Eggshell Crack Classification Using a Hybrid Texture-Color Descriptor with Machine Learning Methods

Aji Setiawan
Department of Information Technology, Faculty of Engineering, Darma Persada University, Jakarta, Indonesia
aji_setiawan@ft.unsada.ac.id (corresponding author)

Suzuki Syofian
Department of Information Technology, Faculty of Engineering, Darma Persada University, Jakarta, Indonesia
suzuki_syofian@ft.unsada.ac.id

Adam Arif Budiman
Department of Information Technology, Faculty of Engineering, Darma Persada University, Jakarta, Indonesia
adam_arif_budiman@ft.unsada.ac.id

Received: 11 October 2025 | Revised: 20 November 2025 and 2 December 2025 | Accepted: 3 December 2025

 DOI: https://doi.org/10.48084/etasr.15450

ABSTRACT

This study presents a lightweight and efficient method for classifying eggshell fractures as an alternative to deep learning-based approaches. The dataset was compiled from public and field sources, resulting in a balanced collection of 3,700 RGB images (1,850 fractured and 1,850 normal). The images were preprocessed by cropping, resizing to 256×256, and enhancement using CLAHE and GrabCut segmentation. Color (HSV), edge (Canny), and texture (GLCM and Gabor) features were combined into a hybrid descriptor and classified using Support Vector Machine (SVM), Random Forest (RF), XGBoost, and a stacking ensemble under 5-fold cross-validation. RF demonstrated the strongest robustness to nonlinear and texture-rich attributes, achieving the highest accuracy of 81.2%. The results confirm that the proposed hybrid descriptor offers a computationally lightweight, interpretable, and CPU-friendly alternative to CNN-based systems. Feature importance and PCA analysis further reveal how color and texture descriptors jointly contribute to decision-making, supporting its practicality for real-time industrial deployment.

Keywords-eggshell crack classification; GrabCut segmentation; GLCM–Gabor texture features; random forest classifier

I. INTRODUCTION

The global demand for eggs continues to increase due to their nutritional value, affordability, and contribution to food security across countries, including Indonesia [1-5]. However, maintaining egg quality remains a major challenge, as manual inspection methods are still prevalent in most agricultural processes [6]. These conventional inspection techniques are not only time-consuming but also subjective and prone to human error, allowing cracked eggs to enter the market and potentially transmit pathogens such as Salmonella, Escherichia coli, Mycoplasma synoviae, and Chlamydia psittaci [7-10]. Such microbial contamination risks emphasize the importance of implementing automated inspection systems to ensure product safety, quality consistency, and consumer trust.

In recent years, various studies have proposed machine vision and deep learning approaches for the detection of eggshell cracks. In [11], the Convolutional Encoder-Decoder Network (CedNet) was introduced for pixel-level crack segmentation and geometric measurement, but although it achieved high detection accuracy, it required extensive computational resources and a complex calibration process. In [12], an enhanced MobileNetv3-based model was presented for detecting preserved egg cracks, offering improved efficiency but still relying on large neural networks with millions of parameters. In [13, 14], traditional image processing methods

were explored, combining thresholding, edge detection, and segmentation, demonstrating fast processing speeds but limited adaptability under diverse lighting and texture conditions. More recently, in [15, 16], CNN-based architectures were utilized for automated egg classification and defect identification, achieving high accuracy but at the cost of increased computational load and dependency on GPU-based environments.

Despite these developments, some significant obstacles persist in the domain of automated eggshell examination. Many deep learning-based methods, despite attaining remarkable accuracy, require significant processing resources, limiting their use in low-power or real-time industrial settings. Moreover, the interpretability of these models is constrained as their decision-making processes generally operate as a "black box," complicating the justification of predictions in quality assurance workflows. Ultimately, model generalization remains a significant challenge, since numerous algorithms trained on static datasets exhibit subpar performance when faced with novel situations, including fluctuating lighting, reflections, or varying eggshell surface patterns.

These constraints underscore the necessity for a lightweight, interpretable, and data-efficient alternative to deep learning-based crack detection. This paper presents a hybrid texture-color descriptor combined with classical machine learning for the classification of eggshell cracks to solve this gap. The proposed approach integrates manually generated HSV color components with Gray Level Co-occurrence Matrix (GLCM) and Gabor texture descriptors, assessing them through ensemble classifiers such as Random Forest (RF) and XGBoost. This approach seeks to harmonize precision and computational efficiency, facilitating real-time execution on conventional CPU hardware while preserving model transparency and interpretability.

This study contributes to existing research by specifically addressing the practical limitations of previous CNN-based methods, which frequently demand substantial processing resources, lack interpretability, and exhibit sensitivity to fluctuations in datasets. The suggested hybrid descriptor offers competitive accuracy in a lightweight format without necessitating GPU acceleration. Moreover, feature importance and Principal Component Analysis (PCA) theoretically elucidate the interaction between color and texture cues, with color representing global illumination and reflection, while texture descriptors encapsulate localized fracture patterns. A comparative assessment of SVM, RF, XGBoost, and stacking ensembles illustrates the superiority of tree-based models in this mixed-distribution feature space, providing enhanced efficiency, interpretability, and deployability for practical egg inspection systems.

