An Analysis of COVID-19 using X-ray Image Segmentation based Graph Cut and Box Counting Fractal Dimension

Faiz Ainur Razi

Mathematics Department, State University of Surabaya, Surabaya, Indonesia
E-mail:faizainurrazi9c@gmail.com

ARTICLE INFO

ABSTRACT

COVID-19 is a disease that spreads relatively quickly. So that many victims are infected by this virus. There are various ways to diagnose the body's infection with the coronavirus. One of them with X-ray results. Detecting COVID-19 with the help of an X-ray sometimes has problems determining the location of the lesion because it is possible because of the large amount of noise in the image. Therefore, the X-ray results will be segmented images using the graph cut algorithm to analyze normal lungs and lungs infected with COVID-19. After obtaining the segmentation results in the form of binary images, the next step is to analyze using the box-counting method's fractal dimensions. From the fractal Dimension results, normal lungs have an average dimension of 1.7890, and lungs infected with COVID-19 have an average dimension of 1.5834. Normal lungs have dimensions larger than lungs infected with the coronavirus due to the lungs' covering by lesions or abnormal conditions in body tissues. This is what causes COVID-19 patients to have complaints of difficulty breathing.

INTRODUCTION

Corona Virus Disease 2019 or COVID-19 is a virus that emerged at the end of 2019. This virus first broke out in Wuhan City, China. Globally in the latest cases, COVID-19 cases remain at around 4.1 million new cases, while recent deaths continue to increase to more than 73,000 from data taken on January 24, 2021. So that cumulatively there are more than 98.2 million cases reported and over 2.1 million deaths globally since the start of the pandemic (World Health Organization, 2021).

Currently, several test tools can be used to detect COVID-19. Combining a swab test and CT scan / X-ray can help individuals get certainty of a more accurate, negative, or positive test result (Xie et al., 2020). From chest CT results, the lesions (abnormal tissue conditions in the body) are usually scattered along with the bronchial bundles or subpleural areas in both lungs (Lin et al., 2020). The results of this X-ray image can solve the problem of image analysis to identify lung diseases, such as COVID-19 (Rahmadewi & Kurnia, 2016).

Edge detection has a boundary between two regions that have different grey levels. The idea used by most edge detection techniques is to use local derivative operator calculations (Setiawaty & Sinurat, 2017). Therefore, medical image segmentation is currently in great demand by researchers and clinicians. There are 5 categories of segmentation methods that are often used, including segmentation based on graph theory, deformation, threshold, area growth, and machine learning (Feng-Ping & Zhi-Wen, 2019).

The size, spatial location, and state of the lesion can be accurately determined through medical image processing. It helps medical personnel qualitatively and even accurately measures diseased tissues and organ problems (Xia et al., 2019). Lung disease will usually be diagnosed using the help of computed tomography (CT) images. This tool has an important role in an efficient medical screening test. Doctors use CT images to analyze and diagnose lung tissue (Ait Skourt et al., 2018). In image segmentation, there is one method for separating the desired object, namely using the graph cut algorithm (Syakrani et al., 2018). In the graph cut algorithm, the application of MAP-MRF and the min-cut/ max-flow algorithm can optimally solve image segmentation. (Lei et al., 2021; Li et al., 2019; Peng et al., 2019).

After determining image segmentation, this study will use the role of fractal geometry to analyze infections that occur in the lungs due to COVID-19. There is a discussion of fractal dimensions in fractal geometry, which has an important role in analyzing complex objects (Cimen et al., 2021; Nayak et al., 2019). This method is also useful for analyzing a disease in the medical world when the Euclidian geometry cannot be used to complete an image analysis.

One of the roles of fractal geometry used to analyze image segmentation results is by using the box-counting method. This algorithm uses the fractal dimensions' calculated square measure to determine the number of cubes needed to cover the totality. We will get the fractal dimensions from this algorithm to analyze the desired object (Davies et al., 2015).

