Copyright © The Author(s) 2027
Jurnal Artificial, Vol 5 No 1, April 2027

Comparison of the Surface Roughness of Crushed Stone from Sea and Lansot Based on GLCM and SVM

Pasquila V. Jade [1*], Luther A. Latumakulita [2], Siska Ayu Widiana [3], Winsy Ch. D. Weku [4], Edwin Tenda [5], Rillya Arundaa [6]

[1, 2, 3, 4, 5, 6] Sam Ratulangi University, Indonesia

* pasquilajade106@student.unsrat.ac.id

Abstract

Conventional assessment of crushed-stone surface roughness frequently relies on subjective visual observations, limiting the consistency and quantitative representation of aggregate surface characteristics. This study aims to compare the image-derived surface-texture characteristics of crushed stone from Sea and Lansot, North Sulawesi, using Gray-Level Co-occurrence Matrix (GLCM) features and evaluate the performance of Support Vector Machine (SVM) in classifying surface-texture categories. A quantitative experimental approach was employed using 200 smartphone images, comprising 100 images from each location, equally distributed into Smooth and Rough classes based on a five-indicator visual scoring rubric. Image preprocessing involved resizing to 256 × 256 pixels, RGB-to-HSV conversion, contrast stretching, and object cropping. Six GLCM features were extracted using three pixel distances (1, 3, and 5) and four orientations (0°, 45°, 90°, and 135°). Classification was implemented in Python using scikit-image and scikit-learn, with a 75:25 training–testing split, GridSearchCV, and five-fold cross-validation. The results showed that Lansot samples exhibited higher Contrast, Dissimilarity, and Entropy values but lower Homogeneity, Energy, and Correlation values than Sea samples, indicating greater image-texture variation. Object cropping increased classification accuracy from 50.00% to 96.00%. The optimized Radial Basis Function (RBF)-SVM achieved 93.33% cross-validation accuracy and 96.00% test accuracy, with precision, recall, and F1-score of 96.00%. RBF also outperformed the Linear kernel in repeated validation (91.05% versus 89.20%; p = 0.0026). These findings demonstrate the potential of GLCM–SVM as a quantitative supplementary approach for preliminary aggregate surface-texture characterization, although direct physical roughness validation remains necessary.

Keywords: *Crushed Stone, Gray-Level Co-Occurrence Matrix, Image Classification, Support Vector Machine, Surface Roughness*

Introduction

The characterization of crushed-stone surface texture is essential in construction-material assessment because aggregate morphology influences mechanical interlocking, aggregate–binder interaction, and the performance of concrete and asphalt mixtures. Surface roughness, angularity, and particle geometry can affect material bonding and concrete mechanical behavior (Al-Rousan et al., 2007; Feng et al., 2020; Zhou et al., 2021). However, conventional surface-texture assessment frequently relies on visual observation, which may produce subjective judgments and inconsistent classifications. Differences in observation conditions and assessor experience can reduce reproducibility, emphasizing the need for objective and quantitative characterization techniques (Tafesse et al., 2013).

https://doi.org/10.54065/artificial.5.1.2027.2027

Digital image analysis offers a complementary approach for translating visible aggregate characteristics into measurable texture descriptors.

Aggregate characteristics are influenced by geological origin, mineral composition, particle morphology, and physical structure. Previous studies have demonstrated relationships between aggregate surface texture, concrete compressive strength, and mechanical performance, supporting the importance of aggregate characterization in construction-material evaluation (Güçlüer, 2020; Zhou et al., 2021). Petrographic investigations have further demonstrated the relevance of mineralogical composition and particle morphology to aggregate behavior (Petrounias et al., 2022). In North Sulawesi, crushed stone obtained from Sea, Pineleng Sub-district, Minahasa Regency, and Lansot, Kema Sub-district, North Minahasa Regency, represents two local construction-material sources. The broader North Sulawesi region exhibits complex geological and magmatic characteristics that may contribute to lithological heterogeneity (Zhang et al., 2022). Nevertheless, source-specific geological differences between Sea and Lansot require direct verification. Comparing image-derived texture characteristics may therefore provide preliminary information about differences between aggregates from these locations without assuming a confirmed geological cause.

Conventional characterization techniques include visual inspection, microscopic observation, and contact or non-contact surface measurements. Although direct methods can quantify physical surface irregularities, irregular aggregate geometry may complicate measurement procedures. Image-based methods provide an alternative approach by extracting quantitative morphological information from digital photographs. Previous investigations have demonstrated the applicability of image analysis to aggregate shape, angularity, and surface-texture evaluation (Roy & Kuna, 2023; Zhang et al., 2019). Computer-vision techniques have also been used to characterize coarse-aggregate morphology quantitatively, supporting the potential of digital imaging for material characterization (Liu et al., 2021). However, image-derived texture descriptors should not be interpreted as direct physical roughness measurements because they can be influenced by illumination, image resolution, and acquisition conditions.

