

A Computer Vision-Based Framework for Multilevel Egg Fertility Classification Using L*a*b* Color Segmentation, Wavelet Texture Features, and Support Vector Machine

Sriyadi^{1*}, Jufriadif Na'am², Mochamad Wahyudi³, Sumanto⁴

^{1,2}Doctoral Program of Informatics, Universitas Nusa Mandiri, Jakarta, Indonesia

¹Department of Information System, Universitas Bina Sarana Informatika, Jakarta, Indonesia

^{3,4}Master of Information Technology Program, Universitas Bina Sarana Informatika, Jakarta, Indonesia

Received:

June 30, 2026

Revised:

September 5, 2026

Accepted:

October 1, 2026

Published:

October 24, 2026

Corresponding Author:

Author Name*:

Sriyadi

Email*:

sriyadi.sry@bsi.ac.id

DOI:

10.63158/journalisi.v8i5.1882

© 2026 Journal of Information Systems and Informatics. This open access article is distributed under a (CC-BY License)



Abstract. Egg fertility assessment is essential for improving hatchery efficiency, reducing incubation costs, and minimizing resource waste caused by infertile eggs. This study aims to develop a computer vision-based framework for three-level egg fertility classification, namely Highly Fertile, Fertile, and Infertile. A balanced dataset consisting of 1,500 manually collected candling images (500 images per class) was validated by three poultry experts. The proposed method involves L*a*b* color-space transformation, K-Means clustering for region-of-interest segmentation, Discrete Wavelet Transform (DWT) using the LL sub-band for texture feature extraction, and multiclass Support Vector Machine (SVM) classification. The experimental results show that the proposed model achieves 90.9% training accuracy, with precision, recall, and F1-score of 92.0%, 90.9%, and 90.5%, respectively. For the independent testing dataset, the model achieves 89.7% accuracy and 89.4% F1-score, indicating good generalization capability with only a 1.2% difference between training and testing accuracy. The novelty of this research is the development of a three-level egg fertility classification framework that extends conventional binary fertility detection by integrating L*a*b*-based segmentation, LL-wavelet statistical texture features, and multiclass SVM classification.

Keywords: egg fertility classification; egg candling; computer vision; image classification; machine learning; agricultural automation; image processing.

1. INTRODUCTION

The fertility of eggs and their classification are crucial aspects of the poultry industry because they directly affect hatchability, production efficiency, and overall egg product quality. The early detection of egg fertility enables poultry farmers to separate fertile from infertile eggs before incubation, thereby reducing resource waste and improving productivity [1], [2], [3], [4], [5]. Non-destructive testing techniques have received considerable attention in monitoring the quality of food and agricultural products, with hyperspectral imaging being an important technique because it provides both spatial and spectral information simultaneously [6]. Advances in machine learning techniques for rapid analysis and higher classification accuracy have further increased the potential of these methods for food-related applications, including egg classification [6]. In this context, the development of an automated system based on image processing and machine learning is essential to replace manual inspection methods, which are often time-consuming, subjective, and prone to human error [7]. For instance, eggshell crack detection has been considered important for improving fresh egg consumption, breeding egg hatchability, and overall egg product quality [8]. Therefore, research on egg fertility classification using intelligent computational approaches is highly relevant for supporting modern hatchery industries that require efficient, accurate, and objective fertility assessment systems.

Although various methods have been developed for egg quality detection and classification, several significant challenges remain and require further investigation. One of the main challenges is the availability of adequate and high-quality datasets. In many related application domains, the absence of standardized public datasets becomes a serious limitation; for example, in automotive surface defect detection, no public dataset is available for classification, resulting in limited data samples for model development [9]. This dataset issue is also relevant to egg fertility classification, where variations in lighting conditions, egg position, and embryonic development stages introduce complexities that are not fully represented in existing datasets. A previous study reported that fertile and infertile egg classification achieved only 84% accuracy on the fourth day of incubation [6], indicating that existing methods do not consistently provide satisfactory results. In related domains, several classification approaches have also shown similar limitations; for example, the region growing algorithm achieved only 78% accuracy in follicle recognition

[10], SVM-based mammogram classification achieved 80%–83% accuracy [11], and ANN and PCA-ANN approaches for brain tumor image classification produced 75%–90% accuracy [12]. Furthermore, classification using dragonfly clusters with KNN obtained the lowest accuracy of 75.5%, whereas firefly clusters with NBC achieved 75.75% [13]. Similarly, Wavelet scattering with SVM as weak classifiers achieved average accuracies of only 61% and 66%, respectively [14]. These findings indicate that single-method approaches are often insufficient for handling the complexity of biological image data, and a more comprehensive integration of methods is required. Based on previous studies, the classification performance on egg objects and egg quality characteristics still varies considerably, particularly in studies reporting accuracies below 85%. This indicates that previous methods have limitations in feature representation, classification techniques, and model capability in handling variations in egg characteristics. Relevant previous studies are summarized in Table 1.

Table 1. Summary of Previous Studies

No	Researcher (Year)	Number of Classes	Segmentation Methods	Feature Extraction Methods	Classification Methods	Accuracy
1	Saha & Manickavasagan (2021) [15]	2: Fertile, infertile	K-means clustering	—	SVM	84% on day 4
2	Shi et al. (2022) [9]	2: cracked, intact	—	Wavelet features	SVM, LDA, Decision Tree	82.5%
3	Elibol et al. (2020)[16]	4: extra large, large, medium, small	—	DWT	SVM, LBNC	80.4%
4	Adegbenjo et al. (2020)[17]	2–3: fertile, infertile, dead embryo	Conventional candling / machine vision	Spectral and textural features	Various conventional methods	<85%

Therefore, in this study, K-means clustering was employed as an initial segmentation method to separate the region of interest from egg images prior to feature extraction. This segmentation stage is important to ensure that the features extracted in the

subsequent stage focus more specifically on image regions that are relevant to egg fertility indicators.

K-means clustering has several fundamental advantages that make it suitable for segmentation in egg fertility classification. Computationally, K-means is a simple, efficient, and easy-to-implement algorithm for partitioning data into groups based on pixel similarity [11],[19]. In egg image applications, segmentation is a critical stage because egg images captured using candling techniques may contain complex intensity variations among the embryo region, air cell, yolk, and eggshell. K-means can automatically separate the region of interest, particularly the embryo-related area as the main fertility indicator, from the background and other egg components without prior knowledge or initial labels [8], [18]. Another advantage of K-means is its ability to handle significant pixel differences between the target region and the surrounding areas, allowing the clustering process to segment the desired region more accurately [18], [19], [20], [21], [22]. In the context of egg fertility classification, the intensity differences between fertile eggs with developing embryos and infertile eggs without visible embryo development create distinct pixel patterns. These differences can be utilized by K-means to produce more precise segmentation results. In addition, histogram-based centroid initialization can be used to reduce the randomness of initial centroid selection, which often causes convergence to local solutions, thereby improving the consistency of segmentation results [18]. K-means has also been proven effective in medical and biological image segmentation tasks that share similar characteristics with egg images, such as tumor segmentation in MRI images [18], [19] and follicle detection in ovarian ultrasound images [8]. Therefore, the ability of K-means to separate relevant image regions makes it a strong foundation for the subsequent feature extraction stage [19].

However, K-means segmentation alone is insufficient to produce accurate classification results because the segmented output is still an image that requires a discriminative numerical feature representation to distinguish fertility levels. During this stage, the discrete Wavelet transform (DWT) plays a crucial role as a bridge between segmentation and classification [12], [23], [24], [25], [26]. The DWT can extract multi-resolution information from images by decomposing them into approximation and detail components, including horizontal, vertical, and diagonal details at different levels. This enables the extraction of textural, edge, and subtle pattern characteristics that may not be visible at a single

resolution [12], [23]. In the context of egg fertility classification, differences between fertile and infertile eggs often appear in fine textural patterns and micro-intensity variations, which can be captured more effectively through a multi-scale analysis provided by the Wavelet transform [9], [12]. The DWT also has a useful time–frequency localization property, meaning that the Wavelet can simultaneously identify where and at what frequency significant changes occur in an image [19]. The integration of K-means and wavelets is therefore important because K-means first isolates the relevant region of interest, and the DWT then extracts discriminative features from that region with less contamination from irrelevant background pixels. This combination produces cleaner and more informative features than applying the Wavelet directly to the raw image [11], [18], [19], [22], [27], [28]. Moreover, removing the Wavelet transform module from a classification pipeline has been shown to significantly reduce the accuracy, confirming the essential role of the Wavelet in feature extraction [9]. The resulting Wavelet features then serve as suitable inputs for a support vector machine (SVM), which is known for its high generalization capability in high-dimensional feature spaces and limited datasets [9], [15].

