Region-Specific Adaptive GLCM Framework for Surface Scratch Detection in Heterogeneous Chocolate Products

Ruth Moly Benjamin[1], Abraham Chandy D.[2], Hepzibah Christinal A.[3]

[1,2]Department of Electronics and Communication Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India.
[3]Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India.

E-mail: [1]mbruth17@gmail.com, [2]abrahamdchandy@gmail.com, [3]christyhep@gmail.com

Orcid ID: [1]0000-0002-2272-4304, [2]000-0002-0655-4544, [3]0000-0003-3965-3183

Abstract

Automated defect detection on the surface of chocolate products is one of the most important techniques that must be used to ensure the quality and consistency of chocolate products during their manufacturing processes. However, due to the variation in the shapes, textures, and reflectivity of different brands of chocolates, this process can become difficult. To address this issue, a region-based adaptive GLCM approach is proposed for the analysis of textures on the surface of chocolates. Local variance analysis is carried out to segment the images into texture homogeneous regions. Following feature extraction, Principal Component Analysis (PCA)-based transformation is performed for feature transformation and then classified using a Support Vector Machine (SVM). A dataset consisting of 233 original chocolate images from 14 commercial brands, together with augmented training samples, was used to evaluate the proposed framework. The proposed method demonstrated an accuracy of 93.59%, recall of 96.15%, precision of 86.21%, F1-score of 90.91%, and AUC of 97.78% during the multi-brand evaluation. Leave-One-Brand-Out validation results achieved a mean recall of 93.42%, demonstrating its robustness in cross-brand defect detection. The proposed framework provides reliable automated chocolate scratch defect detection while maintaining low computational complexity and high interpretability.

Keywords: Adaptive Feature Extraction, Chocolate Defect Detection, Gray-Level Co-occurrence Matrix (GLCM), Region-wise Segmentation, Support Vector Machine (SVM)

1. Introduction

Surface quality is an important factor in chocolate production as it affects consumer appeal and the value of the product. Surface scratches are among the commonly occurring imperfections frequently introduced during the manufacturing stage, storage, and distribution. These defects do not affect the internal structure, but they reduce visual quality, hence making automated quality inspection necessary in industrial applications. Due to variations in texture and reflectance among different brands, scratch, defect detection in chocolate is challenging.

DOI: https://doi.org/10.36548/jiip.2026.3.013
Received: 16.05.2026, received in revised form: 14.06.2026, accepted: 05.07.2026, published: 16.07.2026

Scratches are usually minor and confined, presenting as fine irregularities on predominantly smooth surfaces. In most cases, scratches are subtle and confined to small regions on smooth and uniform surfaces. As a result, the effectiveness of traditional image analysis techniques is reduced. Among texture analysis techniques used for surface defect analysis, the Gray-Level Co-occurrence Matrix (GLCM) is one of the most prominent and widely adopted methods [1]. The traditional GLCM technique uses global texture analysis with constant parameter values under the assumption that the surface is composed of homogenous texture. Such an approach yields low sensitivity since defects present in the local area are overshadowed by nearby smooth regions. An attempt is made here to overcome the limitations mentioned above through the application of an adaptive GLCM approach for chocolate surface textures. The use of localized segmentation along with adaptive distances helps in accurately capturing local texture features.

2. Related Work

The most successful technique used to measure textures is the Gray Level Co-occurrence Matrix (GLCM). The use of GLCM-based texture features with classifiers in an automated banana grading system has been successfully proposed by Ebenezer O. Olaniyi et al. [2]. The findings showed that the GLCM feature is effective at representing the variations in texture features, resulting in improved classification performance compared to the traditional human inspection approach. Similarly, in another study, GLCM-based texture features were used by Chimango Nyasulu et al [3] for the classification of tomatoes affected by various fungal diseases through the use of different machine learning approache. The study concluded that GLCM-based features are highly effective at distinguishing visually similar disease characteristics. Recently, Pratomo Setiaji et al [4] proposed a hybrid GLCM-CNN model for automated egg quality grading using GLCM and deep learning approaches such as VGG19, ResNet50, and VGG16. For a spatial relationship characterized by distance d and orientation θ, the Gray-Level Co-occurrence Matrix (GLCM) is formulated as P (i, j | dθ) where:

i and j correspond to gray-level values of the reference and adjacent pixels

d represents the pixel displacement distance.

θ specifies the orientation angle.

The normalized matrix is expressed in Eq. (1):

$$\sum_{i,j} P(i,j) = 1 \tag{1}$$

Using symmetric and normalized configurations, the co-occurrence matrices describe the probability distribution of gray-level intensity pairs [5]. While these methods offer advantages and are effective, reliance on global texture-based methods with direct feature extraction has limited ability in representing minor defects and variations. For the current study, the traditional GLCM method is first implemented to create a baseline for comparison with the proposed methodology.

3. Proposed Methodology

3.1 Problem Formulation

Consider an image $I \in R^{H \times W}$ indicates a sample of a chocolate surface. The detection problem is framed as a two-class classification problem as shown in Eq. (2).

$$y \in \{0,1\} \tag{2}$$

where:

- y=0 corresponds to GOOD representing a surface free of scratch defects
- y=1 indicates BAD, representing the existence of a scratch defect
- The study aims to design a feature extraction framework focused on texture sensitivity.

3.2 System Overview

The proposed defect identification framework is illustrated in Figure 1. In the first stage, the chocolate image is preprocessed using grayscale conversion, image resizing, Gaussian smoothing, and contrast enhancement using CLAHE. In the next stage, the chocolate surface is partitioned into smooth and irregular texture regions using a variance-based segmentation framework. Subsequently, GLCM-based texture features are computed to characterize local texture variations associated with surface scratch defects. Following feature extraction, normalization and Principal Component Analysis (PCA)-based dimensionality reduction are performed to retain the most informative features. Finally, the processed feature vectors are classified using a Support Vector Machine (SVM) classifier to categorize chocolate samples as defective or non-defective. The combination of region-based texture analysis with machine learning, the proposed method accurately detects localized surface scratches across diverse chocolate samples [6].

