Aksaray University
Journal of Science and Engineering

e-ISSN: 2587-1277
http://asujse.aksaray.edu.tr

Aksaray J. Sci. Eng.
Volume 9, Issue 2, pp. **95-99.**

Available online at DergiPark

Research Article

Handcrafted Feature Pipelines for Image Classification: A Comparative Study of HOG, LBP, and FFT within a Transfer-Learning-Style Workflow

Emre Avuçlu[1*], Filiz Sari[2]

[1]*Aksaray University,Faculty of Engineering, Software Engineering Department, Turkey.*
[2]*Aksaray University, Faculty of Engineering, Electrical and Electronics Engineering Department, Turkey.*

Keywords
Transfer Learning, Machine Learning, Feature Extraction, Image Classification, HOG, LBP, FFT

Article information
Received: Oct 30, 2025
Revised: Dec 17, 2025
Accepted: Dec 17, 2025
Online: Dec 22, 2025

doi: **10.29002/asujse.1813481**

Abstract
This study compares three handcrafted feature pipelines HOG + histogram, LBP + statistics, and FFT + edge-density within a transfer-learning-based workflow that employs an SVM classifier. Evaluation is performed on a balanced three-class image dataset (N = 1050; 350 images per class) using stratified 5-fold cross-validation and standard metrics (Accuracy, Precision, Recall, F1, ROC-AUC). Results indicate clear performance differences. The LBP-based pipeline achieves the highest overall accuracy (99.52%) and the most consistent class-wise behavior (macro-AUC ≈ 0.996), reflecting strong texture discrimination. The HOG-based pipeline attains robust performance (93.71% accuracy; macro-AUC ≈ 0.953), especially where edge and shape cues dominate. In contrast, the FFT-based pipeline is less effective overall (76.95% accuracy; macro-AUC ≈ 0.827), with reduced separability for texture-complex or low-contrast images. ROC-AUC analyses corroborate these findings across all classes, confirming the superiority of LBP features in this setting. Collectively, the results clarify when texture-centric, edge-centric, or frequency-centric descriptors are most advantageous and provide empirical guidance for feature selection in transfer-learning-style image classification pipelines, particularly under computational or data constraints.

1. Introduction

Transfer learning has emerged as a pivotal paradigm in computer vision, enabling the adaptation of knowledge from pre-trained models to new domains with limited data [1,2]. While deep convolutional neural networks (CNNs) have demonstrated remarkable success in image classification tasks, the integration of handcrafted feature extraction methods within transfer learning frameworks remains an active area of research [3] . Traditional feature descriptors such as Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and Fast Fourier Transform (FFT) continue to provide valuable insights into image characteristics and can complement deep learning approaches [4-6].

Recent studies have shown that CNN-derived features typically achieve higher classification accuracy than handcrafted methods in most scenarios [7]. However, handcrafted features offer several advantages including computational efficiency, interpretability, and robustness in resource-constrained environments [8]. The comparative analysis of these methods within a unified framework provides insights into their relative strengths and limitations.

Transfer learning has revolutionized computer vision by enabling the reuse of knowledge from pre-trained models. [7] demonstrated that CNN-derived features through transfer learning typically outperformed traditional handcrafted methods in facial emotion recognition tasks. Their study showed that features extracted from intermediate layers of pre-trained networks achieved superior accuracy while requiring less computational time compared to HOG and LBP descriptors.

HOG features capture edge and shape information by computing gradients in localized regions of an image [9]. Recent research has shown that HOG performs particularly well in tasks where shape and edge information are discriminative. LBP performed well in distinguishing paintings from digital images, demonstrating its effectiveness in shape-dominated classification tasks [10].

LBP is a texture descriptor that encodes local spatial patterns by comparing pixel intensities with their neighbors [11]. Studies have consistently shown LBP's effectiveness in texture-rich applications, particularly in facial recognition and medical image analysis [12]. The method's computational efficiency and rotation invariance make it suitable for real-time applications.

