K-MEANS CLUSTERING FOR EGG EMBRYO'S DETECTION BASED-ON STATISTICAL FEATURE EXTRACTION APPROACH OF CANDLING EGGS IMAGE

Shoffan Saifullah
Department of Informatics, Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia

Abstract
This research discusses the detection of embryonic eggs using the k-means clustering method based on statistical feature extraction. The processes that occur in detection are image acquisition, image enhancement, feature extraction, and identification/detection. The data used consisted of 200 egg image data, consisting of 100 test data and 100 new test data. The acquisition process uses a smartphone camera by capturing candled egg objects. The results of image acquisition become a reference in the process of image enhancement and feature extraction using Statistical Feature Extraction. The statistical feature extraction applied is the Gray Level Co-occurrence Matrix (GLCM) method, which consists of 6 features, namely Energy, Contrast, Entropy, Variance, Correlation, and Homogeneity. The results of feature extraction (6 features) are grouped by the K-means Clustering method. The clustering process uses Euclidean distance calculations to determine the proximity of features. The results of grouping and testing give the best average results with an accuracy of ≈ 74% from several test samples.

INTRODUCTION

Object detection is a process that is used to find out the object's existence. Object detection can be processed with signal processing [1][2] and image processing [3][4]. Image processing has been widely studied in various fields, one of which is in the field of animal husbandry. In this field, the existence of image processing will help in the process of hatching eggs.

Hatching eggs requires embryo detection checks on certain days of the incubator [5]. The process aims to check the initial success of the process of hatching eggs. Conventional and computerized methods have carried out this process. The conventional technique is the use of light illuminating the egg, either by sunlight or the light and saw with eyes. If the visible presence of the embryo, then it will continue to be hatched as for the computerization process using the same initial method, then the process of identifying embryos using computers with various methods.

The methods used are image processing [4, 6, 7] and texture analysis [8].

Detection of embryos in eggs using Near-Infrared Hyperspectral Imaging (NIHsI) [9] and thermal Imaging [10] can detect the presence of embryos based on the temperature emitted by the eggs. Embryo detection based on NIHsI classifies eggs in an incubator machine. Feature extraction in this study uses the Gabor filter method and region of interest. Meanwhile, the classification uses the K-means method. Classification data shows that in the first two days, embryonic development is difficult to detect [9]. Thus, based on the detection process, the best results are around one week in the hatching process. Besides, the detection process uses thermal imaging on embryonic development for 16 days. The detection process uses analysis of cold curves, cold regions, development, oval shape, and egg morphology. The extraction method uses the region of interest. The process

of detecting egg embryos on the 4th day was 89.6% undetectable, and on the 16th day, 96.3% of the embryos were detected.

Feature extraction for egg fertility identification using the GLCM [8,11,12] shows that the image of egg fertility can be analyzed by feature extraction. The identification of egg fertility with GLCM and Backpropagation (BP) based on the results of grayscaling gives a system accuracy of 82.35% [12]. The development of GLCM and BP [8] is an image enhancement process for preprocessing and image segmentation. The comparison with the K-means Clustering method in the process of identifying unsupervised learning. The results of BP and K-means clustering give an accurate comparison of 93% and 74% of each method.

Researchers have researched image processing of egg images, including segmentation, identification, and other analyzes of egg physical. In this research, the researcher explains the use of K-means clustering in the detection of egg embryos based on the Approach Feature Extraction Approach from the acquisition of Candling egg images. The statistical feature extraction used is GLCM. The features of GLCM used are Energy, Contrast, Entropy, Variance, Correlation, and Homogeneity.

Based on the introduction, which contains background and related articles in this section, this article is organized into four sections. Section 2 describes the methods and stages carried out in the embryo detection process, and Section 3 discusses the results regarding the application and experimentation of the process and data. Finally, part 5 is the conclusion of this research.

