PAPER • OPEN ACCESS

Detection anemia based on conjunctiva pallor level using *k-means* *a*lgorithm

To cite this article: N Sevani *et al* 2018 *IOP Conf. Ser.: Mater. Sci. Eng.* **420** 012101

IOP Conf. Series: Materials Science and Engineering **420** (2018) 012101 doi:10.1088/1757-899X/420/1/012101

Detection anemia based on conjunctiva pallor level using *k-means* *a*lgorithm

N Sevani[1], Fredicia[1], G B V Persulessy[2]

[1]Faculty of Engineering and Computer Science, Departement of Informatics Engineering, UKRIDA, Jl Tanjung Duren Raya, Jakarta, Indonesia
[2] Bachelor Candidate in Faculty of Engineering and Computer Science, Departement of Informatics Engineering, UKRIDA, Jl Tanjung Duren Raya, Jakarta, Indonesia

nina.sevani@ukrida.ac.id, fredi.cia@ukrida.ac.id, grensya.2014tin006@civitas.ukrida.ac.id

Abstract. Anemia is one of the most common types of diseases experienced by people in developing countries but gets less attention to deal. Currently, smartphones applications can be utilized to support the process of anemia detection using an image of the conjunctival pallor. The K-means algorithm can be applied to cluster the pixels of conjunctiva pallor images which are represented in digital characters with red (R), green (G), and blue (B) formats. The clustering results from the under eyes image will be represented by the hemoglobin concentration, which is used to determine whether a person suffers from anemia or not. The function of this application is facilitate the user to perform anemia detection anytime and anywhere. Test results from this application indicate that the detection process gives the accuracy of 90%. This means that the result of detection through the application has a similarity of 90% with the results of detection using the EasyTouch®GCHb hemoglobin test.

1.Introduction

Anemia is one of the world's health problem, especially in the developing countries [1][2][3]. The problem of anemia can occur at various ranges of age from infants, pregnant women, children, and adults. One of the types of anemia is Iron Deficiency Anemia (IDA), which is the most common nutritional deficiency problem in the world [4], especially in developing countries including Indonesia. IDA was caused by an impact from a deficiency of iron in the human body. There are several ways to check the probability that sustain anemia or not by get the medical check-up in alaboratories or go to the doctor to get a consultation. The main problem in some developing countries is cost problem, which makes the people ignored to get a medical check-up [4].

Therefore is needed other inexpensive way and without doing blood checking in a laboratory to know the condition that has anemia or not. Another alternative that can be used to detect anemia is by looking at a person's clinical signs based on the pallor level of the membrane under the eyes, palms, or nails [3]. However, among all the clinical sign options, the eye membrane is a part that has a high accuracy than the others [5]. In addition, paleness of the membrane of the under eyes is part of the body that is generally same in all people in the world and not be influenced by the skin color of the various races and ethnic groups that exist [6].

Determination of paleness level from the under eyes membrane can be done by looking for the correlation between under eyes color and the Hb concentration in the body. That process was done by using the computer technology which can be represented by the RGB (Red, Green, Blue) format from the under eye images [6][7]. The K-means algorithm could be used to grouping the picture of the under eyes membrane to determine or detect anemia [6].

Smartphone technology that has expanded and reached various age ranges for children, adults, to the elderly can be used as an application medium to detect anemia [8][9]. IDA of detection application that can be accessed through the smartphone will make it easier for people to be ably detecting the possibility of experiencing IDA anytime and anywhere without having to go to the laboratory or visit a doctor. So by knowing his health condition as early as possible, they can pay more attention to the iron content of food or drink consumed or immediately visit the doctor to get further treatment if anemia. The creation of anemia detector applications based on conjunctiva pallor level can be used to detect anemia for adult users, who generally busy and difficult to perform medical examinations to a doctor. Therefore, by knowing the condition of their bodies, it was expected to help raise public awareness to maintain and improve healthy lifestyles.

