ORIGINAL PAPER

SARS-CoV-2: enhancement and segmentation of high-resolution microscopy images—Part I

Roberto Rodríguez[1] · Brian A. Mondeja[2] · Odalys Valdés[3] · Sonia Resik[3] · Ananayla Vizcaino[2] · Emilio F. Acosta[2] · Yorexis González[2] · Vivian Kourí[3] · Angelina Díaz[2] · María G. Guzmán[3]

Abstract

Possibly, and due to poor eating habits and unhealthy lifestyle, many viruses are transmitted to human people. Such is the case, of the novel coronavirus SARS-Cov-2, which has expanded of exponential way, practically, to whole world population. For this reason, the enhancement of real microscopic images of this coronavirus is of great importance. Of this way, one can highlight the *S-spikes* and visualizing those areas that show a high density, which are related to active zones of viral germination and major spread of the virus. The SARS-Cov-2 images were captured from nasopharyngeal samples of Cuban symptomatic individuals (RT-PCR positives for SARS-CoV-2) and processed via scanning electron microscopy. However, many times these microscopic images present some blurring problems, and the *S-spikes* do not look well defined. Therefore, the aim of this work is to propose new computational methods to carry out enhancement and segmentation of SARS-Cov-2 high-resolution microscopic images. The proposed strategy obtained very satisfactory results, and we validated its performance, together with specialist physicians, on a set of 1005 images. Due to the importance of the obtained results, this first work will be addressed to the application of the proposed algorithm. A second paper will deeply analyze the theory related to these algorithms.

1 Introduction

Probably, due to poor eating habits (for instance; ingestion of semi-processed or not very well-cooked foods) and unhealthy lifestyle (few hand hygiene, unnecessary crowding of people, and no use of masks, among others), many types of viruses are transmitted to humans. Thus, this novel virus, named coronavirus due to visual appearance —under electron microscopy— similar to a crown [1], is the cause of an infectious disease by severe acute respiratory syndrome (SARS-Cov-2) [2], and named by the World Health Organization (WHO): *COVID-19*. This disease has expanded of exponential way, practically, to the whole world population, and already today, it is considered a pandemic.

The fight against *COVID-19* has motivated many scientific institutions and researchers from all the specials in the world to seek effective methods and techniques to help end this pandemic. In this direction, the computer vision community has not been left behind, and many articles have been published, mainly for the improvement of computed tomography (CT) images using deep learning [3–7].

In this sense, it draws attention the small number of published papers for the improvement of microscopic images of the novel coronavirus SARS-Cov-2, when is known the blurring problems inherent in these types of images, which often makes it difficult to analyze correctly them. Therefore, the aim of this work is to propose new computational methods to carry out enhancement and segmentation of SARS-Cov-2 high-resolution microscopic images, which will allow the *S-spikes* to be highlighted and visualizing those areas that show a high density, which are related to active zones of viral germination and major spread of the virus.

✉ Roberto Rodríguez
   rrm@icimaf.cu

[1] Institute of Cybernetics, Mathematics and Physics (ICIMAF), La Habana, Cuba

[2] Center for Advanced Studies of Cuba (CEA), Valle Grande, Cuba

[3] Institute of Tropical Medicine ¨Pedro Kourí ¨ (IPK), Ciudad de La Habana, Cuba

It is known that the microscopy images are widely used for the diagnosis of various diseases, such as infections and tumors. These facilitate basic biomedical research that has been continuously generating new knowledges into the causes of human diseases, and for that reason, they are of great importance in improving of quality of health. However, the obtaining high-quality microscopic images are not always a trivial process, even in the best microscopes, due to physical phenomena that originate between the glass coverslip, light, and microscope optics. Some of these phenomena are scattering, out-of-focus, imperceptible vibrations, voltage disturbances, among others, which causes certain blurring in the microscopic images. Examples of some of them can be seen in [8–10].

With the rapid development of computer vision, where its application has covered practically all fields of modern life, and basically, digital image processing which has been widely utilized in the medical field, including structure segmentation and image enhancement, has provided great support for diagnosis and better medical prognosis [11, 12].

The importance of this study is that the new developed algorithms will be applied to the enhancement of microscopic images of the novel coronavirus SARS-Cov-2, which contributes, due to very little information appears in the literature on this issue, to the novelty of this work.

This first research will be addressed to application of the proposed algorithm, which it obtained very satisfactory results, and we validated its performance, together with specialist physicians, on a set of 1015 images. A second paper will deeply analyze the theory related to these algorithms, and we will carry out a quantitative evaluation of them. Here, by space problems are not possible.

The rest of the paper is organized as follows: In Sect. 2, the materials and methods are given. Section 3 slightly outlines some theoretical and algorithmic aspects. A second paper will deepen into these issues. Section 4 contains the obtained results and discussion. We describe our conclusions in Sect. 5.

2 Materials and methods

2.1 Medical methodology

We captured the samples from nasopharyngeal swabs collected from Cuban individuals with *COVID-19* symptomatic and RT-PCR positive for SARS-CoV-2, and we treated the same (samples) with an aldehyde solution and processed by scanning electron microscopy [13]. Finally, we completed a database of more than 1015 microscopic images.

