained an Optimum feature set through Chain-Like Agent

Genetic Algorithm (CAGA) with Support Vector Machine (SVM). The created algorithm was examined on 68 BCH images that contain more than 3600 cells. The results of empirical proof the sensitivity of mean segmentation is 91.53% (74.05 %) and specificity is 91.64%. The performance of classification for malignant and natural images cell can realize 99.05% (70.27%) of sensitivity, 93.33% (70.81%) of specificity and 96.19% (70.31%) of accuracy.

The authors [9] proposed an enhanced nucleus segmentation algorithm implementing Fuzzy C Mean (FCM) clustering method and Back Propagation Neural Network (BPNN). The existent method depending in FCM cluster was enhanced through obtaining optimal cluster rather than fixed clusters. moreover, from each region, features based on shape that act as input to the BPNN have extracted regions classify such as non-nuclei or nuclei. Therefore, the detected regions false have eliminated to produce the exact division for the nuclei. The proposed methods are estimated on the public Herlev dataset available. Empirical outcomes proof Performance improvement (accuracy, retrieval and dice factor) for segmentation of nucleus of 1%, 7%, and 5%, respectively, compared to current work. The optimal selection for segmentation of clusters, proposed algorithm implemented a threshold parameter. BPNN learning discards the false detected nucleus regions through example negative and positive.

In 2017, Alex Skovsbo presented a framework for ROI including either benign or cancerous colon tissue taken away of the entire slide H & E color images implementing features of cell nuclei. The author extracted an aggregate of 1,596 ROIs of 87 full spots H & E (43 cancers and 44 benign). A segmentation algorithm of nuclei of the cell including of color disintegration, local adaptive thresholding, k-means cluster, and separation of cell is implemented within the ROIs to take away feature of nuclei of the cell. An aggregate of 750 features and texture were based on the intensity taken away from the segmented cell nuclei structures for the ROIs classification. The nine features of most cell nuclei were discriminatory utilized in an irregular categorization of forest to specify if the ROIs included cancer or benign tissue. Classification of the region of interest acquired the Area Under Curve (AUC)

of (0.88%), the specificity of (0.96%), the sensitivity of (0.92%), and accuracy of (0.91%) implementing a threshold modified. The framework has evolved proof of promising outcomes in implementing feature of nucleus cell for classifying region of interest to contain cancer or benign samples of the tissue in H & E stained tissue [10].

In 2017, Khin Yadanar Win and King Mongkut proposed a method for automated fragmentation of nucleus of cell in images of cytology is a pleural effusion that contains tactile and interrelated cells. First, the preprocessing step is executed to minimize noise and improve the contrast implementing CLAHE and the median filter respectively. K-Mean Cluster algorithm implementing, the nuclei of the cell are segmented in the laboratory color space. Then, the border is adjusted, and the region of non-nuclei is removed through the process of morphological. At the end, the nested cell nucleus is separated according to subsequently, method of watershed, boundary of separated cell nucleus is predicted utilizing ellipse fitting technique. The proposed algorithms are evaluated on the local dataset that contains 35 images of pleural cytology with benign and natural cancer cells. The results of experimental yield precision accuracy = 0.90, F = 0.89, call = 0.89, dice similarity = 94% and Piccard = 89%, respectively. The conclusion is compared and confirmed with the truth images manually described through the experts [11].

In the studies above, there is no clear optimum segmentation method that can be applied for nuclei cell to classify the cells into normal and malignant type.

III. METHODOLOGY

As we mentioned in the literature survey section about the main drawback of the created systems, we overcome this drawback using a novel segmentation method and we evaluate our method using the MSE value between the segmented image and its label

We implemented different steps to perform an optimum analysis of nuclei cell image as shown in algorithm I and the following section. Figure 1 presents the input image.

A. The data set

The experimental data set is (UCSB-58) obtained from David Rimm Laboratory in Department of Pathology at Yale University of California, Santa Barbara. The data set is selected from the UCSB including 58 set of histopathological images 200 × 200 size of pixels consist of 63 benign and 48 malignant tissue images. The size of pixel in this image is 896 × 768. Additionally, each of this UCSB-58 data set images are corrupted labeled by pathologists to become 200 × 200 pixels images with ground truth. Figure 1 presents the input image.

