ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/378387765

HeLa cell segmentation using digital image processing

Article *in* Heliyon · February 2024

DOI: 10.1016/j.heliyon.2024.e26520

CITATION
1

READS
151

5 authors, including:

Edgar Duque
University of Guanajuato
4 PUBLICATIONS **2** CITATIONS
SEE PROFILE

Raul E. Sanchez-Yanez
University of Guanajuato
85 PUBLICATIONS **851** CITATIONS
SEE PROFILE

Noé Saldaña Robles
University of Guanajuato
61 PUBLICATIONS **314** CITATIONS
SEE PROFILE

Fabiola León
University of Guanajuato
43 PUBLICATIONS **1,038** CITATIONS
SEE PROFILE

All content following this page was uploaded by Raul E. Sanchez-Yanez on 28 February 2024.

The user has requested enhancement of the downloaded file.

Heliyon 10 (2024) e26520

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

HeLa cell segmentation using digital image processing

Edgar F. Duque-Vazquez [a], Raul E. Sanchez-Yanez [b], Noe Saldaña-Robles [a], Ma. Fabiola León-Galván [a], Jonathan Cepeda-Negrete [a,*]

[a] Universidad de Guanajuato DICIVA, Ex Hacienda El Copal km 9; carretera Irapuato-Silao; A.P. 311, Irapuato, 36500 Guanajuato, Mexico
[b] Universidad de Guanajuato DICIS, Carretera Salamanca - Valle de Santiago km 3.5 + 1.8 Comunidad de Palo Blanco, Salamanca, 36885 Guanajuato, Mexico

ARTICLE INFO

Dataset link: https://www.ebi.ac.uk/empiar/EMPIAR-10478/

Keywords:
Morphological operations
Digital image processing
Cancer
Nucleus
SBF-SEM

ABSTRACT

Computational cell segmentation is a vital area of research, particularly in the analysis of images of cancer cells. The use of cell lines, such as the widely utilized HeLa cell line, is crucial for studying cancer. While deep learning algorithms have been commonly employed for cell segmentation, their resource and data requirements can be impractical for many laboratories. In contrast, image processing algorithms provide a promising alternative due to their effectiveness and minimal resource demands. This article presents the development of an algorithm utilizing digital image processing to segment the nucleus and shape of HeLa cells. The research aims to segment the cell shape in the image center and accurately identify the nucleus. The study uses and processes 300 images obtained from Serial Block-Face Scanning Electron Microscopy (SBF-SEM). For cell segmentation, the morphological operation of erosion was used to separate the cells, and through distance calculation, the cell located at the center of the image was selected. Subsequently, the eroded shape was employed to restore the original cell shape. The nucleus segmentation uses parameters such as distances and sizes, along with the implementation of verification stages to ensure accurate detection. The accuracy of the algorithm is demonstrated by comparing it with another algorithm meeting the same conditions, using four segmentation similarity metrics. The evaluation results rank the proposed algorithm as the superior choice, highlighting significant outcomes. The algorithm developed represents a crucial initial step towards more accurate disease analysis. In addition, it enables the measurement of shapes and the identification of morphological alterations, damages, and changes in organelles within the cell, which can be vital for diagnostic purposes.

1. Introduction

Computational cell analysis of images, performed by computational algorithms, is a problem of great interest for the biotechnology and clinical sector. This is due to the significant knowledge gained by understanding cell behavior and morphology [1].

Manual cell detection is a labor-intensive process. The use of computer systems can remove complexity, especially for the laboratories that lack sophisticated equipment or staff without programming knowledge. Segmentation of cells using computational algorithms represents a significant challenge due to several factors. These factors include the identification of multiple objects of

* Corresponding author.
E-mail address: j.cepeda@ugto.mx (J. Cepeda-Negrete).

https://doi.org/10.1016/j.heliyon.2024.e26520
Received 29 May 2023; Received in revised form 28 November 2023; Accepted 14 February 2024

Available online 21 February 2024
2405-8440/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

cells in an image [2], the variety of images captured by different types of microscopy, the cellular phenotype, and the manipulation of heterogeneous shapes that undergo dynamic changes [3].

Cancer cells are of particular relevance in computational segmentation, and detecting their distinctive morphology compared to other cells can be crucial for early disease detection. Analysis of their external morphology and the nucleus shape is of great importance, as they are characterized by atypical sizes and shapes. In addition, the segmentation of their shape and nucleus is essential for those seeking to understand cancer biology and behavior.

The study of cancer typically involves the use of cell lines [4]. A cell line is a population of cells that have been adapted and immortalized through a virus that allows their continuous growth [5]. The HeLa cell line is one of the most commonly used for studying cancer, thanks to its ease of cultivation and constant cell division capacity [6].

Different types of algorithms have been used for the segmentation of HeLa cells, with deep learning algorithms being the most commonly used and the U-Net architecture being the most employed [7] [8]. Convolutional neural networks [9] [10] or image processing libraries such as FIJI, which is based on ImageJ [11], have also been used. The works that make use of deep learning in the segmentation of HeLa cells attempt to solve issues such as clustering [9], shape segmentation [12] [8], cell quantification [11], and segmentation of various organelles [10], or focusing particularly on one organelle, such as the segmentation of mitochondria [13].

Although the use of deep learning algorithms has gained great popularity in image segmentation, it is important to mention that its implementation requires significant computational power and large amounts of sample data. Therefore, there are alternatives such as traditional computer vision techniques, which are barely employed in HeLa cell segmentation. Few works where HeLa cells have been segmented using digital image processing. Chiang et al., [14] used phase contrast microscopy images and techniques such as Otsu, morphological operations, and the Watershed algorithm for the segmentation of vacuoles. Boukari and Makrogiannis [15] proposed a joint space-time diffusion approach and a region-based ensemble level optimization for moving cell segmentation.