II. METHOD

A. Data Acquisition

The dataset used in this study was obtained from a primary field acquisition at a Serayu Farm and supplemented with a public dataset. The field data acquisition process involved the capture of egg images under controlled illumination conditions to reduce glare and shadows. In addition, publicly available data from [17] were included. The original field data is not available for public release due to ownership and privacy agreements. However, representative samples and processed datasets are available from the corresponding author upon reasonable request.

The dataset is divided into two primary categories: Normal and Cracked eggs. The total number of egg images is 3700, with 1850 images in each category. This balanced composition ensures a reasonable performance evaluation and unbiased classification for both categories. A controlled imaging apparatus, which comprised a webcam and conveyor system, was employed to compile the dataset from field data. Image processing and categorization of egg defects were conducted through multiple methodical processes, as seen in Figure 2.

TABLE I. DATASET SAMPLE

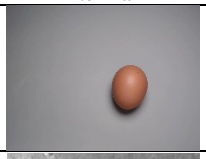
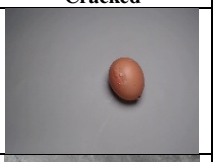
Dataset	Normal	Cracked
[17]		
Own		

Fig. 1. Field acquisition setup showing the controlled illumination environment and conveyor-based imaging process used for capturing egg images.

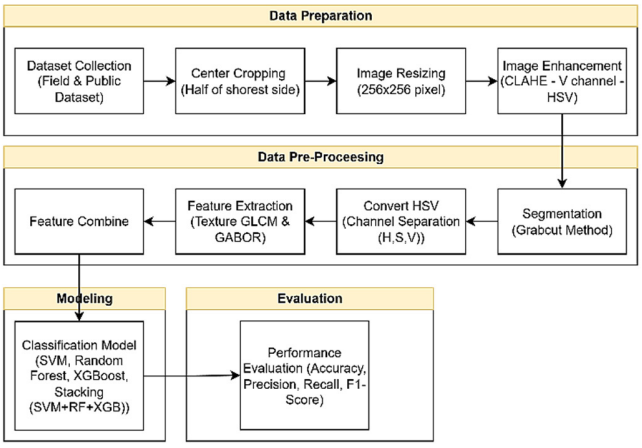


Fig. 2. Method.

B. Data Preparation

1) Center Image Cropping

Each image is cropped to the middle to half of the lesser dimension between its height and width. This cropping intends to focus on the egg while minimizing extraneous background, hence decreasing computing demand and the likelihood of background noise infiltrating the feature extraction phase [18].

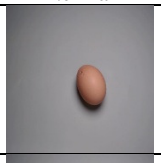
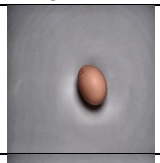
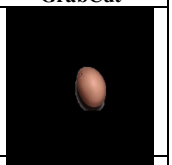
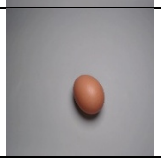
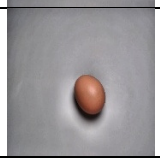
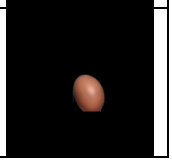
2) Image Resizing

After cropping, the image is resized to a consistent resolution of 256×256 pixels, down from its original dimensions of 640×640 pixels. The resizing phase is crucial to standardize the input size throughout the dataset, therefore enhancing the model training process in the subsequent stage [19].

3) Image Enhancing

The resized image undergoes processing via the CLAHE method applied to the Value (V) channel inside the HSV color space. Table II shows that CLAHE adaptively increases contrast in localized regions of the image, rendering details such as roughness and fissures more discernible while minimizing the prominence of noise [20]. This approach is appropriate for mitigating lighting discrepancies in various collected datasets [21-22] under varying environmental circumstances, including underwater scenarios [23].

TABLE II. IMAGE ENHANCING

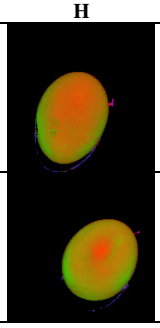
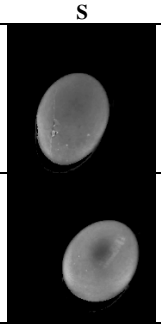
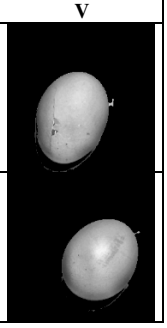
Class	Normal	CLAHE	GrabCut
Crack			
Normal			

C. Data Preprocessing

The contrast-enhanced image is subsequently processed via the GrabCut method. This approach employs the initial bounding box to ascertain the item's position, subsequently distinguishing the foreground (egg) from the background via iterative modeling of background and object colors. The preliminary output from GrabCut is a mask, further refined into a binary mask with a value of 1 representing the foreground and 0 denoting the background. The segmentation procedure seeks to ensure that the retrieved features originate solely from the eggs, excluding the background.

