Zero : Jurnal Sains, Matematika, dan Terapan
E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 9, Number 2, 2025
DOI: 10.30829/zero.v9i2.26263
Page: 587-599

Egg Quality Classification Using Support Vector Machine Based on Image and Non-Image Fusion

[1] Faizal Abrolillah
Department of Computer Science, State Islamic University of Maulana Malik Ibrahim, Malang, 65144, Indonesia

[2] Irwan Budi Santoso
Department of Computer Science, State Islamic University of Maulana Malik Ibrahim, Malang, 65144, Indonesia

[3] Totok Chamidy
Department of Computer Science, State Islamic University of Maulana Malik Ibrahim, Malang, 65144, Indonesia

Article Info

Article history:

Accepted, 30 October 2025

Keywords:

Egg Quality;
Feature-level Fusion;
Poultry Industry.

ABSTRACT

Egg production and consumption in Indonesia continue to rise, highlighting the need for accurate egg quality assessment. This study evaluated egg quality using a Support Vector Machine (SVM) model that integrates image and non-image features through feature-level fusion. A total of 750 eggs were analyzed based on external characteristics (shell color, cleanliness, texture, weight, and images) and internal characteristics (odor, albumen, yolk, black spots, images). Image data were reprocessed through grayscale conversion, resizing, and texture extraction using the Gray Level Co-occurrence Matrix (GLCM). Both linear and polynomial SVM kernel with varying degrees were tested, and the polynomial kernel (degree 6) achieved the best, with 86% accuracy, 91% precision, and 87% recall. These results demonstrate that integrating image and non-image features significantly enhances egg quality classification compared to using either data type alone. These findings provide valuable insights for developing automated egg grading system in the poultry industry.

Corresponding Author:

Faizal Abrolillah
Department of Computer Science
State Islamic University of Maulana Malik Ibrahim
Email: abrorif906@gmail.com

1. INTRODUCTION

Chicken egg production and consumption in Indonesia have steadily increased, with production rising from 5.14 to 5.15 million tons and consumption form 1.6 to 1.7 million tons between 2020 and 2021[1]. Eggs are an affordable and nutrient-rich source of animal protein, vitamins (A, D, E, B12), and mineral such as iron and selenium, playing an important role in public health both in Indonesia and globally [2]. While egg production is achievable, the factors ensuring consistent high quality remain unclear. Quality deterioration, including external defects or internal spoilage, reduces market value, undermines consumer satisfaction, and may pose health risks

[3]. Egg quality, comprising external factors (shell colors, cleanliness, thickness, strength) and internal factors (albumen, yolk index, albumen quality, and blood spot) is influenced by biological and environmental conditions [4]. Normally, traditional assessment has been performed with visual examination manually/egg candling and these methods are subjective, variable, and inefficient [5]. This restriction highlights the requirement for automatic, objective and reproducible methods [6].

In the past two decades, various studies have explored machine learning and image processing techniques for egg quality assessment [7]. Despite the range of alternative methods to characterize the texture, image-based methods are commonly used as they include external as well as internal characteristics. Several tools have been exploited for such purpose, such as, Gray Level Co-Occurrence Matrix (GLCM) that derives textu re features such as shell patterns and albumen features [8] [9] [10]. Because they are promising, however, such image-based methods are often sensitive to illumination, restricted by image quality, and based on a small range of features. Other studies have focused on non-image data, including egg weight, shell cleanliness, yolk quality, and albumen quality [11] [12]. Although also less complicated and not dependent on high-end imaging, these methods do not use the large amount of structural information that can be collected from egg images. Some works tried to consider both internal and external factors, however, not in image-based feature extraction [13]. Although promising, such studies did not sufficiently leverage the power of multi-modal data fusion.

Both internal and external factors have been widely studied. Internal features, such as those from SPF chicken embryo eggs, can be analyzed using image segmentation techniques to detect embryos. [11] Indicators like yolk color also reflect egg quality. Studies with large datasets (up to 10.000 samples) reported high accuracy (up to 98,4%), although relying solely on internal features sometimes reduced accuracy to (67,3%). [9] External factors including shell color, cleanliness, cracks, and also weight, have been used for classification, with digital imaging methods outperforming simple thresholding under controlled lighting [14], [15]. Poor shell quality significantly reduces market value, reinforcing the need for comprehensive assessment [16], [17].

Research combining internal and external factors shows that internal factor-based methods achieve high accuracy but are sensitive to imaging and feature extraction variations, while external factor-based methods are simpler yet less effective in detecting internal quality determinants. Combined approaches are emerging but remain limited in sample size and attribute diversity. For example, the identification of egg quality using 100 digital image samples and the Support Vector Machine (SVM) method achieved an accuracy of 84.57%. Combining internal parameters (such as blood spots and blood vessels) with external parameters (color, size, shell cracks) improved classification results compared to using either data type alone. This approach also mimics the manual evaluation process performed by egg inspectors but is more objective due to its basis on image processing and classification algorithms.