SVM is an efficient machine-learning algorithm for decision-making and pattern recognition, with a key advantage in its regularization parameter, which helps reduce the risk of overfitting [18]. SVM performs well in small-sample classification and high-dimensional feature spaces, where its computational complexity is not directly dependent on the dimensionality of the input space. In addition, multi-class SVM is considered robust in handling overfitting problems [29], [30], [31]. In the context of limited datasets, SVM is suitable because its computation depends mainly on the number of support vectors rather than the total number of samples [9]. SVM has also been shown to produce more accurate classification results than other methods, such as neural network learning and k-nearest neighbors, in several biomedical image classification applications [10]. The combination of these three methods—K-means for segmentation, Wavelet for feature extraction, and SVM for classification—has demonstrated strong effectiveness in various domains, including skin disease identification with an accuracy of 97.1% [18], eggshell crack detection with an accuracy above 99% [8], and medical image classification [23], [32].

Previous studies have demonstrated the potential of computer vision and machine learning techniques for automated egg fertility assessment. However, most existing studies have primarily focused on binary classification between fertile and infertile eggs

rather than assessing fertility levels. Pham (2023) [6] investigated fertile and infertile egg classification during incubation and achieved 84% accuracy; however, the study was limited to two-class classification. Other studies have applied image processing, machine learning, and deep learning methods for egg fertility detection and egg quality assessment [7], [8], [9], [10] but these approaches mainly addressed binary decisions or general quality evaluation rather than fertility level grading. For example, Saha and Manickavasagan [15] investigated fertile and infertile egg classification using SVM and achieved 84% accuracy on day 4 of incubation. Similarly, other machine vision-based approaches have mainly focused on binary classification or general egg quality assessment with limited classification levels. These findings indicate that existing approaches still have limitations in representing different fertility conditions and developmental stages required for multilevel egg fertility assessment.

Although previous studies have demonstrated the potential of computer vision and machine learning for egg fertility assessment, most existing approaches are still limited to binary classification between fertile and infertile eggs. However, binary classification provides only a general fertility decision and cannot represent different fertility conditions and developmental stages that are important for more precise hatchery management and decision-making. Therefore, this research proposes a three-level egg fertility classification framework consisting of highly fertile, fertile, and infertile categories.

The novelty of this research lies in developing a three-level egg fertility classification framework that integrates L^* -based K-Means segmentation, Wavelet-based texture feature extraction, and multiclass Support Vector Machine classification. In this framework, L^* -based K-Means is used to extract relevant egg regions, Wavelet transformation is applied to obtain discriminative texture features, and multiclass SVM is employed to classify eggs into highly fertile, fertile, and infertile categories. Unlike previous binary fertility detection approaches, the proposed framework provides a more detailed fertility assessment for objective hatchery screening. The objectives of this research are to develop an image segmentation approach for extracting relevant egg regions, extract discriminative texture features using Wavelet transformation, and evaluate the performance of multiclass SVM for three-level egg fertility classification.

2. METHODS

The overall workflow consists of five main stages: dataset collection and validation, data splitting, L*a*b*-based K-Means segmentation, Wavelet-based feature extraction, and multiclass SVM classification. The complete workflow of the proposed framework is illustrated in Figure 1, the research process begins with egg image collection and expert-based label validation to ensure the quality and correctness of the data used. The dataset is then divided into training and testing datasets so that model development and evaluation can be carried out proportionally. In the next stage, egg images are processed using K-Means to obtain a relevant region of interest, and image features are extracted using a Wavelet transform to represent texture and intensity patterns more discriminatively. The extracted features are subsequently used as inputs for a support vector machine to classify the eggs into three categories: highly fertile, fertile, and infertile. Through this workflow, the proposed method is expected to provide a more systematic and consistent classification process to support image-based egg fertility analysis.

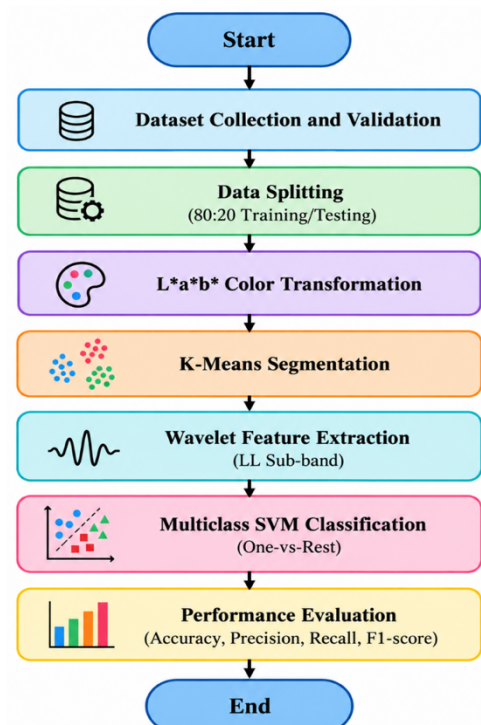


Figure 1. General Research Flowchart of the Proposed Multilevel Egg Fertility Classification Framework