B. Preprocessing

Usually, Digital preprocessing images need to eliminate unfavorable features, noise, and validate illumination artifacts. The pre-processing contains three steps: contrast improvement, adaptive of threshold, morphological filter. A global threshold is not able to create precise binary image outcomes due to the divergence background and presence of the dark nucleus, Therefore, the adaptive threshold created in [12] was utilized. We calculate the adaptive threshold of every pixel as shown in Eq. 1.

$$T(x,y) = mean_{|i-x|\le w,\ |j-k|\le h}\, I(i,j) \qquad (1)$$

where *h* and *w* are the height and weight for local window. In this work, we set the size of the local window experimentally So that the size of the window is larger than the size of the biggest cell. The tiny fractions of noises are removed using the opening of the morphological process (radius equal five pixels)), and the holes on the nuclei is filled implementing the closing of morphological process (radius equal five pixels) [13]. Figure 2 presents the input image in gray scale level.

C. Nuclei segmentation

To cope with the nuclei segmentation, a two-step segmentation procedure is proposed. In the first step, Global threshold using Otsu's method, Poisson

equation for image reconstruction. In the second step, we compute mean square error between ground truth image and image segmentation.

1) **Otsu Thresholding:** Otsu's thresholding technique gathers the difference between groups as a standard into choose the optimum threshold depending on the histogram of gray-level image [14]. This technique is citation as an efficient and widespread threshold method implemented in tasks of the real threshold. Provided an image f (x, y) acted through (L) gray levels, the average living image g (x, y) may be performed through gray levels as shown in Eq.2. below.

$$g(x,y) = \frac{1}{2 \times m + 1} \sum_{\Delta x=-m}^{m} \sum_{\Delta y=-m}^{m} f(x + \Delta x, y + \Delta y) \qquad \textbf{(2)}$$

For this work, *m*=1.

Therefore, for every pixel in the image of original, we may acquire a pair (i, j), that is collected according to the original density i and the mean density j. Let f_{ij} f displays the pair frequency (i, j) shown in the image, let M express the size of the image, and can calculate the probability of the joint as in Eq.3.

$$P_{ij} = \frac{f_{ij}}{M} \qquad (3)$$

Where, $\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} f_{ij} = M$ and $\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{ij} = 1.\, P_{ij}$ is the 2D image histogram. View higher than 2D histogram is presented by Fig.1. It overlays a square area for size L × L. The x-coordinate (i) denotes gray level and the y-coordinate (j) denotes gray level local average. Presented a pair arbitrary threshold (s, t), the 2D histogram is clustered into four regions.

Algorithm 1 Segmentation of Nuclei Cell

1: **procedure** INPUT IMAGE
2: convert input image to gray scale
3: **end procedure**
1: **procedure** IMAGE PREPROCESSING
2: contrast enhancement
3: adaptive threshold
4: morphological filtering
5: **end procedure**
1: **procedure** NUCLEI SEGMENTATION
2: Otsu thresholding
3:define the structuring element
4: Poisson equation

$$-a\nabla^2 u = f \text{ in } \Omega$$

5: **end procedure**
1: **procedure** EVALUATION METHOD
2: MSE

$$MSE = \mathbf{E}\left\{ \left\| \hat{h} - h \right\|^2 \right\} = \mathbf{E}\left\{ \hat{h}^T \hat{h} \right\} - 2\mathbf{E}\left\{ \hat{h}^T \right\} h + h^T h$$

