OPEN ACCESS
J DATA SCI APPL, VOL. 3, NO. 1, PP. 019-030, JANUARY 2020

E-ISSN 2614-7408
DOI: 10.34818/JDSA.2020.3.38

JOURNAL OF DATA SCIENCE AND ITS APPLICATIONS

Snakebite Classification Using Active Contour Model and K-Nearest Neighbor

Chiara Janetra Cakravania, Dody Qori Utama*

School of Computing, Telkom University
Jalan Telekomunikasi No. 1 Bandung, Indonesia

*dodyqori@telkomuniversity.ac.id

Received on 26-08-2020, revised on 08-01-2020, accepted on 11-05-2020

Abstract

The risk of death from snakebite has not been medically handled. The traditional treatment and lack of knowledge of the society to distinguish visually venomous snakebites from the non-venomous one, especially misidentification in visualization, are mains causes of it, particularly in rural areas. This study aims to develop a snakebite identification system using Active Contour Model, K-Nearest Neighbor, image processing tools and confusion matrix. Image processing tools is used in processing of 20 images of snakebite. The Active Contour Model helped to detect bite points on the image. The K-Nearest Neighbor method was used to classify snakebite images into venomous and non-venomous classes whereas confusion matrix is used for performance measurement. Based on parameter testing on K-Nearest Neighbor, we found that the best distance rule is correlation distance with K = 3 and not using the distance weight helps to avoid a poor system performance. Implications of the study include: (a) further development of snakebite visual detection to reduce the risk of death; (b) development of appropriate technology to increase people knowledge on how to handle snakebites accurately. The K-Nearest Neighbor method is more efficient and faster in big data processing according to the needs of society.

Keywords: active contour, K-Nearest Neighbor, death

I. INTRODUCTION

Indonesia, as one of the largest tropic countries, has a high risk of snakebites. Moreover, there is a high number of Indonesian citizens who make their living in agriculture resulting in a high snakebite risk too. Based on data, there were 12,739 – 214,883 snakebite cases in 2017 which caused estimated deaths of 20 – 11,581 victims. This number surely is only based on medical reports from particular hospitals and perhaps different from the actual number. One factor is that many cases occurred in rural areas thus were only treated traditionally. Hence, the estimated number of snakebite cases maybe does not represent the reality [1].

The main cause of death from snakebite cases is by reason of the venom squirted from snake's canine teeth. However, improper treatment as well as poor understanding in identifying whether the snakebites are originated from venomous or non-venomous snakes, are also contributing factors. Regarding that, there is currently no technology-based system to help people to deal with this matter.

Snakebite marks have a great importance to help medical personnel for the identification process [2] [3], and good results will be obtained if the snakebite identification is done by an expert. A technology-based system is needed to facilitate the process of identifying snakes through the bite marks. In this study, we discovered a classification system to identify whether the snakebites are caused by venomous or non-venomous snakes using Active Contour Model in preprocessing and K-Nearest Neighbor in classifying. The Active Contour Model method was chosen based on research conducted by Haijun Li on the "Contour Extraction of Skeletal Based Hand-on Active Contour Model" [4]. We chose K-Nearest Neighbor as the classification method because KNN is very fast in conducting training data, simple and easy to learn, and effective in the process of large-scale training data [5]. Moreover, we also used confusion matrix as the performance measurement method in this study [6].

Theoretical and practical implications of this study are: (a) further development of snakebite visual detection to reduce the risk of death among society; (b) development of simple and appropriate technology to expand society's insight, especially rural citizens, on how to handle snakebites in which the consequences could be fatal.

As for the limitation of this study is that the data consists of 20 images of snakebite marks; 13 of them are originated from venomous snakes while 7 of them are caused by non-venomous snakes. The images have uneven quality and condition. The images were acquired from the preceding author from this study [7]. All images have been cut off beforehand at the bite area with the size of 400x400 pixel. This study focuses on establishing a system that is accurate on preprocessing so that the result will turn out good after classification system.

The next parts of this article include related studies and some theories as well as literature review in the second part. The third part explains the system building process in preprocessing, characteristics extraction and classification. Examination and system analysis process are in the fourth part. Lastly, the fifth part will cover the conclusion and suggestions relating this study.

