International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(9s) 2026 pp. 530-567
https://doi.org/10.70917/ijcisim-2026-3455
Article

A Novel Frequency-Domain Feature Extraction Framework for Machine Vision Applications

Sandeep Bagul, Aarti Sharma[1]

University, Indore (Madhya Pradesh), 452020-India
Department of IET – AS Sage University, Indore (Madhya Pradesh), 452020-India.

Abstract: This research addresses the challenge of brain tumor classification in MRI, specifically distinguishing low-grade gliomas (LGG) from high-grade gliomas (HGG). It presents a frequency-domain feature extraction framework that incorporates advanced preprocessing, optimization-based segmentation, and deep learning to enhance classification accuracy. Key stages include: mean filtering for noise reduction, Kapur's multilevel thresholding for tumor region segmentation, and a 2D Discrete Fourier Transform (DFT) for robust feature extraction. Optimized features are processed through a dual deep learning architecture, combining EfficientNetB3 for semantic feature extraction and Convolutional Neural Networks (CNNs) for preserving spatial hierarchies. The framework achieved over 95% classification accuracy on MRI datasets of LGG and HGG, demonstrating significant improvements over baseline methods. Its computational efficiency and innovative features offer a decision-support tool for clinicians in neuro-oncology, enhancing treatment planning and prognosis assessment.

Keywords: Fourier Transform, Frequency-domain feature extraction, Brain tumor classification, EfficientNetB3, Convolutional Neural Networks, Mean filtering, MRI imaging, LGG/HGG classification, Deep learning

1. Introduction

1.1 Overview of Image Processing

Image processing techniques are systematically classified into two primary domains: spatial domain methods and frequency domain methods. Spatial domain techniques manipulate pixel values directly within the image plane, employing operations such as convolution, morphological operations, and intensity transformations. In contrast, frequency domain methods transform images into spectral representations, manipulate the Fourier coefficients, and subsequently transform the modified spectrum back to the spatial domain to generate enhanced images. The selection between these approaches depends upon the specific application requirements, computational constraints, and desired outcomes. Some advanced enhancement strategies operate on combinations of these methodologies to achieve optimal results.

The significance of image processing in machine vision applications has expanded exponentially with the proliferation of data-driven approaches. Recent developments emphasize the importance of learning robust feature representations that capture both local details and global context across multiple domains. Additionally, the computational efficiency of image processing algorithms has become increasingly critical in edge computing environments, where processing must occur closer to data sources to minimize latency and conserve bandwidth, particularly in IoT and real-time systems requiring instant visual data processing.

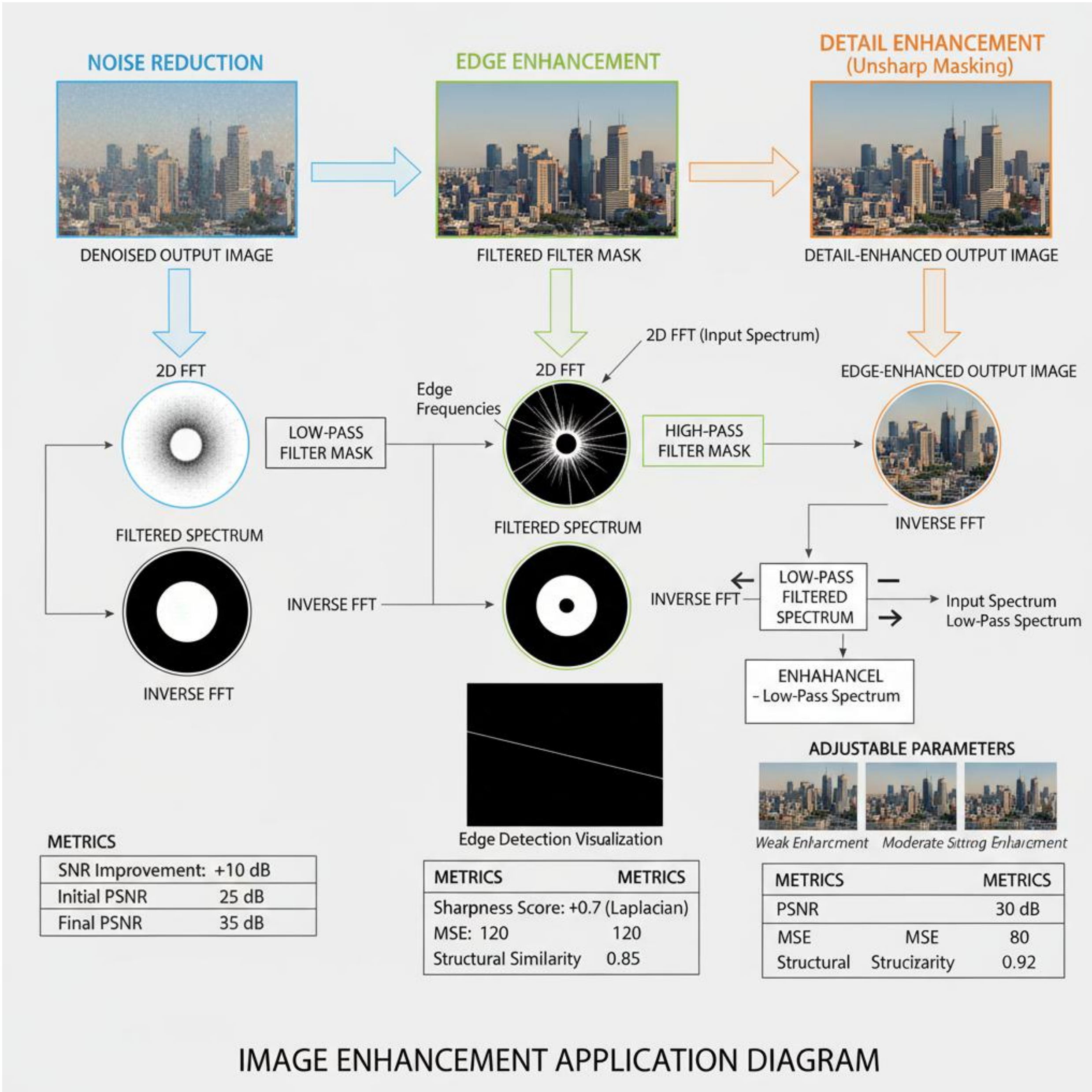

Figure 1: Image Enhancement Applications: Noise Reduction, Edge Enhancement, Detail Extraction

1.2 Role of Fourier Transform in Frequency-Domain Analysis

The Fourier Transform enables the characterization of images by their spatial frequency content, serving as the natural domain for analyzing space-invariant linear processes. A fundamental property of Fourier analysis is that Fourier bases function as eigenfunctions of all space-invariant linear operators; consequently, when complex exponentials are subjected to any linear, space-invariant operator, they remain complex exponentials of the same frequency but with potentially modified amplitude and phase characteristics. This mathematical elegance makes frequency domain analysis particularly suitable for designing filters and understanding linear image transformations.

In frequency domain representation, the origin of the transform corresponds to low-frequency components that describe slowly varying features and overall image structure. As the analysis moves away from the origin, progressively higher frequencies encode increasingly rapid gray-level changes and fine spatial details within the image. This spatial-frequency relationship provides intuitive insight into filter design: low-frequency components can be attenuated to reduce blur and enhance edges, while high-frequency components can be suppressed to reduce noise.

Chakraborty et al. (January 2026) demonstrated that complex-valued convolutional neural networks operating entirely in the frequency domain outperform traditional spatial domain approaches, particularly when preserving phase information through properly designed activation functions. Their work on fully complex-valued residual CNNs revealed that operating directly on frequency domain representations of k-space (Fourier frequency domain used in MRI acquisition) data, utilizing magnitude and phase components simultaneously, achieves superior performance with reduced computational overhead compared to methods requiring domain transitions through FFT/IFFT operations.

The frequency domain has proven invaluable for numerous applications beyond traditional image enhancement. In medical imaging, Tang et al. (2025) presented rigorous mathematical foundations for understanding Fourier transforms in magnetic resonance imaging, emphasizing the critical role of frequency domain analysis in signal reconstruction. Furthermore, Liu et al. (2024) introduced faster Fourier convolution networks that expand receptive fields by leveraging the entire frequency spectrum according to the spectral convolution theorem, enabling superior capture of long-range dependencies and redundant information in multi-coil k-space data.

Recent advances in frequency-domain feature extraction have demonstrated remarkable improvements in machine vision applications. Han et al. (2024) developed self-organizing broad networks incorporating frequency-domain analysis parameters to achieve representative feature extraction superior to traditional randomization approaches, validating effectiveness across diverse engineering applications including sunspot prediction and water quality monitoring. Additionally, Lin et al. (2025) proposed frequency-guided dual-collapse Transformers combining spatial domain processing with frequency-domain guidance through mixed residual fast Fourier transform blocks, achieving state-of-the-art low-light image enhancement results.

1.3 Importance of Mathematical Transformations in Image Enhancement

Jenifera et al. (2025) demonstrated the combined utility of multiple mathematical domains in audio signal classification, extracting time domain features (amplitude envelope, crest factor) and frequency domain features (spectrogram, spectral centroids) through a 2D lightweight CNN achieving 93.75% accuracy on the GTZAN dataset. Their work emphasized that leveraging strengths of both time and frequency domain analysis, capturing both local and global patterns simultaneously, produces superior classification results compared to single-domain approaches.

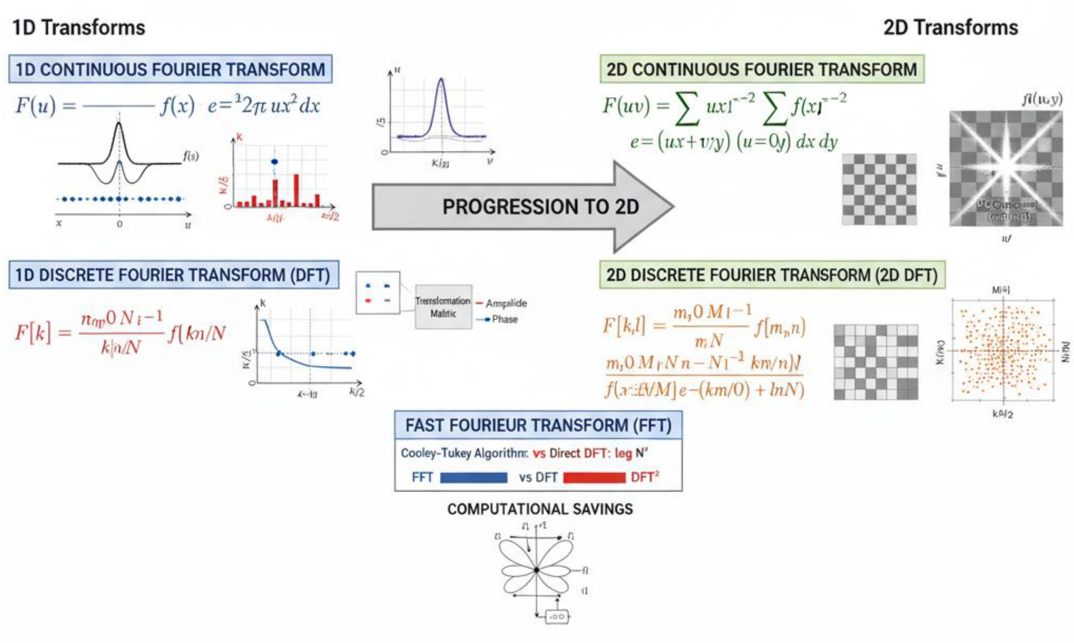

Figure 2: Mathematical Formulation: 1D to 2D Fourier Transform and FFT Algorithm

In magnetic resonance imaging, Oh et al. (2024) developed domain transformation learning networks that map directly from undersampled k-space input (frequency domain) to final images (spatial domain), effectively learning the mathematical transformation inherent in inverse Fourier transforms while simultaneously performing reconstruction. This approach, termed dual-input ETER-net, incorporates both frequency and image domain information as supplementary inputs, achieving superior reconstruction performance across multiple acceleration factors through intelligent exploitation of dual-domain characteristics.

The wavelet domain has emerged as particularly significant for image processing applications requiring simultaneous spatial and frequency localization. Tong et al. (2022) proposed hybrid image-wavelet domain reconstruction networks (HIWDNet) for fast MRI reconstruction, employing distinct reconstruction network architectures optimized for different transform domains rather than applying identical operations across domains. Their results demonstrated superior performance on SSIM and PSNR metrics, validating that domain-specific architectural designs substantially improve enhancement outcomes.

For medical image analysis and segmentation, researchers have integrated frequency domain representations directly into convolutional neural network architectures. Han et al. (2023) developed hybrid domain feature learning modules based on windowed fast Fourier convolution, combining global features with wide-range receptive fields in frequency domain with local features across multiple scales in spatial domain. This hybrid domain approach outperformed spatial-domain-only methods on medical image classification tasks, confirming the complementary value of diverse mathematical domain representations.

Wu et al. (2020) advanced understanding of graph Fourier transforms for analyzing multivariate neuroimaging signals, applying graph signal processing frameworks to functional MRI data using structural connectivity graphs as supports. Their approach achieved superior classification accuracy compared to functional connectome-based methods, illustrating how extending Fourier analysis concepts to graph structures enables more sophisticated representations of complex biological data.

The mathematical rigor and computational efficiency enabled by transformation frameworks make them indispensable for addressing contemporary challenges in image enhancement. Integration of multiple complementary transforms—combining properties of frequency domain analysis, wavelet localization, and graph-based representations—continues to drive advancement in machine vision applications, enabling more sophisticated extraction of meaningful features and more effective enhancement of image quality.

1.4 Limitations of Spatial Domain Techniques

While spatial domain methods offer computational efficiency and intuitive pixel-level manipulation, they encounter fundamental limitations that restrict their effectiveness for complex image enhancement tasks, particularly when dealing with global features, noise reduction, and feature extraction in machine vision applications. Understanding these limitations provides essential motivation for complementary frequency-domain and hybrid approaches.

Spatial domain techniques operate by directly manipulating pixel values through local operations defined by convolution kernels and neighborhood processing. A primary limitation of these methods is their inherently restricted receptive field: spatial domain convolutional operations, even with large kernel sizes, cannot simultaneously capture information from distant spatial locations without significant computational overhead. In contrast, frequency-domain operations naturally integrate global information through Fourier coefficients that each depend upon all spatial locations in an image. This fundamental difference creates inherent constraints for spatial-only approaches in capturing long-range spatial dependencies critical for understanding global image structure.

