Unsupervised Image Segmentation Algorithm using Superpixel and Cosine Similarity

Wahyu S J Saputra[1], C Aji Putra, Yisti Vita Via
Faculty of Computer Science
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Surabaya, Indonesia
[1]wahyu.s.j.saputra.if@upnjatim.ac.id

Abstract— **In computer vision, image segmentation is a process of dividing image to get several segments of image. Image segmentation aim to divide image into simple section that meaningful and easy to analyze. Image segmentation regularly use to locate boundary of object in image, so object in image can analyzed. Object area boundary recognized from edge of object, and the edge of object in the image can be recognized from high discoloration. Superpixel is one of popular image segmentation methods. Superpixel has commonly partitioning image into simply part and reduce computation in various computer vision task. On the other side cosine similarity is a method used to analyze the similarity of two objects based on its features. This paper proposed an image segmentation using Superpixel and the sine similarity. Superpixel used to process segmentation and partitioning image into mere parts. Each segment labeled and process to compare with each neighbor using the cosine similarity. The experiment result shows that cosine similarity can be used to recognize similar segment from superpixel segmentation and make the boundary of object more significant.**

Keywords—image; segmentation; superpixel; cosine similarity; unsupervised

I. INTRODUCTION

In computer vision, image segmentation is a process of dividing image to get several segments of image. Image segmentation aims to get simple section of image that is easy to analyze. There are several algorithms to do image segmentation. From the majority of practical, image segmentation can be categorized into supervised and unsupervised image segmentation. Supervised image segmentation mean segmentation processed based on datasets that have used for previous training process, while unsupervised image segmentation no need training process. From image data, image segmentation can categorize into color image and grayscale image oriented. Grayscale image oriented is used to reduce the computation process and make segmentation run faster. Image in computer vision is a kind of data stored in a matrix called pixels, and each element store value of color of each dot pixel in image. Therefore, segmentation in computer vision aims to automatically label the pixels of an image having specific common characteristics. Based on segmentation results, it becomes possible to get regions of interest or object in an image.

Despite recent of supervised image segmentation is used to obtain great result of area identification, such as Q-shift Dual-Tree Complex Wavelet Transform coefficient combined with color component to do a segmentation process based on texture [1] and RACE-Net successfully used to process biomedical data segmentation [2], unsupervised image segmentation have to challenge in a promising research direction. This is related to how the human eye works to do image segmentation in an unsupervised way, on the other unsupervised image segmentation no need to have a vast variety of labeled sample dataset. Cause of that unsupervised image segmentation is more efficient than supervised image segmentation [3].

Image segmentation is an essential step in numerous imaging studies, for example in medical imaging studies or for the object recognition process. There is no a lot of dataset in medical imaging studies thus unsupervised image segmentation best implemented, like Local Center of Mass image segmentation [4] and Spherical K-Means [5]. Majority of medical image is a grayscale image because the medical sensor is X-Ray sensor that use to capture a picture of the patient's body that needs to be analyzed for example bone. Grayscale image segmentation in the medical image can be used to detect brain tumor [6] using Support Vector Machine [7]. In addition to the medical imaging studies, image segmentation is also used in various science.

Unsupervised image segmentation has used in various research, and many of them use a clustering method to do so. K-Means and C-Means algorithm is easy to implement in unsupervised image segmentation [8]. Fuzzy is also clustering algorithm which can be applied in image segmentation method [9, 10]. Grayscale image segmentation has continues to research and use in various with a combination of the algorithm of the firefly and Gaussian Mixture [11]. Unsupervised color image segmentation technique extensively can be categorized into threshold-based method, edge-based method, clustering-based algorithm, region-based approaches, graph-based approaches and hybrid approaches. The color image segmentation simple method is separately first of channel Red, Green, and Blue, after that unsupervised clustering method can be implemented in each channel as grayscale image segmentation using threshold-based method [12]. The threshold-based method is easier to use and apply, a rigid threshold method can cause multiple points scattered in output, which reduces image segmentation result. The region-based method usually uses a set of pixels as the initial seed to grow into a larger and homogeneous image area. However, segmentation efficiency is very dependent on the choice of the initial seed [13]. The edge-based algorithm detects boundaries of the object in images based on the highest

discoloration of gray level [14]. One common problem of the edge-based segmentation algorithm is there are some object or image with gradient color, and this case can lead to under-segmentation. Clustering technique based on color image segmentation is commonly involved with clustering algorithms like k-means or C-means. Clustering technique approach to analyzing pixels as data to be cluster and output of this technique will be clustered pixels in the image as segmentation area of the image [15].