Corresponding Author: emreavuclu@aksaray.edu.tr 0000-0002-1622-9059

Articles in this journal is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

FFT-based features capture frequency domain characteristics of images, providing information about periodic patterns and global structures [13]. Research by various authors has shown that frequency descriptors can be particularly effective in specific domains, such as spam detection and certain medical imaging applications, where spectral content provides discriminative information [6,14].

This study addresses the gap in comprehensive comparative analysis of HOG, LBP, and FFT features within a transfer learning context. While previous research has often omitted frequency-domain descriptors or focused on limited datasets, our work provides a systematic evaluation of all three methods using consistent evaluation protocols.

This study has four main objectives to guide the investigation. First, it compares the performance of three widely used feature extraction methods HOG, LBP, and FFT for image classification. Second, each method is assessed using a comprehensive set of performance metrics to ensure a fair and reproducible evaluation. Third, it provides empirical evidence to inform feature selection within transfer learning pipelines, clarifying when handcrafted descriptors can complement or replace learned representations. Lastly, the study highlights the practical strengths and limitations of each approach along with implementation considerations relevant to real-world applications.

2. Materials and Methods

This study emulates the common “feature extraction + classifier” paradigm typically found in CNNs, but employs handcrafted descriptors in place of learned convolutional filters. Three distinct feature extraction models were applied to the same image dataset, and their classification performances were compared using a uniform evaluation protocol. The features of each model were used to train an SVM classifier. The model performance was assessed using 5-fold cross-validation.

This study used a multi-class image dataset containing three distinct categories. The dataset provides sufficient diversity to evaluate the effectiveness of different feature extraction approaches across various image characteristics[15-17]. Sample images of the corn types are shown in Figure 1. The dataset comprised 1050 images evenly distributed across three classes (350 images per class). Stratified 5-fold cross-validation was employed to preserve the class balance within each fold.

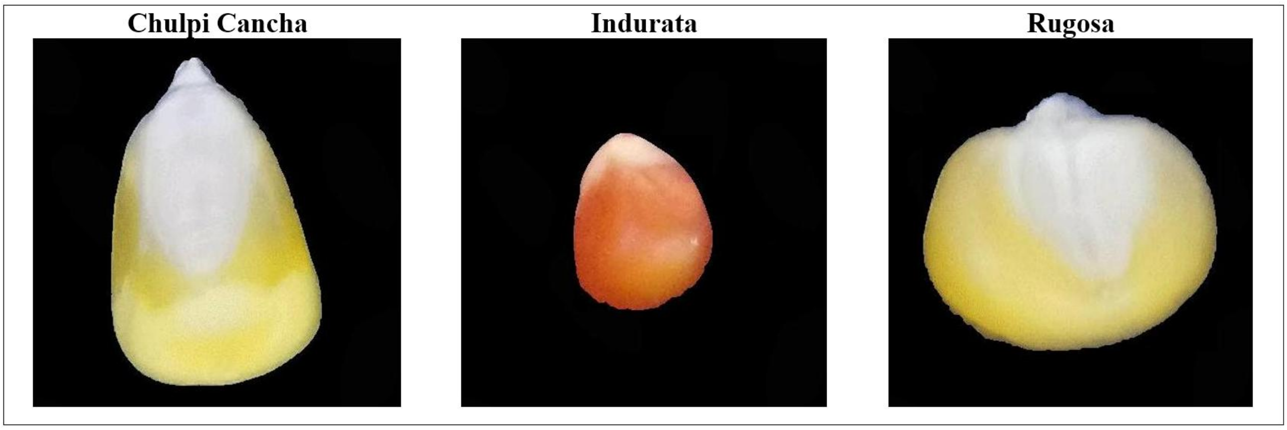

Figure 1. Sample of images

2.1. Feature Extraction Methods

The first model combined HOG features with intensity histograms and was named Model1_HOG. This model concatenates the HOG with global intensity histograms. HOG encodes the local edge-orientation structure, similar to edge-based feature extraction in ResNet-like architectures. This combination provides both structural and intensity-based image representations.