3: **end procedure**
1: **procedure** EVALUATION METHOD
2: Rational error

$$r_e = \frac{MSE_{lab} - MSE_{seg}}{MSE_{lab}} \quad 100\%$$

3. **End Procedure**

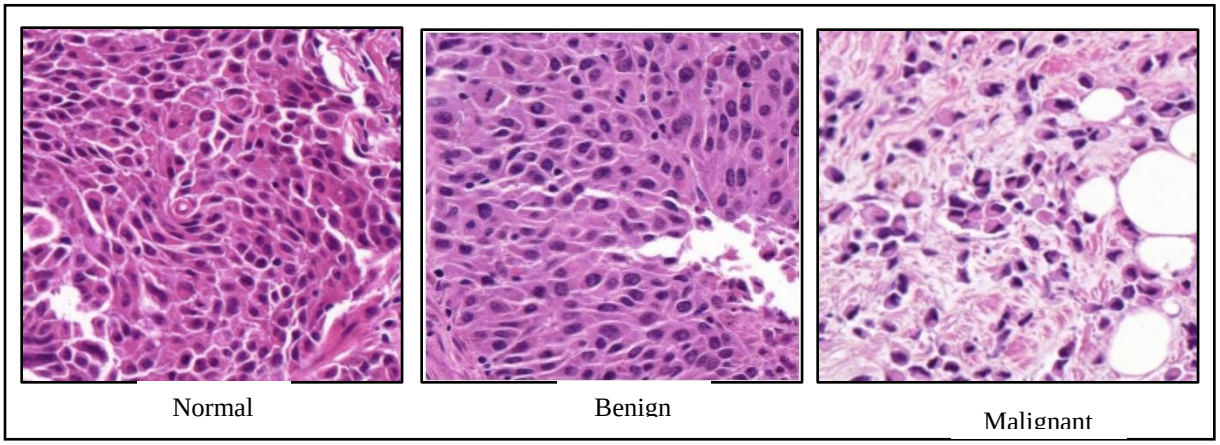

Fig. 1. Input images

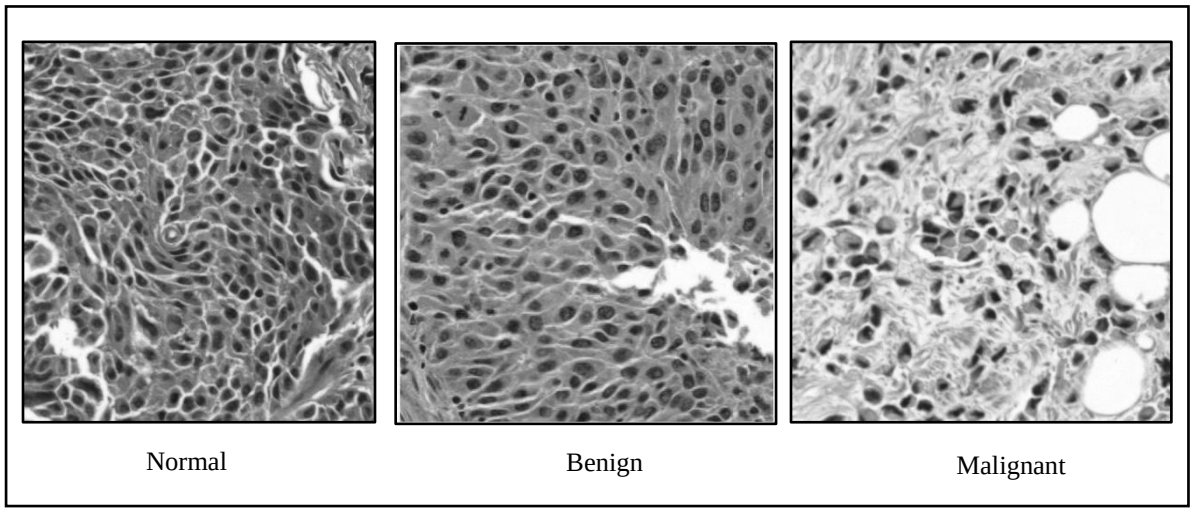

Fig. 2. Input images in gray scale

areas A and C refer to object and background respectively, and areas B and D refer to noise and edge, respectively. Let the C0 and C1 groups represent the object and the background, then the probability of two sets may be presented as shown in Eq.4.