Building upon these limitations, the present work introduces two main contributions: (i) it proposes a multimodal integration pipeline that has been less explored in the previous works and (ii) it compares linear and polynomial SVM kernels to yield the classification performance. [11] The proposed approach aims to provide a more accurate and practical automated egg-quality classification system for industry applications, integrating GLCM-based texture extraction from both internal and external egg images, as illustrated in Figures 1 and 2. [12]

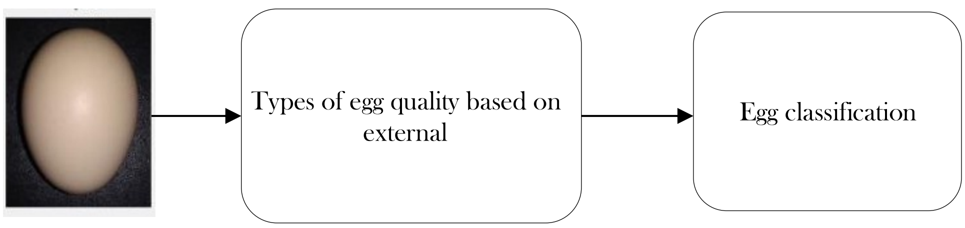

Figure 1. Egg quality based on External

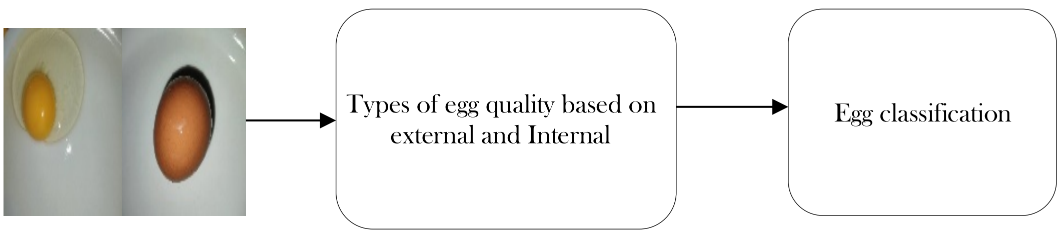

Figure 2. Egg quality based on External & Internal

This study addresses the existing gap by focusing on the classification of egg quality usiang a multimodal approach that combines image and non-image features with a Support Vector Machine (SVM). Image and non-image features included characteristics of eggs (egg weight, shell cleanliness, albumen firmness, yolk index) and

GLCM features (energy, contrast, homogeneity, correlation). After preprocessing, the two datasets were fused at the feature level and used as input to the SVM classifier. The study sistematically compares linear and polynomial SVM kernels, highlighting their ability to handle both linear and non linear relationship in the data [8].

The novelty of this study lies in its integration of multimodal data, capturing both internal and external egg characteristics, and in systematically evaluating SVM kernel function to optimize classification performance. This study aims to determine whether a multimodal SVM model combining image and non-image features can improve egg quality classification accuracy and robustness. We hypothesize that the polynomial kernel will outperform the linear kernel and models using only single-feature types, providing superior classification results by effectively balancing model flexibility and generalization.

2. RESEARCH METHOD

In this study, egg quality classification used a multimodal approach combining image (GLCM-based texture) and non-image features (weight, shell cleanliness, albumen firmness, and yolk index) were fused at the feature level by concatenating their vectors. This early fusion strategy allows the SVM classifier to learn from both types. The system was designed so that all components interact and work synergistically, as illustrated in Figure 3, which shows the detailed training and testing workflow from start to finish.

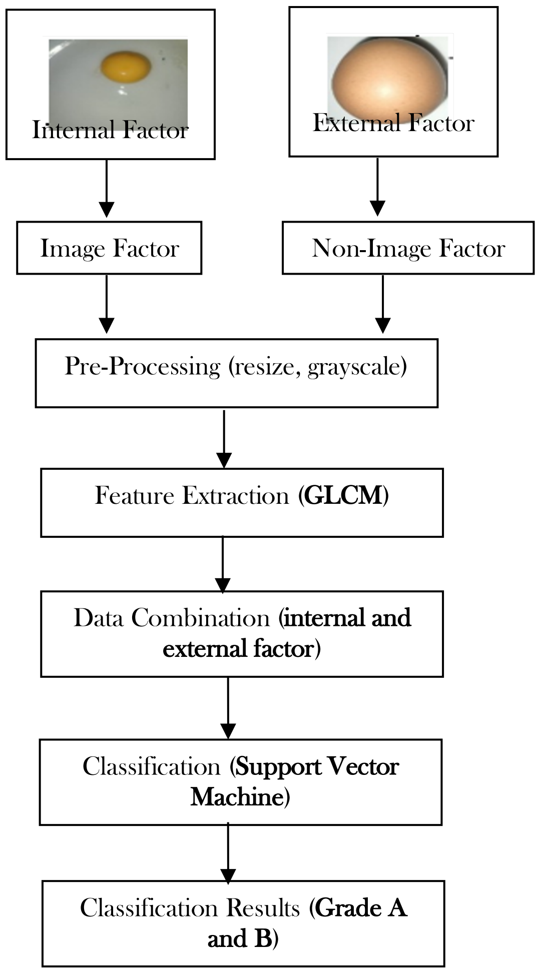

Figure 3. Training and testing scheme

Data collection is very important to obtain information that will be processed into data relevant to our needs. The data used in this study is primary data, where data collection involves several components as special tools, namely the Realme 9 Pro 5G smartphone to take pictures with a 50-megapixel (MP) camera, a TNW Overhead phone holder, and an LED light included in the tool, a gram scale as an egg weighing tool, specifically an SF 400 electronic kitchen scale. A plate with a white background is used to take photos of eggs based on external and internal tools, as shown below.