Superpixel segmentation is one of edge-based image segmentation algorithm. Superpixel has commonly used to reduce computation in image segmentation [16]. Superpixel does the segmentation by detecting the edge of a region in image. Region can be extracted from boundaries of edge detected by superpixel. Superpixel also can use in Free-viewpoint image synthesis [17]. Like clustering algorithm commonly, a region that will be detected by superpixel must be set before process. Common problem is image will be segmented in the same numbers as the number of regions set firstly, even though there is just two object in the image. Single object will be segmented in more than one region depending on gray level of the object area. Therefore, superpixel image segmentation requires further process to complete image segmentation for example superpixel combined with Convolutional Neural Network [18] or with Conditional Random Field [19].

Cosine similarity is a clustering method that commonly used for document clustering. Cosine similarity can be used to detect similarities between two object with multi-dimensional feature such as a document that has a multi-dimensional feature obtained from each word contained in the document [20, 21]. Cosine similarity starts to be used for clustering in the digital image [22]. Cosine similarity can be used in image clustering with pixels as a feature to detect similarities image.

This paper proposes an image segmentation algorithm using superpixel and cosine similarity. The process will start with superpixel to detect the edge of the object in the image. Edge of object will guide to regions from boundaries, and then each region labeled. Each region will process to detect similarity with other regions in the image. This process will create similarity matrix between one object to other. Threshold value will be set to decide two regions have similar or not. Cosine similarity detection selected and pixel frequency is used as a feature. Regions that has a high similarity value above the threshold will be labeled as one cluster.

II. RELATED WORK

A. *Digital image*

A digital image is numeric representation two-dimension image. Size of numeric representation is depends on image resolution. Numeric representation is commonly as a matrix, that have rows and column number same as resolution of image. Grayscale image is represented as two-dimensional matrix. Therefore color image is represented as three-dimensional matrix where each channel is represented for each layer of three-dimensional matrix. There is a need to convert an analog image into digital because of ease in storing and processing it using a computer, called digitalization. Two steps of digitalization is

sampling and intensity digitalization. Sampling is treated to get a resolution of the digital image from the continuous image, for example, an image having a size of 10cm in width and 20cm in height will be digitalization in 200 x 400 resolution so each pixel in digital image can represent every 0.1 cm in width and 0.1 cm in height. The intensity of each pixel get from the average value of intensity in those 0.1cm x 0.1cm regions called intensity digitalization.

Image data is enormous, it needs to be stored in a particular format, so that it can be retrieved and process with ease. There are some image format to save in computer out of which bitmap (BMP) format is the most commonly used format [23]. The color image is represented using three dimension matrix and has three coordinate system such as RGB, YIQ, UVW, U*V*W*, etc. beside bitmap format still has another form used to store and display it such as gif, tiff, pix, jpeg. Digital image processing has become an essential tool in the application as follows: satellite, medical, industrial, and robotics.

B. *Image Segmentation*

Many kinds of digital image format can be used to image segmentation processes such as grayscale image or colored image. Image segmentation is a process to labeled pixel into a group of region in Fig 1. Labeled pixel commonly have similar characteristic and can identify as one region. Segmentation output image is usually represented as colored image that has one color for each region selected as one segment.

Fig. 1. Image Segmentation

Each color of an image segmentation output is represented of one segment. Image segmentation specifically used to locate object and boundaries in image [24]. The result of image segmentation is a set of segments that collectively cover the entire image, see Fig.1. Or a set of contours extracted from the image as the result of edge detection. Each pixel in one region as one segment are similar with related to some characteristic or computed property such as color, intensity, or texture. In some of the digital imaging study, image segmentation can use to build 3D reconstruction with the help of an interpolation algorithm.