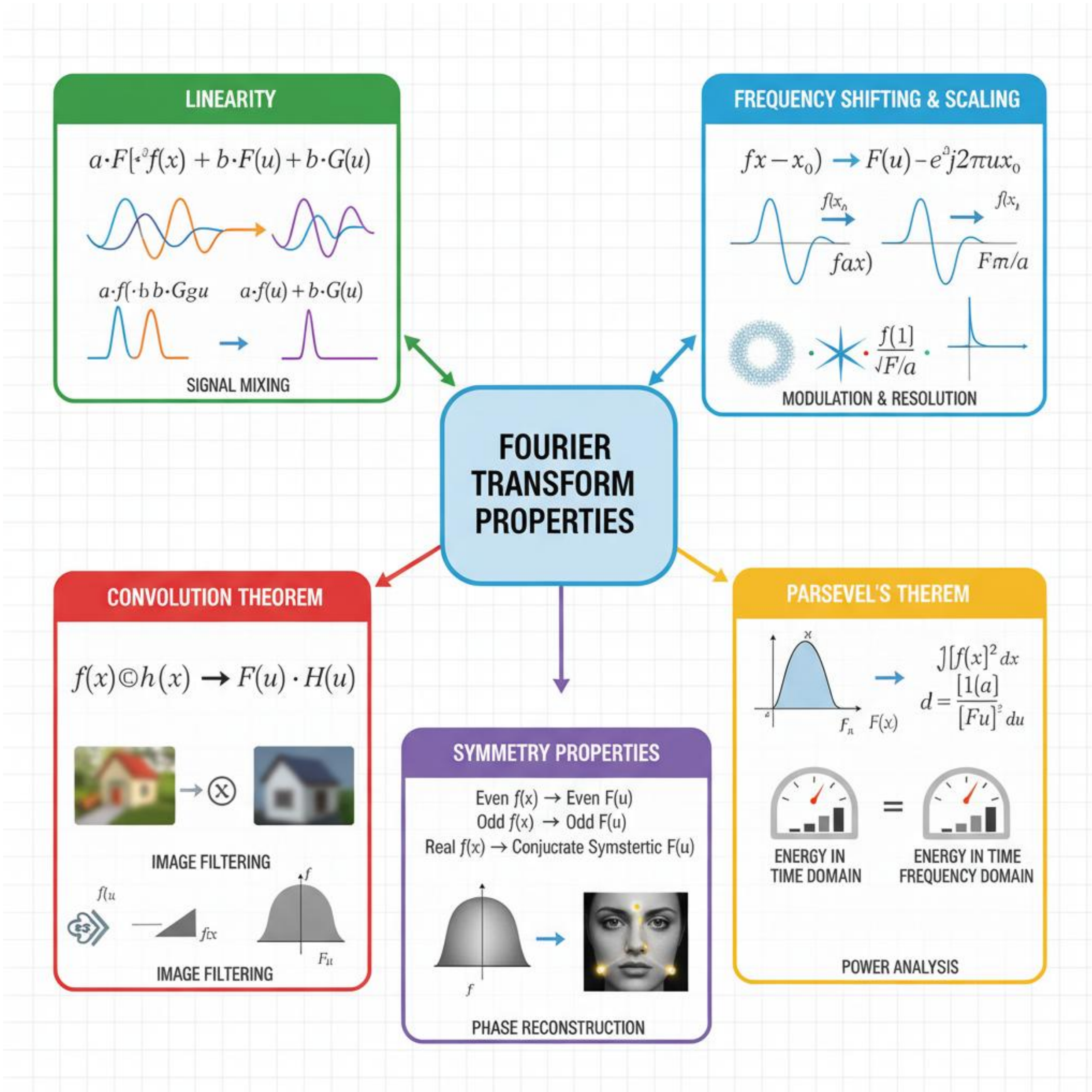

Figure 3: Fourier Transform Properties: Linearity, Convolution, Shifting, Parseval, Symmetry

Spatial domain histogram equalization, though foundational in image enhancement, demonstrates significant limitations in preserving image structure while enhancing contrast. Global histogram equalization distributes gray levels but can enhance noise and fail to reveal fine details within locally homogeneous regions. Local histogram equalization using small neighborhoods can reveal fine structures but introduces undesirable checkerboard artifacts and noise enhancement. These limitations stem from spatial domain methods' inability to distinguish between important structural details and artifacts through spectral analysis.

Han et al. (2024) demonstrated that self-organizing broad networks incorporating frequency-domain parameter calculation achieved superior prediction performance compared to traditional spatial domain randomization approaches, reducing redundant feature extraction and enabling more representative feature learning. Their results highlighted that frequency-domain analysis enables extraction of truly representative sample features that pure spatial domain methods cannot efficiently identify, particularly for complex nonlinear systems exhibiting multiscale behavior.

Malathi et al. (2024) addressed noise reduction limitations of spatial domain techniques through FFT-based approaches combined with deep learning, achieving 97% quality enhancement accuracy compared to spatial-only methods. Their work illustrated that spatial domain denoising using techniques like discrete wavelet transforms and morphological operations, while computationally faster, cannot match the selective noise reduction capabilities of frequency domain filtering that can preserve edge information while removing noise across frequency bands.

The limited interpretability of spatial domain feature representations presents additional constraints for machine vision applications. Spatial domain pixel-level features often fail to capture semantic meaning, requiring extensive feature engineering or deep learning to extract meaningful representations. Yang et al. (2025) demonstrated that efficient 3D MRI reconstruction through 2D transformers with attention-based fusion could reduce computational demands compared to 3D U-Net spatial domain processing while maintaining superior reconstruction quality by leveraging transformer architectures that incorporate frequency information implicitly.

Furthermore, spatial domain techniques exhibit poor performance when confronted with aliasing artifacts and reconstruction challenges in undersampled or compressed data. Traditional spatial domain regularization methods cannot effectively distinguish between meaningful image structure and aliasing artifacts that manifest globally. The inverse Fourier transform understanding, mathematically central to solving inverse imaging problems, cannot be efficiently approximated through pure spatial convolution operations without explicitly incorporating frequency-domain insights.

These fundamental limitations—restricted receptive fields, computational inefficiency for global feature capture, poor noise selectivity, limited interpretability, and inadequacy in handling inverse problems—collectively underscore the necessity for complementary frequency-domain and hybrid approaches in modern machine vision applications. The integration of spatial and frequency domain techniques addresses these constraints systematically, enabling more robust, efficient, and interpretable image enhancement and feature extraction frameworks.

2. Background and Mathematical Foundations

2.1 Mathematical Formulation of 1D and 2D Fourier Transform

The Fourier Transform represents one of the most fundamental mathematical tools in signal processing and image analysis, enabling the decomposition of signals into their constituent frequency components. The transform exists in multiple forms depending on whether the input signal is continuous or discrete, and whether it operates in one or two dimensions. Understanding these distinctions is essential for applying Fourier analysis to machine vision applications.

2.1.1 Continuous Fourier Transform

The Continuous-Time Fourier Transform (CTFT) operates on continuous signals and provides the theoretical foundation upon which digital implementations are based. For a continuous 1D signal $f(x)$, the CTFT is mathematically defined as:

$$F(u) = \int_{-\infty}^{\infty} f(x)\, e^{-j2\pi ux} dx$$

where u represents the frequency variable, and $j = \sqrt{-1}$ is the imaginary unit. The inverse Continuous Fourier Transform (ICFT) that recovers the spatial domain signal from its frequency domain representation is given by:

$$f(x) = \int_{-\infty}^{\infty} F(u)\, e^{j2\pi ux} du$$

Using Euler's formula, $e^{-j2\pi ux} = \cos(2\pi ux) - j\sin(2\pi ux)$, these transforms can be decomposed into real and imaginary components, where the cosine terms represent even symmetry and sine terms represent odd symmetry components of the signal.

For two-dimensional continuous signals $f(x, y)$, the 2D Continuous Fourier Transform extends naturally to:

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, e^{-j2\pi(ux+vy)} dx\, dy$$

The corresponding 2D inverse transform is:

$$f(x, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u, v)\, e^{j2\pi(ux+vy)} du\, dv$$

These formulations establish the fundamental relationship between spatial and frequency domains for continuous signals. The continuous transform serves as the theoretical basis, though digital image processing requires discrete implementations.

2.1.2 Discrete Fourier Transform (DFT)

Digital image processing operates on sampled signals, necessitating the Discrete Fourier Transform (DFT). For a 1D discrete sequence $f[n]$ where $n = 0,1, \dots, N-1$, the DFT is defined as:

$$F[k] = \sum_{n=0}^{N-1} f[n]\, e^{-j2\pi kn/N}, k = 0,1, \dots, N-1$$

The inverse DFT (IDFT) that recovers the spatial domain sequence from frequency domain samples is:

$$f[n] = \frac{1}{N}\sum_{k=0}^{N-1} F[k]\, e^{j2\pi kn/N}, n = 0,1, \dots, N-1$$

For digital images, the 2D DFT operates on an $M \times N$ image matrix $f[m, n]$ where $m = 0,1, \dots, M-1$ and $n = 0,1, \dots, N-1$. The 2D DFT is defined as:

$$F[k, l] = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[m, n]\, e^{-j2\pi(km/M+ln/N)}, k = 0,1, \dots, M-1; l = 0,1, \dots, N-1$$

The corresponding 2D inverse DFT is:

$$f[m, n] = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} F[k, l]\, e^{j2\pi(km/M+ln/N)}$$

A critical computational consideration is that the DFT requires $O(N^2)$ complex multiplications to compute all N frequency components. For an image with $M \times N$ pixels, direct DFT computation requires approximately M^2N^2 operations. This computational complexity becomes prohibitive for large images, motivating the development of faster algorithms.

2.2.1 Linearity

The linearity property states that the Fourier Transform of a weighted sum of signals equals the weighted sum of their individual Fourier Transforms. For continuous signals, this property is expressed as:

$$\mathcal{F}\{af_1(x) + bf_2(x)\} = aF_1(u) + bF_2(u)$$

where a and b are arbitrary constants, and $f_1(x)$, $f_2(x)$ are continuous signals with Fourier Transforms $F_1(u)$, $F_2(u)$ respectively.

For discrete signals, the linearity property is similarly stated:

$$\mathcal{F}\{af_1[n] + bf_2[n]\} = aF_1[k] + bF_2[k]$$

Application in Image Enhancement: Linearity enables the decomposition of complex image enhancement operations into separate frequency-domain operations that can be applied independently. For instance, contrast enhancement through frequency-domain filtering can be decomposed into independent operations on different frequency bands, with the final enhanced image obtained through linear combination of individual processing results.

2.2.2 Convolution Theorem

The convolution theorem establishes a fundamental relationship between spatial domain convolution and frequency domain multiplication, making it essential for understanding filter design and implementation. For 1D continuous signals, the convolution theorem states:

$$\mathcal{F}\{f(x) * h(x)\} = F(u) \cdot H(u)$$

where $*$ denotes convolution, and $\cdot$ denotes pointwise multiplication. The inverse relationship is:

$$\mathcal{F}\{f(x) \cdot h(x)\} = \frac{1}{2\pi}[F(u) * H(u)]$$

This duality between convolution and multiplication is profound: a computationally expensive convolution operation in the spatial domain can be replaced by efficient multiplication in the frequency domain.

For 2D signals, the convolution theorem extends to:

$$\mathcal{F}\{f(x,y) * h(x,y)\} = F(u,v) \cdot H(u,v)$$

For discrete signals and circular convolution, the discrete convolution theorem is expressed as:

$$\mathcal{F}\{f[n] \circledast h[n]\} = F[k] \cdot H[k]$$

where $\circledast$ denotes circular convolution.

Computational Efficiency: The convolution theorem enables efficient image filtering through frequency domain multiplication. A spatial domain convolution with an $M \times M$ kernel requires $O(NM^2)$ operations for an N-pixel image. Using FFT-based convolution via the convolution theorem requires $O(N\log N)$ operations, providing substantial speedup for large kernels.

Practical Application: Liu et al. (2024) leveraged the spectral convolution theorem to develop faster Fourier convolution networks for MRI reconstruction, expanding receptive fields by enabling the network to utilize information from the entire frequency spectrum according to the convolution theorem, achieving superior performance compared to spatial-domain-only approaches

2.2.3 Frequency Shifting & Scaling

Frequency Shifting Property: The frequency shifting property (also called modulation property) describes how spatial domain phase shifts correspond to frequency domain shifts. For 1D signals:

$$\mathcal{F}\{f(x)e^{j2\pi u_0 x}\} = F(u - u_0)$$

This property indicates that multiplying a spatial signal by a complex exponential shifts its frequency spectrum by u_0.

Spatial Scaling Property: The spatial scaling property describes how scaling the spatial domain signal affects the frequency domain representation:

$$\mathcal{F}\{f(ax)\} = \frac{1}{|a|} F\left(\frac{u}{a}\right)$$

This relationship is inverse: spatial compression (|a| > 1) corresponds to frequency expansion, and spatial expansion (|a| < 1) corresponds to frequency compression. This property is fundamental to understanding multi-scale image processing approaches where features at different scales require different frequency resolutions.

Application in Image Rotation: In 2D, frequency shifting enables efficient image rotation analysis. Rotations in the spatial domain correspond to identical rotations in the frequency domain (log-polar representation), allowing rotation-invariant feature extraction through appropriate frequency domain transformations.

2.2.4 Parseval's Theorem

Parseval's theorem establishes energy conservation between spatial and frequency domains, fundamentally important for understanding signal energy distribution and power spectral analysis. For continuous 1D signals, Parseval's theorem states:

$$\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \int_{-\infty}^{\infty} |F(u)|^2 \, du$$

This identity indicates that the total energy in a signal is preserved under Fourier transformation—energy computed in the spatial domain equals energy computed in the frequency domain.

For discrete signals, Parseval's theorem is expressed as:

$$\sum_{n=0}^{N-1} |f[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |F[k]|^2$$

For 2D discrete images, the theorem generalizes to:

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |f[m,n]|^2 = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} |F[k,l]|^2$$

Physical Interpretation: $|f[n]|^2$ represents the power of the signal at position n, while $|F[k]|^2$ represents the power spectral density at frequency k. Parseval's theorem guarantees that total power is conserved, enabling analysis of how energy is distributed across different frequencies.

Application in Feature Extraction: Parseval's theorem justifies frequency domain feature extraction by ensuring that features derived from frequency magnitude spectrum retain complete energy information. This property enables robust feature extraction approaches that work equally well in spatial or frequency domains, with complementary advantages in each domain.

2.3 Fourier Transform of Digital Images (2D DFT)

Digital images are inherently 2D signals where pixel intensity values are sampled at discrete spatial locations. The 2D DFT transforms an $M \times N$ spatial domain image $f[m,n]$ into a frequency domain representation $F[k,l]$, revealing the frequency content and spatial periodicity patterns within the image.