Figure 4. Realme 9 Pro 5G

Figure 5. HP TNW Overhead Telephone Stand

Figure 6. Scales brand electronic kitchen scale type SF 400

Figure 7. Plate on a white background

This research classifies egg quality based on external and internal features using a dataset of 750 eggs. This sample size was selected to provide sufficient statistical power, while ensuring a balanced distribution across quality classes. Features include egg weight, shell cleanliness, albumen firmness, yolk index, and GLCM-based image descriptors. An example of dataset is shown in the table below.

Table 1. Egg quality attributes based on external dan internal factors

Factor	Attribut	Variabel	Measurement Scale
Eksternal	Shell colour	X1	1. Brown 2. Cream 3.White
	Shell cleanliness	X2	1. Dirty 2. Not dirty
	Shell Texture	X3	1. Smooth 2. rough
	Egg Weight	X4	Weight in Grams
	Egg Surface Image	X5	100 x 100 Piksel (RGB)
Factor	**Attribut**	**Variabel**	**Measurement Scale**
Internal	Smell	X6	1. Smells fishy 2. Foul-smelling
	Egg white	X7	1. Thick 2. Medium 3. Liquid
	Egg Yolk	X8	1. Broken 2. Unbroken
	Black Spots	X9	1. There is 2. None
	Image of Egg Filling	X10	100 x 100 Piksel (RGB)

In this study, egg quality is divided into several parts between external and internal, below are some of the parts that are the benchmarks in this study as follows. External factors Eggshells or shells in the data that we display related to egg quality based on external factors include the fineness of the egg shell according [11] the smoother and the state of the egg intact and not c=racked, the colour of the egg that makes whether the egg is quality, The ideal egg weight shown above is categorized as follows, with the division of Grade A <50 grams, Grade B 50-60 grams, and GradeC> 50 grams [12], eggs in terms of the cleanliness of the egg shell, namely the presence or absence of dirt attached to the egg shell, and the egg surface image using GLCM [8]. Egg odour is included in the internal quality factor of the egg where if the egg smells bad then it is certain that the quality of the egg is bad or not worth trading, the colour of the yolk is one of the factors of egg quality [18], the value of the egg white to determine the quality of the egg [11], and the last bitnik blood or blood spot which is the same as the external quality of the egg surface image using GLCM features.

2.1 Pre-processing

At this stage, the pre-processed data is converted into a grey-scale image format. Next, features are extracted using the Grey Level Co-occurrence Matrix (GLCM) method by taking four main parameters, namely energy, contrast, homogeneity and correlation. To ensure the accuracy of the data, the following tests were conducted.

Resize

Resizing is the initial stage in digital image processing, which serves to standardize image size before further analysis. For example, 32 × 32 pixels is used to ensure consistent processing of input into the system. Standardizing image size through resizing also aims to reduce computational complexity without eliminating important features required for classification. This stage is commonly used in image processor-based and computer vision research, including eggshell quality detection using digital images [19], a microcontroller-based egg classification system [20], and the development of an automated system for grading and detecting egg damage using computer vision [21].

Grayscale

Grayscaling process is the stage of converting a color image (RGB) into a grayscale image by maintaining the light intensity at each pixel and ignoring its color information. The intensity value of a grayscale image is generally represented in the range 0–255 [19]. According to [22], the color-to-grayscale conversion method can affect image recognition results. One formula commonly used in this process is:

$$graysacale = 0.299\text{R} + 0.578\text{G} + 0.114{*}\text{B} \tag{1}$$

This formula takes into account the sensitivity of the human eye to color, where the green channel has the greatest influence on the perception of intensity, followed by red, and the least blue (Kanan & Cottrell, 2012). RGB primaries in the range 0 to 1 and $Clinier$ is the intensity value in the 0 to 1 field with the conversion obtained using the $f(x)$ function. Function (x) converts RGB values into grey scale values by summing the R, G, and B components [19] [22]with results as shown in Figures 8 and 9 below.

Figure 8. Egg quality based on External and Internal

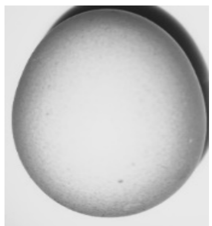

Figure 9. Egg quality based on External

Gray Level Co-Occurance Matrix

In 1973, Haralick, Shanmugam, and Dinstein presented the Gray Level Co-occurrence Matrix (GLCM), a technique for extracting texture features. After determining the frequency of occurrence of pixel pairs with a particular gray level at a given distance and orientation, GLCM generates number of statistical parameters. Contrast, correlation, energy, and other homogeneity are some crucial parameters that are frequently employed. Here, $(i,)$ is a probability element in the GLCM matrix, wherea μ and σ Each displays the standard deviation and mean value. Numerous images processing applications, including pattern recognition, object segmentation, and classification, have found success with this technique. According to recent studies, [8] has demonstrated a high degree of accuracy in detecting egg fertility non-destructively during the incubation period by using GLCM parameters as input to the Support Vector Machine (SVM) algorithm. The GLCM formula is as below.