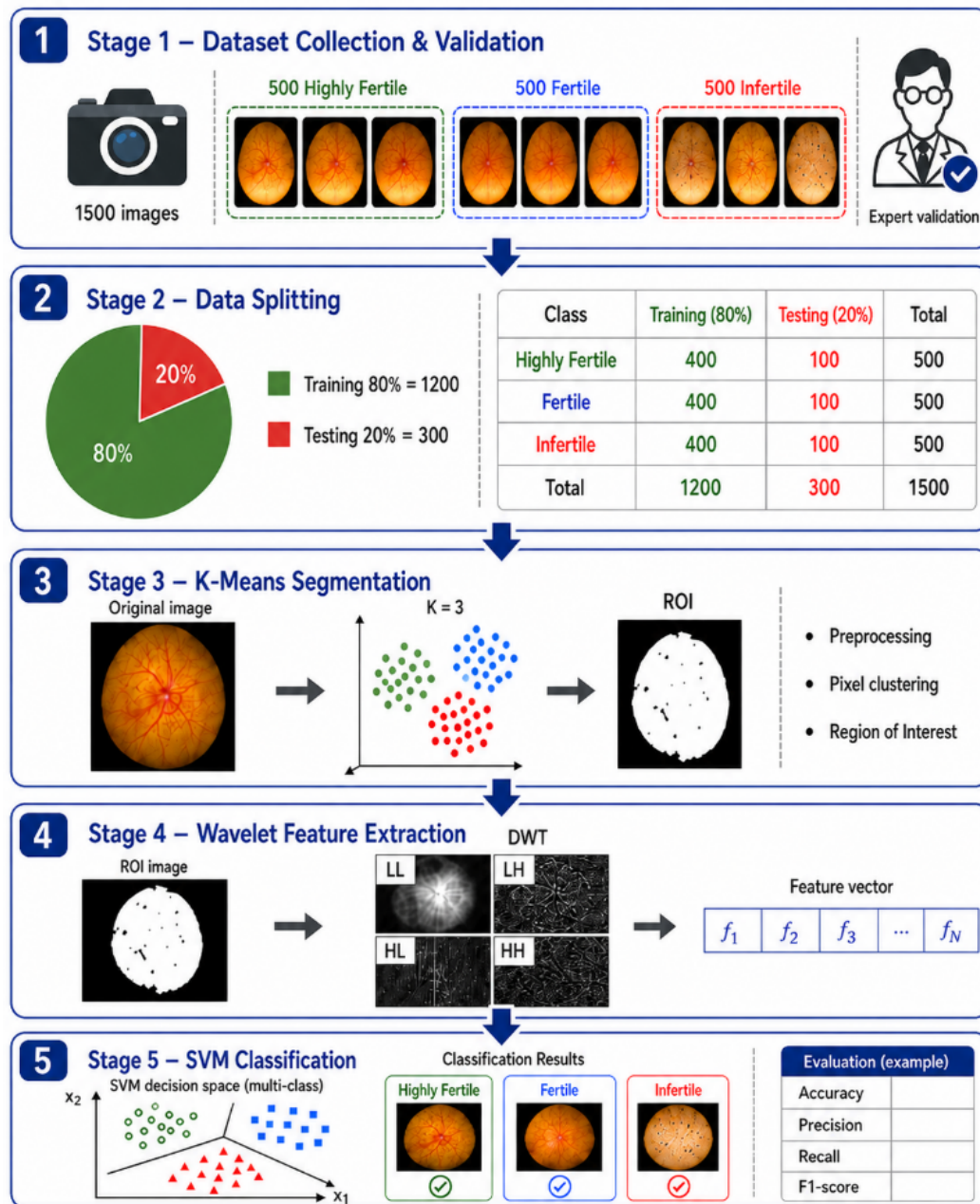


Figure 2. Research Stages

This study proposes an image-based egg fertility level classification framework by combining K-means segmentation, Wavelet feature extraction, and support vector machine (SVM) classification. The research method consists of five main stages: dataset collection and validation, data splitting, image segmentation using K-means, feature extraction using wavelets, and classification using SVM. This workflow is designed to ensure that the classification process is carried out systematically, starting from the use of an expert-validated dataset and culminating in the prediction of egg fertility classes.

2.1. Dataset Collection and Validation

The first stage involves the collection and validation of the egg image dataset. The dataset was collected independently, considering image acquisition conditions such as lighting, egg position, camera distance, and image background. These acquisition conditions were controlled to ensure that the captured images had relatively consistent quality and could be used effectively in the image processing stage.

The dataset used in this study comprised 1,500 egg images divided into three categories: highly fertile, fertile, and infertile eggs. Each category comprised 500 images, resulting in balanced class distribution. After image acquisition, each image was labeled according to its fertility. The image labels were validated by three poultry experts based on the visual fertility indicators observed during candling. In cases of disagreement, the final label was determined by expert consensus. Expert validation is important because the label quality directly affects the success of model training and testing.

Table 2. Dataset Distribution

No	Egg Fertility Category	Number of Images	Description
1	Highly Fertile	500	Validated by experts
2	Fertile	500	Validated by experts
3	Infertile	500	Validated by experts
Total		1500	Balanced dataset

The total number of images can be expressed using Equation (1).

$$N = \sum_{i=1}^C N_i \quad (1)$$

where N represents the total number of images, C represents the number of classes, and N_i represents the number of images in the i -th class. In this study, three classes were used, namely Highly Fertile, Fertile, and Infertile. Therefore, the total number of images is calculated as follows:

$$N = 500 + 500 + 500 = 1500$$

$$C = 3$$

2.2. Data Splitting

The second stage involved data splitting. The validated dataset was divided into training and testing datasets using an 80:20 ratio, with 80% of the images used for training and 20% for testing. The training data were used to build the classification model, whereas the testing data were used to evaluate the final performance of the model on new data that had not been used during the training process. Data splitting was performed proportionally for each class to prevent the model from being biased toward a particular fertility category. Therefore, each class had the same data distribution for training and testing. Of the total of 1,500 images, 1,200 images were used as training data and 300 images were used as testing data. Each class consisted of 500 images, with 400 images used for training and 100 images used for testing. The data division equation can be expressed as follows:

$$\begin{aligned} N_{train} &= 0.80 \times N \\ N_{test} &= 0.20 \times N \end{aligned}$$

With a total dataset of 1,500 images, the following results were obtained:

$$\begin{aligned} N_{train} &= 0.80 \times 1500 = 1200 \\ N_{test} &= 0.20 \times 1500 = 300 \end{aligned}$$

For each class containing 500 images, the data distribution per class is:

$$\begin{aligned} N_{train}^{class} &= 0.80 \times 500 = 400 \\ N_{test}^{class} &= 0.20 \times 500 = 100 \end{aligned}$$

2.3. Image Segmentation Using K-means

The third step involved image segmentation using the K-means clustering algorithm. Segmentation was performed to isolate the important parts of the egg image or regions of interest (ROI) from other irrelevant areas, such as the background, shadows, and noise [33]. In egg fertility classification, not all parts of the image contribute equally. Therefore, segmentation is necessary so that the feature extraction process can focus more on image areas related to egg fertility indicators. In this study, the K-means algorithm was applied using the L*a*b* color space to improve segmentation quality. The L*a*b* color space consists of three components: L* represents lightness, while a* and b* represent chromatic information. This representation separates brightness information from color

components, allowing pixel differences to be represented more effectively during the clustering process. The $L^*a^*b^*$ color space was chosen because it separates the lightness (L^*) component from the a^* and b^* color components, allowing the differences in color characteristics within the egg images to be represented more clearly. The use of the $L^*a^*b^*$ color space helps K-means cluster pixels based on color similarity more stably, especially in images with varying lighting conditions. Thus, $L^*a^*b^*$ -based K-means is expected to produce cleaner and more relevant ROI segmentation.

In this study, RGB images were first converted to the $L^* a^* b^*$ color space. Each pixel is represented as a vector as shown in the equation (2) Each pixel is represented as a three-dimensional feature vector consisting of L^* , a^* , and b^* values, which are used as input features for the K-Means clustering algorithm:

$$x_i = [L_i^*, a_i^*, b_i^*] \quad (2)$$

K-means is used to group pixels into several clusters by minimizing the distance between the pixels and centroids. The K-means objective function is expressed in Equation (3) as follows:

$$J = \sum_{j=1}^K \sum_{x_i \in C_j} \|x_i - \mu_j\|^2 \quad (3)$$

where J represents the objective function, K represents the number of clusters, C_j represents the j -th cluster, x_i represents the i -th pixel, and μ_j represents the centroid of the j -th cluster. The distance between each pixel and the centroid is calculated using Euclidean distance, as shown in Equation (4).

$$d(x_i, \mu_j) = \sqrt{\sum_{m=1}^M (x_{im} - \mu_{jm})^2} \quad (4)$$

where M represents the number of pixel features, x_{im} represents the value of the m -th feature of the i -th pixel, and μ_{jm} represents the value of the m -th feature of the j -th centroid. After each pixel is assigned to the nearest cluster, the centroid is updated by calculating the average of all pixels within the corresponding cluster, as shown in Equation (5):

$$\mu_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i \quad (5)$$

where $|C_j|$ represents the number of pixels in the j -th cluster. The final output of this stage is a segmented image that represents the region of interest (ROI), which is then used as input for the Wavelet feature extraction process.

The clustering process is performed iteratively by assigning each pixel to the nearest centroid, recalculating centroid values, and repeating the process until the centroid values converge. The final cluster representing the relevant egg region is selected as the region of interest (ROI) for further feature extraction.

2.4. Feature Extraction Using Wavelets

The fourth step is feature extraction using the Wavelet transform, specifically the discrete Wavelet transform (DWT). After the egg image has been segmented, the ROI image must be converted into a numerical representation for use by the classification algorithm [34]. Wavelets are used because they can analyze images in a multiresolution domain, thereby capturing information about the texture, edges, and intensity patterns at various levels of detail. In one-level two-dimensional DWT decomposition, the ROI image $I(x, y)$ is decomposed into four sub-bands, as shown in Equation (6):

$$I(x, y) \rightarrow \{LL, LH, HL, HH\} \quad (6)$$

The LL sub-band represents the approximation component or low-frequency information, while LH , HL , and HH represent the horizontal, vertical, and diagonal detail components, respectively. In this study, the LL sub-band was selected because it retains the main structural and intensity information of the egg image. This component provides a more stable representation of fertility-related patterns while reducing the influence of high-frequency noise. In general, the Wavelet Transform can be expressed by the following equation:

$$W_{\psi}(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \quad (7)$$

where $W_{\psi}(a, b)$ represents the Wavelet coefficient, $f(t)$ represents the input signal, $\psi(t)$ represents the mother wavelet, a represents the scale parameter, and b represents the translation parameter.

