Chicken Egg Fertility Identification using FOS and BP-Neural Networks on Image Processing

Shoffan Saifullah¹, Andiko Putro Suryotomo²
¹,²Department of Informatics, Universitas Pembangunan Nasional Veteran Yogyakarta
¹shoffans@upnyk.ac.id, ²andiko.ps@upnyk.ac.id

Abstract

This article aims to test FOS (first-order statistical) in extracting features of embryonated eggs. This test uses the initial step of image processing to get the best input image in feature extraction. The image processing method starts from the image acquisition process, then improves with image preprocessing and segmentation. Image acquisition in this study uses the concept of egg candling in a dark place captured with a smartphone camera. The acquisition results are improved by image preprocessing using gray scaling, image enhancement (by Histogram Equalization), and segmentation of chicken egg image. The segmentation results were extracted using FOS with five parameters: mean, entropy, variance, skewness, and kurtosis. Based on the calculation of these parameters, it is graphed and shows the difference in patterns between fertile and infertile eggs. However, some eggs have a similar pattern, thus affecting the identification process. The identification process used neural networks by the backpropagation method for training and testing. The training results provide an accuracy value of 100% of all training data; however, 80% of the new test data obtained test results at testing. This test is carried out with 100 data, 50 each for training and test data. Based on the test results, which significantly affect the level of accuracy is the feature extraction method. FOS pattern in detecting the fertility of chicken eggs by BP Neural Network is still categorized as low, so it is necessary to improve methods to get maximum results.

Keywords: Backpropagation, Feature Extraction, First Order Statistical, Classification, Image Processing

1. Introduction

Identifying chicken eggs fertility in hatching chicken eggs is exciting to increase chicken production results. The identification process is carried out using the observation method using light directed at the egg so that the contents of the egg can be seen (candling) [1]–[3]. This identification is still being made manually (done by humans) [1], [4], or there are already several tools from various studies [5], [6], both simulators and implementations [7]–[9]. Even though the tools already exist, there are still many researchers doing development to get better results. This condition is growing along with the development of technology. Technology in identifying the fertility of chicken eggs has been developed in various ways, one of which is the application of image processing to identify the fertility of chicken egg images, either using image segmentation [8], [9] or classification or identification [10].

Image processing in the identification of egg embryos begins with image acquisition which has different methods and tools [4], [11], [12]. Several acquisition processes were carried out using thermal cameras [6], [13], hyperspectral [6], [7], [10], and digital cameras [4], [14]. This study uses a smartphone camera and a candling process to obtain a digital image. Using this tool is easy to obtain, even most people have it, and it is user-friendly by various groups of people.

The acquired digital image was done by image preprocessing and segmentation to get the input image in feature extraction. Image preprocessing in identifying embryos/fertility eggs is carried out in various ways, such as segmentation [15], feature extraction [4], [16], and classification. Segmentation in this detection has been done by methods such as k-means, thresholding. In feature extraction, this detection uses GLCM with various features, including five or seven features. The results of feature extraction are inputted into fertility classification (embryo presence) using the K-means method [7], [9], [17]. Backpropagation, SVM [18], and Naïve Bayes. In this study, the segmentation used is when cropping the image according to the egg object in the image preprocessing. In addition, image preprocessing uses resizing and image enhancement.
The resizing method aims to reduce the image and speed up processing. The image enhancement uses the concept of histogram equalization and its derivatives (i.e., CLAHE and HE [19], [20]).

In identification, some apply object feature extraction to obtain different patterns that indicate the characteristics of a particular object. Feature extraction in this study uses first-order feature extraction with five parameters [21], namely Mean, Entropy, Variance, Skewness, and Kurtosis. This feature is able to classify omega 3 and leghorn eggs with an accuracy of 91% [22]. In previous studies [4], [12], [17], it was stated that existing features needed to be reduced. Thus, this study reduced GLCM to FOS (using five parameters). In addition, the identification uses an artificial neural network using the Backpropagation (BP) method. This BP uses forward and backward in its training and testing. BP was used in this research because this method can get the best result in our previous research (93% accurate). This study uses BP as a development of our previous research with supervised and unsupervised methods. Although SVM is better than BP (92.5% vs. 87.5%) [23], from our previous study [12], the BP method has a success percentage equal to the presence of preprocessing in image processing (93%). While Naïve Bayes’s highest accuracy is 94%, the lowest is 78% with fluctuating accuracy data [24].