Some application is use image segmentation in steps, such as content-based image retrieval, computer vision, medical imaging, object detection, recognition task, traffic control system, and video surveillance. Several general-purpose algorithm and technique have developed for image segmentation and combined to be useful with a specific domain.

Threshold-image segmentation method is the most basic image segmentation task. This method based on a clip-level (or threshold value) to get the binary image from gray-image. The key of this method is the selected threshold value (or values when multiple levels is selected). Several popular method is otsu's method (maximum variance) and k-means clustering. Recently, methods have developed for thresholding-computed tomography (CT) images. The key data is different from otsu's method, because the threshold is derived from the radiographs instead of the (reconstructed) image [25, 26].

Clustering method is one method of image segmentation that use the value of K-cluster to identify segment number of image. Basic of clustering method image segmentation is pick K-cluster for the goal of segment number. Assign each pixel in image to the cluster that minimized the distance between the pixel and cluster. The distance of the pixel and cluster can be calculated using each distance algorithm such as Euclidean distance with the formula as follow:

$$d\left(p,q\right)=\sqrt{\sum_{i=1}^{n}\left(q_i-p_i\right)^2} \tag{1}$$

Where p and q represent pixel and cluster, q_i and p_i as feature of a pixel and cluster. Next process of clustering method segmentation is, Re-compute the cluster and averaging all pixel in the cluster until convergence attained.

Beside Euclidean distance, there is some of the distance measurement can be used in a multi-dimensional feature called cosine similarity distance with the formula as follow:

$$similarity=\cos\left(\theta\right)=\frac{A.B}{\|A\|\|B\|}=\frac{\sum_{i=1}^{n}A_iB_i}{\sqrt{\sum_{i=1}^{n}A_i^2}\sqrt{\sum_{i=1}^{n}B_i^2}} \tag{2}$$

Where A_i and B_i are components of vector A and B respectively. The resulting similarity ranges from −1 meaning exactly different, to 1 meaning exactly the same, with 0 indicating orthogonality or decorrelation, while in-between values indicate intermediate similarity or dissimilarity. In the case of information retrieval, the cosine similarity of two documents will range from 0 to 1, since the term frequencies (TF–IDF weights) cannot be negative. The angle between two term frequency vectors cannot be greater than 90° [27]. For text matching, the attribute vectors A and B are usually the term frequency vectors of the documents. Cosine similarity can see as a method of normalizing document length during comparison. This paper, cosine similarity used in digital image processing, therefore vector A and B are pixel frequency of region selected by superpixel segmentation. Basically this paper use cosine similarity to detect similar regions are selected with superpixel segmentation as one region big, with assumption is each object have similar characteristic of pixel in digital image.

C. Superpixel segmentation

Superpixel segmentation is some of clustering method image segmentation. Using K-cluster set firstly, superpixel will detect boundaries of the region as seen in Fig 2. Superpixel detects edge and labeled area with boundaries.

Fig. 2. Superpixel image segmentation

Fig. 2 is superpixel segmentation with 500 cluster set. There is one big car object that scatters in many regions, thus happened because superpixel segmentation is set to clustering image in 500 clusters. There are many kinds of superpixel algorithm, such as normalized cuts [28], was initially the purpose in 2000 [29], graph-based approach [30], a quick shift that can be categorized as gradient ascent method and is a mode –seeking algorithm [31]. Turbopixels [32], Simple Linear Iterative Clustering (SLIC) [33], Compact Superpixels and Constant Intensity Superpixels both approaches defined on grayscale image [34]. Another method of superpixel algorithm are Entropy Rate Superpixel [35], Superpixels via pseudo-Boolean optimization [36], Superpixels extracted via Energy-Driven Sampling [37], Topology Preserved Superpixels [38], and Contour Relaxed Superpixels [39]. This paper use SLIC method for Superpixel segmentation, this algorithm is often used as a baseline and particularly interesting because of its simplicity. SLIC implements a local K-means clustering to generate a superpixel segmentation with K superpixel. Therefore, SLIC can be categorized as gradient accent method.
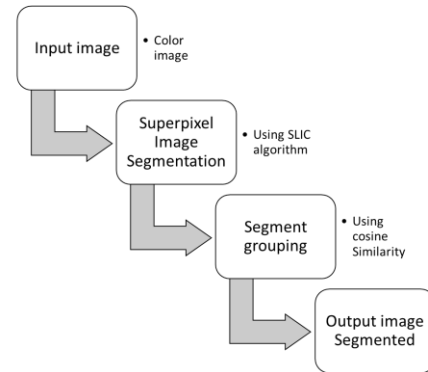
III. PROPOSED ALGORITHM

