Comparative Analysis of Classical and Deep Learning Features for Texture Image Classification

Murad Hossen
Department of Mathematics
University of Houston
Email: mhossen2@cougarnet.uh.edu

***Abstract*—A comparative evaluation of classical statistical features and deep learning-based features extracted from multiple CNN architectures for texture classification on the Describable Textures Dataset (DTD). Classical features form a 22-dimensional vector from GLCM, LBP, Gabor filters, and Haar Wavelets. Deep features are obtained from pretrained models including VGG-16 (4096D), ResNet-18 (512D), EfficientNet-B0 (1280D), and DenseNet-121 (1024D). Hybrid features concatenate classical and deep vectors. Supervised classifiers (SVM, KNN, Random Forest) with hyperparameter tuning are used to assess performance via accuracy and F1-score on a stratified 80/20 split. Results show hybrid features with DenseNet-121 achieving the highest performance (up to 0.723 accuracy with SVM), demonstrating the benefits of fusion for improved robustness in real-world texture tasks. Visualizations like confusion matrices and PCA plots enhance interpretability.**

***Index Terms*—Texture classification, classical features, deep learning features, hybrid fusion, DTD dataset**

I. Introduction

Texture classification is essential in computer vision for applications such as medical imaging and remote sensing. Classical features like GLCM [1] and LBP [2] provide interpretability and efficiency, while deep convolutional neural networks (CNNs) such as ResNet-18 and DenseNet-121 [5] offer superior hierarchical representations. This work compares classical, deep, and hybrid features on the DTD dataset [6], focusing on classification performance and feature separability to identify optimal configurations for high-performance computing contexts.

A. Motivation and Contributions

The DTD dataset's real-world variability challenges feature robustness. Our contributions include: (1) Extraction and evaluation of features from multiple CNNs, (2) Hybrid fusion analysis showing consistent improvements, and (3) Hyperparameter-tuned classification results highlighting DenseNet-121 hybrids as top performers.

II. Methods

A. Dataset Preparation

The DTD dataset [6] comprises 5640 color images across 47 classes (120 per class) with natural variations. Images are preprocessed: grayscale and resized to 200×200 for classical features; RGB, resized to 224×224, and normalized (ImageNet stats) for deep features. A stratified 80/20 split yields 4512 training and 1128 test images, ensuring 96/24 per class.

B. Feature Extraction

Classical Features (22D): Derived from grayscale images: 4 GLCM features (contrast, dissimilarity, homogeneity, correlation [1]), 10 LBP uniform histogram bins [2], 8 Gabor mean/std across 4 orientations [3], and 4 Haar Wavelet means [4].

Deep Features: Extracted from pretrained CNNs on RGB images: - VGG-16: 4096D from fc7 layer. - ResNet-18: 512D from global average pooling [**?**]. - EfficientNet-B0: 1280D from final pooling. - DenseNet-121: 1024D from final pooling [5].

Hybrid Features: Concatenation of classical (22D) with each deep vector, yielding dimensions: 4118D (VGG-16), 534D (ResNet-18), 1302D (EfficientNet-B0), 1046D (DenseNet-121).

C. Classification and Evaluation

Features are normalized using StandardScaler. Classifiers (SVM, KNN, RF) are tuned via GridSearchCV (5-fold CV) on training data. Metrics: accuracy, weighted F1-score, confusion matrices. Best parameters are selected for test evaluation.

D. Visualization

PCA scatter plots and violin distributions (on 1000-sample subset) illustrate feature separability. Comprehensive plots (Fig. 1) compare accuracies and F1-scores across configurations.

III. Results and Discussion

Classification results across feature sets and classifiers are summarized in Table I. Hybrid features consistently outperform standalone classical or deep features, with the DenseNet-121 hybrid achieving the highest accuracy (0.723 with SVM) and F1-score (0.723), highlighting the synergy of classical interpretability and deep robustness. PCA plots (1000 samples) show tighter clusters for hybrids (65-70

PCA plots reveal tighter clusters for hybrids, confirming enhanced separability. Classical features yield low performance due to limited capture of color/variability, while deep features improve it, and hybrids provide marginal but consistent boosts (e.g., 0.025 accuracy gain for DenseNet). This aligns with tactical edge constraints, where efficient hybrids minimize compute needs.

