Mobile-based Fish Quality Detection System Using K-Nearest Neighbors Method

by Suhadi Suhadi

Mobile-based Fish Quality Detection System Using K-Nearest Neighbors Method

Conference Paper · November 2020
DOI: 10.1109/ICIC50835.2020.9288544

6 authors, including:

Suhadi Suhadi
STMIK Bani Saleh

Prima Atika
Universitas Bhayangkara Jakarta Raya

Sugiyatno Sugiyatno
Universitas Bhayangkara Jakarta Raya

Ahmad Panogari
STMIK Bani Saleh

Mobile-based Fish Quality Detection System Using K-Nearest Neighbors Method

Suhadi Suhadi
Informatics Engineering
STMIK Bani Saleh
Bekasi, Indonesia 17113
hadims71ndl@gmail.com

Prima Dina Atika
Informatics Engineering
Universitas Bhayangkara Jakarta Raya
Bekasi, Indonesia 17121
prima.dina@dsn.ubharajaya.ac.id

Sugiyatno Sugiyatno
Informatics Engineering
Universitas Bhayangkara Jakarta Raya
Bekasi, Indonesia 17121
sugiyatno@dsn.ubharajaya.ac.id

Ahmad Panogari
Informatics Engineering
STMIK Bani Saleh
Bekasi, Indonesia 17113
apans.p@gmail.com

Rahmadya Trias Handayanto
Computer Engineering
Universitas Islam 45 Bekasi
Bekasi, Indonesia 17113
rahmadya.trias@gmail.com

Herlawati Herlawati
Informatics Engineering
Universitas Bhayangkara Jakarta Raya
Bekasi, Indonesia 17121
herlawati@ubharajaya.ac.id

***Abstract*—Mujair fish (Oreochromis Mossambicus) is a popular consumption fish in Indonesia. The fish is a freshwater fish found in rivers, ponds, and lakes, with salinity of less than 0.05% for breed. Mujair fish is widely consumed by the public as a cheap and tasty fish that are often found in traditional markets or modern markets. The fish is often sold in a fresh condition (fresh) as well as through the process of freezing (frozen). However, consumers sometimes do not fully know the information about the fish condition. Too long storage process causes the physical changes of the fish into blurry eyes, colors tend to fade, soft fish meat texture, and unpleasant smell. In this research, the process to visualize fresh fish images (suitable for consumption) and not fresh (not suitable for consumption) can be detected using K-Nearest Neighbor (K-NN) using a Smartphone. The purpose of this study is to determine the results of the calculation accuracy using the K-Nearest Neighbor (K-NN) Algorithm with the method of digital image processing of Mujair fish, so consumers can choose whether the fish is suitable for consumption or not, using a smartphone as visualization that also can be implemented as additional facility in online shop.**

Keywords— image processing, K-Nearest Neighbor, online shop, android application

I. INTRODUCTION

Small and Medium Enterprise (SME) in Indonesia plays significant role for supporting the domestic economy [1]. Business process has been shifted from conventional markets to online shop from small industries, services, to household product [2]. However, the online shop problems related to the quality of product make the people trust decrease [3]. The study proposed a mobile-based application to detect the quality of product, as a case study, the Mujair fish quality based on the image capture.

Fish is supplied from fishermen, fisheries companies, and fisheries households. The fish should be sold at the places which are located at the local fishing port and after being auctioned, it will be distributed to markets, supermarkets or even to stalls around it. Such fishes can be seen with the naked eye (visually) in good condition or damaged.

Mujair fish is a type of freshwater fish commonly consumed by Indonesian people, the natural distribution of this fish in African waters and in Indonesia was first discovered at the Serang River, south coast of Blitar, East Java in 1939. Its scientific name is Oreochromis mossambicus, and in English known as Mozambique tilapia, or Java tilapia. These fish are freshwater fish, namely fish that spend part or all of their life in fresh water, such as rivers and lakes, with a salinity of less than 0.05%, making it easy to cultivate them [7].