The segmentation outcomes are transformed back into the HSV color space. The division into three channels (Hue, Saturation, and Value) facilitates independent analysis or processing of each channel. The H channel denotes the egg's base color, the S channel indicates the level of color saturation, and the V channel contains brightness information. The script can analyze and visualize the attributes of each color component in cracked and normal eggs independently by isolating the channels, as demonstrated in Table III.

TABLE III. HSV COLOR FEATURE EXTRACTION

Class	H	S	V
Crack			
Normal			

This study employed two methods for texture feature extraction: GLCM and the Gabor filter. GLCM analyzes second-order statistical texture patterns by calculating the co-occurrence relationships of pixel intensities at specified distances and directions, yielding metrics such as contrast, homogeneity, energy, and correlation. The Gabor filter independently derives multi-orientation and multi-frequency texture responses that highlight directional fracture patterns on the eggshell. Together, these two descriptors offer complementary information regarding texture, as illustrated in Figure 3 (d–e).

(a) (b) (c) (d) (e)

Fig. 3. Sequential preprocessing stages for a cracked egg image: (a) original RGB image, (b) CLAHE-enhanced image, (c) GrabCut segmented egg, (d) local GLCM contrast map, and (e) Gabor filter magnitude response illustrating textural extraction.

D. Modeling

All retrieved image features were subsequently utilized as input for the machine learning model. Four classification algorithms were evaluated: Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and a Stacking Classifier that integrates SVM, RF, and XGBoost with a Logistic Regression (LR) meta-learner. The efficacy of each model was assessed using accuracy, precision, recall, and F1-score, supplemented by a confusion matrix analysis to elucidate the distribution of predictions across both classes.

Each model was trained using the hybrid feature set combining HSV color components, GLCM texture attributes, and Gabor filter responses. Before training, the dataset was divided using stratified 5-fold cross-validation to ensure balanced representation of cracked and normal eggs in every fold. Hyperparameter tuning was performed using grid search within the same cross-validation framework. SVM tuned parameters included the penalty parameter C and the RBF

kernel parameter γ. RF tuning covered the number of trees and the maximum tree depth. XGBoost searched the space that included the number of estimators, learning rate, and maximum depth. The stacking ensemble employed SVM, RF, and XGB as base learners, with LR serving as the meta-classifier. All models were executed on standard CPU hardware without GPU acceleration to reflect real-world deployment constraints in small- and medium-scale egg processing environments.

E. Evaluation

The confusion matrix concept was used to determine whether the constructed model can effectively learn to classify natural and unnatural objects, helping to determine the F1-score, recall, accuracy, and precision metrics for model assessment. True Positive (TP) is the number of cracked eggs classified as Cracked, False Positive (FP) is the number of normal eggs classified as Cracked, False Negative (FN) is the number of cracked eggs classified as Normal, and True Negative (TN) is the number of normal eggs classified as Normal.

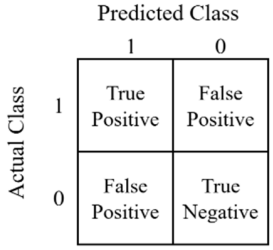


Fig. 4. The confusion matrix structure used to derive accuracy, precision, recall, and F1-score for Cracked and Normal egg classification.

III. RESULTS AND DISCUSSION

The primary aim of faulty egg classification is to facilitate real-time object counting systems in both natural and artificial environments, encompassing cracked eggs, normal eggs, and eggs exhibiting various surface imperfections. This study aimed to evaluate four machine learning classifiers using features extracted from HSV color components, Canny edges, GLCM, and Gabor descriptors. The classification model testing technique involved partitioning the dataset into training and test subsets.

The preliminary phase entailed enhancing image quality by CLAHE to rectify the intensity distribution on the egg's surface. This technique mitigates over-amplification in regions of high contrast by employing a clip limit of 2.0 and a grid size of (8, 8). Subsequently, segmentation is executed with GrabCut, which isolates the egg object from the background by modeling the distribution of foreground and background pixels through a Gaussian Mixture Model (GMM). This delineation is crucial to ensure that the feature extraction operation is conducted solely on the pertinent object region. The segmentation results are converted to the HSV color space, and the S (Saturation) channel is extracted. The S channel is selected because of its stability against variations in illumination while maintaining relevant color information. The average RGB and HSV values are selected to ensure that the color depth corresponds to the class values in Table IV.