The mathematical definition has been presented in Section 2.1.2. However, several important implementation considerations deserve emphasis:

Periodicity Property: The 2D DFT exhibits periodicity in both frequency dimensions with period M and N respectively:

$$F[k+M,l] = F[k,l], F[k,l+N] = F[k,l]$$

This periodicity means that the frequency domain representation repeats outside the primary range $[0, M - 1] \times [0, N - 1]$. When computing the inverse DFT, only the primary frequency range contains unique information.

Zero-Frequency Component (DC Component): The element $F[0,0]$ represents the average intensity of the entire image:

$$F[0,0] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f[m,n] = MN \times (\text{mean intensity})$$

The DC component carries no spatial detail information but represents the overall brightness level of the image.

Frequency Localization:

The frequency components $F[k,l]$ correspond to specific spatial frequencies:

- Low-frequency components (k, l near 0) represent large-scale image structure and smooth variations
- High-frequency components (k, l far from 0) represent fine details, edges, and rapid intensity changes

Computational Implementation: 2D FFT is typically computed using the row-column decomposition approach:

1. Apply 1D FFT to each row independently
2. Apply 1D FFT to each column of the intermediate result

This decomposition reduces computational complexity to $O(MN\log(MN))$ and enables efficient software implementation through library functions like FFTPACK, FFTW, or hardware-optimized implementations.

Frequency Domain Visualization: Due to the periodicity of DFT, the origin (representing zero frequency) is typically positioned at the center of the displayed spectrum for intuitive visualization. Centering is achieved through a shift operation:

$$F_{\text{centered}}[m,n] = F\left[\left(m+\frac{M}{2}\right) mod M, \left(n+\frac{N}{2}\right) mod N\right]$$

This centering repositions the low-frequency components to the image center, creating visually interpretable frequency domain displays.

2.4 Magnitude Spectrum and Phase Spectrum

The 2D DFT output $F[k,l]$ is a complex-valued matrix that can be decomposed into magnitude and phase components, each carrying distinct information about the image characteristics

2.4.1 Magnitude Spectrum

The magnitude spectrum represents the amplitude of frequency components and is defined as:

$$| F[k,l] | = \sqrt{\text{Re}(F[k,l])^2 + \text{Im}(F[k,l])^2}$$

where re$(F[k,l])$ and Im$(F[k,l])$ denote the real and imaginary components respectively.

The magnitude spectrum provides amplitude information: larger magnitude values indicate stronger presence of those particular spatial frequencies in the image. Low-frequency magnitudes are typically much larger than high-frequency magnitudes because images generally contain more energy in slowly-varying components than in rapidly-varying details.

Power Spectral Density: The squared magnitude spectrum represents the power distribution across frequencies:

$$P[k,l] = |F[k,l]|^2$$

Power spectral density is frequently used for spectral analysis and feature extraction, as it indicates the contribution of each frequency component to the overall image energy.

Logarithmic Scaling: For visualization, the magnitude spectrum is typically displayed using logarithmic scaling:

$$\text{Log-Magnitude} = \log(1+|F[k,l]|)$$

The logarithmic scale compresses the wide dynamic range of magnitude values (often spanning several orders of magnitude), making both weak and strong frequency components simultaneously visible in a single display. The constant "1" prevents undefined logarithm values when magnitudes are zero

2.4.2 Phase Spectrum

The phase spectrum represents the phase angle of each frequency component and is defined as:

$$\Phi[k,l] = \arctan\left(\frac{\mathrm{Im}(F[k,l])}{\mathrm{Re}(F[k,l])}\right)$$

The phase spectrum provides information about the spatial positioning and alignment of image features. Phase information determines where features are located spatially, while magnitude information determines their amplitude.

Phase Importance for Image Structure: A counterintuitive but important finding is that phase information contains more spatial detail and structural information than magnitude information alone. Image reconstruction using only magnitude spectrum (setting phase to zero) produces severely degraded results lacking spatial localization of features. Conversely, using only phase information (setting magnitude to unity) produces a structured outline showing feature positions, though with lost intensity information.

Example of Phase Significance: Chakraborty et al. (January 2026) demonstrated that in complex-valued CNNs for frequency-domain processing, jointly using magnitude and phase information achieved superior performance with magnitude serving as the primary contributor to accuracy (approximately 70-80% contribution) and phase providing critical secondary enhancement (approximately 20-30% contribution). They introduced a Log-Magnitude activation function specifically designed to preserve phase information while introducing non-linearity, outperforming split-type adaptations of real-valued activation functions that distorted phase components.

Magnitude-Phase Decomposition Utility

The decomposition into magnitude and phase enables independent manipulation of different image attributes:

$$F[k,l] = |F[k,l]|\, e^{j\Phi[k,l]}$$

This representation allows:

- Contrast Enhancement: Multiply magnitude spectrum while preserving phase
- Spatial Repositioning: Modify phase while preserving magnitude
- Noise Reduction: Attenuate high-frequency magnitude components
- Feature Extraction: Extract features from either magnitude or phase spectrum depending on application requirements

2.5 Sampling, Aliasing & Resolution in Frequency Domain

2.5.1 Sampling Theorem and Nyquist Frequency

Digital image processing requires sampling of continuous images into discrete pixel arrays. The fundamental principle governing this sampling is the Nyquist-Shannon sampling theorem, which establishes the relationship between sampling rate and frequency content to prevent aliasing artifacts. Nyquist-Shannon Sampling Theorem: A continuous signal with maximum frequency component $f_{\max}$ must be sampled at a rate f_s satisfying:

$$f_s \geq 2f_{\max}$$

The Nyquist frequency is defined as half the sampling rate:

$$f_N = \frac{f_s}{2}$$

For spatial sampling of images with sampling interval Δx (the distance between adjacent pixels), the spatial Nyquist frequency is:

$$f_N = \frac{1}{2\Delta x}$$

This means that image features smaller than $2\Delta x$ (twice the sampling interval) cannot be accurately represented and will be lost or distorted during sampling. Conversely, any features larger than $2\Delta x$ can be perfectly reconstructed from sampled data.

Practical Sampling Criterion: In practice, to ensure adequate sampling for high-resolution imaging without oversampling, a sampling interval between 2.5 to 3 samples per smallest resolvable feature is recommended. For an optical microscope with Abbe resolution limit of approximately 0.22 micrometers, the digitizer must sample at intervals corresponding to 0.11 micrometers or less in specimen space.

2.5.2 Aliasing and Aliasing Artifacts

Aliasing Definition: Aliasing occurs when a signal is sampled at a rate lower than required by the Nyquist criterion. High-frequency components that exceed the Nyquist frequency are incorrectly represented as low-frequency components in the sampled signal.

Aliasing Frequency Calculation: When a continuous signal contains a frequency component f greater than the Nyquist frequency f_N, the aliasing frequency f_a at which it appears in the sampled signal is:

$$f_a = \mid kf_s - f \mid$$

where k is chosen as the integer giving the minimum value. Equivalently:

$$f_a = f - kf_s \text{ where } k = \left\lfloor \frac{f}{f_s} \right\rceil$$

Example: Consider a signal sampled at $f_s = 100$ Hz (Nyquist frequency = 50 Hz). If the signal contains frequency components at 25 Hz, 70 Hz, 160 Hz, and 510 Hz:

- 25 Hz: Sampled correctly (below Nyquist frequency)
- 70 Hz: Aliases to |70 - 100| = 30 Hz
- 160 Hz: Aliases to |160 - 100| = 60 Hz
- 510 Hz: Aliases to |510 - 500| = 10 Hz

Only the 25 Hz component is correctly represented; all others appear as erroneous lower-frequency components.

Aliasing in Image Processing: In image processing, aliasing manifests as moiré patterns, jagged edges (jagging), and loss of fine details when images are subsampled or when acquisition resolution is insufficient. The periodicity of the DFT means that high-frequency components that exceed the Nyquist limit fold back into the low-frequency region, corrupting valid low-frequency information.

Anti-Aliasing Strategies:

1. Band-Limiting Filters: Apply low-pass filters before sampling to remove frequency components beyond the Nyquist limit
2. Oversampling: Sample at rates significantly exceeding the minimum Nyquist rate (typically 2-3 times higher)
3. Interpolation: Use appropriate interpolation methods during image resizing to minimize aliasing artifacts

2.5.3 Frequency Domain Resolution

Frequency domain resolution refers to the ability to distinguish different frequency components and is determined by the spatial sampling properties.

Frequency Resolution (Frequency Spacing): For an image of size $M \times N$ with sampling intervals Δx and Δy, the frequency resolution (spacing between adjacent frequency components) is:

$$\Delta f_u = \frac{1}{M\Delta x}, \Delta f_v = \frac{1}{N\Delta y}$$

Frequency resolution is inversely proportional to spatial extent: larger images (larger M or N) provide finer frequency resolution. Conversely, smaller images provide coarser frequency resolution, making it difficult to distinguish closely-spaced frequency components.

Trade-off Between Spatial and Frequency Resolution: The uncertainty principle in Fourier analysis establishes a fundamental trade-off: improving frequency resolution requires larger spatial extent (larger M, N), while improving spatial resolution requires larger sampling rate (smaller $\Delta x, \Delta y$). These are constrained by the fundamental relationship:

$$\Delta x \cdot \Delta f_u = 1, \Delta y \cdot \Delta f_v = 1$$

This trade-off means that a localized spatial feature (high spatial resolution) necessarily produces a broad frequency spectrum (low frequency resolution), and conversely, a narrow frequency spectrum (high frequency resolution) corresponds to extended spatial extent.

Numerical Example: For a 512×512 image with pixel spacing $\Delta x = \Delta y = 1$ micrometer:

- Total spatial extent: 512 micrometers
- Frequency resolution: $\Delta f = 1/(512 \text{ micrometers}) = 1.95 \times 10^{-3}$ micrometers^{-1}
- Nyquist frequency: $f_N = 1/(2 \times 1 \text{ micrometer}) = 0.5$ micrometers^{-1}
- Maximum distinguishable spatial wavelength: 2 micrometers

If the same image is padded to 1024×1024 through zero-padding:

- Frequency resolution improves to: $\Delta f = 1/(1024 \text{ micrometers}) = 0.98 \times 10^{-3}$ micrometers^{-1}
- Improved frequency resolution by factor of 2, but Nyquist frequency unchanged

Resolution in Machine Vision Applications: Xu et al. (2025) addressed resolution in the context of region-specific fidelity for MRI reconstruction, optimizing k-space (frequency domain) sampling patterns to maintain high-fidelity reconstruction in clinically critical regions while using lower resolution in less important regions. Their approach achieved SSIM scores of 0.9764 for global reconstruction and Dice scores of 0.9606 for critical regions, demonstrating that frequency domain resolution allocation can be optimized based on application-specific requirements.

3. Image Filtering in Frequency Domain

3.1 Types of Filters

3.1.1 Low-Pass & High-Pass Filters

Low-pass filters (LPF) remove high frequencies preserving low-frequency image structure: ideal LPF with cutoff frequency D_0:

$$H(u,v) = \begin{cases} 1 & \text{if } D(u,v) \le D_0 \\ 0 & \text{if } D(u,v) > D_0 \end{cases}$$

where $D(u,v) = \sqrt{u^2 + v^2}$ is distance from frequency origin. Results: noise reduction (uncorrelated noise broadly distributed in frequency), image smoothing/blurring (loss of sharp edges).

High-pass filters (HPF) retain high frequencies attenuating low frequencies:

$$H(u,v) = \begin{cases} 0 & \text{if } D(u,v) \le D_0 \\ 1 & \text{if } D(u,v) > D_0 \end{cases}$$

Results: edge enhancement/sharpening (edges concentrate high-frequency energy), removal of image smoothness variations, noise amplification (noise contains substantial high-frequency content).

3.1.2 Band-Pass, Band-Stop Filters

Band-pass filters select specific frequency range $D_1 < D(u,v) < D_2$, preserving frequency bands while attenuating both lower and higher frequencies. Applications: texture analysis (specific texture frequencies), oriented feature extraction (rotated band-pass filters detect specific orientations).

Band-stop (notch) filters remove specific frequency band, preserving everything except target band. Applications: power-line noise removal (50/60 Hz fundamental frequency), periodic interference elimination.

3.1.3 Gaussian & Ideal Filters

Ideal filters exhibit sharp frequency cutoff (step transition) producing severe ringing artifacts in spatial domain (Gibbs phenomenon). Gaussian filters employ smooth frequency transition:

H(u,v)=e-D2(u,v)/(2D02)

reducing ringing artifacts at cost of less sharp frequency cutoff. Trade-off: ideal filters achieve precise frequency isolation with spatial domain artifacts; Gaussian filters sacrifice frequency precision for smooth spatial response

3.2 Image Enhancement

3.2.1 Noise Reduction in Frequency Domain

Noise reduction exploits spectral characteristics: uncorrelated noise distributes uniformly across all frequencies, while diagnostic image information concentrates in specific frequency bands. Low-pass filtering removes high-frequency noise but sacrifices edge information. Wiener filtering optimally balances noise reduction and detail preservation:

$$H_w(u,v) = \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_n(u,v)/S_f(u,v)}$$

where S_n is noise power spectrum, S_f is original image power spectrum

3.2.2 Edge Enhancement using High-Frequency Components

Edge enhancement emphasizes high-frequency components: unsharp masking computes $f_{\text{enhanced}} = f - \alpha \cdot \text{LPF}(f)$ (subtracting blurred version amplifies edge details). In frequency domain: high-pass filter applied emphasizing edges, with tunable enhancement strength parameter.

3.3 Image Compression

3.3.1 Transform Coding

Transform-based compression exploits energy compaction: mathematical transforms (DCT, FFT) concentrate image energy into few coefficients, enabling lossy compression through selective coefficient quantization. General framework:

1. Block decomposition: divide M×N image into smaller blocks (8×8 or 16×16 pixels)
2. Transform application: apply orthogonal transform (DCT or FFT) to each block
3. Quantization: reduce coefficient precision, discard small coefficients
4. Entropy coding: apply Huffman or arithmetic coding

Inverse operations reconstruct image from compressed representation.

3.3.2 FFT-based Compression Techniques

FFT enables efficient image compression by exploiting frequency domain sparsity (most energy concentrated in low frequencies). DCT (Discrete Cosine Transform), closely related to FFT, provides superior energy compaction for natural images compared to pure FFT, achieving compression ratios 10-20:1 while maintaining diagnostic quality in medical imaging.

3.4 Image Restoration

3.4.1 Inverse Filtering

Degradation model: $g(x,y) = h(x,y) \otimes f(x,y) + n(x,y)$ where g is observed degraded image, h is degradation kernel (blur), f is original image, n is noise. Inverse filtering applies $H^{-1}(u,v)$ to reverse degradation. Problem: noise amplification at frequencies where H approaches zero, producing severe artifacts.[4][11]

3.4.2 Wiener Filtering

Optimal restoration filter minimizing mean-squared error:

$$H_w(u,v) = \frac{H^*(u,v)}{|H(u,v)|^2 + S_n(u,v)/S_f(u,v)}$$

Requires knowledge of signal and noise power spectra, providing superior restoration compared to inverse filtering with reduced noise amplification.