II. RELATED WORK

There are various kinds of snakes spread in the world, and the snakes are divided into two categories, namely, venomous snakes and non-venomous snakes. To distinguish between the two snakes, we can see from the teeth structure of the snake. Venomous snakes have two canines to squirt their venom while non-venomous snakes do not have any specific teeth structure [8]. Snakebite cases are very common, it is reported [9] that more snakebite cases occur than what we have seen only based on hospital reports. That is because many cases occur in rural areas, or did not get treated in health facilities.

Bite mark is a physical change in the body that is caused by contact or interdigitation between the upper and lower teeth of human or animal, this leads to a wounded tissue. Bite marks caused by wild animals are mainly canines or cone-shaped canines [10].

In 2019, Astrima et al. [11] conducted a study about electrocardiogram signal classification using Principal Component Analysis and Levenberg Marquardt Backpropagation to detect ventricular tachyarrhythmia, and the results obtained are good to be used in a medical examination.

In 1995 in Brazil, a study was established to identify whether the snakebite is originated from venomous or non-venomous snakes. They examined 42 images visually with the result of being able to recognize venomous snakes with a positive predictive value of 89%, and a positive predictive value of 100% of non-venomous snakebites [2]. It is currently necessary to develop research on the classification of venomous and non-venomous snakebites by using image processing. Feature extraction is required in image processing to detect snakebites through images.

Chomboon et al. [12] conducted a study related to various distance rules used in the K-Nearest Neighbor method. There are some distance rules that they examined, then it was concluded that the distance rule techniques of City-block, Chebychev, Euclidean, Mahalanobis, Minkowski, and Standardized Euclidean Technique are distance rules with the best accuracy.

To determine the initial contour on the Active Contour Model is very necessary as it is needed as the first step of the method. Hongshe Dang et al. [13] compared three methods to determine the initial contour automatically on the Geometric Active Contour Model (GACM): watershed method, half manual method and classical segmentation method. The study aimed to compare the segmentation effects and efficiency on the applied model. The results show that the number of evolution based on initial contour using the watershed model is much lower than the other two methods. The more segmentation efficiency is improved, the better the segmented image, even with a complex background.

Adiwijaya et al. [14] found that KNN method can work well with further observation regarding the parameters and attributes that will be used in distance rules in KNN.

Active Contour Model [15] or Snake model is a model used to do object segmentation in images. There is a curve minimizing process in Active Contour Model and this is where this model differs from other segmentation models. The formation of contour curves depends on the accumulation of three types of energy (E_{snake}) as follows.

$$E_{snake} = E_{internal} + E_{external} + E_{constraint} \tag{1}$$

In the processing, Active Contour Model will create an initial contour which surrounds the object, the energy of the $E_{external}$ image affects the shrinkage of the curve, where the pattern follows the object pattern. In the next step, the curve moves toward the object and adjusts to the shape of the object. The snake curve consists of $E_{external}$ which causes the adjustment. The energy is the main factor that changes the shape of the curve in accordance with the desired object shape. The energy calculation uses the following formula.

$$F(C) = \mu.\, length(C) + v.\, area(inside\ C) \tag{2}$$

The variable area $(inside\ C)$ is used when curve change process that only moves inward is desired. To adjust the elasticity of the curve shape we need a variable that is variable μ. As an illustration, for instance variable C is a snake curve, variables c_1 and c_2 represent the image average value (μ_0) of the inner and outer part of the curve, respectively. Then it is assumed that (a) the image (μ_0) is divided into two regions, namely (μ_0^i) (the object to be segmented), (b) (μ_0^o) (background object), and (c) the object pattern C, then the calculation of the external energy of the image can be calculated with the following formula.

$$F_1(C) + F_2(C) = \int_{inside(C)} |\mu_0 - C_1|^2\, dx\, dy + \int_{outside(C)} |\mu_0 - C_1|^2\, dx\, dy \tag{3}$$

Based on the value of the internal energy and th energy from the segmented object (C), the C curve will become larger or smaller. The snake curve shifts by the following condition.

$$\inf_{C} \{F_1(C) + F_2(C)\} \approx 0 \approx F_1(C) + F_2(C) \tag{4}$$

It can be concluded that segmentation will be completed when $F_1(C) \approx 0$ dan $F_2(C) \approx 0$, or when the C curve is exactly at the segmented area. The curve stops changing its shape as the energy value $C = C'$. It is also based on two formulas above that the equation of energy total used to minimalize curve form is as follows.