Among available texture-extraction methods, the Gray-Level Co-occurrence Matrix (GLCM) is appropriate because it quantifies spatial relationships between neighboring gray-level pixels. Unlike conventional intensity histograms, GLCM incorporates pixel-position relationships, enabling the characterization of local texture variation and spatial dependency (Haralick et al., 1973). Alternative approaches, such as Local Binary Patterns (LBP), describe local intensity-transition patterns, whereas GLCM produces interpretable statistical measurements associated with contrast, uniformity, correlation, and randomness (Ojala et al., 2002; Singh et al., 2019). Comparative investigations of computer-based surface-texture characterization have also demonstrated that GLCM can provide useful information about perceived and physical surface coarseness, although its performance depends on material characteristics and imaging conditions (Ghodrati et al., 2018). Accordingly, this study extracts six GLCM features—Contrast, Dissimilarity, Homogeneity, Energy, Correlation, and Entropy—using pixel distances of 1, 3, and 5 and orientations of 0°, 45°, 90°, and 135° to represent texture characteristics across multiple spatial configurations.

The extracted features are classified using Support Vector Machine (SVM), which identifies separating boundaries between classes and supports nonlinear classification

through kernel functions (Cortes & Vapnik, 1995). The Radial Basis Function (RBF) kernel is particularly relevant when texture-feature relationships cannot be adequately represented by a linear boundary. Previous research has demonstrated the applicability of SVM to image-based limestone classification and granite identification using morphological, color, and texture features (Bianconi et al., 2012; Chatterjee, 2013). GLCM-based features have also been combined with SVM in other image-classification applications (Naveen et al., 2020). Alternative image-classification approaches include convolutional neural networks and transfer learning, which can extract complex visual representations but require appropriate preprocessing, model adaptation, and validation. Previous studies have examined transfer-learning approaches for image classification and demonstrated the application of convolutional neural networks to medical-image recognition, providing methodological examples from domains outside aggregate characterization (Sarkar, 2024; Wafda, 2025). Nevertheless, GLCM–SVM is selected for the present study because it combines interpretable statistical texture descriptors with supervised classification, allowing both descriptive comparisons between aggregate sources and quantitative evaluation of model performance.

Several previous studies provide methodological foundations for the present investigation. Rock-fabric classification using GLCM and Gaussian Mixture Models demonstrated the applicability of spatial texture descriptors to heterogeneous rock structures, although the classification approach and research objectives differed from those used in the present study (Wang & Sun, 2022). Texture-feature extraction combined with machine-learning algorithms has also been applied to rock-grain segmentation, demonstrating the potential of computational texture analysis for identifying geological-material characteristics (Nurzynska & Iwaszenko, 2020). Furthermore, image-based limestone classification using multiclass SVM achieved 96.2% accuracy, providing evidence of the suitability of supervised learning for rock classification, although its feature representation and classification objectives differed from those of crushed-stone surface-texture assessment (Chatterjee, 2013). These studies support the feasibility of digital rock characterization but do not directly investigate the combined comparison of GLCM texture characteristics and binary surface-texture classification for crushed stone obtained from Sea and Lansot.

The research gap therefore concerns the limited integration of source-based crushed-stone texture comparison with supervised classification using multi-distance and multi-orientation GLCM features. Furthermore, classification performance obtained from a particular dataset may not remain stable when the characteristics of new observations differ from those of the training data. Previous machine-learning research has demonstrated the importance of evaluating model stability and generalization under distribution shifts, emphasizing the limitations of relying exclusively on performance measured under a single data distribution (Rayhan et al., 2026). This consideration reinforces the importance of evaluating classification performance while acknowledging the need for external validation under different image-acquisition conditions. The novelty of this study lies in combining six statistical GLCM descriptors extracted across multiple spatial configurations with RBF-based SVM classification, supported by image preprocessing, hyperparameter optimization, and quantitative performance evaluation. Accordingly, this study aims to (1) compare the image-derived surface-texture characteristics of crushed stone from Sea and Lansot using GLCM features and (2) evaluate

the capability of RBF-based SVM to classify visually defined Smooth and Rough categories using accuracy, precision, recall, and F1-score. The findings are expected to provide a reproducible supplementary framework for preliminary aggregate characterization without replacing standardized physical surface-roughness measurements.

Method

This study employed a quantitative experimental approach based on digital image processing to characterize and classify the surface texture of crushed stone obtained from Sea, Pineleng Sub-district, Minahasa Regency, and Lansot, Kema Sub-district, North Minahasa Regency, Indonesia. The research consisted of six stages: image acquisition, class labeling, preprocessing, Gray-Level Co-occurrence Matrix (GLCM) feature extraction, Support Vector Machine (SVM) classification, and performance evaluation. The analysis focused on image-derived surface-texture characteristics rather than direct physical roughness measurements.

Data Collection and Labeling

The target materials were crushed stones originating from the two study locations, while the units of analysis were digital images of individual aggregate samples. The dataset comprised 200 images of crushed stone with particle sizes greater than 4.76 mm, consisting of 100 images from Sea and 100 from Lansot. Each location contributed 50 images classified as Smooth and 50 classified as Rough, resulting in balanced class representation. Images were acquired using a smartphone camera under relatively consistent lighting, camera distance, and object-positioning conditions to minimize variations during texture extraction. The dataset was organized to represent both source locations and surface-texture categories. However, the total population of aggregate particles, the number of unique stones photographed, and a probabilistic sampling framework were not documented; therefore, the findings are limited to the observed image dataset.