METHOD
This section explains materials and methods regarding the flow of the egg embryo detection process. Besides, this study will examine the accuracy of the system using the accuracy of the confusion matrix. In this research, the material used is chicken eggs in hatching. Equipment needed is a smartphone camera, flashlight, and darkroom. Besides, image processing-based detection requires a computer. The software used is Matlab Simulink, Delphi program, and the MySQL database. The main stages of the detection process are shown in Figure 1, namely: image acquisition, image processing, and identification (detection).

Image Acquisition Process
The acquisition process is the initial process of creating a digital image [8]. This process is used to determine the required data

and the method of taking digital images [13]. This stage of the process requires objects, tools, and images. Imaging is converting visible images (such as photographs, drawings, paintings, landscapes, sculptures) into digital images [7].
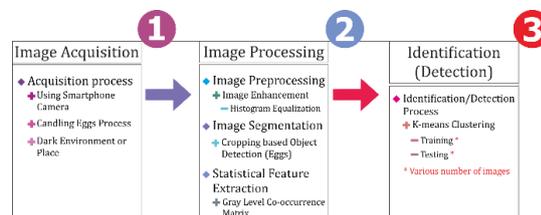

Figure 1. Process of Egg Fertility Detection

The design of image acquisition in this research is shown in Figure 2. The design has three factors, such as s a smartphone camera, an LED lamp for candling, and dark environment conditions [14][15]. The results of image acquisition produce color images (RGB) [4, 7, 8, 11] from eggs. The condition of an egg that has an embryo will be marked by images such as branches or roots.
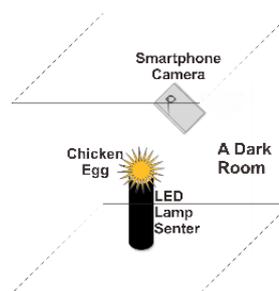

Figure 2. Design of The Image Acquisition [12]

Image Processing
Image processing is processed using image acquisition results. This process aims to process images into useful information. However, to process the image needs to be done on the preprocessing stage. The preprocessing stage functions to prepare images that are ready to be processed, one of which is image enhancement to improve the image quality for further processed.

Image is an image or visual image of an object or several objects [16] such as photographs, images, medical photographs. In principle, the image is divided into three types, namely binary (monochrome) images, grayscale images, and color images. Binary image or black and white image; namely, images have pixel values in the form of numbers 0 or 1 only, or 0 and 255. Binary images function in the process of segmentation (separating objects with their background).

Grayscale imagery uses shades of gray (a combination of black and white). The grayscale image intensity value is between 0 and 255 (256 gray levels) [17] in integers. A grayscale image is more needed for certain purposes. So, the results of the process of color image acquisition will be converted into grayscale images. The conversion process can be done in 2 ways [16], one of which is (1).

$$R' = 0,2989 * R + 0,587 * G + 0,1141 * B \qquad (1)$$

R 'equals G' equals B' is the value of the intensity of the grayscale image. The new R, G, and B color components (R', G', and B') are filled with the same value. The values are obtained by multiplying each color component of R, G, and B (1).

The color image is an image that represents the visual state of objects seen by the eye with the color of the object recorded in the image. The color image is called an RGB image of 3 color components, namely red (R), green (G), and blue (B). The three color components represent each pixel of the image.

An image histogram is a diagram that represents the frequency of occurrence of each intensity value in the image pixel [4][7] . A large value (highest histogram) shows that the intensity value often appears. An image histogram has a function in observing the spread of color intensity (increasing brightness, stretching contrast, or the color distribution of an image) and is used in decision making. Besides, the segmentation process can be used in determining the boundaries of object and background separation.

Image enhancement can be done by modifying the histogram with the histogram equalization method. Histogram equalization is a method used to obtain a histogram that has a uniformly distributed intensity value on the image. A histogram equalization process is performed to obtain a wider gray cedar in areas that have a lot of pixels and narrow the gray cedar in areas that have little pixels. Histogram accumulation for pixels that have a level of k shown in (2).

$$c[k+1] = \sum_{i-1}^{k} hist[k+1] \qquad (2)$$

Furthermore, the gray level is replaced by the value of a with the provision (3).

$$a_k = round\left((L-1)\frac{c[k+1]}{N}\right) \qquad (3)$$

Where c [k + 1] is several pixels that have the same value as k+1, k is the initial gray level value, with a value of k = 0,1,2, ..., L-1 (L is the maximum gray level value). hist [k + 1] is initial histogram. a_k is the new pixel value and N is the number of pixels in the image.

