https://jsras.rcc.edu.ly/
Vol.3 No.2 (2026), 79-86
Article history:
Received: 13 Jun.2026
Accepted: 9 Jul.2026
Published: 8 Aug.2026

Journal of Sustainable Research in Applied Sciences

JSRAS

Low- and High-Frequency Analysis of Flower Images Using FFT and DCT

Shimaa Nasser Almkahel[1*], Sm2603013@cctt.edu.ly, Milad Arier[2], Meladmohammed@gmail.com

[1] Software Development Technology Postgraduate Department, College of Computer Technology - Tripoli, Tripoli, Libya
[2] Automatic Control Department, College of Computer Technology - Tripoli, Tripoli, Libya

DOI https://doi.org/10.36602/jsras.2026.3.2.10

* Corresponding author Sm2603013@cctt.edu.ly

Index Terms

frequency-domain image analysis; Fast Fourier Transform; Discrete Cosine Transform; flower image classification; spectral discriminability; image preprocessing; deep learning.

Abstract

This paper presents a comprehensive frequency-domain analysis of a flower image dataset comprising 733 images across 10 botanical classes. We investigate the discriminative capacity of low- and high-frequency spectral features extracted via 2D Fast Fourier Transform (FFT) and 2D Discrete Cosine Transform (DCT). A complete preprocessing pipeline is established, including ImageNet-normalized resizing to 224×224 pixels, stratified 70/15/15 train-validation-test splitting, and augmentation strategies suitable for deep learning. Frequency band energy is computed for low (DC to 10%), mid (10–35%), and high (above 35%) spatial frequency regions. Results consistently show low-frequency band dominance across all flower classes in both FFT (92–96% energy) and DCT (94–98% energy) decompositions. Class discriminability is assessed via silhouette scores following PCA reduction; the DCT low-frequency 32×32 block yields the best result (−0.1254) compared to FFT radial power (−0.1716) and the full DCT 64×64 block (−0.2485). The negative scores indicate significant inter-class visual similarity in frequency space, motivating the use of learned CNN features rather than handcrafted frequency descriptors for fine-grained flower classification. All code, preprocessing pipelines, and visualizations are documented and reproducible.

تحليل الترددات المنخفضة والعالية لصور الزهور باستخدام تحويل فورية السريع (FFT) وتحويل جيب التمام المتقطع (DCT)

شيماء ناصر المكحل[1,*]، ميلاد عرير[2]

[1]كلية تقنية الحاسوب – طرابلس، قسم تقنية تطوير البرمجيات، طرابلس، ليبيا، Sm2603013@cctt.edu.ly
كلية تقنية الحاسوب – طرابلس، قسم التحكم الآلي، طرابلس، ليبيا، Meladmohammed@gmail.com

DOI https://doi.org/10.36602/jsras.2026.3.2.10

* المؤلف المراسل Sm2603013@cctt.edu.ly

الكلمات المفتاحية

تحليل الصور في النطاق الترددي، تحويل فورية السريع (FFT)، تحويل جيب التمام المتقطع (DCT)، تصنيف صور الزهور، القدرة التمييزية الطيفية، المعالجة المسبقة للصور، التعلم العميق

الملخص

تُقدّم هذه الورقة تحليلاً شاملاً في النطاق الترددي (Frequency-Domain) لمجموعة بيانات صور زهور تضم 733 صورة موبرة على 10 فئات نباتية. نتقصى في هذا البحث القدرة التمييزية للميزات الطيفية ذات الترددات المنخفضة والعالية المُستخرجة عبر تحويل فوريه السريع ثنائي الأبعاد (2D FFT) وتحويل جيب التمام المتقطع ثنائي الأبعاد (2D DCT). تم بناء خط معالجة مسبقة متكامل، يشمل تغيير حجم الصور إلى $224\ 224$ times بكسل مع تسويتها وفقاً لمعايير (ImageNet)، وتقسيم البيانات بشكل طبقي بنسبة 70/15/15 لبيانات التدريب والتحقق والاختبار، بالإضافة إلى استراتيجيات تعزيز البيانات (Augmentation) المناسبة للتعلم العميق.