This article discusses feature extraction with first-order statistics as input for the identification process. The first section describes the background of the problem, research review (gaps/differences of others), and research objectives. Next, the second section describes proposed methods and theories related to image processing, feature extraction, and classification. The third section discusses the experimental results. Finally, the fourth section concludes the results obtained from this study.

2. Research Method

This study used the image processing method in the classification process. The primary image processing used is image acquisition, image preprocessing, and image segmentation [25]. The development used feature extraction and calculation [26], [27] using the backpropagation neural networks method. This research process in detail used the following stages in Figure 1.

This study uses the concept of image processing and its development, namely feature extraction and classification. The first step in image processing is to take a picture first (image acquisition), and the acquisition results become input for the next steps on image processing. The three main stages in this research are image processing, feature extraction, and classification.

2.1. Image Processing

Image processing is an image process in which input and output are images with the initial image acquisition process. Image acquisition is used to obtain digital images. In this study, the acquisition process used a smartphone camera to take pictures of chicken eggs irradiated with LED lights (candling). The description of this acquisition process is shown in Figure 2. (a).

The smartphone camera takes pictures without using a flash, so the digital image obtained emits light from the LED lamp (from the candling process), as shown in Figure 2. (b). The concept of acquisition in this study is to use reflections from real objects, namely chicken eggs. The resulting image is with several color components, namely Red, Green, Blue (RGB).

Image preprocessing aims to improve digital images into better images as input for the following process. Image preprocessing used in this study includes cropping, grayscaling, and image enhancement. Cropping is image preprocessing to get the egg object image for detailed processing. The cropping process will get the image according to the object. In this study, the process detects images from 4 sides, namely, top, left, bottom, and right, to detect objects and select different points that are not close to black (this process used the segmentation method). The cropping result is a color image (RGB) that needs to be converted into a grayscale image (greyscaling process). Grayscale can be processed....
using conventional formulas to change each color component (R: Red, G: Green, and B: Blue) using (1) [28].

\[
Gray = 0.2989xR + 0.5877xG + 0.1141xB \tag{1}
\]

Image enhancement is processed using grayscale image input. In this study, image enhancement used the histogram equalization (HE) method. HE is a method used to obtain an image by histogram alignment. In addition, we used the development of HE (Contrast Limited Adaptive Histogram Equalization/CLAHE). [29]. Image enhancement studies using the CLAHE-HE hybrid [19], [20], [30].

2.2. First Order Statistics (FOS) Feature Extraction

FOS is a texture feature extraction method [31] using an image histogram for the calculation process [32]. This first-order feature extraction [33] uses five main parameters, namely Mean, Entropy, Variance, Skew, and Kurtosis [34], [35]. These features can classify the omega-3 and leghorn eggs more than 90% accurately [22]. Thus, this research used the reference to get features of eggs fertility on classification.

\[
\mu = \sum_{n=0}^{N} f_n p(f_n) \tag{2}
\]

Mean (\(\mu\)) is a measurement that shows the size of the image dispersion. This calculation uses calculations based on histogram and grayscale values (2).

\[
H = -\sum_{n=0}^{N} P(f_n) \log P(f_n) \tag{3}
\]

Entropy (H) is a measurement used to determine the irregularity of the shape of an image. This measurement uses the calculation of the histogram value and the calculation of its logarithm (3).

\[
\sigma^2 = \sum_{n=0}^{N} (f_n - \mu)^2 P(f_n) \tag{4}
\]

Variance (\(\sigma^2\)) shows the variation of elements in the histogram of an image. The variance is also called the standard deviation, calculated based on the grayscale, histogram, and mean (4) values.

\[
a_3 = \frac{\sum_{n=0}^{N} (f_n - \mu)^3 P(f_n)}{\sigma^3} \tag{5}
\]

Skewness (\(a_3\)) calculates the relative slope or slope of the histogram curve of an image. This calculation is calculated using components (variance and the cube of variance) and the division of the standard deviation (5).

\[
a_4 = \frac{\sum_{n=0}^{N} (f_n - \mu)^4 P(f_n)}{\sigma^4} - 3 \tag{6}
\]

Kurtosis (\(a_4\)) is a calculation of the relative sharpness of the histogram curve of an image using components such as the standard deviation to the power of 4 and the standard deviation divisor (6).

