

# Morphological Features and Divergence of SARS-CoV 2 and SARS-CoV Virus Using TEM Images

**Bakr Ahmed Taha**

Universiti Kebangsaan Malaysia <https://orcid.org/0000-0002-8922-3993>

**Yousif Al Mashhadany**

University of Anbar

**Qussay Al-Jubouri**

University of Technology

**Mohd Saiful Dzulkefly Bin Zan**

UKM: Universiti Kebangsaan Malaysia

**Yunhan Luo**

Jinan University

**Li Min**

Wuhan University of Technology

**Norhana Arsad** (✉ [noa@ukm.edu.my](mailto:noa@ukm.edu.my))

UKM: Universiti Kebangsaan Malaysia

---

## Research Article

**Keywords:** SARS-CoV 2, SARS-CoV, Feature extraction, Morphology, TEM

**Posted Date:** February 23rd, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1336182/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The COVID-19 pandemic has created a global health catastrophe affecting every aspect of human life. In addition, it has weighed heavily on almost all countries, especially those with less advanced healthcare systems. Improving the accuracy of epidemic prediction models is critical to gaining insight into the impact and spread of this deadly disease. The current circumstances require rapid advances in detecting systems such as in vitro, environmental (e.g., surfaces, sanitation, water, etc.), and vaccine development. A computerized tomography (CT) scan and X-ray images have three issues: ground-glass opacities, consolidation, and crazy paving patterns. These issues can be seen alone or in conjunction with one another. In this study, we focus on the highlight following: TEM morphological features for SARS-CoV 2 virus identification, the proposed framework of extraction features for SARS-CoV 2 virus identification, Comparison features between SARS-CoV 2 and SARS-CoV, Envelope diameter, spike length, roundness, circularity, and area size of SARS-CoV 2 is found for both viruses. The results were that the average diameter for SARS-CoV 2 and SARS-CoV was 97 nm and 102 nm, in comparison, and the main length of spike protein was 11.5 nm and 11.2 nm, respectively. Likewise, sequentially, the mean roundness of SARS-CoV 2 and SARS-CoV was 89.90 nm and 91.65 nm. Consequently, these characteristics improve the sensitivity and specificity of tool detection, identifying the age of viruses, life cycles, and disease progression in humans.

## 1. Introduction

The novel coronavirus disease (COVID-19) epidemic originated worldwide from the SARS-CoV-2 virus and has become an unprecedented public health crisis affecting people's lives and causing a large number of deaths. SARS-CoV-2 belongs to the same beta coronavirus subgroup as SARS-CoV and MERS-CoV (Tang et al. 2020). A coronavirus is associated with an acute respiratory disease. In addition, it can spread to other parts of the body from no symptoms to death; The severity of the disease varies greatly (Martines et al. 2020). In addition to the false negative test problems, COVID-19 tests can produce a limited number of false positive results due to the unspecific detection of other viruses that are almost related to other respiratory diseases, making diagnosis difficult (Lansbury et al. 2020), (Ai et al. 2020). SARS-CoV-2 has a genetic similarity of 79% to SARS-CoV and 50% to MERS-CoV. Despite this, the coding regions of SARS-CoV-2 shared a genetic structure with bat coronaviruses and SARS-CoV. As far as we know, the only significant variation between SARS-CoV-2 and SARS-CoV and MERS-CoV is the length of the spike protein (Lu et al. 2020). It is necessary to extract specific virus traits with diagnostic value from electron microscopic images for classification and identification in order to speed up image analysis (Zhang and Yan 2020). Morphology of SARS-CoV 2 in Vero-E6 cells with infection samples from Brazilian patients using TEM and real-time reverse transcriptase chain reaction (RT-PCR) to study the biology of SARS-CoV 2 (Barreto-Vieira et al. 2021). Assess data on the in vivo death pattern of SARS-CoV infected individuals come from morphological and biochemical studies in SARS-CoV infected VeroE6 cells (Yan et al. 2004). SARS-CoV-2 was assessed using TEM to directly distinguish virus types by providing morphological information and counts of all virus particles, whether infectious or not. This method can be helpful to

identify the cellular target and further explore viruses(Zhao et al. 2020). In the research, cryo-EM has shown the SARS-CoV-2 molecule's molecular structure. As a result, small compounds with fusion-inhibiting potential may design and tested using atomic-level detail, as demonstrated by Spike. Furthermore, antiviral treatments and vaccine design will be accelerated by this information, making medical countermeasures more effective(Zhao et al. 2020). Although cryo-EM is the best method for studying virus ultrastructure, conventional plastic-embedded EM is still valid, especially for looking at samples that cannot be easily analyzed, such as complex objects or pathological material from patients. The molecular construction of the actual SARS-CoV-2 virus using cryo-EM. The virus spike protein is 30 kb long single-segmented DNA into the 80 nm diameter(Yao et al. 2020). An in situ structural analysis of S can be performed using cryo-electron tomography, sub-tomogram averaging, and molecular dynamics simulations. The glycosylation of viral S was higher than that of recombinant S, and most of the viral S was found in the pre-fusion closed conformation(Turoňová et al. 2020). Structures and distributions on intact virions of the spiking proteins of SARS-CoV 2 virion morphology were analyzed using cryo-EM imaging. In terms of diameter, the mean and standard deviation for every three preparations ranges  $91 \pm 11$ ,  $94 \pm 9$ , and  $92 \pm 8$  nm(Ke et al. 2020). The life cycle of SARS-CoV in host cells and the pathogenesis of SARS in vitro is determined by electron microscopy (Qinfen et al. 2004). Several critical issues, in particular, are rarely adequately addressed, such as the identification and prediction of viruses on surfaces, particularly those that are highly contagious. To address these questions, SARS-CoV 2, SARS-CoV, and compared were studied using TEM morphological features to estimate the diameter and length of spike proteins. Importantly, the molecular architecture of SARS-CoV-2 was characterized by imaging combined assessments of intact virions for two viruses, which can provide a new viewpoint for biosensor diagnostic tools.

## 2. Sars-cov-2 Compared To Sars-cov

Coronavirus Disease 2019 (COVID-19) is an infectious disease caused by SARS-CoV-2. The average incubation period is five days, but those who develop symptoms go through infection for around 12 days (Lauer et al. 2020). Thus, a considerable percentage of viral transmission may occur before infected individuals exhibit symptoms. In addition, infection is typically diagnosed after the beginning of symptoms according to clinical factors (He et al. 2020). The composition of a virus consists of spike (S) protein, membrane (M) protein, envelope (E) protein and nucleocapsid (N) protein. The detection limit of COVID-19 molecular diagnostic techniques is (100, 2000, 20, and 1000) copies respectively(Weissleder et al. 2020). It has the largest genome (26.4–31.7 kb) of all known RNA viruses. Additionally, the virus's genetic material includes single-stranded RNA(Mousavizadeh and Ghasemi 2020). Coronavirus diseases are heavily influenced by receptor recognition. Angiotensin-converting enzyme 2(ACE2) is the specific surface protein that allows virus to enter human cells(Tanonaka and Marunouchi 2016),(Renhong et al. 2020). The first distinction between the two SARS-CoV is that the COVID-19 receptor-binding domain (RBD) has a greater ACE2-binding affinity, resulting in more efficient cell penetration. The ACE2-binding affinity of the complete S protein, on the other hand, appears to be equivalent to or even lower than that of the whole SARS-CoV S protein. This observation implies that, despite its greater potency, COVID-19

RBD is likely lower exposed than SARS-CoV RBD(Shang et al. 2020). The majority of the COVID-19 was in the "lying-down" posture, linked with ineffective receptor binding. In comparison to the spike, cryo-EM investigations revealed that the RBD is predominantly in the "standing-up" form in SARS-CoV spike protein(Letko et al. 2020),(Ou et al. 2020). The differences between the two are mainly in the S protein, ORF8 protein and ORF3b protein, as can be seen in Figure 1.