TABLE IV. CHARACTERISTIC FEATURES' MEAN RGB AND HSV

Class	R	G	B	H	S	V
Normal	26.20	15.54	12.61	2.20	20.74	26.20
Crack	33.45	22.21	18.58	1.77	21.13	33.45

Image processing for texture feature extraction utilizing GLCM on segmented images commences with the *crop_center()* function, which crops the image's center based on a specified percentage. Subsequently, *preprocess_image()* resizes the cropped image to 256×256 pixels, converts it to grayscale, and enhances local contrast through CLAHE.

Subsequently, the textural features are derived from the image utilizing Gabor filters, which are frequently employed to identify particular patterns, orientations, and frequencies within images. The initial stage involves converting the color image to grayscale, followed by the creation of four Gabor kernels that vary in the sigma (scale) and theta (orientation) parameters, specifically combining tiny scale 1 with large scale 25, and low orientation 1 with high orientation 25 degrees. These kernels operate as sinusoidal wave filters modulated by Gaussian functions, attuned to certain pattern orientations and dimensions. Subsequently, each kernel is applied to the image, resulting in four filtered images, each emphasizing distinct texture aspects based on the Gabor parameters. The data are ultimately shown in grayscale, accompanied by a title that elucidates the filter scale and orientation, thus enhancing the visual examination of the Gabor response. The method is highly effective for pattern identification or object classification based on distinct textural attributes, such as cracks in eggshells, and the results are presented in a visually clear manner, enhancing their interpretability (Figure 5).

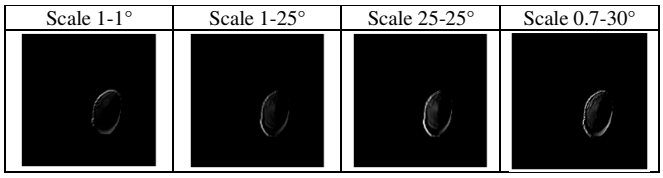


Fig. 5. Sample Gabor filter responses illustrating multi-scale and multi-orientation texture patterns extracted from the eggshell surface.

The CLAHE approach effectively enhanced the contrast of the egg images, making surface features, fissures, and color variations more discernible. The segmentation phase employing the GrabCut algorithm effectively isolated the eggs from the background with adequate precision; however, the efficacy of the segmentation was significantly dependent on the sharpness of the object edges and the contrast with the background. The segmentation results are crucial since they reduce interference from non-object regions in the subsequent analysis phase. Moreover, the utilization of the Gabor filter yields a unique texture representation, allowing for the visualization of the patterns and structures of the egg surface through filter magnitude responses across different orientations and scales.

The original images, segmentation outcomes, and texture extraction results clearly demonstrate the effectiveness of the

preprocessing phase on the quality of the input data for the classification stage. The amalgamation of CLAHE, GrabCut, and Gabor filters can provide more informative images, therefore improving the accuracy in detecting egg cracks during the classification phase using machine learning models, as shown in Figure 6.

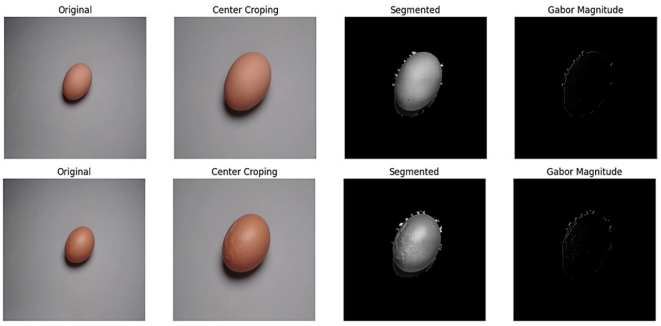


Fig. 6. Preprocessing outputs, including original image, cropped region, segmentation results, and Gabor texture responses.

The feature extraction process commences with color extraction via HSV, proceeds to Canny edge detection, and concludes with texture extraction via GLCM and Gabor, acquiring the image attributes as detailed in Tables V and VI.

TABLE V. FEATURE EXTRACTION GLCM

IMG	Contrast				Correlation			
	0°	45°	90°	135°	0°	45°	90°	135°
A	70	99	37	79	0.993	0.991	0.996	0.993
B	55	98	84	101	0.984	0.972	0.976	0.971

Energy				Homogeneity				Class
0°	45°	90°	135°	0°	45°	90°	135°	
0.664	0.663	0.665	0.663	0.852	0.819	0.837	0.811	Normal
0.852	0.850	0.851	0.850	0.887	0.877	0.884	0.879	Crack

TABLE VI. FEATURE EXTRACTION GABOR

Img	Amplitude		Phase		Energy	Entropy	Class
	Mean	Std dev	Mean	Std dev			
A	-109	0.102	-0.675	126	-106	109	Normal
B	0.79	-0.270	-0.968	-176	0.723	-0.775	Crack

TABLE VII. MODEL HYPERPARAMETERS

Model	Hyperparameter	Set
SVM	C	1,10,100,500
	Gamma	1,0.1,0.01,0.001
	kernel	Rbf
RF	n_estimator	100,200,300
	max_depth	None, 10,20,30
	min_samples_split	2,5,10
XGBoost	n_estimator	100,200,300
	max_depth	0.01,0.1,0.2
	min_samples_split	3,5,7
Stacking model (SVM+RF+XGB)	voting	Soft
	weights	[1,3,1] (svm, rbf,xgb)
	estimators	[('svm', best_svm), ('rf', best_rf), ('xgb', best_xgb)]

The accuracy values derived from the testing set may slightly differ from the true prediction accuracy, depending on the number of observations utilized for training and testing or the degree of bias inherent in the model. The selected models for each network dataset underwent cross-validation following the adjustment of model complexity and hyperparameters.