2.Method

2.1. Problem and knowledge identification

This stage was done to know how to detect the probability of anemia in the body by conducting literature studies to determine the criteria for hemoglobin conditions in a person if they have anemia or not. Standardization of hemoglobin level in this research is using EasyTouch® GCHb hemoglobin test. EasyTouch® GCHb is a multiparameter blood test tool that can be used to check blood sugar (Glucose), cholesterol (Cholesterol), and Hemoglobin. This test kit was commonly used as a test medium for monitoring health [10]**.**

2.2. Creating set image training.

In this stage, a database collected by checking hemoglobin for thirty-six volunteers to examine the Hb content in their blood using the GCHb EasyTouch® hemoglobin test and retrieving the membrane under his eyes using the mobile camerawith a minimum specification of 8 MP with f/2.7 aperture. The results of the collected image will get selection and cropping processof the image to select the membrane below the eye that will be process, with the provisions of a predetermined resolution. Then the selected part will be cluster using the k-means algorithm and the clustering result will be saving into the database. Figure 1 is an image of the lower membrane of the eye for anemia detected, and Figure 2 is a cropped image of a previous process with dimensions of 164x104 and a resolution of 72 dpi both vertical and horizontal, to be processed to detect anemia.

Figure 1. Under eye image who will be processed.

Figure 2. Image after manual selection process.

2.3 Clustering

There was a stage to implemented the k-means algorithm to classify pixel values that have been obtained from the previous process. The k-means was used to group each pixel of the image used in the color cluster that has a similarity intensity [11].

The red, green, or blue pixel values will be clustered using the k-means algorithm by segmenting each color that dominates the image. The process of segmentation was done by calculating the value of the centroid of the image from the under the eye to be processed. The closest centroid value is the solution, then all datasets of the centroid members will be measured in distance with another dataset, using equation (2) to find the Euclidean distance. The Euclidean distance value will determine the cluster color dominance of the existing pixels. That will give the value red, green, or blue from the color of domination that segmented to be stored into the existing database. The workings of the k-means algorithm in this study are as follows:

i. Specifies the number of the cluster to be classified according to the research needs. In this study we used three clusters to determine the dominant color representation from the image. Based on the previous study of detecting anemia from the color of the membrane level under the eye, three clusters are obtained from the value of RGB image segmentation [6]. Another research also revealed that to determine the proximity of the distance between the cluster, pixel image is grouping into three clusters to give the result of a matching picture [12][13][14]. In order to segment the dominant RGB color pixel group in the image, we need to determine the exact number of cluster required according to the RGB image representation [15]. The used of three clusters also considered the computation time for the matching process, as more cluster will certainly need more time to process and it can slow down the computation time and reduce the efficiency of the system [14].

ii. Perform the calculation process to get the center value (centroid) cluster by using Equation (1), where $C_k$is the centroid cluster value, $n_k$cluster data total value, and d_i is the distance between each cluster. An example of a calculation process can be seen in Table 1, which shows the initial pixel color values and table 2 which shows the initial cluster value which is the value of the centroid.

$$C_k = \left(\frac{1}{n_k}\right) \sum d_i \quad (1)$$

Table 1. Color Pixel Value

Data	R	G	B
1	179	159	185
2	174	139	161
3	186	175	202
4	188	156	190
5	180	124	149

To get the centroid value, the middle value of each color attribute will be used. Example of the color attribute R with Id 1 is 179: 2 = 89.5. The step has done when each centroid value of each color data was obtaining.

Table 2. First Cluster Value

Cluster	*Initial Centroid*		
	R	**G**	**B**
1	89.5	79.5	92.5
2	87	69.5	80.5
3	93	87.5	101

iii. *Euclidean distance c*alculation for every cluster. Calculation of Euclidean distance with equation (2), where xi is the value of the$x$ centroid cluster, and yi is the value of the y centroid cluster.

$$d_{(x,y)=}\sqrt{(xi-yi)^2+(xi-yi)^2} \quad (2)$$

For example, using pixel data in table 1 and cluster data in table 2, the calculation for cluster 1 is:

$d_{1\text{-}1} = \sqrt{(179-89.5)^2+(159-79.5)^2+(185-92.5)^2}$ = 151,28

data 1 with*cluste*r 2,

$d_{1\text{-}2} = \sqrt{(174-87)^2+(139-69.5)^2+(161-80.5)^2}$ = 137,41

data 1 with*cluster* 3,

$d_{1\text{-}3} = \sqrt{(186-93)^2+(175-87.5)^2+(202-101)^2}$ = 162,81

Based on above result from the calculation, Euclidean distance at data 1 on cluster 1 (d 1-1) is 151.28, while data 1 on cluster 2 (d 1-2) is 137.41. The steps will be repeated for each cluster of each data. The example from the calculation distance of each cluster with the initial data in Table 1 and Table 2 result will show in Table 3.

Table 3. *Cluster* gap

Data	**C1**	**C2**	**C3**
1	151.28	137.41	162.81
2	135.25	150.45	157.73
3	160.35	170.23	150.65
4	154.91	156.39	173.97
5	144.23	163.35	142.42

iv. Moving data to the cluster that has a minimum distances. Based on the results in table 3, the data was exist in cluster 1 (C1) because C1 has the minimum distance when compared with C2 and C3.

v. Calculating centroid of the new cluster. We also have to search the members of the other clusters with the same process. For example for C1 that already get the members, the centroid of the new cluster can be calculated by finding the average value of each data from each color value. This step will be finished when all cluster members were fulfilled and the position of the data cluster is fixed.

vi. Repeat step 3 to step 5 if the new data position for the new centroid value is different. If the new data position for the new centroid value has the same result, then the process is stopped. The processed data will be inserted into the cluster where the data position is located.

After the clustering process is complete, there will be three clusters of dominant colors from the image training.Each cluster has value Red, Green, Blue, and it will be stored in the database.

2.4. Images matching process

It is a process of matching the Red, Green, and Blue pixel value of the dominant pixel colors obtained in the previous process against the set of training data images in the database. This matching process begins with the clustering process to find the dominant color on the inputted images for anemia detection. After getting the RGB set value cluster of the image, the next step is to find the minimum Euclidean distance from the RGB centroid value of the inputted image with the existing RGB set value in the database.

3.Result and Discussion

3.1. Flowchart

Flowchart of the anemia detections application system has show in figure 3. This flowchart illustrates the ongoing process in the anemia detections application system. The process will start with inputthe image of the membrane under the eyes to be processed and the gender. After inputting the image and select the gender, the system will start to clustering the inputted image.

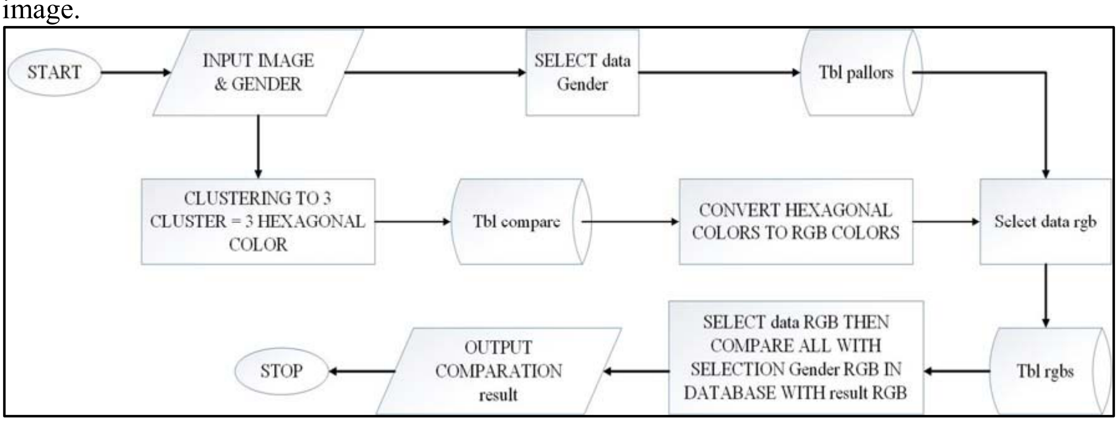

Figure 3. Application flowchart.