$$w_o = \sum_{i=0}^{s-1}\sum_{j=0}^{t-1} P_{ij} \quad , \quad w_1 = \sum_{i=s}^{L-1}\sum_{j=t}^{L-1} P_{ij} \quad (4)$$

The two clusters density means a vector of the value can be calculated as shown in Eq.5, eq.6, and Eq.7.

$$\mu_0 = (\mu_{00}, \mu_{01})^T = \left[\sum_{i=0}^{s-1}\sum_{j=0}^{t-1} \frac{iP_{ij}}{w_0}, \sum_{i=0}^{s-1}\sum_{j=0}^{t-1} \frac{jP_{ij}}{w_0}\right]^T \quad (5)$$

$$\mu_1 = (\mu_{11}, \mu_{10})^T = \left[\sum_{i=s}^{L-1}\sum_{j=t}^{L-1} \frac{iP_{ij}}{w_1}, \sum_{i=s}^{L-1}\sum_{j=t}^{L-1} \frac{jP_{ij}}{w_1}\right]^T \quad (6)$$

$$\mu_t = (\mu_{ti}, \mu_{tj})^T = \left[\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} iP_{ij}, \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} jP_{ij}\right]^T \quad (7)$$

most cases, the likelihood of getting away from the diagonal may be small. Therefore, it is easy to verify as shown in Eq.8.

$$w_0 + w_1 = 1, \mu t = w_0\mu_0 + w_1\mu_1 \quad \ldots(8)$$

Also, we define, the between-class discrete matrix as expressed in Eq.9

$$S_b = \sum_{k=0}^{1} [(\mu_k - \mu_t)(\mu_k - \mu_t)^T] \quad \text{… (9)}$$

The trance of discrete matrix is computed as shown in equation below.

$$r_{trance} = w_0 \left[(\mu_{0i} - \mu_{ti})^2 + (\mu_{0i} - \mu_{tj})^2\right] + w_1 \left[(\mu_{1i} - \mu_{ti})^2 + (\mu_{1i} - \mu_{tj})^2\right]$$

The optimum threshold of 2D Otsu method is the threshold that maximizes r_{trance} (S_b).

After the Otsu method, the strel MATLAB function is used to define the structuring element. Figure 3 shows the segmented image using Otsu method.

2) Poisson equation: In this paper, a gradient-based image completion algorithm by solving Poisson equation is presented. Poisson's equation expressed below in Eq. 10 is a partial differential equation of elliptic type with broad utility in mechanical engineering and theoretical physics. It arises, for instance, to describe the potential field caused by a given charge or mass density distribution; with the potential field known, one can then calculate gravitational or electrostatic field [15].

$$-a\nabla^2 u = f \text{ in } \Omega \qquad (10)$$

Where, Ω, f, a, and u denote an open subset of $\mathbb{R}^d$(set space) for d = 1, 2 or 3, a source term, the coefficient, a scalar potential respectively.

In two phases, our image completion method is proceeded. first, the angle maps of the harmed district are recreated utilizing a model put together a system which depends with respect to another fix coordinating basis consolidating both the shading and the inclination factors. Subsequent to indicating the angle maps, the fulfillment result is registered from the slopes by illuminating a poisson condition. Figure 4 shows the last step of the segmentation method using the Poisson equation.

IV. MEAN SQUARE ERROR (MSE) FOR EVALUATION

After obtaining a segmented image, we evaluate our system using the MSE method. In this method we compared the error in distance between the centroid of the segmented image and the centroid its ground truth. The equation 12 below expresses the MSE computation. We specified a threshold for the MSE to be as a criterion to deiced if the segmented image is accurate or not. This threshold is measured as an average of the total number of nuclei images in the database. The MSE could be written in a generic way as shown in Eq.10 [16].

$$MSE = \mathbf{E}\left\{ \left\| \hat{h} - h \right\|^2 \right\} = \mathbf{E}\left\{ \hat{h}^T \hat{h} \right\} - 2\mathbf{E}\left\{ \hat{h}^T \right\} h + h^T h \tag{10}$$

where, ĥ and h denotes the position of the centroid of the segmented image and its ground truth respectively.