In most studies where HeLa cells are segmented, the focus is on detecting the cell coverage without targeting any specific organelle. This represents an important opportunity for those who study cells by analyzing the structure of their organelles. Therefore, our work has taken the challenge of developing an algorithm to accurately and quickly segment a HeLa cell and its nucleus. In order to detail images of cellular structures, a microscopy like Serial Block Face Scanning Electron Microscopy (SBF-SEM) is required, which offers a detailed image quality that other microscopy techniques cannot provide.

Our work is closely related to that proposed in [16]. Where an algorithm using traditional computer vision techniques to segment HeLa cells using images captured from SBF-SEM was proposed. Such an algorithm segments cellular nuclei, and in another of their works, the authors use the obtained segmentation to reconstruct a nucleus in 3D [17]. The results of the nucleus segmentation show good accuracy. However, we have encountered difficulties in the segmentation when the shape of the cell and nucleus is complex. These difficulties can be addressed by using verification stages that ensure the accuracy of the segmented objects as nuclei. Additionally, developing an algorithm that approximates the shape of the cells more accurately would lead to better results in the segmentation of cells. Therefore, our work focuses on improving that limited segmentation and reducing the computational cost of the process.

The article is divided into four sections following this introduction. The next section details the methodology used to develop the proposed algorithm. Then, the evaluation of our proposal is presented and the results are discussed. Finally, the conclusions and perspectives of the work are exposed.

2. Methodology

The algorithm was developed following several general steps, which are depicted in the block diagram in Fig. 1. An overall description is as follows. First, the methodology used for cell segmentation is described, which consists of an algorithm that allows approximating the ideal shape of the original cell from a previously binarized and eroded cell. Then, the procedure used for nucleus segmentation is explained, along with the different verification stages proposed to ensure the accuracy of the outcome. Finally, the metrics used in the evaluation of the developed algorithm are introduced. Together, these stages allow for obtaining an accurate segmentation of the cell and nucleus, which is essential for the subsequent identification and analysis of the features of interest.

The methodology introduced below has been specifically designed to segment a HeLa cell and its nucleus, which are located in the center of an image. The dataset utilized in this work was obtained from the article published by Karabağ et al., [18] and consists of 300 images with their respective ground truth. The original images and ground truth are available at https://dx.doi.org/10.6019/EMPIAR-10478 and https://zenodo.org/record/4590903#.Y8cUNHbMKrw. The images were acquired using an SBF-SEM technology. Each image has a size of 2000×2000 pixels and its color format is grayscale.

The dataset is organized sequentially and in Fig. 2, three representative images are shown to depict this sequence. In the first image (Fig. 2a), a fragment of a cell is shown at the center of the image. In each subsequent image, the shape of the cell and the location of the nucleus are increasingly defined in the center of the image, until is reached a point where the cell and its nucleus are defined in the image (Fig. 2b). After this image, the following ones show how the shape of the cell disintegrates until it reaches a similar appearance to the initial state with which the sequence begins (Fig. 2c).

Before describing the methodology used, it is important to understand the concept of binary object segmentation, as it will be a key element throughout our explanation. Segmentation divides a grayscale digital image into regions or segments with meaningful characteristics. The result of this process is an image with two colors: white, which represents the segmented object, and black, denoting the background.

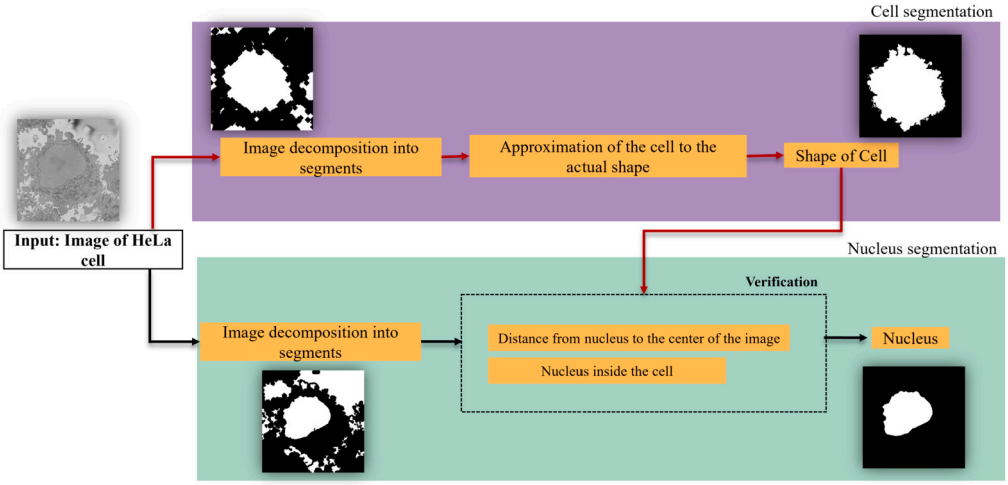

Fig. 1. Block diagram of the developed algorithm seen from a general perspective.

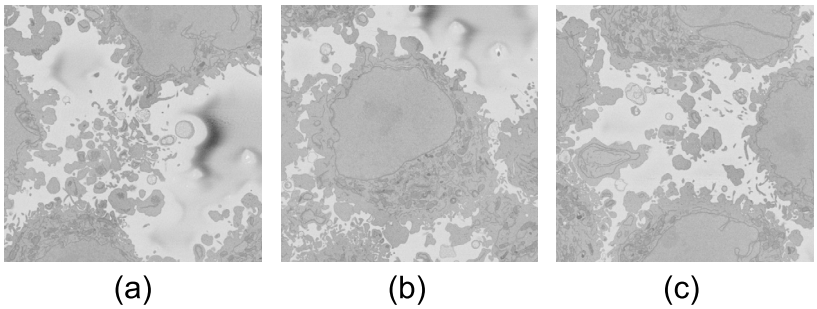

Fig. 2. Example of three images from the dataset used. (a) The first image in the dataset, showing fragments of a cell in the center. (b) A well-defined cell with a nucleus. (c) The last image in the dataset, and like (a), it shows fragments of a cell in the center.