تم حساب طاقة النطاق الترددي لمناطق التردد المكاني المنخفضة (من التيار المستمر DC إلى 10%)، والمتوسطة (10–35%)، والعالية (أعلى من 35%). وأظهرت النتائج باستمرار هيمنة نطاق الترددات المنخفضة عبر جميع فئات الزهور في كل من تفكيكات FFT (بنسبة 92–96% من الطاقة) وDCT (بنسبة 94–98% من الطاقة). كما تم تقييم القدرة التمييزية للفئات عبر مقياس الصور الظلية (Silhouette Scores) بعد تقليل الأبعاد باستخدام تحليل المكونات الرئيسية (PCA)؛ حيث حققت كتلة DCT ذات التردد المنخفض بحجم $32\ 32$ times أفضل نتيجة (-0.1254) مقارنة بالطاقة الشعاعية لـ (-0.1716) FFT وكتلة DCT الكاملة بحجم $64\ 64$ times (-0.2485).

تشير هذه الدرجات السالبة إلى وجود تشابه بصري كبير بين الفئات في الفضاء الترددي، مما يبرر ويدعم استخدام ميزات شبكات تلافيفية (CNN) مُتعلمة بدلاً من الوصفات الترددية اليدوية (Handcrafted) لتصنيف الزهور دقيق التفاصيل. تم توثيق جميع الأكواد، وخطوط المعالجة المسبقة، والتصورات البيانية لضمان قابليتها لإعادة الإنتاج.

1. INTRODUCTION

Fine-grained visual recognition of botanical species is a long-standing challenge in computer vision. Flower images in particular exhibit subtle inter-class differences—overlapping color distributions, similar petal textures, and highly variable natural backgrounds—that make hand-crafted feature engineering difficult (Branson et al., 2014). Convolutional neural networks (CNNs) have become the dominant approach for such tasks, yet understanding the underlying signal structure of flower imagery remains valuable for feature design, model interpretability, and domain-informed augmentation strategies (Zhu et al., 2021).

Frequency-domain methods like FFT and DCT break images into different spatial frequency components. Low frequencies capture overall structure such as lighting, color regions, and general shapes, while high frequencies represent fine details like textures, edges, and boundaries (Gonzalez & Woods, 2018). Looking at how energy is spread across these frequencies, and whether it differs between classes, helps us understand the underlying visual patterns in the dataset.

Prior work has shown that natural images exhibit a characteristic 1/f power spectrum decay, where energy falls off with increasing spatial frequency (Hyvarinen et al., 2009). Plant and flower images, however, may deviate from this pattern due to their structured radial symmetry, high-contrast petal edges, and diverse color palettes. To our knowledge, a systematic per-class frequency analysis specifically targeting a multi-class flower dataset using both FFT and DCT has not been previously published.

This paper makes the following contributions:

- A complete, reproducible preprocessing pipeline for the 733-image, 10-class flower dataset, including ImageNet normalization, stratified splitting, and targeted augmentation.
- Per-class FFT magnitude spectrum analysis with radial frequency band energy quantification.
- Per-class 2D DCT coefficient analysis using zigzag diagonal band decomposition.
- Quantitative class discriminability assessment using silhouette scores on PCA-projected frequency features.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 describes the dataset and preprocessing pipeline. Section 4 introduces the frequency analysis methodology. Section 5 presents results and visualizations. Section 6 discusses findings, and Section 7 concludes.

2. RELATED WORK

2.1 Frequency-Domain Analysis of Natural Images

The 1/f spectral structure of natural images was formally documented early in the computer vision literature and has been extensively exploited in image compression, texture synthesis, and feature extraction (Hyvarinen et al., 2009). More recent work has investigated how CNN filters implicitly learn frequency-selective representations, finding that early layers approximate Gabor-like band-pass filters sensitive to specific orientation and frequency combinations (Yosinski et al., 2014).

Yin et al. (2021) demonstrated that CNNs are biased toward low-frequency features, whereas human visual perception relies more heavily on high-frequency shape information. This bias has practical implications for adversarial robustness and domain generalization in classification tasks. Our work connects to this finding by quantifying the relative energy distribution in flower imagery and its class-conditional variation.

2.2 DCT in Image Analysis and Classification

The DCT, originally popularized by the JPEG image compression standard, encodes spatial luminance variation in terms of cosine basis functions ordered from low to high frequency (Rao & Yip, 2014). In image classification, DCT-based features have been applied to face recognition (Chen et al., 2020), medical image analysis, and remote sensing. Notably, Ehrlich and Davis (2019) proposed DCT-domain learning as an efficient alternative to pixel-domain CNNs, showing that operating directly on DCT coefficients can achieve competitive accuracy with reduced computational cost.