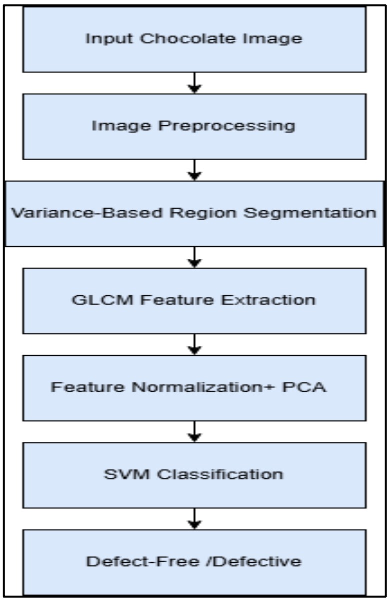

Figure 1. System architecture of the proposed Region-Specific adaptive distance GLCM framework for chocolate scratch defect detection

3.3 Dataset Overview

The data used in this project is based on chocolates sampled from 14 different commercial chocolate brands. The dataset includes both defective and non-defective chocolate samples. There are a total of 233 original images of chocolate samples, taken in a controlled light environment using a high-resolution mobile camera. Of these, 154 are non-defective (GOOD) samples and 79 are defective (BAD) samples with surface scratches. The dataset is further classified into independent training and testing datasets. There are 102 GOOD and 53 BAD original images in the training dataset, and 52 GOOD and 26 BAD original images in the testing dataset. To prevent data leakage, the train-test split was applied prior to performing any data augmentation. Data augmentation was done only to the training set. This ensures that no augmented samples of testing data are present in the training set, thus facilitating generalization assessment.

To provide balanced representation of classes in the training set, 318 augmented GOOD samples were randomly chosen from the 1,428 augmented samples created were randomly chosen and combined with 318 augmented BAD samples. The training set then contained 420 GOOD samples and 371 BAD samples, totaling 791 training samples.

Table 1. Dataset distribution

Dataset Category	Good	Bad	Total
Original Training Images	102	53	155
Original Testing Images	52	26	78
Augmented Images Used for Training	318	318	636
Final Training Dataset	420	371	791
Final Testing Dataset	52	26	78

Figure 2 shows sample images from the dataset, highlighting defect-free(GOOD) and defective chocolate products from multiple commercial brands.

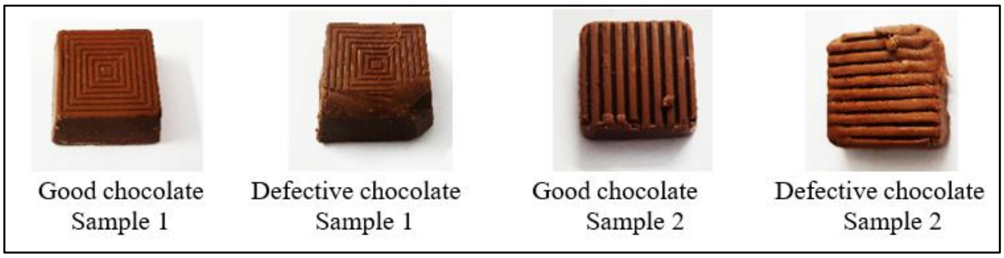

Figure 2. Sample images from the dataset showing good and defective chocolates

3.4 Defect Simulation Protocol

Due to the shortage of naturally developed scratches, defects were artificially introduced. Scratches were deliberately introduced using a fine cutting tool to mimic real world industrial wear. Variations were implemented in terms of:

- Scratch orientation
- Scratch length
- Intensity variation
- Spatial surface location

The generated defects primarily modified local texture characteristics while preserving the overall shape and geometric structure of the samples. As a result, the proposed framework was designed to detect texture variations associated with surface scratches while reducing the impact of geometric variations. Although controlled scratch generation enabled a structured evaluation of the texture descriptors, it may not fully reflect the complete diversity of real-world industrial scratch patterns.

3.5 Image Preprocessing

The acquired chocolate images were first converted to grayscale using the OpenCV framework (Figure 3). Color features were not included in the analysis, as the descriptors generated through GLCM analysis rely on spatial intensity relationships. A fixed image resolution of 128 × 128 pixels was maintained for all images in the dataset. Preliminary experiments demonstrated that this resolution preserved the key texture characteristics for scratch detection while minimizing computational complexity and extraction expenses [7].

Prior to texture analysis, Gaussian filtering with a 3 × 3 kernel was performed to reduce image noise. Subsequently, Contrast Limited Adaptive Histogram Equalization (CLAHE) (clip limit = 2.0, tile grid size = 8 × 8) was applied to improve the visibility of local texture details. A 7 × 7 local window was applied to estimate local texture variations across the chocolate surface. The variance-based representation was then used to differentiate low-variance and high-variance texture regions, enabling adaptive GLCM feature extraction (Figure 3).

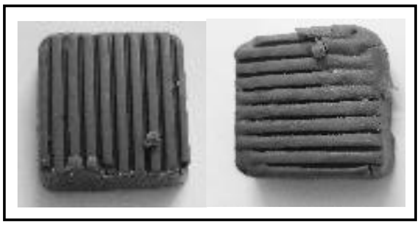

Figure 3. Grayscale conversion of input chocolate images in the preprocessing stage.

3.5 Traditional Global GLCM (Baseline Model)

The initial approach relies on classical global GLCM texture modelling. In this approach, the entire image is assumed to represent as one homogeneous texture region. For this study, fixed spatial parameters were employed:

Distances: {1,2}

Angles: {0°, 45°, 90°, 135°}

Four Haralick texture descriptors are obtained from each derived GLCM matrix:

- Texture contrast
- Texture homogeneity
- Texture energy
- Pixel intensity correlation

Feature extraction is carried out across:

- 4 texture properties
- 2 spatial distances
- 4 directional orientations

The resulting feature becomes, Eq. (3):

$$4 \times 2 \times 4 = 32 \quad (3)$$

3.7 Region-Wise Fixed Distance GLCM

Scratch defects are usually confined to specific localized areas within the chocolate surface [8]. When GLCM features are calculated across the whole image, uniform regions dominate the extracted texture features, lowering the significance of defect-related information. This issue is addressed by applying texture modelling independently across regions.

3.7.1 Local Texture Segmentation

To identify regions with variations in texture characteristics local variance is computed. The variance within the local neighborhood for each pixel location (x,y), is given by Eq. (4).

$$V(x,y) = \mathrm{Var}(N(x,y)) \quad (4)$$

where N(x,y) corresponds to the local neighborhood window surrounding the pixel. An adaptive threshold was determined by combining the mean local variance with 0.6 times its corresponding standard deviation. The pixels were partitioned into high-variance and low-variance regions based on the adaptive threshold, with values above the threshold assigned to high-variance regions, whereas those with values less than or equal to the threshold were assigned to low-variance regions.