2.1. Segmentation of cell shape

2.1.1. Image decomposition into objects

The first part of the developed algorithm consisted of cell segmentation (Fig. 3). To do this, we aimed to segment all cells present in the image (Fig. 3a) using the Otsu thresholding method (Fig. 3b). As a result, we obtained an image with two colors: white, representing the background, and black, representing the cell. The image obtained in this process was inverted, so we applied the negative of the image (Fig. 3c). To separate the cell from other cellular shapes present in the image, the morphological operation of erosion was used (Fig. 3d). Before erosion, the holes inside the cell of interest were filled (Fig. 3e). To perform erosion, a structuring element with the shape of a sphere was used to preserve the round geometry of the cell of interest. Since we worked with an image size of 2000×2000, a structuring element size of 58 pixels was selected. This value was determined through experimental testing, trying values in a range from 40 to 60 pixels. This value allowed for the separation of most cells from other cellular shapes with minimal change in their shape. It was found that a value between 56 and 60 resulted in a better separation of the cell of interest from other incomplete cells. And we took the number 58 as an intermediate value from the range that provided the best results in separating the cells (Fig. 3f). The previous process divided the image into several objects, one of which corresponds to the shape of the cell of interest. To distinguish the cell of interest from the other objects, the distance from the center of each object to the center of the image was measured. Considering these distances, the object closest to the center would be identified as the cell of interest (Fig. 3g).

Up to this point, the shape of the cell of interest has been obtained after being eroded. Now, it is necessary to restore the original shape of the cell. To achieve this, the eroded image was dilated in such a way that its size slightly exceeded the size of the original cell. This slight expansion allowed us to obtain the membrane of cell. This process will be explained in more detail in the following section.

2.1.2. Approximation of the cell to the actual shape

The goal of the algorithm developed in this stage was to approximate the edges of the dilated cell from the previous step to the edges of the original cell. This procedure was crucial to obtain a realistic structure of the membrane.

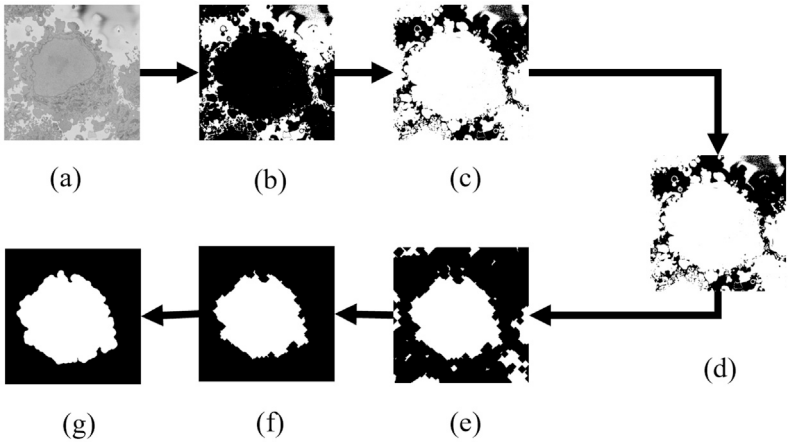

Fig. 3. Process carried out in the division of the cell of interest and the other objects. (a) Input image, (b) Thresholded, (c) Negative, (d) Fill holes in the image, (e) Eroded image, (f) Delete peripheral, (g) Dilated image.

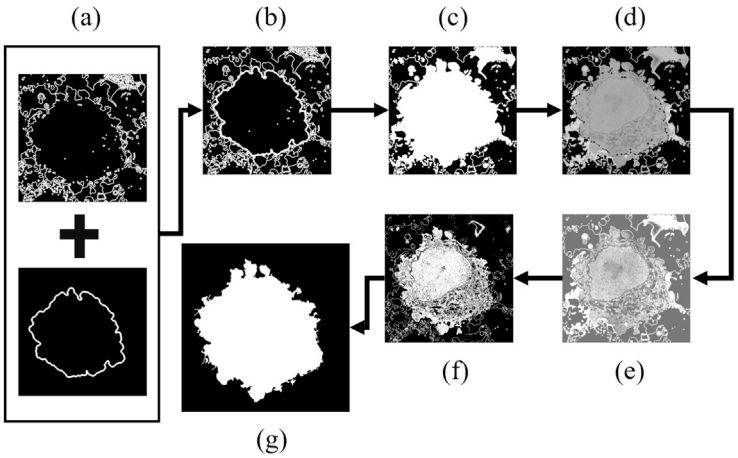

Fig. 4. Approximation of the cell to its actual shape using image processing techniques. (a) Sum of the edges of the original thresholded image and the obtained eroded image, (b) Image obtained from (a), (c) Hole-filling applied to (b), (d) Original image overlaid onto (c), (e) Multithresholding of (d), (f) Pixel limitation of (e), (g) Hole-filling applied to (f) and deletion of small and large objects.

The process involved the edges of both the binarized original image and the dilated image (Fig. 4a), and then fusing both images to obtain a more accurate approximation of the actual shape of the cell (Fig. 4b). Furthermore, a hole filling algorithm was used to complete any missing section (Fig. 4c). To further refine the process, the tonalities of the original image were overlapped on the previous image (Fig. 4d) and multithresholded into 5 levels (Fig. 4e). Based on the experimentation, the regions of interest correspond to levels 3 and 4. Therefore, we only preserve the tonalities associated with these two labels (Fig. 4f). Finally, a hole filling algorithm was used and objects that were far away from the center of the image were eliminated using the minimum distance (Fig. 4g).

2.2. Segmentation of nucleus

2.2.1. Image decomposition into objects

To obtain the segmentation of the nucleus of interest, several steps were applied to the original image (Fig. 5). Firstly, a standard deviation filtering was applied to enhance the edges of the nucleus in the image. Then, the Canny algorithm was used to detect the edges.

Next, the shape of the nucleus was obtained by inverting the colors of the image. To achieve this, the edges obtained by Canny were dilated to close the shape of the nucleus, and the colors of the resulting image were inverted. This process generated several objects, including the nucleus of interest that was sought to be segmented.