$$F(C, c_1, c_2) = \mu.\, length(C) + v.\, area(inside\ C) + \int_{inside(C)} |\mu_0 - C_1|^2\, dx\, dy + \int_{outside(C)} |\mu_0 - C_1|^2\, dx\, dy \tag{5}$$

K-Nearest Neighbor (K-NN) is a classification method of a set of data based on a data learning that has been classified previously. The working principle of K-NN is to acquire the nearest distance between test data and 'k' nearest neighbors in train data. Train data is projected into a multidimensional space where each dimension represents the features of the data. The space is divided into parts based on train data classification [16].

In K-NN, there are some distance measurements that can be used, they are as follows [17] [12]:

1. Euclidean Distance,

$$d_{st}^2 = (x_s - y_t)(x_s - y_t)' \tag{6}$$

$where$

$d_{st} = distance\ between\ vector\ x\ to\ y$
$x = data\ x\ vector$
$y = data\ y\ vector$

2. City-block Distance,

$$d_{st} = \sum_{j=1}^{n} |x_{sj} - y_{tj}| \tag{7}$$

$where$

$x = data\ x\ vector$
$y = data\ y\ vector$
$j = sum\ of\ points$

3. Correlation Distance,

$$d_{st}^2 = \left(1 - \frac{(x_s - x'_s)(y_t - y'_t)'}{\sqrt{(x_s - x'_s)(x_s - x'_s)'}\sqrt{(y_t - y'_t)(y_t - y'_t)'}}\right) \quad (8)$$

$where$

$$x'_s = \frac{1}{n}\sum_j x_{sj}$$
$$y'_t = \frac{1}{n}\sum_j y_{tj}$$
$$x = data\ x\ vector$$
$$y = data\ y\ vector$$

III. RESEARCH METHOD

This study used 20 images; 14 images for train data and 6 images for test data. Train data consists of 9 images of venomous snakebite and 5 images of non-venomous snakebite, while test data consists of 4 images of venomous snakebite and 2 images of non-venomous snakebite. The feature from each image was extracted and then divided into train data and test data. Classification class is divided into two, V (venomous) and NV (non-venomous). Figure 1 shows image example of venomous and non-venomous snakebite.

(a) (b)

Fig 1. (a) venomous snakebite, (b) non-venomous snakebite [7]

A. System Design

On this research, system designed to identify snakebite marks used digital image processing with the input of snakebite image and the output of the results of the classification of input data. The system consists of three stages: preprocessing stage, feature extraction stage and classification stage. In this study, we used Active Contour Model and K-Nearest Neighbor classification method. Figure 2 shows the system block diagram.

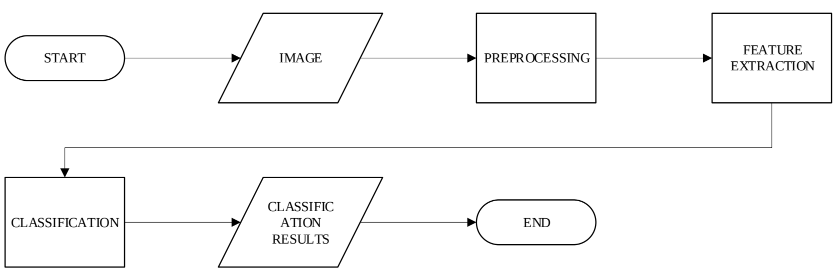

Fig 2. System block diagram

B. System Flow

The system flow consists of two stages: train stage and test stage. Snakebite image is required as the input in the initial process, then the preprocessing stage uses Active Contour (Snake) method, which then the image feature is extracted. The image feature will be stored in database train whereas test image feature will be stored in database test which will be used as data in the classification process. Figure 3 shows the system flow.