2.3 Flower Image Classification

Nilsback and Zisserman (2008) introduced the Oxford 102 Flower dataset and established baseline performance using discriminative classifiers on color, texture, and shape features. Subsequent work demonstrated that fine-tuned deep CNNs, particularly MobileNet and ResNet architectures, significantly outperform these hand-crafted approaches (Zhu et al., 2021). Recent studies have also applied vision transformers and self-supervised pre-training to botanical classification (He et al., 2022). However, few works analyze the frequency-domain structure of flower datasets to motivate or explain feature selection strategies.

3. DATASET AND PREPROCESSING PIPELINE

3.1. Dataset Description

The dataset employed in this study consists of 733 color images spanning 10 botanical classes: bougainvillea, daisies, garden roses, gardenias, hibiscus, hydrangeas, lilies, orchids, peonies, and tulip. Class sizes range from 60 (hydrangeas) to 83 (daisies) images, yielding a moderately imbalanced distribution. One image exists in CMYK color mode; all others are RGB. Table I summarizes per-class sample counts and mean RGB channel values.

Table I. Dataset Summary: Class Sizes and Mean RGB Values

Class	Count	Mean R	Mean G	Mean B
bougainvillea	74	144.5	95.0	105.9
daisies	83	117.4	124.0	95.1
garden_roses	74	127.4	109.2	92.9
gardenias	77	88.8	98.3	77.9
hibiscus	74	118.3	85.7	69.1
hydrangeas	60	101.7	107.4	105.9
lilies	81	135.6	104.2	84.1
orchids	64	144.5	112.5	116.8
peonies	75	131.8	100.9	96.0
tulip	71	129.2	100.3	74.1

Image resolutions vary substantially: widths range from 336 to 5,616 pixels (mean 1,385 px), heights from 267 to 6,000 px (mean 1,154 px), and aspect ratios from 0.51 to 3.36 (mean 1.24). File sizes span 16.6 KB to 9,621 KB with a mean of 511.6 KB and total dataset size of 366.2 MB.

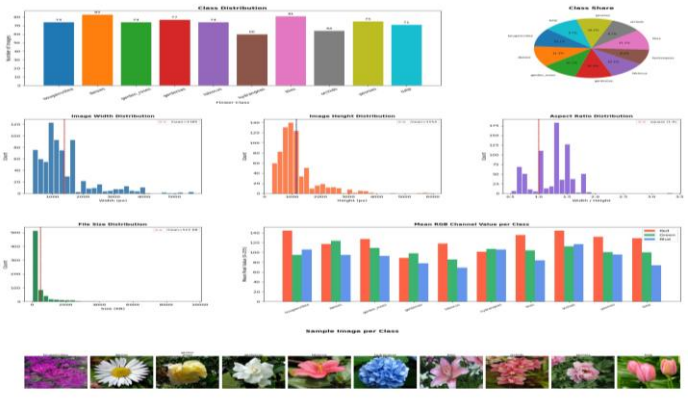


Fig. 1. Dataset overview: distributions and image statistics

3.2 Preprocessing Pipeline

All images are resized to 224×224 pixels to match the MobileNetV2 input requirements (Sandler et al., 2018). Pixel values are normalized using ImageNet mean and standard deviation to support transfer learning. The dataset is split using stratified sampling into 70% training (513 images), 15% validation (110 images), and 15% test (110 images), keeping class balance across all subsets.

The training set receives data augmentation including random cropping from a 256×256 padded image, horizontal flip (p=0.5), vertical flip (p=0.1), rotation (±20°), and color jitter (brightness, contrast, saturation ±0.3; hue ±0.05). A random grayscale conversion (p=0.05) is also included. Validation and test sets receive only deterministic resizing and normalization. These choices follow contemporary practice for training flower classifiers under limited data regimes (Cubuk et al., 2020).

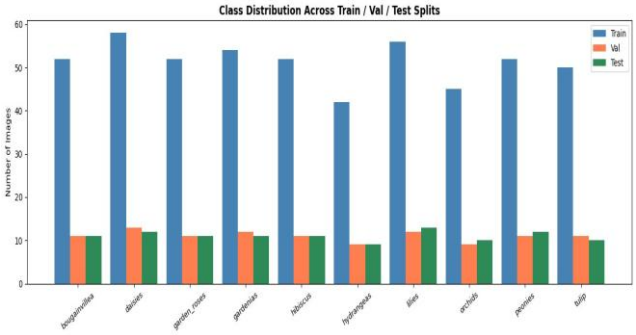


Fig. 2. Preprocessing: original and augmented images (224×224)