2.2.2. Verification steps

The verification procedure consisted of separating the nucleus of interest from other objects in the image, and for this two verification stages were proposed.

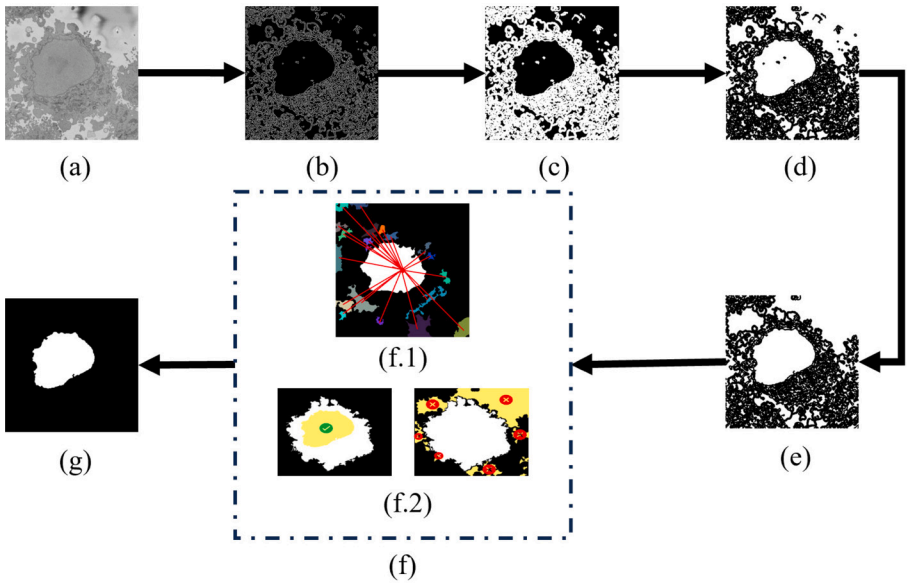

Fig. 5. Process carried out to segment the nucleus of HeLa cell. (a) Input image, (b) Std filter and Canny, (c) Dilate edges, (d) Negative, (e) Fill holes, (f) Verification steps (f.1) Distance from nucleus to the center of the image, (f.2), (g) Detected nucleus.

- Distance from nucleus to the center of the image: The first stage verified that the distance of the objects to the previously detected cell center was very short. To do this, the centroid of each object in the image was calculated, and the Euclidean distance from the center of the image was obtained. A representative value of 100 pixels was taken, and if the object had a greater distance, it was eliminated because it could not be a nucleus.
- Nucleus inside the cell: The second verification stage was developed considering that a nucleus should be inside a cell. In this stage, the pixel-to-pixel correspondence of each object in the image and those of the cell detected in the previous section was checked. For this, the coordinates of each pixel of the objects were obtained and compared with the value of the pixel of the cell in the same position. If the value of each pixel of the objects coincided in its entirety with the value of the pixel of the cell in the same position, it was determined that it was a nucleus and that it belonged to the cell.

2.3. Metrics

To evaluate the accuracy of the algorithm developed it is necessary to compare our outcome images against ground truth/reference images. In our study, four segmentation metrics were proposed: the Normalized Probabilistic Rand Index (NPR), the Jaccard index, the DICE coefficient, and the F1-score. The use of multiple metrics was proposed to eliminate bias in a comparative evaluation, as the use of a single metric can be influenced by specific factors of the dataset or methodology. In order to avoid possible manipulations, we chose to use the functions of the metrics implemented in MATLAB. This ensures objectivity and reliability in the comparison of the results obtained. However, in order to guarantee a rigorous and detailed presentation of the methodology employed, the metrics used are described below.

2.3.1. Jaccard similarity index

The Jaccard index quantifies the similarity between two sets by dividing the size of their intersection by the size of their union. In mathematical terms, it is represented by the equation (1).

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

Here, A and B denote two sets (representing the segmented regions of an image and the reference mask), and |A| and |B| indicate the number of pixels in each set [19].

2.3.2. Normalized Probabilistic Rand (NPR) index

The Normalized Probabilistic Rand (NPR) index gauges the similarity between two partitions of a dataset. The formula for NPR calculation is expressed as follows (refer to equation (2)):

$$NPR = \frac{I - I_E}{I_m - I_E} \tag{2}$$

Here, I is the actual index, I_E is the expected value of the Rand index, and $I_M - I_E$ represents the maximum possible difference between two indices [20].

2.3.3. DICE similarity index

Similar to the Jaccard index metric, DICE quantifies the similarity between two sets. The DICE index is determined as twice the intersection of two sets A and B (where A is a segmented image and B is the reference mask) divided by the sum of their sizes [21]. In mathematical terms, it is articulated as follows (see equation (3)):

$$DICE = 2 \cdot \frac{|A \cap B|}{(|A| + |B|)} \tag{3}$$

2.3.4. F1 score

The F1 score is a widely used metric for assessing the accuracy of a model in binary classification tasks. F1 amalgamates Precision and Recall parameters into a singular performance measure. Precision evaluates the ratio of true positive results among all positive results, while Recall gauges the ratio of actual positive cases correctly identified [22]. F1 is computed as follows (refer to equation (4)):

$$F1 = \frac{2 \cdot (Precision \cdot Recall)}{(Precision + Recall)} \tag{4}$$

Each of the values resulting from the described metrics varies between 0 and 1, where 1 indicates perfect segmentation (i.e., A and B are identical) and 0 indicates a complete lack of similarity.

3. Results

In this section, the results obtained after applying the methodology described in the previous section will be presented. First, the evaluation of the segmentation of the cell and its nucleus shape, carried out using the developed algorithm, will be shown. Finally, the results will be visually shown by figures, and described by tables. Together, these results will allow the analysis of the effectiveness and efficiency of the developed algorithm.