### 3. Methodology

Coronaviruses disease (COVID-19) is a pathogenic virus which might be transmitted by surfaces, air, and human contact. Otherwise, the spike protein of the COVID-19 is a key part of the cell host infection. It is responsible for the interaction of viral and cellular membranes to cell receptors and their subsequent fusion. Therefore, it is necessary to understand the outside appearance and features of viruses such as envelope and spike to provide new insights into detection devices. Dataset collected was SARS-CoV 2, and SARS-CoV-2 isolated from SARS Coronavirus 2 Italy-INMI1 (SARS-CoV-2)(Colavita et al. 2020), SARS Coronavirus Frankfurt 1 (SARS-CoV)(Thiel et al. 2003) alternately in ultrathin plastic sections of infected Vero cell cultures. Transmission electron microscopy was used to identify SARS-CoV-2 after inoculating cells with 2% paraformaldehyde and 2.5 percent glutaraldehyde. The structure of SARS-CoV-2 was discovered by analyzing infected cells three days after infection. In addition, the coronavirus-specific morphology of SARS-CoV-2 was disclosed by electron microscopy, with virus particle sizes ranging from 70 - 90 nm being detected. The application of computer vision technologies is critical in developing these systems, for objectives such as object detection, tracking, identification, image retrieval, and different angles of view of the identical scene or object (Kim et al. 2014). Electron microscopy is an effective method to identify the structural features of viruses where the size varies from micrometers to nanoscales. It can give researchers important information about images(Li et al. 2018). In addition, advances in computer tools and mathematical methods have led to the development of software that can replace human vision. The extraction of virus features is the common thread that runs through them all. The procedure sketched in framework proposal for identifying virus applies to three critical aspects of decision making include (diagnostic, life cycle of virus, and disease progression). It would be interesting to examine the benefits of these features introduces more possibilities to result in more accuracy in identifying the virus, as shown in Figure 2.

### 3.1. Dataset collection

A data set is a collection of raw statistics and data obtained from a research study. Datasets created by government agencies or nonprofits are often available for free download. In our analysis we used nine datasets from the Zenodo website as indicated in Table 1, we have discussed the dataset collection which includes Transmission Electron Microscopy SARS-CoV-2 and SARS-CoV images. It contains 586 transmission electron microscopy photos of extracellular virus particles in Vero cell cultures through ultra-thin (60-70 nm) plastic discs. Images with a resolution of 1376-1032 pixels were collected and saved in 16-bit TIF format. A picture viewer can also read 16-bit images for all photos in order to see them (e.g.

Irfan-View). Image files are size calibrated and can be viewed in ImageJ or Fiji with the required size calibration using the bio-format importer.

Table 1

A description of the data sets that were used for the virus particles SARS-COV2 and SARS-COV.

| Location of virus isolation | Virus categories | Count of images | Plastic thin | dimensions pixel | pixel size | File type     |
|-----------------------------|------------------|-----------------|--------------|------------------|------------|---------------|
| Germany /Frankfurt          | SARS-CoV         | 126             | 60-70 nm     | 1376 x 1032      | 0.64 nm    | Tif, 16-bits  |
| Italy-INMI                  | SARS-CoV 2       | 122             | 60-70 nm     | 1376 x 1032      | 0.64 nm    | Tif, 16-bits  |
| Germany /Frankfurt          | SARS-CoV         | 111             | 45 nm        | 1376 x 1032      | 0.54 nm    | Tif, 16-bits  |
| Italy-INMI                  | SARS-CoV 2       | 134             | 45 nm        | 1376 x 1032      | 0.54 nm    | Tif, 16-bits  |
| Italy-INMI                  | SARS-CoV 2       | 66              | 110 nm       | 1376 x 1032      | 0.64 nm    | Tif, 16-bits  |
| Germany /Frankfurt          | SARS-CoV         | 11              | tomograms    | 1900 x 1900      | 0.57 nm    | Tif, 8 - bits |
| Italy-INMI                  | SARS-CoV 2       | 11              | tomograms    | 1900 x 1900      | 0.57 nm    | Tif, 8 - bits |
| Italy-INMI                  | SARS-CoV-2       | 85              | 65 nm        | 1376 x 1032      | 0.64 nm    | Tif, 16-bits  |
| Italy-INMI                  | SARS-CoV-2       | 101             | 85 nm        | 1376 x 1032      | 0.64 nm    | Tif, 16-bits  |

## 3.2. Preprocessing

Image enhancement is an essential step in image analysis. Contrast enhancement is used to increase the quality of an image to make it more suitable for a particular purpose. The purpose of image improvement is to make the digital image more attractive to visual system, for instance, by smoothing or intensifying the images. This is a significant area of study in image analysis. Many studies suggested that histogram equalization (HE) is an essential and easy technique for increasing contrast and image quality (Khan et

al. 2015),(Vidyasaraswathi, H. N., & Hanumantharaju 2020). Spatial correlation is more significant than signal strength in scientific imaging, like distinguishing genetic material of specific length. A low signal-to-noise ratio generally makes visual identification difficult. Histogram equalization frequently results in false pictures; nevertheless, it is highly beneficial in scientific images such as thermal, satellite, or x-ray imaging. In addition, the histogram's intersection match value is enhanced in the background by one pixel only if the pixel has the same color in the image and the number of pixels of that color in the object is less than the number of pixels in the picture. Correspondingly, normalizing a histogram is a mathematical operation that converts the discrete distribution of intensities to a discrete probability distribution. The enhance contrast of the histogram of images based on the band pass filter such as normalize histogram, and equalize histogram to enhance edges while reducing noise.

### **3.3. Morphology analysis of structural difference.**

In this work, two isolates of the virus were in use: SARS-CoV Frankfurt and SARS-CoV 2 Italy-INMI1 (SARS-CoV-2). (Colavita et al. 2020),(Thiel et al. 2003). We adjusted brightness/contrast, and the threshold image was changed before processing, making envelope and spike density very clear. Beside, using the straight-line tool function in Image J software, we assessed the morphology of the particle size distribution of the SARS-CoV and COVID-19 viruses. (Yao et al. 2020),(Abràmoff et al. 2004),(Laue et al. 2021). We are using Fiji software to determine the particle size by the "geometric selection" tool and the measured set "fit ellipse" to choose the outer layer of the viral membrane. Consequently, we calculate the fitted ellipse's max / mini diameters shape features such as surface area and circularity. Then, the most significant length of the spikes related to a viral particle was determined using a step-wise (nanometer) extension of the geometrical selection used to find the maximal diameter of each virus particle. Next, individual line measurements of spike length confirm the technique's accuracy. Finally, we used MS Excel to compile and analyze the data. In figure 3, we measured the envelope protein diameter, which was a random selection of the virus particle, using the scale bar tool in "ImageJ software." We adjusted it according to global size in the literature. However, we need to measure the diameter of different particles from varied images to prove the distribution of average diameters envelop (E) protein for COVID-19 and SARS-CoV. Likewise, we extracted features of the spike(S)protein length. In addition, we measured the diameter using a cross form due to the non-circular shape. Interestingly, when compared to other traditional CT scans and X-ray analysis used to classify human infections, the observed diversity of the envelope diameter and the tip length of SARS-CoV 2 and SARS-CoV an attractive one. This morphological variance may be due to fixation artifact in clinical specimens. This problem can be successfully solved by imaging the virus produced from cell cultures. The imaging of a few virus particles is intrinsically restricted(Aggarwal et al. 2012).

### **3.4. Extraction features of COVID-19**

All viruses are classified into four main categories: filamentous, icosahedral, enveloped, and head and tail. Morphology methods can categorize these types and address the possibility of dependency between different of the viruses. In addition, it is possible to divide viruses based on many different characteristics, such as their nucleic acid, the shape of their capsids, and even whether or not they possess an exterior

envelope. The classification of viral core structures uses genetic material (DNA or RNA), single/double-stranded, linear/circular, and segmented or not. (Crick and Watson 1956). Many studies and reviews have been performed to examine the characterization and structures of COVID-19. However, the shape of the virus under the electron microscope of negatively filled particles was spherical. The diameter was between (60 - 140) nm and has distinctive peaks with an average length of about (9 -12) nm, which gave the virus the formation of a solar corona. Researchers discovered extracellular free virus particles and bodies filled with virus particles in membrane-bound vesicles in ultra-thin sections of the human airway epithelium. This morphology is the characteristic of the coronavirus family (Zhu et al. 2020a). In addition, it had a volume of approximately  $\sim 10^6 \text{ nm}^3 = 10^{-3} \text{ fL}$ , and mass of  $\sim 10^3 \text{ MDa} \approx 1\text{fg}$ . A recent study reported a frequency of COVID-19 spike proteins of  $\sim 90$  spikes per particle (Zhu et al. 2020a). A feature can be described as a fascinating building block or important focus in an image. In addition, a well-defined location displays extensive visual information such as shape, color and texture [39]. A rather consensual idea is the basic properties that such image features should have. They should stand out clearly from the background, the associated values of interest should have a meaning, potentially useful for additional interpretability, also independent of radiometric and geometric distortions (invariance), robust against image noise and distinguishable from other points. In addition to extract the characteristic, the points of interest, key points, corners, affine regions, invariant regions are also used [40]. Our investigation involved segmentation and analysis using the trainable weka segmentation in the figure 4. This Fiji plugin combines machine learning algorithms with selected image features to produce pixel-based segmentations. It can be helpful to identify critical keypoints of the sars-cov2 virus such as Spikes, Envelop, roundness, and nucleocapsid.