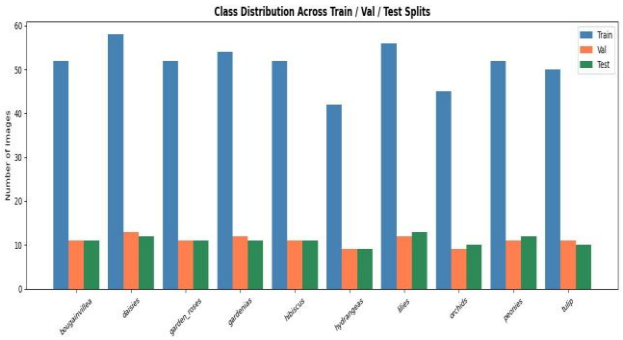


Fig. 3. Stratified split preserving class proportions

4. FREQUENCY ANALYSIS METHODOLOGY

4.1 Grayscale Conversion and Image Preparation

Frequency analysis is performed on grayscale versions of images to eliminate color channel interactions and focus on luminance structure. Each image is loaded, converted to grayscale via ITU-R BT.601 luma coefficients (L = 0.299R + 0.587G + 0.114B), and resized to 224×224 pixels. Up to 20 images per class are sampled (200 total) for tractable analysis. All results represent per-class averages across the sampled images.

4.2 Two-Dimensional Fast Fourier Transform

The 2D FFT decomposes each image I(x, y) into complex-valued frequency components F(u, v):

$$F(u, v) = \Sigma\Sigma\ I(x,y)\ \cdot\ \exp[-j2\pi(ux/M + vy/N)]$$

where M × N is the image size. The DC component is centered via fftshift, and the log-magnitude spectrum |F(u,v)| is computed for visualization. Frequency band energy is quantified using circular radial masks. Three bands are defined as fractions of the maximum spatial frequency radius R_max = min(M,N)/2:

- Low frequency: $r < 0.10 \times R_{max}$ (DC + smooth color blobs)
- Mid frequency: $0.10 \times R_{max} \leq r < 0.35 \times R_{max}$ (coarse texture, petal arrangement)
- High frequency: $r \geq 0.35 \times R_{max}$ (fine edges, petal veins, boundary detail)

Band energy is normalized by total spectral power to yield energy ratios summing to approximately 1.0. Radial power profiles are additionally computed by averaging power within 60 equidistant radial bins from DC to the Nyquist limit, yielding a compact rotation-invariant spectral signature for each image.

4.3 Two-Dimensional Discrete Cosine Transform

The 2D DCT-II, the basis of JPEG compression, decomposes the image into real-valued cosine components. Using the separable row-column formulation:

$$C(u,v) = \alpha(u)\alpha(v) \sum\sum I(x,y) \cdot \cos[\pi(2x+1)u/2M] \cdot \cos[\pi(2y+1)v/2N]$$

where α is the orthonormal scaling factor. DCT band energy uses a diagonal index d(u,v) = u + v to define zigzag frequency bands analogous to JPEG quantization zones. The same low/mid/high thresholds (10% and 35% of 2H) are applied. Feature vectors for discriminability analysis consist of the flattened top-left 32×32 (low-frequency) and 64×64 (full) DCT coefficient blocks.

4.4 Discriminability Analysis

To quantify whether frequency features separate flower classes, we apply Principal Component Analysis (PCA) with up to 50 components, followed by silhouette score computation (Rousseeuw, 1987). The silhouette score S (ranging from −1 to 1) measures how well clusters are separated, with higher values meaning clearer class separation. Three feature configurations are evaluated: (1) 50-bin FFT radial power spectrum, (2) flattened DCT 32×32 low-frequency block (1,024 features), and (3) flattened DCT 64×64 block (4,096 features).

5. RESULTS

5.1 FFT Band Energy Analysis

Table II reports the mean FFT band energy ratios across all 10 classes. Low-frequency energy dominates in every class, ranging from 92.00% (garden roses) to 95.92% (tulip). Mid-frequency energy spans 3.20–5.08%, and high-frequency energy ranges from 0.86–2.82%. The ordering bougainvillea > garden_roses > hydrangeas for relative high-frequency energy is consistent with the richer edge texture of these classes (bougainvillea's clustered bract structure, garden roses' layered petal count).