The resulting variance map, illustrated in Figure 4, delineates:

- Low-variance areas corresponding to smooth surface areas
- High-variance areas indicating scratch-induced surface irregularities

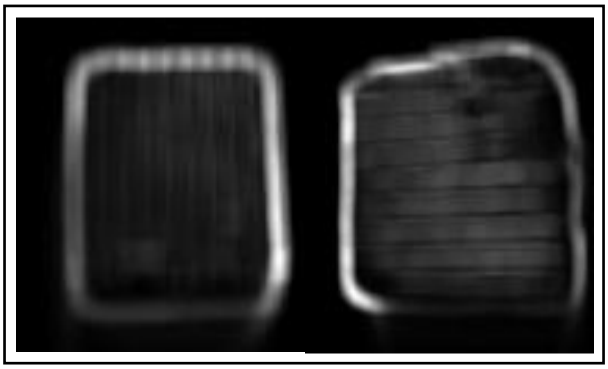

Figure 4. Representation of a local variance map for good and defective chocolate samples

3.7.2 Feature Extraction

Each segmented region is examined individually by GLCM-based texture descriptors, allowing for region-specific analysis, as shown in Figure 5.

For each segmented region:

- GLCM matrices were generated for both low-variance and high-variance texture regions using fixed spatial distances of {1, 2}
- Texture features were extracted using four directional orientations: 0°, 45°, 90°, and 135°.
- Haralick-based texture features, namely contrast, homogeneity, energy, and correlation, were then computed.

The final feature vector was constructed by combining the feature vectors obtained from the low-variance and high-variance regions, as described in Eq. (5):

$$f_{region} = [f_{low}, f_{high}] \tag{5}$$

As each region contributes 32 features, the total feature dimension is Eq. (6):

$$32 \times 2 = 64 \tag{6}$$

By preserving spatial characteristics, the proposed method preserves spatial information while improving the detection of localized scratch defects.

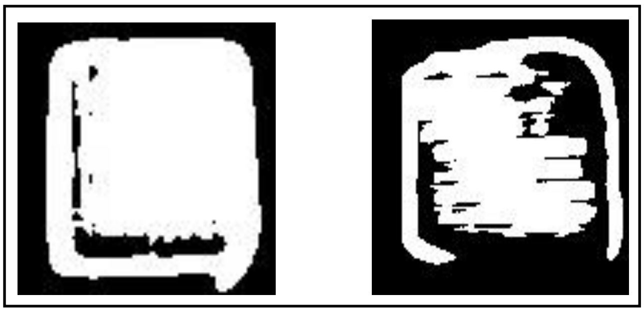

Figure 5. Region-based segmentation for good and defective samples.

3.8 Region-Specific Adaptive Distance GLCM

GLCM-based methods are often constrained by fixed spatial scales, limiting their ability to represent diverse texture patterns across different levels of granularity. Smooth chocolate surfaces have a fine grain structure, while scratched surfaces show a coarse grain structure. To solve this problem, an adaptive distance strategy is used.

3.8.1 Adaptive Distance Strategy

Every identified region was classified into variance categories according to its mean value. The low-variance categories were set with short spatial distances {1, 2} since smooth chocolate structures had fine grain textures associated with short-range spatial dependencies.

Large-scale spatial distances {3, 4} were set to represent large-scale textures caused by scratches.

Regions with low variance (smooth surfaces), Eq. (7):

$$d = \{1,2\} \tag{7}$$

High-variance regions, Eq. (8):

$$d = \{3,4\} \tag{8}$$

The adaptive distance selection algorithm allows the GLCM computation scale to depend on local texture properties, thus improving the modeling of uniform surfaces and scratches. By applying specific spatial distances for different regions, multi-level texture description is possible without loss of computational performance.

3.8.2 Feature Construction

For each segmented region, the GLCM was estimated based on the respective spatial parameters. The Haralick texture features were then combined into one feature vector.

- Feature extraction was performed using:
- Distances {1,2}
- Orientations {0°, 45°, 90°, 135°}
- Four Haralick texture descriptors: contrast, homogeneity, energy, and correlation

Feature vectors having dimensions of 32 were generated for each of the segments. The final feature vectors were then combined together using the low and high variance segments. The process is captured in Eq. (9).

$$32 \times 2 = 64 \tag{9}$$

The suggested framework used a feature vector of size 64 and modified the texture analysis depending on local surface changes through variance-guided segmentation and adaptive distance selection.

3.9 Region-Specific Adaptive Multi-Scale GLCM

The framework was also improved to develop a multi-scale version to enhance texture representation at various spatial scales. Regions are analyzed at multiple distance values instead of considering a fixed value of spatial distances. The resulting GLCM-based feature are combined to form a multi-scale framework. As in the previous method, spatial distances {1, 2, 3} were considered for low-variance regions and {3, 4, 5} for high-variance regions. For each region, GLCM matrices were generated using the same four directional orientations {0°, 45°, 90°, 135°} from which the four Haralick texture descriptors—contrast, homogeneity, energy, and correlation—were extracted [9]. For each segmented region, four Haralick texture descriptors were extracted over three spatial distances and four directional orientations, resulting in a 48-dimensional feature vector. The feature vectors obtained from the low-variance and high-variance regions were subsequently concatenated to construct a 96-dimensional multi-scale texture representation (48×2=96).

3.10 Orientation-Aware Adaptive GLCM

Directional properties of scratch defects are analyzed using a GLCM-based framework in which gradient angles are obtained using Sobel filters as in Eq. (10).

$$G_x = \frac{\partial I}{\partial x}, \; G_y \; \frac{\partial I}{\partial y} \tag{10}$$

The magnitude and orientation of the gradient are determined as Eq. (11).

$$\emptyset = \operatorname{arctan2}(G_y, G_x) \tag{11}$$

A gradient orientation histogram is obtained to identify the dominant directions. Gradient orientations were quantized into eight angular bins covering the full range of orientation values $[-\pi, \pi]$. The two dominant orientation bins were selected and employed in the subsequent GLCM feature extraction process [10].

3.11 Feature Normalization and Dimensionality Reduction

Once features are extracted, statistical preprocessing is conducted.

3.11.1 StandardScaler

The features are normalized using, Eq. (12):

$$z = \frac{x - \mu}{\sigma} \tag{12}$$

where x denotes the original feature value, μ represents the feature mean, and σ denotes the feature standard deviation. This ensures balanced feature contributions while enhancing classifier robustness.