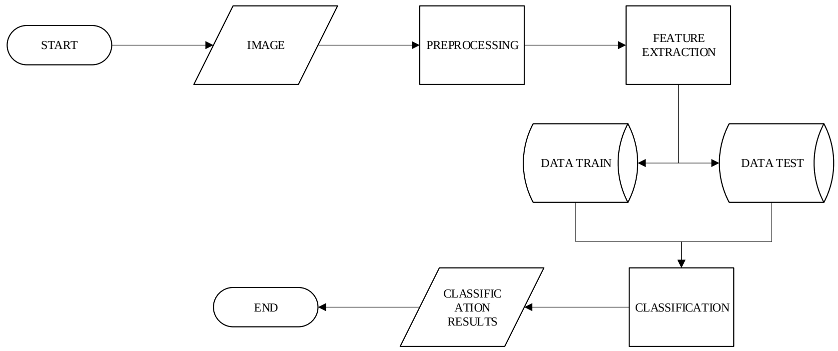

Fig 3. Flow system diagram

Figure 3 illustrates the system flow of stage process. All images that have passed the train and test stage will go through the process of preprocessing and feature extraction. All image features will be stored in database (train data or test data), where the database will be used for reference or comparison data for classification.

C. Preprocessing Stage

Preprocessing stage flow diagram is shown in Figure 4.

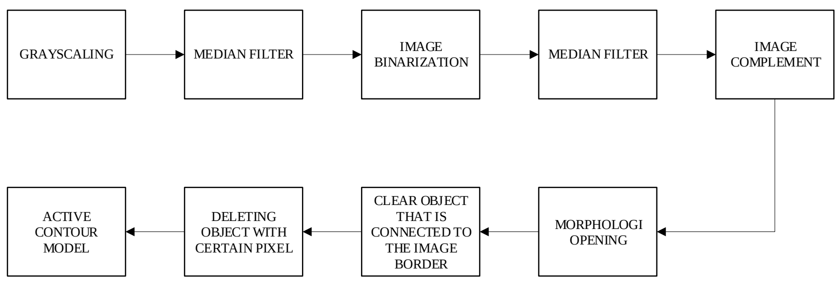

Fig 4. Preprocessing flow diagram

1. Before going through the preprocessing stage, all input images will be converted to grayscale so that the image matrix will have a 2-dimensional size, this stage is called *Grayscaling*.
2. The image will be filtered to create a smoother image data. The filter used is the Median Filter.
3. Binarization is used to change the image pixel value to 0 or 1 only. This stage uses the *imbinarize* function with the adaptive threshold in MATLAB programming.
4. Image complement will change the image value of 1 to 0, vice versa. This process will make the recognized object have the value of 1.
5. *Morphology* with the open operation is a *Morphology* that uses erosion operation then followed by dilation operation, this process is used for separating objects that are neighboring.
6. Objects that are attached to the edge of the image will be removed because it can affect the recognition of objects or bite marks by using the *imclearborder* function in MATLAB programming.
7. Object removal to eliminate noise that is still recognizable. In this stage, objects that are considered noise will be eliminated. The object to be deleted is an object sized between 50 and 1450 pixels.
8. In the final stage, Active Contour method is used to identify bite marks. The Active Contour stage starts with the initial contour, then evolution or initial contour iteration will shrink and detect objects in the image. Active Contour Model is used to help object or bite points detection.

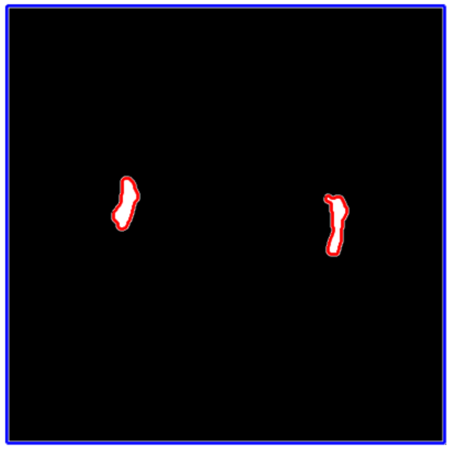

Fig 5. Active contour model process result

Figure 5 shows the result of the Active Contour Model process applied at the preprocessing step. The blue lines indicate initial contour, while the red lines indicate final contour that detect the objects in the image.

After all the preprocessing step was done, an additional step during the process was added to maximize the results of the preprocessing step. However, this additional step was only done to the venomous snake bite images, with only taking two objects that have been detected by the previous processing step. This additional step was not applied to the non-venomous snake bite images. Afterwards, the image was saved in .tiff format.