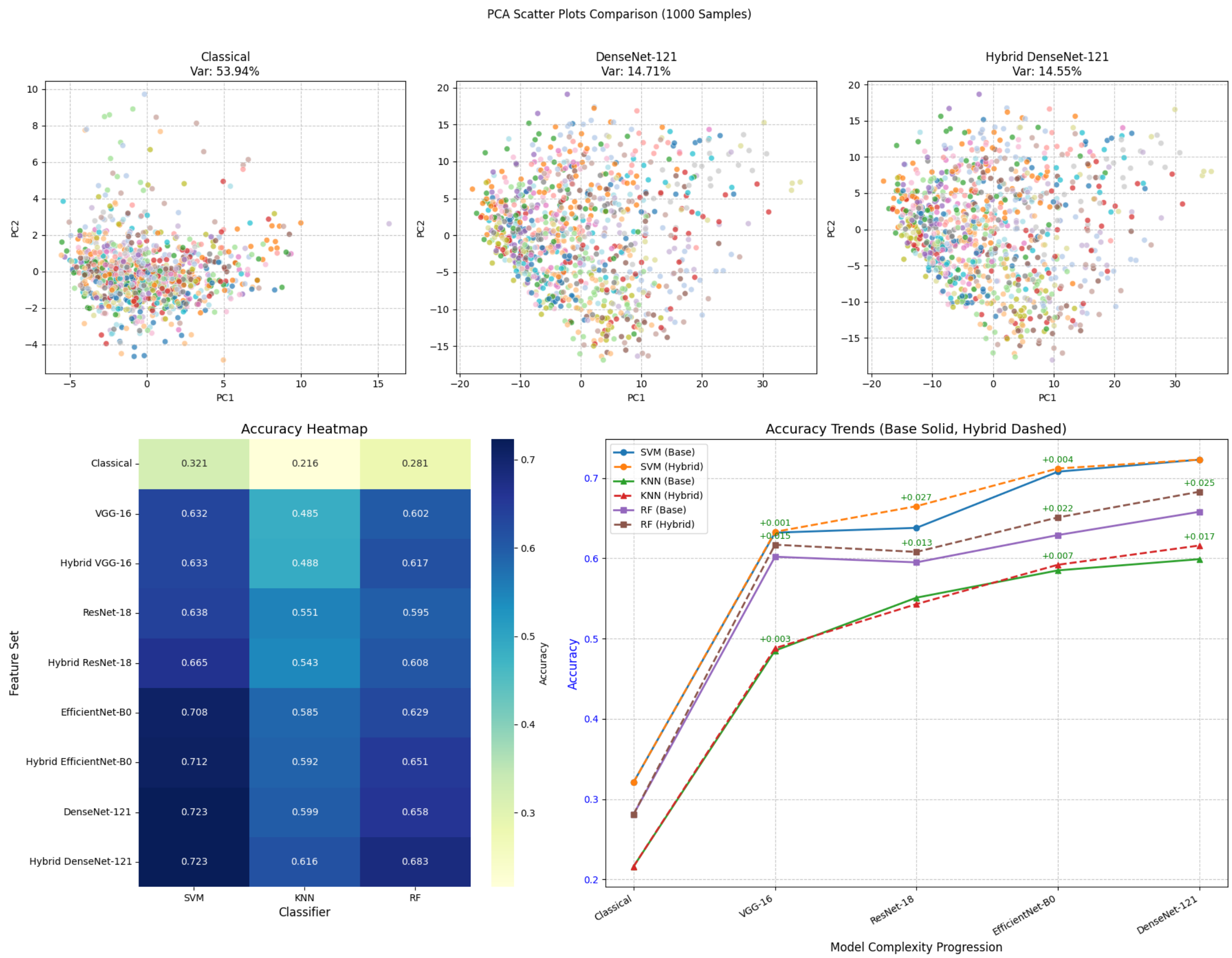

Fig. 1. PCA comparison (1000 samples): Classical, DenseNet-121, Hybrid DenseNet-121 with variance (top); Comprehensive results visualization: Accuracy heatmap (top-left), F1-score heatmap (top-right), Trends in accuracy (bold, primary axis) and F1-score (faint, secondary axis) with base (solid) vs. hybrid (dashed) lines (bottom).

TABLE I
CLASSIFICATION PERFORMANCE (ACCURACY / F1-SCORE) ACROSS FEATURE SETS AND CLASSIFIERS.

Feature Set	**SVM**	**KNN**	**RF**
Classical	0.321 / 0.310	0.216 / 0.213	0.281 / 0.269
VGG-16	0.632 / 0.637	0.485 / 0.487	0.602 / 0.597
Hybrid VGG-16	0.633 / 0.637	0.488 / 0.489	0.617 / 0.614
ResNet-18	0.638 / 0.638	0.551 / 0.545	0.595 / 0.587
Hybrid ResNet-18	0.665 / 0.665	0.543 / 0.539	0.608 / 0.598
EfficientNet-B0	0.708 / 0.711	0.585 / 0.577	0.629 / 0.623
Hybrid EfficientNet-B0	0.712 / 0.714	0.592 / 0.586	0.651 / 0.645
DenseNet-121	0.723 / 0.722	0.599 / 0.599	0.658 / 0.653
Hybrid DenseNet-121	**0.723 / 0.723**	**0.616 / 0.619**	**0.683 / 0.679**

IV. CONCLUSION

Hybrid features, particularly with DenseNet-121, outperform classical and standalone deep features on DTD, achieving up to 0.723 accuracy. Future work includes transformer models and self-supervised learning for further gains in extreme computing environments.

REFERENCES

[1] 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.

[2] Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1), 51–59.

[3] Daugman, J. G. (1988). Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7), 1169–1179.

[4] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674–693.

[5] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4700–4708.

[6] Cimpoi, M., Maji, S., Kokkinos, I., Mohamed, S., & Vedaldi, A. (2014). Describable Textures Dataset (DTD). https://www.robots.ox.ac.uk/ vgg/data/dtd/