RESEARCH METHODS

The method is used to analyze lung infections due to COVID-19 has four steps, as shown in figure 1. In the first step, we collected a dataset of images and information needed for the experiment. The second step is preprocessing. This second step aims to reduce the noise from the dataset. Furthermore, the image segmentation will be carried out using the graph cut algorithm. After that, the image that has been segmented is analyzed using the box counting method's fractal geometry. The results will be obtained from these steps. As shown in figure 1.
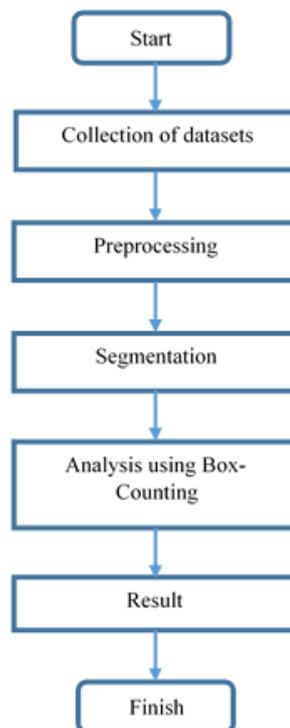

Figure 1. Flowchart of research

1. **Information of Dataset**

    Chest X-ray (CXR) images in this research were used because radiologists widely used these radiographic images in clinical diagnosis (Chowdhury et al., 2020). There are 2 datasets used to form one database, namely COVID-19 Radiography and the patient's lungs. From the two datasets, chest X-rays of COVID-19 infected and normal lungs have been shown.

    a. COVID-19 Radiography

    The COVID-19 Radiography dataset has 1200 positive images of COVID-19 and 1341 normal images (Rahman, 2020). Also, there are 1345 viral pneumonia images, but this data is not shown in this study because it will be focused on analyzing COVID-19 infected lungs and normal lungs. Figure 1 shows a sample from the dataset of normal and positive COVID-19 images.

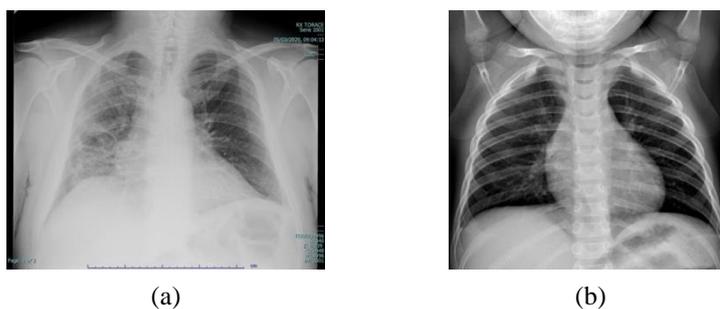(a)   (b)

Figure 1. (a) X-ray results of the patient infected with COVID-19. (b) X-ray results with normal patient's lungs

    b. Patients Lungs

    The Kaggle CXR database is the most popular. In the patient's lungs dataset, there are 70 positive COVID-19 images and 28 CXR images of normal lungs (Nabeel, 2020). For more details, see Table 1.

Table 1. Datasets for experiment

Dataset	Modality	Quantity	
		COVID-19	Normal
COVID-19 Radiography	X-ray	1200	1341
Patients Lungs	X-ray	70	28

2. **Preprocessing**

    In segmenting an image, it is important to pay attention to the quality of the dataset used. This section improves the CXR image's specific parameters and produces an image display by removing unwanted noise in unwanted parts of the background (Cadena et al., 2018).

3. **Segmentation**

    After we arrange the image from the dataset in such a way, the next step segment between the 2 datasets, normal lungs, and lungs infected with COVID-19. Medical image segmentation is expected to have low noise so that the decision-making process can be carried out optimally. Therefore, this research will use the Graph cut segmentation method. The graph cut algorithm's application has a good performance when used for image segmentation (Awate et al., 2019; Jaeger et al., 2014; Kim et al., 2018).