The second model integrates the LBP with statistical measures. LBP features extract local texture patterns, which are analogous to texture-based feature extraction in VGG-style networks. Statistical Features include the mean and standard deviation of the pixel intensities. This approach emphasizes the texture and local statistical properties.

The third model combines frequency-domain analysis with edge detection. The FFT Features capture frequency domain information, similar to spectral analysis in EfficientNet architectures. Edge Ratio Features quantify edge density using Canny edge detection. This model focuses on spectral characteristics and structural complexities.

2.2. Classification and Validation

All extracted features were classified using a Support Vector Machine (SVM) with optimized hyperparameters. The SVM serves as the final classifier, analogous to the fully connected layers in CNN architectures.

The dataset was randomly divided into five folds for cross-validation. Each fold used four parts for training and one part for testing. This process was repeated five times with different fold combinations. The final results represent the average performance across all the folds.

A comprehensive set of metrics is reported to facilitate a fair comparison, given and formulated in Figure 2.

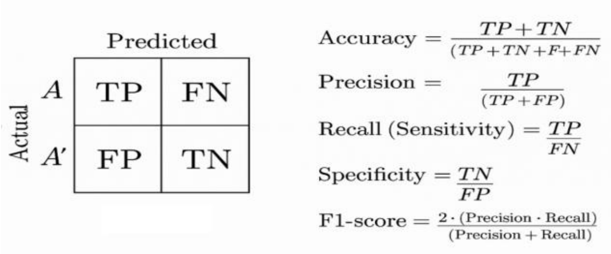

Figure 2. Performance metrics

A workflow summary for corn classification is presented in Table 1.

Table 1. Workflow summary

Stage	Operation	Approximate analogue in a CNN pipeline
1	Image loading and preprocessing	Input layer
2	Feature extraction (HOG / LBP / FFT, etc.)	Convolutional feature extraction
3	SVM training	Fully connected + softmax decision stage
4	5-fold cross-validation	Generalization assessment
5	Metric aggregation (fold averages)	Model evaluation
6	ROC curves and confusion matrices (CM)	Performance visualization

3. Results and Discussion

The novelty of this study lies in its holistic and fair comparison of spatial (HOG, LBP) and frequency-based (FFT) feature extraction methods under a transfer learning simulation approach using the same classifier and dataset conditions. Furthermore, the experimental demonstration that the combination of LBP and statistical features, when used with SVM, significantly increases both accuracy and discriminative power in terms of ROC-AUC represents a unique contribution to this field.

The experimental results revealed clear performance differences among the three feature extraction methods. Figure 3 shows the confusion matrices, and Table 2 summarizes the comprehensive performance metrics for all models. Because the dataset was class-balanced (350 images per class; N = 1050), the macro and micro metrics were comparable and were not confounded by class imbalance.