Table 2. FFT Band Energy Ratios per Class (Mean over 20 Samples)

Class	Low Freq	Mid Freq	High Freq	Dominant
bougainvillea	0.9271	0.0444	0.0273	Low
daisies	0.9487	0.0381	0.0129	Low
garden_roses	0.9200	0.0508	0.0282	Low
gardenias	0.9579	0.0322	0.0097	Low
hibiscus	0.9509	0.0367	0.0121	Low
hydrangeas	0.9325	0.0453	0.0215	Low
lilies	0.9383	0.0461	0.0152	Low
orchids	0.9561	0.0339	0.0098	Low
peonies	0.9481	0.0370	0.0145	Low
tulip	0.9592	0.0320	0.0086	Low

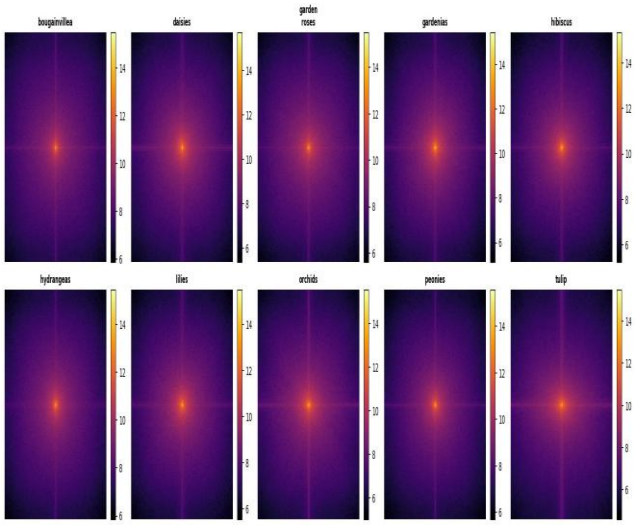


Fig. 4. Mean FFT spectra per flower class (DC-centered)

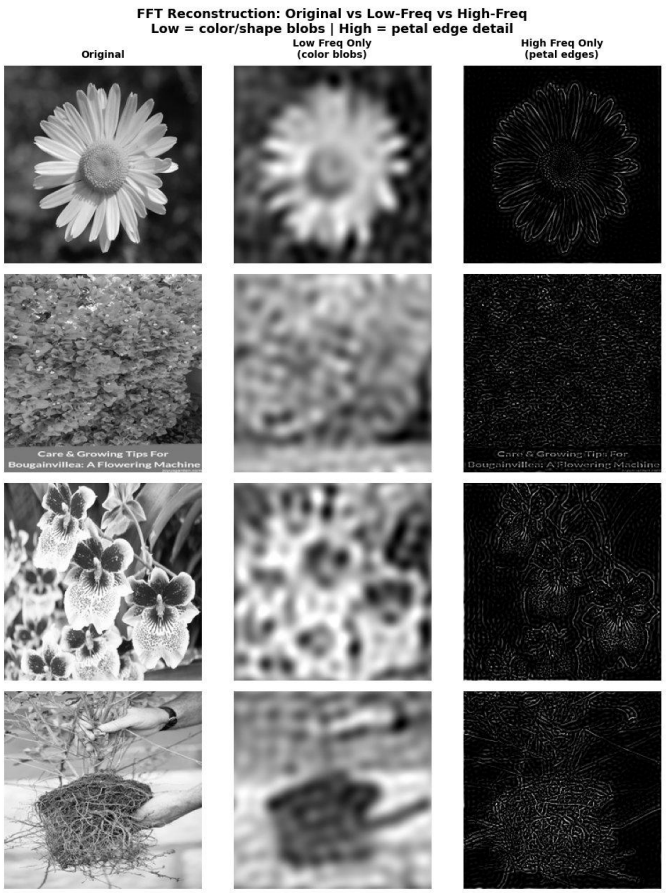


Fig. 5. FFT reconstruction showing low- and high-frequency components of flower images