Image segmentation based on graph cut method relates background and object to be a segment, which can be employed as a binary graph labeling problem (Mamatha & Krishnappa, 2019). A set of $\mathcal{L} = \{0,1\}$, where 0 is the background and 1 is the object. Given $\mathcal{P}$ is the set of all the pixels in an image. Then given $\mathcal{N}$ as standard 4 or 8 neighbourhood system. There are pairs of the pixel (p, q), $p < q$. For $f_p \in \mathcal{L}$ be the label to pixel p, and $f = \{f_p | p \in \mathcal{P}\}$ is a set of all label placements. In accordance with the Markov Random Field (MRF) minimization equation, the functions used for segmentation are as follows:

$$E(f) = \sum_{p \in \mathcal{P}} D_p(f_p) + \lambda \sum_{(p,q) \in \mathcal{N}} V_{pq}(f_p, f_q) \qquad (1)$$

In equation (1), there is a term called regional or data. This equation is used to measure the suitability of pixels on an object or background. A simulation image can be seen in Figure 1. $D_p(f_p)$ shows the assignment of label f_p to pixel p. The greater the probability of f_p for pixel p, the smaller $D_p(f_p)$. The second addition of equation (1) is called the boundary. The segmentation limit occurs when there are two different neighboring pixels. $V_{pq}(f_p, f_q)$ is label assignment f_p and f_q to neighbouring pixels. The closest pixels will be labelled the same. If the neighbouring pixels do not have the same label then that label will not be selected. Specifically, $V_{pq}(f_p, f_q) = w_{pq} \cdot I(f_p \neq f_q)$ where $I(\cdot)$ is 1 if $f_p \neq f_q$ and 0 otherwise. In image segmentation, the boundary values are taken from the selection of edges from different angles. So that the boundary is positioned at the edge. Therefore, the selection of edge weight becomes (Boykov, 2001):

$$w_{pq} = e^{-\frac{(I_p - I_q)^2}{2\sigma^2}} \qquad (2)$$

Where $\lambda \geq 0$ and σ as the standard deviation of the intensity between neighboring pixels. Based on the research of Boykov (2001), Eriksson et al. (2006), and Kolmogorov & Zabih (2002), show that making graphics and labeling according to the minimum cut is to optimize the energy in equation (1). And broadly speaking, binary calculation segmentation can be optimized using the graph cut algorithm. For an illustration of image segmentation using the graph cut algorithm, it can be seen in Figure 2.
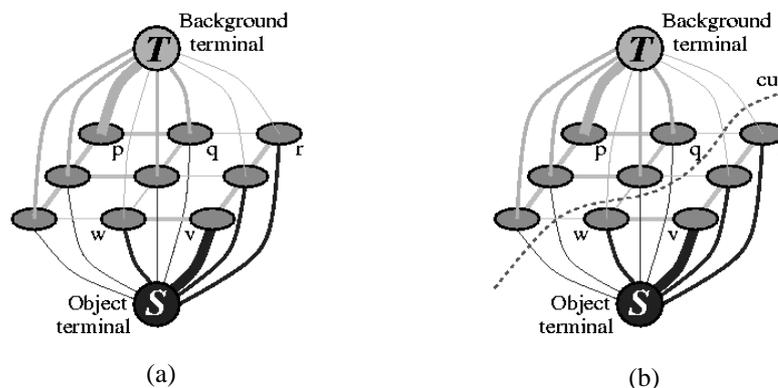

Figure 2. Graph cut illustration for segmentation image. (a) without cutting
and (b) with cutting.

From the dataset that has been obtained, then image segmentation will be carried out. In Figure 3, part (a) is the X-ray result of the lungs taken from the dataset. Then in part (b), using Matlab R2019b, the color grouping is carried out to obtain the desired image segmentation. The green mark

