Human height and weight classification based on footprint using gabor wavelet and K-NN methods

Ryan Bagus Wicaksono[1], Suci Aulia[2*], Sugondo Hadiyoso[3], Bambang Hidayat[4]
[1,4]School of Electrical Engineering, Telkom University, Indonesia
[2,3]School of Applied Science, Telkom University, Indonesia
Bandung 40257, Indonesia
*Corresponding email: suciaulia@telkomuniversity.ac.id

Abstract —In the forensic field, height and weight are one of the parameters in identifying a person. Identifying height and weight is usually done manually using scales and height measuring instruments. However, problems are often found when a person's body is separated from one another. One way to estimate weight and height is to measure foot length. A relationship between height, and leg length can be expressed in the correlation coefficient (r), the same as for weight. Therefore, in this study, a system for measuring human height and weight based on images of the footprint is implemented on Android. The methods used in this study are Gabor Wavelet and k-Nearest Neighbor (k-NN). The simulation results generate the best accuracy of 75%. The system can process images with an average computation time of 8.92 seconds. This proposed application is expected to assist forensic experts in identifying a person's height and weight if their body parts are separated from one another. The system can also be used to categorize the ideal body level according to the Body Mass Index (BMI).

Keywords – Feet, height, weight, Gabor Wavelet, Android

I. INTRODUCTION

Human footprints have mainly been known for a variety of reasons. Scientists concentrated their efforts on footprints analysis. Their research such as morphometric or orthopedic diagnosis [1], sex determination [2]-[4], anatomical and anthropological [5]-[7], Pleistocene [8], biometrics [9]-[12], and forensic oriented studies of footprints [13], [14]. Forensic science has several ways to determine a human's height and weight if their bodies have been separated. Human height and weight can be determined by analyzing the soles of the feet. In some countries, footprints are commonly found at crime scenes. If carefully examined, footprints can be found at almost any crime scene [15]. Numerous studies have simulated a system for estimating height based on footprint measurements [16]. However, only a few studies have examined the relationship between height, weight, and footprint. An individual's stature is well understood to be related to body parts such as the head, trunk, limbs, feet, and footprints [14], [17]. There must be a link between a person's body height and weight with their footprints. In the past, some forensic studies have confirmed this relationship.

However, some researchers have studied the relationship between anthropometric foot dimensions with bodyweight [18]-[21]. They used linear regression for weight estimation. Their study did not address height estimation. Therefore, this study proposed a system to classify the human height and weight based on footprint using the android application. This study was carried out in two stages: the training and testing processes. As for training, we use 51 footprint images and 12 footprint images. Gabor wavelet was used in the feature extraction process, and K-Nearest Neighbor was used in the classifying process. Gabor wavelets are a widely used mathematical tool for analyzing mathematical functions and determining the best solutions. Object tracking [22], image compression [23], fabric classification [24], text classification [25], and faces classification [26] are some of its applications. Meanwhile, KNN is the simplest algorithm among other algorithms predicting a class in an imaging sample [27]-[29].

As a guide, this paper is organized as follows. Section II contains methods that been used . Section III contains explanation of the results achieved and. Finally, section IV is conclusion, limitations, future work.

II. MATERIAL AND METHOD

Figure 1 shows an overview of the designed system for estimating human height and weight based on the footprint. This system consists of image acquisition, preprocessing, and feature extraction, and the last stage is height and weight estimation using a classifier. The following subsection contains details of each stage and the methods used in this study.

Fig.1. The proposed system

A. Data Collection

Image acquisition is where the feet' soles are stamped onto an A4 size paper using red ink. The footprint was taken from the right foot, normal or not deformed with an upright tasting position. An example of the footprint image result is shown in Figure 2a, and Figure 2b shows how to measure the foot length. The formula for determining body weight and height using foot length as Equation 1 and 2 follows [21]:

Body weight = 1.29 x $T2_{-Length}$ x 30,0 kg   (1)

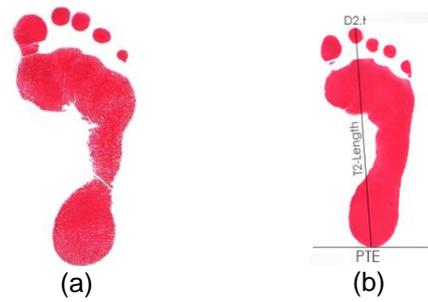Body height = 3.864 x $T_{-2Length}$ + 77.783 cm (2)