3.5 Pattern Recognition and Feature Extraction

Frequency domain features provide invariance properties valuable for pattern recognition: spectral moments (center of mass, variance) characterize frequency distribution; radial power spectrum describes object complexity;

texture features quantify roughness and periodicity. Machine learning classifiers operate on feature vectors extracted from frequency domain.[2][16]

3.6 Applications in Medical Imaging

3.6.1 MRI Fourier Reconstruction

MRI acquisition operates in frequency domain (k-space), directly acquiring Fourier coefficients of object. 2D FFT reconstructs spatial image from k-space data, with acquisition parameter control over frequency resolution and field-of-view.

3.6.2 CT Image Frequency Analysis

Frequency analysis of CT images reveals artifact characteristics (metal artifacts concentrated in specific frequency bands), enabling targeted artifact reduction through frequency-domain filtering.[2][1]

4. Image Acquisition and Preprocessing

Image acquisition converts continuous optical scene into discrete M×N pixel array through lens optics and sensor photoelectric conversion. Anti-aliasing filters (optical or electronic) limit frequency content to Nyquist frequency before sampling, preventing aliasing artifacts. Preprocessing includes: noise reduction, illumination normalization, geometric correction, color space conversion (RGB to grayscale if necessary).

4.1 Computation of 2D FFT and Spectrum Visualization

Preprocessing image padded to power-of-2 size (e.g., 256×256, 512×512) for efficient FFT computation. 2D FFT implemented via sequential 1D FFTs on image rows then columns. Frequency spectrum computed as magnitude $|F(u,v)|$ displayed as 2D image with brightness proportional to spectral power. Visualization typically applies logarithmic scaling $\log(1+|F(u,v)|)$ to compress dynamic range enabling simultaneous visualization of strong low-frequency and weak high-frequency components.

4.2 Frequency-Domain Filtering Process

Filtering implemented as element-wise multiplication in frequency domain: $G(u,v) = H(u,v) \cdot F(u,v)$ where H is filter transfer function. Inverse 2D FFT reconstructs filtered spatial image: $g(x,y) = \text{IFFT}(G(u,v))$. Filter design involves specifying H (u, v) (via mathematical formula or numerical specification) balancing frequency selectivity against spatial domain artifact generation.

4.3 Implementation of Transform-Based Enhancements

Enhancement techniques systematically applied: noise reduction (LPF), sharpening (HPF or unsharp masking), edge detection (band-pass or directional filtering), contrast enhancement (frequency-dependent gain modulation).

4.4 Performance Metrics

4.4.1 PSNR (Peak Signal-to-Noise Ratio)

$$\text{PSNR} = 10\log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}}\right)\ \text{dB}$$

where MAX is maximum pixel value (255 for 8-bit images), MSE is mean squared error.

4.4.2 MSE (Mean Squared Error)

$$\text{MSE} = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[f(m,n)-\hat{f}(m,n)]^2$$

quantifies pixel-level differences between original and processed images.

4.4.3 SSIM (Structural Similarity Index)

$$\text{SSIM} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

measures structural similarity considering luminance

Although structural MRI scans give information regarding brain tumors on a cellular, functional, metabolic, as well as vascular level [10], radiologists support these scans for brain tumor diagnosis. Multi-orientation and multi-modality MRI can provide a reliable diagnosis of brain tumor tissue. The multi-modal MRI includes T1, T1-Gadolinium (T1-Gd), T1-Weighted (T1-W), T2, T2-W, Fluid Attenuated Inversion Recovery (FLAIR), and additional modalities [11, 12]. The various tumor tissues such as edema, necrosis, enhancing, and non-enhancing tissues are examined with these modalities. Active tumors can be detected with T1Gd, the tumor core can be identified with T2, and the entire tumor can be identified with the T2 FLAIR modality [13]. T2-W MRI signifies that the edema tissue around the tumor is shining, as seen in the MRI image. Furthermore, axial, coronal, and sagittal orientations of MRI are presented in the MRI images [14]. The various modalities offer different biological data that may be evaluated using basic quantitative metrics for disease progression, treatment planning, and tumor diagnosis in clinical routine [15].

5. Motivation

Larger datasets are required to assess the efficiency of various techniques for diagnosing brain tumors. A few models have increased computational complexity, which reduces the overall efficacy of identification. A few brain tumor classification algorithms were developed to reduce overfitting in the dataset; however, they generated an enormous quantity of noise during training. During the process of segmenting the MRI image, some significant limitations included time complexity, high processing costs, and the necessity for additional parameters. Several treatments are less effective when the dimensions, as well as the position of brain tumors, are not accurately identified. Brain tumors frequently exhibit inhomogeneous intensities in their MRI images, which are limited by imaging principles. The MRI image reveals overlapping and unclear borders between the tumor and the surrounding tissues. To overcome this kind of issue, the proposed technique is developed to provide efficient brain tumor segmentation and classification. The main contribution of the suggested method is discussed below.

- To enhance the image quality, efficient pre-processing techniques such as Upgraded Mean Filtering (UP-MFil), image resizing, and HSV color channel conversion are used.
- To segment the pre-processed image, Gannet-based Kapur's Thresholding (GKT) scheme is used based on WT, TC, and ET.
- To extract the optimal features, the Gannet Optimal EfficientNetB3 (GO_ENetB3) model is used, and it also reduces dimensionality issues.
- To classify the tumor as HGG and LGG, the Fine-Tuned Hybrid Deep Convolutional Neural Network (FT-HDCNNet) model is utilized.

The remaining content of the paper is organized in the following manner: Section 2 includes the survey of some related techniques over multi-modal brain tumor segmentation and classification. Then, the techniques utilized in the proposed methodology are described briefly in Section 3, with Figures. The Results obtained by the proposed method are analyzed and discussed in Section 4. Finally, the overall conclusion of the paper is shown in Section 5.

Table 1: Analysis of some related techniques

Author name and reference	Technique use	Dataset used	Limitations	Performance
M.O. Khairandish et al. [28]	Hybrid CNN-SVM	BRATS 2015 dataset	Limited features are considered	97.43% accuracy
Ilyasse Aboussaleh et al. [29]	CNN	BraTS 2017 dataset	It only considers FLAIR modalities	91% accuracy
Nagwan Abdel Samee et al. [30]	DCNN	BRATS 2015 dataset	Low performance	88.6% accuracy

Author name and reference	Technique use	Dataset used	Limitations	Performance
R. Pitchai et al. [31]	ANN-Fuzzy K means algorithm	BRATS dataset	It is not affected by over-segmentation	94% accuracy
Angona Biswas and Md. Saiful Islam, [32]	ANN	Figshare dataset	Computation complexity	95.4% accuracy
Amran Hossain et al. [33]	BINet model	Real-time dataset	High computational complexity	89.33% accuracy and 93.10% Dice Score
R. Poonguzhali et al. [34]	TSO-GRU model	FigShare dataset	High time consumption	97.48% accuracy
Hanene Sahli et al. [35]	Combined ResNet-SVM model	MRI dataset	89.36% accuracy	Hard to handle a larger dataset

The requirement for larger datasets renders many strategies for detecting disease less effective. Certain models have a worse overall efficacy for brain tumor identification due to their high computational complexity. While numerous brain tumor classification methods have been created to reduce overfitting in datasets, the training process generates a large amount of noise. Some notable problems encountered while segmenting the MRI image included temporal complexity, high processing costs, and the need for extra parameters. Insufficient identification of dimensions, as well as the location of disease, reduces the efficiency of several methods. Disease frequently exhibits inhomogeneous intensities in their MRI images, which are limited by imaging principles. The MRI image reveals overlapping and unclear borders between the tumor and the surrounding tissues.

Image resizing

The process of comprising the pixels of images is known as image resizing, and each pixel in the image has three color combinations such as red, blue, and green. It is used to reduce the size of the image as well as reduce the requirement of many bytes to store the image. The size of the image in the dataset is which resized into for efficient brain tumor classification.

5.1. HSV color channel conversion

This technique is employed to reduce the size of MRI images to avoid classification errors, as shown below.

$$C = \cos^{-1} \frac{\frac{1}{2[(R-G)^2 + (R-B)]}}{\sqrt{(R-G)^2 + ((R-B)(G-B))}} \tag{1}$$

$$A = 1 - \frac{3[Min(R,G,B)]}{R+G+B} \tag{2}$$

$$L = \frac{R+G+B}{3} \tag{3}$$

Here, the red, green, and blue colors are represented as R , G and B . This technique is utilized to trim the edges of each input image, convert the color channel of the image, and improve its quality.

5.2. Segmentation using GKT

The Segmentation stage is used to group pixels based on their similarities and to segment the tumor in an MRI image. Hence, the GKT model was used for segmentation, and its efficiency was improved using Gannet optimization. It is utilized to segment Disease based on WT, TC, and ET.

5.2.1 Gannet-based Kapur's thresholding

The segmentation is performed based on a threshold-based technique to evaluate the optimum threshold. The probability distribution and image histogram entropy are required for the calculation. This technique establishes the optimum threshold for entropy. The bi-level threshold is evaluated by objective utility, which is described in the below equation.

$$BTO_{kap}(a) = S_1 + S_2$$

(4)

Here, the mathematical expression of S_1 and S_2 described in the equation below.

$$S_1 = \sum_{v=1}^{a} \frac{x_v}{e_0} 1s\left(\frac{x_v}{e_0}\right) \tag{5}$$

$$S_2 = \sum_{v=a+1}^{u} \frac{x_v}{e_1} 1s\left(\frac{x_v}{e_1}\right) \tag{6}$$

Here, the probability distribution of grayscale intensity level is denoted as for and class model. Then, and are represented by and probability distribution. The entropy-based model is sufficiently malleable for multilevel thresholding. Here, it is utilized to divide the MRI image into class labels using the threshold numbers. The variations of objective values are described in the below equation.

$$BTO_{kap}(A) = \sum_{u=1}^{s} S_u \tag{7}$$

Here, a vector involves numerous threshold numbers are represented as. The entropies are termed discretely with the corresponding threshold value denoted as. It is modified for the entropy, which is described in the below equation.

$$S_s^t = \sum_{x=a_{s+1}}^{K} \frac{xp}{e_{s-1}} 1s\left(\frac{xp}{e_{s-1}}\right) \tag{8}$$

Here, the probability rate for classes is represented as , and the threshold numbers are optimized utilizing the gannet optimization method.

5.2.2 Gannet optimization algorithm

A metaheuristic Gannet optimization algorithm was utilized to regulate the optimal threshold for efficient segmentation [36]. The mathematical expression to update the position is described in the below equation.

$$GO_a(y+1) = \begin{cases} O_a(y) + p_1 + p_2, & n \geq 0.5 \\ O_a(y) + q_1 + q_2, & n < 0.5 \end{cases} \tag{9}$$

$$p_2 = R * (O_a(y) - O_a(y)) \tag{10}$$

$$q_2 = Q * (O_a(y) - O_a(y)) \tag{11}$$

$$R = (2 * i_4 - 1) * h \tag{12}$$

$$Q = (2 * i_5 - 1) * t \tag{13}$$

Here, the random numbers between 0 and 1 are represented as and. The random number between as well as is described as, similarly, the random number among as well as is represented as. The distinct in current inhabitants is denoted as. Then, and are arbitrarily selected distinct and average locations of distinct in the current inhabitants, which is described in the below equation.

$$O_k\left(y\right)=\frac{1}{Z}\sum_{a=1}^{Z}O_a\left(y\right) \tag{14}$$

Here, Gannet-based Kapur's Thresholding is utilized to perform segmentation based on intensity, texture, or color. The Gannet optimization is utilized to choose the optimal threshold to provide efficient segmentation.

5.3 Feature extraction and selection performed by the GO_ENetB3 model

The feature extraction is performed by the GO_ENetB3 model, which is the combination of the EfficientNet B3 model and the EGO algorithm. In this technique, the EfficientNet B3 model is utilized to remove the features, and the EGO algorithm is utilized to select optimal features and solve the dimensionality issues.

5.3.1 Feature extraction using EfficientNetB3

The EfficientNet is a type of CNN that includes several convolutional layers, such as Mobile Inverted Residual Bottleneck Convolutional (MBConv) blocks and fully connected layers. The Efficient is smaller than other models, such as ResNet50 and EfficientNet-B0, which includes 23,534,592 and 5,330,564 boundaries. These models have high computational power; hence, EfficientNetB3 is utilized to extract the features [37].

MBConv is the fundamental component of the EfficientNet paradigm and is included in MobileNet-V2. Initially, the EfficientNet model employs a stem architecture, which is shared by all eight models and their final layers. There are seven blocks available, each with a unique sub-block. It moved from the EfficientNetB0 to the EfficientNetB7 model by increasing the number of iterations. EfficientNetB0 and EfficientNetB7 each have 237 and 813 layers, respectively. The second module finds the first sub-block of seven main blocks, with the exception of the first block. The skip connection connects Module 3 to each sub-block, and it is the first sub-block to be merged with Module 4. In Module 5, each sub-block is connected to the one behind it using a skip path. Finally, these modules are merged and used in precise ways within the blocks to form sub-blocks. The architecture of the GO_ENetB3 model utilized for feature extraction, as well as optimal feature selection, is represented in Figure 4.

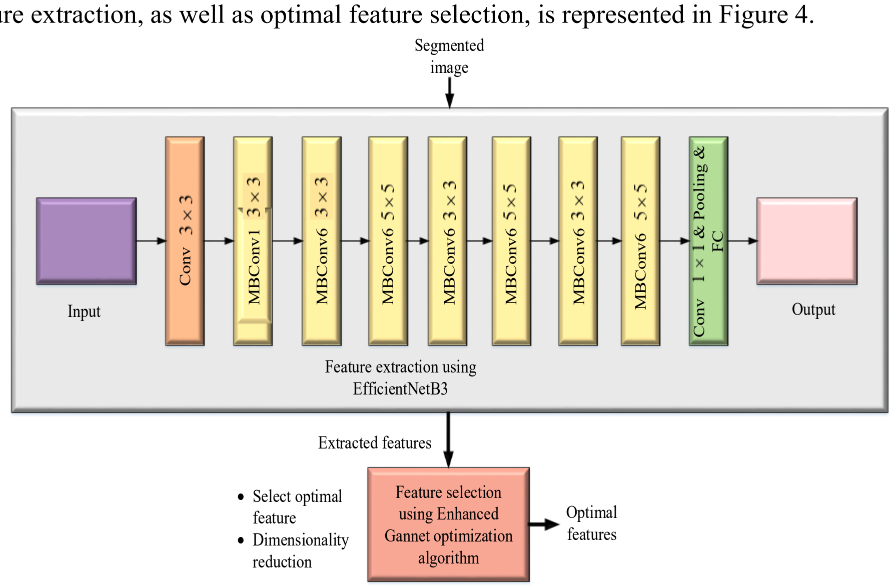

Figure 4: Architecture of GO_ENetB3 model

Feature selection using EGO algorithm

The Gannet optimization algorithm (GoA), a metaheuristic optimization method, was replaced by the EGO. It takes its cues from the gannet bird's natural predation behavior, which emulates the concepts of genetic inheritance and fittest survival to gradually evolve the population of solutions toward the ideal or nearly ideal state. There are two phases to it: exploration and exploitation. The mechanism operates on the basis of abrupt rotation, random wandering, and U-shaped and V-shaped diving modes.