Image Segmentation

The indicator of the right image is a result of segmentation. Segmentation is a stage that aims to partition/divide the image [18][19] into several main parts that contain important information. Recognition of objects in the image requires parameters that can characterize the object [6]. Characteristics that can be used as distinguishing objects from one another are shape, size, geometry, texture, and color.

In this research, the segmentation process will make a cropping image of the egg image by the egg object. Then, it will identify the object and background to distinguish and select the object that the identification process will use. So that this segmentation will choose a single egg image for further detection using feature extraction with GLCM.

Statistical Feature Extraction

The statistical feature extraction method used is GLCM, namely second-order statistics. The intensity co-occurrence matrix is a matrix in which the occurrence of a pair of two pixels at a certain intensity, distance, and direction in the image [20].

GLCM matrix is a matrix that describes the relationship between neighbors in the image in various directions (θ) and distance (d). The GLCM matrix of an image $f(x, y)$ is a two-dimensional matrix in which each matrix element represents the probability of the intensity levels of x and y at certain distances (d) and angles (θ). In making the GLCM matrix, there are four angular directions used, namely 0°, 45°, 90°, and 135°.

Texture analysis on the co-occurrence matrix is calculated using multiple feature calculations [8, 12, 21], such as in (4)-(13).

$$En = \sum_{i_1} \sum_{i_2} p^2(i_1, i_2) \qquad (4)$$

$$Ct = \sum_{i_1} \sum_{i_2} (i_1 - i_2)^2 \, p(i_1, i_2) \qquad (5)$$

$$Et = -\sum_{i_1} \sum_{i_2} p(i_1, i_2) \, log \, p(i_1, i_2) \qquad (6)$$

$$V = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - \mu)^2 * p(i, j) \qquad (7)$$

$$Cr = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j) * \frac{(i - \mu_x)(i - \mu_y)}{\sigma_x \sigma_y} \qquad (8)$$

$$\mu_x = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} i * p(i,j) \qquad (9)$$

$$\mu_y = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} j * p(i,j) \qquad (10)$$

$$\sigma_x = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i-\mu)^2 * p(i,j) \qquad (11)$$

$$\sigma_y = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (j-\mu)^2 * p(i,j) \qquad (12)$$

$$H = \sum_{i_1} \sum_{i_2} \frac{p(i_1,i_2)}{1+|i_1-i_2|} \qquad (13)$$

Based on these formulas, information from (4)-(13) in general can be shown in Table 1.

Table 1. Description of the formulas (4-13)

Variable	Description
En (Energy)	measures the uniformity of pixels in an image
Ct (Contrast)	measures the strength of the difference in intensity in the image
Et (Entropy)	measures the randomness of the intensity distribution
V (Variance)	measures pixel values distribution (based on a combination of the mean of the reference and neighboring pixels)
Cr (Correlation)	describes the linear dependence of grayscale (based on a matrix of reference pixels and neighboring pixels)
H (Homogeneity)	measures the homogeneity of variations in intensity in the image
$p(i,j)$	elements of a normalized GLCM matrix or probability distribution of pixel pairs
(i,j)	Gray level, i_1 = 0, 1, 2, 3, ..., m and i_2 = 0, 1, 2, 3, ..., n
N_g	many gray levels
μ_x	the average value of the column element in the matrix $p(i,j)$
μ_y	the average value of row elements in matrix $p(i,j)$
σ_x	the standard value of column element deviation in matrix $p(i,j)$
σ_y	the standard value of line element deviation from matrix $p(i,j)$

Identification (Detection)

Identification or detection is the final step to find out the object based on its characteristics [22]. This process can be done by conducting training and testing [23][24]. This process requires parameter values of object characteristics to be included in groups/classes. In the training process, training data that contains characteristic parameters are used to distinguish between objects. The training process will map training data towards training targets by the specific algorithm used.