D. Feature Extraction

Feature extraction is used to extract the features from each image, those features then will be used in the classification stage. The feature is the mean distance between each object or snakebite marks and the number of objects or snakebite marks in the image. The feature will be stored into two variables. The first variable is used to store the mean distance between objects, whereas the second variable is used to store the number of objects. Figure 6 illustrates the feature extraction flow diagram.

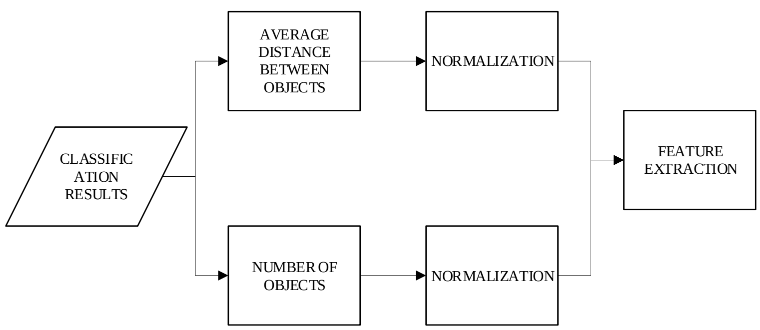

Fig 6. Feature Extraction flow diagram

There is a normalization stage in the feature extraction process. Normalization is the process of scaling the attribute values from the data so that it can lie in a certain range [18]. Several methods for data normalizing are as follows:

1. Min-Max Normalization, is a normalization method by performing a linear transformation of the original data so that it will produce a balance between each data [19].

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{9}$$

$where$

$x_{norm} = normalisation\ value$
$x = normalised\ value$
$\min(x) = minimal\ value\ of\ data$
$\max(x) = maximal\ value\ of\ data$

2. Z-score Normalization, is a normalization method that is based on data mean and standard deviation [18].

$$x_{norm} = \frac{x - \text{mean}(x)}{\text{stdev}(x)} \tag{10}$$

$where$

$x_{norm} = normalisation\ value$
$x = normalised\ value$
$\text{mean}(x) = mean\ value\ of\ data$
$\text{stdev}(x) = standard\ deviation\ of\ data$

3. Decimal Scaling Normalization, is a normalization method by moving the decimal value of the data in the desired direction [20].

$$x_{norm} = \frac{x}{10^i} \tag{11}$$

$where$

$x_{norm} = normalisation\ value$
$x = normalised\ value$
$i = direction\ value$

We used Min-Max Normalization method in this study because its application is very suitable with this research.

E. Classification Stage

The next stage is the classification process to group images based on the features obtained so that classification class can be determined. Classification is divided into two class: venomous (V) and non-venomous (NV). Whereas the classification method used in detecting snakebites is K-Nearest Neighbor. This method works by identifying unknown data classes based on data from the nearest neighbor whose class is known [21].

The obtained distance values are then sorted from the smallest to the largest distance values. Then, we select the K top data has been sorted based on the smallest distance value, the class is determined by the majority of the K top data.

F. System Performance Measurement

System performance measurement in this study uses confusion matrix. Confusion matrix will take record of classification result for each class and then the system will determine sensitivity, specificity and accuracy value. Sensitivity value shows the value of positive-valued class data which is predicted accurately to be positive, while specificity shows the value of negative-valued class data which is predicted accurately to be negative. The accuracy value indicates overall system performance in classifying test data.

$$Sensitivity = \frac{tp}{tp + fn} \tag{12}$$

$$Specificity = \frac{tn}{tn + fp} \tag{13}$$

$$Accuracy = \frac{tp + tn}{tp + fn + tn + fp} \tag{14}$$

True positive (tp) is the number of venomous snakebites that are accurately classified as venomous snakebites, True negative (tn) is the number of non-venomous snakebites that are accurately classified as non-venomous snakebites, False positive (fp) is the number of non-venomous snakebites that are inaccurately classified as non-venomous snakebites, and False negative (fn) is the number of venomous snakebites that are inaccurately classified as venomous snakebites.

In this study, sensitivity value shows the system ability to classify images into venomous snakebite class, whereas specificity shows the system ability to classify images into non-venomous snakebite.

IV. RESULTS AND DISCUSSION

System testing was done to obtain the best parameter value to produce an optimum system. Table I shows the testing scenarios done in this research.