5.4 Initialization

The population of potential solutions is randomly generated by every individual. These individuals denote the possible solution for optimization issues. The initial population is assumed by the gannets, which are initialized randomly with the boundary values, and it is described in the below equation.

$$g_{x,y} = rand_1 \times \left(B_y - b_y\right) + b_y \tag{15}$$

here, represents the position of the search agent in dimension. Then, and are denoted as upper and lower boundaries. The random value is denoted as which lies between 0 and 1.

The memory matrix is represented by the term in this process; the gannet's position is changed over the iteration, which is returned in the memory matrix. The matrix value is modified based on the fitness measurements.

5.4.1 Exploration

During the exploration phase, gannets search the water's surface for prey. Then, it adjusts their diving patterns to match the water depth. U-shaped and V-shaped dives are the two different forms, as the equation below illustrates.

$$U = 2 * \cos(2 * \pi * srand_2) * g \tag{16}$$

$$V = 2 * P\left(2 * \pi * srand_3\right) * g \tag{17}$$

$$g = 1 - \frac{gc}{gm} \tag{18}$$

In equation (18), the maximum iteration count and current iteration are represented as and. Then, and are denoted as random values that lie between 0 and 1. The updated position based on two dive criteria is described in the below equation.

$$D_j\left(g+1\right) = \begin{cases} D_j\left(g\right) + S_1 + S_2, & x \geq 0.5 \\ D_j\left(g\right) + T_1 + T_2, & x < 0.5 \end{cases} \tag{19}$$

here, and.

The mathematical expression for terms and is described in the below equation.

$$u = (2 * srand_4 - 1) * U \text{ , } v = \left(2 * srand_5 - 1\right) * V \tag{20}$$

Here, random values between 0 and 1 are denoted as the variables and. Then, denotes the current search agent and represented a random selection of the population. The average position is denoted as correspondingly, and the evaluation of the average position is described in the below equation.

$$D_S\left(g\right) = \frac{1}{L}\sum_{j=1}^{L} D_j\left(g\right) \tag{21}$$

5.4.2 Exploitation

The gannet catches the prey after the search agents rush to the surface of the water, which requires more energy. The enormous energy is utilized by the gannets to capture the prey. The gannet can effectively catch its prey when it is supplied with enough energy. Based on the strategy of capturing, the updated position is described in the below equation.

$$D_j\left(g+1\right) = \begin{cases} g * \delta * \left(D_j\left(g\right) - D_{bt}\left(g\right)\right) + D_j\left(g\right), & Rx \geq c \\ D_{bt}\left(g\right) - \left(D_j\left(g\right) - D_{bt}\left(g\right)\right) * M * g, & Rx < c \end{cases} \tag{22}$$

here, denotes the constant parameters, which are fixed as 0.2, and the term is evaluated based on the below-described equation.

$$Rx = \frac{1}{X * g2} \tag{23}$$

$$g2 = 1 + \frac{gc}{gm} \tag{24}$$

$$X = \frac{mS * eS^2}{d} \tag{25}$$

$$d = 0.2 + (2 - 0.2) * srand_6 \tag{26}$$

here, denotes the random value between 0 to 1, and mass as well as the velocity of the gannet is represented as as well as , which is set to be 2.5 Kg and 1.5 m/s. The mathematical expression of the term is shown below.

$$\delta = Rx * \left| D_j(g) - D_{bt}(g) \right| \tag{27}$$

$$M = Levy(K) \tag{28}$$

Here, the best value is obtained, and this algorithm is improved to identify the issues of optimizing transportation and placing pooling centers. The EGO algorithm is utilized to improve the performance by finding the optimum value. The conventional GoA method is integrated to evaluate the new equation. The GoA eliminates the problems with local optima and offers good processing capability and a higher convergence rate. However, the GoA yields extra complexity by employing a large number of parameters to lower the EGO. The major issues of GoA are parameter setting and its adaptation. Hence, the delta parameter derived from equation (22) is replaced with the below-described equation.

$$\delta = \frac{\beta}{\gamma} \tag{29}$$

$$\gamma = \frac{\beta}{\alpha} \tag{30}$$

Here, the mathematical expression for the term and is described in the equation below.

$$\alpha = \frac{L}{100} \tag{31}$$

$$\beta = \frac{L}{20} \tag{32}$$

Here, the total number of the population is denoted as the term. Hence, the updated or modified delta value is utilized in the equation (22) to obtain a better solution. The pseudo-code of the EGO algorithm is represented in Table 3.

5.4.3 Classification using FT-HDC2Net

The brain tumor is classified into glioma, pituitary, and meningioma tumors using the FT-HDC2Net model, which is a combination of the CNN and capsule networks. Then, the hyperparameter utilized in the model is tuned using the WSO algorithm.

The multidimensional data, which includes the pooling layer, convolutional layer, and ReLU layer used in this model, is processed by CNN [38]. From the preceding layers, the convolutional layer is used to ascertain the local connectivity of the feature. It also maps their information in a specific feature map. The convolution of input with filter which is shown below.

$$(X * L)_{a,b} = \sum_{h=-p_1}^{p} 1 \sum_{k=-p_2}^{p} 2L_h, \quad {}_k X_{a-h,b-1} \tag{33}$$

here, denotes the non-linearity activation function, which is utilized to generate feature maps with convolutional layers. Similar features are combined using the max-pooling layers conveyed from the previous layer. The down sampling operation is carried out by the max-pooling layers by determining the maximum value of every field that overlaps the filter on the feature map.

The CNN structure from the fully connected (FC) layer to the classification is similar to the multilayer perceptron neural network (MLP). In the MLP layer, the process of both the FC layer and the hidden layer are the same. CNN structure includes one or more FC layers, and it connects the neuron in the next layer to the previous layer. The non-normalized values of the previous layer in CNNs are usually matched to a probable distribution across projected class scores using the softmax function. The following equation describes the mathematical role of the CNN's softmax function.

$$\sigma(i_n) = \frac{e^{i_n}}{\sum_{m=1}^{H} e^{i_n}}, \quad m = 1,2,\ldots,H \tag{34}$$

Here, softmax output for every is represented by , and the values of the input vector are denoted as.

The training time of CNNs is reduced using the batch normalization layers and sensitivity to the network initialization. Hence, the normalization process is selected for this layer. The normalized activations with input, mini-batch mean , and mini-batch variance are evaluated using the below-described equation.

$$\hat{i}_n = \frac{i_n - B_M}{\sqrt{B_V^2 + t}} \tag{35}$$

Here, the constant is represented as which develops the numerical state when is very small. The mathematical equations of and are described in the below equations.

$$B_M = \frac{1}{a}\sum_{x=1}^{a} i_n \tag{36}$$

$$B_V = \frac{1}{a}\sum_{i=1}^{a} (i_n - B_M)^2 \tag{37}$$

Finally, it calculates the activations in the batch normalization layer using shift, and the mathematical description of the scale operation is shown in the equation below.

$$j_n = x\hat{i}_n + y \tag{38}$$

Here, the balance and scale factors are represented as and. During the training process, these factors are learnable variables that are utilized to update the most appropriate values. The vectorial formulation has the ability to provide a capsule to identify features and also recognize changes to be a feature. It is an important property that distinguishes the capsule as an effective but lightweight feature representation approach.

5.4.4 Capsule Network

The vectorial capsule representations serve as the core neurons in capsule networks, which describe the features of entities. More exactly, a capsule's starting parameters describe the entity's inherent characteristics, but the capsule's length represents the entity's probability of existence. Conventional convolution procedures and capsule convolution processes are very different from one another. The entire input to a capsule in a capsule's convolution layer. The mathematical equation for the total input to the capsule is described below.

$$A_y = \sum_x p_{xy} \cdot U_{xy} \tag{39}$$

Here, total input to the capsule is denoted as. The coupling coefficient is dynamically determined as , which indicates the contribution degree of prediction from the capsule in the previous layer. The prediction form capsule is represented as which is described in the below equation.

$$U_{xy} = V_{xy} \cdot U_x \qquad (40)$$

Here, the output of the capsule is represented as , and the transformation matrix acts as a feature mapping function is represented as. For instance, all the connected capsules of the capsule in the layer have coupling coefficients that add up to 1. The enhanced dynamic routing process is utilized to determine the coupling coefficients, which are effectively utilized to construct a deep capsule network.

According to the length-based probability encoding method, capsules that are used more often should forecast with greater probability, while shorter capsules should contribute less to probability estimations. Lastly, to normalize the capsule's output, the squashing function is used as an activation function. The following equation describes the squashing function's mathematical expression.

$$U_y = \frac{\|A_y\|^2}{1+\|A_y\|^2} \cdot \frac{A_y}{\|A_y\|} \qquad (41)$$

Here, the original input, as well as the normalized output of the capsule, are represented as and. By normalizing, long capsules are shrunk to a length close to one, casting high predictions. In contrast, short capsules are repressed to almost zero length to provide fewer contributions. Figure 5 illustrates the FT-HDC2Net model's architecture.

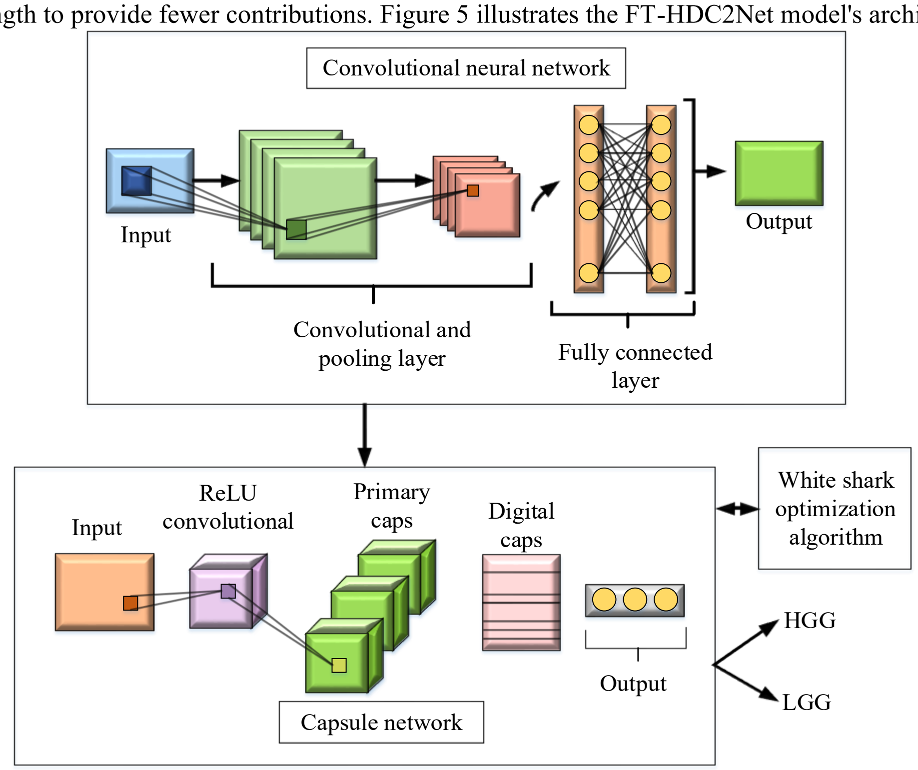

Figure 5: Architecture of FT-HDCNNet model

Hyperparameter tuned using WSO algorithm.

The hyperparameters are tuned using the WSO method and are modeled after the hunting and tracking behaviors of white sharks during foraging [39].

Initialization

The WSO is a population-based method that generates a pool of starting solutions to optimization problems at random to start the optimization process. The population of white sharks in -dimensional search space with every white shark's position is known as the candidate solution of the problem. It can be described mathematically in the matrix, as shown below in the equation.

$$s = \begin{bmatrix} s_1^1 & s_2^1 & \cdots & \cdots & s_d^1 \\ s_1^2 & s_2^2 & \cdots & \cdots & s_d^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s_1^x & s_2^x & \cdots & \cdots & s_d^x \end{bmatrix} \quad (42)$$

Here, the location of every white shark in the search space is denoted as , and the amount of decision variables is represented as. Then, the position of the white shark in is denoted as. The equation below describes the uniform random initialization used to generate the initial population in the search domain.

$$s_j^i = h_j + rand \times \left(UB_j - LB_j\right) \quad (43)$$

Here, the initial vector of the white shark in dimension is represented as. The upper as well as lower bounds of search space in dimension is denoted as as well as. Then, the random number generated in the interval [0, 1] is represented as.

Every candidate solution for a white shark's new position is evaluated according to a fitness function. The former is renovated if it is determined that the new position is superior to the current one. If the white shark's current position in the WSO simulation is better than its new position, then it will stay in the current position.

Speed movements toward the prey

Since white sharks are creatures with a need to survive, they spend a great deal of time hunting and pursuing prey. They generally use all available techniques, such as extraordinary senses of smell, sight, and hearing, to track and track their prey. A white shark swims in an erratic rhythm in search of its prey, identifying it by listening for waves to break while it is moving. The mathematical equation for the velocity vector is described below.

$$c_{l+1}^i = \mu\left[c_l^i + f_1\left(s_{gbest_l} - s_l^i\right) \times u_1 + f_2\left(s_{best}^{c_l^i} - s_l^i\right) \times u_2\right] \quad (44)$$

Here, the index of white sharks for a population size is represented as. The new velocity vector of the white shark in step is represented as. The current speed vector of the white shark in step is represented as. The global optimal location vector is attained by the shark in iteration and is denoted as. The existing location of the white shark in the step is represented as , and the index vector of white sharks is denoted as , which searches the optimal position using the equation (45). Here, the two uniformly generated random numbers are represented as and. The force of white sharks is represented as and which utilized to control the effect of and on using (46) and (47). The constriction factor is represented as which utilized to analyze the shark's convergence behavior using equation (48).

$$c = \left[x \times R(1,\, x)\right] + 1 \quad (45)$$

Here, the vector of a random number generated with uniform distribution in the range [0, 1] is represented as.

$$f_1 = f_{\max} + \left(f_{\max} - f_{\min}\right) \times e^{-(4l/L)^2} \quad (46)$$

$$f_2 = f_{\min} + \left(f_{\max} - f_{\min}\right) \times e^{-(4l/L)^2} \quad (47)$$

Here, existing as well as the total number of the iterations are represented as and. The initial and subordinate velocities are represented as and utilized to achieve better motion for white sharks. After rigorous analysis, the values of and are set to be 0.5 and 1.5.

$$\mu = \frac{2}{\left|2 - \tau - \sqrt{\tau^2 - 4\tau}\right|} \quad (48)$$

Here, the acceleration coefficient is denoted as is equal to 4.125.