Fig.2. (a) The footprint image, (b) How to calculate the footprint length

Digital images are obtained by scanning the footprint and stored in JPEG format. The image is carried out through the preprocessing stage to have good quality and to standardize to produce high accuracy. Furthermore, identification simulation is carried out through two processes, training and testing, as presented in Figure 3. These two phases have a similar process, but in the training phase, there is a process of storing the feature extraction results in a numeric value. The result of feature extraction is then stored in a database which will be used as comparison test image data in the classification process [30]. In the training stage, the pixel value is carried out as a database reference matched in the classification. Meanwhile, the test phase is the process used to test the image so that the proposed system can classify it. In the training stage, used 51 training images. Meanwhile, in the test phase, used 12 testing images.
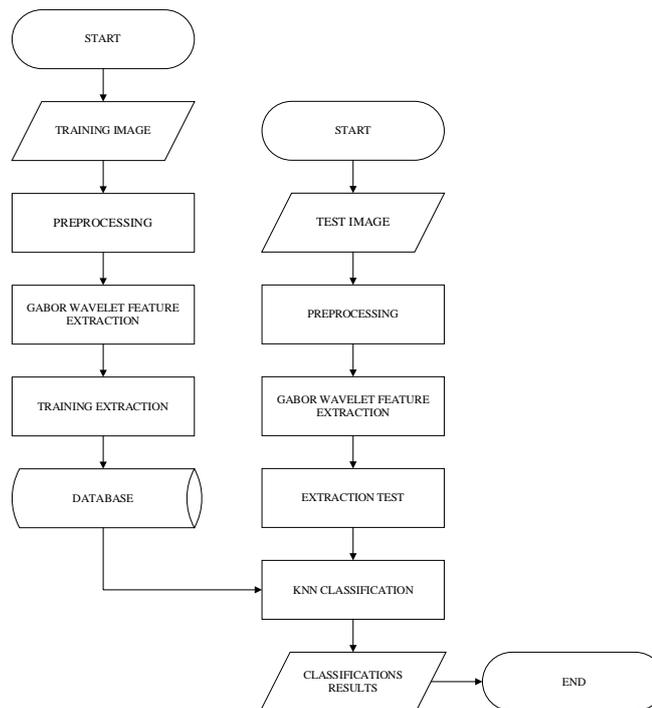

Fig.3. Flowchart of the training and testing phases of the proposed system

B. Image Pre-processing

This research, preprocessing is used to improve the image quality to be processed at the next stage. This process consists of grayscale and black-white conversion using histograms, as shown in Figure 4. Specifically for the black and white conversion, the Otsu thresholding process is carried out in determining the threshold to obtain dynamic values.
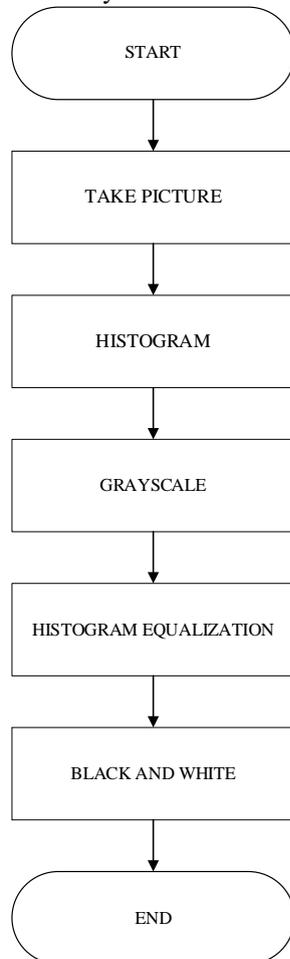

Fig.4. Image preprocessing based on histogram equalization

C. Feature Extraction

In this phase, feature extraction is carried out to obtain characters from the processed image. The feature extraction method used in this study is the Gabor wavelet transform. In this process, there are complex numbers, so in working on the Gabor wavelet transformation, real and imaginary numbers must be separated. Here are the steps to generate a feature vector using Gabor wavelets [31]-[32]:

1. Generating Kernel

The kernel generating process is carried out separately according to the data type (real or complex). This process will produce a feature vector. The feature vector generated by the Gabor wavelet is a combination of the orientation and frequency values.

2. Kernel Convolution

Convolution will be carried out if the Gabor kernel has been formed. This Gabor kernel will then be convoluted with the pixel value of each part of the scanned image of the sole. This process starts from the top-left pixel to the bottom right pixel.