In this research, the algorithm used is K-means Clustering based on the physical characteristics of the object. Besides, this research will conduct a testing process. In the testing process, a formula based on the training process is used to map test data. Thus, the testing process will produce output data that will be compared with the test target to get the level of accuracy of the testing process.

K-means Clustering

K-means clustering is a block model method [25], which is used for grouping data. Raw data (without labels) can be processed and accepted for the clustering process. This method is different from supervised learning [26][27], which receives input in the form of vectors (x_1, y_1), (x_2, y_2), …, (x_i, y_i), where x_i is data of the training data and y_i is the class label for xi. This method is very popular, fast, and simple [28].

This learning process uses a computer to classify data randomly. This process is usually known as the concept of unsupervised learning [29][30]. The input is data or objects and clusters. This algorithm works to group data or objects into certain groups, as in Figure 3, where each cluster has a cluster center [31].
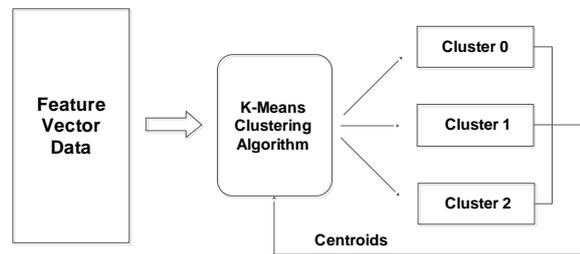

Figure 3. The basic flow of the k-means clustering algorithm with 3 clusters

The K-means clustering algorithm [30][31] is performed using five main steps, and it is repeated according to the conditions of this algorithm. In detail, the steps are as follows:
1. Determine the number of clusters K according to conditions
2. Determine the point K centroid from the cluster randomly
3. The process of grouping data and producing cluster K, which has a centroid point for each cluster based on the previous centroid point
4. The process of updating the centroid point value
5. Repeat the process in steps 2 and 3 until there is no change in the centroid point value.

Data grouping in one cluster is calculated with the closest distance from the centroid point. This calculation uses the Euclidean distance

method between two data. The distance calculation formula [34] is shown in (14) and (15).

$$d(i,j) = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2} \qquad (14)$$

$d(i,j)$ is the object distance from i to j, p is the data dimension. x_{ik} is the object coordinates-i on the dimension k. And x_{ik} is object coordinates-j on the dimension k.

$$\mu_k = \frac{1}{N} \sum_{q=1}^{N_k} x_q \qquad (15)$$

Based n the K-cluster, μ_k is the centroid point. N_k is the amount of data. And, x_q is q data.

RESULTS AND DISCUSSION

This section explains the processes and identification/detection of egg embryos. The process that occurs, starting from image acquisition, image processing, and embryo detection based on statistical characteristics (according to Figure 1). The image acquisition process produces an RGB color image as in studies that have been done. The results of the acquisition are shown in Figure 4, as researchers have done in previous studies [12].
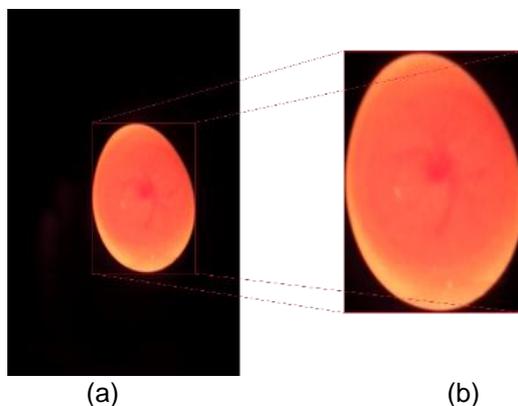

(a)   (b)

Figure 4. Image of acquisition results (a) and cropping results (b) based on the egg object

The image acquisition result of Figure 4 (a) will be carried out the cropping process using image segmentation so that the results are obtained as in Figure 4 (b). This process serves to eliminate the blank background in the image so that the image will provide maximum processing. This process is like what has been done by researchers in previous studies [35]. In addition, the resize process of image size is carried out to speed up the detection process.