3.1. Evaluation of algorithm proposed

The evaluation of the cell and nucleus segmentation was done by comparing the results obtained from the proposed algorithm and the algorithm developed in Karabağ et al. [16]. Considering the entity responsible for providing the dataset, there is no record of other algorithms that have developed a similar solution. Our results were compared using the ground truth previously introduced. The similarity between the results of our proposal and the results of the method implemented in [16] was calculated using the aforementioned four metrics, namely Jaccard index, NPR, DICE, and F1, with the ground truth as the reference. The use of multiple metrics was done to evaluate the accuracy of our method in a more comprehensive and robust manner. Furthermore, using multiple metrics eliminates bias in the evaluation, as the use of a single metric may be influenced by specific factors in the dataset or methodology. To summarize the values obtained for each metric, four statistical measures were used: mean, median, standard deviation, and Z-test. The latter measure is used to show the significance of the presented results. Additionally, for better understanding and analysis, Figs. 6 and 7 were used to visualize the data obtained from the metrics. These figures display both the specific results of the metrics and the overall trends, which facilitates observation and enables a more detailed analysis of the performance of the algorithms in question. In the graphs, the trends of the proposed algorithm results are represented by a dashed blue line, while the results of the reference algorithm are shown with a dashed red line. The trends of both algorithms are also plotted, with the proposed algorithm displayed using a blue line and the reference algorithm with a different line.

3.2. Cell shape segmentation assessment

The Jaccard (Fig. 6a), NPR (Fig. 6b), DICE (Fig. 6c), and F1 (Fig. 6d) metrics were utilized to assess the degree of similarity between the resulting images and the ground truth. In general, in each graph of the Fig. 6, similar results can be observed. However, starting from image 150, a significant improvement in favor of the proposed algorithm becomes evident.

In the methodology, the sequence followed by the images was introduced, which can be summarized in three stages: the formation of the cell of interest in the center of the image (Slides 0 to slide 40 approximately), the fully formed cell (Slides 41 to slide 240 approximately), and the cell deformation again (Slides 241 to slide 300 approximately). In the interpretation of the results, these stages directly affect the results presented for both the proposed algorithm and Karabağ's algorithms, as both have difficulties detecting the first and third stages. Despite this, in segmentation can be observed in the third stage with the proposed algorithm, an improvement as reflected in the trend line.

Statistical analyses were conducted to summarize the values obtained for each of the metrics, which are presented in the Table 1. Our algorithm shows favorable results in each statistical measure: mean, median, and standard deviation, indicating better performance compared to the evaluated method. To determine the significance of the obtained results, the Z-test was applied. The test results were significant for all metrics, favoring the proposed algorithm. This allows us to confidently assert that our algorithm shows better performance compared to the reference method evaluated in the study.

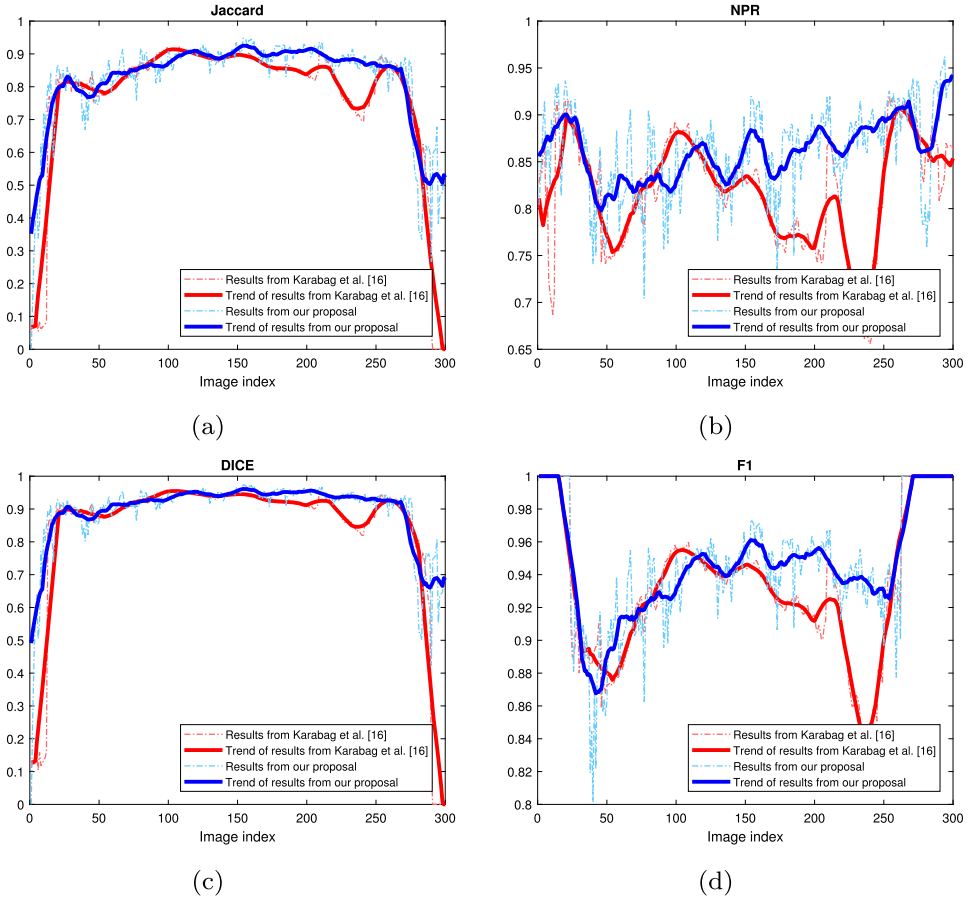

Fig. 6. Comparison of the results of the proposed algorithm for cell shape segmentation and Karabağ et al. [16] against the ground truth using a trend. a) Jaccard index, b) NPR, c) DICE, d) F1.

Table 1
Summary of the statistical results obtained in shape cell segmentation of the proposed algorithm and Karabağ et al. [16] against the ground truth using different similarity metrics.