The structure identification of the SARS-CoV-2 virus is essential toward understanding its historical resemblance to other coronaviruses (mode of infection, mechanism of

entry at tissue site of infection, and the replication process in the infected cells) [14]. Particularly, high-resolution microscopic studies are essential in order to identify the etiological agent of several outbreaks. In the case of SARS-CoV-2, the Scanning Electron Microscopy (SEM) images provide fundamental data of the structural aspects of the virus and can be a guiding point in therapeutic development [13].

The specialists inactivated clinical specimens for 12 h in a solution of 25% formaldehyde and 5% glutaraldehyde before microscopy study. Finally, we processed the inactivated samples by SEM. The coverslips were sputter-coated with gold (thickness of 10 nm) and viewed under the MIRA3-TESCAN scanning electron microscope at 10 kV [13].

### 2.2 Characteristics of studied images

The dataset has more than 1015 images, but by space problem, we only give eight examples. Figure 1 shows typical coronavirus images. Figure 2 shows other examples.

In Figs. 1 and 2, one can observe several notable characteristics of these microscopic images, which are common to typical microscopic images we encounter in the study of the novel coronavirus. For example, in Fig. 3, one can see a horizontal profile through images of Figs. 1b and 2d. This is a plot of the pixel intensities along a single row, where one can note:

1. The intensities are slightly clearer within the virus than in the local surrounding background (Observe in Fig. 3, the "y" axis).
2. One can observe high local variation of intensities both, within the coronavirus and within the background (see Fig. 3).
3. It is possible to see (Figs. 1 and 2) high variability of coronavirus in both, size and shape.
4. The contrast of intensities between the coronavirus and the background is low. This is evident from the intensities of histogram (see Fig. 4). This histogram is unimodal, reflecting the low separation in average intensities of the virus and background.

Due to the acquisition protocol and to the already mentioned physical phenomena, a lot of noise corrupts the microscopic images. One can see the ripples in the histograms and profiles (see Figs. 3 and 4), and the blurring that exists in all these images. Across the entire image, changes in intensity levels cannot reliably distinguish *S-spikes*, due to possible non-uniformity of the average background intensity and the low contrast between virus and background. However, within a region of microscopic images, changes in intensity levels can effectively distinguish *S-spikes*, since locally a virus is clearer than its surrounding background.

Fig. 1 **a**-**d** Some examples of original microscopic images of the novel coronavirus SARS-Cov-2
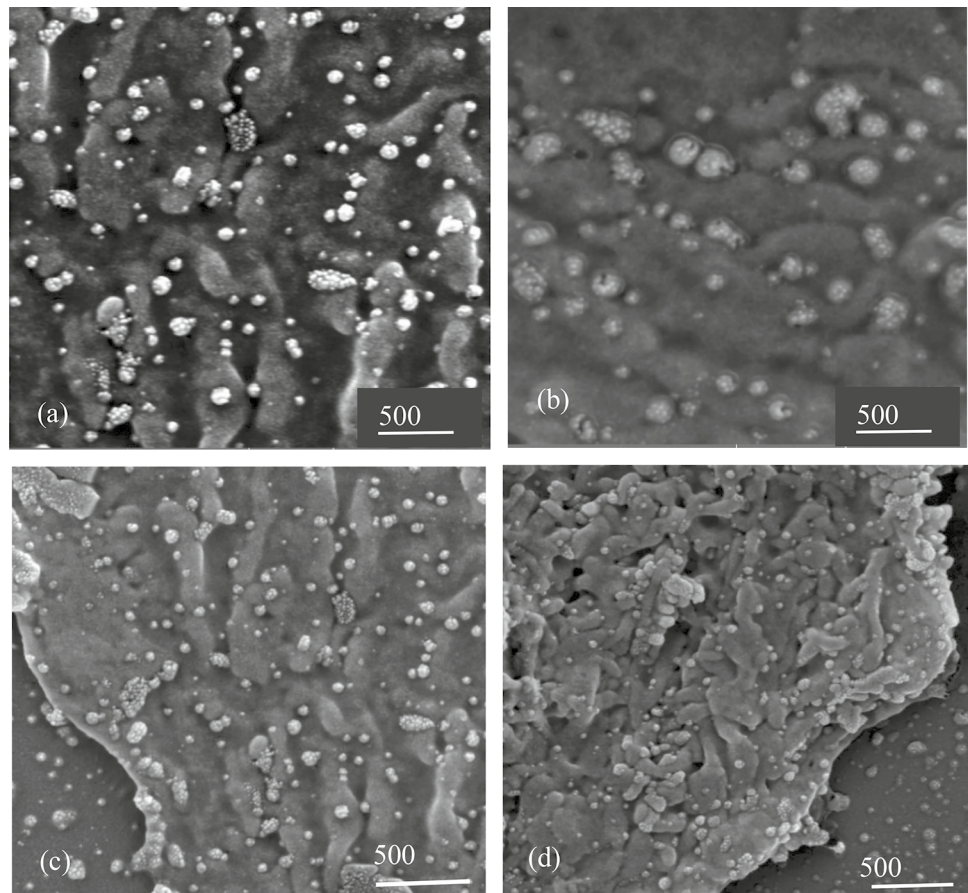
In other words, one can notice that coronavirus areas (the *S-spikes*) are hyper-dense regions.