Statistical features can be calculated from the Wavelet coefficients. Several features used in this study include the mean, standard deviation, energy, and entropy. The mean value is calculated using Equation (8).

$$Mean = \frac{1}{n} \sum_{i=1}^n c_i \quad (8)$$

The standard deviation was calculated using Equation (9).

$$Std = \sqrt{\frac{1}{n} \sum_{i=1}^n (c_i - \bar{c})^2} \quad (9)$$

Energy is calculated using Equation (10).

$$Energy = \sum_{i=1}^n c_i^2 \quad (10)$$

Entropy is calculated using Equation (11).

$$Entropy = -\sum_{i=1}^n p_i \log_2(p_i) \quad (11)$$

where c_i represents the i -th Wavelet coefficient, n represents the number of coefficients, $\bar{c}$ represents the mean of the Wavelet coefficients, and p_i represents the probability of occurrence of the i -th coefficient value. The features extracted from each sub-band are then combined into a feature vector, as expressed in Equation (12).

$$F = [f_1, f_2, f_3, \dots, f_n] \quad (12)$$

This feature vector serves as a numerical representation of the egg image and is used as input in the classification stage using an SVM.

2.5. Classification Using Support Vector Machines

The fifth step involves classifying egg fertility levels using a Support Vector Machine (SVM). The feature vectors extracted via the Wavelet method are used as inputs for the SVM model to classify egg images into three classes: highly fertile, fertile, and infertile. The SVM works by finding an optimal hyperplane that can separate data across classes

with the maximum margin. For binary classification, the SVM decision function can be expressed as shown in Equation (13)[18]:

$$f(x) = w^T x + b \quad (13)$$

Class predictions were determined based on

$$y = \text{sign}(w^T x + b)$$

where x represents the input feature vector, w represents the hyperplane weight vector, b represents the bias, and y represents the class label. The optimization objective of SVM is to minimize the following function:

$$\min \frac{1}{2} \|w\|^2$$

subject to:

$$y_i(w^T x_i + b) \geq 1$$

For data that cannot be linearly separated, soft-margin SVM is used:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i$$

subject to:

$$\begin{aligned} y_i(w^T x_i + b) &\geq 1 - \xi_i \\ \xi_i &\geq 0 \end{aligned}$$

where C represents the regularization parameter and ξ_i represents the slack variable. Because this study uses three classes, an SVM is applied using a multiclass approach. One approach that can be used is one-vs-rest, in which each class is compared with the combination of the remaining classes. In the One-vs-Rest strategy, each SVM classifier is trained to distinguish one fertility category from the other two categories. During prediction, the decision scores from all classifiers are compared, and the class with the highest decision value is selected as the final prediction. The number of SVM models in the one-vs-rest approach is equal to the number of classes:

$$M = C$$

Since there are three classes in this study, then:

$$M = 3$$

The model that was created is:

SVM_1 : Highly Fertile vs. Others

SVM_2 : Fertile vs. Others

SVM_3 : Infertile vs. Others

The final class is determined based on the maximum value of the decision function:

$$\hat{y} = \arg \max_{k=1}^C f_k(x)$$

where $\hat{y}$ represents the predicted class, C represents the number of classes, and $f_k(x)$ represents the decision function value for the k -th class. Model evaluation was performed using testing data that were not used during the training process. The evaluation metrics used in this study include accuracy, precision, recall, and the F1-score. Accuracy is calculated using Equation (14).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (14)$$

Precision is calculated using Equation (15).

$$Precision = \frac{TP}{TP+FP} \quad (15)$$

Recall is calculated using Equation (16).

$$Recall = \frac{TP}{TP+FN} \quad (16)$$

The F1 score was calculated using the following equation (17):

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (17)$$

where TP represents true positive, TN represents true negative, FP represents false positive, and FN represents false negative.

3. RESULTS AND DISCUSSION

This section presents the results of each research stage in sequence, starting with the distribution of the dataset, segmentation results using K-means, feature extraction using wavelets, classification using support vector machines, and ending with the evaluation of model performance. The results are presented in accordance with the research methodology to ensure that the relationships between each stage are clear and systematic.

3.1. Dataset Distribution

The dataset used in this study consists of 1,500 egg images that have been validated by experts. The dataset is divided into three fertility categories: highly fertile, fertile, and infertile. Each category contains the same number of images (500). After the validation process, the dataset was split into training and test sets at an 80% to 20% ratio.

Table 3. Distribution of the Research Dataset

Category	Total Images	Training	Testing
Highly Fertile	500	400	100
Fertile	500	400	100
Infertile	500	400	100
Total	1500	1200	300

The data distribution in Table 3 shows that the number of images in each class is balanced. This data balance is important for reducing potential model bias toward any one fertility class. With this distribution, the model has sufficient data for training, can still be validated during development, and can be tested using data that were not used during training.

3.2. K-means Clustering

Image segmentation in this study was performed using the K-means algorithm to separate the region of interest (ROI) from egg images, allowing the analysis process to focus on areas directly related to fertility indicators. The segmentation stage is important because the acquired egg images may contain not only the main object but also background elements, shadows, light reflections, and intensity variations that can interfere with the

feature extraction process. To obtain more optimal segmentation results, K-means was applied in the $L^*a^*b^*$ color space. This color space was selected because it separates the lightness component (L^*) from the color components (a^* and b^*), allowing visual differences in egg images to be represented more clearly than when using the RGB color space. The separation between lightness and color information helps the K-means algorithm group pixels more stably based on color similarity, particularly in images with lighting variations.

Using this approach, pixels with similar characteristics can be grouped into specific clusters, whereas irrelevant regions can be separated from the main object. As shown in Figure 2, the segmentation results obtained using $L^*a^*b^*$ -based K-means indicate that the main egg region can be separated from the background and other unnecessary image components. This result serves as an important initial stage before feature extraction because the quality of the Wavelet features is strongly dependent on the image region used as input. If the processed image still contains considerable noise, shadows, or background information, the extracted features may become less representative of egg fertility characteristics. Therefore, the use of $L^*a^*b^*$ -based K-means segmentation helps produce cleaner, more focused, and more relevant input images for the subsequent Wavelet feature extraction stage.

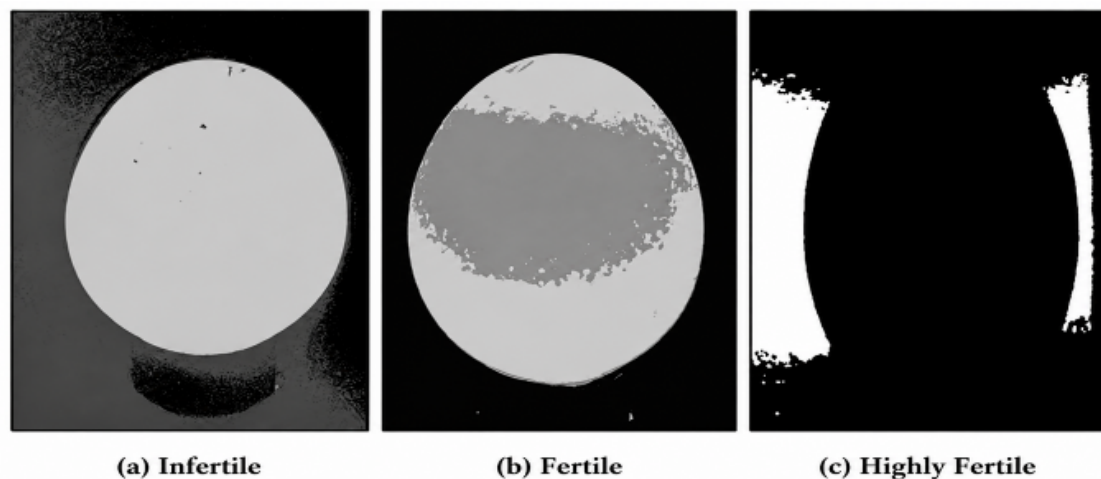


Figure 3. Segmentation Results Using K-means

As shown in Figure 3, the segmentation results indicate that lab-based K-means is capable of generating more precise ROIs for each egg fertility category. These results

suggest that the segmentation stage has provided more relevant image input for the subsequent Wavelet Feature extraction process.