Measure	Jaccard Index		NPR Index		DICE		F1	
	Karabağ	Proposal	Karabağ	Proposal	Karabağ	Proposal	Karabağ	Proposal
Median	0.8437	**0.8724**	0.8290	**0.8556**	0.9152	**0.9318**	0.9336	**0.9465**
Mean	0.7706	**0.8229**	0.8194	**0.8618**	0.8420	**0.8939**	0.9333	**0.9457**
Standard Deviation	0.2239	**0.1442**	0.0606	**0.0465**	0.2280	**0.1150**	0.0440	**0.0371**
z-test		✓		✓		✓		✓

3.3. Nuclei segmentation assessment

Unlike the cell shape segmentation, in Fig. 7a-d, the nucleus segmentation shows a very similar trend between the proposed algorithm and the reference algorithm in all four graphs. However, the trend line of our algorithm is slightly above the trend of the results of Karabağ et al., indicating that our results are closer to the ground truth reference.

Table 2 provides a summary using statistical measures of the similarity metric results. The comparison is between the proposed algorithm and the reference work and is focused on cell nucleus segmentation. When analyzing the values in Table 2, it can be appreciated that our algorithm shows better performance. This is due to its ability to detect poorly defined nuclei. The Z-test between the results obtained and those of the reference algorithm reveals a significant difference at a 95% level of significance. Therefore, we can conclude that the proposed algorithm emerges as an attractive alternative in nucleus detection.

3.4. Visual results

In statistical terms, we have verified that our algorithm is efficient segmenting both the nucleus and the shape of the cell. However, visual observation is also of great importance for understanding the contribution of this work.

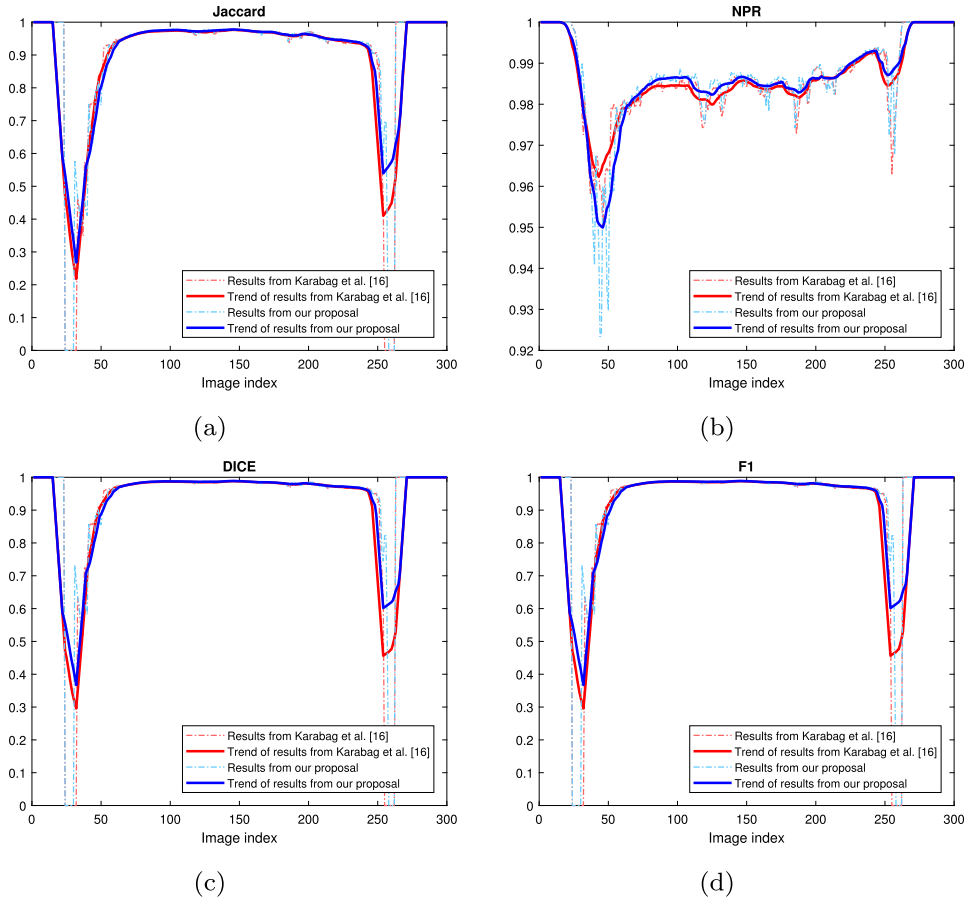

Fig. 7. Comparison of the results of the proposed algorithm for nuclei segmentation and Karabağ et al. [16] against the ground truth using a trend. a) Jaccard index, b) NPR, c) DICE, d) F1.

Table 2
Summary of the statistical results obtained in nuclei segmentation of the proposed algorithm and Karabağ et al., [16] against the ground truth using different similarity metrics.

Measure	Jaccard Index		NPR index		DICE		F1	
	Karabağ	Proposal	Karabağ	Proposal	Karabağ	Proposal	Karabağ	Proposal
Median	0.9689	**0.9702**	0.8290	**0.8556**	0.9802	**0.9809**	0.9842	**0.9849**
Mean	0.8891	**0.8977**	0.8194	**0.8618**	0.9079	**0.9191**	0.9119	**0.9231**
Standard deviation	0.2403	**0.2170**	0.0606	**0.0465**	0.2301	**0.2007**	0.2341	**0.2047**
z-test		✓		✓		✓		✓

Fig. 8 shows three examples where the results obtained by our algorithm are compared to with the ground truth. In the images shown, the segmentation of the cells is accurate for the most part. Whereas in the segmentation of the nucleus, the ground truth and outcomes are almost identical. Nevertheless, the difference between pixels is minimal.