Feature extraction, however, is a component of the dimensionality reduction technique that breaks down an initial collection of raw data and reduces it to more manageable groupings. This simplifies the processing. The most striking characteristic of this large amount of data is its abundance of variables. Processing these variables requires considerable computing power. Thus, feature extraction enables the best features to be extracted from large data sets by separating variables into elements and combining them, greatly reducing the amount of data. Although extensive research has been carried out on the characteristics of COVID-19, there is still much scope to understand the external shape of the virus on surfaces before it enters the human body, in order to monitor changes in composition and the life cycle of the detected virus. Figure 5, illustrates the extraction characteristics of the SARS-CoV 2 virus.

## 4. Results And Discussion

Viruses appear in various sizes and forms, but each viral family has unique features. SARS-CoV-2 is similar in architecture to other known human CoVs. For example, a nucleic acid genome is inside a protective coating of proteins known as a capsid in every virus except the spike portion responsible for binding the host cell-surface receptor during host cell entry. Results show the spike protein and envelope components as significant features in the coronavirus family (Yutin et al. 2021). In some cases, viral may look similar, but their structure (envelope diameter, length, and density of spike portion) differs. A

published report shows that SARS-CoV-2 has 79.5% similarity to SARS-CoV and 50% similarity to MERS-CoV(Zhu et al. 2020b).It appears to be only slightly different from SARS-CoV in terms of clinical features. However, it is spreading even faster [40],[41]. Virus particles have rounded or oval shapes in their surface appearance. However, in each image, some particles were a different, irregular, or deformed shape; for this reason, they were excluded from the morphometric assessment. There were differences in the size distribution of SARS-CoV and SARS-CoV2. It is vital to emphasize that in many countries, computed tomography is the most commonly utilized medical imaging modality for diagnosing SARS-CoV-2; due to its widespread use, databases are now available to researchers. However, there are many difficulties, such as ground-glass opacities, consolidation, and bizarre paving patterns. These problems can be viewed alone or in combination with one another. It suggests that while chest CT is very sensitive, it is not highly specific for COVID-19. Therefore, patients who get these findings should have a more thorough clinical examination and laboratory tests to rule out other possible reasons(Schultz et al. 2020),(Chen et al. 2020),(Huang et al. 2021). In this work, we use ImageJ (Fiji) software extraction to analyse these features, a broad selection method focusing on biological-image analysis to enable fast image processing methods. (Schindelin et al. 2012), We examined a dataset collection of TEM images containing 515 images of SARS-CoV2 and 248 SARS-CoV particles in ultrathin plastic slices of infected Vero cultured cells.

## 4.1. Assessment of envelope diameter

A study evaluated that the SARS-CoV 2 diameter varied from about 60 to 140 nm as observed under a transmission electron microscope(Zhu et al. 2020a). In contrast to what was seen before, our assessment differs considerably in identifying the envelope diameter of the virus. Extracting the features of the virus diameter is essential for the SARS-CoV2 identifiers in our system and we measured the diameter sizes from various EM images using ImageJ software. In order to evaluate images, we may have to compensate for recording errors. Noise, uneven lighting, and background fluorescence can cause a variety of image processing problems. In the first stages, we apply enhancement contrast on the images to show the shape outlier of a virus, then applied the scale bar tool and the measurement option closely fit to choose the outer layer of the viral envelope. Figure 6, shows the TEM morphology features of the envelope protein distribution diameter for both SARS-CoV2 and SARS-CoV viruses, and we adjusted scale bars "Straight Line" tool at 100 nm. Consequently, the average diameter for SARS-CoV2 and SARS-CoV was 97 nm and 102 nm, respectively. The Gaussian Distribution uses the numerical correlation that explains values in a data collection, and measurements approximate this connection as sample size expands. The diameter of an envelope is based on the mean and standard deviation of a given Gaussian distribution. The mean defines the location of the curve's center, whereas the standard deviation determines its apparent breadth.

## 4.2. Assessment of length spike protein

In this part we extracted the average length peak of SARS-CoV2 and SARS-CoV. We also measured the tip length with the ImageJ software with the scale bar tool at 12 nm. Many studies described the length of

virus spikes protein, around (9-12) nm in size, which gave virions the appearance of a solar corona (Zhu et al. 2020a). In addition, the architecture observed is compatible with the coronavirus family. Cryo-EM images of isolated SARS-CoV-2 virions revealed the existence of virus tips and membranes. About 20%-30% of virions have multiple spines around the membrane, while most other virus particles have few spikes(Liu et al. 2020). However, we tested many tip lengths of various images to prove the main length using the scale bar tool in ImageJ software. Figure 7, Illustrate the distribution average spike length using TEM images and the adjusted scale bars "Straight Line" tool at 12 nm and the Gaussian distribution of SARS-CoV 2 and SARS-CoV spike protein with probability density of spikes length. The results indicate that the average spike length for SARS-CoV 2 SARS-CoV was 11.5 nm 11.2 nm, respectively.

### **4.3. Assessment of roundness and circularity**

We calculate the shape of the virus using the following descriptors: Circularity, aspect ratio (AR), Round (roundness), and Solidity (area/convex area). Comparing the roundness of a virus is another factor that may be a good indicator of quality. Roundness can be calculated mathematically using the following formula(  $Roundness = 4\pi \times area / area / perimeter^2$ ). The roundness criteria describe how perfect the form of a virus is oval, ellipse, spherical, circular, and irregular. In addition, the study confirmed the findings of the loss function, which primarily focuses on surrounding protein projections that appeared in electron microscopy pictures as weighting about spherical physical shape, which seems in electron microscopy images as an approximately round polygon(Zhang and Yan 2020). Describe SARS-CoV 2 in situ using cryo-EM images are roughly spherical particles with varied sizes centered around 100 nm (Klein et al. 2020). The shape of the selected virus particle with spikes was measured using the "freehand selection" tool. The results of this analysis indicate the average round of SARS-CoV 2 and SARS-CoV was 89.90 nm and 91.65 nm sequentially, as shown in figure 8.

Maximal and minimal fitting ellipse and shape descriptors, such as aspect ratio and circularity, were determined. In figure 9, the black line represents the distribution circularity of the SARS-CoV and the red line for the SARS-CoV 2. A value of 1.0 indicates a perfect circle. The results show that the value approaches 0.0, an increasingly elongated, which would make the shape of both viruses non-circular.