Surface-texture labeling was conducted manually using a study-specific composite scoring rubric comprising five visual indicators: (1) surface pores and voids, (2) light reflection, (3) granular texture, (4) fracture edges and angularity, and (5) micro-crack patterns. Each indicator received a binary score of 0 for characteristics associated with smoother surfaces and 1 for characteristics associated with rougher surfaces. The total score ranged from 0 to 5, with scores of 0–2 categorized as Smooth and scores of 3–5 categorized as Rough. Although SNI 7619:2012 was referenced in the original assessment approach, the five-indicator rubric and its classification thresholds were developed specifically for this study and do not constitute a standardized roughness measurement prescribed by that standard.

The labeling results were reviewed through discussion with one experienced crushed-stone worker to assess the plausibility of the assigned categories. This procedure provided practical assessment support but did not constitute formal expert validation or an independent inter-rater reliability test. Consequently, no statistical agreement coefficient was calculated, and potential subjectivity in the reference labels is acknowledged as a methodological limitation.

Image Preprocessing

Image preprocessing was performed to standardize image dimensions and improve the visibility of surface-texture patterns. Each RGB image was resized to 256 × 256 pixels and converted into the Hue, Saturation, and Value (HSV) color space. The Value (V) channel was selected as the grayscale intensity representation for subsequent texture analysis. Contrast stretching was then performed using the 2nd and 98th intensity percentiles to expand the intensity distribution and enhance local texture differences.

The stone object was subsequently cropped to minimize interference from background features that could affect GLCM measurements. The resulting images were used for feature extraction and classification. The contribution of object cropping to classification performance was examined by comparing results obtained from uncropped and cropped images, with the outcomes presented in the Results section.

GLCM Feature Extraction

Surface-texture features were extracted using the GLCM method implemented through the graycomatrix() function in the Python scikit-image library. GLCM matrices were generated using three pixel distances (d = 1, 3, and 5) and four orientation angles (0°, 45°, 90°, and 135°), producing 12 distance–orientation combinations for each image. The matrices were configured as symmetric and normalized to represent spatial gray-level relationships.

Six statistical features were extracted: Contrast, Dissimilarity, Homogeneity, Energy, Correlation, and Entropy. The first five features were calculated using the graycoprops() function, while Entropy was calculated separately using the Shannon entropy formulation. For each image, feature values were averaged across the 12 GLCM configurations, producing a six-dimensional texture-feature vector. The resulting features were standardized using StandardScaler as part of the classification workflow to reduce differences in numerical scale.

SVM Classification and Validation

The extracted feature dataset was divided into 75% training data (150 images) and 25% testing data (50 images) using stratified sampling based on class labels, with random_state = 42 to support reproducibility. SVM classification was performed using the Radial Basis Function (RBF) kernel, which accommodates nonlinear relationships between texture features.

Hyperparameter optimization was conducted using GridSearchCV with 5-fold cross-validation. The evaluated parameter ranges were C = {0.1, 1, 10, 100} and gamma = {0.001, 0.01, 0.1, 1}. The optimal parameter combination was selected based on cross-validation accuracy. To compare kernel performance, the RBF and Linear kernels were additionally evaluated using Repeated Stratified 5-Fold Cross Validation with 10 repetitions, producing 50 paired fold-level evaluations. A paired-samples t-test was used to examine differences in classification accuracy between kernels. Because repeated cross-validation scores are not fully independent, the statistical comparison was interpreted as supporting evidence rather than definitive proof of kernel superiority.

Model Performance Evaluation and Data Analysis

Classification performance was evaluated using accuracy, precision, recall, F1-score, and a confusion matrix. Accuracy represents the proportion of correctly classified samples, whereas precision measures the proportion of predicted positive samples that are correctly classified. Recall indicates the proportion of actual positive samples correctly identified, while F1-score represents the harmonic mean of precision and recall. These metrics were calculated using the following equations:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

$$\text{Precision} = TP / (TP + FP)$$

$$\text{Recall} = TP / (TP + FN)$$

$$\text{F1-score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. Macro-average precision, recall, and F1-score were obtained by averaging the corresponding class-specific values. The confusion matrix was used to identify classification errors between Smooth and Rough samples. In addition, decision scores were obtained using the SVM decision_function() to examine class separation; these scores were interpreted as decision-boundary outputs rather than calibrated probabilities.

For source-based texture characterization, the mean values of the six GLCM features were compared descriptively between Sea and Lansot samples. This analysis was intended to identify differences in image-texture patterns rather than establish statistically significant differences or causal geological relationships. The computational workflow employed Python-based image-processing and machine-learning tools, including scikit-image and scikit-learn. Model effectiveness was assessed through classification metrics, cross-validation, and kernel comparison, while formal product-feasibility testing was not applicable to this image-classification research design.