is the foreground or object. Meanwhile, red is the background of the selected image. After determining the foreground and background in Figure 3 (c), a segmentation is formed using the graph cut. Then the segmented image is converted into a binary image, shown in Figure 3 (d).
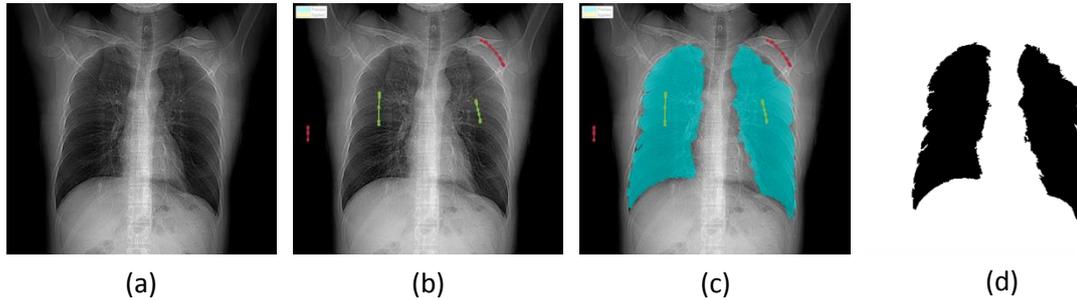

(a)   (b)   (c)   (d)

Figure 3. Image segmentation of lungs using graph cut

4. Analysis

a. Fractal Dimension

Apart from Euclide Geometry, currently developing rapidly in geometry research, namely fractal geometry introduced by Mandelbrot (Mandelbrot, 1967). In analyzing the fractal data, we can take advantage of the fractal Dimension. There are many kinds of fractal dimensions, such as the information dimension, Hausdorff dimension, box-counting, Hurst exponent, etc. Of the many types of fractal dimensions, box-counting is the most widely used to find the self-similarity of data (Ayad et al., 2019; Khotimah & Juniati, 2018; Nayak et al., 2018).

b. Box Counting

The box-counting method for analysis can estimate the fractal dimensions statistically by calculating the self-similarity pattern. The fractal dimensions calculation using the box-counting method is done by covering the object with a square box. In more detail, equation (3) can be used to find the fractal dimensions using the box-counting method:

$$\text{Dim} = \frac{\log(N(r))}{\log\left(\frac{1}{r}\right)} \qquad (3)$$

Where $N(r)$ lots of squares on that cover the image, and r is a scale.

The analysis will be carried out using the box-counting method from the segmentation results in Figure 4 (d). The software used to determine the fractal dimensions is image. In figure 5, the fractal Dimension is 1.6471, with the selection of many boxes covering the object is 256 boxes.
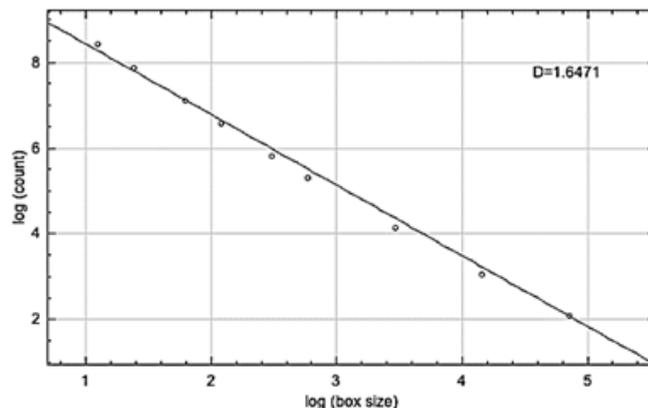

Figure 5. The results of the fractal dimension calculation using the box-counting method

EXPERIMENTAL RESULT

After the segmentation process and analysis using the box-counting method, the next step is to collect fractal dimension value. For each existing image, the segmentation process and dimension calculations are carried out using the box-counting method.

Table 2. The result of the fractal dimension calculation uses box-counting

	COVID-19	Normal Lung
Average	1.5834	1.7890
Max	1.7945	1.8461
Min	0.8239	1.7571

In Table 2 can be seen that the calculation result of fractal dimensions in patients with COVID-19 has an average value of 1.5834. Meanwhile, normal lungs have an average fractal dimension of 1.7890. The maximum value of the fractal dimensions of patients with COVID-19 is 1.7945, and the minimum value of fractal dimensions is 0.8239. For normal lungs, the maximum value for fractal dimensions is 1.8461, and the minimum value is 1.7571. It should be noted that the maximum value of COVID-19 sufferers is that the experiment results are great value. It can happen when the X-ray results of a COVID-19 sufferer have few lesions so that it is not too different from normal lungs. Meanwhile, the average reduction in dimensions between normal lungs and infected lungs with COVID-19 was 11.49%.