A movement toward optimal prey

White sharks spend most of their time seeking prey that are either ideal or suboptimal. The white sharks constantly change postures and travel towards the prey based on the wave created by the prey and its smell. Sometimes, the prey escapes from the white shark when the prey moves based on the flow of waves. The white shark can track the prey by its smell even when it leaves their position. A white shark's approach to searching for prey in random locations is comparable to how fish schools search for food. The white shark's approach to prey is characterized by the position-updating strategy outlined in the equation below.

$$s_{l+1}^{i} = \begin{cases} s_l^i \cdot \neg \oplus s_0 + UL \cdot p + LL \cdot q; & R < mf \\ s_l^i + c_l^i / w; & R \geq mf \end{cases} \tag{49}$$

Here, the Updated position vector of the white shark in the iteration step is represented as and the negation operator is denoted as. The one-dimensional binary vectors are denoted as well as , which is determined by the equations (50) and (51). Equation (52) was utilized to evaluate the frequency of wavy motion in white sharks. The random number generated in the range [0, 1] is denoted as , and the movement force is denoted as , which increases with the number of iterations. It is performed by the white shark approaching prey, and it is described in the equation below.

$$p = \operatorname{sgn}\left(s_l^i - UL\right) > 0 \tag{50}$$

$$q = \operatorname{sgn}\left(s_l^i - LL\right) < 0 \tag{51}$$

$$s_0 = \oplus\left(p,\ q\right) \tag{52}$$

Here, a bit-wise xor operation is represented as.

In order for support solutions to behave randomly in the search space, these equations (50) and (51) must be satisfied. It is also utilized to assist the shark in searching every potential area of search space.

$$w = w_{Min} + \frac{w_{Max} - w_{Min}}{w_{Max} + w_{Min}} \tag{53}$$

Here, minimum as well as maximum frequencies of the rolling signal are represented as and. A random number is represented as . Then, and value is set to be 0.07 and 0.75 to handle the problem. These numbers can be readily altered for different situations when necessary because they were selected with thorough research and testing for a variety of problems.

$$mf = \frac{1}{\left(p_0 + e^{(H/2-h)/p_1}\right)} \tag{54}$$

Here, the two positive constants are represented as and which are used to manage both exploration as well as exploitation phases. Then, a parameter utilized to determine the white shark's strength for hearing, sense as well as smell.

5.4.5 A movement toward the best white shark

When the white shark is near its prey, it maintains its position. The equation below describes how this behavior occurs.

$$s_{k+1}'^{i} = s_{gbest_l} + rand_1 \vec{I}_s \operatorname{sgn}\left(rand_2 - 0.5\right) \qquad rand_3 < g_g \tag{55}$$

Here, the updated position of white sharks based on the position of prey is denoted as. Then, gives either 1 or -1 to modify the search direction; the random numbers are denoted as, and which lies among [0, 1]. The distance between the prey as well as the white shark is represented as using equation (56). The parameter is denoted as to determine the white shark's strength for smell and sight when it went close to the optimal prey, which is described in equation (57).

$$\vec{I}_s = \left| R \times \left(s_{gbest_l} - s_l^i \right) \right| \quad (56)$$

The random number is represented as lies in the range of [0, 1], as well as the current position of white share based on is represented as.

$$u_u = \left| 1 - e^{(-p_2 \times h/H)} \right| \quad (57)$$

Here, the positive constant is represented as which is used to control both exploration as well as exploitation, and it is set to be 0.0005 for every addressed test problem.

5.4.5.1 Fish school behavior

Maintain the top two optimal solutions and adjust the placements of the remaining sharks according to the best positions determined by analyzing the behavior of white sharks in groups. The following equation describes the mathematical behavior of fish groups in white sharks.

$$s_{l+1}^i = \frac{s_l^i + s_{l+1}'^i}{2 \times R} \quad (58)$$

Here, a random number distributed frequently within the interval [0, 1] is denoted as. Equation (57) is used to update shark positions so that they are in line with the ideal white shark position, which is close to the prey.

At some point, great white sharks may locate themselves inside the search space, often in close proximity to the best prey. The behavior of fish schools and the migration of white sharks to the best white sharks can be used to predict the collective behavior of WSO, which raises the possibility of more aspects of exploration and exploitation.

5.4.5.2 Exploration phase

Some of the parameters utilized for the exploration are described as follows.

μ : It is utilized to control exploration behavior, avert early convergence, as well as prevent the solution from descending into local optima. Its value depends on the parameter.

f_1 as well as : These parameters, which are iteration functions that determine the velocity of white sharks, will be adjusted during WSO's global and local searches to ensure an equal balance of exploration and exploitation. WSO's global and local searches, which are iteration-dependent parameters, control the white sharks' velocity updates and maintain a constant balance of exploration and exploitation. Therefore, during the WSO iterations, these parameters are flexible to enable reasonable exploration ability with accurate results.

: This parameter was chosen based on experimental assessments; it shouldn't be greater than 1.0 as this would result in a decline in exploration performance. In the initial iterations, all white sharks are found far from their target. Using to update the location of white sharks based on equation (50) facilitates both local and global space searching for WSO. The stronger the exploration capability as well as the higher the coefficient in . Similarly, a lower indicates a better capability for exploitation and a low capacity for exploration. Due to this, the value was appropriately adjusted to increase the capabilities for both exploration and exploitation. Based on test results, it is discovered that values that are primarily greater than 20 for and more than 100 for promote search agents to look for boundaries.

$\text{sgn}(rand_2 - 0.5)$ In Equation (55): It is utilized to control the direction of exploration.

$rand_1$ In Equation (55): It is utilized to assist the solution in searching space randomly, and variable variation provides efficient exploration.

5.4.5.3 Exploitation phase

Some parameters utilized to control exploitation and location search are described as follows.

: It is utilized for an intensity amount of exploitation, and its value relies on the parameter that reduces the likelihood of slack in local optima solutions.

and : These factors affect WSO's performance, which is demonstrated by their efforts to strike a balance between local and global search activity.

: Exploration decreases, and exploitation increases with the number of iterations. In the final iterations, when a white shark is close to a prey, will assist the WSO in executing a local search for the prey, which will result in exploitation. Additionally, the exploitation feature of is controlled by coefficients a0 and a1. By creating around the optimal answer, it determines the extent of exploitation.

in equation (56): It controls exploitation by determining the local search's direction since, because it exists in [0, 1], there is an equal chance of positive and negative signs.

in equation (56): It enhances the search space's solution uncertainty, enabling a sensible exploitation process.

Complexity analysis

The computational complexity of an algorithm can be represented as a function that links the method's run-time to the size of the input required to solve the problem. This is accomplished by utilizing the well-known Big-O notation, which is explained below, to calculate the computational complexity of the WSO in both time and space.

Time complexity

This complexity problem formed based on the number of white sharks, the number of iterations, the given problem's dimension , and the cost of function assessment. The mathematical equation for the time complexity of WSO is described below.

$$C(wso) = C(prob.def) + C(init.) + (\cos t\ \ function) + (U\_sol) \tag{59}$$

Here, the component's time complexity is described in equation (60), which is described as follows.

Initializing the problem definition demands time.

Initialized population generation demands time.

Assess cost function demands time.

Calculate updated solution demands time.

Based on these, the total time complexity of WSO is described in the equation below.

$$C(wso) = C(1 + xd + Hux + Hxd) \tag{60}$$

As,, and , then equation (60) is reduced to equation (61).

$$C(wso) \cong C(Hux + Hxd) \tag{61}$$

It clearly defines the WSO's time complexity as polynomial order.

Space complexity

The parameters for the number of white sharks, as well as the problem's dimensions, determine the WSO's space complexity in relation to memory space. This indicates the amount of space needed for the planned WSO's initialization.

$$C(xd) \tag{62}$$

The pseudo-code summarizing the iterative optimization process of the WSO algorithm is described in Table 2.

Table 2: Pseudo code of WSO algorithm

Initialize problem parameter
Initialize parameter of WSO
Generate the initial position of WSO randomly
Initialize velocity of every initial population
Evaluate its position
while $h < H$ **do**
Update $c, u_1, u_1, p, q, s_o, w, mf \; and \; g_g$ using equation (45), (46), (47), (48), (50), (51), (52), (53), (54) and (57)
for $i = 1 \; to \; x$ **do**

$$c_{l+1}^{i} = \mu\left[c_l^i + f_1\left(s_{gbest_l} - s_l^i\right)\times u_1 + f_2\left(s_{best}^{c_l^i} - s_l^i\right)\times u_2\right]$$

end for
for $i = 1 \; to \; x$ do
if $R < mf$ **then**

$$s_{l+1}^{i} = s_l^i \cdot \neg \oplus s_0 + UL \cdot p + LL \cdot q$$

else

$$s_{l+1}^{i} = s_l^i + c_l^i / w$$

end if
if $R < g_g$ **then**

$$s_{l+1}^{i} = s_{gbest_l} + rand_1 \vec{I}_s \operatorname{sgn}\left(rand_2 - 0.5\right)$$

else

$$s_{l+1}'^{i} = s_{gbest_l} + rand_1 \vec{I}_s \operatorname{sgn}\left(rand_2 - 0.5\right)$$

$$s_{l+1}^{i} = \frac{s_l^i + s_{l+1}'^{i}}{2 \times R}$$

end if
end if
end for
adjust white sharks' position beyond the boundaries
calculate and update the new position

$h = h + 1$

end while

Return optimal solution

Results

Several similar techniques, including CNN, Residual Network (ResNet) 18, Long Short-Term Memory (LSTM), and AlexNet, were used to examine the overall performance of the suggested method. A number of performance indicators are used to evaluate the effectiveness of the strategies examined. The effectiveness of the suggested technique is expressively described when contrasted with the performance of the current model.

Dataset description

The multi-modal brain tumor image segmentation benchmark (BRATS) dataset 18 was utilized in the proposed technique. It includes various images gathered from 285 glioma patients. Nearly 210 images are HGG, and the remaining images belong to the LGG cases. BraTS 2018 includes 66 patients with unknown grades. The storage capacity of this dataset is 28.95 GB, the usability of the proposed technique is 1.88, and it is the publicly available dataset. The training data's ground truth is labeled by the expert, which is not a publicly available dataset. This dataset includes four main classes: Enhancing tumor, Edema, Healthy tissue, and Necrosis/non-enhancing tumor. This BraTS dataset 2018 dataset can be assessed through https://www.kaggle.com/datasets/sanglequang/brats2018.

Performance metric and its formulation

The overall performance of the proposed technique analyzed its performance using various performance metrics with its formulation are described as follows.

5.4.6 Accuracy

Accuracy is the performance metric utilized to evaluate the accurate detection of the model, which is described below.

$$accuracy = \frac{tp + tn}{tp + fp + tn + fn} \tag{63}$$

Here, the measurement of true positive and true negative is represented as and. The values of false positive as well as false negative are denoted as as well as.

Precision: The performance parameter known as precision is evaluated using the equation that follows, which represents the ratio of positive prediction to total positive prediction.

$$precision = \frac{tp}{tp + fp} \tag{64}$$

Recall: The ratio of positive prediction analyzed based on the sum of true positives as well as false negative prediction is known as recall or sensitivity, which is described in the equation below.

$$recall = \frac{tp}{tp + fn} \tag{65}$$

F1-score

F1-score is measured based on the average precision and recall, which is described in the equation below.

$$F1 = \frac{precision \times recall}{precision + recall} \tag{66}$$

Specificity

Specificity is a performance utilized to evaluate the ratio of actual negative cases from correctly predicted data, which is described in the below equation.

$$specificity = \frac{tn}{tn + fp} \tag{67}$$

MAE

The Mean Absolute Error (MAE) is evaluated as the difference between the actual as well as predicted value, which is described in the below equation.

$$mae = \frac{1}{x} = \sum_{i=1}^{x} \left(e_1 - \hat{e}_i\right) \tag{68}$$

Here, the actual and predicted values are represented as and.

MSE

The Mean Square Error (MSE) evaluates the average square difference between actual and predicted values, which is shown below.

$$mse = \frac{1}{x} = \sum_{i=1}^{x} \left(e_1 - \hat{e}_i\right)^2 \qquad (69)$$

RMSE

The root mean square error (RMSE) is the square root of MSE, which is described in the equation below.

$$rmse = \frac{1}{x} = \sum_{i=1}^{x} \left(e_1 - \hat{e}_i\right)^2 \qquad (70)$$

Performance evaluation of the proposed technique

The overall performance of the suggested technique was analyzed using various related techniques, such as LSTM, ResNet18, AlexNet, and CNN.

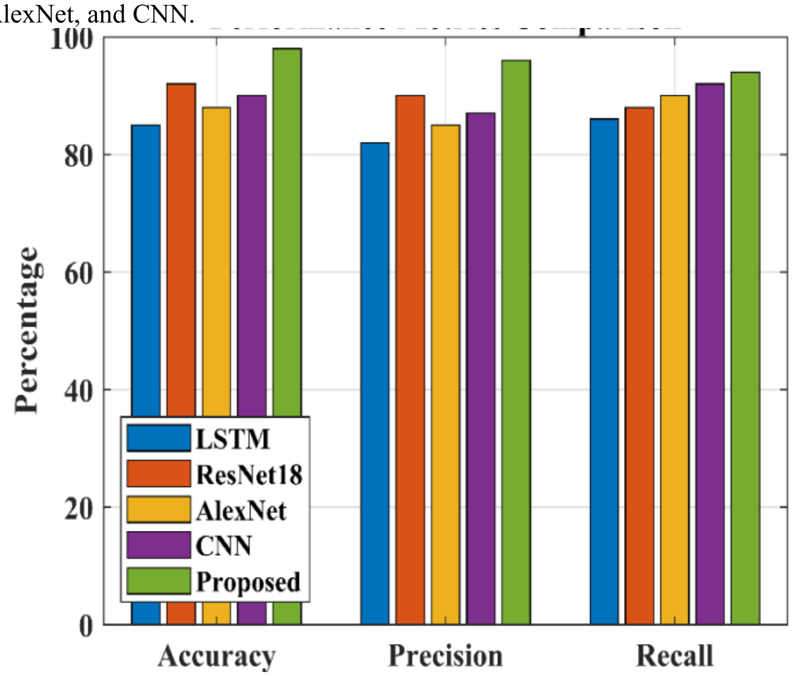

Figure 6: Comparison of Accuracy, Precision and Recall

By contrasting it with these methods, the suggested technique's effective performance is explained. The comparison of accuracy, precision, and recall utilizing the suggested and related methodologies is shown in Figure 6.