Table VII elucidates the hyperparameter search in conjunction with stratified 5-fold cross-validation (primary metric: accuracy), followed by the refitting of the selected models on the complete training subset before evaluation on the reserved test set. The grid for XGBoost encompassed $n\_estimators \in \{100,200,300\}$, $learning\_rate \in \{0.01, 0.1, 0.2\}$, and $max\_depth \in \{3, 5, 7\}$ (27 configurations), optimizing boosting capacity, shrinkage, and interaction depth to address bias variance trade-off. RF examined $n\_estimators \in \{100, 200, 300\}$, $max\_depth \in \{None, 10,20,30\}$, and $min\_samples\_split \in \{2,5,10\}$ (36 configurations), which dictate ensemble size, representational capacity, and node-level regularization, respectively. For the RBF-kernel SVM, the margin-smoothness trade-off was explored with C values of $\{1, 10, 100, 500\}$ and gamma values of $\{1, 0.1, 0.01, 0.001\}$ (16 configurations). $Probability = True$ was activated to generate calibrated probabilities necessary for soft voting.

The search comprised a total of 27×5+36×5+16×5=395 fits. The final ensemble integrated the optimal SVM, RF, and XGBoost models using soft voting with weights [1, 3, 1], with the RF model's superior validation performance being a testament to the success of the model selection process. This model also maintained the complementary decision boundaries of the margin-based and gradient-boosting algorithms. Random seeds were established to ensure reproducibility, and detailed logging maintained a verifiable record of the search.

Cross-validation is essential to generate an impartial model and facilitate the calculation of prediction errors and accuracies that more closely represent the true value [24]. *K*-fold is a widely utilized cross-validation method that employs distinct test data for each iteration, ensuring a rigorous assessment of the classification model's success rate, with as many *K* values as required to replicate the experiment [25].

The evaluation results in Table VIII demonstrate a substantial performance disparity between techniques. RF exhibits the highest performance, attaining an accuracy of 81.2%, an average precision of 0.812, an average recall of 0.812, and an average F1-score of 0.810. This indicates that RF more efficiently captures texture and color patterns derived from GLCM and Gabor feature extraction compared to other models, mostly due to its capacity to handle non-linear connections and leverage voting from numerous decision trees. Figure 7 presents a visual position of the models.

The XGBoost model achieved a second-place ranking with an accuracy of 71.0%, while maintaining balanced scores across other metrics (precision, recall, and F1-score were all 0.710). This result suggests that the gradient boosting method is proficient in data modeling, although it remains inferior to RF, likely due to its susceptibility to noise and the need for further parameter tuning.

TABLE VIII. CLASSIFICATION REPORT

Model	Accuracy	Precision (avg)	Recall (avg)	F1-score (avg)
SVM	0.638	0.645	0.638	0.616
RF	0.812	0.812	0.812	0.810
XGBoost	0.710	0.710	0.710	0.710
Stacking model (SVM+RF+XGB)	0.783	0.782	0.783	0.782

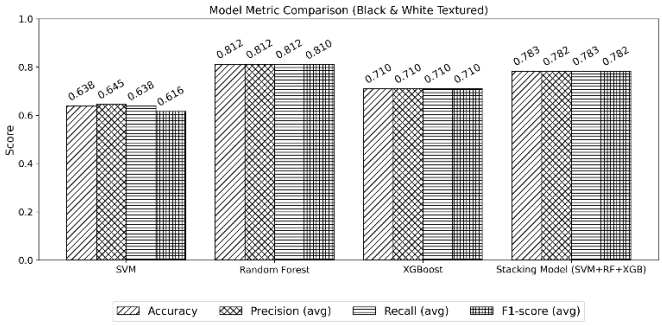


Fig. 7. Comparative performance of SVM, RF, XGBoost, and Stacking Classifier across accuracy, precision, recall, and F1-score.

The Stacking Classifier, integrating SVM, RF, and XGBoost with an LR meta-learner, attained an accuracy of 78.3%, with similar precision, recall, and F1-score (0.782–0.783). Despite its performance being inferior to that of pure RF, the stacking model demonstrates advantages over SVM and XGBoost individually, highlighting the merits of model integration, although the contribution weight of the superior model (RF) does not entirely govern the final outcomes. The SVM model showed potential for improvement, achieving an accuracy of 63.8%, an average precision of 0.645, an average recall of 0.638, and an F1-score of 0.616. The suboptimal performance of SVM is likely attributable to the intricate nature of the data distribution, which is not entirely linearly separable, rendering the employed kernel insufficient for effectively distinguishing the two classes. Further model refinement could potentially overcome these challenges.