3.11.2 Principal Component Analysis (PCA)

The Principal Component Analysis (PCA) technique was used to achieve a lower dimensional feature set while preserving the most discriminant texture information. In the proposed Adaptive Distance GLCM methodology, the generation of the 64 dimensional feature set using GLCM features is done on multiple distance, orientations, and regions of interest, and may cause redundancy and correlation among features. Principal Component Analysis (PCA) maps the original feature space into orthogonal principal component space capturing the maximum informative variance in the data. Components with relatively less variance were ignored, thus, reducing the dimensionality of the feature set while preserving discriminative texture information. In the proposed methodology, Principal Component Analysis (PCA) was applied in such a way that 98% cumulative variance of the original feature set was maintained.

3.12 Classification

Classification was performed using a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel. RBF kernel is formulated as Eq. (13):

$$K(x_i, x_j) = \exp\left(-\gamma ||x_i - x_j||^2\right) \tag{13}$$

The use of an RBF kernel provides the ability to learn nonlinear decision boundaries that facilitate the modeling of texture distribution. To deal with imbalanced classes, penalty terms based on class weighting were incorporated into the SVM classifier. GOOD and BAD classes were assigned weights of 1.0 and 1.8, respectively, to penalize defective samples more severely during misclassification. The parameters used for the SVM classifier include C=5, γ=scale, and an RBF kernel.

3.13 Validation Strategy

The independence of the performance evaluation was guaranteed, and any data leakage was avoided by dividing the original chocolate images into two independent training and testing datasets before augmenting the data.

3.13.1 Single Brand Experiments

Single-brand experiments were performed using a number of chocolates belonging to a commercial brand to evaluate the performance of the proposed texture analysis techniques.

Evaluation Framework:

- Independent train–test evaluation
- Training performed using original and augmented samples
- Testing performed using unseen original samples

3.13.2 Multi-Brand Evaluation

This analysis was extended to include a multi-brand dataset comprising chocolate samples from 14 different brands. In comparison with the single-brand analysis, the multi-brand analysis includes several differences related to product appearance, surface texture, structural geometry, manufacturing aspects, and, in particular defects. The main goal of this analysis is to evaluate of the efficiency and generalization ability of the proposed frameworks in various real-life settings.

3.13.3 Leave-One-Brand-Out (LOBO) Validation

Generalization performance of the proposed technique on cross-brands was evaluated using the Leave-One-Brand-Out (LOBO) validation strategy. The evaluation process involved the following steps:

- A brand was excluded from training and considered only for testing.
- Training samples were generated using all brands except the one excluded.

- Feature normalization, feature dimensionality reduction via PCA, and classification via SVM were carried out using the training samples.
- The model was tested on the excluded brand.
- This process was repeated until each brand had been left out as the testing set once.

The recall of defective samples was computed in each round of LOBO, and then the average recall and standard deviation of the recall values were calculated for all the brands.

4. Results

The experimental implementation was done using the Python programming language with the aid of the OpenCV, scikit-image, and scikit-learn libraries. Before the feature extraction process, all images were converted to grayscale and resized to 128×128 pixels to achieve the same spatial resolution. In the calculation of the GLCM, 256 gray levels were used, which matches the gray-level intensity of the grayscale images. For the computation of the GLCM, symmetric and normalized GLCMs were used since the objective was to compute the probability distribution function of the gray levels in the image. The features from the different spatial distances and directional orientations were converted into one-dimensional feature vectors by concatenation. Classification was performed using an RBF SVM. To solve the class imbalance problem, weighted class penalty parameters of {GOOD: 1, BAD: 1.8} were considered. The standardization of the feature vectors was achieved using StandardScaler, after which PCA was applied to retain 98% of the total variance.

4.1 Single-Brand Experimental Evaluation

Experiments were initially conducted using samples from one commercial brand to assess the performance of the proposed texture-based defect detection frameworks. Performance assessment employed an independent train–test protocol, where only the training set was augmented while the testing set comprised only unseen original chocolate images. Three texture modelling strategies were investigated: Traditional Global GLCM, Region-Wise GLCM, and the proposed Region-Specific Adaptive Distance GLCM framework.

4.1.1 Traditional Global GLCM (Baseline)

In the initial experiment the traditional global GLCM approach was applied in which the whole chocolate surface image is considered a single uniform texture region.

Classification Pipeline: StandardScaler → PCA (98% variance) → RBF-SVM (C = 5, γ = scale)

The classification results of the Traditional Global GLCM baseline model in the single-brand experimental setup are presented in Table 2.

Table 2. Performance of the traditional global GLCM baseline model.

Method	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)	AUC (%)
Global GLCM	100	100	100	100	100

In controlled individual brand conditions the conventional global GLCM technique achieved perfect classification performance. However, the extracted texture representation was derived from the complete chocolate surface and did not explicitly focus on localized scratch regions. As a result, texture features associated with defects may be weakened by neighboring homogeneous surface regions. The effect of this limitation was negligible in the controlled single-brand study; but its impact becomes more pronounced when greater texture variability is encountered in multi-brand experiments.

4.1.2 Region-Wise Fixed GLCM

To resolve the drawbacks of global feature-based analysis, spatial information was incorporated by separating images into texture-based regions. Very small regions (fewer than 50 pixels) were removed to maintain stable texture measurements.

The classification results of the Region-Wise Fixed GLCM model in the single-brand experimental setup are presented in Table 3.

Table 3. Performance of the Region-Wise fixed GLCM model

Method	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)	AUC (%)
Region-Wise Fixed GLCM	100	100	100	100	100

The inclusion of spatial localization enhances classification performance by maintaining texture features associated with defects.

4.1.3 Region-Specific Adaptive Distance GLCM

The limitations of global texture analysis were addressed by partitioning the chocolate surface into texture-based regions for spatially aware analysis. A 7 × 7 neighborhood window was employed to characterize local texture variations across the image. The image was partitioned into two regions using an adaptive variance threshold determined by the mean local variance and 0.6 times its standard deviation. The classification results of the Region-Specific Adaptive Distance GLCM model in the single-brand experimental setup are presented in Table 4.

Table 4. Performance of the Region-Specific adaptive distance GLCM

Method	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)	AUC (%)
Region-Specific Adaptive Distance GLCM	100	100	100	100	100

Spatial localization enhanced defect-related texture representation by independently analyzing smooth and irregular surface regions. The preservation of localized texture information improves the framework's sensitivity to scratch-induced texture variations while retaining the computational simplicity of fixed-distance GLCM analysis. The perfect classification performance observed under the single-brand setting indicates that defect and non-defect samples were highly separable under controlled conditions.