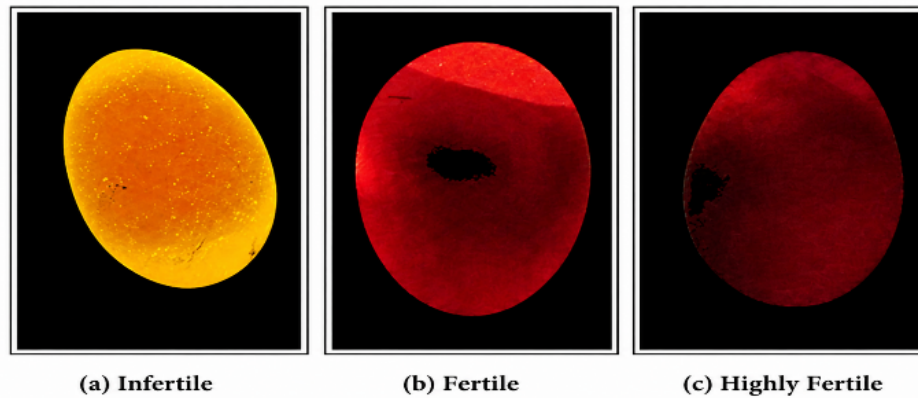


Figure 4. Segmentation Results Using K-means-Lab

3.3. Wavelet Feature Extraction

After K-means segmentation, the egg images underwent Wavelet transformation to obtain a more informative numerical feature representation. As shown in Figure 4, the Wavelet processing results for the three egg categories reveal distinct visual characteristics between the infertile, fertile, and highly fertile classes. In the unfertilized class, the image appears brighter and relatively homogeneous, whereas in the fertilized class, dark areas with certain variations in intensity begin to appear. The highly fertilized class shows more complex textural patterns and intensity distributions. These differences indicate that the Wavelet transform can capture important information in the form of intensity patterns, textures, and changes in the visual structure of egg images. The results of the Wavelet processing are shown in Figure 4.

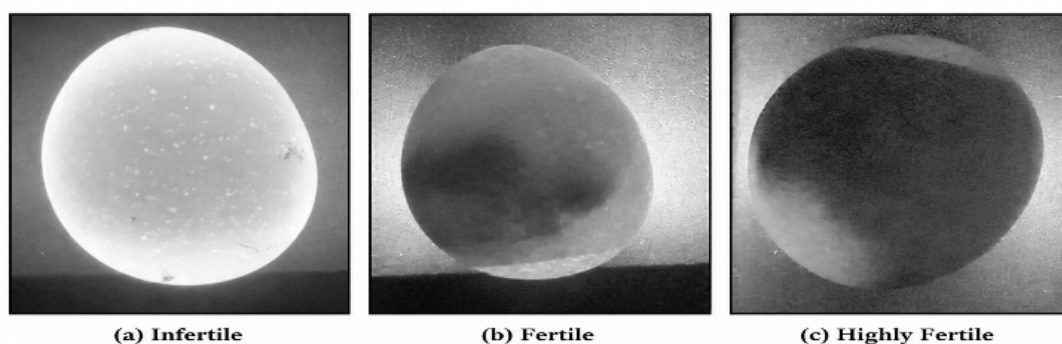


Figure 5. Feature Extraction Results Using Wavelet Transform (LL Sub-band)

The results of the Wavelet feature extraction serve as a crucial foundation for the classification process because the image is no longer represented merely as a visual image but is transformed into numerical feature values, such as the mean, standard deviation, energy, and entropy. These features are used to describe the characteristics of each egg fertility class in a more quantifiable manner. With this feature representation, the SVM classification model can learn the differences in patterns between classes more systematically. Therefore, the Wavelet feature extraction stage acts as a bridge between the image segmentation results and the classification process by providing the discriminative information needed to distinguish between infertile, fertile, and highly fertile categories. The results of each feature extraction, averaged across all images, are presented in Table 4.

Table 4. Wavelet Feature Extraction Results

Category	Mean (Average)	Standard Deviation (Avg)	Energy (Average) 10^{10}	Entropy (Average)
Fertile	116.1557	22.7301	3.5097	5.1994
Infertile	116.7357	25.2094	4.1353	5.1276
Highly Fertile	118.7481	21.2167	4.6269	4.8324

Based on the Wavelet feature extraction results on the LL sub-band, each egg fertility class exhibited different feature values. The Fertile class attained a mean value of 116.1557, a standard deviation of 22.7301, an energy value of 3.5097×10^{10} , and an entropy value of 5.1994. The higher entropy value in the Fertile class indicated that images in this category contained a relatively greater variation of information or textural irregularity compared to those in the other classes. This suggests that the intensity pattern of fertile egg images still has a relatively complex textural distribution.

In the Infertile class, the mean value was 116.7357, standard deviation was 25.2094, energy value was 4.1353×10^{10} , and entropy value was 5.1276. The standard deviation in this class is the highest among the three classes, indicating greater pixel intensity variation in infertile egg images. This condition suggests that infertile egg images have a more widely distributed brightness variation, although their entropy value is slightly lower than that

of the Fertile class. Therefore, the standard deviation can be considered one of the distinguishing features between the Infertile and other classes.

Meanwhile, the Highly Fertile class attained the highest mean value, 118.7481, with a standard deviation of 21.2167, an energy value of 4.6269×10^{10} , and an entropy value of 4.8324. The highest energy value indicates that images in the Highly Fertile class have stronger dominant intensity in the LL sub-band. In contrast, the lowest entropy value suggests that the image patterns in this class tend to be more structured or have a lower information distribution than those in the other classes. The differences in mean, standard deviation, energy, and entropy among the three classes indicate that the LL sub-band provides distinguishable numerical patterns across egg fertility levels, supporting its use as an input for the SVM classifier. These extracted statistical features play an important role in the SVM classification process because they provide a numerical representation of the intensity distribution and texture characteristics of each fertility category. The differences in feature patterns among Highly Fertile, Fertile, and Infertile classes allow the SVM classifier to construct a decision boundary for separating the three fertility levels. Therefore, the Wavelet feature extraction stage contributes directly to the classification performance by providing discriminative information rather than only reducing image data into numerical values.

This study uses only the LL sub-band because it represents the approximation component or the main information of the wavelet-transformed image. The LL sub-band stores low-frequency information related to the global structure, intensity distribution, and main shape of the egg object. In the context of egg fertility level classification, this global information is considered more relevant because the differences among infertile, fertile, and highly fertile classes are mainly reflected in the dominant intensity changes, bright-dark regions, and overall visual patterns of the egg images. Therefore, the LL sub-band was selected to obtain features that are more stable and less affected by noise or irrelevant fine details.

In contrast, the LH, HL, and HH sub-bands mainly represent high-frequency detail components, such as horizontal, vertical, and diagonal edges. Although these components may provide additional information, the detail regions in the acquired egg images are often affected by noise, shadows, eggshell texture, light reflections, and segmentation

artifacts. If the features from LH, HL, and HH are used directly, the extracted features may capture not only fertility-related characteristics but also visual disturbances unrelated to egg fertility levels. Thus, selecting the LL sub-band helps obtain a simpler, more stable, and more focused feature representation based on the main intensity patterns of egg images. This decision also reduces feature complexity before SVM classification, particularly because this study uses statistical features such as mean, standard deviation, energy, and entropy. However, the LH, HL, and HH sub-bands can still be explored in future studies to evaluate the contribution of edge and detailed texture features in improving egg fertility classification performance.

3.4. SVM Model Performance Evaluation

The evaluation of the SVM classification model was conducted in four stages: training evaluation, testing evaluation, confusion matrix analysis, and computational efficiency analysis. This sequence was used to evaluate the model learning capability, generalization performance, classification behavior, and computational requirements. Egg fertility level classification was performed using Support Vector Machine (SVM) based on the features extracted from the Wavelet process. In this stage, the numerical features obtained from feature extraction were used as input to distinguish egg images into three categories: Highly Fertile, Fertile, and Infertile. SVM was selected because it can form an optimal decision boundary between classes based on the given feature vectors. Thus, the model can learn the characteristic patterns of egg images from the training data and use them to predict the fertility class of new data.