To assess whether the proposed hybrid descriptor truly improves classification performance, several baseline feature settings were evaluated using the RF classifier. As shown in Table IX, models trained using only raw RGB features or single-feature extractors (HSV-only, GLCM-only, or Gabor-only) achieved substantially lower accuracy than the hybrid feature set. The hybrid descriptor improved accuracy by 13–27 percentage points relative to single-feature baselines, confirming that the integration of color and texture information provides complementary discriminatory power.

TABLE IX. BASELINE VS HYBRID FEATURE PERFORMANCE (RF)

Feature Set	Accuracy	Feature Set	Accuracy
Raw RGB only	0.580	HSV + GLCM	0.742
HSV only	0.642	HSV + Gabor	0.708
GLCM only	0.671	Hybrid (HSV + GLCM + Gabor)	0.812
Gabor only	0.623		

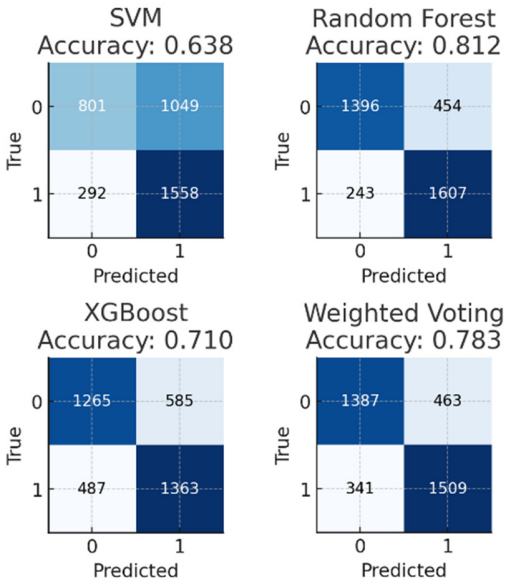


Fig. 8. Confusion matrices.

The proposed model is regarded as lightweight due to its minimal computational requirements relative to alternative deep learning methods. The RF classifier requires approximately 12.4 MB of storage space and executes end-to-end inference in approximately 45 ms per image on a standard Intel i7 CPU (16 GB RAM), with 5-fold cross-validation concluded in less than three minutes without GPU acceleration. For illustrative purposes, compact CNNs such as MobileNetV2 and EfficientNet-B0 require between 14 and 29 MB of memory and a processing time ranging from 65 to 110 ms per image within an identical CPU environment. These distinctions affirm that the suggested hybrid texture color descriptor integrated with RF delivers a markedly more effective and cost-efficient solution, suitable for implementation within small- to medium-scale egg inspection frameworks. The findings further demonstrate that RF exhibits the highest performance due to the texture features derived from GLCM and Gabor, with subsequent performance levels observed in the stacking classifier, XGBoost, and SVM.

PCA was implemented to conduct feature importance analysis and dimensionality reduction to bolster the theoretical justification for the interaction between color and texture features in the proposed method. The global spectral distribution on the eggshell surface is conceptually represented by color features (HSV and RGB), which include brightness, saturation, and color differences that can indicate cracks or variations in reflectivity. On the other hand, texture features (GLCM and Gabor) capture micropatterns and local variations in the surface structure, including the smoothness, orientation, and direction of fractures. Color features supply global context, while texture features facilitate the identification of local details that cannot be conveyed by color alone. These two feature types interact in a complementary manner. Color features dominate the model's performance, accounting for approximately 76% of the total importance weight, according to the feature importance analysis conducted using the RF algorithm. GLCM texture features and Gabor features follow at

±16% and ±8%, respectively. The features that make the most significant contributions are Smax (0.1133), Bmax (0.0741), Gmax (0.0739), and Phase Mean (0.0368). This suggests that color intensity and frequency information are crucial in the differentiation of fractured and intact eggs. PCA revealed that the initial five principal components accounted for 92.98% of the overall data variation, demonstrating that the suggested color-texture feature combination is both information-rich and computationally efficient.

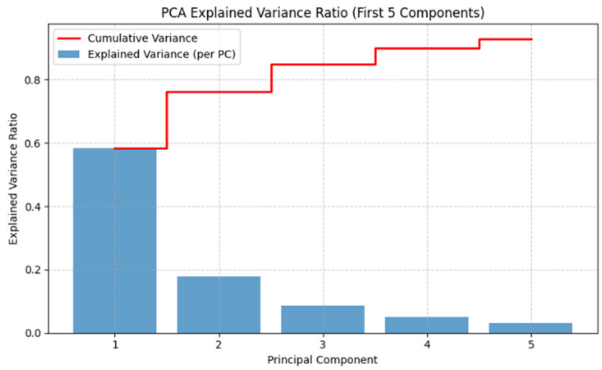


Fig. 9. Explained variance ratio for the first five principal components, demonstrating that 92.98% of total feature variance is captured in them.