### **4.3. Assessment of the Area size of virus.**

TEM may be a vital technique for demonstrating viral infection, but it must be used with caution when interpreting cytoplasmic features to detect viral particles appropriately (Goldsmith et al. 2020),(Ogando et al. 2020). Electron microscopy can assist in the quick identification of viral illnesses since it can be conducted in a matter of hours. Analytical differences of morphological preservation attained by the various fixation and embedding procedures might explain the disparities in results. However, identify the source of the material and the patient's symptoms since they will lead to suggestions of prospective agents while ruling out others. The preparation technique is chosen based on the consistency of the sample, extraction, concentration, and tissue culture amplification. It must avoid false positives by distinguishing viruses from cell organelles or detritus, bacterial contamination (Miller 1986),(Kim et al. 2020). Hence, extracting the main characteristics of the Outside Shape Virus provides insight into its life

cycle and how to reduce pandemic outbreaks. Our results have been remarkably close to those expected and showed that the average area size of SARS-CoV 2 and SARS-CoV are comparable in morphology and size, about 80%, based on the close taxonomic relatedness of the two viruses and the reports available on the virus. In addition, the SARS-CoV 2 has a broad shifting in the shape it, which proves improving Mutation Capabilities as shown in figure 10. Finally, Training the Artificial intelligence techniques required a database based on TEM morphology features can enable accelerated decision-making and improve understanding of how viruses spread. Furthermore, improved diagnostic tools and accuracy enabled new efficient therapeutic breakthroughs and identified most residents at risk as possibly having physiological traits (Alimadadi et al. 2020),(Alazab et al. 2020). In future work, we aim to integrate an optical sensor based on image processing by creating an algorithm model to identify and detect a virus in public spaces using TEM morphological features. Thus, that is the best answer to contain the pandemic and save doctors' time(Taha et al. 2020), (Taha et al. 2021).

## 5. Conclusion

COVID-19 has had a significant impact on more than 200 nations that have reported cases of infection; several festivals, gaming events, competitions, and exhibitions have been postponed or canceled. In addition, the epidemic has had a significant impact on education worldwide, with many institutions and universities closing. As a result, the world economy has been seriously affected. The virus can be transmitted to things we touch and then to us through direct contact or the air. Hence, we need to develop inventive techniques to avoid touching things and keep surfaces clean to prevent it. However, living our lives makes it difficult to prevent all forms of contact, so we need to recognize the valid dangers in our surroundings. The extraction of image features is an important field in computer vision research. Specific properties are extracted from a massive set of pixel data. In this work, we proposed modeling to extract shape features of the virus from TEM images. The results indicate 97 nm and 102 nm of SARS-CoV 2 and SARS-CoV; in contrast, the main length of spike protein was 11.5 nm and 11.2 nm, respectively. In the future, we will integrate this model with a fiber optic sensor to improve the accuracy in identifying and detecting the virus on the surfaces. In this way, we can potentially connect hospitals to public places to assess the environmental risk and collect the data remotely.

## Declarations

### Authors contribution:

B.A.T., wrote the initial manuscript draft, Y.A.M., Q.A.J., Conceptualization and methodology, investigation and, N.A., supervision and validation with the support of M.S.D.B.Z., Y.H.L., L.M., All authors have reviewed and accepted the published version of the manuscript.

### Funding:

Universiti Kebangsaan Malaysia (UKM) funded this study through grant number (GUP-2019-010 and GP-2019-K013436).

**Institutional Review Board Statement:**

Not applicable.

**Informed Consent Statement:**

Not applicable.

**Data Availability Statement:**

Datasets are accessible in the Zenodo data repository: <https://zenodo.org>:Dataset1: <https://doi.org/10.5281/zenodo.3985098>; Dataset3:<https://doi.org/10.5281/zenodo.3985110>; Dataset 6: <https://doi.org/10.5281/zenodo.3986526>; Dataset :<https://doi.org/10.5281/zenodo.3986580>; Dataset 8: <https://doi.org/10.5281/zenodo.4275703> ;Dataset 9 <https://doi.org/10.5281/zenodo.4275728> ; Dataset 10: <https://doi.org/10.5281/zenodo.4275742>; Dataset 11: <https://doi.org/10.5281/zenodo.4275769>.

**Acknowledgements:**

The authors gratefully acknowledge the financial support of the Department of Electrical, Electronic and Systems Engineering/Faculty of Engineering and Built Environment/Universiti Kebangsaan Malaysia (UKM) for their encouragement and grant support (GUP-2019-010 and GP-2019-K013436).

**Conflicts of Interest:**

The authors declare no conflict of interest.

**Ethics and Permission to Participate:**

This manuscript has not been previously released and is not now under consideration by any journal for publication.

## References

1. Abràmoff MD, Magalhães PJ, Ram SJ (2004) Image processing with imageJ. *Biophotonics International* 11:36–41. <https://doi.org/10.1201/9781420005615.ax4>
2. Aggarwal A, Mehta S, Gupta D, et al (2012) Clinical & immunological erythematosus patients characteristics in systemic lupus Maryam. *Journal of dental education* 76:1532–9. <https://doi.org/10.4103/ijmr.IJMR>
3. Ai T, Yang Z, Hou H, et al (2020) Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* 296:E32–E40.

<https://doi.org/10.1148/radiol.2020200642>

4. Alazab M, Awajan A, Mesleh A, et al (2020) COVID-19 prediction and detection using deep learning. *International Journal of Computer Information Systems and Industrial Management Applications* 12:168–181
5. Alimadadi A, Aryal S, Manandhar I, et al (2020) Artificial intelligence and machine learning to fight covid-19. *Physiological Genomics* 52:200–202.  
<https://doi.org/10.1152/physiolgenomics.00029.2020>
6. Barreto-Vieira DF, da Silva MAN, Garcia CC, et al (2021) Morphology and morphogenesis of sars-cov-2 in vero-e6 cells. *Memorias do Instituto Oswaldo Cruz* 116:1–6. <https://doi.org/10.1590/0074-02760200443>
7. Chen SG, Chen JY, Yang YP, et al (2020) Use of radiographic features in COVID-19 diagnosis: Challenges and perspectives. *Journal of the Chinese Medical Association* 83:644–647.  
<https://doi.org/10.1097/JCMA.0000000000000336>
8. Colavita F, Lapa D, Carletti F, et al (2020) SARS-CoV-2 isolation from ocular secretions of a patient with COVID-19 in Italy with prolonged viral RNA detection. *Annals of Internal Medicine* 173:242–243.  
<https://doi.org/10.7326/M20-1176>
9. Crick FHC, Watson JD (1956) Structure of small viruses. *Nature* 177:473–475
10. Goldsmith CS, Miller SE, Martines RB, et al (2020) Electron microscopy of SARS-CoV-2: a challenging task. *The Lancet* 395:e99. [https://doi.org/10.1016/S0140-6736\(20\)31188-0](https://doi.org/10.1016/S0140-6736(20)31188-0)
11. He X, Lau EHY, Wu P, et al (2020) Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine* 26:672–675. <https://doi.org/10.1038/s41591-020-0869-5>
12. Huang S, Yang J, Fong S, Zhao Q (2021) Artificial intelligence in the diagnosis of covid-19: Challenges and perspectives. *International Journal of Biological Sciences* 17:1581–1587.  
<https://doi.org/10.7150/ijbs.58855>
13. Ke Z, Oton J, Qu K, et al (2020) Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* 588:498–502. <https://doi.org/10.1038/s41586-020-2665-2>
14. Khan MF, Khan E, Abbasi ZA (2015) Image contrast enhancement using normalized histogram equalization. *Optik* 126:4868–4875. <https://doi.org/10.1016/j.ijleo.2015.09.161>
15. Kim JM, Chung YS, Jo HJ, et al (2020) Identification of coronavirus isolated from a patient in Korea with covid-19. *Osong Public Health and Research Perspectives* 11:3–7.  
<https://doi.org/10.24171/j.phrp.2020.11.1.02>
16. Kim YD, Park JT, Moon IY, Oh CH (2014) Performance analysis of ORB image matching based on android. *International Journal of Software Engineering and its Applications* 8:11–20.  
<https://doi.org/10.14257/ijseia.2014.8.3.02>
17. Klein S, Cortese M, Winter SL, et al (2020) SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. *Nature Communications* 11:1–10. <https://doi.org/10.1038/s41467-020-19619-7>