4.2 Multi-Brand Experimental Evaluation

Further analysis of the proposed framework was conducted for the multi-brand chocolate dataset to assess its effectiveness and performance under conditions of high texture variation and defect representation. The achieved performance in this test provides a more difficult evaluation of the algorithm's defect detection ability than that in the controlled single-brand environment.

4.2.1 Traditional Global GLCM (Multi-Brand Baseline)

The performance of the Traditional Global GLCM algorithm showed 85.90% accuracy and 80.77% defect recall. Although the algorithm can successfully capture the general texture features of images, the limitation of representing only one global texture feature made it less efficient in capturing localized scratch patterns of different chocolates. Therefore, the performance of defect detection in this case was lower than that of the controlled single-brand test. The classification results of the Traditional Global GLCM algorithm in the multi-brand environment are shown in Table 5.

Table 5. Performance of the traditional global GLCM (Multi-Brand Baseline)

Method	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)	AUC (%)
Global GLCM	85.90	80.77	77.78	79.25	93.49

4.2.2 Region-Specific Adaptive Distance GLCM

The suggested Adaptive Distance GLCM framework achieved higher scores compared to the classical Global GLCM approach based on all used metrics. Accuracy was raised from 85.90% to 93.59%, and defect recall was raised from 80.77% to 96.15%. These results show that the suggested framework allowed increasing defect detections while simultaneously reducing false negatives. The improvement in performance can be explained by the suggested region-based approach and the adaptive distance calculation, which allows obtaining an accurate texture description according to the particularities of the surface structure. Due to the assignment of different distances for smooth and scratch-affected regions, the framework allows better distinguishing of fine texture details and defect patterns compared to global approaches to texture modeling. To present the workflow of the suggested framework, Figure 6 shows the intermediate stages of processing of a defective chocolate sample.The local variance image represents texture irregularities in the vicinity of scratch defects, and variance-based segmentation separates smooth and irregular surface regions.

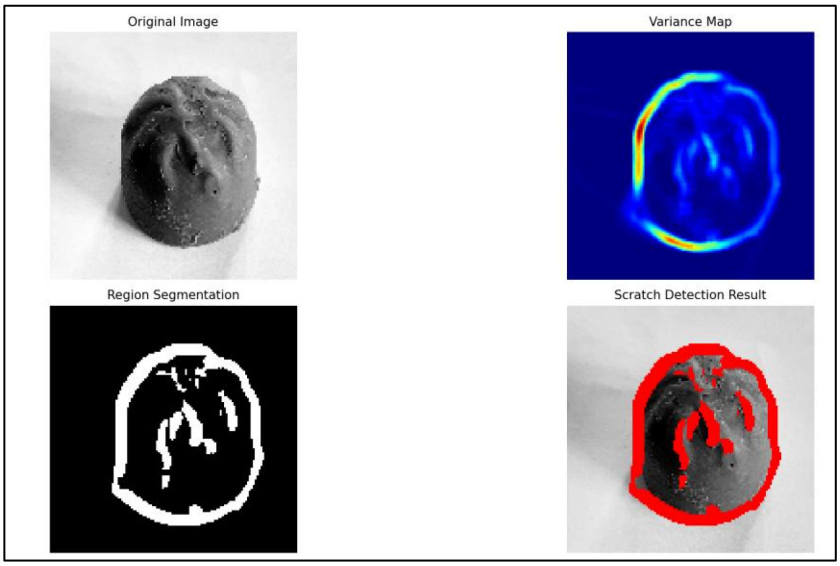

Figure 6. Intermediate steps involved in the Region-Specific adaptive distance GLCM framework including (a) the original image of chocolate, (b) local variance map, (c) regional segmentation based on variance map, and (d) scratch localization.

A high AUC value of 97.78% reveals the excellent discrimination ability of the proposed approach to discriminate defective from non-defective chocolates that have distinct surface textures and defects.

The classification outcomes of the proposed region-specific adaptive distance GLCM approach for multi-brand experiments are given in Table 6.

Table 6. Performance of the region-specific adaptive distance GLCM

Method	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)	AUC(%)
Adaptive Distance GLCM	93.59	96.15	86.21	90.91	97.78

4.2.3 Region-Specific Adaptive Multi-Scale GLCM

To obtain multiscale texture data, the adaptive technique was modified to compute the GLCM feature extraction at different spatial distances for each region segmentation.

The classification outcomes of the region-specific adaptive multiscale GLCM in multi-brand experiment setting can be seen in Table 7.

Table 7. Performance of the region-specific adaptive multi-scale GLCM

Method	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)	AUC(%)
Adaptive Multi-Scale GLCM	92.31	96.15	83.33	89.29	97.78

An accuracy of 92.31%, accompanied by a recall value of 96.15%, was achieved using the Adaptive Multi-Scale GLCM method. However, although the recall was similar to that of the Adaptive Distance GLCM approach, there were small decreases in accuracy, precision, and F1 score. In this way, the results suggest that the use of more distances did not bring much more discriminating power to the task of detecting the particular chocolate scratch defects.

4.2.4 Orientation-Aware Adaptive GLCM

A model based on orientation-sensitive GLCM was studied for analyzing the directional nature of scratches. The classification performance of Orientation-Sensitive Adaptive GLCM in the multiple brands scenario is provided in Table 8.

Table 8. Performance of the Orientation-Aware adaptive GLCM

Method	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)	AUC (%)
Orientation-Aware Adaptive GLCM	88.46	88.46	79.31	83.64	94.90

The performance obtained from the orientation-sensitive method was slightly lower than that of the adaptive distance GLCM method, indicating that the investigated scratches have no strong directional consistency.

4.3 Additional Analysis of the Proposed Adaptive Distance GLCM Framework

4.3.1 Confusion Matrix and Performance Analysis

Additional evaluation metrics, including accuracy, precision, F1-score, specificity, AUC, and confusion matrix analysis, were computed to further evaluate the effectiveness of the proposed Adaptive GLCM framework. The results obtained are summarized in Table 9.

Table 9. Performance metrics obtained for the adaptive GLCM framework

Metric	Value
Recall	96.15%
Accuracy	93.59%
Precision	86.21%
F1-Score	90.91%
Specificity	92.31%
AUC	97.78%

The classification performance of the proposed Adaptive GLCM framework is illustrated in Table 9. With an accuracy value of 93.59% and specificity of 92.31%, the proposed Adaptive GLCM framework demonstrates its efficiency in distinguishing non-defective chocolate samples. A high recall value of 96.15% indicates the framework's high sensitivity to scratch defects, minimizing the risk of false negatives. With a precision of 86.21%, it can be concluded that most samples marked as defective were actually defective. An F1-score value of 90.91% also reveals that the proposed framework maintains a good trade-off between precision and recall values. The AUC value of 97.78% illustrates good discrimination between defect-free and defective samples.