These results indicate that the principal discriminatory information is contained in a limited number of orthogonal dimensions, illustrating the interplay between global information (color) and local features (texture). The hybrid integration of color and texture descriptors enhances crack classification accuracy and establishes a theoretical foundation for merging these two feature domains, thus improving model robustness and generalization to variations in lighting, reflectance, and eggshell surface characteristics.

IV. CONCLUSIONS

This study developed a methodical framework for egg quality classification using digital images. During the data preparation phase, egg images were sourced from direct field collection and a public dataset. Subsequently, these images were center-cropped to emphasize the egg object, resized to a consistent dimension of 256×256 pixels, and enhanced in quality utilizing the CLAHE method applied to the Value channel of the HSV color space. The data preprocessing phase encompasses egg object segmentation via the GrabCut algorithm to distinguish foreground from background, conversion to HSV color space with the separation of H, S, and V channels, and the extraction of texture features utilizing GLCM and Gabor filters, which are subsequently amalgamated into the final feature representation.

The feature extraction outcomes were subsequently employed in the classification phase, utilizing four methods: SVM, RF, XGBoost, and Stacking Classifier. The performance evaluation results indicate that the RF model achieved superior performance, with an accuracy of 81.2%, an average precision of 0.812, an average recall of 0.812, and an average F1-score of 0.810. This illustrates RF's capacity to manage the intricacies of non-linear feature distributions and its robustness through the aggregation of numerous decision trees. XGBoost attained a second-place ranking with an accuracy of 71.0%, while the Stacking Classifier, which integrates SVM, RF, and XGBoost, achieved 78.3%. This performance is marginally worse than that of pure RF but remains superior to all individual models except RF. SVM exhibited the poorest performance, achieving an accuracy of 63.8%, which underscores its constraints in distinguishing data with distributions that are not entirely linearly separable.

This work demonstrates that RF is the most reliable model for distinguishing normal and cracked eggs using GLCM and Gabor texture feature extraction, although model combinations via stacking have yet to exceed the efficacy of standalone RF. Consequently, it may be inferred that the bagging-based ensemble method is more appropriate than boosting or stacking for egg image classification.

ACKNOWLEDGMENT

This work is supported by the National Science Foundation KEMENRISTEK DIKTI (Grant number: 124/C3/DT.05.00/PL/2025 and Derivate Grant number: 0969/LL3/AL.04/2025).

REFERENCES

[1] K. El-Sabrout, S. Aggag, and B. Mishra, "Advanced Practical Strategies to Enhance Table Egg Production," *Scientifica*, vol. 2022, pp. 1–17, Oct. 2022, https://doi.org/10.1155/2022/1393392.

[2] S. Molnár and L. Szőllősi, "Sustainability and Quality Aspects of Different Table Egg Production Systems: A Literature Review," *Sustainability*, vol. 12, no. 19, Sept. 2020, Art. no. 7884, https://doi.org/10.3390/su12197884.

[3] J. Miranda *et al.*, "Egg and Egg-Derived Foods: Effects on Human Health and Use as Functional Foods," *Nutrients*, vol. 7, no. 1, pp. 706–729, Jan. 2015, https://doi.org/10.3390/nu7010706.

[4] Z. Kralik, G. Kralik, M. Košević, O. Galović, and M. Samardžić, "Natural Multi-Enriched Eggs with n-3 Polyunsaturated Fatty Acids, Selenium, Vitamin E, and Lutein," *Animals*, vol. 13, no. 2, Jan. 2023, Art. no. 321, https://doi.org/10.3390/ani13020321.

[5] M. Pal and J. Molnár, "The Role of Eggs as an Important Source of Nutrition in Human Health," *International Journal of Food Science and Agriculture*, vol. 5, no. 1, pp. 180–182, Mar. 2021, https://doi.org/10.26855/ijfsa.2021.03.023.

[6] D. Tabernik, S. Šela, J. Skvarč, and D. Skočaj, "Segmentation-based deep-learning approach for surface-defect detection," *Journal of Intelligent Manufacturing*, vol. 31, no. 3, pp. 759–776, Mar. 2020, https://doi.org/10.1007/s10845-019-01476-x.

[7] C. Wang, J. Zhou, H. Wu, J. Li, Z. Chunjiang, and R. Liu, "Research on the Evaluation Method of Eggshell Dark Spots Based on Machine Vision," *IEEE Access*, vol. 8, pp. 160116–160125, 2020, https://doi.org/10.1109/ACCESS.2020.3020260.