Gray Level Co-Occurance Matrix (GLCM) Texture Feature Extraction

$$Energy = \sum_{i,j} [p(i,j)]^2 \tag{2}$$

Where:
$\mathcal{P}(i,j)$ = probability value of GLCM at position row-i and column-j
$\mathcal{P}^2$ = Number of grey levels

$$kor = \sum \frac{(i-\mu_j)(j-\mu_j)\mathcal{P}(i,j)}{\sigma_i \times \sigma_j} \tag{3}$$

Where:
$(i-\mu_j)$ = The average value of the intensity of the grey image of a distribution
$(j-\mu_j)$ = Pixel value distribution in the image
(i,j) = row value to - i in the column to – j

$$Hom = \sum \frac{1}{i + |i+j|} (i,j) \tag{4}$$

Where:
(i,j) = the probability value or normalised frequency of the GLCM matrix in row i and column position$j$.
$|i+j|$ = the amount of greyness in the image

$$Con = \sum\sum (i,j)^2 . \mathcal{P}(i,j) \tag{5}$$

Where:
(i,j) = row value to -i and column to - j on GLCM
i,j = The grey amount of the image

Energy in the context of GLCM is one of the features extracted to describe the texture of the image. In the feature energy is measured as the sum of squares of the elements in the normalised GLCM. Correlation measures the tendency of pairs of pixels with the same intensity to appear close to each other. Homogeneity in the context of Gray-Level Co-Occurrence Matrix (GLCM) is a measure that describes the extent to which the intensity of pixels in the image is close to the same intensity value with degrees of 0°, 45°, 90°, 135°. The higher the homogeneity value, the more uniform and consistent the texture of the image. Finally, contrast in the context of Gray-Level Co-Occurrence Matrix (GLCM) is a measure that describes the degree of difference between the intensities of adjacent pixels in the image.

Data Non-Image

In addition to image data, this study also incorporates non-image attributes derived from both external and internal egg characteristics, such as shell color and cleanliness, texture, egg weight, odor, the condition of the albumen and yolk, and the presence of blood spots. To ensure that all these variables can be optimally processed by the Support Vector Machine (SVM) algorithm, several preprocessing steps were carried out as follows:

Nominal Categorical Variable Encoding

Coding was used to process nominal categorical variables, including the presence of blood spots (present, absent), cleanliness (clean, dirty), odor (fishy, rotten), and also shell color (brown, cream, white). This approach was chosen because it gives the machine learning algorithm a more accurate representation of the data by giving equal weight to each category without regard to their order. One of the best method for managing nominal category data to maintain class neutrality is one-hot coding, according to a recent study [23].

Ordinal Variabel Encoding

Label encoding was used to process ordinal variables, such as albumen consistency (thin, medium, thick) and shell texture (rough, smooth), while accounting for the logical order between categories. This method preserves information about the quality level, which is in line with standard practice in ordinal data processing. According to medical research, using order-preserving encoding techniques to process ordinal data can increase model accuracy and data validity [24].

Normalitation of Numerical Variabel

Continuous numerical variables, such as egg weight (grams), were normalized using Min-Max scaling into the [0–1] range. Normalization is important because distance-based algorithms like SVM are highly sensitive to differences in feature scales. With normalization, the contribution of each variable becomes more balanced [25].

After normalization, all non-image features were appended together with the image texture features extracted from GLCM: energy, contrast, homogeneity and correlation). The SV M model with linear and polynomial kernel were used as input of merged files. By doing so, the fusion of image and non-image information was performed in a proportional manner, and the classifier is expected to achieve better classification.

2.2 Support Vector Machine

This section describes the architecture of the linear kernel-based Support Vector Machine (SVM) model. Support Vector Machine (SVM) was chosen as the classification method for this study due to its proven effectiveness in handling high-dimensional and complex datasets. SVM is well-suited for cases with a large number of features relative to the number of samples, as it searches for optimal hyperplane that maximizes the interclass margin, thus improving generalization performance [22]. Previous studies have also examined the effectiveness of SVM in image-based classification but not of eggs, such as the classification of mangrove genera using UAV multispectral data [27], which further supports its application in egg quality evaluation. Creating a classification model for evaluating egg quality based on internal and external factors is the goal. Meanwhile, the polynomial kernel uses polynomial functions to handle nonlinear patterns, the linear kernel is used for data that is linearly separable. Additionally, the definitions of the linear and polynomial kernels as well as the basic formula of SVM are provided as the theoretical foundation for the classification procedure. Image-based features $X_5 - X_{10}$ were extracted using the Gray-Level Co-occurrence Matrix (GLCM) method and non-image features $X_1 X_4$ and $X_6 - X_9$ (such as shell color, weight, cleanliness, and texture) as input data for SVM.

$$\mathcal{S}(X) = sign\ (\mathcal{w}^{\mathcal{T}} X + a) \tag{6}$$

Where:
$\mathcal{S}(X)$ = The result of the decision function that determines the class of the input data X
X = Feature vectors from input data (e.g. image feature extraction results)
$\mathcal{w}$ = Weight vector obtained from the SVM training process
a = Bias or offset constant
$sign$ = A sign function that produces a value of +1 or -1 to determine the class (Grade A or Grade B)

Where, the parameters $\mathcal{w}$ and a are used in the kernel function during training, transformation of (X) is done through $\varphi(X)$, y which is a feature space mapping to improve data separation. Then the classification function becomes as below.

$$\mathcal{S}(X) = sign\ (\mathcal{w}^{\mathcal{T}} \varphi(X) + a) \tag{7}$$

Where:
$\varphi(X)$ = Mapping function
$\mathcal{w}$ = Weight vector in the mapping result space