Results

Effect of Preprocessing

Image preprocessing substantially influenced the performance of GLCM-based surface-texture classification. The SVM model achieved a test accuracy of 50.00% using raw (uncropped) images, whereas accuracy increased to 96.00% after object cropping. This 46-percentage-point improvement indicates that restricting feature extraction to the crushed-stone region was associated with substantially better discrimination between the Smooth and Rough classes. Background textures may introduce irrelevant gray-level relationships that interfere with the representation of stone-surface characteristics. The observed improvement therefore suggests that object isolation is an important component of the classification workflow.

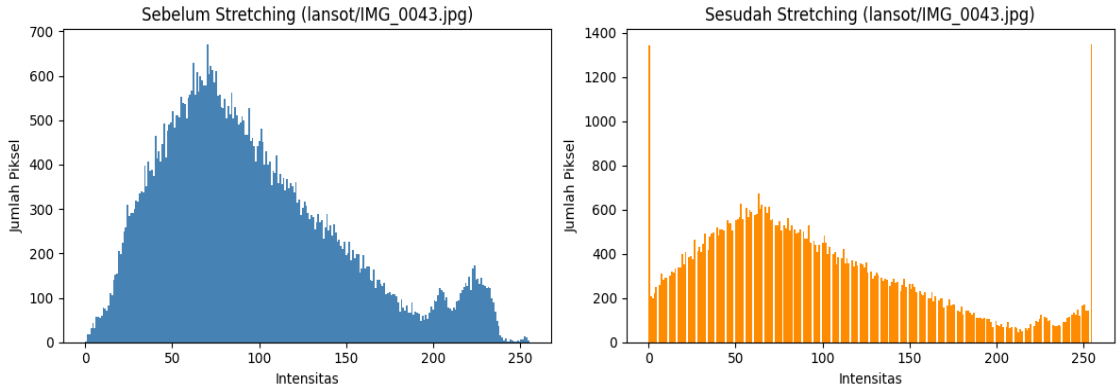


Figure 1 Pixel-intensity histogram before (left) and after (right) contrast stretching on the HSV Value channel of a sample image.

Figure 1 illustrates changes in pixel-intensity distribution following contrast stretching using the 2nd and 98th percentiles. Before preprocessing, intensity values were concentrated within a relatively narrow range around 70, whereas the processed image exhibited a broader distribution approaching the available 0–255 intensity range. This transformation enhanced intensity differences between neighboring pixels, potentially improving the representation of Contrast and Dissimilarity in GLCM calculations. Nevertheless, the separate contributions of cropping and contrast stretching cannot be determined without additional controlled comparisons of these preprocessing operations.

Kernel Comparison

The RBF and Linear kernels were compared using Repeated Stratified 5-Fold Cross Validation with 10 repetitions, producing 50 paired fold-level evaluations. The RBF kernel achieved higher mean performance across all classification metrics. Mean accuracy was 91.05% for RBF and 89.20% for Linear, representing a difference of 1.85 percentage points. A paired-samples t-test yielded $t = -3.1687$ and $p = 0.0026$, indicating a statistically significant difference under the applied testing procedure. The detailed comparison is presented in Table 1.

Table 1. Repeated Stratified 5-Fold Cross Validation Results (RBF vs. Linear)

Metric	RBF (mean ± SD)	Linear (mean ± SD)
Accuracy	91.05% ± 5.05%	89.20% ± 4.94%
Precision	91.49% ± 4.74%	89.94% ± 4.45%
Recall	91.05% ± 5.05%	89.20% ± 4.94%
F1-score	91.01% ± 5.11%	89.12% ± 5.06%

Table 1 shows that the RBF kernel consistently obtained higher mean accuracy, precision, recall, and F1-score than the Linear kernel. The improvement suggests that nonlinear relationships among the extracted GLCM features may contribute to distinguishing the two surface-texture categories. However, the relatively similar standard deviations indicate comparable variability between kernels. Therefore, although RBF performed better in the reported evaluations, the magnitude of its advantage remained relatively modest.

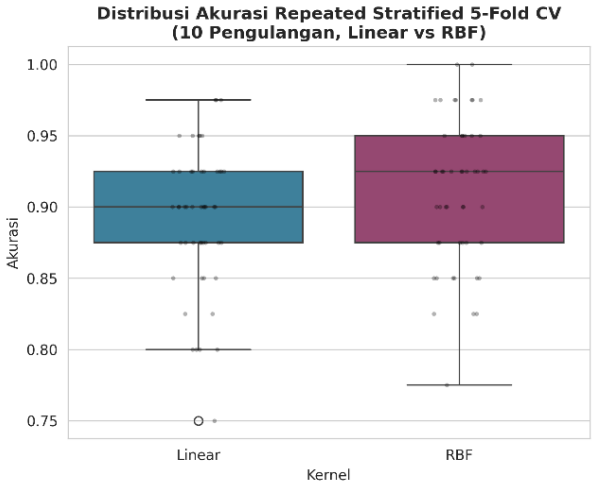


Figure 2. Distribution of accuracy across 50 repeated stratified 5-fold cross-validation runs, RBF vs. Linear kernel.