Previous study used the image processing-based to detect the fish disease [4]. After fish diagnostic through the microscope the disease can be identified. Then, the drug can be found for the treatment from two database (Pathogen DB and Drug DB). The processes consisted of preprocessing through principal component analysis (PCA), correlation matching statistically by pattern matching to the disease database and send the result to the farmers through their mobile phones.

Mobile device needs an efficient classification method, e.g. k-nearest neighbors [5], [6], support vector machine [7], neural network [8], [9], etc. The proposed application in this study used K-Nearest Neighbor (KNN) on Android-based mobile device. The main reason of KNN method use in the mobile device implementation is the fact that k-nearest neighbor method is the simplest method that only uses average distance calculation, without for example fuzzification, hierarchy, or other complex calculation found in other methods

The rest of the paper is organized as follows. After discussing the research methodology, the system is designed through the standard software engineering method. The system is tested to check the accuracy and the acceptance to the users.

II. RESEARCH METHODOLOGY

A. Image Processing

Before doing the image processing, a standard methodology was used as shown in Fig 1. It starts from problem definition, approach, development of mobile application, to public implementation, measure the accuracy and conclusion.

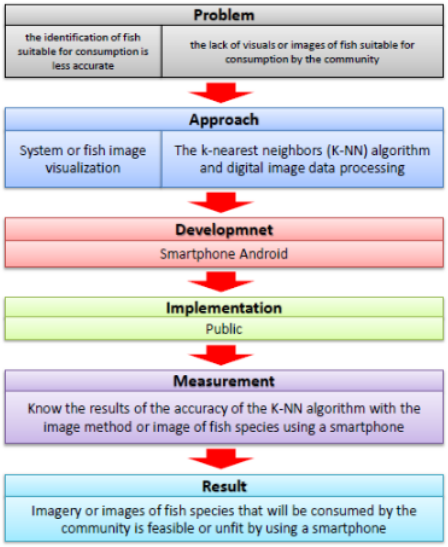
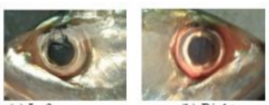


Fig. 1. Research Metodology

RGB (Red, Green, Blue) is an additive color model, where three beams of light are added together, by adding wavelengths, to create the final color. The RGB function is used to analyze the coverage of the fish eye being analyzed. This provide a value on the focus of the fish eye as shown in Fig 2.



(a) Left eye (b) Right eye

Fig. 2. Eye Function Focus

Equation (1) – (5) shows the normalized RGB calculation:

$$\bar{X} = \sum_{n=1}^{1} Xi \quad (1)$$

$$\bar{X}_y = \frac{\left[\frac{R_y + G_y + B_y}{3}\right]_2 + \left[\frac{R_y + G_y + B_y}{3}\right]_R}{2} \quad (2)$$

$$R_y = \left[\frac{(R_i + R_r)}{2}\right] \quad (3)$$

$$G_y = \left[\frac{(G_i + G_r)}{2}\right] \quad (4)$$

$$B_y = \left[\frac{(B_i + B_r)}{2}\right] \quad (5)$$

Variable R, G, B are red value, green value, and blue value, respectively with the standard value as shown in Table I.

TABLE I. RGB COLOR FUNCTION

Color	Nominal Range	White	Yellow	Oyan	Green	Magenta	Red	Blue	Black
R	0 to 255	255	255	0	0	255	255	0	0
G	0 to 255	255	255	255	255	0	0	0	0
B	0 to 255	255	0	255	255	255	0	255	0

RGB result values can be seen in Table II with membership functions for low and high range.

TABLE II. RGB RESULT VALUE

Method	Input	Range	
		Low	High
$\bar{X}_{RGB}$	Eye	89-122	104-136
	Gill	71-115	94-132
RGB_{EG}	Red eye	122-164	145-182
	Green eye	73-110	97-151
	Blue eye	60-98	89-116
	Red gill	87-142	126-172
	Green gill	85-118	61-107
	Blue gill	64-109	85-119

Table III shows the range value for low, medium, and high. This output range limit value makes the calculation simple with the membership function that is usually implemented in fuzzy method.