3.2. Clustering image trainning

The design of the anemia detection application based on the pallor level of the under eye membrane begins with the clustering stage set of image training that performed on the back end side. The aim of the clustering process is to store the RGB image data that inputted by the user, obtain the RGB value from the image and store the value in the database. Figure 4 displays the application interface for inputting the image**.**

IOP Conf. Series: Materials Science and Engineering **420** (2018) 012101 doi:10.1088/1757-899X/420/1/012101

Figure 4. *Input Set Image Trainning Process.*

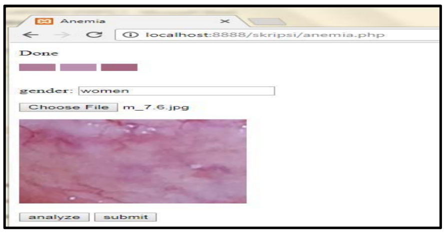

Figure 5. *Input Analysis Process.*

The user also needs to input the gender, as shown in figure 5 and clicking the analyze button to start the system to perform clustering the dominant color in hexadecimal format. The next step is to converse the hexadecimal value into RGB value and store the value to the database, that done by clicking the submit button. The conversion process can be shown in figure 6 and the database can be shown in figure 7**.**

Figure 6. *Hexadecimal Conversion Process.*

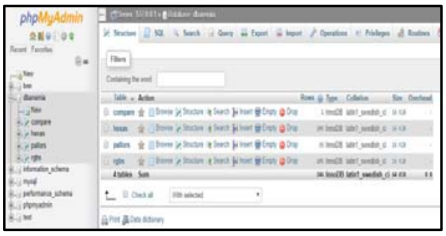

Figure 7. *Database.*

3.3.Evaluation

In the evaluation phase, we use black box testing, white box testing, and accuracy testing by compare the result of the detection process between using application and GCHbEasyTouch® hemoglobintest kits. The black box testing made by displaying the application and distribute the questionnaire about the application to thirty six users. There are three evaluation factors in the questionnaire, such as application display (user interface), functionality of the application, the overall application. The result of the black box testing is shown in figure 8.

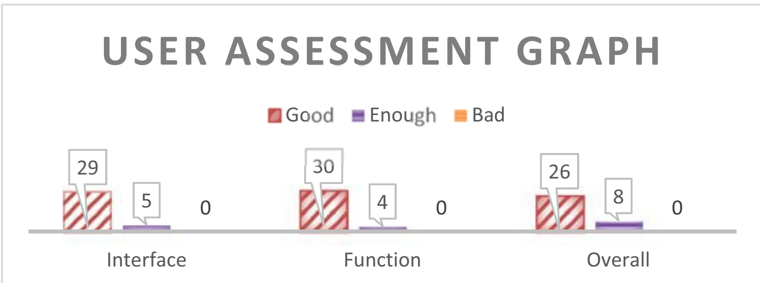

Figure 8. User rating graph.

White box testing is done by testing the logical flow in the application, such as looping and process in the system. There are four categories tested by using white box testing. Table 4 show the result of the white box testing

IOP Conf. Series: Materials Science and Engineering **420** (2018) 012101 doi:10.1088/1757-899X/420/1/012101

Table 4. White Box Testing Table

Examination	**Result**
Input image from the phone gallery	Success
Detection process of the input image	Success
Display the detection result	Success
Give a reference base on the detection result	Success

The accuracy testing is done by comparing the result from the application with the result of the GCHb EasyTouch® hemoglobintest kits. The test involved ten respondents, that consisting of four female and six male. The result of the female respondents show 100% accuracy, but the result for male respondents show there is one case mismatch. Based on the results of the accuracy tests, it can be shown that the accuracy of the anemia detection application is 90%, due to from ten data testing there is one mismatch result.