The previous paragraph points out the need of working in the region of interest, that is, locally. This strategy ensures that one taken into account the correlation among pixels of the *S-spikes*, which is of great importance for getting good results, and in addition, it helps notably in the design of the algorithms [15].

In Fig. 5, we represent a *3D* projection of a region of interest (ROI) from Figs. 1b and 2d. One can see that the crests are smooth and peaks are not well distinguished, which it indicates the blurring existing in these microscopic images.

The characterization of the study images is of great importance, and a basic principle to have in mind. Many times, this determines the strategy to follow [16].

3 Some theoretical and algorithmic aspects

In this section, due to the importance of the obtained results, we only outline lightly some theoretical and algorithmic aspects. A second paper will deeply analyze the theory related to the proposed algorithm.

3.1 Some definitions

Definition 1 (Regional maximum) *A regional maximum at altitude h of grayscale image I is a connected component C of $T_h(I)$ such that $C \cap T_{h+1}(f) = \varphi$.*

We consider the successive thresholds $T_h(I)$ of I, for $h = 0$ to L-1 [16],

$$T_h(I) = \{p \in D_I / I(p) \leq h\} \qquad (1)$$

They constitute the threshold decomposition of I, where these sets satisfy the following inclusion relationship:

$$T_h(I) \subseteq T_{h-1}(I) \quad \forall \quad h \in [0, L-1], \qquad (2)$$

where L is the number of gray levels.

In Fig. 6, one can observe a representation of regional maxima in a grayscale image. In this case, we considered the regional maximum as the *S-spikes*, where one can see that are justly the hyper-dense areas (see Figs. 1 and 2).

Definition 2 (Grayscale reconstruction) *The grayscale reconstruction $\rho_I(J)$ of I from J obtained by iterating*

Fig. 2 **a-d** Other examples of original microscopic images of the novel coronavirus SARS-Cov-2
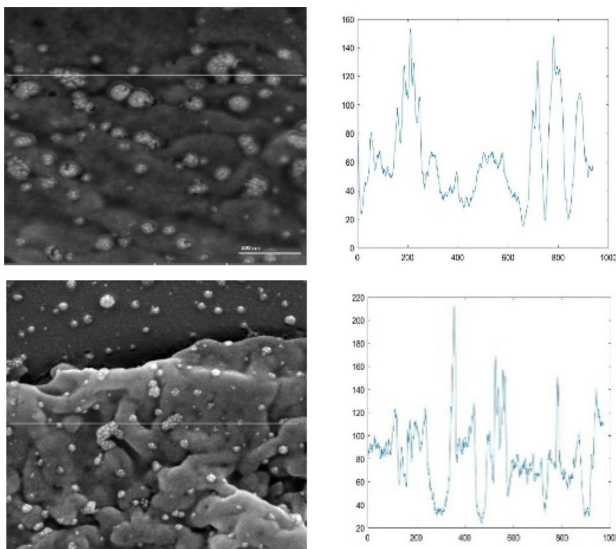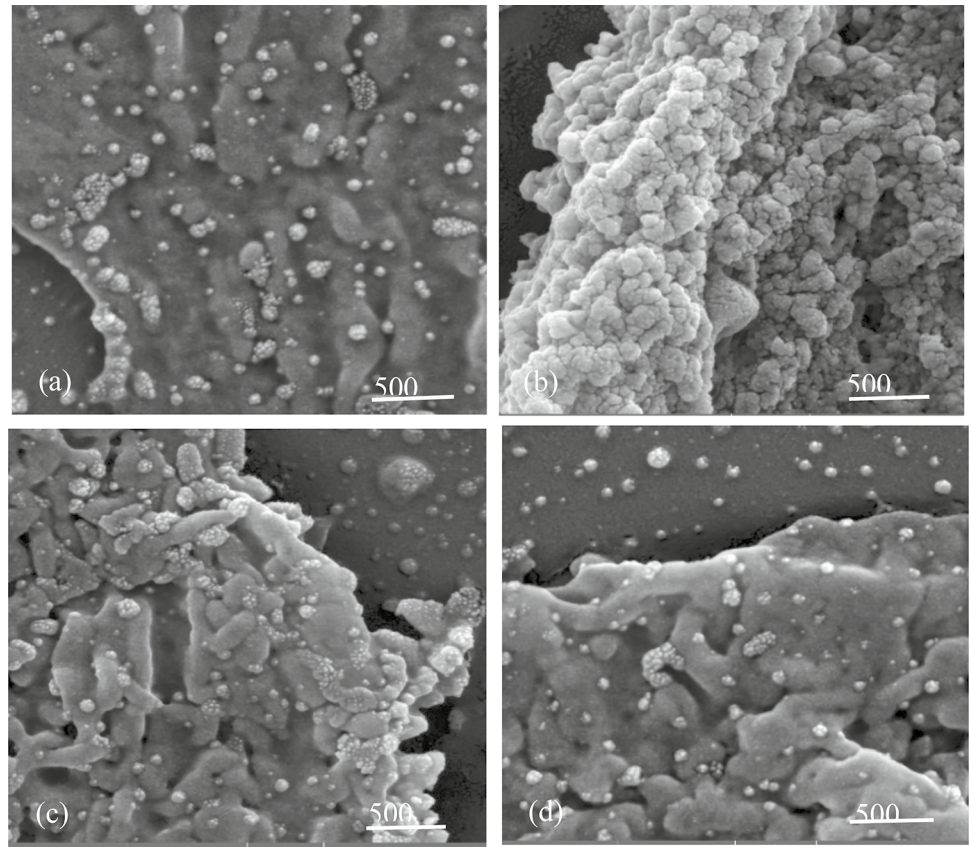