Figure 7 illustrates the confusion matrix calculated for the proposed adaptive GLCM framework in the multi-brand evaluation process.

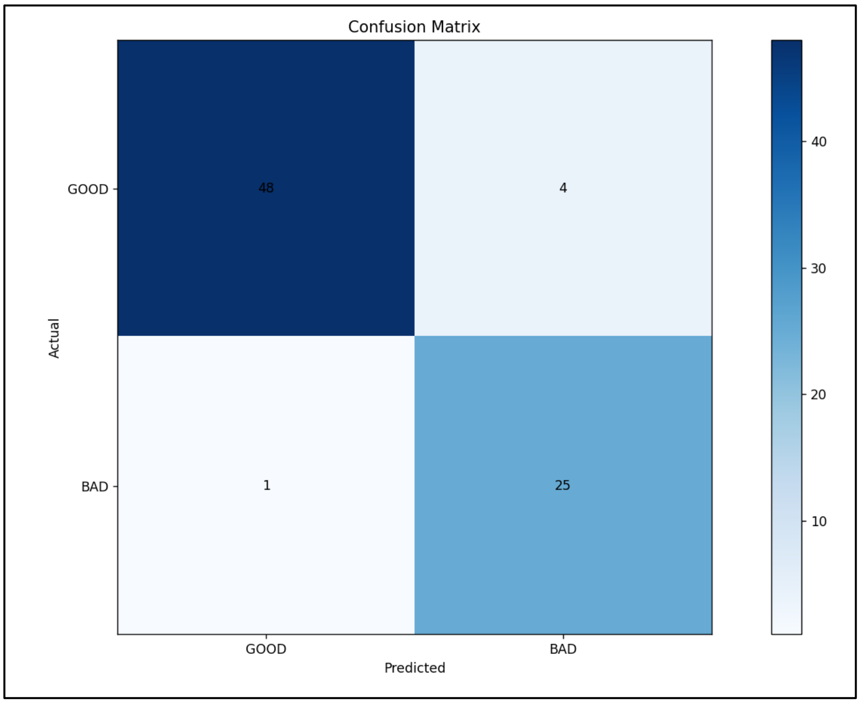

Figure 7. Confusion matrix obtained for the proposed adaptive GLCM framework

From the confusion matrix, it is clear that 48 non-defective and 25 defective chocolate samples were classified successfully. Misclassification only happened in four cases where non-defective chocolates were classified as defective while one defective chocolate was classified as non-defective. Because of the low rate of false negatives, the proposed framework achieved a recall rate of 96.15%.

4.3.2 False Positive Analysis

The confusion matrix revealed four false positives, where non-defective chocolates were incorrectly classified as defective. Examination of the misclassified samples revealed embossed surface structures, minor boundary irregularities, localized surface roughness, and illumination-related intensity variations. Despite being defect-free, their surface characteristics produced high-variance texture patterns that resembled scratch defects. As a result, the extracted adaptive texture descriptors showed characteristics similar to those of defect-related patterns, resulting in false-positive classifications. A specificity of 92.31% was achieved, with only four of the fifty-two defect-free samples reflecting the framework's ability to distinguish genuine scratch defects from normal surface texture variations.

4.3.3 Comparison with MobileNetV2

The performance of the proposed Adaptive Distance GLCM framework was evaluated against MobileNetV2, which served as the benchmark deep learning model. To ensure a fair performance comparison, both approaches were trained and evaluated using identical data partitions. Table 10 summarizes the performance comparison of the proposed Adaptive Distance GLCM framework and MobileNetV2

Table 10. Comparison between the proposed adaptive GLCM framework and MobileNetV2 baseline model

Method	Accuracy (%)	Recall (%)	Precision(%)	F1-Score (%)	AUC (%)
Proposed Adaptive GLCM	93.59	96.15	86.21	90.91	97.78
MobileNetV2	92.31	80.77	95.45	87.50	96.75

The proposed adaptive GLCM framework achieved higher accuracy (93.59%), recall (96.15%), F1-score (90.91%), and AUC (97.78%) compared to the MobileNetV2 baseline model. Although the MobileNetV2 recall is lower at 80.77%, its precision rate of 95.45% means more chocolate defects were erroneously detected as good chocolates. Relative to the benchmark method, the developed architecture can successfully classify 25 out of 26 defective chocolate samples.

From these experimental results, it can be seen that the proposed region-specific adaptive texture descriptors are quite effective in representing the discriminative texture features associated with scratches on the surface.

4.3.4 Leave-One-Brand-Out (LOBO) Validation

The cross-brand performance of the proposed Region-Specific Adaptive Distance GLCM framework was assessed using Leave-One-Brand-Out (LOBO) validation. During each validation iteration:

- One brand was designated as the test set
- The other 13 brands were used for training
- Model training and evaluation were performed
- The same procedure was applied for each brand

The Leave-One-Brand-Out (LOBO) recall performance for the proposed Adaptive GLCM framework across 14 chocolate brands are shown in Table 11.

Table 11. Leave-One-Brand-Out (LOBO) recall performance

Metric	Value
Mean Recall	93.42%
Standard Deviation	11.99%
Highest Recall	100.00%

The framework maintained consistent defect detection performance across multiple chocolate brands with diverse surface textures. Variations in recall across different chocolate brands indicate the inherent challenges of cross-brand texture generalization in automated chocolate defect inspection systems. Table 12 shows the Leave-One-Brand-Out (LOBO) recall performance of all 14 chocolate brands.

Table 12. Leave-One-Brand-Out (LOBO) recall performance

Brand ID	Recall (%)
Brand 1	100.00
Brand 2	90.48
Brand 3	90.48
Brand 4	96.43
Brand 5	95.24
Brand 6	100.00
Brand 7	97.14
Brand 8	100.00
Brand 9	52.38
Brand 10	89.29
Brand 11	96.43
Brand 12	100.00
Brand 13	100.00
Brand 14	100.00

Note: To preserve commercial confidentiality, the identities of the chocolate brands are anonymized and reported as Brand 1–Brand 14.

The results of the brand-wise LOBO analysis indicate that the proposed Adaptive Distance GLCM framework maintained reliable defect detection performance across most chocolate brands. Eight of the evaluated chocolate brands exhibited recall values greater than 95%, with six brands demonstrating perfect recall. Variations in performance across chocolate brands resulted in a mean recall of 93.42% with a standard deviation of 11.99%, highlighting the influence of variations in surface texture characteristics and defect appearance.