Figure 2 illustrates the distribution of classification accuracy across repeated validation folds. The RBF kernel exhibited a higher median accuracy of approximately 0.925 compared with 0.900 for the Linear kernel. Although the distributions overlapped, the RBF results were generally shifted toward higher accuracy values, while the Linear kernel included a lower outlier of approximately 0.75. These findings indicate a consistent performance advantage for RBF within the evaluated dataset. Nevertheless, because repeated cross-validation scores are statistically dependent, the paired t-test result should be interpreted cautiously rather than as definitive evidence of universal kernel superiority.

Final Classification Performance

Hyperparameter optimization using GridSearchCV identified C = 1 and gamma = 1 as the optimal parameter combination for the RBF kernel, producing a cross-validation accuracy of 93.33%. Evaluation on the held-out test set containing 50 images yielded an accuracy of 96.00%. Macro-average precision, recall, and F1-score also reached 96.00%, as presented in Table 2.

Table 2. SVM Model Evaluation Results (RBF Kernel, C = 1, gamma = 1) on the Test Set

Metric	Value
Accuracy	96.00%
Precision (macro avg)	96.00%
Recall (macro avg)	96.00%
F1-score (macro avg)	96.00%

Table 2 indicates that the optimized RBF-SVM model achieved balanced classification performance across the evaluated metrics. The similarity between accuracy, precision, recall, and F1-score suggests that high overall accuracy was not accompanied by a substantial imbalance in classification performance between the two categories. However, the evaluation was limited to 50 test images, and model generalizability to other aggregate sources or image-acquisition conditions requires further investigation.

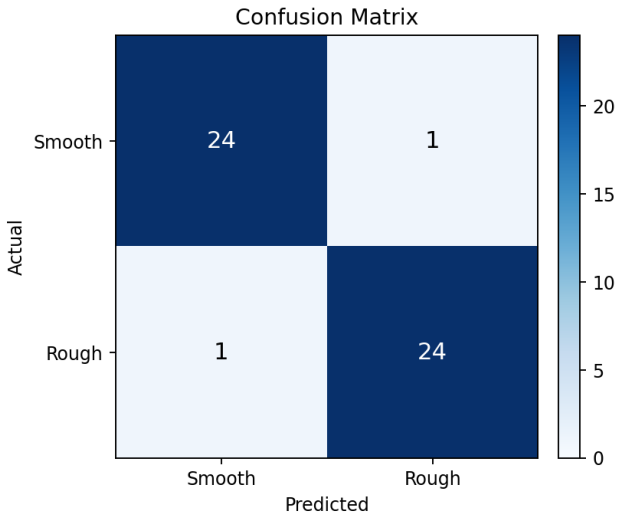


Figure 3. Confusion matrix of the final SVM model on the test set (Smooth vs. Rough).

Figure 3 shows that 24 of the 25 Smooth samples and 24 of the 25 Rough samples were correctly classified. Only two misclassifications occurred, consisting of one Smooth sample predicted as Rough and one Rough sample predicted as Smooth. The corresponding confusion-matrix values were 24 true positives, 24 true negatives, one false positive, and one false negative, with Rough treated as the positive class. This symmetric error distribution indicates no apparent class preference within the test set. Furthermore, the SVM decision function produced a maximum positive score of 1.580106 for a correctly classified Rough sample and a minimum negative score of −1.473807 for a correctly classified Smooth sample. These values indicate separation from the decision boundary for the corresponding observations but should not be interpreted as calibrated classification probabilities.

Descriptive Comparison of GLCM Features Between Sea and Lansot

The six GLCM features were compared descriptively to identify differences in image-derived surface-texture characteristics between crushed-stone samples from Sea and Lansot. The comparison showed that Lansot samples exhibited higher Contrast, Dissimilarity, and Entropy values, whereas Homogeneity, Energy, and Correlation were lower than those observed in Sea samples. The detailed results are presented in Table 3.

Table 3. Descriptive GLCM Feature Values for Crushed-Stone Images from Sea and Lansot

GLCM Feature	Sea	Lansot
Contrast	2431.63	2999.98
Dissimilarity	32.31	39.85
Homogeneity	0.0907	0.0507
Energy	0.0198	0.0131
Correlation	0.760	0.612
Entropy	13.54	14.03

Table 3 demonstrates distinguishable image-texture patterns between the two aggregate sources. Higher Contrast and Dissimilarity values in Lansot samples indicate greater differences between neighboring pixel intensities, while higher Entropy suggests greater randomness in the spatial gray-level distribution. Conversely, higher Homogeneity, Energy, and Correlation values in Sea samples indicate more uniform and spatially

organized intensity relationships. These differences provide quantitative evidence of variation in image-derived texture characteristics. However, because the comparison was descriptive, the results do not establish statistical significance or confirm differences in physical surface roughness.

To integrate the principal findings across preprocessing, texture extraction, classification, and performance evaluation, Table 4 summarizes the quantitative evidence and its technical interpretation. This table provides an integrated assessment appropriate for the quantitative research design rather than conventional qualitative triangulation.