The results of the image in Figure 4 (b) are preprocessing to improve image quality. Image enhancement is processed after the input image is the grayscaling process using (1). The image using image enhancement is shown in Figure 5. The image shows the detail of the image processed by its feature extraction.
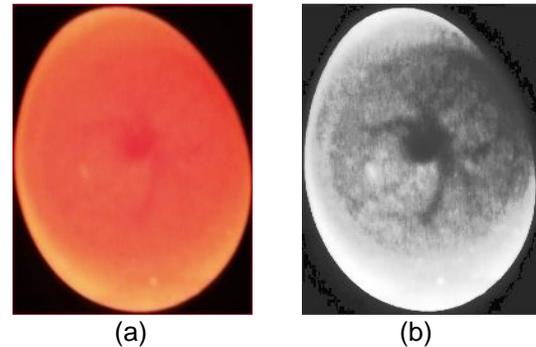

(a)   (b)

Figure 5. Image of acquisition results that have been carried out by cropping (a) and grayscaling results that have been processed with image enhancement (b) of eggs having egg embryos

Based on Figure 5, the image used for feature extraction is the grayscale image (Figure 5 (b)). The grayscale image will be an image of n x m size and needs to be converted to a matrix containing numeric values with each pixel valued between 0-255. For example, in Figure 5 (b), when the original size is 801 x 657, then resizing to 8 x 6, the image matrix that can be displayed is as in Table 1.

Table 1. Example of a grayscale image matrix whose size is changed to 8 x 6

0	47	116	89	15	0
21	122	131	134	95	6
63	124	123	125	132	55
87	126	126	127	136	109
100	138	136	137	144	140
84	158	148	152	170	142
25	151	183	188	198	72
1	36	122	147	79	1

Based on Table 1, the GLCM matrix is used to calculate GLCM features. In this research, GLCM calculations use six features. Thus, one image will have all six of those features. These features are used as the material that has been delivered to the Statistical Feature Extraction: energy (En), contrast (Ct), entropy (Et), variance (V), the correlation (Cr), and homogeneity (H). Based on one of the processed images, the results of the calculation of the GLCM features are obtained as in Table 2.

Table 2 is the result of the GLCM calculation of eggs categorized in embryonated eggs. In this calculation, GLCM has six features with each degree. Table 2 is an example of a normalized calculation where the value ranges from 0 to 1. Normalization results make it easier to calculate for the detection process, both in the training process and testing.

The calculation of each feature with their respective degrees (0, 45, 90, and 135) by the calculation, each feature is calculated on average. The final results used in the training and testing process are with data such as the GLCM sample values from ten data. The ten data consists of fertile and infertile eggs with five samples in each, as shown in Table 3.

The results of the GLCM sample calculation in Table 1, the pattern analysis process made from the six features, can be seen in Figure 6. Embryo and Figure 7, non-embryo eggs have distinct line patterns. Thus, different patterns can be used for the detection process, both in the training and testing processes.

Table 2. The results of calculating the value of GLCM features with each normalized degree

Degree	En	Ct	Et	V	Cr	H
0	0.055850	0.016277	0.049044	0.052592	0.037419	0.059007
45	0.044491	0.042648	0.054171	0.052592	0.037362	0.047268
90	0.048179	0.028561	0.052723	0.052592	0.037389	0.050687
135	0.044620	0.039862	0054157	0.052592	0.037378	0.047324

Table 3. Sample of process statistical feature extraction using GLCM method's

No	Eggs Image		Statistical Feature Extraction with GLCM											
	Fertile	Unfertile	Fertile						Unfertile					
			En	Ct	Et	V	Cr	H	En	Ct	Et	V	Cr	H
1			0.0482	0.0318	0.0525	0.0525	0.0373	0.0510	0.0407	0.0475	0.0501	0.0339	0.0579	0.0493
2			0.0442	0.0328	0.0530	0.0666	0.0295	0.0517	0.0571	0.0276	0.0469	0.0286	0.0687	0.0500
3			0.0478	0.0398	0.0519	0.0568	0.0345	0.0509	0.0865	0.0415	0.0406	0.0144	0.1352	0.0514
4			0.0491	0.0338	0.0519	0.0703	0.0279	0.0517	0.0732	0.0627	0.0426	0.0228	0.0856	0.0509
5			0.0513	0.0470	0.0519	0.0700	0.0280	0.0516	0.0390	0.0242	0.0496	0.0196	0.0999	0.0496