TABLE I
SYSTEM TESTING SCENARIO

No.	Testing Scenario	Aim
1	Change the distance functions, by comparing classification results using several distance functions (Euclidean, City Block, Correlation)	Understand and analyze the best distance function
2	Search for the best nearest neighbor (K) value in the classification system with the K-Nearest Neighbors method. The K values used were 1, 3, 5, 7, 9	Understand and analyze the best nearest neighbor (K) value
3	Understand the effect of the usage of distance weight (1/distance) in the K-Nearest Neighbors classification system	Understand the effect of distance weight

First Scenario

The first test scenario was done to understand the effect of various distance functions in the K-Nearest Neighbors method on the accuracy of the system. The distance functions used in this research are Euclidean, city block and correlation. The objective of this testing is to determine the best distance rule used in this study. Testing was done using 14 train data and 6 test data that were previously selected. Train data consisted of 9 images of venomous snake bites and 5 images of non-venomous snake bites, while test data consisted of 4 images of venomous snake bites and 2 images of non-venomous snake bites. The calculation of sensitivity, specificity and accuracy of the testing is shown in Table II.

TABLE II
SYSTEM TESTING RESULTS USING VARIOUS DISTANCE FUNCTIONS, (A) EUCLIDEAN, (B) CITY BLOCK, (C) CORRELATION

(A)

K	Euclidean		
	Sensitivity	Specificity	Accuracy
1	100%	100%	100%
3	100%	100%	100%
5	100%	50%	83.33%
7	100%	50%	83.33%
9	100%	50%	83.33%

(B)

K	City Block		
	Sensitivity	Specificity	Accuracy
1	100%	100%	100%
3	100%	100%	100%
5	100%	50%	83.33%
7	100%	50%	83.33%
9	100%	50%	83.33%

(C)

K	Correlation		
	Sensitivity	Specificity	Accuracy
1	100%	0%	66.67%
3	100%	100%	100%
5	100%	100%	100%
7	100%	100%	100%
9	100%	100%	100%

Second Scenario

The second test scenario was done to understand the effect of K values used in the K-Nearest Neighbors method on the accuracy of the system. The Correlation distance function was used together with the K-Nearest Neighbors method in the second scenario. The test was performed to determine the change in K value made in this study, and to determine the best K value. The variation of K values used in the second scenario were odd numbers between 1 and 10. Testing was done both on the test data and train data, which consisted of 14 train data and 6 test data.

TABLE III
TABLE OF TESTING RESULTS OF THE VARIATION OF K VALUES ON THE SYSTEM ACCURACY

K	Correlation		
	Sensitivity	Specificity	Accuracy
1	100%	0%	66.67%
3	100%	100%	100%
5	100%	100%	100%
7	100%	100%	100%
9	100%	100%	100%

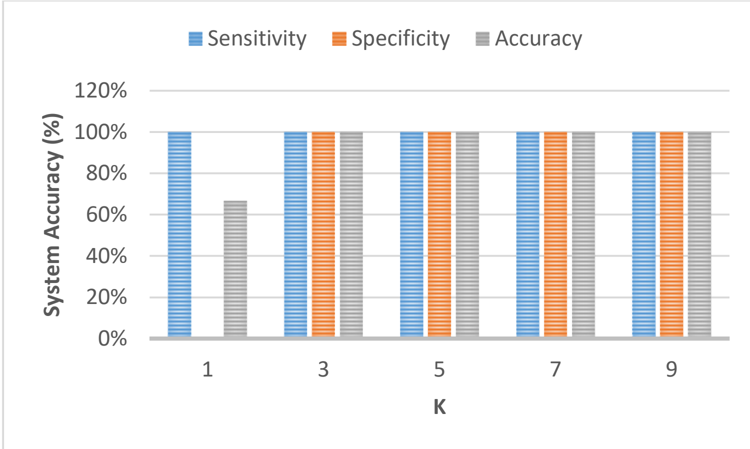

Fig 7. Graph of testing results of the variation of K values on the system accuracy

Third Scenario

The third test scenario was done to understand the effect of distance weight used in the K-Nearest Neighbors method on increasing the system accuracy. Testing by distance weight is intended to increase the level of accuracy in this study. The distance weight used in the third scenario was 1/distance. The distance set-up technique used in this scenario was correlation and the K values used in this scenario were odd number between 1 and 10. Testing was done to 14 train data and 6 test data. The 'No' on Table IV implies that the testing of the classification system did not use distance weight, while the 'Yes' on Table IV implies that the classification system used distance weight.