4.3.5 Comparative Performance Analysis

Table 13 provides the recall performance of all evaluated models and demonstrates the effectiveness of adaptive texture analysis for chocolate scratch defect detection.

Table 13. Performance comparison across the various models

Model	Evaluation	Recall
Global GLCM	Multi-brand Evaluation	80.77%
Region-Specific Adaptive Distance GLCM	Multi-brand Evaluation	96.15%
Region-Specific Adaptive Multi-Scale GLCM	Multi-brand Evaluation	96.15%
Orientation-Aware Adaptive GLCM	Multi-brand Evaluation	88.46%
Region-Specific Adaptive Distance GLCM	LOBO Validation	93.42%

The proposed Region-Specific Adaptive Distance GLCM framework demonstrated superior defect detection performance by achieving the highest overall defect recall in the multi-brand evaluation while maintaining reliable cross-brand performance in the LOBO evaluation.

5. Discussion

The results demonstrate that effective modelling of spatial texture information significantly contributes to the performance of texture-based scratch defect detection. The study systematically investigated GLCM-based strategies, encompassing approaches ranging from conventional global modelling to sophisticated region-aware and adaptive frameworks.

5.1 Limitations of Global Texture Modeling

The traditional global GLCM approach generated an average recall of 80.77% for the multi-brand experiment. Although the traditional global GLCM model is capable of capturing the global texture properties of the surface, it performed worse compared to the proposed adaptive models. The traditional global GLCM approach captures global texture properties but cannot capture the spatial details to detect fine defect structures. This disadvantage becomes more visible in multi-brand settings, where there are significant differences between the appearances and textures of chocolate surfaces. These findings demonstrate that accurate scratch defect detection relies on both texture characterization and the retention of localized defect-related information.The findings highlighted the need for region-based and adaptive texture representation, leading to the development of the proposed methods.

5.2 Impact of Region-Wise Texture Modelling

The use of region-wise texture analysis resulted in improved defect detection performance. By employing local variance mapping and region-wise segmentation, the proposed framework enhanced the effectiveness of feature extraction by preserving localized texture variations. While both methods demonstrated identical performance in the single-brand evaluation, region-wise segmentation enabled a more effective representation of localized defect-related texture information. This finding highlights the importance of maintaining spatial texture heterogeneity for accurate defect detection. The benefits of region-wise modelling were more clearly demonstrated in the multi-brand evaluation. Region-wise modelling replaces the global texture representation providing greater emphasis on defect-related regions during feature extraction.

5.3 Effectiveness of Adaptive Distance Selection

Further improvements in defect detection performance are achieved using the developed region-specific adaptive GLCM method. The adaptive framework assigns spatial distances to each segmented region based on its local texture characteristics. Smaller pixel distances are better suited for effective characterization of fine-scale textures on smooth chocolate surfaces. Unlike smooth regions, scratch defects exhibit broader structural variations that are more accurately captured using greater spatial distances. Through adaptive selection of spatial distances based on local texture characteristics, the framework increases the responsiveness of texture descriptors to better represent defect structures. The use of adaptive modeling improved recall performance, with values increasing from 80.77% (global GLCM) to 96.15% in the multi-brand setting, indicating that adaptive spatial modelling contributes positively to defect detection performance.

5.4 Analysis of Multi-Scale Texture Modeling

To further study the influence of spatial scale variations, a multi-scale GLCM framework was assessed. In this method, features extracted at different spatial distances were integrated into a higher-dimensional feature representation. Despite achieving the same recall (96.15%) as the adaptive distance GLCM framework, no further improvement in defect detection performance was observed. A greater number of scales leads to increased feature dimensionality but yields only marginal improvements in discriminative capability [13].

5.5 Directional Orientation of Scratch Defects

The mean recall value dropped to 88.46% when compared to the Adaptive Distance technique. The findings suggest that there is no clear directional orientation present in the data set. Scratch defects possess fine irregularity patterns with arbitrary orientation directions, and therefore multi-directional texture analysis would be better suited for them, hence the slight drop in accuracy rate [14].

5.6 Ability to Generalize Across Different Chocolate Brands

The leave-one-brand-out (LOBO) validation technique was used to test the ability of the framework to generalize across different chocolate brands. The adaptive framework generated an average recall of 93.42% using the leave-one-brand-out (LOBO) validation technique. It is clear that the drop from 96.15% recall when evaluating across the multi-brands is rather insignificant, suggesting that the framework is able to capture defect-specific texture features and not brand-specific visual features [15].

5.7 Computational Complexity and Runtime Analysis

The main sources of computational complexity are in computing the local variance and extracting features from the GLCMs. In an image of size N pixels, pre-processing and variance estimation steps are linear O(N). Calculating the GLCM matrix takes O(N x D x A), where D and A represent the spatial distance and angle respectively. The computational complexity of the feature extraction step depends on the number of features to be computed and different parameters of GLCM. In order to evaluate the actual computational efficiency, timing experiments were conducted based on the proposed experimental design. The most computationally demanding step was the feature extraction step which took 220.41 s for the entire dataset. In contrast, feature normalization and PCA required only 0.014 s, while SVM training required 0.048 s. The average prediction time was 0.00092 s, corresponding to approximately 1.18×10^{-5} s per image.

Table 14 shows the runtime analysis of the proposed method. The results indicate that feature extraction dominates the overall computational cost, whereas the classification stage incurs minimal computational overhead. The use of a compact 64-dimensional feature vector and an efficient SVM classifier provides rapid computation while maintaining strong defect detection capability.

Table 14. Runtime analysis of the proposed adaptive distance GLCM framework

Operation	Time (s)
Feature Extraction	220.41
Feature Normalization + PCA	0.014
SVM Training	0.048
Average Prediction Time	0.00092
Average Prediction Time per Image	1.18×10^{-5}

5.8 Practical Implications for Industrial Inspection

The proposed method provides multiple advantages that enhance its suitability for industrial automated chocolate inspection systems. The method uses interpretable texture features that create a straightforward relationship between the structure of defects and classification outcomes. Moreover, GLCM-based feature extraction is computationally very efficient 16]. Combining the region-wise analysis with adaptive spatial modeling improves the ability to detect local surface defects on heterogeneous chocolate pieces. The proposed framework showed reliable defect detection performance for various chocolate brands. Nevertheless, more research is required to evaluate the robustness of the proposed framework for actual industrial image acquisition settings. The framework combines region-wise texture analysis and adaptive texture characterization allowing for the localized analysis of surface defects. Compared to the proposed framework, deep learning methods usually provide less interpretable feature representations from large-scale datasets [17]. The experiment was performed in a controlled laboratory imaging setting without conveyor-belt simulation and variable illumination. This proposed approach is able to achieve uniform performance in defect detection regardless of the different chocolate brands. In order to maintain uniformity, the surface scratch defects used in this experiment were generated manually to mimic realistic scratch defects. However, while these defects are manually created with variation in intensity, orientation, and spatial distribution, there is no guarantee that they are a true representation of the defects found in real-life situations such as on a conveyor belt and packaging [18].