Table 4. Results of the training and testing process of embryo detection images using the k-means clustering method

Epoch	Training Data		Testing Data		Succeed		Failed		Succeed (%)
	Fertile	Unfertile	Fertile	Unfertile	Fertile	Unfertile	Fertile	Unfertile	
1	10	10	90	90	60	53	30	37	62,78
2	20	20	80	80	52	48	28	32	62,50
3	30	30	70	70	44	46	26	24	64,29
4	40	40	60	60	45	42	15	18	72,50
5	50	50	50	50	38	36	12	14	74,00
6	50	50	60	60	47	42	13	18	74,17
7	50	50	70	70	54	49	16	21	73,57
8	50	50	80	80	61	58	19	22	74,38
9	50	50	90	90	69	64	21	26	73,89
10	50	50	100	100	76	73	24	27	74,50

The process of detecting egg embryos is carried out using the k-means clustering method. Following the method, the beginning of this process will group eggs into two groups randomly based on the pattern of the GLCM and the calculation of the distance (Euclidean distance). The testing process is carried out with several variations of testing and training data. The results are seen in Table 4.

Table 4 shows that the testing process with low success, which at epoch 1 to 5 is only able to give success around 63%. Whereas on the 4th epoch, the test results increased to 72.5%. Subsequent epochs to 10 showed fluctuating success with results around of 74%.
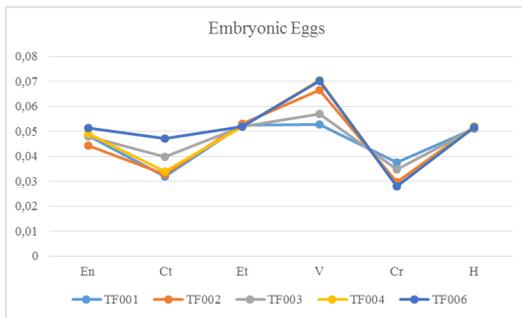
Figure 6. Analysis of the pattern of the value of the features produced to distinguish the characteristics of features embryo eggs
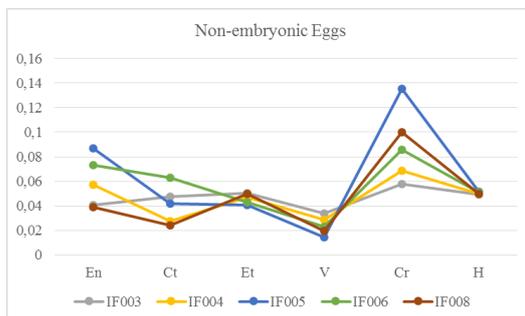

Figure 7. Analysis of the pattern of the value of the features produced to distinguish the characteristics of features non-embryos

CONCLUSION

The result of egg fertility detection is only able to detect egg fertility with an accuracy of around 74%. This detection process is carried out using images of any size. The condition happens because there is an image resizing process. In this study, learning needs to be done in the training process so that it can be applied to Supervised learning in combination with k-means.

ACKNOWLEDGMENT

REFERENCES

[1] S. Sunardi, A. Mahfurdz, and S. Saifullah, "Green turtle and fish identification based on acoustic target strength," *Int. J. Adv. Intell. Informatics*, vol. 4, no. 1, p. 53, Mar. 2018, DOI: 10.26555/ijain.v4i1.147.

[2] S. Jazayeri et al., "Automatic object detection using dynamic time warping on ground penetrating radar signals," *Expert Syst. Appl.*, vol. 122, pp. 102–107, May 2019, DOI: 10.1016/ j.eswa.2018.12.057.