2.3 Model Validation

In order to guarantee the credibility of the classification outcomes, the k-fold cross-validation was used in this study. The reason for employing this methodology is that it gives a more reliable estimate of model performance than using a single split of training and test data. In the implementation, the 750 samples were first partitioned into the same number of folds, through a random way. The eachfold was using by turns as testing set and also the remaining folds were training set. This was done iteratively for each fold as testing set. Finally, the data from each iteration, which is accuracy, precision, and recall, were averaging to result the total predicted ability of the model. In mathematical terms the mean metrics for cross-validation can be written as:

$$M_{cv} = \frac{1}{k} \sum_{i=1}^{k} M_i \tag{8}$$

Where:
M_{cv} = the average cross-validation metric (accuracy, precision, or recall).
(k) = the number of folds used in the validation.
(M_i) = the metric value for the $-i$ th fold

The use of cross-validation methods such as this is common in machine learning-based research because it can reduce bias due to data distribution that happens to be less representative [28]. Thus, the test results obtained in this study can be more reliable and provide a consistent picture of the performance of the classification model.

2.4 Evaluation Model

In this study to evaluate researchers using egg quality classification evaluation measurements. The confusion matrix is a testing method used for accuracy values based on the cumulative calculation of the number of correct classifications divided by the number of correct and incorrect classifications or comparing the results of system predictions with the actual data available, [29] confusion matrix includes True Positive (TP), True Negative (TN), and False Positive (FP) and the last False Negative (FN). Below are important formulas related to the confusion matrix, namely formulas for calculating precision, recall, and accuracy. These three metrics are very useful for assessing the performance of a classification model thoroughly and objectively.

$$\mathcal{P}rcn = \frac{TP}{TP+FP} \tag{9}$$

$$\mathcal{R}cal = \frac{TP}{TP + FN} \tag{10}$$

$$\mathcal{A}ccry = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

As a complement, the evaluation results are not only presented in tabular form but also visualized using a confusion matrix, so that the distribution of correct and incorrect predictions can be seen more clearly and facilitate the assessment of model performance. In addition to the tabular presentation, the results are visualized using a confusion matrix to clearly show the distribution of correct and incorrect predictions, thus facilitating the assessment of model performance.

3. RESULT AND ANALYSIS

In this section, classification results are presented using the Support Vector Machine (SVM) method with two kerne. l types: linear and polynomial. The classification process was performed on test data that had undergone thorough preprocessing and also preparation. Next, SVM testing was performed on both kernels to evaluate and compare their performance in classifying the data. The results show that accuracy values differ between kernels, reflecting the characteristics and capabilities of each kernel in handling the classification problem. Comparing different kernels allows for an assessment of which kernel is more suitable for capturing patterns in the dataset, especially when combining internal and external egg features.

Table 2. The results of egg quality support vector machine with linear and polynomial kernel

SVM	Data Image			Data Non-Image			Data Image dan Non-Image		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Linear	58	68	70	80	90	85	85	90	85
Poly (Deg = 2)	60	66	71	70	65	78	73	81	81
Poly (Degr =3)	62	69	75	71	72	81	78	83	81
Poly (Deg=4)	65	74	80	73	79	84	81	87	72
Poly (Deg=5)	70	79	82	73	77	87	84	94	88
Poly (Deg= 6)	70	56	62	83	81	90	86	91	87
Poly (Deg= 7)	59	63	71	73	80	87	83	85	81

The above table demonstrates the results of the two methods on the differences between their average results while using the two kinds of data mentioned above (i.e image based (IB) and non-image based data (NIB data)), where we can observe that the polynomial kernel enables to reach a little higher accuracy compared with the linear kernel (only about 1% higher) with both kind of the used data. The best performing results are in bold. In addition, the comparison of the performance achieved and the table comparing SVM with linear/polynomial kernel with test data utilizing imagebased inputs are given. The experiment result shows that testing by the only image data is not good enough and the accuracy is even lower than the combination of the image and non-image data. One reason for such a result is that only 4 GLCM parameters are considered in this paper, by increasing the GLCM parameters to more than 4, a better result can be ensured.

With respect to the performance of image data (linear kernel accuracy), poor quality of the tested images might be one of the reasons for the low performance. Higher quality and higher pixel images should result in higher accuracy, the reason for it may be that only 4 parameters of GLCM were used, images were taken at 30cm distance and light condition was well-optimized.

Accuracy is much higher when both non-image and image data are combined, indicating the fusion with non-image would have positive impact on model performance. This study utilises both types of data, therefore, we performed additional testing on the non-image data set, using the same approach, but with different kernels. The visualization of the best accuracies and the accuracy of the testing results in either the polynomial or the linear approach are compared below to complement the result analysis in the literature. Figure 10 shows performance, recall and precision results.

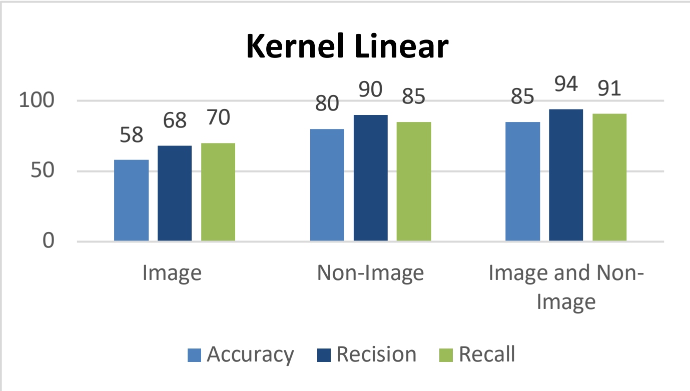

Figure 10. Comparison of test results between data and using a linear kernel

The image above shows that image-based data alone still does not have a high level of accuracy, but Non-image data has quite high results, while the combined data between Image and Non-image produces a higher level of accuracy than those that only use images and Non-Image only, the accuracy value of the combined data between Image and Non-image is 85%, recall 91% precision is 95% while from Non-image data of 80% the accuracy value is 90% precision, 85% for recall while for Image data the accuracy value is only 58% precision 68% the recall gets a value of 70%.