3. Feature Vector

After the convolution is done, each segmented part will produce a feature vector with the mean, variance, and deviation after the convolution is done.

D. Classification using K-Nearest Neighbor (KNN)

KNN works based on the nearest neighboring distance between objects in the following way [24], [33]:

1) It is calculating the distance from all training vectors to test vectors,
2) Take the K value that is closest to the vector value,
3) Calculate the average value. If the value of k = 1, the object is assumed to be a class member of its nearest neighbor [34], [35].

The best value of k depends on the amount of data. In general, the higher the value of k, the lower the noise effect on the classification process. All the parameters resulting from feature extraction will be used as predictors in the classification process. The k-NN classification process in the study uses k = 1 with the type of city block neighbors.

E. Application design

The user interface application is designed to run on the Android mobile platform. The following Figure 5 shows the design of the proposed application.
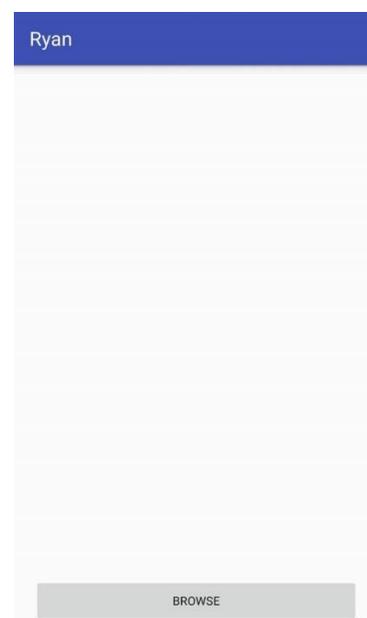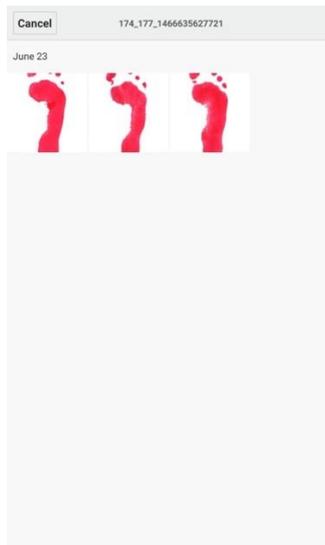
Fig.5. Height and weight detector apps. for Android

K Value	Accuracy	Computing Time (s)
1	42%	8,53
2	42%	8,72
3	42%	8,72
4	42%	8,88
5	42%	8,9
10	75%	8,92

F. System performance evaluation

The proposed system can be evaluated by using formulas that are described in Eq.3 and Eq.4:

1. Accuracy

$$Accuracy = \frac{The\ number\ of\ data\ is\ correct}{Total\ data\ count} \times 100\%$$

(3)

2. Computing Time

$$Computing\ Time = Time's\ up - Start\ Time$$

(4)

III. RESULTS

The test was carried out in three scenarios to determine the performance generated by the system by changing each parameter of the method used. In the first scenario, an experiment was conducted on the k-NN method using the Euclidian distance measurement with k = 1-5 and 10. Then in the second scenario, using a distance city block with k = 1- 5 and 10. In the third test scenario, experiments were also carried out against the k-NN method using the Chebyshev distance with k = 1-5 and 10. The results of the three distances are then compared, and find the best k value from the simulation. Another test parameter is to change the pixel size to 1500x1060. For each test, 12 test images were used.

A. Scenario-1 Test Result

In the first scenario testing, testing using Euclidean distance. The k values used are 1, 2, 3, 4, 5, and 10. The system is tested with an image with a pixel size of 800x566. Table 1 is the test result, including the value of accuracy and computation time.

Table 1. Scenario-1 test result

B. Scenario-2 Test Result

The test is carried out in this scenario using the distance city block type with k and image size equal to scenario-1. The results of the scenario-2 test are shown in Table 2.

Table 2. Scenario-2 test result

K Value	Accuracy	Computing Time (s)
1	41,66%	8,87
2	41,66%	8,65
3	41,66%	8,8
4	41,66%	8,84
5	41,66%	8,97
10	41,66%	9,22

C. Scenario-3 Test Result

They are testing the third scenario using the type distance Chebyshev. With k and the image size is the same as the previous scenario. The results of the scenario-3 test are shown in Table 3.