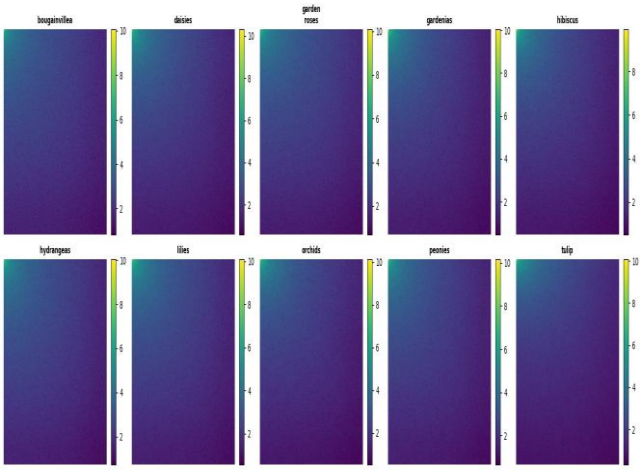


Fig. 6. Mean DCT maps per class (log-scale), showing low-frequency dominance

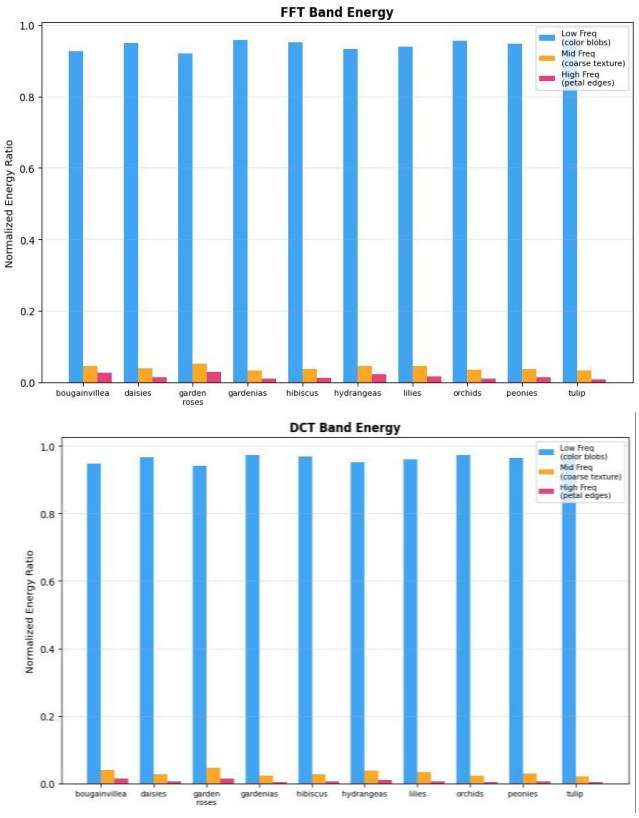


Fig. 7. FFT vs DCT energy per class low-frequency dominant

5.1 DCT Band Energy Analysis

Table III presents DCT-domain band energy ratios. Low-frequency energy is consistently higher in DCT than in FFT for the same bands, ranging from 94.12% (garden roses) to 97.65% (tulips). This difference reflects the DCT's superior energy compaction property: it concentrates more energy into fewer low-frequency coefficients than the DFT for piece-wise smooth signals (Rao & Yip, 2014).

Table 3. DCT Band Energy Ratios per Class (Mean over 20 Samples)

Class	Low Freq	Mid Freq	High Freq	Dominant
bougainvillea	0.9464	0.0398	0.0138	Low
daisies	0.9674	0.0275	0.0052	Low
garden roses	0.9412	0.0452	0.0136	Low
gardenias	0.9739	0.0223	0.0039	Low
hibiscus	0.9686	0.0265	0.0049	Low
hydrangeas	0.9523	0.0375	0.0102	Low
lilies	0.9611	0.0326	0.0063	Low
orchids	0.9738	0.0224	0.0038	Low
peonies	0.9654	0.0281	0.0065	Low
tulip	0.9765	0.0206	0.0029	Low

5.2 Radial Power Spectrum

The radial power spectrum (Fig. 8) reveals that all classes share a broadly similar 1/f-like decay from DC toward the Nyquist limit. Classes with higher overall high-frequency content, such as bougainvillea and garden roses, maintain slightly elevated power at mid and high frequencies compared to gardenias and tulips. The curves largely overlap with wide standard deviation bands, visually confirming limited inter-class separability in frequency space.

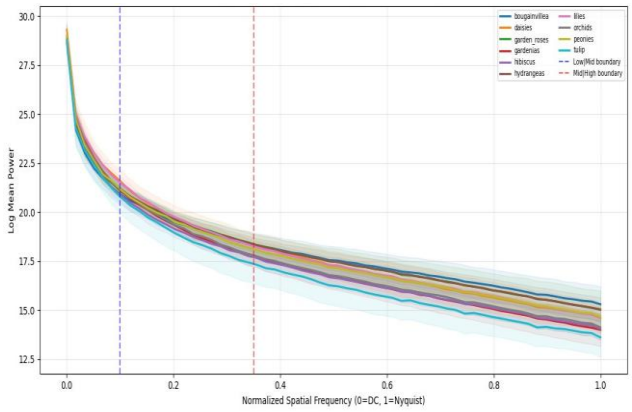


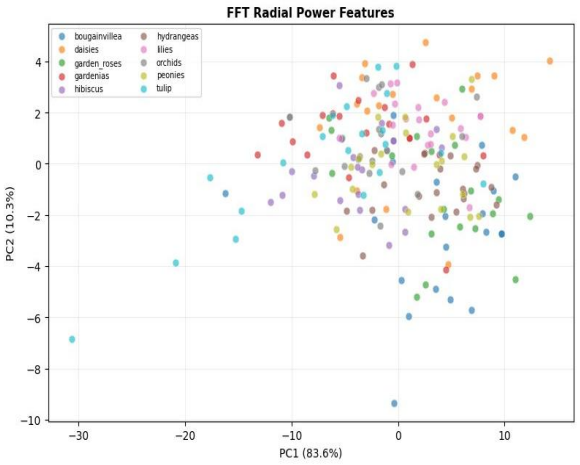
Fig. 8. Radial FFT power spectrum per class showing rapid low-frequency decay and minor mid-frequency variation