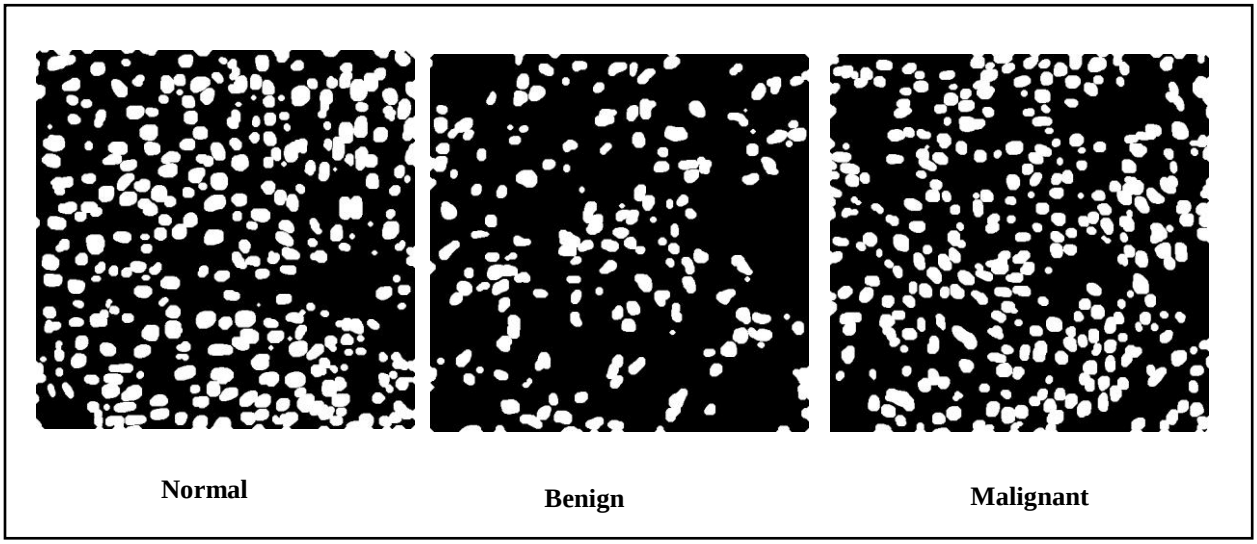

Fig. 3. Segmented image Global thresholding using Otsu

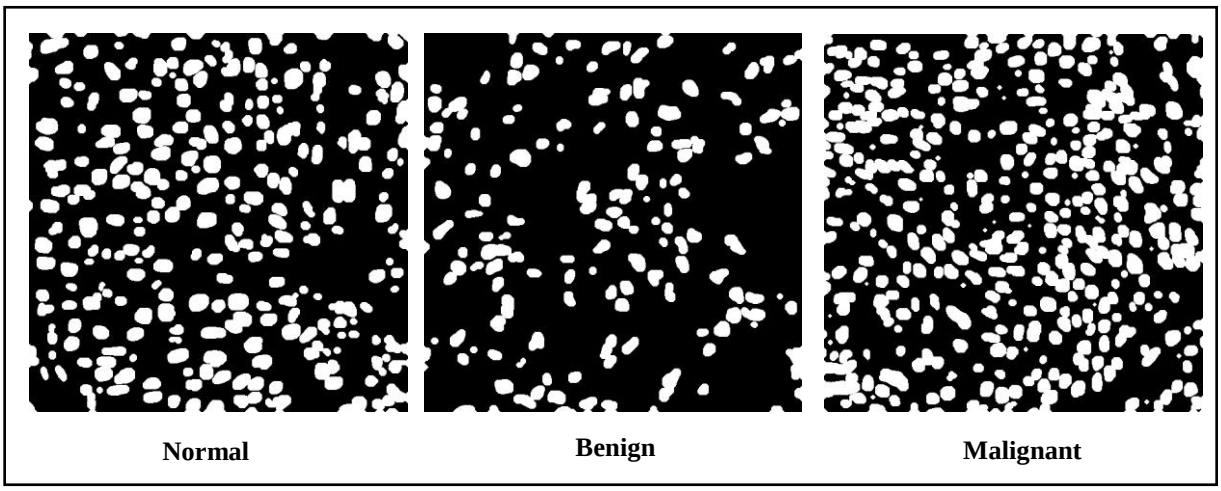

Fig. 4. Segmented image Poisson Equation.