Fig. 3 An intensities profile through the images. We indicated the profile by a white line

grayscale dilations of J "*under*" I until stability is reached, that is,

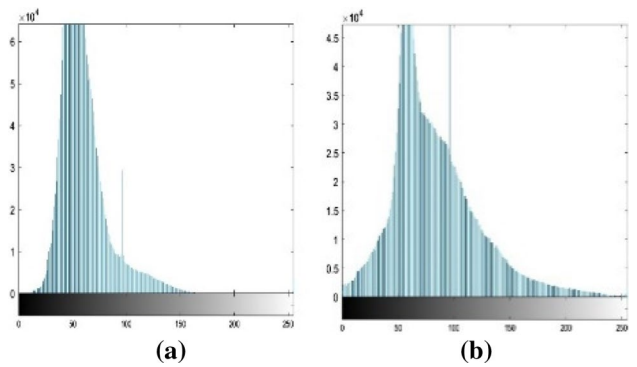$$\rho_I(J) = \bigvee_{n \geq 1} \delta_I^{(n)}(J) \tag{3}$$

Fig. 4 Histograms. **a** from Fig. 1, **b** from Fig. 2 d

Definition 3 *The h-dome image $D_h(I)$ of the h-domes of a grayscale image I given by*

$$D_h(I) = I - \rho_I(I - h) \tag{4}$$

The *h-dome* transformation extracts light structures without involving any size or shape criterion. The only parameter (h) related to the height of these structures. In the case of segmentation of the S-spikes of novel coronavirus, this parameter played a vital role.
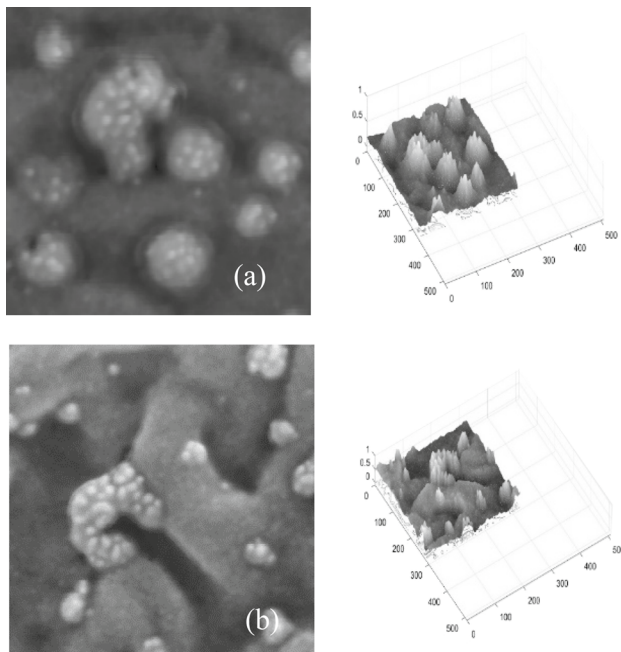
Fig. 5 **a** A ROI from Fig. 1 (b). **b** A ROI from Fig. 2 (d). Observe the blurring in the microphotographs, and the smooth of peaks
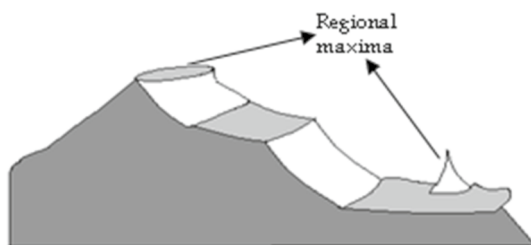

Fig. 6 Regional maxima in a grayscale image

3.2 A vision on the strategy

All researches carried out with the proposed algorithm worked locally. This was the principal characteristic of designed algorithms for segmenting and enhancement the microphotographs of the novel coronavirus.

In effect, when locally working the spatial information is taken in consideration (and implicitly the correlation among the pixels), and one obtains a major stability in performance of strategies. One of the problems that have the classical thresholding operators, it is to assume that the pixels are statistically independent [15].

Algorithm No.1

Data: *Original image **A**, parameter (h)*;
Result: *Enhanced or segmented image. Let **B** be*

    1. *Initialize: **h = 80, window size = 3x3, sigma (σ) = 1.5**,*
    2. *Pass Gaussian filter to the original image, or once the Mean-Shift Iterative Algorithm (MSHi) [16]. Let **C** be the result*
    3. *Subtract the **h** parameter from **C**. Let **D** be the result*
    4. *Accentuate high spatial frequencies: **E = D − C***
    5. *Apply an extreme filter to **E** [17]. Let **F** be the results*
    6. *If it is necessary, go to step 4. Let **B** be.*
    7. *Return **B***
   end

A deep explanation of the theory of the *Algorithm No.1* will appear in a second paper. Here, we will give a brief review of it.