TABLE III. OUTPUT RANGE LIMIT VALUE

Output	Range
Low	0 - 0.4
Medium	0.2 – 0.8
High	0.6 - 1

Fish sample then analyzed for training data used for classification as shown in Table IV. The study used 25 fish sample with various class based on the fish day.

TABLE IV. RESULT TEST VALUE

Sample	Days				
	1	2	3	4	5
1	0.83	0.92	0.716	0.5	0.343
2	0.96	0.25	0.5	0.5	0.343
3	0.5	0.5	0.5	0.5	0.39
4	0.94	0	0.719	0.1	0.025
5	1	0.25	0.5	0.75	0.5
6	0.885	0.648	0.698	0.945	0.055
7	1	0.75	0.353	0.03	0.135
8	1	0.5	0.98	0.925	0.688
9	0.5	0.5	0.5	1	1
10	0.365	0.5	0.5	0.75	0.025
11	1	0.698	0.75	0.25	0.5
12	0.5	0.7	0.75	0.75	0.04
13	0.5	0.5	1	0.5	0.5
14	0.25	0.5	0.5	0.662	0.615
15	0.25	0.75	0.75	0.925	0.5
16	0.25	0.75	0.995	0.5	0.716
17	0.25	0.335	0.5	0.5	0.33
18	0.5	0.25	0.89	0.5	0.25
19	0.5	0.75	0.5	0.25	1
20	0.11	0.5	0.385	0.5	0.5
21	0	0.645	0.342	0.5	0.3

Sample	Days				
	1	2	3	4	5
22	0.679	0.64	0.745	0.662	0.92
23	0.75	0.95	0.666	0.25	0.338
24	0.61	0.89	0.653	0.925	0.615
25	0.5	0.5	0.5	0.5	0.37

The K-NN algorithm is a method for classifying objects based on learning data which is the closest distance to the object [10]. The formula for finding the K-NN algorithm is as shown in equation (6).

$$u(x, c_i) = \frac{\sum_{k=1}^{K} u(x_k, c_i) * d(x, x_k)^{\frac{-2}{(m-1)}}}{\sum_{k=1}^{K} d(x, x_k)^{\frac{-2}{(m-1)}}} \quad (6)$$

Where T is new case, S is cases in data store, n represents the attribute number in each case, I is individual attribute from 1 to n, f is similarity function of I attribute between case T and S, and w represents weight of attribute i.

Digital image is a two-dimensional image that can be displayed on a computer screen as a set/discrete digital values called pixels/picture elements. In a mathematical view, image is a continuous function of light intensity in a two-dimensional plane [11]. the formulas for finding a digital image are shown (7) – (11):

$$\boldsymbol{r} = \frac{\boldsymbol{R}}{\boldsymbol{R+G+B}}, \boldsymbol{g} = \frac{\boldsymbol{G}}{\boldsymbol{R+G+B}}, \boldsymbol{b} = \frac{\boldsymbol{B}}{\boldsymbol{R+G+B}} \quad (7)$$

$$v = \max(r, g, b) \quad (8)$$

$$S = \begin{cases} 0, & jika\ V = 0 \\ 1 - \frac{\min(r,g,b)}{v}, & V > 0 \end{cases} \quad (9)$$

$$H = \begin{cases} 0, & jika\ S = 0 \\ \frac{60*(g-b)}{S*V}, & jika\ V = r \\ 60 * \left[2 + \frac{b-r}{S*V}\right], & jika\ V = g \\ 60 * [4] + \frac{r-g}{S*V}, & jika\ V = b \end{cases} \quad (10)$$

$$H = H + 360\ jika\ H < 0 \quad (11)$$

Where each element in a digital image (meaning matrix element) is called an image element, picture element or pixel or mop. Therefore, an NM-sized image has a pixel NM. For example, suppose a unit is 256x256 pixels and is represented numerically by a matrix consisting of 256 rows (indexed from 0 to 255) and 256 columns (indexed from 0 to 255).

Adobe OpenCV (Open Computer Vision), is an API (Application Programming Interface) specifically designed to develop digital image processing and machine learning applications. There are many libraries and algorithms that have been provided by OpenCV such as face recognition, character recognition, object tracking, machine learning, object classification, tracking camera movements, extracting 3D models, etc. [12].