6. Conclusion

Texture analysis is used in the suggested method for detecting surface scratch defects in various chocolate products. Since the existing traditional global GLCM approach gives only a mean recall of 80.77%, it is probable that the cause for such poor performance is this algorithm’s inability to properly model local defects. To improve performance, regional-based texture modeling is used first and then the region-specific distance-based GLCM algorithm is applied. The improvement in defect sensitivity comes from adaptive selection of distance in relation to the texture of different regions, thus giving a mean recall of 96.15% on a multi-brand dataset. The suggested method performed better than MobileNetV2 in terms of recall, F1-score and AUC score, while remaining computationally efficient and interpretable.

References

[1] Singh, Shruti, Divya Srivastava, and Suneeta Agarwal. "GLCM and its Application in Pattern Recognition." In 2017 5th International Symposium on Computational and Business Intelligence (ISCBI), IEEE, 2017, 20-25.

[2] Olaniyi, Ebenezer O., Adefemi A. Adekunle, Temitope Odekuoye, and Adnan Khashman. "Automatic System for Grading Banana using GLCM Texture Feature Extraction and Neural Network Arbitrations." Journal of food process engineering 40, no. 6 (2017): e12575.

[3] Nyasulu, Chimango, Awa Diattara, Assitan Traore, Cheikh Ba, Papa Madiallacké Diedhiou, Yakhya Sy, Hind Raki, and Diego Hernán Peluffo-Ordóñez. "A Comparative Study of Machine Learning-Based Classification of Tomato Fungal Diseases: Application of GLCM Texture Features." Heliyon 9, no. 11 (2023).

[4] Setiaji, Pratomo, Kusworo Adi, and Bayu Surarso. "Development of Classification Method for Determining Chicken Egg Quality Using GLCM-CNN Method." Ingenierie des Systemes d'Information 29, no. 2 (2024): 397.

[5] Benco, Miroslav, Robert Hudec, Patrik Kamencay, Martina Zachariasova, and Slavomir Matuska. "An Advanced Approach to Extraction of Colour Texture Features Based on GLCM." International Journal of Advanced Robotic Systems 11, no. 7 (2014): 104.

[6] Nielsen, Birgitte, Fritz Albregtsen, and Håvard E. Danielsen. "Low Dimensional Adaptive Texture Feature Vectors from Class Distance and Class Difference Matrices." IEEE transactions on medical imaging 23, no. 1 (2004): 73-84.

[7] Kumar, Chandan, Siddharth Chauhan, and R. Narmadha Alla. "Classifications of Citrus Fruit using Image Processing-GLCM Parameters." In 2015 International Conference on Communications and Signal Processing (ICCSP), IEEE, 2015, 1743-1747.

[8] Barburiceanu, Stefania, Romulus Terebes, and Serban Meza. "3D Texture Feature Extraction and Classification using GLCM and LBP-Based Descriptors." Applied Sciences 11, no. 5 (2021): 2332.

[9] Zhang, Shu, Jinru Wu, Enze Shi, Sigang Yu, Yongfeng Gao, Lihong Connie Li, Licheng Ryan Kuo, Marc Jason Pomeroy, and Zhengrong Jerome Liang. "MM-GLCM-CNN: A Multi-Scale and Multi-Level Based GLCM-CNN for Polyp Classification." Computerized Medical Imaging and Graphics 108 (2023): 102257.

[10] Rampun, Andrik, Harry Strange, and Reyer Zwiggelaar. "Texture Segmentation using Different Orientations of GLCM Features." In Proceedings of the 6th international conference on computer vision/computer graphics collaboration techniques and applications, 2013, 1-8.

[11] Gustian, Dinda Aulia, Nur Laelatul Rohmah, Guruh Fajar Shidik, Ahmad Zainul Fanani, and Ricardus Anggi Pramunendar. "Classification of Troso Fabric using SVM-RBF Multi-Class Method with GLCM and PCA Feature Extraction." In 2019 International seminar on application for technology of information and communication (iSemantic), IEEE, 2019, 7-11.

[12] Preethi, G., and V. Sornagopal. "MRI Image Classification using GLCM Texture Features." In 2014 international conference on green computing communication and electrical engineering (ICGCCEE), IEEE, 2014, 1-6.

[13] Rafi, Mudassir, and Susanta Mukhopadhyay. "Texture Description using Multi-Scale Morphological GLCM." Multimedia Tools and Applications 77, no. 23 (2018): 30505-30532.

[14] Mu'jizah, Hanimatim, and Dian Candra Rini Novitasari. "Comparison of The Histogram of Oriented Gradient, GLCM, and Shape Feature Extraction Methods for Breast Cancer Classification Using SVM." Jurnal Teknologi dan Sistem Komputer 9, no. 3 (2021): 150-156.

[15] Malegori, Cristina, Laura Franzetti, Riccardo Guidetti, Ernestina Casiraghi, and Riccardo Rossi. "GLCM, An Image Analysis Technique for Early Detection of Biofilm." Journal of Food Engineering 185 (2016): 48-55.

[16] Ali, Syed Sumera Ershad, and Sayyad Ajij Dildar. "An Efficient Quality Inspection of Food Products using Neural Network Classification." Journal of Intelligent Systems 29, no. 1 (2019): 1425-1440.

[17] Zhang, Lei, Junfeng Jing, and Hongwei Zhang. "Fabric Defect Classification Based on LBP and GLCM." Journal of Fiber Bioengineering and Informatics 8, no. 1 (2015): 81-89.

[18] Kabir, Masud, Fatih Unal, Tahir Cetin Akinci, Alfredo A. Martinez-Morales, and Sami Ekici. "Revealing GLCM Metric Variations Across a Plant Disease Dataset: A Comprehensive Examination and Future Prospects for Enhanced Deep Learning Applications." Electronics 13, no. 12 (2024): 2299.