The *Gaussian* filter was used to diminishing the noise in the original microscopic images. However, the *MSHi* algorithm has a performance more precisely [16], but it consumes more computational time. In the case of the *Gaussian* filter, the best performance was obtained using $\sigma = 1.5$. The used window size was 3×3. A larger dimensional window caused a loss of information in the microphotographs.

The value of the h parameter determines the processing characteristics. For a small value of h, the *S-spikes* were improved (more contrasted) in relation to the original image. While for a large value of h, the *S-spikes* were isolated (segmented), where the *S-spikes* appeared clear on a black background. The importance of improving or segmenting the *S-spikes* is because the areas that showed an *S-spike* high density are related to an active zone of viral germination [13].

4 Experimental results. Discussion

In computer vision, image enhancement and particularly segmentation are recognized as one of the most important steps in any image analysis system.

In Fig. 7, we show a first example of the obtained result of applying the *Algorithm No.1* for different values of h parameter on the image of Fig. 1a.

One can see (in Fig. 7) that all the processed images were enhanced; that is, contrast increased, and the *S-spikes* are evident and virion particles show the urchin-shape. Note that for small values of h, we obtained a better-contrasted image. However, as we increased the h value, the *S-spikes* begin to segment, where one can see white areas on a black background. It is possible to observe the edges more defined. Of this way, we can to carry out an automatic analysis of the density of *S-spikes* (number of *S-spikes*/ area unit); that is, automatic count of white points on a black background area.
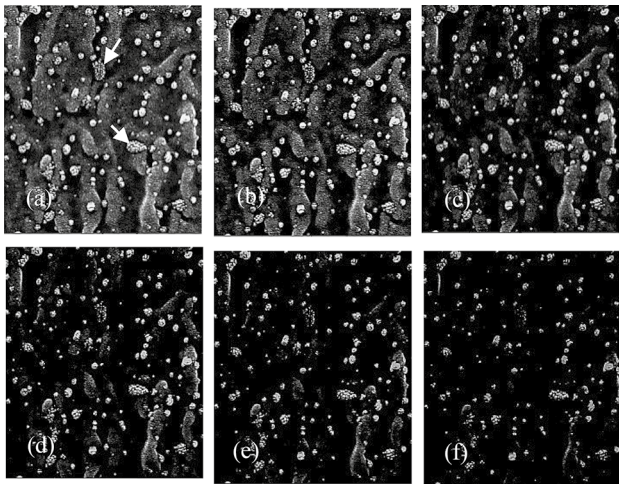
Fig. 7 Application of the *Algorithm No. 1* on the image of Fig. 1 (a) at different values of the *h* parameter. **a** $h=15$, **b** $h=30$, **c** $h=60$, **d** $h=80$, **e** $h=100$, **f** $h=140$. In (**a**), the arrows show white dots, which representing the *S-spikes*

Fig. 9 Application of the *Algorithm No. 1* on the image of Fig. 1 (b) for three values of *h* parameter. **a** $h=15$, **b** $h=60$, **c** $h=80$. We show the original image for a better visual comparison
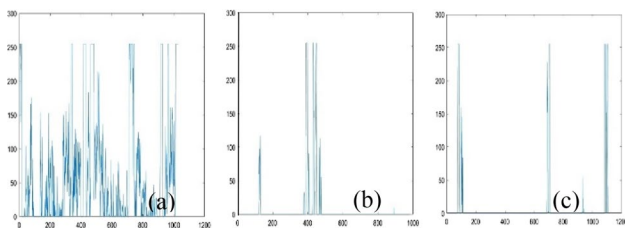

Fig.8 An intensity profile through a processed image from Fig. 3. **a** $h=15$, **b** $h=80$, **c** $h=140$. Observe that the peaks of profiles begin to flattening (see profiles of Fig. 3)

When analyzing the profiles of the processed images there are remarkable differences with those of Fig. 3. In effect, one can observe in Fig. 8 that the peaks of the profiles, for different *h* values, begin to flattening. In other words, peaks with parallel lines to the "*x*" axis are indicative of equal intensity levels (segmented zone). In addition, distanced peaks show the isolation of the *S-spikes* in the different areas where the coronavirus is. The intensity of background reached zero value.

Moreover, on Fig. 7, one can have the following question: Which image to choose as the best-processed image? The answer to this question is not direct, because it depends on the aim of observer or from the application itself; that is, segmentation is heavily dependent on the application [18]. A theoretical analysis of the influence of *h* parameter will appear in an upcoming publication. Figure 9 shows the obtained result of processing the image of Fig. 1b for three values of *h* parameter.