4.Conclusion

Detecting anemia based on the membrane level under the eye using the K-means algorithm can be done by making an image database of anemia that has been done by clustering to get the dominant color. The use of the K-means algorithm will get the dominant color clustering in the image of the under eyes membrane, so it can be searched the similarity of dominant cluster value with the existing data in the database to get the conclusion of the detection result. From the test results,it can be seen that the results of detection using anemia detection application provides the accurate rate of 90%.

5.References

[1] Critchley, Julia, and Imelda Bates. "Haemoglobin colour scale for anemia diagnosis where there is no laboratory: a systematic review." *International Journal of Epidemiology,* 2005: Vol 34:1425-1434.

[2] Khan, Adeel Ahmed, and Muhammad Masood Kadir. "Accuracy and Use of WHO Hemoglobin Color Scale for Diagnosis of Anemia Among Pregnant Women by Health Care Provider in Periurban Setting in Karachi, Pakistan." *Asia-Pacific Journal of Public Health,* 2015: Vol 27(6): 610-619.

[3] Marn, Heiko, and Julia Alison Critchley. "Accuracy of the WHO Haemoglobin Colour Scale for the diagnosis of anemia in primary health-care setting in low income countries: a systematic review and meta-analysis." *Institute of Tropical Medicine and International Health,* 2016: Vol 4(2): 51-65.

[4] N. Sinha, P. R. Deshmukh and B. S. Garg. "Evaluation of WHO haemoglobin colour scale & palmar pallor for screening of anaemia among children (6-35 months) in rural Wardha, India". *Indian Journal Medis,* 2008: 278-281.

[5] Sheth, Tarang N, Niteesh K. Choudhry, Matt Bowes, and Allan S. Detsky. "The Relation of Conjunctival Pallor to the Presence of Anemia". *Journal Gen Intenr Med,* 1997: 102-106.

[6] Chen, Yi Ming, Shaou Gang Miaou, and Hongyu Bian. "Examing parpebral conjunctiva for anemia assesment with image." *Computer Methods and Programs in Biomedicine,* 2016.

[7] Suner, Selim, Gregory Crawford, John McMurdy, and Gregory Jay. “Non-invasive Determination of Hemoglobin by Digital Photography of Palpebral Conjunctiva.” *The Journal of Emergency Medicine*, 2007: Vol. 33(2): 105-111.

[8] Liu, Na and Ruifeng Yu. ”Identifying Design Feature Factor Critical to Haemoglobin Colour Scale for the diagnosis of anaemia in primary health-care setting in low income countries: a systematic review and meta-analysis.” *Institute of Tropical Medicine and International Health,* 2016: Vol 4(2): 51-65.

[9] Tsai, Tsai-Hsuan, Kevin C. Tseng, Yung-Sheng Chang. “Testing the Usability of Smartphone Surface Gesture on Different Sizes of Smartphones by Different Age Groups of Users.” *Computer in Human Behavior,* 2017.

[10] www.easytouchmeter.com di akses 4 Februari 2018.

[11] Rao, Karri Tata, and Ch.Ramesh. “Baseline JPEG Image Compression with K-Means Clustering Based Algorithm.” *International Journal of Computer Science and Mobile Computing,* 2015: Vol 4: 379-386.

[12] Gamino-Sanchez, Fernando, et al. “Block-Matching Fuzzy C-Means clustering algorithm for segmentation of color images degraded with Gaussian noise.” *Engineering Applications of Artificial Intelligence,* 2018: Vol 73: 31-49.

[13] Dabov, Kostadin, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarin. “Image denoising by sparse 3-D transform-domain collaborative filtering.” *IEEE Transactions on image processing,* 2007: Vol 16: 2080-2095.

[14] Le Capitaine, Hoel, and Carl Frelicot. “A fast fuzzy c-means algorithm for color image segmentation.” *EUSFLAT,* 2011: (pp. 1074-1081).

[15] Dhanachandra, Nameirakpam, Khumanthem Manglem, and Yambem Jina Chanu.“Image segmentation using K-Means clustering algorithm and subtractive clustering algorithm.” *Procedia Computer Science,* 2015: Vol 54: 764-771.