Table 3. Scenario-3 test result

K Value	Accuracy	Computing Time (s)
1	33,33%	8,75
2	33,33%	8,54
3	33,33%	8,65
4	33,33%	8,77
5	33,33%	8,98
10	41,66%	9,22

D. Scenario-4 Test Result

In the fourth test scenario, an experiment was conducted on the pixel size of the image, which previously used size of 800x566. In this scenario, it was changed to 1500x1060. Then ,the calculation of the distance used is using Euclidean. The results of the scenario-4 test are shown in Table 4.

Table 4. Scenario-4 test result

K Value	Computing Time (s)
1	28,73
2	28,84
3	28,91
4	29,78
5	30,12

E. Scenario-5 Test Result

In testing the fifth scenario, it compares the best results from testing scenarios 1, 2, and 3. The highest accuracy in testing on Android applications can be seen in Table 5 shows the results of the scenario-5 test.

Table 5. Scenario-5 test result

Distance Type	K Value	Accuracy
Euclidean	10	75,00%
City Block	10	41,66%
Chebyshev	10	42%

IV. DISCUSSION

From Table 1, it is known that the fastest computation time in this test is with a value of k = 1 with a computation time of 8.53 seconds. Meanwhile, the best accuracy in this test is with a value of k = 10 with an accuracy of 75%, even though it generates a higher computation time with a delta of not more than one second. From Table 2, it is known that the fastest computation time is 8.65 seconds, with an accuracy value of 41.66% achieved by all k values. From Table 3, it can be seen that the fastest computation time is 8.54 seconds, with the highest accuracy of 41.66% by k = 10. Meanwhile, in scenario-4, it can be concluded that using large pixels will also get a large computation time. This scenario proves that the pixel image size of 1500x1060 produces a large computational time from scenario one testing. In this test, the average computation time is 29.27 seconds. While the test results in the combined scenario, as presented in Table 5, show that the highest accuracy achieved is 75% using the Euclidean distance type. Meanwhile, distance city block and Chebyshev produced 41.66% and 42% accuracy.

The results of several test scenarios presented show that the best combination of parameters from scenarios 1, 2, and 3 succeeded in increasing the accuracy significantly to 75%. In this case, the operation of image morphology, including dilation and erosion on black and white images, greatly affects the estimation of the geometric distance measurement of the foot. Errors in this process significantly contribute to estimating height and weight and can substantially reduce accuracy. At least, the proposed system can prove a correlation between foot size and human height and weight. In the forensic field, this can be useful for describing human posture. The performance validation in this study is considered more valuable when compared to the study by Eboh and Ewamayinma, which only investigated the relationship of height to footprint [36]. This preliminary study provides extensive research opportunities to explore other methods to improve detection accuracy.

V. CONCLUSION

This study has designed a system to estimate height and weight based on foot size. Based on the tests carried out on this system, the following conclusions are obtained. Notably, (1) the best result of the proposed system is 75% accuracy with a computation time of 8.92 seconds, and (2) the system is implemented on Android with a minimum version of 4.4 (KitKat). With this application, it is hoped that it can assist medical forensics in identifying the human's weight and height using an Android mobile device. However, this study still has limitations where the number of samples is relatively small. Therefore, future research is still very wide open to exploring other image processing methods to improve detection accuracy with a larger sample.

ACKNOWLEDGMENT

REFERENCES

[1]. Domjanic, J., Seidler, H., & Mitteroecker, P. (2015). A combined morphometric analysis of foot form and its association with sex, stature, and body mass: Morphometric Analysis of Footprint Form. *American Journal of Physical Anthropology*, *157*(4), 582–591. https://doi.org/10.1002/ajpa.22752

[2]. Sharma, S. K., Jain, S. K., Garg, R., & Sharma, N. (2016). Prediction of Height From Foot Length Among Western Uttar Pradesh Population. *International Journal of Anatomy and Research*, *4*(4.3), 3289–3293. https://doi.org/10.16965/ijar.2016.463

[3]. Abledu, J. K., Abledu, G. K., Offei, E. B., & Antwi, E. M. (2015). Determination of Sex from Footprint Dimensions in a Ghanaian Population. *PLOS ONE*, *10*(10), e0139891. https://doi.org/10.1371/journal.pone.0139891