[3] M. P. Bose and S. K. Bandyopadhyay, "Object Detection, Extraction and Classification using Image Processing Technique," *J. Res.*, vol. 03, no. 08, pp. 9–19, 2017.

[4] A. Yudhana, Sunardi, and S. Saifullah, "Segmentation comparing eggs watermarking image and original image," *Bull. Electr. Eng. Informatics*, vol. 6, no. 1, pp. 47–53, 2017, DOI: 10.11591/eei.v6i1. 595.

[5] L. K. S. Tolentino et al., "Development of Fertile Egg Detection and Incubation System Using Image Processing and Automatic Candling," *in IEEE TENCON 2018*, Oct. 2018, pp. 0701–0706, DOI: 10.1109/ TENCON.2018. 8650320.

[6] L. Huang et al., "A Multi-Feature Fusion Based on Transfer Learning for Chicken Embryo Eggs Classification," *Symmetry (Basel).*, vol. 11, no. 5, p. 606, May 2019, DOI: 10.3390/sym11050606.

[7] Sunardi, A. Yudhana, and S. Saifullah, "Identity analysis of egg based on digital and thermal imaging: Image processing and counting object concept," *Int. J. Electr. Comput. Eng.*, vol. 7, no. 1, pp. 200–208, 2017, DOI: 10.11591/ijece.v7i1.12718.

[8] S. Saifullah and V. A. Permadi, "Comparison of Egg Fertility Identification based on GLCM Feature Extraction using Backpropagation and K-means Clustering Algorithms," in *2019 5th International Conference on Science in Information Technology (ICSITech)*, 2019, pp. 140–145.

[9] L. Liu and M. O. Ngadi, "Detecting Fertility and Early Embryo Development of Chicken Eggs Using Near-Infrared Hyperspectral Imaging," *Food Bioprocess Technol*, vol. 6, pp. 2503–2513, 2013.

[10] L. Hai-ling et al., "Research on the Discrimination of Hatching Eggs Activity Based on Thermal Imaging: A Food Nondestructive Testing Practice," *Int. J. Smart Home*, vol. 10, no. 2, pp. 175–186, 2016.

[11] N. A. Fadchar and J. C. Dela Cruz, "Prediction Model for Chicken Egg Fertility Using Artificial Neural Network," *in 2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA)*, Apr. 2020, pp. 916–920, DOI: 10.1109/ICIEA49774.2020.9101966.

[12] Sunardi, A. Yudhana, and S. Saifullah, "Identification of Egg Fertility Using Gray Level Co-Occurrence Matrix and

Backpropagation," *Adv. Sci. Lett.*, vol. 24, no. 12, pp. 9151–9156, 2018, DOI: 10.1166/asl.2018.12115.

[13] V. G. Narushin et al., "Digital imaging assisted geometry of chicken eggs using Hügelschäffer's model," *Biosyst. Eng.*, vol. 197, pp. 45–55, Sep. 2020, DOI: 10.1016/j.biosystemseng.2020.06.008.

[14] W. Fang, T. Zuojun, C. Fei, and W. Youxian, "Detecting Cracks of Preserved Egg Based on Polarized Light," *in 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics*, Aug. 2013, pp. 66–69, DOI: 10.1109/IHMSC.2013.163.

[15] C. Pan et al., "Light Optimization for a LED-based Candling System and Detection Combined with Egg Parameters for Discrimination of Egg Fertility," *Trans. ASABE*, p. 0, 2020.

[16] S. Das et al., "Defect detection of jute fabric using image processing," *Econ. Aff.*, vol. 61, no. 2, p. 273, 2016, DOI: 10.5958/0976-4666.2016.00035.8.

[17] A. B. W. Putra, S. Supriadi, A. P. Wibawa, A. Pranolo, and A. F. O. Gaffar, "Modification of a gray-level dynamic range based on a number of binary bit representation for image compression," *Sci. Inf. Technol. Lett.*, vol. 1, no. 1, pp. 9–16, Apr. 2020, DOI: 10.31763/sitech.v1i1.17.