[8] G. Kulshreshtha, L. D'Alba, I. C. Dunn, S. Rehault-Godbert, A. B. Rodriguez-Navarro, and M. T. Hincke, "Properties, Genetics and Innate Immune Function of the Cuticle in Egg-Laying Species," *Frontiers in Immunology*, vol. 13, Feb. 2022, Art. no. 838525, https://doi.org/10.3389/fimmu.2022.838525.

[9] M. Cisneros-Tamayo *et al.*, "Investigation on eggshell apex abnormality (EAA) syndrome in France: isolation of Mycoplasma synoviae is frequently associated with Mycoplasma pullorum," *BMC Veterinary Research*, vol. 16, no. 1, Dec. 2020, Art. no. 271, https://doi.org/10.1186/s12917-020-02487-0.

[10] B. Ahmed, C. De Boeck, A. Dumont, E. Cox, K. De Reu, and D. Vanrompay, "First Experimental Evidence for the Transmission of *Chlamydia psittaci* in Poultry through Eggshell Penetration,"

Transboundary and Emerging Diseases, vol. 64, no. 1, pp. 167–170, Feb. 2017, https://doi.org/10.1111/tbed.12358.

[11] S. Li and X. Zhao, "Automatic Crack Detection and Measurement of Concrete Structure Using Convolutional Encoder-Decoder Network," *IEEE Access*, vol. 8, pp. 134602–134618, 2020, https://doi.org/10.1109/ACCESS.2020.3011106.

[12] W. Tang, J. Hu, and Q. Wang, "High-Throughput Online Visual Detection Method of Cracked Preserved Eggs Based on Deep Learning," *Applied Sciences*, vol. 12, no. 3, Jan. 2022, Art. no. 952, https://doi.org/10.3390/app12030952.

[13] B. Purahong, V. Chaowalittawin, W. Krungseanmuang, P. Sathaporn, T. Anuwongpinit, and A. Lasakul, "Crack Detection of Eggshell using Image Processing and Computer Vision," *Journal of Physics: Conference Series*, vol. 2261, no. 1, June 2022, Art. no. 012021, https://doi.org/10.1088/1742-6596/2261/1/012021.

[14] Y. M. Valencia *et al.*, "A Novel Method for Inspection Defects In Commercial Eggs Using Computer Vision," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIII-B2-2021, pp. 809–816, June 2021, https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-809-2021.

[15] 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, July 2023, Art. no. 2354, https://doi.org/10.3390/ani13142354.

[16] Y. Huang *et al.*, "Damage Detection of Unwashed Eggs through Video and Deep Learning," *Foods*, vol. 12, no. 11, May 2023, Art. no. 2179, https://doi.org/10.3390/foods12112179.

[17] "Broken Eggs." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/frankpereny/broken-eggs.

[18] C. E. Ko, P. H. Chen, W. M. Liao, C. K. Lu, C. H. Lin, and J. W. Liang, "Using A Cropping Technique or Not: Impacts on SVM-based AMD Detection on OCT Images," in *2019 IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, Hsinchu, Taiwan, Mar. 2019, pp. 199–200, https://doi.org/10.1109/AICAS.2019.8771609.

[19] D. Occorsio, G. Ramella, and W. Themistoclakis, "An Open Image Resizing Framework for Remote Sensing Applications and Beyond," *Remote Sensing*, vol. 15, no. 16, Aug. 2023, Art. no. 4039, https://doi.org/10.3390/rs15164039.

[20] S. Rani, Y. Chabrra, and K. Malik, "An Improved Denoising Algorithm for Removing Noise in Color Images," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8738–8744, June 2022, https://doi.org/10.48084/etasr.4952.

[21] A. Setiawan, K. Adi, and C. E. Widodo, "Comparative Analysis of Deep Convolutional Neural Network for Accurate Identification of Foreign Objects in Rice Grains.," *Engineering Letters*, vol. 32, no. 2, 2024.

[22] A. Setiawan, K. Adi, and C. E. Widodo, "Rice Foreign Object Classification Based on Integrated Color and Textural Feature Using Machine Learning," *Mathematical Modelling of Engineering Problems*, vol. 10, no. 2, pp. 572–580, Apr. 2023, https://doi.org/10.18280/mmep.100226.

[23] G. Ulutas and B. Ustubioglu, "Underwater image enhancement using contrast limited adaptive histogram equalization and layered difference representation," *Multimedia Tools and Applications*, vol. 80, no. 10, pp. 15067–15091, Apr. 2021, https://doi.org/10.1007/s11042-020-10426-2.

[24] D. Chyzhyk, G. Varoquaux, M. Milham, and B. Thirion, "How to remove or control confounds in predictive models, with applications to brain biomarkers," *GigaScience*, vol. 11, Mar. 2022, Art. no. giac014, https://doi.org/10.1093/gigascience/giac014.

[25] T. T. Wong and P. Y. Yeh, "Reliable Accuracy Estimates from *k* -Fold Cross Validation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 8, pp. 1586–1594, Aug. 2020, https://doi.org/10.1109/TKDE.2019.2912815.