Initially, accuracy, precision, as well as recall of the proposed technique, were analyzed, and different existing techniques were compared to determine efficient performance. Figure 4 clearly shows that the suggested method achieves higher accuracy than other related methods. The proposed method achieves 98% accuracy, 96.52% precision, and 94.03% recall. The proposed technique provides efficient brain tumor segmentation and prediction by achieving this performance. The accuracy of other related techniques is 90% for CNN, 88.34% for AlexNet, 92.01% for ResNet18, and 85.34% for LSTM. Similarly, the precision and recall rate of these existing techniques is lower than that of other techniques. The comparison of F1-score, Specificity, and Kappa are analyzed and represented in Figure 7.

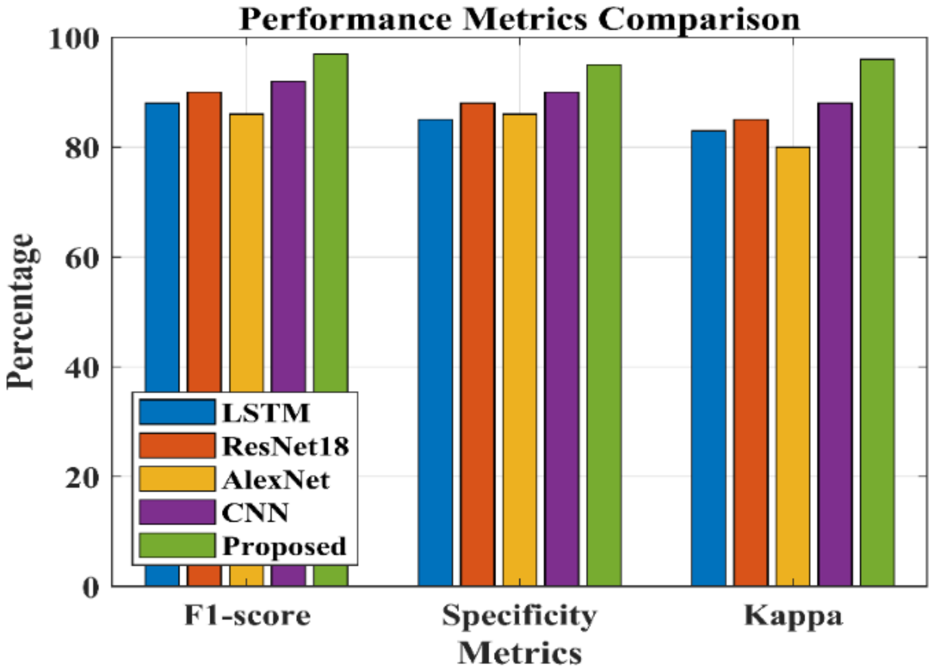

Figure 7: Comparison of F1-score, Specificity and Kappa

The F1-score, Specificity, and Kappa are analyzed using various existing techniques to determine how well the proposed approach performs. By examining the accuracy, Figure 5 illustrates how the suggested strategy outperforms other relevant strategies. The proposed approach yields 97% F1-score, 95% Specificity, and 96% Kappa. By accomplishing these goals, the proposed approach offers effective brain tumor segmentation and prediction. The F1-score of the proposed technique achieves efficient performance compared to LSTM (88%), ResNet18 (90%), AlexNet (86%), and CNN (92%). Similarly, the Specificity and Kappa rate of the suggested technique is higher than other related current methods. Figure 8 denotes accuracy as well as loss rate of training as well as testing.

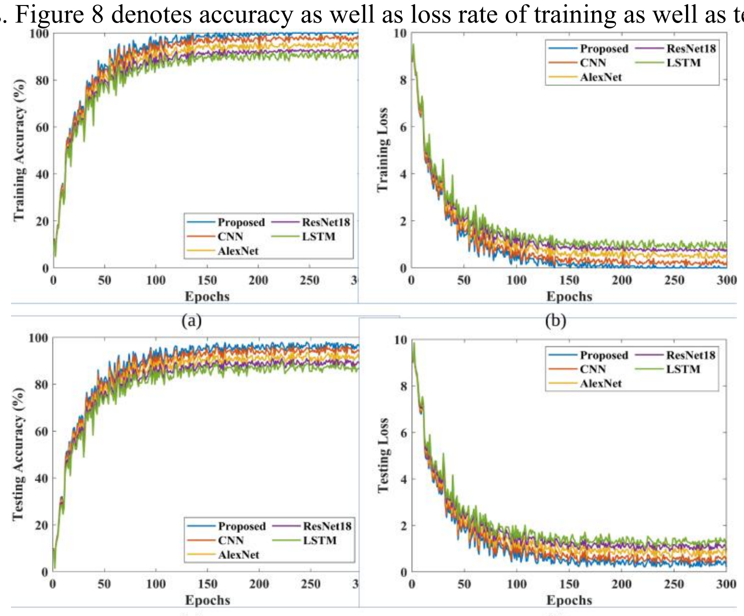

Figure 8: Performance of (a) Training Accuracy, (b) Training

The accuracy, as well as loss rates of the performance, are analyzed for both training and testing to detect the brain tumor effectively. In this work, the data is split into a 70:30 ratio for training as well as testing. It describes that

70% of data from the dataset are utilized for training the model, and the remaining 30% is utilized for testing. The entire effectiveness of the suggested technique will be assessed, and its superiority will be established based on this training and testing process. When compared to other related procedures, Figures 6(a) and (b) clearly show that the accuracy of both training and testing marginally improves its performance. In the same way, the suggested technique's loss is minimized during testing and training, indicating its effectiveness. The comparison of the MAE, MSE, and RMSE analyses is shown in Figure 7.

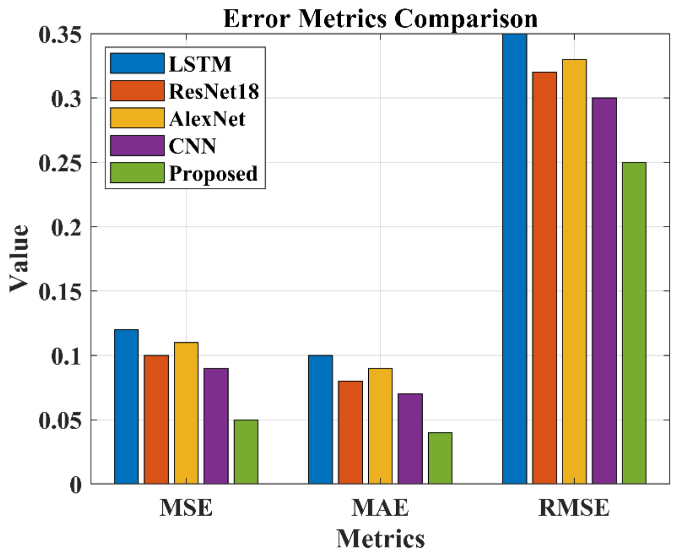

Figure 9: Comparison of MSE, MAE and RMSE

The performance of error metrics, such as MSE, MAE, and RMSE, was analyzed using various related techniques to determine the efficient performance of the suggested method. The error rate of the suggested technique was analyzed using related techniques, including LSTM, ResNet18, AlexNet, and CNN. The proposed technique achieves 0.05 MSE, 0.04 MAE, and 0.25 RMSE, which are lower than those of other related techniques. Figure 7 clearly describes the efficiency of the suggested approach compared to other related techniques. The comparison of the ROC curve is represented in Figure 8.

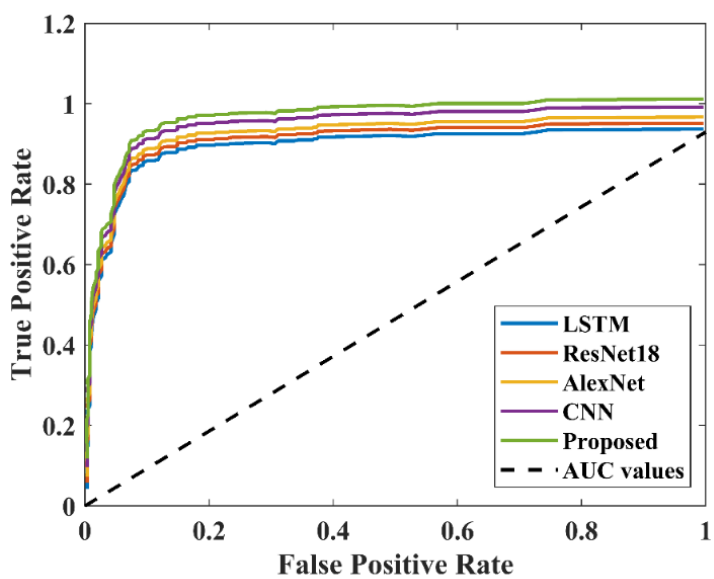

Figure 10: Comparison of ROC curve

The Receiver operating characteristic (ROC) curve of the proposed technique is analyzed based on the AUC value, which is evaluated based on the Area under the ROC curve (AUC) rate. Then, the comparison is conducted to evaluate the efficiency of the suggested technique with other related techniques such as LSTM, ResNet18, AlexNet, and CNN. Figure 8 clearly describes the superiority of the suggested approach as well as other related approaches by comparing the AUC value. The confusion matrix of the suggested technique is represented in Figure 11.

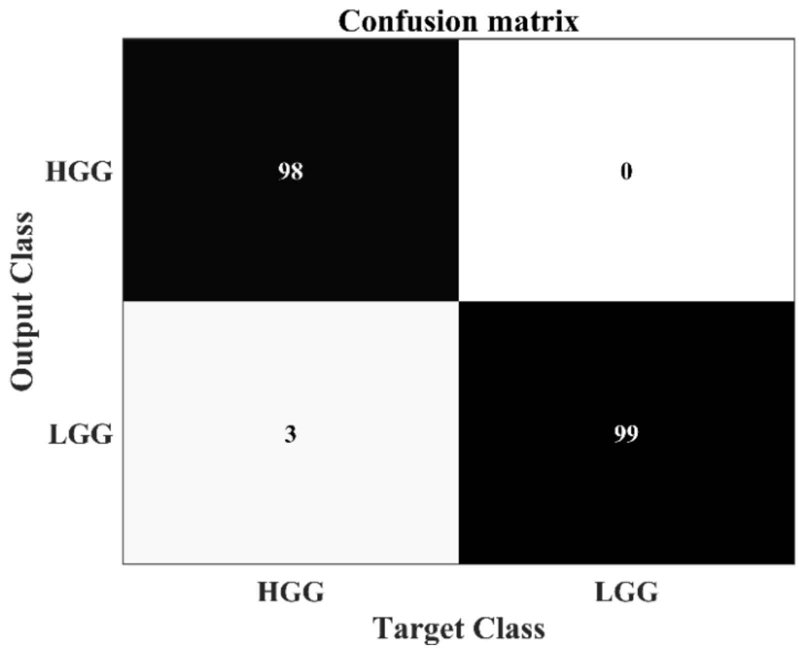

Figure 11: Confusion matrix of the suggested technique

Summary of Key Findings

This study presented a novel frequency-domain feature extraction framework tailored for advanced machine vision applications, leveraging the Fourier Transform to capture discriminative spectral characteristics of visual data. The results demonstrated that frequency-domain features effectively represent global image structures, periodic textures, and repetitive patterns that are often inadequately captured by purely spatial-domain techniques. Experimental evaluations confirmed improved feature robustness, noise tolerance, and classification accuracy when frequency-domain descriptors were integrated with machine learning models, making the proposed framework suitable for tasks such as object recognition, texture analysis, and industrial inspection.

Limitations of Fourier-Based Analysis

Despite its strengths, Fourier-based analysis exhibits certain limitations. The primary drawback is the loss of spatial localization, as standard Fourier Transform techniques do not preserve precise positional information of image features. This limitation affects performance in applications requiring fine spatial detail or localized anomaly detection. Additionally, Fourier-based methods assume signal stationarity, which may not hold true for complex real-world images with abrupt transitions. High computational cost for large-scale images and sensitivity to boundary effects also pose challenges in real-time machine vision systems.

Ethical and Societal Impact

The adoption of frequency-domain feature extraction in machine vision systems has significant ethical and societal implications. Enhanced vision-based automation improves manufacturing efficiency, quality control, and safety monitoring, contributing positively to industrial productivity. However, its use in surveillance and biometric systems raises concerns related to data privacy, algorithmic transparency, and responsible deployment. Ensuring ethical usage requires adherence to fair data practices, explainable algorithms, and compliance with regulatory standards to avoid misuse or biased decision-making.

Future Research Opportunities

Future research can extend this work by integrating Fourier-based features with time–frequency methods such as Wavelet Transform, Short-Time Fourier Transform (STFT), and hybrid deep learning architectures. Combining frequency-domain representations with convolutional neural networks (CNNs) and transformer-based vision models can further enhance adaptability and accuracy. Additionally, exploring real-time optimization, edge-computing implementations, and domain-specific customization for medical imaging, autonomous systems, and smart manufacturing presents promising directions for continued advancement.

Concluding Statements

In conclusion, the proposed frequency-domain feature extraction framework demonstrates strong potential in addressing the growing complexity of modern machine vision tasks. By exploiting spectral information, the framework enhances feature discrimination, robustness, and generalization capability across diverse applications. While limitations remain, strategic integration with complementary methods can significantly mitigate these challenges.

Concluding Remark

Overall, "A Novel Frequency-Domain Feature Extraction Framework for Machine Vision Applications" establishes a solid foundation for next-generation vision systems by effectively harnessing Fourier-based spectral analysis. The framework contributes meaningful insights to both academic research and practical deployment, paving the way for scalable, intelligent, and ethically responsible machine vision solutions in future technological ecosystems.

Compliance with Ethical Standards

Funding

No funding is provided for the preparation of the manuscript.

Conflict of Interest

Authors declare that they have no conflict of interest.

Ethical Approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Consent to participate: All the authors involved have agreed to participate in this submitted article.

Consent to Publish: All the authors involved in this manuscript give full consent for publication of this submitted article.

Authors Contributions: All authors read and approved the final manuscript.

Data Availability Statement: Data sharing is not applicable to this article.