18. Lansbury L, Lim B, Baskaran V, Lim WS (2020) Co-Infections in People with COVID-19: A Systematic Review and Meta-Analysis. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3594598>
19. Laue M, Kauter A, Hoffmann T, et al (2021) Morphometry of SARS-CoV and SARS-CoV-2 particles in ultrathin plastic sections of infected Vero cell cultures. *Scientific Reports* 11:1–11. <https://doi.org/10.1038/s41598-021-82852-7>
20. Lauer SA, Grantz KH, Bi Q, et al (2020) The incubation period of coronavirus disease 2019 (CoVID-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine* 172:577–582. <https://doi.org/10.7326/M20-0504>
21. Letko M, Marzi A, Munster V (2020) Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiology* 5:562–569. <https://doi.org/10.1038/s41564-020-0688-y>
22. Li W, Field KG, Morgan D (2018) Automated defect analysis in electron microscopic images. *npj Computational Materials* 4:1–9. <https://doi.org/10.1038/s41524-018-0093-8>
23. Liu C, Mendonça L, Yang Y, et al (2020) The Architecture of Inactivated SARS-CoV-2 with Postfusion Spikes Revealed by Cryo-EM and Cryo-ET. *Structure* 28:1218-1224.e4. <https://doi.org/10.1016/j.str.2020.10.001>
24. Lu R, Zhao X, Li J, et al (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet* 395:565–574. [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8)
25. Martines RB, Ritter JM, Matkovic E, et al (2020) Pathology and pathogenesis of SARS-CoV-2 associated with fatal coronavirus disease, united states. *Emerging Infectious Diseases* 26:2005–2015. <https://doi.org/10.3201/eid2609.202095>
26. Miller SE (1986) Detection and identification of viruses by electron microscopy. *Journal of Electron Microscopy Technique* 4:265–301. <https://doi.org/10.1002/jemt.1060040305>
27. Mousavizadeh L, Ghasemi S (2020) Genotype and phenotype of COVID-19: Their roles in pathogenesis. *Journal of Microbiology, Immunology and Infection* 0–4. <https://doi.org/10.1016/j.jmii.2020.03.022>
28. Ogando NS, Dalebout TJ, Zevenhoven-Dobbe JC, et al (2020) SARS-coronavirus-2 replication in Vero E6 cells: Replication kinetics, rapid adaptation and cytopathology. *Journal of General Virology* 101:925–940. <https://doi.org/10.1099/jgv.0.001453>
29. Ou X, Liu Y, Lei X, et al (2020) Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature Communications* 11:.. <https://doi.org/10.1038/s41467-020-15562-9>
30. Qinfen Z, Jinming C, Xiaojun H, et al (2004) The life cycle of SARS coronavirus in Vero E6 cells. *Journal of Medical Virology* 73:332–337. <https://doi.org/10.1002/jmv.20095>
31. Renhong Y, Yuanyuan Z, Yaning L, et al (2020) Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* 367:1444–1448

32. Schindelin J, Arganda-Carreras I, Frise E, et al (2012) Fiji: An open-source platform for biological-image analysis. *Nature Methods* 9:676–682. <https://doi.org/10.1038/nmeth.2019>
33. Schultz MJ, Sivakorn C, Dondorp AM (2020) Challenges and opportunities for lung ultrasound in novel Coronavirus disease (COVID-19). *American Journal of Tropical Medicine and Hygiene* 102:1162–1163. <https://doi.org/10.4269/ajtmh.20-0323>
34. Shang J, Wan Y, Luo C, et al (2020) Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences of the United States of America* 117:.. <https://doi.org/10.1073/pnas.2003138117>
35. Taha BA, Al Mashhadany Y, Bachok NN, et al (2021) Detection of covid-19 virus on surfaces using photonics: Challenges and perspectives. *Diagnostics* 11:1119. <https://doi.org/10.3390/diagnostics11061119>
36. Taha BA, Mashhadany Y Al, Mokhtar MHH, et al (2020) An analysis review of detection coronavirus disease 2019 (Covid-19) based on biosensor application. *Sensors (Switzerland)* 20:1–29. <https://doi.org/10.3390/s20236764>
37. Tang D, Comish P, Kang R (2020) The hallmarks of COVID-19 disease. *PLoS Pathogens* 16:1–24. <https://doi.org/10.1371/journal.ppat.1008536>
38. Tanonaka K, Marunouchi T (2016) Angiotensin-converting enzyme 2. *Folia Pharmacologica Japonica* 147:120–121. <https://doi.org/10.1254/fpj.147.120>
39. Thiel V, Ivanov KA, Putics Á, et al (2003) Mechanisms and enzymes involved in SARS coronavirus genome expression. *Journal of General Virology* 84:2305–2315. <https://doi.org/10.1099/vir.0.19424-0>
40. Turoňová B, Sikora M, Schürmann C, et al (2020) In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* 370:203–208. <https://doi.org/10.1126/science.abd5223>
41. Vidyasaraswathi, H. N., & Hanumantharaju MC (2020) Low-Contrast Image Enhancement Histogram method using Grey Wolf Optimization search algorithm. 15:19–28
42. Weissleder R, Lee H, Ko J, Pittet MJ (2020) COVID-19 diagnostics in context. *Science Translational Medicine* 12:1–6. <https://doi.org/10.1126/scitranslmed.abc1931>
43. Yan H, Xiao G, Zhang J, et al (2004) SARS coronavirus induces apoptosis in Vero E6 cells. *Journal of Medical Virology* 73:323–331. <https://doi.org/10.1002/jmv.20094>
44. Yao H, Song Y, Chen Y, et al (2020) Molecular Architecture of the SARS-CoV-2 Virus. *Cell* 183:730–738.e13. <https://doi.org/10.1016/j.cell.2020.09.018>
45. Yutin N, Benler S, Shmakov SA, et al (2021) Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. *Nature Communications* 12:1–11. <https://doi.org/10.1038/s41467-021-21350-w>
46. Zhang L, Yan WQ (2020) Deep Learning Methods for Virus Identification from Digital Images. *International Conference Image and Vision Computing New Zealand 2020-Novem*:6–11. <https://doi.org/10.1109/IVCNZ51579.2020.9290670>

47. Zhao J, Zhou H, Huang W, et al (2020) Cell morphological analysis of SARS-CoV-2 infection by transmission electron microscopy. *Journal of Thoracic Disease* 12:4368–4373. <https://doi.org/10.21037/jtd-20-1368>
48. Zhu N, Zhang D, Wang W, et al (2020a) A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* 382:727–733. <https://doi.org/10.1056/nejmoa2001017>
49. Zhu Z, Lian X, Su X, et al (2020b) From SARS and MERS to COVID-19: A brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. *Respiratory Research* 21:1–14. <https://doi.org/10.1186/s12931-020-01479-w>