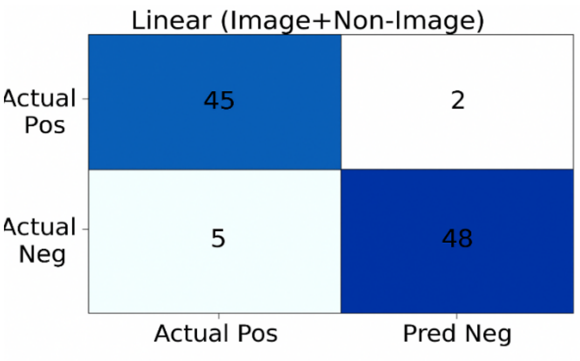

Figure 11. Confusion Matrix of SVM classification results with a Linear kernel using a combination of Image and Non-Image data

The linear kernel confusion matrix for combined image and non-image test data shows 45 positive and 48 negative samples correctly classified, with 7 errors (figure 11). These misclassifications suggest that linear kernel may struggle with non-linear patterns in the data, particularly for borderline cases.

In the use of the Polynomial kernel, it is the same as the previous test, namely using an image database, Non-Image and combining data between the two, only the difference in the kernel used in the test, in the test using the polynomial kernel produces the greatest accuracy in the use of image-based and Non-Image data or a combination of 87% with a process value of 91% and a recall of 87%, while for the use of image-based data alone, the smallest accuracy results are 70% and 83% for Non-Image data, for a comparison of data from the polynomial kernel can be seen in Figure 12 below.

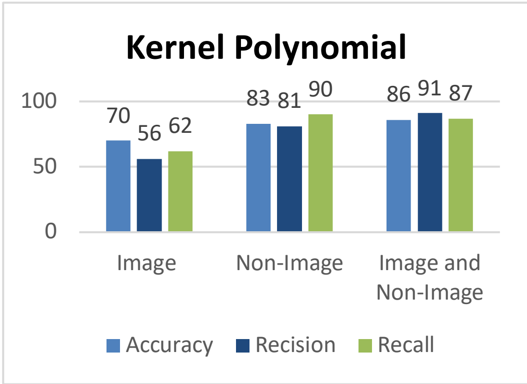

Figure 12. Comparison of test results between data and using the Polynomial kernel

A comparison of the test results with the polynomial kernel SVM is shown in Figure 12 above. Figure 13 below shows the confusion matrix with a 6th-degree polynomial kernel on the combined dataset (image features + non-image features). The model successfully identified 43 positive samples and 46 negative samples, with 11 errors. This finding confirms that the use of polynomial kernels with higher complexity, especially 6th-degree, provides significant performance improvements when combining image- and non-image-derived features, thus strengthening the value of multi-feature integration in egg quality classification. Other research findings also support this finding, for example, a 2nd-degree polynomial kernel was shown to outperform a linear kernel in assessing landslide susceptibility [27].

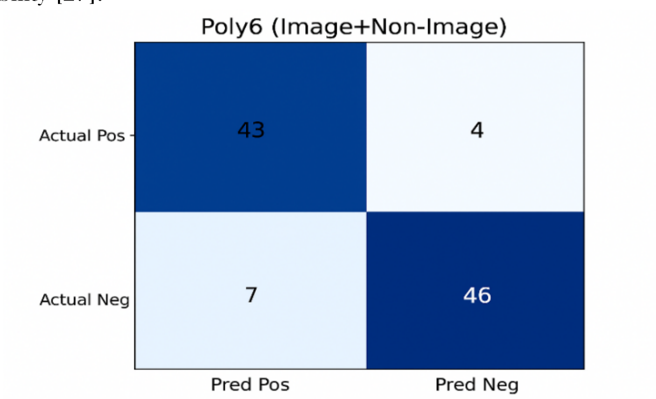

Figure 13. Confusion Matrix of SVM classification results with a Polynomial kernel degree 6 using a combination of Image and Non-Image data.

This study tested image (Image) and non-image (non-Image) data separately, as well as their combination, to examine effects on classification performance. Each SVM kernel was evaluated on these datasets to determine the most effective configuration. The heterogeneity between image and non-image features affects the geometry of the decision boundary: image features often capture complex, high-dimensional pattern, while non-image features are lower dimensional and more linear. Combining them creates a richer feature space, allowing the SVM (particularly with a polynomial kernel) to form more flexible, non-linear boundaries that can better separate classes. This analysis helps determine which kernel and data type combination maximizes classification accuracy.

