Minimax Theory and its Applications

Volume 11 (2026), No. 5s, 163-179

SGLN-Net: Multimodal Feature Learning with Improved Self-Guided Filtering and a Hybrid Improved LeNet–SqueezeNet Architecture for Parkinson's Disease Detection Using MRI

Dr.G. Suresh Reddy[1], P L N Manoj Kumar[2], Dr. Hari Kishan Chapala[3], Dr S V N Sreenivasu[4], K.K.Nithiyanantham[5], Peddireddy Veera Venkateswara Rao[6], Elangovan Muniyandy[7], Mr. Anil Kumar Katta[8], Srija Gundapaneni[9]

[1]Professor, Department of Information Technology, VNR Vignana Jyothi Institute of Engineering & Technology Hyderabad, Telangana, Email : gali.sureshreddy@gmail.com, Orcid id: 0000-0001-5214-2690
[2]Assistant Professor, Department of Advanced Computer Science and Engineering, Vignan's Foundation for Science, Technology & Research, Vadlamudi, Guntur, Andhra Pradesh. Email : manojnanda2000@gmail.com
[3]Professor & Head, Department of CSE - AI & ML St. Ann's College of Engineering & Technology Chirala Bapatla district Andhra Pradesh, Email : drchkishan@gmail.com, Orcid Id:0000-0003-0842-7976
[4]Professor , Dept Of Computer Science And Engineering Narasaraopeta Engineering College Narasaraopet , Email:drsvnsrinivasu@gmail.com, Orcid id 0000-0002-6049-911X
[5]Associate Professor, Department of Aeronautical Engineering, Rajalakshmi engineering college Thandalam, Tamilnadu. Email:nithiyanantham.kk@rajalakshmi.edu.in, Orcid:0009-0007-2094-2694.
[6]Associate Professor, Department of computer science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522502, Guntur District,Andhra Pradesh, India, Email : raovenkat21@gmail.com
[7]Department of Biosciences, Saveetha School of Engineering. Saveetha Institute of Medical and Technical Sciences,Chennai - 602 105, Email: muniyandy.e@gmail.com
[8]Assistant Professor , Department Of Ece Rvr&Jc College Of Engineering, Guntur Pincode: 522019, Email:anilecedepartment@gmail.com, Orcid id: 0000-0003-1148-9107
[9]Assistant professor , Department of CSE -IOT, R.V.R.&J.C College of Engineering ,Chowdavaram Guntur , Email:gundapaneni.srija@gmail.com

Abstract
Structural magnetic resonance imaging (MRI) is widely available and non-invasive, but the changes that Parkinson's disease (PD) produces on conventional T1-weighted scans are subtle, and they are easily masked by Rician noise, intensity inhomogeneity and inter-scanner variability. This study proposes SGLN-Net, a multimodal feature learning framework that couples image enhancement with complementary handcrafted and deep representations. First, an Improved Self-Guided Filter (ISGF) denoises each slice using the slice itself as guidance, with edge-aware adaptive regularization, iterative refinement and explicit correction of the Rician noise bias. Second, three feature modalities are extracted from the enhanced slices: texture features by an Improved Local Phase Optimization (ILPO) scheme that combines multi-scale decorrelated local phase quantization with monogenic local-phase statistics and tunes its parameters with a Grey Wolf Optimizer; shape features from geometric descriptors, Hu invariants and Zernike moments; and deep features from two branches. Third, a hybrid classifier fuses an Improved LeNet branch with residual squeeze-and-excitation blocks, an ImageNet-initialized SqueezeNet branch and the handcrafted vector through learned branch gating, and subject-level predictions are obtained by averaging slice probabilities. On a Shepp–Logan phantom corrupted by Rician noise ($\sigma = 0.05$), ISGF raised PSNR from 24.14 dB to 33.92 dB and SSIM from 0.258 to 0.884 over 10 noise realizations, outperforming Gaussian, non-local means and standard self-guided filtering, with most of the gain attributable to Rician bias correction. A leakage-free evaluation protocol on PPMI T1-weighted scans is defined, with subject-level stratified cross-validation, component-wise ablation and re-implemented reference networks; the classification experiments on patient data are in progress, and the paper states the hypotheses and success criteria against which they will be judged.