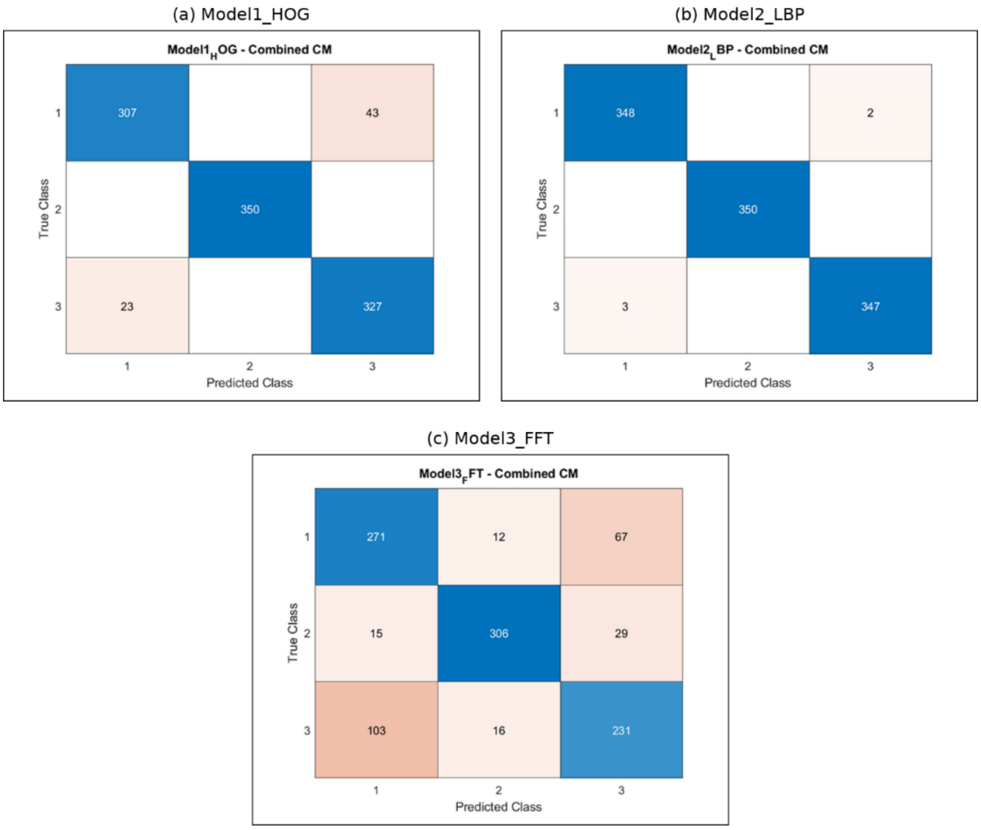

Figure 3. Confusion matrix representation (a) HOG + histogram, (b) LBP + statistics, (c) FFT + edge-density

Table 2. Performance Comparison of Feature Extraction Methods

Metric	Model1_HOG	Model2_LBP	Model3_FFT
Accuracy (%)	93.71	99.52	76.95
Precision (%)	93.80	99.52	77.31
Recall (%)	93.71	99.52	76.95
F1-Score	0.937	0.995	0.770

The LBP-based model exhibited exceptional performance across all metrics. The highest accuracy (99.52%) indicates excellent overall classification capability. The F1-score was 0.995, showing an excellent balance between precision and recall. Consistent performance across different classes indicates a robust feature representation.

The HOG-based model demonstrated a solid performance. It has good accuracy (93.71%) with strong edge sensitivity. It is effective for images with strong edge and shape characteristics.

The FFT-based model exhibited the lowest performance. Moderate accuracy (76.95%) indicates challenges in complex scenarios. Performance limitations in low-contrast and complex-texture scenarios.

The ROC-AUC values provide insights into the discriminative capability of each model across different classes. Figure 4 shows the receiver operating characteristic (ROC) curves, and Table 3 shows the AUC values for each class and model.