The images were combined to obtain three colors: yellow, which represents the pixels where the ground truth and our result coincide (true positive); red, that represents false negatives, and green, that represents the result obtained by our algorithm but is not ground truth (false positives), and the background that represents True negatives (Figs. 9 and 10). In nucleus segmentation, our results closely resemble the ground truth. However, in some cases, the nucleus coverage does not fully match the ground truth. On the other hand, in cell shape segmentation, due to its irregularity, it is more noticeable that some membrane structures are missing in the segmentation. Nevertheless, it is evident that our result closely aligns with the reference truth.

3.5. Speed test

The speed of the algorithm proposed was evaluated by processing the 300 images provided in the dataset. These tests, along with the ones described earlier in the text, were conducted on a workstation with the following specifications: 48 GB of RAM, 2TB

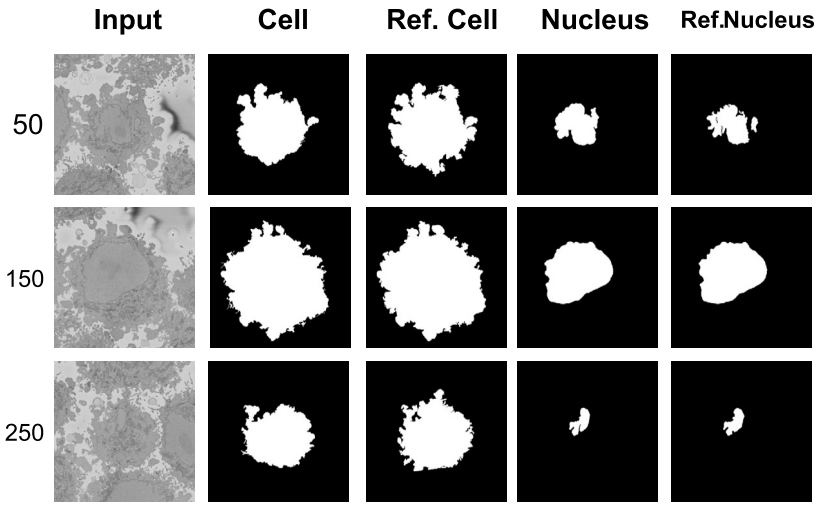

Fig. 8. Visual comparison of the results obtained by the developed algorithm against the reference truth.

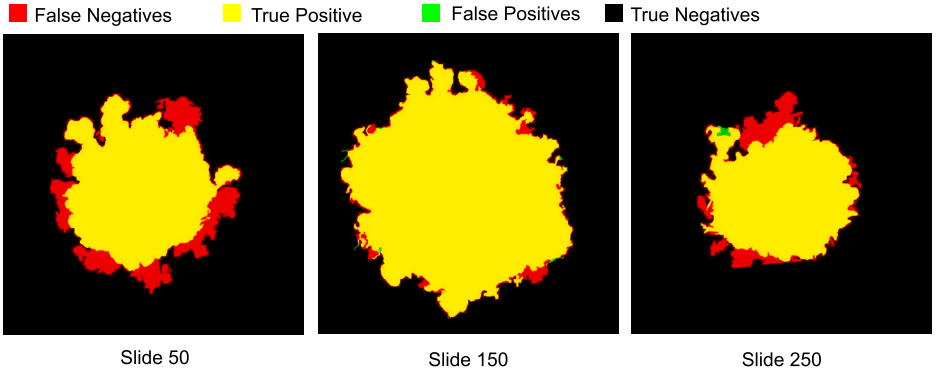

Fig. 9. Combination of the result obtained from cell segmentation overlaid onto the ground truth.

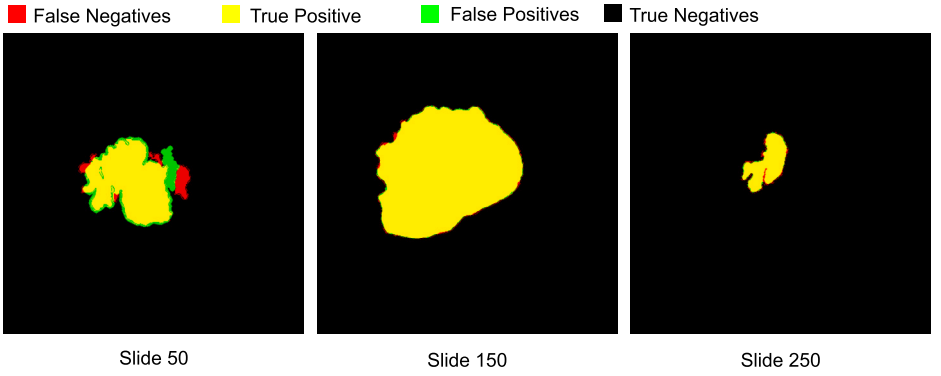

Fig. 10. Combination of the result obtained from nucleus segmentation overlaid onto the ground truth.

+ 3.5 TB hard disk, an NVIDIA GeForce RTX 2060 graphics card, an Intel Core i7-9700K processor, and a Windows 11 operating system, using MATLAB R2022b software. The built-in Matlab function *tic toc* was used to measure the speed. The results showed that our algorithm had a processing time of 818.85 seconds, while the reference algorithm took 1460.89 seconds. Our results demonstrated that our algorithm is highly competitive in terms of processing time and accuracy in segmenting the cellular shape and its nucleus. Importantly, our algorithm is significantly faster than the other algorithm, reducing the processing time to almost the half. Furthermore, our algorithm maintains good segmentation accuracy, making it a very attractive option for future applications.

4. Conclusion

This work successfully segmented the nucleus and shape of a HeLa cell using digital image processing. However, it is important to note that the algorithm currently developed only works for images obtained from SBF-SEM, in which the structure of cell and nucleus is well defined. It is possible that the algorithm may not perform effectively under different conditions.

It is important to highlight that there are still many unknowns to be resolved regarding the behavior of cancer cells exposed to different molecules. Cell segmentation is a starting point that can aid in cell interaction analysis. Therefore, it is essential not only segment the nucleus but also the other organelles within the cell. Obtaining accurate segmentation of these organelles ensures a more precise analysis and increases expectations for successful diagnosis in the future.

