K-Nearest Neighbor method for detecting egg quality conditions using Raspberry Pi

M Ahsan¹*, D A Nugraha¹, A Yunus¹, A E Budianto¹, W Setyaningsih¹, A Firniawan¹ and M Susilowati²

¹ Informatics Study Program, Faculty of Science and Technology, Universitas PGRI Kanjeruha Malang, Jl. S. Supriadi, No. 48 Malang Jawa Timur, Indonesia
² Program Studi Sistem Informasi, Fakultas Sain dan Teknologi, Universitas Ma Chung Malang, Jl. Villa Puncak Tidar Blok N No. 1, Dau, Malang, Jawa Timur, Indonesia

*ahsan@unikama.ac.id

Abstract. Food production in Indonesia is very strong and constrained, such as rice, corn, eggs and so on. Eggs that are consumed every day by Indonesian citizens, many consumers of egg food are not alert to the condition or condition of the eggs, whether they are feasible or expired. There are many studies related to the nature of eggs, egg quality and others, in part there is no modeling tool for detecting egg quality conditions based on raspberry pi using the K-Nearest Neighbor method which can be used to see good or bad eggs. In this study, the Raspberry Pi was used as the main component with the aid of a camera as an image capture on eggs using training data of 100 egg samples consisting of 50 good eggs and 50 bad eggs in order to get maximum accuracy. With the help of 2 led as a light source and the software uses the Python language. The purpose of this study was to obtain the expected accuracy by collecting 20 eggs randomly to get a percentage (17/20) * 100 = 85%.

1. Introduction
Food needs are one of the aspects of basic human needs in everyday life. Increasing population growth and increasing development in society have made food a major commodity in society so that egg breeders are one of the crucial sectors in providing food needs and However, the egg breeder sector is currently experiencing major challenges because the climate, weather and humidity, usually affects the condition of the eggs so that it can cause a decrease in egg quality, so food will become expensive [1].

Eggs are a source of animal protein which is needed by the body in order to maintain the body's metabolism, so eggs are a staple food that is often encountered in Indonesia. Because it often happens when consumers use this food production so that the eggs are not in a proper condition because they have expired [2]. So this study aims to make a modeling tool for detecting egg quality conditions based on raspberry pi using the k-nearest neighbor method which can be used to determine good and bad eggs [3]. Indeed, some previous researchers have also made modeling of egg quality, but this is usually intended to determine whether or not they can be hatched. So what will be picked up is the results of research that can be built modeling a tool to detect the quality condition of eggs based on raspberry pi with objects in the outer shell of the new eggs from the parent and the reason why the problem is important because if it has been a long time ago or it does not sell, it will affect the health of consumers,
so the nutrition or protein in eggs has become obsolete [2]. To meet these needs, the community innovates by sorting out which eggs are good and which are bad, but the community has not done it optimally because of the lack of support for tools that can help to sort out good and bad eggs, therefore to deal with this problem in order to get good egg results [4], the maximum. With this automatic tool, people or those who have egg breeders can take advantage of technological aspects, one of which is that this technology is connected one device to another to obtain valid data in developing a study. This technology can also be applied in the livestock sector, one of which is to help with the above problems to monitor the egg yield. From the monitoring, it can be used to develop better egg quality results [5].

2. Methods
The method in research procedure is a series of activities carried out regularly and systematically in research to achieve research objectives. This research procedure is useful for solving problems in research and producing perfect results. In the research that will be carried out in this final project entitled a tool to detect egg quality conditions based on Raspberry Pi using the K-Nearest Neighbor method [6].

Hardware design sketch for detecting egg quality conditions based on Raspberry Pi using the K-Nearest Neighbor method [7]. Raspberry Pi 2 Model B A small computer as a provider for access to other devices and as the main board in these devices [8]. The software components needed to make a detection tool for egg quality conditions based on the Raspberry Pi using the K-Nearest Neighbor method is Raspbian Jessie 02-09-2016: As an Operating System on the Raspberry Pi 2 Model B; Open CV: Used for digital image promotion library; Python 2.7: Used to write programs needed on the Raspberry Pi system [9].

K-Nearest Neighbor is a method for making decisions using supervised learning where the results of new input data are classified based on the closest value to the data. The K-Nearest Neighbor (KNN) algorithm is a method for classifying objects based on learning data that is closest to the object [10]. KNN is a supervised learning algorithm where the results of new query instances are classified based on the majority of categories in the KNN algorithm. Where the class that appears the most will be the class resulting from the classification. Proximity is defined in metric distances, such as Euclidean distances. The Euclidian distance can be found using equation 1.

\[ D_{xy} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \]  

(1)

K-Nearest Neighbor methods include: Determine the parameter K (number of closest neighbors); Calculate the square of the Euclid distance (query instance) of each object against the given sample data using equation 1; Then sort the objects into groups that have the smallest Euclid distance; Collecting the Y category (Nearest Neighbor classification); By using the majority of Nearest Neighbor categories, the calculated value of the query instance can be predicted [11].