TABLE IV
THE EFFECT OF THE USAGE OF DISTANCE WEIGHT ON THE CLASSIFICATION SYSTEM

K	No			Yes		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
1	100%	0%	66.67%	100%	0%	66.67%
3	100%	100%	100%	100%	0%	66.67%
5	100%	100%	100%	100%	0%	66.67%
7	100%	100%	100%	100%	0%	66.67%
9	100%	100%	100%	100%	0%	66.67%

Analysis of Testing Results

The usage of Active Contour Model in this research did not show maximum results. The application of this method did not show any significant differences in the system. This is caused by the Active Contour Model that only works well with data with even quality and with single objects. However, the method can still be applied to this system, specifically at the preprocessing step.

It can be seen on Table 3 that there is no significant difference of the sensitivity, specificity and accuracy values with K values 1, 3, 5, 7 and 9 using Euclidean and City Block distance functions. Both distance functions showed identical accuracy values. However, differences in sensitivity, specificity and accuracy values was seen in the usage of the correlation distance function. The correlation technique succeeded in classifying test data with accuracy of 100% at K values K = [3 5 7 9], but failed in classifying data at K=1. Euclidean and City Block techniques could only classify data with accuracy of 100% at K values K = [1 3], this is because in correlation technique the determination of closeness or K value is determined based on the statistical dependency between two vectors and uses modified probabilities.

As shown in Table III and Figure 7, it can be said that the variation of K values affected the system's accuracy in the classification process, this is because the higher the K value the higher the amount of nearest neighbors used in the classification process. K values of 3, 5, 7 and 9 showed sensitivity, specificity and accuracy of 100%, which means that the system has the ability to identify venomous snake bites and non-venomous snakebites correctly.

Table IV shows that distance weight greatly affected the classification system. The sensitivity, specificity and accuracy table shows that the usage of distance weight resulted in unfavorable results. The usage of distance weight resulted in sensitivity of 100%, specificity of 0% and accuracy of 66.67% for all K values, which means all test data in the class 'venomous bite' could be classified correctly, while the class 'non-venomous bite' could not be classified correctly. The classification system showed better results without the application of distance weight in the classification process, with sensitivity, specificity and accuracy of 100% at K values 3, 5, 7 and 9.

V. CONCLUSION

Based on the various test scenarios, it can be concluded that the system has the ability to classify venomous and non-venomous snakebite images with sensitivity, specificity and accuracy up to 100% using the correlation distance function, K=3 value and without applying distance weight. Accuracy of 100% was obtained because of the additional process at the preprocessing step by taking two points or objects of the snake bite detected at the preprocessing step and this process was only done to the venomous snake bite images. This caused the venomous snake bite image to only have two points or objects detected. However, it can be said that the usage of Active Contour Model in this system not maximal, that is because the method did not show a significant difference in detecting the object bite points in the image. The result of a study by Rayiemas [7] shows an effective snakebite detection system with the accuracy value of 80% the value of K = 13.

This research succeeded in obtaining accuracy of 100%. However, the writer is aware that further research must be done to develop an optimal snakebite identification system. The preprocessing step in this research is still considered poor, this is because of the unbalanced data quality which caused the preprocessing algorithm to not have the ability to work maximally to all of the images processed. The result of this study could be used as a reference for further studies so that the more optimal method could be implemented. The writer suggests to use evenly distributed data and use a more optimal preprocessing technique in future research. Future research is also expected to add machine learning algorithms in the preprocessing step that can be useful in detecting bite spots precisely, so that there is no noise detected.

ACKNOWLEDGMENT

The authors are grateful to the Telkom University and other parties who supported this research, which is briefly described in this article.