## Figures

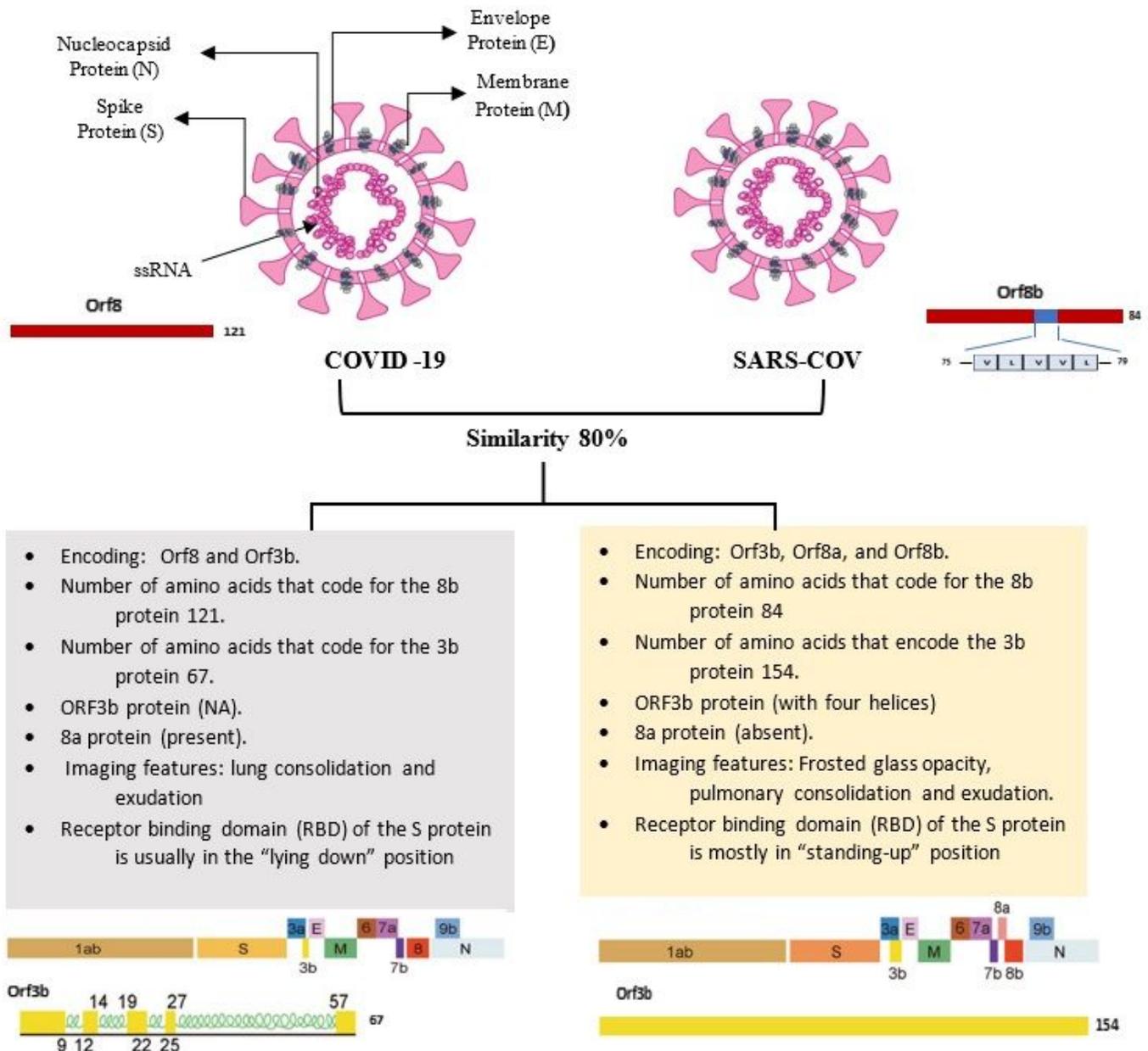


Figure 1

Amino acid content and protein structure differences the different capabilities of SARS-CoV and COVID-19 lie in the S protein, ORF8 protein and ORF3b protein.

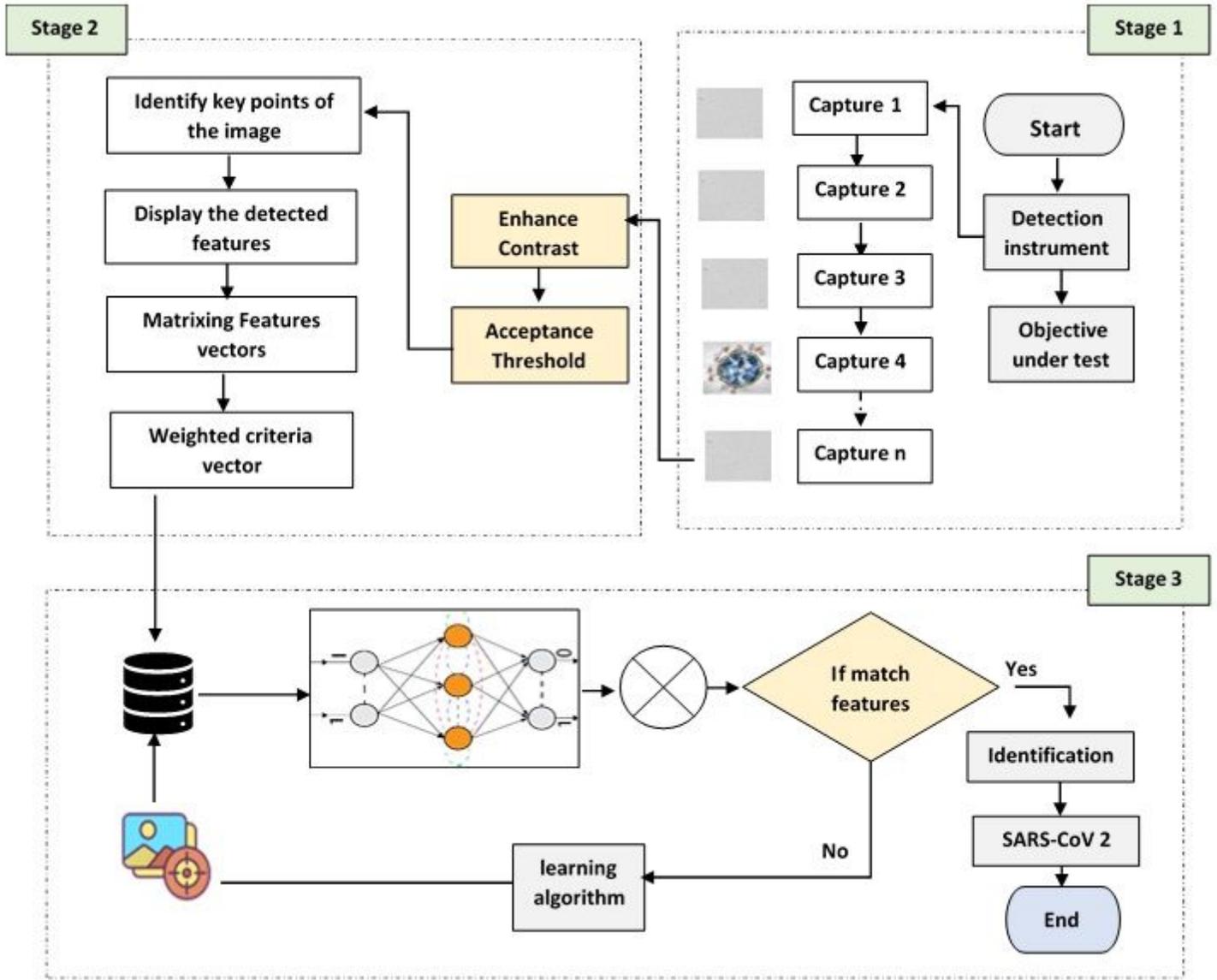
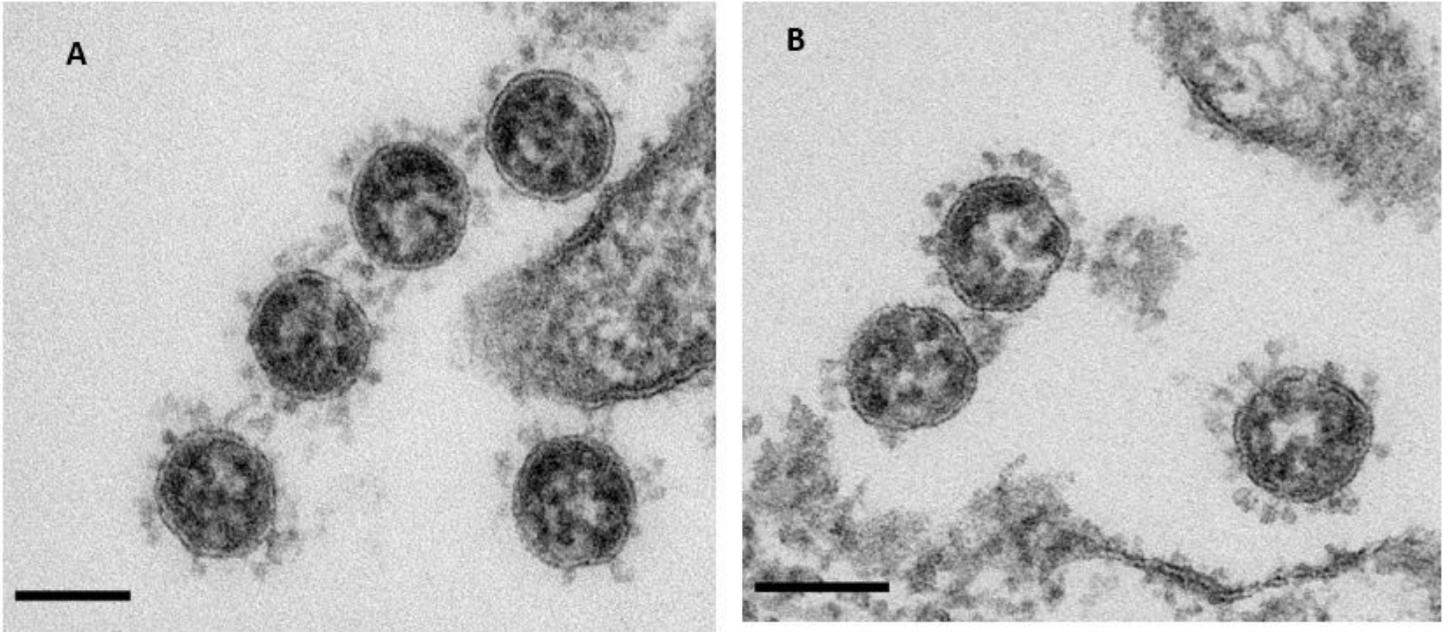