[18] A. Upadhyay and R. Kashyap, "Fast Segmentation Methods for Medical Images," *Int. J. Comput. Appl.*, vol. 156, no. 3, pp. 18–23, Dec. 2016, DOI: 10.5120/ijca2016912399.

[19] G. V. Sivanarayana et al., "Review on the Methodologies for Image Segmentation Based on CNN," *In Communication Software and Networks* 2021, pp. 165–175, DOI: 10.1007/978-981-15-5397-4_18

[20] D. Gadkari, *Image Quality Analysis Using GLCM*. Orlando, Florida, 2004.

[21] X. Zhang et al., "A Study for Texture Feature Extraction of High-Resolution Satellite Images Based on a Direction Measure and Gray Level Co-Occurrence Matrix Fusion Algorithm," Sensors, vol. 17, no. 7, p. 1474, Jun. 2017, DOI: 10.3390/s17071474.

[22] K. Ray et al., "Detection and Identification of Parasite Eggs from Microscopic Images of Fecal Samples," 2020, pp. 45–55.

[23] J.-S. Fang *et al.,* "Path-dependent human identification using a pyroelectric infrared sensor and Fresnel lens arrays," *Opt. Express*, vol. 14, no. 2, pp. 609–624, 2006.

[24] H. Harb and L. Chen, "Gender identification using a general audio classifier," in *2003 International Conference on Multimedia and Expo. ICME'03. Proceedings (Cat. No. 03TH8698)*, 2003, vol. 2, pp. II--733.

[25] A. Žiberna, "K-means-based algorithm for blockmodeling linked networks," *Soc. Networks*, vol. 61, pp. 153–169, May 2020, DOI: 10.1016/j.socnet.2019.10.006.

[26] A. Y. Saleh and E. Liansitim, "Palm oil classification using deep learning," *Sci. Inf. Technol. Lett.*, vol. 1, no. 1, pp. 1–8, Apr. 2020, DOI: 10.31763/sitech.v1i1.1.

[27] A. A. Zarir, S. Bashar, and A. R. Ismail, "Automated image captioning with deep neural networks," *Sci. Inf. Technol. Lett.*, vol. 1, no. 1, pp. 17–23, Apr. 2020, DOI: 10.31763/sitech.v1i1.31.

[28] R. Garcia-Dias, S. Vieira, W. H. Lopez Pinaya, and A. Mechelli, "Clustering analysis," in *Machine Learning*, Elsevier, 2020, pp. 227–247.

[29] M. Caron, P. Bojanowski, J. Mairal, and A. Joulin, "Unsupervised pre-training of image features on non-curated data," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 2959–2968.

[30] B. Nguyen and B. De Baets, "Kernel-Based Distance Metric Learning for Supervised k - Means Clustering," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 10, pp. 3084–3095, Oct. 2019, DOI: 10.1109/TNNLS.2018.2890021.

[31] M. T. Jones, *Artificial Intelligent: A System Approach*. Hingham: David Pallai, 2008.

[32] S. F. Hussain and M. Haris, "A k-means based co-clustering (kCC) algorithm for sparse, high dimensional data," *Expert Syst. Appl.*, vol. 118, pp. 20–34, Mar. 2019, DOI: 10.1016/j.eswa.2018.09.006.

[33] J. S. Low et al., "Seeding on Samples for Accelerating K-Means Clustering," in *Proceedings of the 3rd International Conference on Big Data and Internet of Things - BDIOT 2019*, 2019, pp. 41–45, DOI: 10.1145/3361758.3361774.

[34] R. Loohach and K. Garg, "Effect of distance functions on k-means clustering algorithm," *Int. J. Comput. Appl.*, vol. 49, no. 6, pp. 7–9, 2012.

[35] S. Saifullah, Sunardi, and A. Yudhana, "Analisis Perbandingan Pengolahan Citra Asli dan Hasil Croping Untuk Identifikasi Telur," *J. Tek. Inform. dan Sist. Inf.*, vol. 2, no. 3, pp. 341–350, 2016.