Keywords— Parkinson's Disease Detection; Structural MRI; Self-Guided Filtering; Rician Noise; Local Phase Quantization; Grey Wolf Optimizer; LeNet; SqueezeNet; Hybrid Deep Learning

Introduction

Parkinson's disease (PD) is the fastest-growing neurological disorder worldwide, and its global burden more than doubled between 1990 and 2016 [1]. The disease results from progressive degeneration of dopaminergic

neurons in the substantia nigra and presents with bradykinesia, rigidity, rest tremor and postural instability, together with a broad range of non-motor symptoms [2]. Diagnosis remains clinical and relies on structured examinations such as the Movement Disorder Society-sponsored Unified Parkinson's Disease Rating Scale [3]. These assessments depend on specialist expertise, are subject to inter-rater variability, and are least reliable in early disease, when motor signs are mild and may be confused with normal ageing.

Structural magnetic resonance imaging (MRI) is attractive as an objective complement because it is non-invasive, widely available and already part of the routine work-up of movement disorders. Deep learning has been applied extensively to T1-weighted MRI for PD detection, yet the reported performance varies widely. A large multi-center study that trained an explainable convolutional neural network (CNN) on T1-weighted data noted that published accuracies span roughly 71.5% to 100% [4], and a recent multi-center 3D CNN reached 74% accuracy for moderate-to-severe PD and only 64% for mild PD when imaging was used alone [5]. At the same time, hybrid CNN–transformer models [6] and EfficientNet-attention frameworks [7] report accuracies of about 97% on PPMI data. Part of this spread is explained by evaluation design: when slices of the same subject appear in both training and test sets, 2D CNNs can recognize individuals rather than disease, which inflates accuracy substantially [8].

Image quality is a second, less discussed factor. Magnitude MRI follows a Rician rather than Gaussian noise distribution, which introduces a signal-dependent bias in low-intensity regions [9]. PD-related changes on T1-weighted images, such as subtle atrophy in subcortical and cortical regions, have low contrast, so noise and bias can obscure precisely the textures a classifier needs. Edge-preserving filters such as the guided filter [10] are efficient and preserve anatomical boundaries, but they assume additive noise and use a fixed regularization that over-smooths fine structures in some regions and under-smooths noise in others. Likewise, texture descriptors based on local phase, such as local phase quantization (LPQ) [11], are robust to blur and illumination changes, but they are usually applied at a single, manually chosen scale.

Research Motivation

These observations motivate a framework that treats enhancement and representation as a single design problem. If the enhancement stage removes Rician bias while preserving edges, the subsequent texture, shape and deep features can capture subtle structure rather than noise. If texture is described by local phase at several optimized scales, it can complement the learned features of CNNs, which tend to emphasize intensity patterns. Lightweight networks such as SqueezeNet [12] and LeNet [13] are well suited to the small datasets typical of neuroimaging, provided that their capacity and regularization are adapted to the task.

Research Significance

This research is significant because it addresses two causes of unreliable MRI-based PD detection simultaneously: degraded image quality and optimistic evaluation. The proposed framework combines physically motivated denoising with interpretable handcrafted features and compact deep networks, and its evaluation protocol operates strictly at the subject level so that performance estimates are not inflated by identity leakage. The work also establishes the image-enhancement and hybrid-feature foundation for the subsequent multimodal and explainable stages of this research programme.

Problem Statement

Existing MRI-based PD detectors face four limitations. Preprocessing pipelines rarely model the Rician statistics of magnitude MRI, so the noise bias remains in the images used for learning [9]. Texture descriptors are applied at fixed scales and parameters that are not adapted to the data [11]. Most models rely on a single deep representation, although handcrafted shape and texture information can provide complementary evidence on small datasets. Finally, many studies split data at slice level, which overestimates performance [8], making it difficult to judge whether high reported accuracies would hold for new patients.

Recent Innovation and Challenges

Recent work has introduced transformer–CNN fusion [6], attention-augmented EfficientNet backbones [7] and multi-center 3D CNNs with transfer learning [5]. These advances improve representational power but increase model size and data requirements, and multi-center results show that accuracy drops considerably when evaluation is rigorous and early-stage patients are included [4], [5]. The open challenge is therefore to improve the signal available to the classifier and to combine complementary representations efficiently, while reporting performance that reflects generalization to unseen subjects.

Key Contribution

Proposes an Improved Self-Guided Filter (ISGF) with edge-aware adaptive regularization, iterative refinement and Rician bias correction, and validates it quantitatively against Gaussian, non-local means and standard self-guided filtering.

Introduces Improved Local Phase Optimization (ILPO), a multi-scale decorrelated LPQ and monogenic local-phase descriptor whose scales and weights are tuned by a Grey Wolf Optimizer on training data only.