2.3. Classification

Classification aims to identify an object based on specific characteristics. In this research, the classification used the backpropagation neural network (BP-NN) method. An artificial neural network represents an artificial brain in a computer program that can complete some calculation processes during the learning process [36]. An artificial neural network has several vital components (such as the human brain), namely a neural network consisting of neurons and connections between neurons. These neurons transform the information received from one neuron to another through their outgoing connections. In comparison, the artificial neural network assigns weights to the process, which stores certain valuable information. Figure 3 shows the structure of neurons in an artificial neural network.

![Neuron structure activation process in neural networks](image)

Figure 3. Neuron structure activation process in neural networks [36]

The workings of artificial neurons are similar to biological neurons [36], namely:

1. Information (input) is sent to neurons with a certain arrival weight.
2. The input is processed by a propagation function which adds up the values of all the incoming weights.
3. This sum is compared with a specific threshold value through the activation function of each neuron.
4. If the input passes a specific threshold value, then the neuron is not activated.
5. If a neuron is activated, it sends output through its output weights to all the associated neurons.

BP-NN is a supervised learning algorithm usually used by multi-layered perceptrons to change the weights associated with the neurons in the hidden layer. The backpropagation algorithm uses error output to change the value of the weights in the backward direction. The error value is obtained by doing forward propagation [37], [38] first by activating the neurons using the sigmoid activation function. The BP-NN architecture is shown in Figure 4.

A multi-layer neural network with one hidden layer (unit Z) is shown in Figure 4. The output unit (Y) and the hidden unit have a bias (1). Wok denotes (bias) in the output unit Yk, and Voj denotes the bias in the hidden unit Zj. This bias term acts as a weight on the connection of a unit whose output is always 1.

The backpropagation method’s training process involves three main stages: feedforward, backpropagation, and weight adjustment. During the feedforward process, each input unit (Xi) receives the input signal and forwards the signal to the hidden unit (Z1, ..., Zj). Each hidden unit counts its activation and sends the signal (\(zj\)) to each output unit. Each output unit (Yk) counts activations (\(yk\)) to form the network response to a given input pattern.
During training, each output unit compares the calculated activation \( y_k \) with the target value \( t_k \) to determine the associated error in the pattern with that unit. Based on this error, the factor \( \delta_{k}^j \) is calculated. \( \delta_{k}^j \) is used to distribute the error at the output unit \( Y_k \) back to all units in the previous layer (hidden units connected to \( Y_k \)). It is also used (later) to update the weights between the output and hidden layers. The same way is done to calculate the factor \( \delta_{j}^i \) on each hidden unit \( Z_j \). It is unnecessary to propagate the error back to the input layer, but \( \delta_{j}^i \) is used to update the weights between the hidden and the input layers.

![BP-NN architecture with one hidden layer](image)

During training, each output unit compares the calculated activation \( y_k \) with the target value \( t_k \) to determine the associated error in the pattern with that unit. Based on this error, the factor \( \delta_{k}^{(m)} \) is calculated. \( \delta_{k}^{(m)} \) is used to distribute the error at the output unit \( Y_k \) back to all units in the previous layer (hidden units connected to \( Y_k \)). It is also used (later) to update the weights between the output and hidden layers. The same way is done to calculate the factor \( \delta_{j}^{(p)} \) on each hidden unit \( Z_j \). It is unnecessary to propagate the error back to the input layer, but \( \delta_{j}^{(p)} \) is used to update the weights between the hidden and the input layers.

After the factors have been determined, the weights on all layers are adjusted simultaneously. The \( w_{jk} \) weight adjustment (from the hidden \( Z_i \) unit to the \( Y_k \) output unit) is based on the \( \delta_{k}^{(m)} \) factor and the \( z_j \) activation of the \( Z_j \) hidden unit. The adjustment to the weight of \( v_{ij} \) (from the input unit \( X_i \) to the hidden unit \( Z_j \)) is based on the \( \delta_{k}^{(m)} \) factor and \( x_i \) activation from the input unit. The flowchart of the BP Algorithm is shown in Figure 5.

### 3. Results and Discussion

This research produced several outputs and accurate calculations. The outputs produced include image processing, feature extraction using first-order statistical methods. In feature extraction, analysis is carried out based on each object type (fertile and infertile eggs). In addition to focusing on these two things, the experiment in this study is to identify them. The identification method used is BP-NN.