CONCLUSIONS AND RECOMMENDATIONS

COVID-19 is a severe disease. Because it disturbs the human respiratory system, the more lesions in the human lung, the worse the patient's complaints will be. With X-rays, it can be seen how badly the coronavirus infects the human respiratory system. The Graph cut algorithm's application can efficiently classify lungs that are covered with lesions and those that are not from X-ray images. Then, from the application of the box-counting method, it can be seen how big the resulting dimensions are between normal lungs and lungs infected with COVID-19. For future research, an automatic segmentation method based on fuzzy and X-ray reconstruction using a deep neural network can also be used.

REFERENCES

Ait Skourt, B., El Hassani, A., & Majda, A. (2018). Lung CT image segmentation using deep neural networks. *Procedia Computer Science*, *127*, 109–113. https://doi.org/10.1016/j.procs.2018.01.104

Awate, S. P., Garg, S., & Jena, R. (2019). Estimating uncertainty in MRF-based image segmentation: A perfect-MCMC approach. *Medical Image Analysis*, *55*, 181–196. https://doi.org/10.1016/j.media.2019.04.014

Ayad, A., Amrani, M., & Bakkali, S. (2019). Quantification of the Disturbances of Phosphate Series Using the Box-Counting Method on Geoelectrical Images (Sidi Chennane, Morocco). *International Journal of Geophysics*, *2019*. https://doi.org/10.1155/2019/2565430

Boykov, Y. Y. (2001). *Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images*. *July*, 105–112.

Cadena, L., Zotin, A., & Cadena, F. (2018). Enhancement of medical image using spatial optimized filters and OpenMP technology. *Lecture Notes in Engineering and Computer Science*, *2233*, 324–329.

Chowdhury, M. E. H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. Bin, Islam, K. R., Khan, M. S., Iqbal, A., Al-Emadi, N., & Reaz, M. B. I. (2020). Can AI help in screening viral

and COVID-19 pneumonia? *ArXiv*.

Cimen, M. E., Boyraz, O. F., Yildiz, M. Z., & Boz, A. F. (2021). A new dorsal hand vein authentication system based on fractal dimension box counting method. *Optik*, *226*, 165438. https://doi.org/10.1016/j.ijleo.2020.165438

Davies, N. A., Harrison, N. K., Keith Morris, R. H., Noble, S., Lawrence, M. J., D'Silva, L. A., Broome, L., Brown, M. R., Hawkins, K. M., Williams, P. R., Davidson, S., & Evans, P. A. (2015). Fractal Dimension (df) as a new structural biomarker of clot microstructure in different stages of lung cancer. *Thrombosis and Haemostasis*, *114*(6), 1251–1259. https://doi.org/10.1160/TH15-04-0357

Eriksson, A. P., Barr, O., & Kalle, A. (2006). Image Segmentation Using Minimal Graph Cuts. *Swedish Symposium on Image Analysis*, 45–48.

Feng-Ping, A., & Zhi-Wen, L. (2019). Medical image segmentation algorithm based on feedback mechanism convolutional neural network. *Biomedical Signal Processing and Control*, *53*. https://doi.org/10.1016/j.bspc.2019.101589

Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., Xue, Z., Palaniappan, K., Singh, R. K., Antani, S., Thoma, G., Wang, Y. X., Lu, P. X., & McDonald, C. J. (2014). Automatic tuberculosis screening using chest radiographs. *IEEE Transactions on Medical Imaging*, *33*(2), 233–245. https://doi.org/10.1109/TMI.2013.2284099

Khotimah, C., & Juniati, D. (2018). Iris Recognition Using Feature Extraction of Box Counting Fractal Dimension. *Journal of Physics: Conference Series*, *947*(1). https://doi.org/10.1088/1742-6596/947/1/012004

Kim, T., Ahn, C., & Lee, O. (2018). Image segmentation by Graph cut for radiation images of small animal blood vessels. *Microscopy Research and Technique*, *81*(12), 1506–1512. https://doi.org/10.1002/jemt.23154