Note that the algorithm has very similar behavior to the result of Fig. 7, which shows the stability of this strategy.
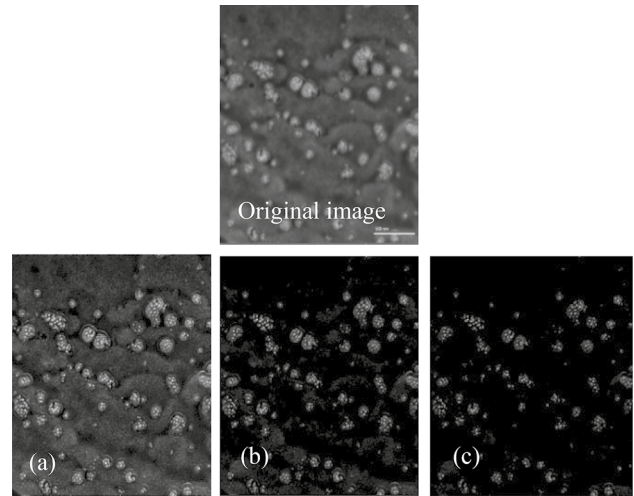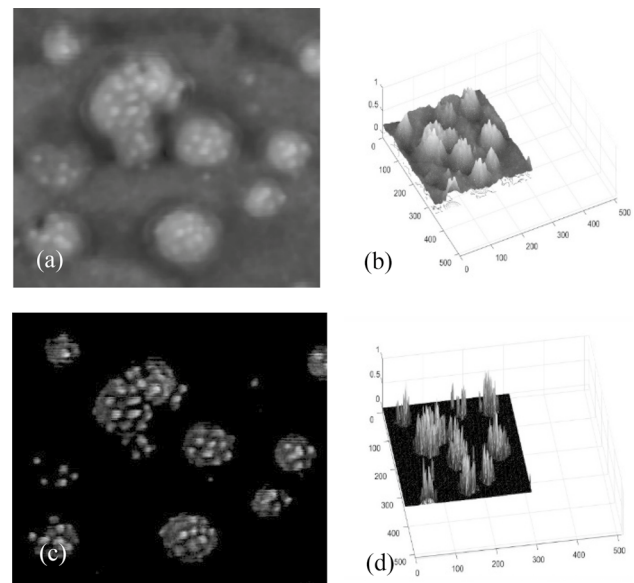

Fig. 10 A region of interest (ROI) from the image in Fig. 1 (b). **a** Original ROI, **b** *3D* representation of ROI, **c** processed image, **d** *3D* representation of the processed ROI

In effect, for small values of the *h* parameter, we achieved a contrast enhancement, while that for high values, we obtained a segmented image. We considered small values of the *h* parameter in the range 10 to 59. Superiors to this interval, we considered high. On the other hands, it is evident that the algorithm, in all cases, improved contrast and accentuated the high spatial frequencies (edges and the *S-spikes*). See the blurring of the original microscopic images with relation to processed images. In a *3D* representation, one can better see this claim (see Fig. 10).

In Fig. 10a, we can better appreciate the blurring in the original image, where the *S-spikes* seem to be out-of-focus. Note in the *3D* representation that the crests are flattened (see image 10 (b)). However, in Fig. 10c, one observes the *S-spikes* and the crown shape is evident in some viral particles, resembling the typical morphology of the coronaviruses [13]. The image has better contrast and edges more defined. The *3D* representation of this processed image shows acute, separated crests and high peaks (see Fig. 10d).

Other processing examples carried out with the *Algorithm No. 1*, for two *h* parameter values, appear in Fig. 11. One can note the stability of this strategy for different images. See that the enhanced and segmented images the *S-spikes* are easy to distinguish. However, in the original microscopic images, one cannot say the same. Here, the *S-spikes* appear very blurred. We carried out more than 750 experiments on different images and we obtained the same results.

4.1 Comparisons with other methods

In this section, we will compare our strategy with two recognized classical techniques. Namely, a histogram equalization (contrast improvement method) and the Otsu's method (segmentation). Such comparison is necessarily incomplete, given the information volume existing in the literature. Instead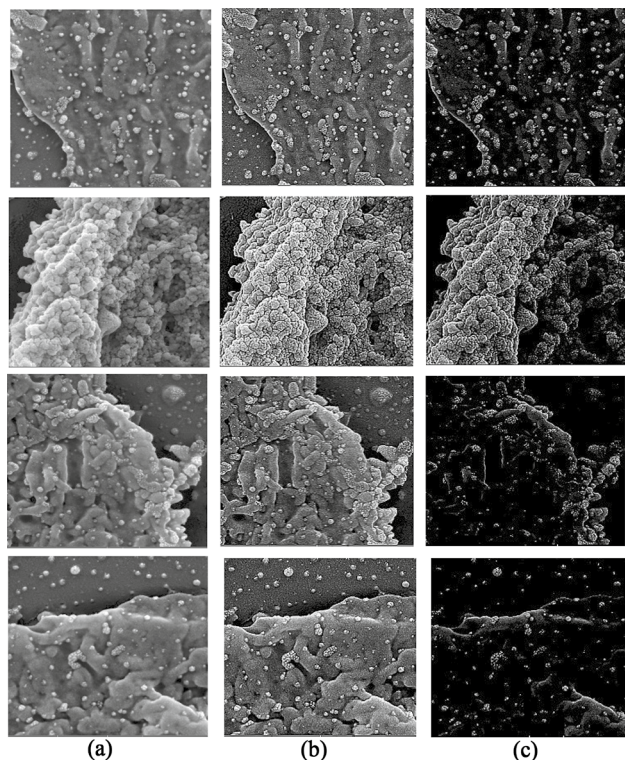, we selected two recognized classical techniques, which are suitable for the kind of used images in our study. In Fig. 12, we show the obtained results with our algorithm and the histogram equalization (HE) and Otsu methods.