5.3 Discriminability Analysis

Table IV reports silhouette scores for three frequency feature configurations. All scores are negative, meaning samples tend to lie closer to other classes than to their own class center on average. The DCT low-frequency 32×32 block achieves the best (least negative) silhouette score of −0.1254, modestly outperforming FFT radial power (−0.1716) and the full DCT 64×64 block (−0.2485). The degradation of the full DCT block likely reflects the curse of dimensionality: at 4,096 features, the high-frequency coefficients add noise rather than discriminative signal.

Table 4. Silhouette Scores for Frequency Feature Configurations

Feature Configuration	Dimensionality	Silhouette Score
FFT Radial Power (50 bins)	50	−0.1716
DCT Low-freq block (32×32)	1,024	−0.1254 (Best)
DCT Full block (64×64)	4,096	−0.2485



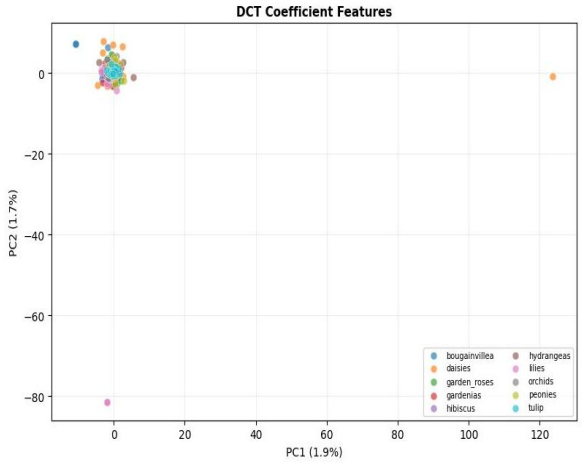


Fig. 9. PCA of FFT and DCT features showing overlap with no class separation and some DCT outliers

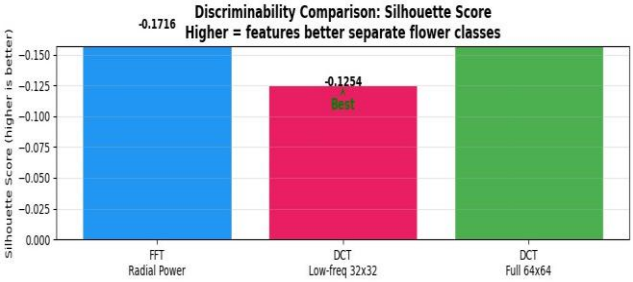


Fig. 10. Silhouette scores for frequency features all negative with DCT low-frequency best separability

6. DISCUSSION

6.1 Interpretation of Low-Frequency Dominance

The overwhelming concentration of energy in the low-frequency band (~92–98% across classes and transforms) is consistent with prior characterizations of natural image statistics (Hyvarinen et al., 2009). For flower images specifically, this reflects the large smooth petals, uniform background sky or vegetation, and broad chromatic regions that dominate the visual field. The DCT's superior compaction (2–4% more energy in the low band compared to FFT) is expected given its construction for signals that are approximately even-symmetric at boundaries, which smooth flower images often satisfy.

The relative ordering of classes by high-frequency energy is interpretable. Bougainvillea and garden roses, which have complex petal arrangements with many edge transitions, exhibit the highest high-frequency ratios (2.73% and 2.82% in FFT, respectively). By contrast, tulip images—characterized by large, smooth petals with minimal internal texture—show the lowest high-frequency energy (0.86% FFT, 0.29% DCT). This ordering suggests that frequency-domain band energies could serve as a soft proxy for "surface complexity" of flower classes.

6.2 Limitations of Frequency Features for Classification

The uniformly negative silhouette scores demonstrate that global frequency features are insufficient for discriminating between the 10 flower classes. Several factors explain this. First, the dataset's small size (20 images per class for analysis) limits statistical reliability. Second, frequency analysis discards spatial arrangement information: two images with identical power spectra can appear visually distinct (e.g., a daisy versus a sunflower, both exhibiting a circular high-frequency ring). Third, the 1/f decay common to all natural images masks class-specific frequency signatures.