Reference

1. Mainak Chakraborty et al., Frequency domain complex-valued convolutional neural network" Expert Systems with Applications Volume 295, 1 January 2026, https://doi.org/10.1016/j.eswa.2025.128893
2. Khadija Slama et al., "Comprehensive review of machine learning and deep learning techniques for epileptic seizure detection and prediction based on neuroimaging modalities" Open Access, 2025. https://doi.org/10.1186/s42492-025-00208-8
3. Albert Dede et al., "Wavelet-Based Feature Extraction for Efficient High-Resolution Image Classification" WILEY, Volume7, Issue2 ,2025, page no.70027.

https://doi.org/10.1002/eng2.70027

4. Kaibo Tang et al., "A Selection of Distributions and Their Fourier Transforms with Applications in Magnetic Resonance Imaging" 2025.
5. Sora Togawa et al., "Understanding convolutional neural networks through Z-transform: Frequency domain analysis of kernel feature extraction" IEICE, vol. 16, no. 4, 2025, pp. 878 -895.
6. Patrik Jan Gallinnis et al., "Investigation of Endogenous Renal CEST Contrast and the Influence of Respiratory Motion on a Clinical 3 Tesla MRI: An In Vivo and In Vitro Study" WILEY, 2025.

https://doi.org/10.1002/mrm.70210

7. A. Emily Jenifera et al., “Enhanced Audio Signal Classification with Explainable AI: Deep Learning Approach in Time and Frequency Domain Analysis Procedia Computer Science 258 ,2025 2372–2381.

8. Tianyi Yang et al., “Efficient 3D magnetic resonance image reconstruction by 2D transformers and attention-based fusion model” Biomedical Signal Processing and Control, Volume 110, Part B, 2025, 108071.
https://doi.org/10.1016/j.bspc.2025.108071

9. Subashini et al., “Applications of Fourier Transforms in Wave Mechanics and Signal Processing” IJSDR, Volume 10, Issue ,2025, ISSN: 2455-2631.

10. Ruru Xu et al., “A reinforcement learning approach for optimized MRI sampling with region-specific fidelity” Neurocomputing, Elsevier, Volume 638, 2025, 130116. https://doi.org/10.1016/j.neucom.2025.130116

11. Honggui Han et al., “Self-organizing broad network with frequency-domain analysis” Engineering Applications of Artificial Intelligence, Volume 137, Part A, 2024, 109155
https://doi.org/10.1016/j.engappai.2024.109155

12. Vaibhav Garg et al., “Spatial spectral contextual image analysis approach to map water quality of inland waterbody using AVIRIS-NG datasets” Advances in Space Research Volume 73, Issue 2, 2024, Pages no. 1549-1559 https://doi.org/10.1016/j.asr.2022.09.012

13. Changheun Oh et al., “Domain transformation learning for MR image reconstruction from dual domain input” Computers in Biology and Medicine, Volume 170, 2024, 108098.
https://doi.org/10.1016/j.compbiomed.2024.108098

14. Xiaohan Liu et al., “Dual-domain faster Fourier convolution based network for MR image reconstruction” Computers in Biology and Medicine, Elsevier, Volume 177, 2024, 108603.
https://doi.org/10.1016/j.compbiomed.2024.108603

15. Haozhong Sun et al., “Fourier Convolution Block with global receptive field for MRI reconstruction” Medical Image Analysis, Elsevier, Volume,2025, 103349. https://doi.org/10.1016/j.media.2024.103349

16. yanting Yang et al ., “Deep learning-based MRI reconstruction with Artificial Fourier Transform Network (AFTNet)” Computers in Biology and Medicine, Elsevier, Volume 192, Part B, 2025, 110224. https://doi.org/10.1016/j.compbiomed.2025.110224

17. Jianpu Lin et al., “Frequency-guided dual-collapse Transformer for low-light image enhancement” Engineering Applications of Artificial Intelligence, Volume 142, 2025, 109906.
https://doi.org/10.1016/j.engappai.2024.109906

18. L. Malathi et al., “FPGA design of FFT based intelligent accelerator with optimized Wallace tree multiplier for image super resolution and quality enhancement” Biomedical Signal Processing and Control, Volume 88, Part B, 2024, 105599.
https://doi.org/10.1016/j.bspc.2023.105599

19. Heng Jiang et al., “High-Fidelity compressive spectral image reconstruction through a novel Non-Convex Non-Local Low-Rank tensor approximation model” Optics & Laser Technology Volume 171, 2024, 110424.
https://doi.org/10.1016/j.optlastec.2023.110424

20. Talha Ilyas et al., “Enhancing medical image analysis with unsupervised domain adaptation approach across microscopes and magnifications” Computers in Biology and Medicine, Elsevier, Volume, 2024, 108055. https://doi.org/10.1016/j.compbiomed.2024.108055

21. Changheun Oh et al., “Domain transformation learning for MR image reconstruction from dual domain input” Computers in Biology and Medicine, Volume 170, 2024, 108098.
https://doi.org/10.1016/j.compbiomed.2024.108098

22. Xuyang Liu et al., “Deep learning in spectral analysis: Modeling and imaging” TRAC Trends in Analytical Chemistry , Elsevier, Volume 172, 2024, 117612. https://doi.org/10.1016/j.trac.2024.117612

23. Masaaki Inoue et al., “Enhanced frequency and time domain feature extraction for communication infrastructure type classification using optical fiber sensing” Optical Fiber Technology, Volume 87, 2024, 103859. https://doi.org/10.1016/j.yofte.2024.103859

24. Anuja Bhargava et al., “ Hyperspectral imaging and its applications: A review” Review article pen access ,Volume 10, Issue 12e 33208, 30, 2024 .
https://doi.org/10.1016/j.heliyon.2024.e33208

25. Qing-Yuan Wu et al., “An edge detail enhancement strategy based on Fourier single-pixel imaging” Optics and Lasers in Engineering, Elsevier, Volume 172, 2024, 107828. https://doi.org/10.1016/j.optlaseng.2023.107828

26. Liu Zhou et al., “RNLFNet: Residual non-local Fourier network for undersampled MRI reconstruction” Biomedical Signal Processing and Control, Elsevier, Volume, 2023, 104632.
https://doi.org/10.1016/j.bspc.2023.104632

27. Oren Solomon et al., “Motion robust magnetic resonance imaging via efficient Fourier aggregation” Medical Image Analysis, Elsevier, Volume 83, 2023, 102638. https://doi.org/10.1016/j.media.2022.102638
28. Qi Han et al., “EHDFL: Evolutionary hybrid domain feature learning based on windowed fast Fourier convolution pyramid for medical image classification” Computers in Biology and Medicine, Elsevier, Volume 152, 2023, 106353.
https://doi.org/10.1016/j.compbiomed.2022.106353
29. Zhengliang Wu et al., “Deep learning based MRI reconstruction with transformer” Computer Methods and Programs in Biomedicine, Elsevier, Volume, 2023, 107452.
https://doi.org/10.1016/j.cmpb.2023.107452.
30. Norikazu Koori et al., “Influence of half Fourier and elliptical scanning (radial scan) on magnetic resonance images” Journal of Magnetic Resonance, Elsevier, Volume 355, 2023, 107560.
https://doi.org/10.1016/j.jmr.2023.107560 .
31. Elena Escobar-Linero et al., “On the feature extraction process in machine learning. An experimental study about guided versus non-guided process in falling detection systems” Engineering Applications of Artificial Intelligence ,Volume 114, 2022, 105170.https://doi.org/10.1016/j.engappai.2022.105170
32. Yang Gaoet al., “Research on algorithm and implementation of Fourier transform in the case of N < M” CISAT 2020, oid:10.1088/1742-6596/1634/1/012067
33. Chuan Tong et al., “HIWDNet: A hybrid image-wavelet domain network for fast magnetic resonance image reconstruction” Computers in Biology and Medicine , Elsevier, Volume 151, Part A, December 2022, 105947 https://doi.org/10.1016/j.compbiomed.2022.105947
34. Elias Rajaby et al ., “A structured review of sparse fast Fourier transform algorithms” Digital Signal Processing, Elsevier, Volume 123,2022, 103403.
https://doi.org/10.1016/j.dsp.2022.103403
35. Muhammad Awais et al., “Energy Efficient Design of Fast Fourier Transform using Reversible Logic” Sustainable Energy Technologies and Assessments, Volume 52, Part C, 2022, 102257.
https://doi.org/10.1016/j.seta.2022.102257
36. Soobia Saeed et al., “Improve correlation matrix of Discrete Fourier Transformation technique for finding the missing values of MRI images” Elsevier, 2022.
37. Dong Wei et al., “Noise reduction method in both time and frequency domains for envelope peak position extraction of pulse-train interferometry measuring distance between interference fringes” Optics Communications, Elsevier, Volume 523, 2022, 128730. https://doi.org/10.1016/j.optcom.2022.128730
38. Shengke Xue et al., “2D probabilistic undersampling pattern optimization for MR image reconstruction” Medical Image Analysis, Elsevier, Volume 77, 2022, 102346. https://doi.org/10.1016/j.media.2021.102346
39. Mohsen Mousavi et al., “Machine learning analysis of features extracted from time–frequency domain of ultrasonic testing results for wood material assessment” Construction and Building Materials, Volume 342, Part A, 2022, 127761.
https://doi.org/10.1016/j.conbuildmat.2022.127761
40. Jiangyi Qin et al., “Interference Signal Feature Extraction and Pattern Classification Algorithm Based on Deep Learning” Elsevier, 2022.
https://doi.org/10.3390/electronics11142251
41. Songyang Zhang et al., “Multilayer graph spectral analysis for hyperspectral images” EURASIP Journal on Advances in Signal Processing, Open Access, 2022.
https://doi.org/10.1186/s13634-022-00926-8
42. Quanshan Gao et al., “Spectral and spatial reduction of hyperspectral image guided by data reconstruction and super pixels” Engineering Applications of Artificial Intelligence, Elsevier , Volume 111, 2022, 104803. https://doi.org/10.1016/j.engappai.2022.104803
43. M. Farooq Wahab et al., “Discrete Fourier transform techniques for noise reduction and digital enhancement of analytical signals” TrAC Trends in Analytical Chemistry. Volume 143, 2021, 116354.
https://doi.org/10.1016/j.trac.2021.116354.
44. Srinivasu Polinati et al., “The Fusion of MRI and CT Medical Images Using Variational Mode Decomposition” 2021. https://doi.org/10.3390/app112210975
45. Amir Aghabiglou et al., “Projection-Based cascaded U-Net model for MR image reconstruction” Computer Methods and Programs in Biomedicine, Elsevier, Volume 207, 2021, 106151.
https://doi.org/10.1016/j.cmpb.2021.106151

46. Venkata Salini Priyamvada Davuluru et al., “Performance Analysis of DFT and FFT Algorithms on Modern GPUs” MDPI, 2021.
47. Carlo Ciulla et al., “Inverse Fourier transformation of combined first order derivative and intensity-curvature functional of magnetic resonance angiography of the human brain” Computer Methods and Programs in Biomedicine, Elsevier, Volume, 2021, 106384.
https://doi.org/10.1016/j.cmpb.2021.106384
48. Andreia Valentina Miclea et al., “On Spectral-Spatial Classification of Hyperspectral Images Using Image Denoising and Enhancement Techniques, Wavelet Transforms and Controlled Data Set Partitioning” MDPI, 2021. https://doi.org/10.3390/rs14061475
49. M. Farooq Wahab et al., “Discrete Fourier transform techniques for noise reduction and digital enhancement of analytical signals” TrAC Trends in Analytical Chemistry. Volume 143, 2021, 116354.
https://doi.org/10.1016/j.trac.2021.116354.
50. Xianlun Tang et al., “Introducing frequency representation into convolution neural networks for medical image segmentation via twin-Kernel Fourier convolution” Computer Methods and Programs in Biomedicine, Elsevier, Volume 205, 2021, 106110.
https://doi.org/10.1016/j.cmpb.2021.106110
51. Nurudeen Alegeh et al., “Feature Extraction of Time-Series Data Using DWT and FFT for Ballscrew Condition Monitoring” Elsevier,2021.
52. Anam Khanzadah et al., “Exploring Digital Image Processing: A Study on Frequency Domain Filtering Techniques” Elsevier,2020.
53. Mohammed H. Rasheed et al., “Image compression based on 2D Discrete Fourier Transform and matrix minimization algorithm” Array, Volume 6, 2020, 100024.
https://doi.org/10.1016/j.array.2020.100024
54. Hadhrami Ab et al., “A review on sparse Fast Fourier Transform applications in image processing” IJECE, Vol. 10, No. 2, 2020, ISSN: 2088-8708, pp1346-1351.
55. Li Wang et al., “Temporal-spatial-frequency depth extraction of brain-computer interface based on mental tasks” Biomedical Signal Processing and Control , Elsevier, Volume 58, 2020, 101845.
56. Lijun Bao et al., “Under sampled MR image reconstruction using an enhanced recursive residual network” Journal of Magnetic Resonance, Volume 305, Elsevier, 2019, Pages 232-246.
https://doi.org/10.1016/j.jmr.2019.07.020
57. Agon Kokaj et al., “Fast Fourier Transform and a Complementary Filter Based Control of a Robotic System” IFAC Papers On Line 51-30 ,2018, 561–564.
58. Yi Lu et al., “On time-frequency domain feature extraction of wave signals for structural health monitoring” Measurement, Volume 114, 2018, Pages 51-59. https://doi.org/10.1016/j.measurement.2017.09.016
59. Mohamed H. M. Ali et al., “Fourier-Transform Infrared Imaging Spectroscopy and Laser Ablation -ICPMS New Vistas for Biochemical Analyses of Ischemic Stroke in Rat Brain” Original Research, 2018
60. M.E. Komlosh et al., “Anisotropic phantom to calibrate high-q diffusion MRI methods” Journal of Magnetic Resonance, Volume 275, 2017, Pages 19-28. https://doi.org/10.1016/j.jmr.2016.11.017
61. V. G. Bochkareva et al., “Methods for Improving Image Quality Using Spatial Spectral Analysis” Elsevier, Vol. 54, No. 6, 2015, pp. 897–904.
62. K.M.M. Prabhu et al., “3-D warped discrete cosine transform for MRI image compression” Biomedical Signal Processing and Control, Elsevier, Volume, 2013, Pages 50-58. https://doi.org/10.1016/j.bspc.2012.04.004
63. Praveen Shrivastava et al., “Fourier Transform and Applications in Image Processing” J. Comp. & Math. Sci. Vol.4 (1), 1-3 ,2013.
64. S. Arunachalam et al., “The Fast Fourier Transform Algorithm and Its Application in Digital Image Processing” ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online) Vol.3, No.6, 2013.
65. Mridul Kumar Mathur et al., “image compression using dft through fast Fourier transform technique” Volume 1, Issue 2,2012.
66. Tomáš Hararik et al., “Frequency analysis of acoustic signal using the Fast Fourier Transformation in MATLAB” Procedia Engineering 48, Elsevier, 2012, 199 – 204.
67. Abiodun M. Aibinu et al., “MRI Reconstruction Using Discrete Fourier Transform: A tutorial” International Journal of Information, Control and Computer Sciences ISSN: 2517-9942, Vol:2, No:6, 2008.