In a visual comparison with specialist physicians (more than 10), one can notice that the histogram equalization was not able to discriminate the *S-spikes* and produced an over-saturation. Very similar happened with the Otsu's method. However, our algorithm accentuated and isolated the *S-spikes*. Here, we will show a qualitative comparison carried out by ten experts (see Table 1). We utilized as qualitative evaluation criteria: good (g), regular (R), and poor (P).

Table 1 reflects that our *Algorithm* had better performance than the classical methods in the enhancement and sharpens of the *S-spikes.* One can see that from a sample of 750 images, the experts rated of "*Good*" to 738 images enhanced with our algorithm.

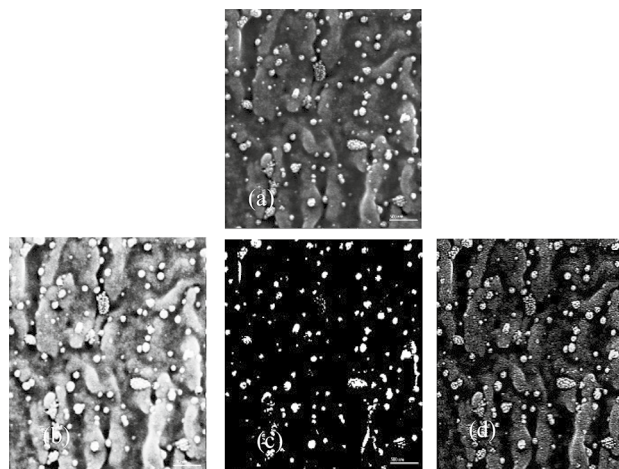Other examples of comparisons with other images appear in Figs. 13 and 14.

Fig. 12 Comparisons with two classical methods. **a** Original microscopic image, **b** Histogram equalization (HE), **c** Otsu's method (OM), **d** Our algorithm. It is evident the obtained result with our strategy

Table 1 Qualitative comparison of our algorithm with the HE and Otsu methods (Shaded cells: number of experts)

Images	HE		Otsu		Our algorithm	
Fig. 12	Poor		Poor		Good	
	10		**10**		**10**	
Fig. 13	Poor		Poor		Good	**9**
	10		**10**		Regular	**1**
Fig. 14	Poor		Regular	**7**	Good	
	10		Poor	**3**	**10**	
Total	G	0	G	10	G	738
Images	R	22	R	40	R	12
(750)	P	728	P	700	P	0

Fig. 11 Enhanced and segmented images applying the *Algorithm No. 1*. **a** Original microscopic images, **b** *h=15*, **c** *h=85*
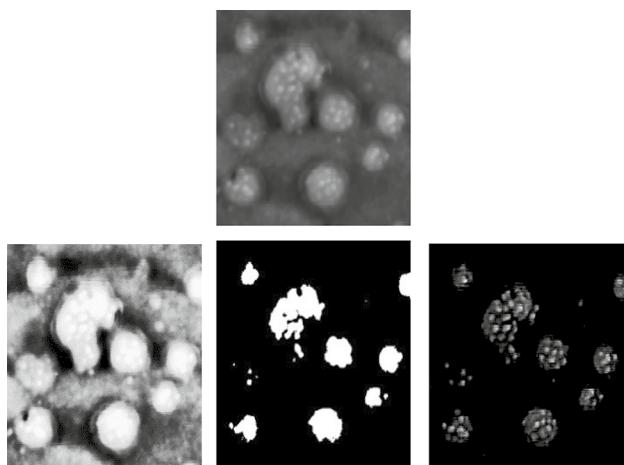
Fig. 13 (Iden. to Fig. 12). In this case, the histogram equalization and the Otsu's method joined more the *S-spikes* and produced more oversaturation. Observe the blurring of the original microscopic image. Our strategy always performed well
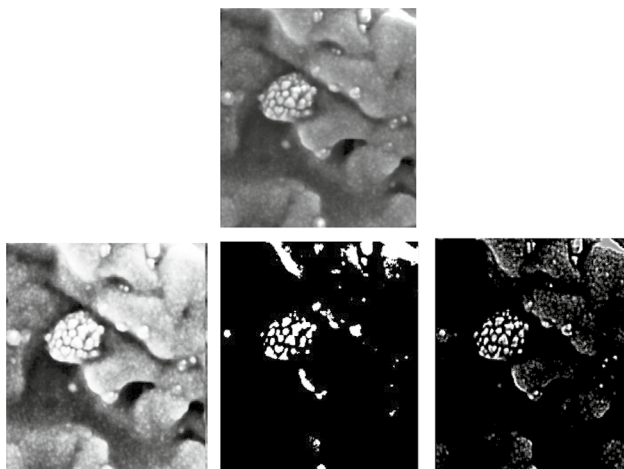