III. Result and Discussion

In this section, the testing process was discussed on the Mujair fish eye. The test aims to determine whether the calculations made can determine the digital image eye class (RGB) as indicated in Fig 3.

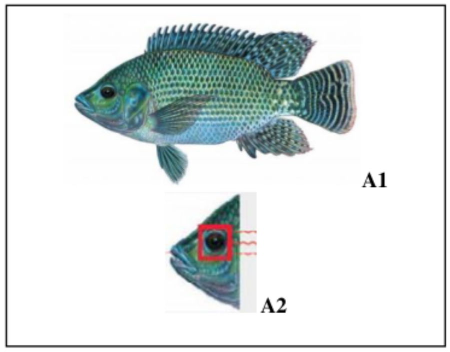


Fig. 3. Focusing eyes on fish

First, the proximity of the new case to A2 (fish eye) for the Red (R) register digital image is calculated if the following values are known: a) The weight of the category attribute proximity (fish eye) = 0, b) Category attribute weight = 0.8, c) Subcategory attribute proximity weight (focus Red) = 0, d) Subcategory attribute weight = 0.7, e) The weight of the proximity of the primary indication attribute (Green) = 0.7, f) The weight of the primary indication attribute = 0.6, g) The weight of the proximity of the secondary indication attribute (Blue) = 0, and h) The weight of the secondary indication attribute = 0.5. Therefore, the R distance can be calculated as follows:

$$Jarak = \frac{(a*b) + (c*d) + (e*f) + (g*h)}{b + d + f + h}$$

$$Jarak = \frac{(0.3*0.8) + (0*0.7) + (0*0.6) + (0*0.5)}{0.8 + 0.7 + 0.6 + 0.5}$$

$$Jarak = \frac{0}{2.6}$$

$$Jarak = 0.00$$

Second, the proximity of the new case to A2 (fish eye) for the digital image of Green (G) register is calculated if the following values are known: a) The weight of the category attribute proximity (fish eye) = 0, b) Category attribute weight = 0.8, c) Subcategory attribute proximity weight (focus Red) = 0, d) Subcategory attribute weight = 0.7, e) The weight of the proximity of the primary indication attribute (Green) = 0.7 f) The weight of the primary indication attribute = 0.6, g) The weight of the proximity of the secondary indication attribute (Blue) = 0, h) The weight of the secondary indication attribute = 0.5. Therefore, the G distance can be calculated as follow:

$$Jarak = \frac{(a*b) + (c*d) + (e*f) + (g*h)}{b + d + f + h}$$

$$Jarak = \frac{(0.3*0.8) + (0*0.7) + (0*0.6) + (0*0.5)}{0.8 + 0.7 + 0.6 + 0.5}$$

$$Jarak = \frac{0}{2.6}$$

$Jarak = 0.00$

Third, the proximity of the new case to A2 (fish eye) for the digital image of Blue (B) register is calculated if the following values are known: a) The weight of the category attribute proximity (fish eye) = 0, b) Category attribute weight = 0.8, c) Subcategory attribute proximity weight (focus Red) = 0, d) Subcategory attribute weight = 0.7, e) The weight of the proximity of the primary indication attribute (Green) = 0.7 f) The weight of the primary indication attribute = 0.6, g) The weight of the proximity of the secondary indication attribute (Blue) = 0, h) The weight of the secondary indication attribute = 0.5. Therefore, the B distance can be calculated as follow:

$$Jarak = \frac{(a * b) + (c * d) + (e * f) + (g * h)}{b + d + f + h}$$

$$Jarak = \frac{(0.3 * 0.8) + (0 * 0.7) + (0 * 0.6) + (0 * 0.5)}{0.8 + 0.7 + 0.6 + 0.5}$$

$$Jarak = \frac{0}{2.6}$$

$$Jarak = 0.00$$

The calculation shows the R, G, B distance values of zero, so the calculation can be known immediately. It means that the outer and inner eye circles are both close so that the value is absolute. And to see the sample tests that have been carried out in tests up to B1-B17, the results can be seen at Table V.