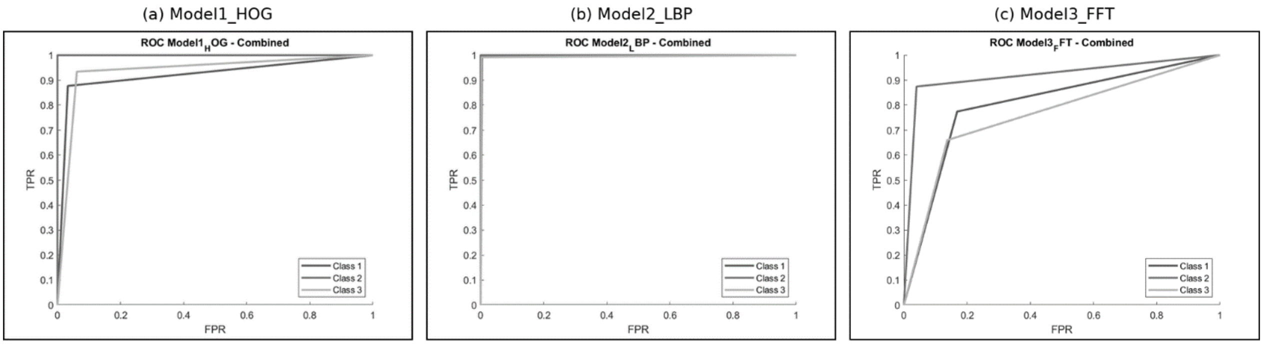

Figure 4. ROC curves of the three models. (a) HOG + histogram, (b) LBP + statistics, (c) FFT + edge density. The sharp structure of the SVM-based ROC curve stems from the fact that the model produces decision function outputs instead of probability outputs, and a limited number of threshold values are formed from these outputs.

ROC analysis further clarified the class-wise discriminative capacity (Figure 4, Table 3). The LBP-based model yielded near-perfect separation across classes, with AUCs exceeding 0.99 and a macro-AUC of 0.996. The HOG-based model achieved an AUC of 1.000 for Class 2 and maintained a competitive performance for the remaining classes (Class 1:0.922; Class 3:0.936), yielding a macro-AUC of 0.953. The FFT-based model displayed greater variability across classes (AUC range: 0.761–0.917) and was notably challenged in Class 3 (AUC = 0.761), with an overall macro-AUC of 0.827.

Table 3. Class-wise AUCs and Macro AUC

Class	Model1_HOG	Model2_LBP	Model3_FFT
Class 1	0.922	0.995	0.803
Class 2	1.000	1.000	0.917
Class 3	0.936	0.994	0.7614
Macro-AUC	0.953	0.996	0.827

The results confirm that texture-based features (LBP) provide more stable and discriminative representations than edge-based (HOG) or frequency-based (FFT) methods for the evaluated dataset.

4. Conclusion

The findings clearly indicate that the LBP-based feature extractor is the most effective model. Model2_LBP exhibited near-perfect separability, with class-wise AUC values exceeding 0.99. The HOG-based model (Model1_HOG) attained high accuracy for edge-salient classes but remained behind LBP on average. The FFT-based model exhibited a comparatively weak performance (76.95% accuracy, AUC ≈ 0.82), with notable information loss in low-contrast images. In terms of F1, the LBP model exhibited strong consistency (0.995). Overall, these results suggest that texture-oriented descriptors (LBP) yield more stable outcomes than classical edge- or frequency-based methods. This study also demonstrates that high accuracy can be achieved within a transfer-learning paradigm without relying on deep learning. This study systematically reveals the crucial role of the feature extraction stage in classification performance, emphasizing that an accurate representation shapes performance regardless of model complexity. Future research examining feature selection/optimization, the impact of classifiers with different kernel functions, and cross-domain generalization capabilities will further strengthen the scope and scientific contribution of the method.

References

[1] Gholizade, M., Soltanizadeh, H., Rahmanimanesh, M., & Sana, S. S. **(2025)**. *A review of recent advances and strategies in transfer learning*. International Journal of System Assurance Engineering and Management, 16(3), 1123–1162. https://doi.org/10.1007/s13198-024-02684-2

[2] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. **(2021)**. *A Comprehensive Survey on Transfer Learning*. Proceedings of the IEEE, 109(1), 43–76. https://doi.org/10.1109/jproc.2020.3004555

[3] Soundarya, B., & Poongodi, C. **(2025)**. *A novel hybrid feature fusion approach using handcrafted features with transfer learning model for enhanced skin cancer classification*. Computers in Biology and Medicine, 190, 110104. https://doi.org/10.1016/j.compbiomed.2025.110104