Model performance was evaluated using two datasets: 1,200 training images and 300 testing images. The training data were used to examine the model's ability to learn patterns from the extracted Wavelet features, whereas the testing data were used to measure the model's generalization ability on images that were not used during the training process. The classification results are presented in the form of confusion matrices to show the relationship between the actual and predicted classes. In addition, model performance was evaluated using accuracy, precision, recall, and the F1-score.

3.5. Training Evaluation

The first evaluation was performed using the training dataset consisting of 1,200 images. This evaluation aims to measure the capability of the SVM model in learning fertility-

related patterns from the Wavelet Features extracted from the egg images. The training performance was analyzed using the confusion matrix and classification metrics, including accuracy, precision, recall, and F1-score.

Table 5. Confusion Matrix Data Training

Actual / Predicted	Highly Fertile	Fertile	Infertile	Total
Highly Fertile	400	0	0	400
Fertile	52	291	57	400
Infertile	0	0	400	400
Total	452	291	457	1200

As shown in Table 5, the SVM model correctly classified 1,091 out of 1,200 training images. The Highly Fertile and Infertile classes achieved perfect classification, with all 400 images correctly predicted. However, several misclassifications occurred in the Fertile class, where 52 images were predicted as Highly Fertile and 57 images were predicted as Infertile. This indicates that the Fertile category has more complex visual characteristics because it represents an intermediate fertility condition between Highly Fertile and Infertile classes.

3.6. Testing Evaluation

After evaluating the training performance, the model was tested using an independent testing dataset consisting of 300 images that were not involved during the training process. This evaluation was conducted to measure the generalization capability of the proposed SVM classifier on unseen egg images.

Table 6. Confusion Matrix Data Testing

Actual / Predicted	Highly Fertile	Fertile	Infertile	Total
Highly Fertile	69	31	0	100
Fertile	0	100	0	100
Infertile	0	0	100	100
Total	69	131	100	300

Based on Table 6, the SVM model correctly classified 269 out of 300 testing images. The Fertile and Infertile classes achieved 100% correct classification, while the main error occurred in the Highly Fertile class, where 31 images were classified as Fertile. This misclassification indicates that Highly Fertile and Fertile eggs may share similar intensity distribution and texture characteristics in the LL Wavelet feature representation.

3.7. Confusion Matrix Analysis

The confusion matrix analysis shows that the dominant classification errors occurred between the Highly Fertile and Fertile categories. Although both categories represent fertile eggs, their visual characteristics during candling observation may have similar intensity distributions and texture patterns. The LL Wavelet features used in this study capture global intensity and structural information; however, some overlapping feature distributions remain between these two fertility levels, making the classification boundary more challenging for the SVM model.

In contrast, the Infertile category achieved complete classification in the testing dataset. This indicates that infertile eggs have more distinctive visual characteristics compared with fertile categories, allowing the extracted Wavelet features to provide clearer separation. Therefore, the confusion matrix analysis confirms that the proposed K-Means–Wavelet–SVM framework is capable of distinguishing three fertility levels, while further improvement may focus on reducing confusion between closely related fertile categories.

3.8. Computational Efficiency Analysis

In addition to classification performance, computational efficiency was evaluated to determine the processing requirements of the proposed framework. Table 7 presents the computation time and classification performance of the SVM model.

Table 7. Result SVM Model Performance Evaluation

Data	Computation Time	Accuracy	Precision	Recall	F1-Score
Training	1.562	0.909	0.920	0.909	0.905
Testing	0.836	0.897	0.921	0.897	0.894

Based on Table 7, the SVM model achieved an accuracy of 0.909 on the training data and 0.897 on the testing data. These results show that the model produced good classification performance on both the training and testing datasets. The difference in accuracy between the training and testing data was relatively small (0.012), indicating that the model did not show significant overfitting. In other words, the model was not only able to learn patterns from the training data but also maintain its performance when applied to new unseen data.

A precision value of 0.921 on the testing data indicates that the model predictions for egg fertility classes had a high level of correctness. Meanwhile, a recall value of 0.897 shows that the model was able to recognize most images from each fertility class. An F1-score of 0.894 indicates a good balance between precision and recall, demonstrating that the model was consistent in recognizing the three fertility categories.

The computational results also show that the proposed framework requires only 1.562 seconds for training and 0.836 seconds for testing, indicating efficient processing capability. These results suggest that the proposed K-Means-Wavelet-SVM framework has potential for practical implementation in automated egg fertility assessment.

3.9. Comparison with Previous Studies

A comparison with previous studies was conducted to demonstrate the position, contribution, and advantages of the proposed method. This study integrates three main stages—K-means segmentation, Wavelet feature extraction, and SVM classification—into a structured workflow. This approach differs from several previous studies that used only partial processing stages separately, such as focusing only on segmentation and classification or using Wavelet features without applying segmentation as an initial step to obtain the region of interest. By applying segmentation before feature extraction, the analyzed images become more focused on regions relevant to egg fertility indicators, making the extracted features more representative for classification. A comparison with previous studies is presented in Table 8.

Table 8. Comparison with Previous Studies

Researcher	Object	Method	Accuracy
Saha & Manickavasagan (2021)	Egg fertility	K-means + SVM	84%
Shi et al. (2022)	Egg quality	Wavelet + SVM/LDA/DT	78.5%–82.5%
Elibol et al. (2020)	Egg classification	DWT + SVM/LBNC	78.21%–80.4%
This study	Egg fertility	K-means + Wavelet + SVM	89.4%

Based on Table 8, the proposed method achieved an accuracy of 89.4%, which is higher than several previous studies using related approaches on egg fertility or egg quality classification. Saha and Manickavasagan (2021) achieved an accuracy of 84% for egg fertility classification using K-means and SVM. Meanwhile, Shi et al. (2022) obtained an accuracy range of 78.5%–82.5% using Wavelet features with several classifiers, while Elibol et al. (2020) achieved 78.21%–80.4% using DWT with SVM/LBNC classification methods. The improved performance of the proposed K-Means–Wavelet–SVM pipeline can be attributed to the sequential integration of image segmentation, feature extraction, and classification processes. Unlike previous approaches that directly utilize the whole image or apply feature extraction without prior region isolation, the proposed method first applies L^*a^* -based K-Means segmentation to reduce irrelevant background information. Furthermore, the LL Wavelet sub-band provides a compact representation of global intensity distribution and texture characteristics related to fertility development, while SVM effectively separates the extracted feature patterns among the three fertility categories. The advantage of the proposed method is the use of K-means as an initial stage to obtain relevant image regions. Therefore, feature extraction is not performed on the entire raw image but on a more focused region of interest. Wavelet Transform is then used to generate more discriminative texture and intensity features, and the SVM performs the classification based on these extracted features. Thus, the K-means–wavelet–SVM approach not only improves the classification performance but also provides a more systematic, objective, and measurable processing workflow for classifying egg fertility levels. The improvement in classification performance can be attributed to the sequential integration of segmentation, feature extraction, and classification stages in the proposed

pipeline. Unlike approaches that directly classify raw images or apply feature extraction without prior region isolation, the proposed framework first applies L^*a^* -based K-Means segmentation to reduce irrelevant background information and obtain a more focused egg region. Subsequently, LL Wavelet feature extraction provides compact statistical representations of intensity and texture characteristics, while SVM classification effectively separates the extracted feature patterns among the three fertility categories.

3.10. Discussion

The results of this study indicate that classifying egg fertility levels requires a structured processing workflow, starting from segmentation, feature extraction, and classification. K-means segmentation plays an important role in separating the region of interest in egg images from the background, shadows, and other irrelevant image regions. This finding indicates that segmentation is not only used to improve visual separation but also reduces irrelevant variations before feature extraction. By focusing the analysis on the egg region, the extracted features become more representative of fertility-related characteristics and improve the reliability of the subsequent classification process.

Wavelet feature extraction contributes significantly by representing texture characteristics, intensity patterns, and visual details of egg images in a multi-resolution manner. The extracted statistical features, including mean, standard deviation, energy, and entropy, provide quantitative representations of intensity distribution and texture characteristics for each fertility category. These feature differences become important inputs for the SVM classifier to construct decision boundaries among Highly Fertile, Fertile, and Infertile classes.