Fig. 14 (Iden. to Fig. 12). In this case, the classic methods performed better, although in both, a bit of over- saturation there was. Notice that the original microscopic image is less blurred. Our algorithm worked well in both situations

5 Conclusions

We presented a strategy for enhancement and segmentation of microscopic images of the novel coronavirus. The basic idea was to take advantage of characteristics of microscopic images of the novel coronavirus (hyperdense zones) for the creation, with a minimum of parameters, of a new algorithm. The proposed strategy enhanced and segmented the *S-spikes*. In this sense, we proposed a variant of the *h-dome* transformation that consisted in subtracting an *h* parameter to separate hyper-dense areas and locally working with high spatial frequencies. The

proposed algorithm showed accuracy and had good performance and stability in the enhancement and isolation of *S-spikes* of the SARS-CoV-2 coronavirus. Finally, we compared our algorithm with two well-established classical methods and proved that in all cases, it was superior in the processing of microscopic images. In that sense, we carried out a qualitative comparison in a database of 750 images, where ten experts evaluated 738 results obtained by our algorithms as "*Good*." A next publication will carry out a theoretical analysis of these results.

References

1. Chen, Y., Liu, Q., Guo, D.: Emerging coronaviruses: genome structure, replication, and pathogenesis. J. Med. Virol. **92**(4), 418–423 (2020)
2. Paules, C.I., Marston, H.D., Fauci, A.S.: Coronavirus infections—more than just the common cold. JAMA **323**(8), 707–708 (2020)
3. Ulhaq, A., Khan, A., Gomes, D., Pau, M.: Computer vision for covid-19 control: a survey (2020). *arXiv preprint* https://arxiv.org/abs/2004.09420
4. Gozes, O., Frid-Adar, M., Greenspan, H., Browning, P. D., Zhang, H., Ji, W., Siegel, E.: Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis (2020). arXiv preprint https://arxiv.org/abs/2003.05037, https://www.researchgate.net/publication/339873521
5. Yan, Q., Wang, B., Gong, D., Luo, C., Zhao, W., Shen, J., You, Z.: COVID-19 chest CT image segmentation--a deep convolutional neural network solution (2020). arXiv preprint https://arxiv.org/abs/2004.10987
6. Fan, D.P., Zhou, T., Ji, G.P., Zhou, Y., Chen, G., Fu, H., Shao, L.: Inf-net: Automatic covid-19 lung infection segmentation from CT images. IEEE Trans. Med. Imaging, 39(8), 2626–2637 (2020). https://doi.org/10.1101/2020.04.22.20074948
7. Amyar, A., Modzelewski, R., Li, H., Ruan, S.: Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. Comput. Biol. Med. **126**, 104037 (2020). https://doi.org/10.1101/2020.04.16.20064709
8. Chhikara, B.S., Rathi, B., Singh, J., Poonam, F.N.U.: Infection, prevention and clinical advances of the prospective chemical drug therapeutics. Chem Biol Lett **7**(1), 63–72 (2020)
9. Fred Plapp: The COVID-19 pandemic: a summary. The Pathologist, 2020. www.thepathologist.com
10. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Tan, W.: A Novel coronavirus from patients with Pneumonia in China, 2019. N. Engl. J. Med. **382**, 727–733 (2020). https://doi.org/10.1056/NEJMoa2001017
11. Gharbi, M., Chen, J., Barron, J.T., Hasinoff, S.W., Durand, F.: Deep bilateral learning for real-time image enhancement. ACM Trans. Gr. (TOG) **36**(4), 1–12 (2017)
12. Liu, X., Guo, S., Yang, B., Ma, S., Zhang, H., Li, J., Fu, Y.: Automatic organ segmentation for CT scans based on super-pixel and convolutional neural networks. J. Digit. Imaging **31**(5), 748–760 (2018)
13. Mondeja, B., Valdes, O., Resik, S., Vizcaino, A., Acosta, E., Montalván, A.: SARS-CoV-2: high-resolution microscopy study in human nasopharyngeal samples. Virol. J. (2020). https://doi.org/10.21203/rs.3.rs-36154/v1
14. Chu, H., Chan, J.F.-W., Yuen, T.T.-T., Shuai, H., Yuan, S., Wang, Y., et al.: Comparative tropism, replication kinetics, and cell

damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. Lancet Microbe. 1(1):e14–23 (2020)

15. Rodríguez, R., Garcés, Y., Torres, E., Sossa, H., Tovar, R.: A vision from a physical point of view and the information theory on the image segmentation. J. Intell. Fuzzy Syst., 37, pp. 2835–2845.

16. Rodríguez, R., Sossa, J. H.: Mathematical techniques for biomedical image segmentation. In: R. Narayan (Ed.), Encyclopedia of Biomedical Engineering, vol. 3, pp. 64–78. Elsevier. ISBN: 9780128048290, Elsevier Inc. All rights reserved. Elsevier, Copyright © 2019.

17. Rodríguez, R., Sossa, J. H. (2012). Procesamiento y análisis digital de imágenes. Alfaomega Grupo Editor, México

18. Roberto Rodríguez, M., Didier, D., Esley, T., & Juan, H. S.: Image segmentation through an iterative algorithm of the mean shift. In: Advances in Image Segmentation. IntechOpen (2012)