TABLE V. SAMPLE TEST VERIFICATION RESULT

No. Sampel	Hasil Verifikasi	Hasil	Nilai
B1	Fresh	Fresh	1
B2	Fresh	Fresh	1
B3	Fresh	Fresh	1
B4	Not Fresh	Not Fresh	0
B5	Fresh	Fresh	1
B6	Fresh	Fresh	1
B7	Fresh	Fresh	1
B8	Not Fresh	Not Fresh	0
B9	Fresh	Fresh	1
B10	Fresh	Fresh	1
B11	Fresh	Fresh	1
B12	Not Fresh	Not Fresh	0
B13	Fresh	Fresh	1
B15	Fresh	Fresh	1
B16	Not Fresh	Not Fresh	0
B17	Fresh	Fresh	1

Result showed the capability of the proposed system in calculating the Mujair quality based on the eye image. The KNN algorithm was used to process the image received. Fig 4 shows the classification of the proposed system through a mobile device.

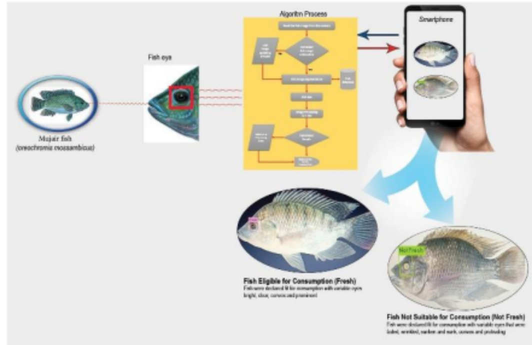


Fig. 4. Output Application Result

The technology used for the implementation is smartphone/mobile-based with the Android Study programming language. Nowadays, the mobile device is the most acceptable for mobile activity. In particular, for the online shop that can help the consumers that do not have the knowledge regarding the quality of Mujair fish. Fig 5 shows the home screen of the propose mobile application.

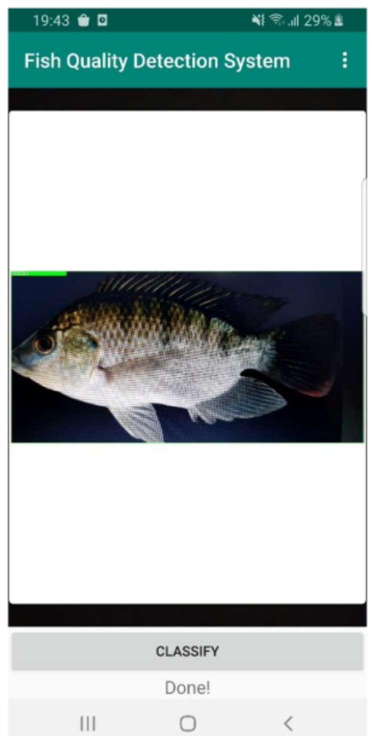


Fig. 5. Result of Design of Fish Detection Application

Testing should be done for the fresh and not fresh Mujair fish. Fig 6 shows when the not fresh Mujair fish was detected.

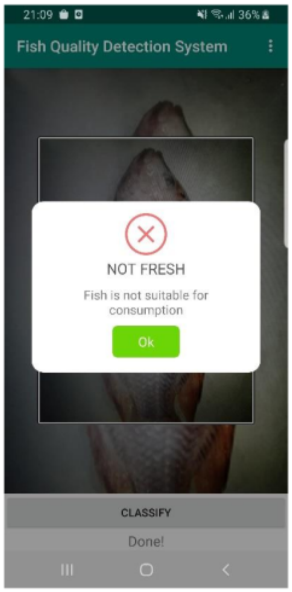


Fig. 6. The Results of the Design of Not Fresh Fish Detection Applications

When the application detected the fresh Mujair fish, it shows the notification. Some error appeared (20%) for ambiguous fish. Additional classes might be needed, instead of just fresh or not fresh fish.