Table 4. Integrated Summary of Quantitative Findings and Technical Interpretations

Analysis Component	Quantitative Findings	Technical Interpretation
Image preprocessing	Accuracy increased from 50.00% to 96.00% after cropping.	Object isolation reduces potentially irrelevant background texture in feature extraction.
Kernel comparison	RBF: 91.05%; Linear: 89.20%; p = 0.0026.	RBF performed better under repeated cross-validation, suggesting possible nonlinear feature relationships.
Final classification	Cross-validation: 93.33%; test accuracy: 96.00%; F1-score: 96.00%.	The optimized GLCM–SVM model classified the visually defined texture categories with high performance on the test dataset.
Confusion matrix	24/25 Smooth and 24/25 Rough samples correctly classified.	Classification errors were balanced between the two categories.
GLCM source comparison	Lansot showed higher Contrast, Dissimilarity, and Entropy; Sea showed higher Homogeneity, Energy, and Correlation.	The aggregate sources exhibited distinguishable image-texture patterns without establishing physical roughness differences.

Table 4 integrates evidence from different analytical stages and demonstrates that classification performance depends on the combined contribution of image preprocessing, texture-feature extraction, and model configuration. The strongest observed performance difference was associated with object cropping, whereas RBF provided a smaller improvement over the Linear kernel. The GLCM comparison further indicates source-related image-texture variations that can be quantified independently of classification metrics. Collectively, these findings support the applicability of digital image analysis as a supplementary method for aggregate characterization, although they do not replace direct material testing.

Discussion

The most important finding of this study is the substantial improvement in classification accuracy following object cropping, from 50.00% to 96.00%. This improvement can be explained by the spatial-statistical characteristics of GLCM, which measures relationships between neighboring pixel intensities. When background information is included, unrelated intensity patterns may influence the extracted features and reduce their ability to represent the actual stone surface. Object cropping restricts feature extraction to the region of interest, potentially improving the discrimination of relevant surface-texture patterns. Previous studies have emphasized the importance of image-analysis procedures in quantitative aggregate characterization and the reduction of subjective morphological assessment (Al-Rousan et al., 2007; Tafesse et al., 2013). Research on material-surface characterization has also demonstrated that image resolution and

texture-extraction conditions can influence the relationship between image-derived measurements and surface characteristics (Ghodrati et al., 2018). Supporting evidence from another image-classification domain indicates that image-quality enhancement, contrast adjustment, and noise reduction can contribute to preparing visual information for feature extraction and classification (Sarkar, 2024). Although that study focused on medical imaging rather than aggregates, it reinforces the methodological importance of preprocessing. Nevertheless, the observed improvement cannot be attributed exclusively to cropping without additional experiments isolating the effects of individual preprocessing operations.

The effectiveness of GLCM is associated with its ability to represent spatial gray-level relationships through statistical descriptors. The multi-distance and multi-orientation approach allows texture characteristics to be examined across different spatial configurations, which is relevant to the heterogeneous appearance of crushed-stone surfaces. Previous investigations have demonstrated that GLCM can characterize differences in digital rock structures and distinguish rock fabrics with varying internal patterns (Singh et al., 2019; Wang & Sun, 2022). Similarly, texture-feature extraction combined with machine learning has been applied to rock-grain segmentation, demonstrating the relevance of computational texture descriptors to geological-material analysis (Nurzynska & Iwaszenko, 2020). However, these studies employed different imaging modalities, classification methods, datasets, and research objectives. The present study extends this analytical approach to the descriptive comparison of locally sourced crushed-stone images and the classification of visually defined surface-texture categories.

The kernel comparison revealed that RBF achieved a higher mean accuracy than the Linear kernel, with values of 91.05% and 89.20%, respectively. The 1.85-percentage-point difference suggests that nonlinear relationships among GLCM features may contribute to separating Smooth and Rough samples. The theoretical ability of SVM to construct nonlinear decision boundaries provides a reasonable explanation for this finding (Cortes & Vapnik, 1995). Previous research on granite-image classification also demonstrated the effectiveness of SVM when model parameters were appropriately optimized (Bianconi et al., 2012). Nevertheless, the improvement obtained using RBF was relatively modest compared with the effect associated with preprocessing, suggesting that the quality of extracted features may have a greater influence on classification performance than kernel selection under the evaluated conditions. Although the paired t-test produced $p = 0.0026$, repeated cross-validation folds are not fully independent. Therefore, the result supports the selection of RBF for the evaluated dataset but does not establish universal superiority over linear classification.

The optimized RBF-SVM model achieved 93.33% cross-validation accuracy and 96.00% test accuracy, with precision, recall, and F1-score also reaching 96.00%. This performance demonstrates the ability of the extracted GLCM features to distinguish the visually defined surface-texture categories within the evaluated dataset. Previous image-based limestone classification achieved 96.2% accuracy using multiclass SVM and a broader combination of image features (Chatterjee, 2013). Although the reported values are numerically similar, direct comparisons are inappropriate because the studies differ in object type, feature representation, dataset size, preprocessing, and validation procedures. Other investigations have demonstrated that combining color and texture features can support natural-stone classification, further indicating that classification performance depends on

the characteristics of the selected visual descriptors (Bianconi et al., 2012). Alternative approaches based on pretrained deep-learning architectures have also been examined for image classification, emphasizing the roles of feature representation, model adaptation, and dataset characteristics in determining classification performance (Wafda, 2025). However, because the present study did not experimentally compare SVM with deep-learning models, its findings should be interpreted as evidence of feasibility within the evaluated dataset rather than proof of superiority over alternative classification methods.