In this study, the LL sub-band was selected because it provides a stable representation of low-frequency information related to the overall intensity distribution and structural characteristics of egg images. The selection of the LL sub-band represents a balance between preserving relevant fertility information and reducing sensitivity to unnecessary high-frequency variations. High-frequency components may contain detailed patterns related to eggshell texture, illumination reflection, or acquisition noise rather than fertility characteristics. Compared with the LH, HL, and HH sub-bands, the LL component is more suitable for this study because egg fertility classification mainly depends on global intensity changes and structural patterns observed during candling. In contrast, LH, HL,

and HH mainly represent high-frequency details, such as edges and local variations, which may introduce less relevant information for fertility-level discrimination. Therefore, LL features provide a more stable and compact feature representation that supports the SVM classification process. In addition, the LL sub-band reduces the influence of high-frequency variations associated with eggshell surface noise, uneven illumination during candling, and reflection artifacts. By minimizing the contribution of these irrelevant variations, LL features provide more consistent texture and intensity information related to fertility characteristics, thereby improving the reliability of the classification process.

Compared with previous studies, the proposed approach provides an improvement through the integration of K-means segmentation, Wavelet feature extraction, and SVM classification into a systematic workflow. Previous research, such as fertile and infertile egg classification on the fourth day of incubation, achieved an accuracy of 84%. The improved performance of the proposed framework can be attributed to the use of La-based K-means segmentation to reduce irrelevant image regions and LL Wavelet features to capture fertility-related intensity and texture characteristics before classification. Although the proposed framework achieved good classification performance, some confusion remained between Highly Fertile and Fertile categories. This limitation is related to the similarity of their visual patterns and overlapping feature distributions in the LL Wavelet representation. These conditions make the SVM decision boundary more challenging to separate the two closely related fertility levels.

Despite achieving promising performance, several limitations should be considered. First, the dataset used in this study was obtained under relatively controlled image acquisition conditions. Therefore, the performance of the proposed framework may vary when applied to more complex industrial hatchery environments. Variations in illumination conditions, camera devices, egg positioning, background conditions, and image acquisition settings may affect the segmentation and feature extraction processes. Second, although the dataset was balanced among the three fertility categories, its size is still limited compared with the diversity of conditions found in real-world hatchery operations. Third, this study focuses on image-based classification performance and does not yet evaluate real-time processing capability in an automated industrial system.

From an industrial application perspective, the proposed $L^*a^*b^*$ -based K-Means segmentation, LL Wavelet feature extraction, and SVM classification framework has potential to be integrated into automated egg-candling conveyor systems for hatchery management. Such implementation can utilize fixed cameras equipped with controlled illumination to automatically capture egg images during the candling process. The captured images can then be processed sequentially through image segmentation, feature extraction, and classification stages to provide automatic fertility-level assessment. This integration has the potential to reduce manual candling workload, improve inspection consistency, and support more efficient decision-making in industrial hatchery operations. For practical deployment, several technical considerations should be addressed, including camera positioning, illumination consistency, synchronization between conveyor movement and image acquisition, and processing speed requirements for real-time inspection. Although the proposed SVM-based framework demonstrates efficient computational performance, further validation using industrial-scale datasets and real-time hardware platforms is required to evaluate its robustness under actual hatchery conditions.

Future studies should further validate the proposed framework using larger datasets collected under diverse hatchery conditions and evaluate its robustness against variations in real-world image acquisition environments. In addition, further investigation involving real-time implementation, optimization of processing speed, and comparison with advanced segmentation and classification approaches can be conducted to improve the generalization capability and practical applicability of the proposed framework.

4. CONCLUSION

This study successfully developed an image-based egg fertility level classification framework by integrating K-means segmentation in the $L^*a^*b^*$ color space, low-low (LL) sub-band Wavelet Transform for feature extraction, and Support Vector Machine (SVM) for classification. Unlike previous studies that mainly focused on binary egg fertility classification, this study classified eggs into three fertility levels: highly fertile, fertile, and infertile. The proposed framework was evaluated using a dataset consisting of 1,500 egg images, with 1,200 images used for training and 300 images used for testing. The experimental results demonstrated that the SVM classifier achieved 90.9% accuracy, 92.0%

precision, 90.9% recall, and 90.5% F1-score on the training dataset, while achieving 89.7% accuracy, 92.1% precision, 89.7% recall, and 89.4% F1-score on the testing dataset. The small performance difference between training and testing results indicates that the proposed framework provides good generalization capability with limited overfitting. The main contribution of this study lies in extending conventional binary egg fertility detection into a multilevel fertility classification framework by combining L*a*-based K-means segmentation, LL Wavelet feature representation, and SVM classification. From a practical perspective, the proposed framework has the potential to support automated hatchery management by enabling more objective fertility screening, reducing manual candling workload, and improving incubation efficiency through earlier identification of fertility levels. However, this study was conducted using egg images acquired under controlled conditions; therefore, further validation using more diverse hatchery environments, including variations in illumination, camera distance, egg orientation, and acquisition conditions, is required to further assess the robustness and practical applicability of the proposed framework.

REFERENCES

- [1] X. Yang, R. B. Bist, S. Subedi, and L. Chai, "A Computer Vision-Based Automatic System for Egg Grading and Defect Detection," *Animals*, vol. 13, no. 14, 2023, doi: 10.3390/ani13142354.
- [2] R. L. M. Bambilla, M. J. B. Lee, and F. L. Valiente, "Freshness Condition Detection, Size Classification with Weight Prediction and Sortation of Philippine Egg Produce Using YOLOv8 and OpenCV," in *2025 17th International Conference on Computer and Automation Engineering, ICCAE 2025*, School of Electrical, Electronics and Computer Engineering, Mapúa University, Manila, Philippines: Institute of Electrical and Electronics Engineers Inc., 2025, pp. 158–164. doi: 10.1109/ICCAE64891.2025.10980499.
- [3] H. A. Muslim, H. Oktavianto, R. T. Widodo, E. Purwantini, and W. T. Sesulihatien, "Implementation of Chicken Eggs Fertility Detection Device Using SSD MobileNet-V2 FPNLite," *2024 Int. Electron. Symp. Shap. Futur. Soc. 5.0 Beyond, IES 2024 - Proceeding*, pp. 171–177, 2024, doi: 10.1109/IES63037.2024.10665784.
- [4] D. Sun, H. Wei, Y. Luo, W. Li, and Y. Yu, "Real-Time Detection Method for Cracked Eggs Based on Improved YOLOv8 and Dual Verification with Triple Classification Granularity," *IEEE Access*, 2026, doi: 10.1109/ACCESS.2026.3686452.

- [5] S. L. de Castro Júnior, A. E. C. M. Cândido, A. C. de S. Silva, and I. J. O. da Silva, "Translucency of chicken eggs: proposal of an automated system for analysis and classification," *Poult. Sci.*, vol. 104, no. 10, 2025, doi: 10.1016/j.psj.2025.105612.
- [6] T. D. Pham, "Classification of Motor-Imagery Tasks Using a Large EEG Dataset by Fusing Classifiers Learning on Wavelet-Scattering Features," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1097–1107, 2023, doi: 10.1109/TNSRE.2023.3241241.
- [7] P. Zeng *et al.*, "High-Accuracy Classification of Fertilized Eggs Using Mamba-Based Model," in *2025 6th International Conference on Artificial Intelligence and Electromechanical Automation, AIEA 2025*, School of Physics and Information Engineering, Guangdong University of Education, Guangzhou, 510303, China: Institute of Electrical and Electronics Engineers Inc., 2025, pp. 92–96. doi: 10.1109/AIEA66061.2025.11159668.
- [8] P. G. Gülhan and G. Özmen, "Follicle Detection for Polycystic Ovary Syndrome by Using Image Processing Methods," *Int. J. Appl. Math. Electron. Comput.*, vol. 8, no. 4, pp. 203–208, 2020, doi: 10.18100/ijamec.803400.
- [9] C. Shi *et al.*, "Nondestructive Detection of Microcracks in Poultry Eggs Based on the Electrical Characteristics Model," *Agric.*, vol. 12, no. 8, p. 1137, 2022, doi: 10.3390/agriculture12081137.
- [10] S. Anantharajan and S. Gunasekaran, "Automated brain tumor detection and classification using weighted fuzzy clustering algorithm, deep auto encoder with barnacle mating algorithm and random forest classifier techniques," *Int. J. Imaging Syst. Technol.*, vol. 31, no. 4, pp. 1970–1988, 2021, doi: 10.1002/ima.22582.
- [11] I. K. G. D. Putra, N. P. A. O. Wiastini, K. S. Wibawa, and I. M. S. Putra, "Identification of Skin Disease Using K-Means Clustering, Discrete Wavelet Transform, Color Moments and Support Vector Machine," *Int. J. Mach. Learn. Comput.*, vol. 10, no. 4, pp. 542–548, 2020, doi: 10.18178/ijmlc.2020.10.4.970.
- [12] D. M. Toufiq, A. M. Sagheer, and H. Veisi, "Brain tumor identification with a hybrid feature extraction method based on discrete wavelet transform and principle component analysis," *Bull. Electr. Eng. Informatics*, vol. 10, no. 5, pp. 2588–2597, 2021, doi: 10.11591/eei.v10i5.3013.
- [13] F. A. Jibon *et al.*, "Cancerous and Non-Cancerous Brain MRI Classification Method Based on Convolutional Neural Network and Log-Polar Transformation," *Healthc.*, vol. 10, no. 9, p. 1801, 2022, doi: 10.3390/healthcare10091801.