REFERENCES

[1] R. Adiwinata and E. J. Nelwan, "Snakebite in Indonesia," *Acta Medica Indones. - Indones. J. Intern. Med.*, vol. 47, pp. 358–365, 2015.
[2] S. D. A. Nishioka, P. V. P. Silveira, and F. A. Bauab, "Bite Marks are Useful for the Differential Diagnosis of Snakebite in Brazil," *J. Wilderness Med.*, vol. 6, pp. 183–188, 1995.
[3] D. A. Warrell, *World Health Organization: Guidelines for the clinical management of snake bites in the Southeast Asian region.*, vol. 30 Suppl 1. 2010.
[4] H. Li, "Contour extraction of hand-wrist skeletal based on active contour model," *2009 WRI World Congr. Softw. Eng. WCSE 2009*, vol. 2, no. 4, pp. 283–286, 2009.
[5] N. Bhatia and C. Author, "Survey of Nearest Neighbor techniques," *Int. J. Comput. Sci. Inf. Secur.*, vol. 8, no. 2, pp. 302–305, 2010.
[6] H. Leidiyana, "Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bemotor," *J. Penelit. Ilmu Komputer, Syst. Embed. Log.*, vol. 1, no. 1, pp. 65–76, 2013.
[7] R. M. Putra, Adiwijaya, and D. Q. Utama, "Snake bite classification using Chain code and K nearest neighbour," *J. Phys. Conf. Ser. IOP Publ.*, vol. 1192, no. 1, p. 012015, 2019.
[8] B. S. Gold, Richard C. Dart, and R. A. Barish, "Bites of Venomous Snake," *English J.*, vol. 347, no. 5, pp. 347–356, 2002.
[9] M. A. Mahmood *et al.*, "Snakebite incidence in two townships in Mandalay Division, Myanmar," *PLoS Negl. Trop. Dis.*, vol. 12, no. 7, pp. 1–12, 2018.
[10] H. Mamile, "Analisis Bite Mark dalam Identifikasi Pelaku Kejahatan," 2015.
[11] A. Manik, A. Adiwijaya, and D. Q. Utama, "Classification of Electrocardiogram Signals using Principal Component Analysis and Levenberg Marquardt Backpropagation for Detection Ventricular Tachyarrhythmia," *J. Data Sci. Its Appl.*, vol. 2, no. 1, pp. 78–87, 2019.
[12] K. Chomboon, P. Chujai, P. Teerarassammee, K. Kerdprasop, and N. Kerdprasop, "An Empirical Study of Distance Metrics for k-Nearest Neighbor Algorithm," pp. 280–285, 2015.
[13] H. Dang, Y. Hong, X. Fang, and F. Qiang, "Initial Contour Automatic Selection of Geometric Active Contour Model," *2009 2nd*

Int. Conf. Intell. Comput. Technol. Autom. ICICTA 2009, vol. 2, pp. 66–69, 2009.
[14] Adiwijaya, M. N. Aulia, M. S. Mubarok, W. Untari Novia, and F. Nhita, "A comparative study of MFCC-KNN and LPC-KNN for hijaiyyah letters pronounciation classification system," *2017 5th Int. Conf. Inf. Commun. Technol. ICoIC7 2017, IEEE*, pp. 1–5, 2017.
[15] Z. H. Pradana, "Beef Cattle Weight Determine by Using Digital Image Processing," pp. 13–16, 2016.
[16] E. Budianita, J. Jasril, and L. Handayani, "Implementasi Pengolahan Citra dan Klasifikasi K-Nearest Neighbour Untuk Membangun Aplikasi Pembeda Daging Sapi dan Babi," *J. Sains dan Teknol. Ind.*, vol. 12, no. Vol 12, No 2 (2015): Juni 2015, pp. 242–247, 2015.
[17] W. Hui, "Nearest neighbors by neighborhood counting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 6, pp. 942–953, 2006.
[18] D. A. Nasution, H. H. Khotimah, and N. Charnidah, "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," vol. 4, no. 1, p. 5, 2019.
[19] T. T. Hanifa, A. Adiwijaya, and S. Al-faraby, "Analisis Churn Prediction pada Data Pelanggan PT . Telekomunikasi dengan Logistic Regression dan Underbagging," vol. 4, no. 2, pp. 3210–3225, 2017.
[20] T4Tutorials, "Normalization with decimal scaling in data mining - examples." [Online]. Available: https://t4tutorials.com/decimal-scaling-normalization-in-data-mining/.
[21] A. J. Corresp and A. James, "Snake classification from images," *PeerJ Prepr.*, 2017.