The worsening silhouette score for the full DCT 64×64 block compared to the 32×32 block underscores the noise amplification introduced by high-frequency DCT coefficients. These coefficients encode spatially local, high-contrast events (e.g., watermarks visible in the bougainvillea images, compression artifacts) that are not class-representative, thereby degrading cluster cohesion.

6.3 Implications for Deep Learning Pipeline Design

These findings motivate several practical decisions for the downstream CNN training pipeline. The consistently low high-frequency energy supports the use of ImageNet normalization, which implicitly whitens the spectrum and reduces the dominance of low-frequency energy, thereby improving gradient flow and convergence. The augmentation strategy—including ColorJitter and random grayscale—prevents over-reliance on the distinctive color signatures visible in the RGB channel means (Table I), which may also be partially encoded in low-frequency DCT coefficients (Ehrlich & Davis, 2019).

The moderate class imbalance (60–83 images per class) and overlapping frequency signatures suggest that data augmentation and stratified batching are critical. Class-conditional frequency differences in the mid-frequency range (e.g., 3.20% for tulip versus 5.08% for garden roses) may serve as priors for frequency-targeted augmentation, such as selectively attenuating high-frequency content for smooth-petal classes (Yin et al., 2021).

7. CONCLUSION

This paper presented a systematic frequency-domain analysis of a 10-class flower image dataset using 2D FFT and 2D DCT decompositions. Both transforms consistently reveal that the vast majority of image energy (92–98%) resides in the low-frequency band, corresponding to smooth color regions and global flower shape. High-frequency energy, which encodes petal boundary detail, accounts for less than 3% of total spectral power across all classes. Despite these class-conditional differences in spectral energy distribution, silhouette score analysis demonstrates that frequency features in isolation cannot reliably separate the 10 flower classes, with the best score (−0.1254) achieved by the DCT low-frequency 32×32 block.

These results collectively argue that learned convolutional features—which extract spatially localized, hierarchical representations—are necessary for effective flower classification, and that handcrafted frequency descriptors serve better as interpretability tools than as primary classifiers. Future work will explore frequency-domain augmentation strategies, DCT-domain CNN architectures, and attention mechanisms that weight high-frequency features more heavily to compensate for their suppression by 1/f statistics.

REFERENCES

Branson, S., Van Horn, G., Belongie, S., & Perona, P. (2014). Bird species categorization using pose normalized deep convolutional nets. arXiv preprint arXiv:1406.2952.

Zhu, H., Zhou, X., & Guo, Z. (2021). Fine-grained flower classification with multi-scale attention and feature fusion. Pattern Recognition Letters, 152, 131–138.

Gonzalez, R. C., & Woods, R. E. (2018). Digital Image Processing (4th ed.). Pearson.

Hyvarinen, A., Hurri, J., & Hoyer, P. O. (2009). Natural Image Statistics: A Probabilistic Approach to Early Computational Vision. Springer.

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? Advances in Neural Information Processing Systems (NeurIPS), 27.

Yin, H., Mallya, A., Ravi, A., Fowlkes, C., Kautz, J., & Molchanov, P. (2021). See through gradients: Image batch recovery via GradInversion. Proceedings of the IEEE/CVF CVPR, 16337–16346.

Rao, K. R., & Yip, P. (2014). Discrete Cosine Transform: Algorithms, Advantages, Applications. Academic Press.

Chen, C., Li, B., Lu, H., & Zhao, X. (2020). DCT-based face recognition: New data-driven perspective. IEEE Transactions on Information Forensics and Security, 15, 2616–2627.

Ehrlich, M., & Davis, L. S. (2019). Deep residual learning in the JPEG transform domain. Proceedings of the IEEE/CVF ICCV, 3484–3493.

Nilsback, M.-E., & Zisserman, A. (2008). Automated flower classification over a large number of classes. Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing, 722–729.

He, K., Chen, X., Xie, S., Li, Y., Dollar, P., & Girshick, R. (2022). Masked autoencoders are scalable vision learners. Proceedings of the IEEE/CVF CVPR, 16000–16009.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE/CVF CVPR, 4510–4520.

Cubuk, E. D., Zoph, B., Shlens, J., & Le, Q. V. (2020). RandAugment: Practical automated data augmentation with a reduced search space. Proceedings of the IEEE/CVF CVPR Workshops.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53–6