[4]. Krishan, K., Kanchan, T., & Passi, N. (2011). Estimation of stature from the foot and its segments in a sub-adult female population of North India. *Journal of Foot and Ankle Research*, *4*(1), 24. https://doi.org/10.1186/1757-1146-4-24

[5]. Mehta, U. K., Dhakal, A., Parajuli, S. B., & Shah, S. K. (2019). Prediction of Stature Based on Foot Length In Medical Students of Birat Medical

College & Teaching Hospital, Morang, Nepal. *Birat Journal of Health Sciences*, *4*(2), 707–711. https://doi.org/10.3126/bjhs.v4i2.25441

[6]. Dey, D. S. M., Dr Sayantani Majumdar, Dr Nidhi Lal &. Dr Arijit. (2017). Estimation Of Height From Foot Length In Southern Parts Of West Bengal. *Indian Journal of Medical Research and Pharmaceutical Sciences*. https://doi.org/10.5281/ZENODO.810285

[7]. An, K., Mk, B., Kr, A., & Lakshmiprabhasubhash. (2016). Study on correlation of foot length with stature. *National Journal of Clinical Anatomy*, *05*(04), 191–195. https://doi.org/10.1055/s-0039-3401613

[8]. Pastoors, A., Lenssen-Erz, T., Breuckmann, B., Ciqae, T., Kxunta, U., Rieke-Zapp, D., & Thao, T. (2017). Experience based reading of Pleistocene human footprints in Pech-Merle. *Quaternary International*, *430*, 155–162. https://doi.org/10.1016/j.quaint.2016.02.056

[9]. Marado, L., & Ribeiro, J. (2018). Biological Profile Estimation Based on Footprints and Shoeprints from Bracara Augusta Figlinae (Brick Workshops). *Heritage*, *1*(1), 33–44. https://doi.org/10.3390/heritage1010003

[10]. Khokher, R., & Chandra Singh, R. (2016). Footprint-Based Personal Recognition using Scanning Technique. *Indian Journal of Science and Technology*, *9*(44). https://doi.org/10.17485/ijst/2016/v9i44/105167

[11]. Bb, M., D, A., K, M., P, S., & PK, C. (2012). Estimation of height of an individual from foot length: A study on the population of Odisha. *Int. J. Rev. Life. Sci.*, *2*(2), 7.

[12]. Kamboj, K., Khan, I., & Panday, K. (2018). A study on the correlation between foot length and height of an individual and to derive regression formulae to estimate the height from foot length of an individual. *International Journal of Research in Medical Sciences*, *6*(2), 528. https://doi.org/10.18203/2320-6012.ijrms20180292

[13]. Barker, S. L., & Scheuer, J. L. (1998). Predictive Value of Human Footprints in a Forensic Context. *Medicine, Science and the Law*, *38*(4), 341–346. https://doi.org/10.1177/002580249803800411

[14]. Nataraja Moorthy, T., Mostapa, A. M. B., Boominathan, R., & Raman, N. (2014). Stature estimation from footprint measurements in Indian Tamils by regression analysis. *Egyptian Journal of Forensic Sciences*, *4*(1), 7–16. https://doi.org/10.1016/j.ejfs.2013.10.002

[15]. Bogusz, M. J. (2015). Quality Assurance of Forensic - Toxicological Analysis: Selected Issues = ضمان الجودة في تحاليل السموم الجنائية :قضايا مختارة. *Arab Journal of Forensic Sciences and Forensic Medicine*, *1*(1), 103–113. https://doi.org/10.12816/0011254

[16]. Geetha, G. (2015). Estimation of Stature From Hand and Foot Measurements in a Rare Tribe of Kerala State in India. *JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH*. https://doi.org/10.7860/JCDR/2015/13777.6582

[17]. Agarwal, P., Singh, A., & Gupta, R. (2018). Estimation of human stature by measurements of foot. *Natl J Integr Res Med*, *9*(5), 7.