Fig. 3. Proposed Algorithm

This paper use colored image as input, and Flowchart of the purposed algorithm is shown in Fig. 3. Input color image is firstly processed. The color image is in Red Green Blue color space, then do superpixel image segmentation, and labeled each area selected. This paper use SLIC superpixel image segmentation because SLIC is easy to implements, and there is a ready function in some programming tools. Despite its simplicity, SLIC patch to boundaries as well as or better than

previous methods. At the same time, it is faster and more memory efficient, better segmentation performance, and is straightforward to extend to superpixel generation [40]. The output of the superpixel segmentation process is marked region with an edge that covered an image. Marked edge data is binary data image, that edge is represented as white color or 1 (one/true). With binary edge, image data area can get from the whole image, and do segment save to next process.

For each segment from superpixel compared to all of another section in the image, and generated confusion matrix of similarity between each region. The similarity is calculated using cosine similarity with Pixel Frequency as a feature for each segment. Pixel Frequency of each segment processed in a selected number of the highest frequency of gray-level color, so not all of gray-level will be used as a feature. Pixel frequency is chosen because the size of each segment is different, but gray-level in each segment is constant from zero to 255. Using cosine similarity will obtain a similarity value about 0 to 1. Zero value if regions compared are different, and one if they have similarity. The image is in RGB space color, so before doing cosine similarity each channel must be extracted as two-dimensional data. For each channel will be calculated of its cosine similarity, and cosine value of segment-compared get from an average of three channel and save to confusion matrix of similarity.

By using threshold value regions with the highest similarity, more than threshold will be clustered as one cluster. Output of this process is segmented image having the biggest area segment compared to the superpixel segmentation process only.

IV. CONCLUSION

Topic of this paper focused on image segmentation process using SLIC Superpixel algorithm combine with cosine similarity as post processing. It is necessary to have a methodology to do segmentation in image processing unsupervised for effectively and efficiency, because supervised image segmentation no need big dataset to train. This paper present new approach to optimization technique for unsupervised image segmentation. The proposed algorithm, which inspired by unsupervised image segmentation using superpixel with Simple Linear Iterative Clustering generates segment in number of K-Cluster, combine with multidimensional of similarity measurement using cosine similarity. This algorithm can do unsupervised image segmentation in colored image and grouping some similar segment, cause area of segment is bigger than superpixel segment. Bigger area mean that an object will detected as one segment.

REFERENCES

[1] P. Kennel, C. Fiorio, and Borne, "Supervised image segmentation using Q-Shift Dual-Tree Complex Wavelet Transform coefficients with a texton approach," F. Pattern Anal. Applic., vol. 20, p. 227. 2017.

[2] A. Chakravarty and J. Sivaswamy, "RACE-net: A Recurrent Neural Network for Biomedical Image Segmentation," IEEE J. Biomed. and Health Inf., July 2018

[3] G. Xua, X. Li, B. Lei, K. Lv, "Unsupervised Color Image Segmentation with Color-alone Feature Using Region Growing Pulse Coupled Neural Network," Neurocomput., vol. 306, pp. 1-16, September 2018.

[4] I. Aganj, M.G. Harisinghani, R. Weissleder and B. Fischl, "Unsupervised Medical Image Segmentation Based on the Local Center of Mass," Sci. Rep., vol. 8, Article number: 13012, 2018.