#### 3.1. Image processing on the first step for FOS input

Image processing obtained in this study was processed according to Figure 1. The initial process is image acquisition to get a digital image of chicken eggs. This image has a single egg object with a black background, as shown in Figure 6. (a), so it is necessary to do a cropping process using the concept of segmentation to detect eggs. The cropping result is an image of an egg with a minimal background according to the edge object. The cropping process will detect the object whether or not to crop the image according to the object. The cropping result is still in a color image (RGB) to convert to a grayscale image. This conversion uses the segmentation method to get the image of the object only. So that image processing is focused on the object only and does not do many background calculations.

The cropped image (Figure 6. (a)) is converted to a grayscale image. This conversion uses (1) to get an image with a single intensity value (gray level: 0-255). This process is done because the FOS calculation is based on the image’s histogram (single) features. Thus, this conversion makes it easy to calculate the parameters; however, it needs to apply image enhancement (CLAHE-HE method) to improve image quality. This method has contributed to segmentation to clarify fertile and infertile eggs (Figure 6. (c)).

The final result of image processing is a segmented image, as shown in Figure 6. (d). This process is the end of image processing, which is then used as input to calculate the value of the FOS parameter. A compliment is carried out in the final process to show the resulting egg image object and clarify the egg’s characteristics. The sample shown in this image is an image of chicken eggs that are categorized as fertile.

#### 3.2. FOS analysis on chicken egg fertility detection by an object of image processing results

The final image results in image processing are used in calculating the parameter values of the FOS. The segmented grayscale image is an image with a value range between 0-255. FOS is calculated based on the existing histogram features [29] and uses (2)-(6) with the results as shown in Table 1. Table 1 is an example of 10 examples of data calculation results using 5 FOS parameters. The first five data shows the calculation results to identify fertile eggs, and the other five shows infertility.

Based on the results of these calculations, the next is the analysis of the patterns generated from these five parameters. From the ten samples (Table 1), the pattern is shown in Figure 7. In Figure 7, there are two main patterns with similar shapes. The shape of the pattern generated from this sample shows that the fertile and infertile egg images have the same pattern, but the limit values obtained from both are different.
Figure 5. Flowchart of the training process using the BP-NN method

Figure 6. Image processing results based on image preprocessing experiments using the following methods: (a) Cropping (based on the concept of segmentation), (b) Grayscale, (c) Image Enhancement (CLAHE-HE), and (d) Image for Feature Extraction (segmentation) along with the results of the complement

DOI: https://doi.org/10.29207/resti.v5i5.3431
Creative Commons Attribution 4.0 International License (CC BY 4.0)
Based on the five parameters, the values of entropy and skewness have almost similar values and can be included in both (the distribution of values). While the other three parameters (mean, variance, and kurtosis) have different values with the value of the fertile egg image < compared to the infertile egg image. Moreover, in detail, the variance value has the most significant difference in the values of the three parameters. The sample we use is 100 data. From the distribution of the resulting values, the data have almost the same values. Thus, in Table 1, we only present 10 data to represent the processed data.

When tested using training data, the training results can recognize all that is being trained. After that, 50 new test data were processed to check the training results carried out. The FOS parameters are input in identifying the fertility of chicken eggs using the BP method, both in the training and testing process. The experimental results with this scheme are shown in Table 2.

Table 1. Part of sample (10 data) calculated from five data for each image of fertile and infertile eggs