The confusion matrix showed balanced classification performance, with 24 of 25 Smooth samples and 24 of 25 Rough samples correctly classified. However, classification reliability depends on the validity of the reference labels. In this study, the labels were assigned using a five-indicator visual rubric and reviewed through discussion with one experienced crushed-stone worker. Because neither independent inter-rater agreement nor direct physical surface-roughness validation was conducted, the reported accuracy represents agreement with the study-defined visual categories rather than confirmed agreement with standardized physical measurements. This distinction is important because visual texture characteristics and measured physical roughness are related but not necessarily equivalent, particularly under varying image-acquisition conditions (Ghodrati et al., 2018).

The descriptive comparison revealed distinguishable image-texture characteristics between Sea and Lansot. Higher Contrast, Dissimilarity, and Entropy values in Lansot samples indicate greater local gray-level variation and less uniform intensity distributions, whereas higher Homogeneity, Energy, and Correlation values in Sea samples suggest comparatively more consistent spatial relationships. These observations are consistent with the application of digital image analysis to identify variations in coarse-aggregate microtexture and rock morphology (Liu et al., 2021; Roy & Kuna, 2023). Regional geological studies also indicate complex lithological and magmatic characteristics in North Sulawesi, suggesting that variations in material origin may plausibly contribute to differences in aggregate morphology (Zhang et al., 2022). However, this explanation remains hypothetical because the study did not perform mineralogical, petrographic, porosity, or direct physical surface-roughness measurements. Therefore, the observed GLCM differences cannot be attributed conclusively to geological origin.

From an applied perspective, combining digital image preprocessing, multi-distance and multi-orientation GLCM feature extraction, and RBF-based SVM classification provides a potentially accessible method for preliminary aggregate-surface screening. Its principal contribution lies in integrating quantitative texture characterization with supervised classification using interpretable image-derived variables. However, the dataset was limited to 200 images from two locations, and generalization to other materials or image-acquisition conditions remains untested. Research on machine-learning model stability has demonstrated that distribution shifts can alter classification performance and that robustness should be evaluated using complementary metrics rather than relying exclusively on a single performance indicator (Rayhan et al., 2026). This finding is relevant because differences in camera characteristics, lighting conditions, aggregate morphology, and geological sources may introduce variations between training and future application datasets. Moreover, advances in transfer learning provide opportunities to investigate alternative feature-extraction and classification strategies, although their effectiveness for crushed-stone characterization requires independent experimental validation (Wafda,

2025). Future studies should therefore incorporate larger datasets, multiple independent assessors, direct surface-roughness measurements, comparative evaluation of classification algorithms, and external validation across different cameras, lighting conditions, aggregate sizes, and geological sources. The proposed approach should be regarded as a supplementary characterization method rather than a replacement for standardized physical, mechanical, or metrological testing.

Conclusion

This study demonstrates that the integration of Gray-Level Co-occurrence Matrix (GLCM) feature extraction and Support Vector Machine (SVM) classification provides a quantitative image-based approach for characterizing crushed-stone surface textures from Sea and Lansot, North Sulawesi. Addressing the first research objective, Lansot samples exhibited higher Contrast (2999.98), Dissimilarity (39.85), and Entropy (14.03) than Sea samples (2431.63, 32.31, and 13.54, respectively), whereas Homogeneity, Energy, and Correlation were lower, indicating greater local intensity variation and reduced image-texture uniformity. Regarding the second objective, the optimized RBF-SVM model achieved 93.33% cross-validation accuracy and 96.00% test accuracy, with precision, recall, and F1-score reaching 96.00%. Object cropping also improved classification accuracy from 50.00% to 96.00%, demonstrating the importance of isolating relevant image regions. These findings indicate that GLCM–SVM has potential as a rapid and relatively consistent preliminary screening method for aggregate surface-texture characterization, although it cannot replace standardized physical or mechanical testing.

The study was limited to 200 images from two locations, with visually assigned class labels reviewed by only one experienced worker and without direct surface-roughness measurements. Future research should incorporate larger and more geographically diverse datasets, multiple independent assessors, and physical measurements such as profilometry to validate image-derived texture characteristics. Comparative evaluation with alternative classification algorithms and different image-acquisition conditions is also recommended to strengthen model generalizability and practical applicability.