[5] T. Moriya, H.R. Roth, S. Nakamura, H. Oda, K. Nagara, M. Oda, and K. Mori, "Unsupervised pathology image segmentation using representation learning with spherical k-means," SPIE Medical Imaging, 2018, Houston, Texas, United States, 2018.

[6] S.N. Nerurkar, "Brain Tumor Detection using Image Segmentation," Int. J. Eng. Res. in Comput. Sci. and Eng. (IJERCSE), vol. 4, Issue 4, pp. 2394-2320, April 2017.

[7] R. Tak, N. Kumar, Satyaki, S. Verma, S. Dixit, "Segmentation of Medical Image Using Region Based Statistical Model," Int. Conf. on Intell. Comput. and Control Syst. (ICICCS), 2017.

[8] D. S. Mane, B.B. Gite, "Brain Tumor Segmentation Using Fuzzy C-Means and K-Means Clustering and Its Area Calculation and Disease Prediction Using Naive-Bayes Algorithm," Int. J. Innov. Res. in Sci., Eng. and Technol., vol. 6, pp. 2319-8753, November 2017.

[9] S.K. Choy, S.Y. Lam, K.W. Yu, L.W. Yan, K.T. Leung, "Fuzzy model-based clustering and its application in image segmentation," Pattern Recognit., vol. 68, pp. 141-157, August 2017.

[10] S. Yin, Y. Qian, M. Gong , "Unsupervised Hierarchical Image Segmentation Through Fuzzy Entropy Maximization," Pattern Recognit. vol. 68, pp. 245-259, August 2017.

[11] D. Giuliani, "A Grayscale Segmentation Approach Using the Firefly Algorithm and the Gaussian Mixture Model," Int. J. Swarm Intell. Res. (IJSIR), vol. 9, pp. 39-57, January 2018.

[12] Y. Du, C.I. Chang, P.D. Thouin, "Unsupervised Approach to Color Video Thresholding," Opt. Eng., vol. 43, pp. 282-289, 2004.

[13] M. Hosseinzadeh, P.A. Khoshvaght, "Comparative Study of Image Segmentation Algorithms," Int. J. Comput., Electrical, Autom. Control and Inf. Eng., vol. 9, pp. 1966-1977, 2015.

[14] M.A. Ruzon and C. Tomasi, "Color Edge Detection With the Compass Operator," IEEE Conf. on Comput. Vis. and Pattern Recognit. (CVPR '99), vol. 2, pp. 160–166, June 1999.

[15] F. Carvalho, "Fuzzy C-Means Clustering Methods for Symbolic Interval Data," Pattern Recognit. Letters, vol. 28, pp. 423–437, March 2007.

[16] Y. Zhang, X. Li, X. Gao, C. Zhang, "A Simple Algorithm of Superpixel Segmentation With Boundary Constraint," IEEE Trans. on Circuits and Syst. for Video Technol., vol. 27, pp. 1502-1514, July 2017.

[17] M.P. Tehrani, T. Tezuka, K. Suzuki, K. Takahashi, and T. Fujii, "Free-viewpoint Image Synthesis Using Superpixel Segmentation," APSIPA Trans. on Signal and Inf. Process., January 2018.

[18] S. Wang, L. Liu, L. Duan, C. Yu, G. Cai, F. Gao, J. Dong, "Accurate Segmentation of Ulva Prolifera Regions with Superpixel and CNNs," Int. Conf. on Secur., Pattern Anal., and Cybern. (SPAC), 2017.

[19] W. Sultani, S. Mokhtari, H.B. Yun, "Automatic Pavement Object Detection Using Superpixel Segmentation Combined With Conditional Random Field," IEEE Trans. on Intel. Transp. Syst., vol. 19, pp. 2076-2085, July 2018.

[20] M. Steinbach, G. Karypis, V. Kumar, "A Comparison of Document Clustering Techniques," [Online]. Available: https://www.matlabi.ir/wp-content/uploads/bank_papers/ipaper/o27_www.Matlabi.ir_A%20Compa rison%20of%20Document%20Clustering%20Techniques.pdf. [Accessed: 26-Aug-2018].