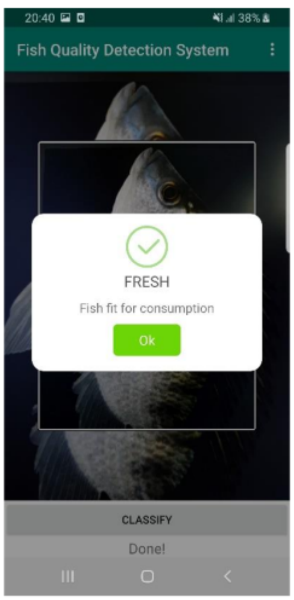


Fig. 7. The Results of the Design Fresh Fish Detection Applications.

A questionnaire was sent to the potential user to check the acceptability of the proposed mobile application. The good accuracy and the user friendly can be seen in Table VI.

TABLE VI. QUESTIONNAIRE AFTER SYSTEM IMPLEMENTATION.

No.	Questions	Percentage
1	Display communicative applications	90,7%
2	Ease of operation of the application	87,5%
3	Application is very helpful in fish detection	80,8%
4	How many times the fish detector has failed	20,%
5	Convenience of using the application as a whole	80,8%

IV. CONCLUSION

Consumers need a guidance in checking the condition of the product, especially the fresh product, for example a fish. Based on the results, the proposed mobile application shows the ability to classify the Mujair fish whether fresh or not fresh based on the eye image detection using the KNN method. The application, of course, can be added for other kind of fish as well as for other fresh product, e.g. vegetables and fruits as well as using other advanced method, e.g. machine learning and deep learning. It also can help the government to keep the consumers health.

REFERENCES

[1] D. Herwanto and A. Suzianti, "Workplace Design Process at Indonesian Manufacturing SMEs," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 847, no. 1, 2020.

[2] I. C. Wu and H. K. Yu, "Sequential analysis and clustering to investigate users' online shopping behaviors based on need-states," *Inf. Process. Manag.*, vol. 57, no. 6, p. 102323, 2020.

[3] Rahmisyari, B. Aditi, T. A. Darojat, K. Kholik, and M. G. Saragih, "Privacy and security analysis of E-commerce on customer re-purchase intention with trust as an intervening variable," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 5 Special Issue, pp. 1578–1589, 2020.

[4] J. S. Park, M. J. Oh, and S. Han, "Fish disease diagnosis system based on image processing of pathogens' microscopic images," *Proc. Front. Converg. Biosci. Inf. Technol. FBIT 2007*, pp. 878–883, 2007.

[5] R. Amat, J. Y. Sari, and I. P. Ningrum, "Implementasi Metode Local Binary Patterns Untuk Pengenalan Pola Huruf Hiragana Dan Katakana Pada Smartphone," *JUTI J. Ilm. Teknol. Inf.*, vol. 15, no. 2, p. 152, 2017.

[6] T. D. Susanto, *Sebuah Kajian Akademik Berdasarkan Dokumen Computing Curricula 2005 Computing sebagai sebuah Rumpun Ilmu*. 2005.

[7] C.~Cortes, C.~Cortes, V.~Vapnik, and V.~Vapnik, "Support Vector Networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273~-~297, 1995.

[8] S. Rizam, F. Yasmin, A. Ihsan, and K. Shazana, "Non-destructive watermelon ripeness determination using image processing and artificial neural network (ANN)," *World Acad. Sci. Eng. Technol.*, vol. 38, pp. 542–546, 2009.

[9] L. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. USA: Prentice-Hall, Inc., 1994.

[10] M. Han, Jiawei; Kamber, *Data Mining, Concept and Techniques*. San Fransisco: Elsevier Inc., 2006.

[11] R. C. Gonzalez and R. E. Woods, *Digital image processing*. Upper Saddle River, N.J.: Prentice Hall, 2008.

[12] R. Shilkrot and D. M. Escrivá, *Mastering OpenCV 4 - Third Edition*. Packt Publishing, 2018.

Mobile-based Fish Quality Detection System Using K-Nearest Neighbors Method

ORIGINALITY REPORT

30%	30%	0%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	repository.unika.ac.id Internet Source	30%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		