The experimental results demonstrate that the polynomial kernel with degree 6 achieved the best overall performance compared to both the linear kernel and polynomial kernel with degree 7. Quantitatively, the degree 6 polynomial kernel yielded 86% accuracy, 91% precision, and 87% recall. In contrast, increasing the polynomial to degree 7 resulted in a decrease in performance, with accuracy dropping to 84%, precision to 88%, and recall to 83%. Mathematically, a polynomial kernel of degree 6 provides an optimal level of model complexity: it is flexible enough to capture the non-linear relationship between image and non-image features while avoiding excessive curvature that can lead the overfitting. Higher-degree kernels (degree 7) introduce more complex decision boundaries, which may fit the training data too closely and reduce generalization on unseen samples. Conversely, lower-degree kernels (including linear) are too simple to capture the intricate interactions among heterogeneous features. Therefore, degree 6 strikes a balance between flexibility and generalization, producing superior classification results.

These findings align with previous research on SVM-based egg quality classification. For instance, applied [8] applied SVM with GLCM parameters to detect egg fertility, achieving up to 84.57% accuracy on internal egg data. Other studies also suggest that combining image and non-image features improve classification compared to using image features alone [30], [31]. However, the extract performance depends on data characteristics, preprocessing, and kernel choice. This study confirms that systematically exploring polynomial degree and applying feature fusion between image and non-image data is crucial for achieving optimal and accurate egg quality classification.

4. CONCLUSION

In the testing phase, the research data was split into three groups: combined Image and Non-Image data, Image data only, and non-Image data only. SVM with a polynomial kernel of varying degrees (2-7) was evaluated. Using degree 2, the accuracy was low, prompting testing up to degree 6. The degree 6 polynomial kernel achieved the highest performance on the combined dataset, with 86% accuracy, 91% precision, and 87% recall, compared to 84% accuracy, 88% precision, and 83% recall with degree 7. Image-only data produced lower accuracy (around 78%), while non-image-only data performed better (approximately 82%), confirming that feature fusion significantly improves classification.

In practical terms, these result offer valuable insight for the poultry industry, particularly in the development of automated egg quality grading systems. The implementation of an SVM-based approach with an optimal polynomial kernel (degree 6) can enhance the accuracy and consistency of egg quality assessments, reduce human error, improve human efficiency in sorting and packaging processes. Consequently, this method has the potential to support data driven quality control and contibute to improving product standardization in modern poultry production.

For future work, researchers recommend exploring more GLCM parameters for image-based features and increasing dataset detail and size to further improve accuracy. Additionally, investigating other kernel types or advanced feature fusion techniques could yield even better performance and robustness in egg quality classification.

5. REFERENCES

[1] Kementan, "Buku Outlook Komoditas Peternakan Daging Domba," 2023.

[2] N. U. Idayanti, Darmawati S, "105435-ID-perbedaan-variasi-lama-simpan-telur-ayam.pdf," *Jurnal Kesehatan*, vol. 2, no. 1. pp. 19–26, 2009.

[3] G. Kralik, Z. Kralik, M. Grčević, and D. Hanžek, "Quality of Chicken Meat," in *Animal Husbandry and Nutrition*, 2018. doi: 10.5772/intechopen.72865.

[4] F. G. Silversides and T. A. Scott, "Effect of Storage and Layer Age on Quality of Eggs.pdf," pp. 1240–1245, 2001.

[5] R. S. Vargas Cruz, L. C. Ruiz Salvador, and M. C. Navas Lema, "Combinando la ovoscopía manual y automática: una soluciónsegura y socialmente responsable," *Enfoque UTE*, no. 2, pp. 70–76, 2018.

[6] T. Eshetie, A. Tesema, D. Bekele, and T. Assefa, "Effect of Genotype, Environment, and Their Interactions on Internal as well as External Egg Quality Traits of Different Chicken Genotypes and Poultry Species in Ethiopia and Beyond: A Review," *Mathews J. Nutr. Diet.*, vol. 7, no. 1, pp. 1–15, 2024, doi: 10.30654/mjnd.10031.

[7] H. E. Samli, A. Agma, and N. Senkoylu, "Effects of storage time and temperature on egg quality in old laying hens," *J. Appl. Poult. Res.*, vol. 14, no. 3, pp. 548–553, 2005, doi: 10.1093/japr/14.3.548.

[8] S. Saifullah and R. Drezewski, "Non-Destructive Egg Fertility Detection in Incubation Using SVM Classifier Based on GLCM Parameters," *Procedia Comput. Sci.*, vol. 207, no. Kes, pp. 3248–3257, 2022, doi: 10.1016/j.procs.2022.09.383.

[9] W. Zhang, L. Pan, S. Tu, G. Zhan, and K. Tu, "Non-destructive internal quality assessment of eggs using a synthesis of hyperspectral imaging and multivariate analysis," *J. Food Eng.*, vol. 157, pp. 41–48, 2015, doi: 10.1016/j.jfoodeng.2015.02.013.

[10] J. H. So, S. Y. Joe, S. H. Hwang, S. J. Hong, and S. H. Lee, "Current advances in detection of abnormal egg: a review," *J. Anim. Sci. Technol.*, vol. 64, no. 5, pp. 813–829, 2022, doi: 10.5187/jast.2022.e56.

[11] J. U. Raji, A. O.; Aliyu, J.; Igwebuike, "Effect of storage methods and time on egg quality traits of laying hens in a hot dry climate," *ARPN J. Agric. Biol. Sci.*, vol. 4, no. 4, pp. 1–7, 2009.

[12] E. Altuntaş and A. Şekeroğlu, "Effect of egg shape index on mechanical properties of chicken eggs," *J. Food Eng.*, vol. 85, no. 4, pp. 606–612, 2008, doi: 10.1016/j.jfoodeng.2007.08.022.