3. Results and discussion
The implementation of the dimensional design of the hardware that is designed to be the object of the application of the Raspberry Pi-based egg quality detection tool is as figure 1:

Figure 1. Looks of the whole tool starting from the media holder, an egg holder, camera, led.
Dark images tend to the left and light images to the right. It should be noted that the histogram takes grayscale images instead of color images [12]. This means that the left area of the histogram shows the number of pixels darker for pixel value 0 and the right area shows the number of pixels that are lighter for pixel value 255. For the middle the pixel value is approximately 127 smaller. So the highest value on the histogram will be taken from the test data which will be entered into the KNN method [10]. For example, in the main egg test, RGB extraction values can be taken but in this test data only uses one data. Following are the results of RGB extraction using a histogram of 0.0.0 for the K value of 3 and for the training data 100 egg image samples consisting of 50 good eggs and 50 bad eggs. The example taken from the extraction of egg sample data like is figure 2 and figure 3:

Figure 2. File name is egg_b1.png. Extraction Results 129,0,0,egg_good.

Figure 3. File name is egg_b2.png; Extraction Results : 225,80,16,egg_bad.

Design implementation is the stage of making something that has a specific purpose. At this stage, it is explained about the implementation of the design of hardware, software and operational testing requirements.

Figure 4. Extraction results :25,80,16,egg_Bad.

Figure 5. Extraction results :130,0,0,egg_Good.

Figure 6. Extraction results :99,0,0,egg_Good.

Figure 7. Extraction results :106,12,0,egg_Good.

Figure 8. Extraction results :255,61,35,egg_Bad.

Figures 4,5,6,7 and 8: Egg testing that has been tested in training, there are 20 eggs whose quality conditions are not known before, after training there are 14 good eggs for 3 eggs detected bad and 3 eggs that have been in the training data are inaccurate and are in the picture This shows that the eggs are ugly but in the test the eggs are good, so if the percentage is calculated the result is (17/20) * 100 = 85%.

The data test will be included in the KNN method. For example, in the main egg test taken RGB extraction value using histogram. It is almost the same as the extras of training data but in this test data uses only one data. Here are the results of RGB extras using histogram 0,0,0 The results of this test data are retrieved from egg data which in the next test data classifier using KNN method. Classification using KNN method for K value is 3 and for training data is 100.

4. Conclusion
The conclusions that can be drawn based on the research of results in this study: Raspberry Pi-based egg quality detection tool by using Raspberry Pi as the main component successfully made and applied. The
tool also uses the camera Pi as an egg quality detection to take pictures of good or ugly eggs whether it is worth consuming; Digital image processing research requires the extraction of RGB color with histogram to get data on the training data image used as many as 100 eggs consisting of 50 good eggs and 50 ugly eggs; Testing with KNN method get the inaccuracy of egg detection from the reading of training data can be seen from the highest value of RGB extraction e.g. 129,0,0, egg_good is 129 and 225,80,16, egg_bad has a result is 321 then the higher the RGB value will produce bad eggs depending on the condition of the quality of the eggs to be tested; and Testing with KNN method to obtain the accuracy of egg detection from the reading of training data can be seen from the test there are 20 eggs that were previously unknown quality condition after training there are 14 good eggs and 3 eggs detected bad and 3 eggs that have been recorded training inaccurately. Then if calculated by percentage the result is (17/20) *100=85%.

References
[1] Sutoyo T 2009 *Teori Pengolahan Citra Digital* (Penerbit Andi)
[2] Ibrahim R, Zin Z M, Nadzri N, Shamsudin M Z and Zainudin M Z 2012 Egg’s grade classification and dirt inspection using image processing techniques *Lect. Notes Eng. Comput. Sci.* **2198** 1179–1182
[3] Jura M, Turner J L, Van Dyk S and Knapp G R 2000 *What Is Hatching in the Egg?* *Astrophys. J.* **528** L105–8
[4] Carissa Y H and Soetedjo I K S A 2018 *Deteksi Kualitas Telur, Fertilitas Beserta Ukuran Telur Berbasis Kamera Untuk Penyortiran Telur Sereal Untuk Beras* **51** 51
[5] He J and Wu X 2020 Survey of canthaxanthin content in commercial chicken eggs sampled from Hunan province, China *IOP Conf. Ser. Earth Environ. Sci.* **559**
[6] Chen S 2018 *K-Nearest Neighbor Algorithm Optimization in Text Categorization* *IOP Conf. Ser. Earth Environ. Sci.* **108**
[7] Prayudani S, Hizriadi A, Lase Y Y, Fatmi Y and Al-Khowarizmi 2019 *Analysis Accuracy of Forecasting Measurement Technique on Random K-Nearest Neighbor (RKNN) Using MAPE and MSE* *J. Phys. Conf. Ser.* **1361**
[8] Budiawan I and Andriana A 2015 *Pengujian Pengenalan Wajah Menggunakan Raspberry Pi J. Otomasi Kontrol dan Instrumentasi* **6** 135
[9] Khabibulloh M A, Kusumawardhani I A, Sc M and Pratama D Y 2012 *Rancang Bangun Sistem Deteksi Embrio Pada Telur Menggunakan Webcam* **1** 1–6
[10] Danil M, Efendi S and Widia Sembiring R 2019 *The Analysis of Attribution Reduction of K-Nearest Neighbor (KNN) Algorithm by Using Chi-Square* *J. Phys. Conf. Ser.* **1424**
[11] Mukahar N and Rosdi B A 2017 *Interval valued fuzzy sets k-nearest neighbors classifier for finger vein recognition* *J. Phys. Conf. Ser.* **890**
[12] Mohamed Z, Universiti Y, Azril M and Universiti H 2015 *Jurnal Teknologi Full paper Migration In Double Porosity Soil Column* *J. Teknol.* **23**–29