- [14] S. K. Prabhakar and D. O. Won, "Performance comparison of bio-inspired and learning-based clustering analysis with machine learning techniques for classification of EEG signals," *Front. Artif. Intell.*, vol. 6, 2023, doi: 10.3389/frai.2023.1156269.
- [15] D. Saha and A. Manickavasagan, "Machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review," *Curr. Res. Food Sci.*, vol. 4, pp. 28–44, 2021, doi: 10.1016/j.crfs.2021.01.002.
- [16] G. Elibol, M. Yumurtaci, S. Ergin, and İ. Yabanova, "Classification of Dynamic Egg Weights Using Feature Extraction Methods," *Eskişehir Tech. Univ. J. Sci. Technol. a - Appl. Sci. Eng.*, vol. 21, no. 4, pp. 499–513, 2020, doi: 10.18038/estubtda.658077.
- [17] A. O. Adegbenjo, L. Liu, and M. O. Ngadi, "Non-destructive assessment of chicken egg fertility," *Sensors (Switzerland)*, vol. 20, no. 19, pp. 1–23, 2020, doi: 10.3390/s20195546.
- [18] D. Kalpana and K. S. Reddy, "SVM for Classification of Brain MRI with K-Means for Detection," *Interantional J. Sci. Res. Eng. Manag.*, vol. 07, no. 11, pp. 1–11, 2023, doi: 10.55041/ijrsrem26641.
- [19] G. Yuan, Y. Liu, W. Huang, and B. Hu, "Differentiating Grade in Breast Invasive Ductal Carcinoma Using Texture Analysis of MRI," *Comput. Math. Methods Med.*, vol. 2020, pp. 1–14, 2020, doi: 10.1155/2020/6913418.
- [20] A. Nanda, R. C. Barik, and S. Bakshi, "SSO-RBNN driven brain tumor classification with Saliency-K-means segmentation technique," *Biomed. Signal Process. Control*, vol. 81, 2023, doi: 10.1016/j.bspc.2022.104356.
- [21] S. Fooladi, H. Farsi, and S. Mohamadzadeh, "Segmenting the Lesion Area of Brain Tumor using Convolutional Neural Networks and Fuzzy K-Means Clustering," *Int. J. Eng. Trans. B Appl.*, vol. 36, no. 8, pp. 1556–1568, 2023, doi: 10.5829/IJE.2023.36.08B.15.
- [22] O. S. Faragallah, H. M. El-Hoseny, and H. S. El-sayed, "Efficient brain tumor segmentation using OTSU and K-means clustering in homomorphic transform," *Biomed. Signal Process. Control*, vol. 84, 2023, doi: 10.1016/j.bspc.2023.104712.
- [23] A. Saday and İ. A. Ozkan, "Classification of Epileptic EEG Signals Using DWT-Based Feature Extraction and Machine Learning Methods," *Int. J. Appl. Math. Electron. Comput.*, vol. 9, no. 4, pp. 122–129, 2021, doi: 10.18100/ijamec.988691.
- [24] I. Zine-Dine, J. Riffi, K. El Fazazy, I. El Batteoui, M. A. Mahraz, and H. Tairi, "A review: Machine Learning Techniques of Brain Tumor Classification and Segmentation," *Mach. Graph. Vis.*, vol. 34, no. 3, pp. 31–55, 2025, doi: 10.22630/MGV.2025.34.3.2.

- [25] A. Tayeh Swadi and F. Sabar Miften, "Detection of attention deficit hyperactivity disorder based on EEG signals using Least Square Support Vector Machine (LS-SVM)," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, vol. 11, no. 6, pp. 2495–2507, 2023, doi: 10.1080/21681163.2023.2243347.
- [26] K. P. Ajitha Gladis, A. Ahilan, N. Muthukumaran, and L. Jenifer, "ECG Arrhythmia Measurement and Classification for Portable Monitoring," *Meas. Sci. Rev.*, vol. 24, no. 4, pp. 118–128, 2024, doi: 10.2478/msr-2024-0017.
- [27] P. Yan, Y. Zhang, S. Hou, W. Zhang, and X. Yang, "Noise reduction method for blasting vibration signals based on improved EMD-wavelet packet," *Zhendong yu Chongji/Journal Vib. Shock*, vol. 43, no. 11, pp. 264-271and287, 2024, doi: 10.13465/j.cnki.jvs.2024.11.029.
- [28] H. Yang, Y. Zhai, J. Liang, D. Guo, X. Liu, and R. Zhang, "Algorithm for Extracting Wear Characteristics of Piston Pins Based on Orthogonal Empirical Mode Decomposition," *Nongye Jixie Xuebao/Transactions Chinese Soc. Agric. Mach.*, vol. 55, pp. 412–419, 2024, doi: 10.6041/j.issn.1000-1298.2024.S1.044.
- [29] K. K. Çevik, H. E. Koçer, and M. Boğa, "Deep Learning Based Egg Fertility Detection," *Vet. Sci.*, vol. 9, no. 10, 2022, doi: 10.3390/vetsci9100574.
- [30] J. L. Dioses, R. P. Medina, A. C. Fajardo, and A. A. Hernandez, "Performance of Classification Models in Japanese Quail Egg Sexing," in *Proceeding - 2021 IEEE 17th International Colloquium on Signal Processing and Its Applications, CSPA 2021*, Graduate School, Technological Institute of the Philippines, Quezon City, Philippines: Institute of Electrical and Electronics Engineers Inc., 2021, pp. 29–34. doi: 10.1109/CSPA52141.2021.9377275.
- [31] Y. F. Wei, P. Yang, Z. Y. Yang, P. Wang, and X. W. Wang, "Fault Detection of Flexible DC Grid Based on Empirical Wavelet Transform and WOA-CNN," *J. Electr. Eng. Technol.*, vol. 20, no. 1, pp. 217–229, 2025, doi: 10.1007/s42835-024-02038-9.
- [32] M. H. Saleem, J. Potgieter, and K. M. Arif, "Automation in Agriculture by Machine and Deep Learning Techniques: A Review of Recent Developments," *Precision Agriculture*, vol. 22, no. 6. pp. 2053–2091, 2021. doi: 10.1007/s11119-021-09806-x.
- [33] S. Saifullah, R. Drezewski, A. Khaliduzzaman, L. K. Tolentino, and R. Ilyos, "K-Means Segmentation Based-on Lab Color Space for Embryo Detection in Incubated Egg," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 8, no. 2, p. 175, 2022, doi: 10.26555/jiteki.v8i2.23724.

- [34] V. Kiruthika and M. M. Ramya, "Automatic segmentation of ovarian follicle using K-means clustering," *Proc. - 2014 5th Int. Conf. Signal Image Process. ICSIP 2014*, pp. 137–141, 2014, doi: 10.1109/ICSIP.2014.27.