[13] D. N. Kanasuah, K. Adomako, B. A. Hagan, and O. S. Olympio, "Influence of Feather Genotype, Storage Duration and Temperature on the External and Internal Qualities of Chicken Table Eggs," *Online J. Anim. Feed Res.*, vol. 15, no. 1, pp. 21–32, 2025, doi: 10.51227/ojafr.2025.4.

[14] A. Gacem *et al.*, "Valorization of eggshell waste as sustainable mechanical reinforcement in biodegradable bio-packaging materials," *J. Agric. Food Res.*, vol. 23, p. 102273, 2025, doi: 10.1016/j.jafr.2025.102273.

[15] T. Rho and B. Cho, "Non-Destructive Evaluation of Physicochemical Properties for Egg Freshness : A Review," *Agricultue*, vol. 14, no. 11, pp. 1–19, 2024, doi: 10.3390/agriculture14112049.

[16] Z. Gao, J. Zheng, and G. Xu, "Research Progress and Technological Application Prospects of Comprehensive Evaluation Methods for Egg Freshness," *Foods*, vol. 14, no. 9, pp. 1–19, 2025, doi: 10.3390/foods14091507.

[17] W. Zhou *et al.*, "Analysis of egg consumption trends: Gen Z egg consumption preferences and the classification of pink-shell eggs in China," *Poult. Sci.*, vol. 104, no. 11, p. 105727, 2025, doi: 10.1016/j.psj.2025.105727.

[18] K. Lokaewmanee, K. Yamauchi, T. Komori, and K. Saito, "Enhancement of egg yolk color by paprika combined with a probiotic," *J. Appl. Poult. Res.*, vol. 20, no. 1, pp. 90–94, 2011, doi: 10.3382/japr.2009-00140.

[19] Maimunah, "Deteksi Kebersihan Kerabang Telur Ayam Berdasarkan Pengolahan Ctra Digital," *J. Penelit. Ilmu Komputer, Syst. Embed. Log.*, vol. 3, no. 1, pp. 41–49, 2015.

[20] Y. D. Distya, Z. Ludfi, D. Sari, B. Cahya, and E. Putra, "' EGG-GRADING ' Mesin Klasifikasi Telur Ayam (Berat Telur dan Telur Rusak) Otomatis Berbasis Microcontroller," *Pros. Nas. Rekayasa Teknol. Ind. dan Inf. XIV Tahun 2019*, vol. 2019, no. November, pp. 380–385, 2019. https://journal.itny.ac.id/index.php/ReTII/article/view/1429

[21] 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.

[22] C. Kanan and G. W. Cottrell, "Color-to-grayscale: Does the method matter in image recognition?," *PLoS One*, vol. 7, no. 1, 2012, doi: 10.1371/journal.pone.0029740.

[23] F. Bolikulov, R. Nasimov, A. Rashidov, F. Akhmedov, and Y. I. Cho, "Effective Methods of Categorical Data Encoding for Artificial Intelligence Algorithms," *Mathematics*, vol. 12, no. 16, 2024, doi: 10.3390/math12162553.

[24] A. Seveso, A. Campagner, D. Ciucci, and F. Cabitza, "Ordinal labels in machine learning: A user-centered approach to improve data validity in medical settings," *BMC Med. Inform. Decis. Mak.*, vol. 20, no. Suppl 5, pp. 1–14, 2020, doi: 10.1186/s12911-020-01152-8.

[25] D. Anelli, P. Morano, F. Tajani, and M. R. Guarini, "The Interpretative Effects of Normalization Techniques on Complex Regression Modeling: An Application to Real Estate Values Using Machine Learning," *Inf.*, vol. 16, no. 6, pp. 1–28, 2025, doi: 10.3390/info16060486.
[26] V. Cortes, Corinna; vapnik, "In silico log p prediction for a large data set with support vector machines, radial basis neural networks and multiple linear regression," *Chem. Biol. Drug Des.*, vol. 74, no. 2, pp. 142–147, 1995, doi: 10.1111/j.1747-0285.2009.00840.x.
[27] Armanda, Syamsul Bahri Agus, and Jonson Lumban Gaol, "Classification and Distribution Of Mangrove Genus Using Multispectral Unmanned Aerial Vehicle (UAV) In The Waters Of Lancang Island, Kepulauan Seribu, Indonesia," *J. Geosci. Eng. Environ. Technol.*, vol. 9, no. 2, pp. 210–220, 2024, doi: 10.25299/jgeet.2024.9.2.17195.
[28] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Stat. Surv.*, vol. 4, pp. 40–79, 2010, doi: 10.1214/09-SS054.
[29] J. Kivimäki, J. Białek, W. Kuberski, and J. K. Nurminen, "Performance Estimation in Binary Classification Using Calibrated Confidence," no. Ml, 2025. https://doi.org/10.48550/arXiv.2505.05295
[30] X. Liu and C. Zhao, "Research on Image Feature Extraction Algorithm of the Egg and Egg White Protein Thermal Gelation Based on PCA/ICA," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–11, 2022, doi: 10.1155/2022/1266332.
[31] I. S. Isa, U. K. Yusof, and M. M. Zain, "Image Processing Approach for Grading IVF Blastocyst: A State-of-the-Art Review and Future Perspective of Deep Learning-Based Models," *Appl. Sci.*, vol. 13, no. 2, pp. 1–22, 2023, doi: 10.3390/app13021195.