V. RESULT AND DISCUSSION

In the segmentation of nuclei in histopathological images novel approach is created using the Otsu thresholding and Poisson equation. All benign and malignant images have been segmented and the centroid position of the segmented images is compared to centroid position of the original images extracted from their labels. The evaluation method used for comparing the segmented image and its label is the MSE method. To obtain an accurate result, we specify an MSE value as threshold.

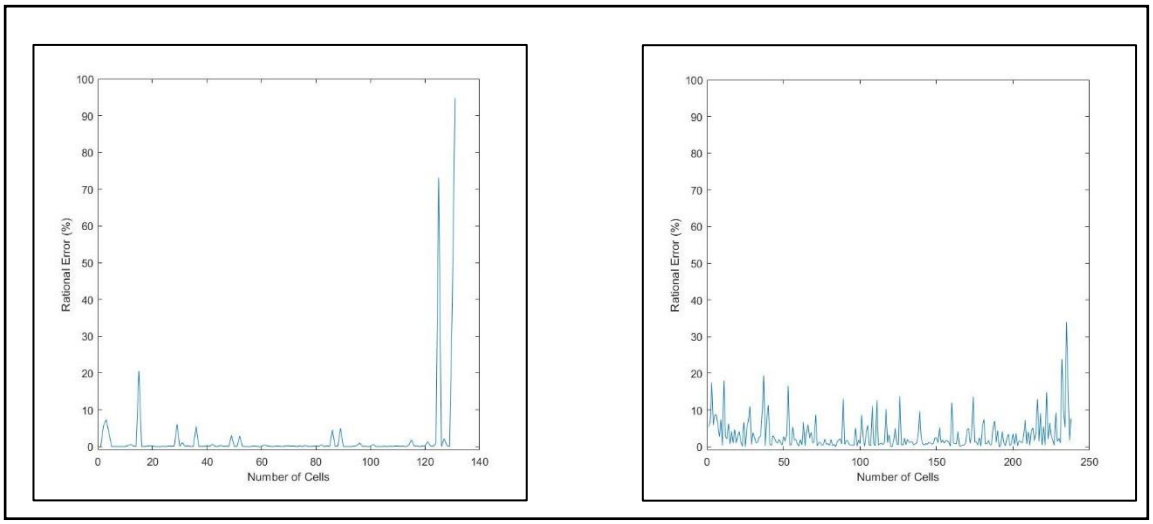

Fig. 5. The rational error between the benign and malignant images and their labels.

The value of this threshold is less than 10 units. So, if the MSE value less than 10, we consider the segmented cell equals to the cell in the label of the original image and the state is true, otherwise we consider the state is falls. In this work, we compute the rational error value of the MSE for the benign and malignant images as shown in equation 11 below.

$$r_e = \frac{MSE_{lab} - MSE_{seg}}{MSE_{lab}} \quad \textbf{100\%} \qquad \textbf{(11)}$$

Where, r_e, *MSElab*, MSE_{seg} denote the rational error, MSE of the label image, and MSE of the segmented image. The average of the rational error of the two types of the images is about 11.361 %. Figure 5 presents the rational error between the benign image and their labels. Figure 6 presents the rational error between the malignant image and their labels.

VI. CONCLUSION

In this paper, we created a novel segmentation algorithm using Otsu thresholding and Poisson equation to detect center of nuclei in histopathological images. To further explore nuclei in data set images, we evaluate the created algorithm using the MSE method to be used between the centroid position of the segmented and label image. Then, we compute the rational error between the MSE of the segmented image and the MSE of the label of the original image. In the final step, we compute the average of the rational error for the two types of images. From the obtained result, we proof that the created segmentation method outperforms most of the segmentation methods used for nuclei cell image.