<table>
<thead>
<tr>
<th>Data</th>
<th>M</th>
<th>E</th>
<th>V</th>
<th>S</th>
<th>Q</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>102.03</td>
<td>6.72</td>
<td>4332.69</td>
<td>-0.50</td>
<td>-1.23</td>
<td>Fertile</td>
</tr>
<tr>
<td>2</td>
<td>103.89</td>
<td>6.70</td>
<td>4447.91</td>
<td>-0.57</td>
<td>-1.27</td>
<td>Fertile</td>
</tr>
<tr>
<td>3</td>
<td>100.18</td>
<td>6.62</td>
<td>4719.25</td>
<td>-0.45</td>
<td>-1.42</td>
<td>Fertile</td>
</tr>
<tr>
<td>4</td>
<td>103.70</td>
<td>6.45</td>
<td>4155.24</td>
<td>-0.69</td>
<td>-1.23</td>
<td>Fertile</td>
</tr>
<tr>
<td>5</td>
<td>100.65</td>
<td>6.35</td>
<td>4291.52</td>
<td>-0.61</td>
<td>-1.35</td>
<td>Fertile</td>
</tr>
<tr>
<td>6</td>
<td>58.55</td>
<td>6.35</td>
<td>7699.94</td>
<td>-0.74</td>
<td>-0.39</td>
<td>Infertile</td>
</tr>
<tr>
<td>7</td>
<td>63.32</td>
<td>6.40</td>
<td>9135.50</td>
<td>-0.68</td>
<td>-0.53</td>
<td>Infertile</td>
</tr>
<tr>
<td>8</td>
<td>59.33</td>
<td>6.41</td>
<td>870.94</td>
<td>-0.57</td>
<td>-0.47</td>
<td>Infertile</td>
</tr>
<tr>
<td>9</td>
<td>58.11</td>
<td>5.84</td>
<td>10212.21</td>
<td>-0.84</td>
<td>-0.86</td>
<td>Infertile</td>
</tr>
<tr>
<td>10</td>
<td>56.24</td>
<td>5.98</td>
<td>10234.41</td>
<td>-0.70</td>
<td>-0.97</td>
<td>Infertile</td>
</tr>
</tbody>
</table>

*M: Mean, E: Entropy, V: Variance, S: Skewness, K: Kurtosis

Table 2. Results of Fertility identification experiments using training and test data

<table>
<thead>
<tr>
<th>Neuron hidden layer</th>
<th>Train &amp; Test (F/IF)</th>
<th>Succeed (F/IF)</th>
<th>Failed (F/IF)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>25/25</td>
<td>23/15</td>
<td>2/10</td>
<td>76</td>
</tr>
<tr>
<td>6</td>
<td>25/25</td>
<td>22/17</td>
<td>3/8</td>
<td>78</td>
</tr>
<tr>
<td>12</td>
<td>25/25</td>
<td>24/16</td>
<td>1/9</td>
<td>80</td>
</tr>
</tbody>
</table>

Based on Table 2, egg fertility testing using the BP method and 5 FOS parameters has the best accuracy of 80%. This result is obtained with the number of hidden layers being one and the neuron hidden layer being 12. Previous research shows that the implementation of hidden layers of more than 12 gets the same results, so it is enough with the maximum value [2], [9]. In addition, this feature has almost the same distribution of values, namely mean, variance, entropy, and kurtosis. Thus, the resulting patterns do not differ (see Figure 7 and Table 1). However, this method has a good accuracy of 80%. In addition, the analysis shows that the accuracy obtained is influenced by FOS parameters which have the same pattern. So that in future improvements, we will reduce or change parameters or update the latest feature extraction methods to improve them. In addition, the acquisition process that is less than optimal also affects image processing. Thus, it is also necessary to use acquisition tools and methods that produce a better image or improve it in the preprocessing stage.

4. Conclusion

The implementation of image processing can identify the fertility of chicken eggs (embryo presence) in the incubator. On the seventh day, image processing starts from image acquisition to identify egg fertility based on feature extraction. Feature extraction in this study is FOS Feature Extraction. FOS has five parameters to form different eggs identified as fertile (embryo) and infertile (no embryo). The resulting pattern has differences in each condition (fertile/infertile). Thus, the FOS and BP-Neural Networks method resulted in accuracy testing 100% data test and 80% on new data. The FOS feature extraction can detect the chicken egg embryo on BP-Neural Networks methods based on the results. However, some of the egg images have a similar pattern (FOS), so the parameters of the features implemented in the BP Neural Networks method are

DOI: https://doi.org/10.29207/resti.v5i5.3431
Creative Commons Attribution 4.0 International License (CC BY 4.0)
needed to be reduced or changed. Thus, we recommended improving by adding other features or using second-order feature extraction. However, this method can generally be referred to for improvement in the detection process based on statistical feature extraction. In the future, we will combine the supervised method to increase the accuracy and look for the best feature extraction parameters.

Acknowledgments
The authors express their gratitude to the Universitas Pembangunan Nasional “Veteran” Yogyakarta, contributing to providing the grand research “PDP” on agreement number B/41/UN.62/PT/IV/2021. In addition, the department and LPPM UPN Veteran Yogyakarta also provided support in publishing this article.

References