[18]. 'Uyun, S., & Efendi, T. (2019). Classification of Human Weight Based on Image. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *13*(2), 105. https://doi.org/10.22146/ijccs.35794

[19]. Atamturk, D., & Duyar, I. (2008). Age-Related Factors in the Relationship Between Foot Measurements and Living Stature and Body Weight. *Journal of Forensic Sciences*. https://doi.org/10.1111/j.1556-4029.2008.00856.x

[20]. Fawzy, I. A., & Kamal, N. N. (2010). Stature and Body Weight Estimation from Various Footprint Measurements Among Egyptian Population: Stature and Body Weight Estimation From Various Footprint Measurements. *Journal of Forensic Sciences*, *55*(4), 884–888. https://doi.org/10.1111/j.1556-4029.2010.01372.x

[21]. Krishan, K. (2008). Establishing correlation of footprints with body weight—Forensic aspects. *Forensic Science International*, *179*(1), 63–69. https://doi.org/10.1016/j.forsciint.2008.04.015

[22]. Chao He, Jianyu Dong, Zheng, Y. F., & Ahalt, S. C. (2001). Object tracking using the Gabor wavelet transform and the golden section algorithm. *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No.01CH37164)*, *2*, 1671–1676. https://doi.org/10.1109/ROBOT.2001.932851

[23]. Fischer, S., & Cristobal, G. (2001). Minimum entropy transform using Gabor wavelets for image compression. *Proceedings 11th International Conference on Image Analysis and Processing*, 428–433. https://doi.org/10.1109/ICIAP.2001.957047

[24]. Pradnyana, G. A., Suryantara, I. K. A., & Darmawiguna, I. G. M. (2018). Impression Classification of Endek (Balinese Fabric) Image Using K-Nearest Neighbors Method. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 213–220. https://doi.org/10.22219/kinetik.v3i3.611

[25]. Ramadhan, Nur Ghaniaviyanto. "Indonesian Online News Topics Classification using Word2Vec and K-Nearest Neighbor." Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi) 5.6 (2021): 1083-1089. https://doi.org/10.29207/resti.v5i6.3547

[26]. Vinay, A., Shekhar, Vinay. S., Murthy, K. N. B., & Natarajan, S. (2015). Face Recognition Using Gabor Wavelet Features with PCA and KPCA - A Comparative Study. *Procedia Computer Science*, *57*, 650–659. https://doi.org/10.1016/j.procs.2015.07.434

[27]. Candrasari, E. B., Novamizanti, L., & Aulia, S. (2019). Discrete Wavelet Transform on static hand gesture recognition. *Journal of Physics: Conference Series*, *1367*, 012022. https://doi.org/10.1088/1742-6596/1367/1/012022

[28]. Geler, Z., Kurbalija, V., Radovanović, M., & Ivanović, M. (2016). Comparison of different weighting schemes for the kNN classifier on time-series data. *Knowledge and Information Systems*, *48*(2), 331–378. https://doi.org/10.1007/s10115-015-0881-0

[29]. Aulia, S., Hadiyoso, S., & Ramadan, D. N. (2015). Analisis Perbandingan KNN dengan SVM untuk Klasifikasi Penyakit Diabetes Retinopati berdasarkan Citra Eksudat dan Mikroaneurisma. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, *3*(1), 75. https://doi.org/10.26760/elkomika.v3i1.75

[30]. Syahrian, N. M., Risma, P., & Dewi, T. (2017). Vision-Based Pipe Monitoring Robot for Crack Detection Using Canny Edge Detection Method as an Image Processing Technique. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 243–250. https://doi.org/10.22219/kinetik.v2i4.243

[31]. Barina, D. (2016). Gabor Wavelets in Image Processing. *ArXiv:1602.03308 [Cs]*. http://arxiv.org/abs/1602.03308

[32]. Tai Sing Lee. (1996). Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*(10), 959–971. https://doi.org/10.1109/34.541406

[33]. Nando, P., Putrada, A. G., & Abdurohman, M. (2019). Increasing The Precision Of Noise Source Detection System using KNN Method. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 157–168. https://doi.org/10.22219/kinetik.v4i2.757

[34]. Zhang, S., Li, X., Zong, M., Zhu, X., & Cheng, D. (2017). Learning *k* for kNN Classification. *ACM Transactions on Intelligent Systems and Technology*, *8*(3), 1–19. https://doi.org/10.1145/2990508

[35]. Saragih, T. H., Fajri, D. M. N., & Rakhmandasari, A. (2020). Comparative Study of Decision Tree, K-Nearest Neighbor, and Modified K-Nearest Neighbor on Jatropha Curcas Plant Disease Identification. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 55–60. https://doi.org/10.22219/kinetik.v5i1.1012

[36]. D. E. O. Eboh and A. R. Ewamayinma, "Stature Estimation using Footprint Dimensions among Ukwuani People of South-South Nigeria," *Int. J Forensic Med Invest.*, vol. 5, no. 2, pp. 27–36, 2020.

107