REFERENCE

1) Dean P, Prabhakar R, Bradley S. Harris, Jason A, Karen J. Meaburn, Masa-Aki Nakaya, Terry P, Tom Misteli, and Stephen J, "**Segmentation of Whole Cells and Cell Nuclei From 3-D Optical Microscope Images Using Dynamic Programming**", IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 27, NO. 5, pp 723-734, 2008.

2) M. E. Plissiti, C. Nikou, and A. Charchanti," Automated **detection of cell nuclei in pap smear images using morphological reconstruction and clustering**", IEEE Trans. Inf. Technol. Biomed., vol. 15, no. 2, pp. 233-241, Mar. 2011.

3) M. E. Plissiti, C. Nikou, and A. Charchanti," Combining **shape, texture and intensity features for cell nuclei extraction in Pap smear images**", Pattern Recognit. Lett., vol. 32, no. 6, pp. 838-853, 2011.

4) X. Zhou, F. Li, J. Yan, and S. T. C. Wong," A **novel cell segmentation method and cell phase identification using markov model**", IEEE Trans. Info. Tech. Biomed., vol. 13, no. 2, pp. 152-157, Mar. 2009.

5) M. N. Gurcan, T. Pan, H. Shimada, and J. Saltz," Image **analysis for neuroblastoma classification: Segmentation of cell nuclei**", in Proc. IEEE Ann. Int. Conf. Eng. Med. Biol. Soc., pp. 4844-4847, 2006.

6) J. Cheng and J. C. Rajapakse," Segmentation **of clustered nuclei with shape markers and marking function**", IEEE Trans. Bio. Eng., vol. 56, no. 3, pp. 741-748, Mar. 2009.

7) X. Yang, H. Li, and X. Zhou," Nuclei **segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy**", IEEE Trans. Ciruits Syst. I: Reg. Papers, vol. 53, no. 11, pp. 2405-2414, Nov. 2006.

8) Pin Wang, Xianling Hu, Yongming Li, Qianqian Liu, Xinjian Zhu," **Automatic cell nuclei segmentation and classification of breast cancer histopathology images**", Signal Processing, 0165-1684, Elsevier, pp 1-13, 2015.

9) Bharti Sharma, Kamaljeet Kaur Mangat," An **Improved Nucleus Segmentation for Cervical Cell Images using FCM clustering and BPNN**", IEEE-ICACCI, 1924-1929, 2016.

10) Alex Skovsbo Jrgensen, Anders Munk Rasmussen, Niels Kristian Makinen Andersen," Using **Cell Nuclei Features to Detect Colon Cancer Tissue in Hematoxylin and EosinStained Slides**", Wiley Online Library, pp 785-793, 2017.

11) Khin Yadanar Win, Kazuhiko HAMAMOTO," **K Mean Clustering Based Automated Segmentation of Overlapping Cell Nuclei in Pleural Effusion Cytology Images** ", International Conference on Advanced Technologies for Communications, IEEE, pp 265-299, 2017.

12) Lindblad J, Wahlby C, Bengtsson E, Zaltsman **A. Image analysis for automatic segmentation of cytoplasms and classification of Rac1 activation**. pp 57(1):22-33, Cytometry A Jan;2004, [PubMed: 14699602].

13) Anoraganingrum D. **Cell segmentation with median filter and mathematical morphology operation.** Proc Int Conf Image Anal Process, pp 1043-1046, 1999.

14) Ningbo Zhu1, Gang Wang1, Gaobo Yang, Weiming Dai, "**A Fast 2D Otsu Thresholding Algorithm Based on Improved Histogram**", IEEE, pp 1-5, 2009.

15) Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe, **"Poisson surface reconstruction"** In Proceedings of the fourth Eurographics symposium on Geometry processing (SGP '06). Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 61-70, 2006.

16) Wackerly, Dennis; Mendenhall, William; Scheaffer, Richard L. **"Mathematical Statistics with Applications"** Belmont, CA, USA: Thomson Higher Education, 2008.