Figure 2

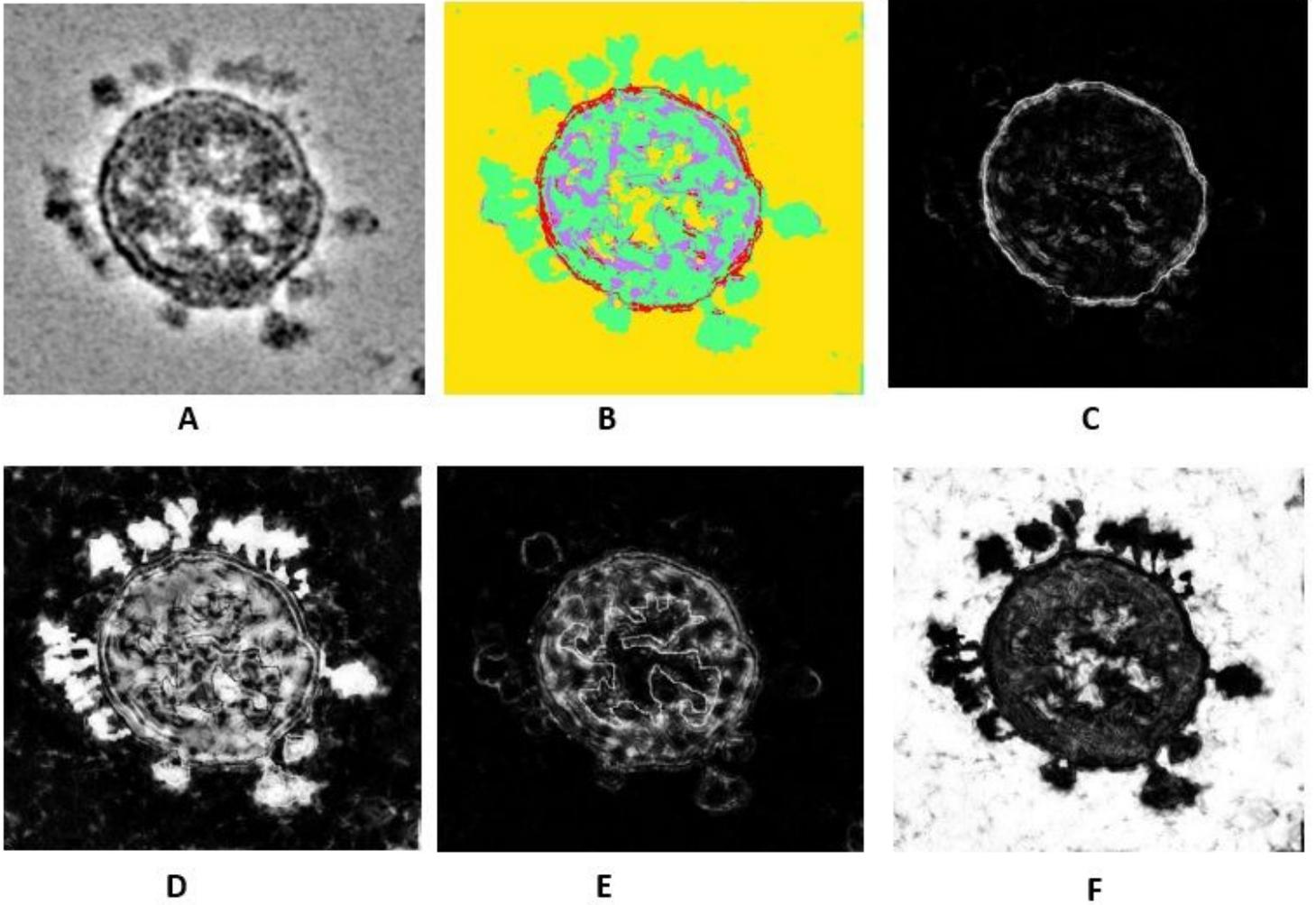
A framework proposal shows for identification of SARS-CoV 2 virus.



**Figure 3**

TEM image of Vero cells infected with a virus: **A.** SARS – CoV and **B.** SARS-CoV 2.

Scale bars = 100 nm

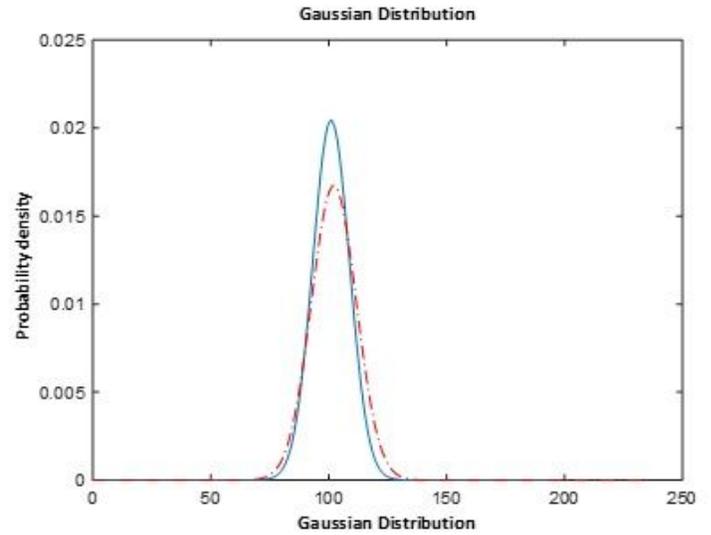
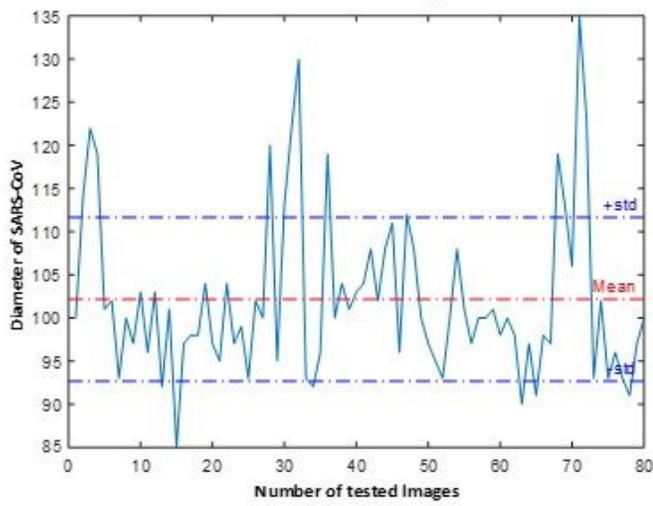
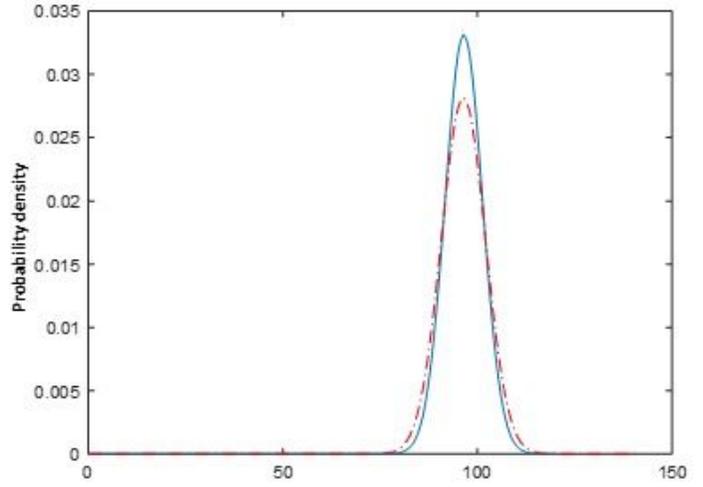
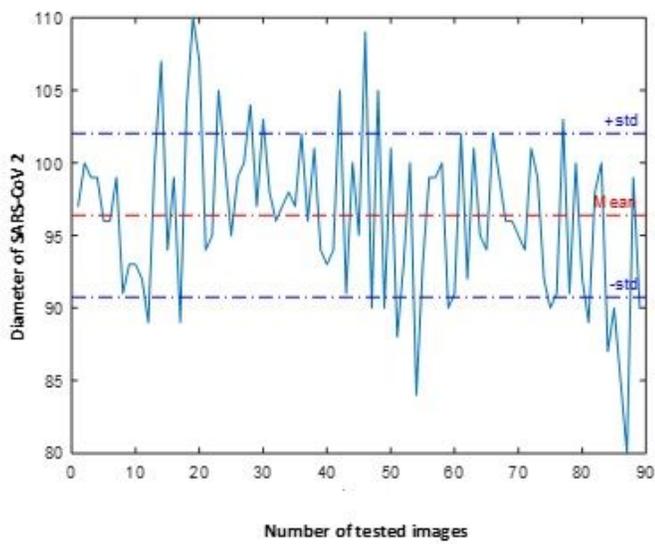