[21] L. Muflikhah, B. Baharudin, "Document Clustering Using Concept Space and Cosine Similarity Measurement," Int. Conf. on Comput. Technol. and Dev., 2009.

[22] Q. Yan, Y. Ding, J.J. Zhang, Y. Xia, C.H. Zeng, "A discriminated Similarity Matrix Construction Based on Sparse Subspace Clustering Algorithm for Hyperspectral Imagery," Cognit. Syst. Res., February 2018.

[23] M.A. Joshi, "Digital Image Processing, an Algorithmic Approach, Second Edition," PHI Learning Private Limited, pp. 58-57, 2014.

[24] L.G. Shapiro and G.C. Stockman, "Computer Vision," New Jersey: Prentice-Hall, 2001, pp 279-325.

[25] K.J. Batenburg, J. Sijbers. "Adaptive Thresholding of Tomograms by Projection Distance Minimization," Pattern Recognit., vol. 42, pp. 2297–2305, 2001.

[26] K.J. Batenburg, J. Sijbers, "Optimal Threshold Selection for Tomogram Segmentation by Projection Distance Minimization," IEEE Trans. on Med. Imaging, vol. 28, pp. 676–686, June 2009.

[27] G.L. Giller ."The Statistical Properties of Random Bitstreams and the Sampling Distribution of Cosine Similarity," Giller Invest. Res. Notes (20121024/1), 2012.

[28] X. Ren, J. Malik. "Learning a Classification Model for Segmentation", IEEE Int. Conf. on Comput. Vis., pp. 10–17, 2003.

[29] J. Shi, J. Malik. "Normalized Cuts and Image Segmentation", Trans. on Pattern Anal. and Mach. Intell, vol. 22, pp. 888–905, 2000.

[30] P.F. Felzenswalb, D.P. Huttenlocher, "Efficient Graph-Based Image Segmentation," Int. J. of Comput. Vis., vol. 59, pp. 167-181, 2004.

[31] A. Vedaldi, S. Soatto, "Quick Shift and Kernel Methods for Mode Seeking," Eur. Conf. on Comput. Vis., vol. IV, pp. 705–718, 2008.

[32] A. Levinshtein, A. Stere, K.N. Kutulakos, D.J. Fleet, S.J. Dickinson, K. Siddiqi, "TurboPixels: Fast Superpixels Using Geometric Flows", Trans. on Pattern Anal. and Mach. Intell., vol. 31, pp. 2290–2297, 2009.

[33] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, "SLIC Superpixels," IEEE Trans. on Pattern Anal. and Mach. Intell., vol. 34, pp. 2274-2282, 2012.

[34] O. Veksler, Y. Boykov, P. Mehrani, "Superpixels and Supervoxels in an Energy Optimization Framework", Eur. Conf. on Comput. Vis., vol. V, pages 211–224, 2010.

[35] M.Y. Lui, O. Tuzel, S. Ramalingam, R. Chellappa, "Entropy Rate Superpixel Segmentation," IEEE Conf. on Comput. Vis. and Pattern Recognit., pp. 2097–2104, 2011.

[36] Y. Zhang, R. Hartley, J. Mashford, S. Burn, "Superpixels via pseudo-boolean optimization," International Conference on Computer Vision, pp. 1387–1394, 2011.

[37] M. Van den Bergh, X. Boix, G. Roig, B. de Capitani, L. Van Gool, "SEEDS: Superpixels Extracted Via Energy-Driven Sampling," Eur. Conf. on Comput. Vis., pp. 13–26, 2012.

[38] D. Tang, H. Fu, X. Cao, "Topology Preserved Regular Superpixel," Int. Conf. on Multimed. and Expo, pp. 765–768, 2012.

[39] C. Conrad, M. Mertz, R. Mester, "Contour-Relaxed Superpixels. Energy Minimization Methods in Computer Vision and Pattern Recognition," Int. Conf. on Energy Minim. Methods in Comput. Vis. And Pattern Recognit., vol. 8081, pp. 280–293, 2013.

[40] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," IEEE Trans. on Pattern Anal. and Machine Intell., vol. 34, pp. 2274-2282, Nov. 2012.