CRediT authorship contribution statement

Edgar F. Duque-Vazquez: Conceptualization, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. **Raul E. Sanchez-Yanez:** Writing – review & editing. **Noe Saldaña-Robles:** Visualization, Writing – review & editing. **Ma. Fabiola León-Galván:** Investigation, Conceptualization. **Jonathan Cepeda-Negrete:** Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no financial or personal conflicts of interest that could influence the results or interpretation of this study.

Data availability

The dataset used in this work is available at https://www.ebi.ac.uk/empiar/EMPIAR-10478/, and the code for the project can be found at https://github.com/efdv/HeLa-cell-segmentation-using-digital-image-processing.

Acknowledgements

Edgar F. Duque-Vazquez thanks the Mexican National Council on Humanities, Sciences and Technology (CONAHCYT) for the scholarship grant 1081409.

References

[1] Yuliang Wang, Zaicheng Zhang, Huimin Wang, Shusheng Bi, Segmentation of the clustered cells with optimized boundary detection in negative phase contrast images, PLoS ONE 10 (6) (2015) e0130178.
[2] Sotiris Dimopoulos, Christian E. Mayer, Fabian Rudolf, Joerg Stelling, Accurate cell segmentation in microscopy images using membrane patterns, Bioinformatics 30 (18) (2014) 2644–2651.
[3] Muhammad Farhan, Pekka Ruusuvuori, Mario Emmenlauer, Pauli Rämö, Christoph Dehio, Olli Yli-Harja, Multi-scale gaussian representation and outline-learning based cell image segmentation, BMC Bioinform. 14 (10) (2013) 1–14.
[4] Hyun Seok Kim, Yeo-Jin Sung, Soonmyung Paik, Cancer cell line panels empower genomics-based discovery of precision cancer medicine, Yonsei Med. J. 56 (5) (2015) 1186–1198.
[5] Rebecca Skloot, The immortal life of Henrietta Lacks, Broadway Paperbacks, 2017.
[6] John R. Masters, HeLa cells 50 years on: the good, the bad and the ugly, Nat. Rev. Cancer 2 (4) (2002) 315–319.
[7] Sherry Lin, Narges Norouzi, An effective deep learning framework for cell segmentation in microscopy images, 2021, pp. 3201–3204.
[8] Ali Ghaznavi, Renata Rychtáriková, Mohammadmehdi Saberioon, Dalibor Štys, Cell segmentation from telecentric bright-field transmitted light microscopy images using a residual attention u-net: A case study on HeLa line, Comput. Biol. Med. 147 (2022) 105805.
[9] Tim Scherr, Katharina Löffler, Moritz Böhland, Ralf Mikut, Cell segmentation and tracking using CNN-based distance predictions and a graph-based matching strategy, PLoS ONE 15 (12) (2020) e0243219.
[10] Larissa Heinrich, Davis Bennett, David Ackerman, Woohyun Park, John Bogovic, Nils Eckstein, Alyson Petruncio, Jody Clements, Song Pang, C. Shan Xu, et al., Whole-cell organelle segmentation in volume electron microscopy, Nature 599 (7883) (2021) 141–146.
[11] Mischa Schwendy, Ronald E. Unger, Mischa Bonn, Sapun H. Parekh, Automated cell segmentation in FIJI® using the DRAQ5 nuclear dye, BMC Bioinform. 20 (1) (2019) 1–9.
[12] Ehab Essa, Xianghua Xie, Phase contrast cell detection using multilevel classification, Int. J. Numer. Methods Biomed. Eng. 34 (2) (2018) e2916.
[13] Christian A. Fischer, Laura Besora-Casals, Stéphane G. Rolland, Simon Haeussler, Kritarth Singh, Michael Duchen, Barbara Conradt, Carsten Marr, Mitosegnet: easy-to-use deep learning segmentation for analyzing mitochondrial morphology, iScience 23 (10) (2020) 101601.
[14] Pei-Ju Chiang, Shao-Ming Wu, Min-Jen Tseng, Pin-Jie Huang, Automated bright field segmentation of cells and vacuoles using image processing technique, Cytometry, Part A 93 (10) (2018) 1004–1018.
[15] Fatima Boukari, Sokratis Makrogiannis, Joint level-set and spatio-temporal motion detection for cell segmentation, BMC Med. Genom. 9 (2) (2016) 179–194.
[16] Cefa Karabağ, Martin L. Jones, Christopher J. Peddie, Anne E. Weston, Lucy M. Collinson, Constantino Carlos Reyes-Aldasoro, Semantic segmentation of HeLa cells: An objective comparison between one traditional algorithm and four deep-learning architectures, PLoS ONE 15 (10) (2020) e0230605.
[17] Cefa Karabağ, Martin L. Jones, Christopher J. Peddie, Anne E. Weston, Lucy M. Collinson, Constantino Carlos Reyes-Aldasoro, Segmentation and modelling of the nuclear envelope of HeLa cells imaged with serial block face scanning electron microscopy, J. Imaging 5 (9) (2019) 75.
[18] Cefa Karabağ, Martin L. Jones, Christopher J. Peddie, Anne E. Weston, Lucy M. Collinson, Constantino Carlos Reyes-Aldasoro, HeLa cell images with four labels (nuclear envelope, nucleus, rest of the cell, and background) for deep learning architecture training, May 2020.
[19] Paul Jaccard, Étude comparative de la distribution florale dans une portion des alpes et des jura, Bull. Soc. Vaud. Sci. Nat. 37 (1901) 547–579.

[20] Ranjith Unnikrishnan, Caroline Pantofaru, Martial Hebert, Toward objective evaluation of image segmentation algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 29 (6) (2007) 929–944.
[21] Amos Tversky, Tversky index, Psychol. Rev. 84 (1977) 327–352.
[22] C.V. Rijsbergen, Information retrieval, 2nd ed, Butterworths, London, 1979, p. 115.

View publication stats