**Figure 4**

Illustrate the identity SARS-CoV 2 feature using trainable weka segmentation: A. input image, B. classified image, C. Envelope, D. spike, E. nucleocapsid, and F. output image.

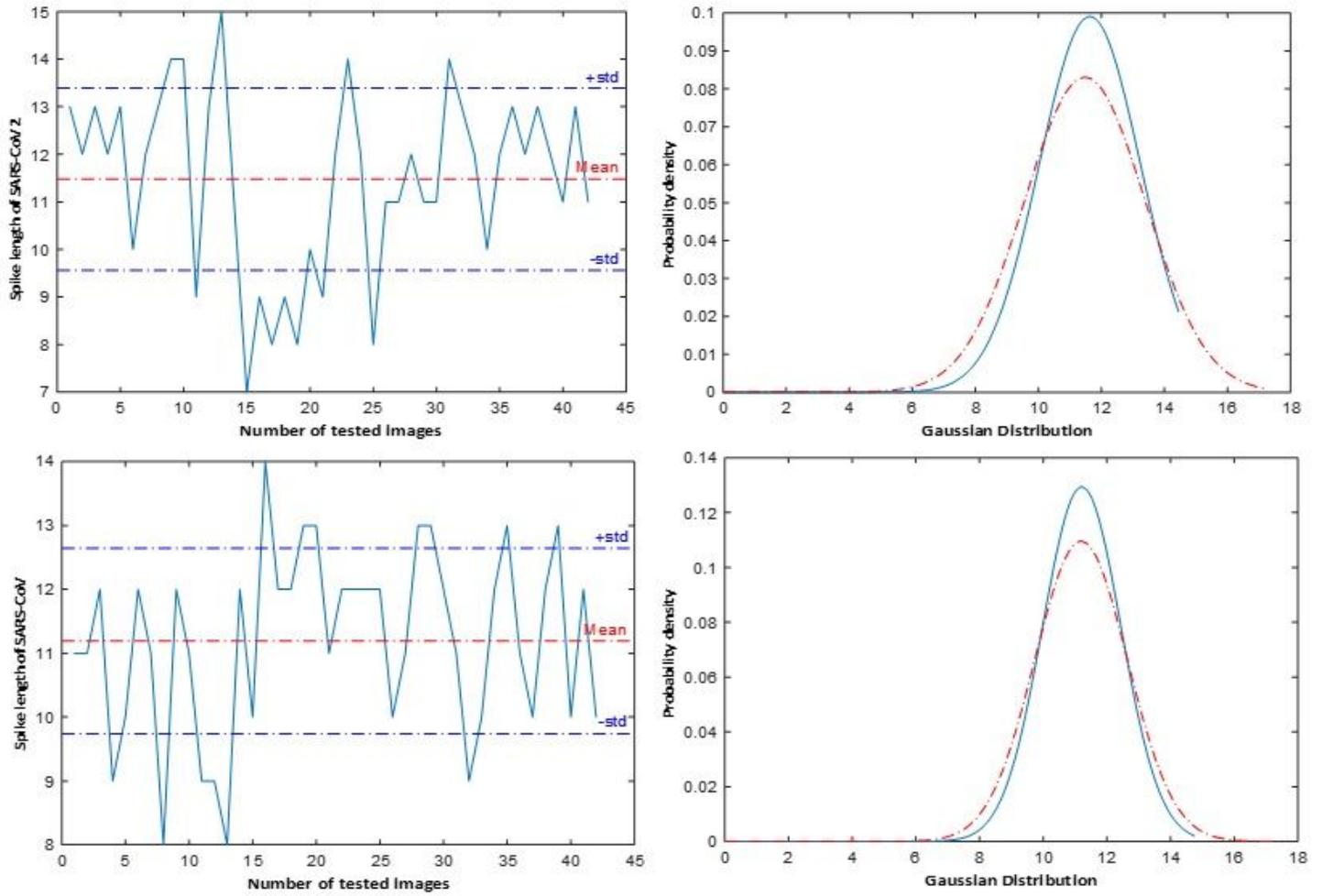
**Figure 5**

Illustrate the framework of extraction features of SARS-CoV 2 virus.



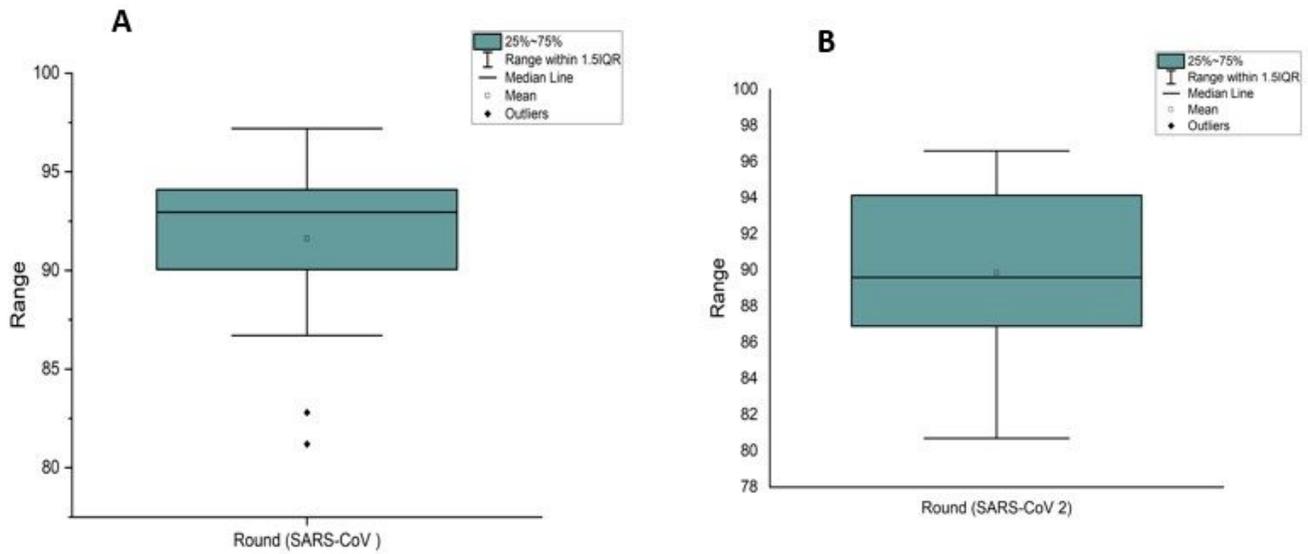
**Figure 6**

Distribution of average diameters Envelop (E) protein of viruses in electron microscopy images at scale bars is 100 nm: A. SARS-CoV2, B. SARS-CoV.



**Figure 7**

Distribution of average length of Spike (S) protein in TEM images at adjusting scale bars is 12 nm: A. SARS-CoV 2, B. SARS-CoV.



**Figure 8**

Illustrate density of the virus roundness: A. SARS-CoV virus. B. SARS-CoV 2 virus.

**Figure 9**

Comparison the circularity between SARS-CoV 2 and another virus.

**Figure 10**

Analysis area size of SARS-CoV 2 and SARS-CoV virus per pixels.