Acknowledgment

-

References

Al-Rousan, T., Masad, E., Tutumluer, E., & Pan, T. (2007). Evaluation of image analysis techniques for quantifying aggregate shape characteristics. *Construction and Building Materials, 21*(5), 978–990. https://doi.org/10.1016/j.conbuildmat.2006.03.005

Bianconi, F., González, E., Fernández, A., & Saetta, S. A. (2012). Automatic classification of granite tiles through colour and texture features. *Expert Systems with Applications, 39*(12), 11212–11218. https://doi.org/10.1016/j.eswa.2012.03.052

Chatterjee, S. (2013). Vision-based rock-type classification of limestone using multi-class support vector machine. *Applied Intelligence, 39*(1), 14–27. https://doi.org/10.1007/s10489-012-0391-7

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273–297. https://doi.org/10.1007/BF00994018

Feng, P., Wang, H., Ding, H., Xiao, J., & Hassan, M. (2020). Effects of surface texture and its mineral composition on interfacial behavior between asphalt binder and coarse aggregate. *Construction and Building Materials, 262*, Article 120869. https://doi.org/10.1016/j.conbuildmat.2020.120869

Ghodrati, S., Gorji Kandi, S., & Mohseni, M. (2018). How accurately do different computer-based texture characterization methods predict material surface coarseness? A guideline for effective online inspection. *Journal of the Optical Society of America A, 35*(5), 712–725. https://doi.org/10.1364/JOSAA.35.000712

Güçlüer, K. (2020). Investigation of the effects of aggregate textural properties on compressive strength (CS) and ultrasonic pulse velocity (UPV) of concrete. *Journal of Building Engineering, 27*, Article 100949. https://doi.org/10.1016/j.jobe.2019.100949

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Liu, Z., Zhang, C., Shao, L., & Wang, J. (2021). Study on quantitative characterization of morphological characteristics and high temperature performance evaluation of coarse aggregate based on computer vision. *Frontiers in Materials, 7*, Article 607105. https://doi.org/10.3389/fmats.2020.607105

Naveen, M., Vidyashankara, M. S., & Kumar, G. H. (2020). Leaf classification based on GLCM texture and SVM. *International Journal of Computer Applications, 177*(35), 18–21. https://doi.org/10.5120/ijca2020919846

Nurzynska, K., & Iwaszenko, S. (2020). Application of texture features and machine learning methods to grain segmentation in rock material images. *Image Analysis and Stereology, 39*(2), 73–90. https://doi.org/10.5566/ias.2186

Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*(7), 971–987. https://doi.org/10.1109/TPAMI.2002.1017623

Petrounias, P., Rogkala, A., Giannakopoulou, P. P., Kalpogiannaki, M., Laskaris, N., & Lampropoulou, P. (2022). The role of the aggregate shape on the compressive strength of concrete using a new micro geo-informatics methodology. *Micron, 161*, Article 103333. https://doi.org/10.1016/j.micron.2022.103333

Rayhan, M. A., Manik, C. F. D., & Putra, A. P. (2026). Analyzing model stability and generalization under distribution shift in real-world machine learning applications. *Dewantara Journal of Technology, 6*(2), 78–89. https://doi.org/10.59563/djtech.v6i2.367

Roy, N., & Kuna, K. K. (2023). Image texture analysis to evaluate the microtexture of coarse aggregates for pavement surface courses. *International Journal of Pavement Engineering, 24*(2), Article 2099854. https://doi.org/10.1080/10298436.2022.2099854

Sarkar, M. R. (2024). Deep learning methods for chest X-ray imaging-based COVID-19 pneumonia detection. *Journal Artificial: Informatika dan Sistem Informasi, 2*(1), 1–9. https://doi.org/10.54065/artificial.487

Singh, A., Armstrong, R. T., Regenauer-Lieb, K., & Mostaghimi, P. (2019). Rock characterization using gray-level co-occurrence matrix: An objective perspective of digital rock statistics. *Water Resources Research, 55*(3), 1912–1927. https://doi.org/10.1029/2018WR023342

Tafesse, S., Robison Fernlund, J. M., Sun, W., & Bergholm, F. (2013). Evaluation of image analysis methods used for quantification of particle angularity. *Sedimentology, 60*(4), 1100–1110. https://doi.org/10.1111/j.1365-3091.2012.01367.x

Wafda, A. (2025). Transfer learning advancements: A comprehensive literature review on text analysis, image processing, and health data analytics. *Journal Artificial: Informatika dan Sistem Informasi, 3*(1), 1–10. https://doi.org/10.54065/artificial.544

Wang, Y., & Sun, S. (2022). A rock fabric classification method based on the grey level co-occurrence matrix and the Gaussian mixture model. *Journal of Natural Gas Science and Engineering, 104*, Article 104627. https://doi.org/10.1016/j.jngse.2022.104627

Zhang, S., Li, R., & Pei, J. (2019). Evaluation methods and indexes of morphological characteristics of coarse aggregates for road materials: A comprehensive review. *Journal of Traffic and Transportation Engineering (English Edition), 6*(3), 256–272. https://doi.org/10.1016/j.jtte.2019.01.003

Zhang, X., Huang, T.-N., Chung, S.-L., Maulana, A., Mawaleda, M., Tien, C.-Y., Lee, H.-Y., & Liu, P.-P. (2022). Late Eocene subduction initiation of the Indian Ocean in the North Sulawesi Arc, Indonesia, induced by abrupt Australian plate acceleration. *Lithos, 422–423*, Article 106742. https://doi.org/10.1016/j.lithos.2022.106742

Zhou, X., Xie, Y., Long, G., & Li, J. (2021). Effect of surface characteristics of aggregates on the compressive damage of high-strength concrete based on 3D discrete element method. *Construction and Building Materials, 301*, Article 124101. https://doi.org/10.1016/j.conbuildmat.2021.124101