[4] Tassiopoulou, S., & Koukiou, G. **(2024)**. *Fusing Ground-Penetrating Radar Images for Improving Image Characteristics Fidelity*. Applied Sciences, 14(15), 6808. https://doi.org/10.3390/app14156808

[5] Mame, A. B., & Tapamo, J. R. **(2023)**. *Parameter optimization of histogram-based local descriptors for facial expression recognition*. PeerJ Computer Science, 9, e1388. https://doi.org/10.7717/peerj-cs.1388

[6] Rey-Díaz, A., Martín-Fernández, I., San-Segundo, R., & Gil-Martín, M. **(2024)**. *Frequency Analysis and Transfer Learning Across Different Body Sensor Locations in Parkinson's Disease Detection Using Inertial Signals*. ECSA-11, 82(1), 32. https://doi.org/10.3390/ecsa-11-20507

[7] Tsalera, E., Papadakis, A., Samarakou, M., & Voyiatzis, I. **(2022)**. *Feature Extraction with Handcrafted Methods and Convolutional Neural Networks for Facial Emotion Recognition*. Applied Sciences, 12(17), 8455. https://doi.org/10.3390/app12178455

[8] Jaruenpunyasak, J., & Duangsoithong, R. **(2021)**. *Empirical Analysis of Feature Reduction in Deep Learning and Conventional Methods for Foot Image Classification*. IEEE Access, 9. https://doi.org/10.1109/access.2021.3069625

[9] Dalal, N., & Triggs, B. **(2005)**. *Histograms of oriented gradients for human detection. Proceedings*. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, I. https://doi.org/10.1109/cvpr.2005.177

[10] Liu, L., Zhao, L., Long, Y., Kuang, G., & Fieguth, P. **(2012)**. *Extended local binary patterns for texture classification*. Image and Vision Computing, 30(2), 86-99. https://doi.org/10.1016/j.imavis.2012.01.001

[11] Yan, W., & Dong, Y. (2024). *Local Directional Difference and Relational Descriptor for Texture Classification*. Mathematics, 12(21), 3432. https://doi.org/10.3390/math12213432

[12] Huang, D., Shan, C., Ardabilian, M., Wang, Y., & Chen, L. **(2011)**. *Local binary patterns and its application to facial image analysis: A survey*. IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews (Vol. 41, Issue 6). https://doi.org/10.1109/tsmcc.2011.2118750

[13] Alessio, S. M. **(2016)**. *Digital Signal Processing and Spectral Analysis for Scientists*. Springer International Publishing. https://doi.org/10.1007/978-3-319-25468-5

[14] Novelli, P., Meanti, G., Buigues, P. J., Rosasco, L., Parrinello, M., Pontil, M., & Bonati, L. **(2025)**. *Fast and Fourier features for transfer learning of interatomic potentials*. Npj Computational Materials, 11(1), 293. https://doi.org/10.1038/s41524-025-01779-z

[15] Avuçlu, E., Taşdemir, Ş., & Köklü, M. **(2023)**. *A new hybrid model for classification of corn using morphological properties*. European Food Research and Technology, 249(3), 835–847. https://doi.org/10.1007/s00217-022-04181-x

[16] Avuçlu, E., & Köklü, M. **(2025)**. *Fast and Accurate Classification of Corn Varieties Using Deep Learning With Edge Detection Techniques*. Journal of Food Science, 90(7), e70439. https://doi.org/10.1111/1750-3841.70439

[17] Yasin, E. T., Ropelewska, E., Kursun, R., Cinar, I., Taspinar, Y. S., Yasar, A., Mirjalili, S., & Koklu, M. **(2025)**. *Optimized feature selection using gray wolf and particle swarm algorithms for corn seed image classification*. Journal of Food Composition and Analysis, 145, 107738. https://doi.org/10.1016/j.jfca.2025.107738