Kolmogorov, V., & Zabih, R. (2002). What energy functions can be minimized via graph cuts? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *2352*, 65–81. https://doi.org/10.1007/3-540-47977-5_5

Lei, L., Xi, F., Chen, S., & Liu, Z. (2021). Iterated graph cut method for automatic and accurate segmentation of finger-vein images. *Applied Intelligence*, *51*(2), 673–689. https://doi.org/10.1007/s10489-020-01828-8

Li, Z., Liu, J., & Cheng, J. (2019). Exploiting Multi-Direction Features in MRF-Based Image Inpainting Approaches. *IEEE Access*, *7*, 179905–179917. https://doi.org/10.1109/ACCESS.2019.2959382

Lin, X., Gong, Z., Xiao, Z., Xiong, J., Fan, B., & Liu, J. (2020). Novel coronavirus pneumonia outbreak in 2019: Computed tomographic findings in two cases. *Korean Journal of Radiology*, *21*(3), 365–368. https://doi.org/10.3348/kjr.2020.0078

Mamatha, S. K., & Krishnappa, H. K. (2019). *Graph Cut Based Multiple Interactive Image Segmentation for Medical Applications*. *7*(3), 567–571.

Mandelbrot, B. (1967). How long is the coast of Britain? Statistical self-similarity and fractional Dimension. *Science*, *156*(3775), 636–638. https://doi.org/10.1126/science.156.3775.636

Nabeel, S. (2020). *COVID-19 Patients Lungs X Ray Images 10000*. https://www.kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images

Nayak, S. R., Mishra, J., Khandual, A., & Palai, G. (2018). Fractal Dimension of RGB color images. *Optik*, *162*, 196–205. https://doi.org/10.1016/j.ijleo.2018.02.066

Nayak, S. R., Mishra, J., & Palai, G. (2019). Analysing roughness of surface through fractal Dimension: A review. *Image and Vision Computing*, *89*, 21–34. https://doi.org/10.1016/j.imavis.2019.06.015

Peng, Z., Qu, S., & Li, Q. (2019). Interactive image segmentation using geodesic appearance overlap graph cut. *Signal Processing: Image Communication*, *78*(April), 159–170. https://doi.org/10.1016/j.image.2019.06.012

Rahmadewi, R., & Kurnia, R. (2016). Klasifikasi Penyakit Paru Berdasarkan Citra Rontgen dengan Metoda Segmentasi Sobel. *Jurnal Nasional Teknik Elektro*, *5*(1), 7. https://doi.org/10.25077/jnte.v5n1.174.2016

Rahman, T. (2020). *COVID-19 Radiography Database*. Kaggle. https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

Setiawaty, I., & Sinurat, S. (2017). Penerapan Algoritma Homogeneity untuk Deteksi Tepi Citra pada Citra Rontgen. *Pelita Informatika Budi Darma*, *16*, 414–417.

Syakrani, N., Widhiyasana, Y., & Efendi, A. A. (2018). Deteksi Tumor Hati dengan Graph Cut dan Taksiran Volume Tumornya. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, *7*(1). https://doi.org/10.22146/jnteti.v7i1.398

World Health Organization. (2021). COVID-19 Weekly Epidemiological Update 22. *World Health Organization*, *January*, 1–3. https://www.who.int/docs/default-source/coronaviruse/situation-reports/weekly_epidemiological_update_22.pdf

Xia, K., Yin, H., Qian, P., Jiang, Y., & Wang, S. (2019). Liver semantic segmentation algorithm based on improved deep adversarial networks in combination of weighted loss function on abdominal CT images. *IEEE Access*, *7*, 96349–96358. https://doi.org/10.1109/ACCESS.2019.2929270

Xie, X., Zhong, Z., Zhao, W., Zheng, C., Wang, F., & Liu, J. (2020). Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: Relationship to Negative RT-PCR Testing. *Radiology*, *296*(2), E41–E45. https://doi.org/10.1148/radiol.2020200343