Designs a hybrid classifier that fuses an Improved LeNet branch, a SqueezeNet branch and a handcrafted texture–shape branch through learned branch gating.

Defines a subject-level, leakage-free evaluation protocol with component-wise ablation and re-implemented reference networks, together with explicit hypotheses and success criteria.

Rest of the Paper

The rest of this paper is organized as follows: Section II reviews recent MRI-based and multimodal PD detection studies; Section III describes the proposed SGLN-Net; Section IV reports the validation of the enhancement and texture stages and the experimental design and expected results for PD classification; and Section V concludes the paper.

Related Works

Chakraborty et al. [14] trained a 3D CNN on T1-weighted MRI from 406 PPMI subjects and reported 95.29% accuracy, 94.3% average recall, 92.7% average precision and an ROC-AUC of 0.98. Their work showed that volumetric networks can learn PD-related patterns directly from structural MRI, but the cohort was drawn from a single database and no external validation was performed.

Solana-Lavalle and Rosas-Romero [15] combined voxel-based morphometry with feature selection and classical machine learning to classify PPMI T1-weighted scans, analyzing male and female subjects separately. The approach provides anatomically interpretable features, but it depends on accurate registration and segmentation, and its handcrafted pipeline does not exploit learned representations.

Erdaş and Sümer [16] proposed a fully automated pipeline in which brain extraction and registration were followed by 2D and 3D CNNs for PD detection and severity prediction from T1-weighted images. The work demonstrated end-to-end automation, although, as with many CNN pipelines, image noise was not modelled explicitly.

Islam et al. [17] combined clinical assessments with neuroimaging samples in a comprehensive artificial intelligence framework, comparing multiple machine learning models and deep networks. Desai et al. [18] classified 3D MRI volumes with deep learning for PD diagnosis. Both studies confirm the potential of neuroimaging features, while highlighting the dependence of performance on dataset size and preprocessing choices.

Camacho et al. [4] trained an explainable CNN on one of the largest multi-center collections of T1-weighted MRI and found that deformation (Jacobian) maps yielded better accuracy than intensity images alone. Basaia et al. [5] trained a multi-center 3D CNN with clinical variables and transfer learning; imaging alone reached 74% accuracy for moderate-to-severe PD versus controls and 64% for mild PD, and accuracy improved when clinical data were added. These two studies provide realistic estimates of what structural MRI can achieve under rigorous, multi-site evaluation.

Dentamaro et al. [19] integrated imaging and clinical data from PPMI in a multimodal deep learning model with explainable AI. Sar et al. [20] proposed MultiParkNet, which fused speech, drawing, neuroimaging and cardiovascular signals and reported 96.74% test accuracy. Hussain et al. [6] fused Swin transformer and CNN features with cosine-similarity-based integration for MRI classification and reported approximately 97% accuracy, and Alrawis et al. [7] combined U-Net segmentation, autoencoder denoising, EfficientNet and attention, reaching 97.2% accuracy on PPMI.

Denoising of magnitude MRI has its own literature. Non-local means [21] averages similar patches across the image and was adapted to 3-D MRI with an optimized blockwise implementation [22] and to spatially varying Rician noise [23]. These methods are effective but computationally heavier than guided filtering, and they are rarely integrated into PD classification pipelines. For texture, local binary patterns [24] remain the most common handcrafted descriptor in medical imaging, whereas local phase descriptors, which are insensitive to blur and smooth intensity changes [11], have received little attention for PD.

Overall, the literature shows a shift from handcrafted morphometry to deep and hybrid architectures, but with two recurring gaps. Denoising, when present, is generic or computationally heavy and is not tailored to the classification task, and the combination of interpretable handcrafted texture and shape descriptors with compact deep networks has rarely been explored. Moreover, the large differences between single-cohort and multi-center results [4], [5] call for subject-level evaluation.

Table 1. Comparison Table

Author	**Method**	**Data**	**Strengths**	**Limitations**
Chakraborty et al. [14]	3D CNN on T1-weighted volumes	PPMI, 406 subjects	High AUC; volumetric context	Single database; no external test
Solana-Lavalle & Rosas-Romero [15]	VBM + feature selection + ML	PPMI T1	Anatomically interpretable	Registration-dependent; no learned features
Erdaş & Sümer [16]	Automated 2D/3D CNN, severity regression	T1 MRI	Fully automated pipeline	Noise not modelled
Camacho et al. [4]	Explainable CNN on Jacobian maps	Large multi-center T1	Realistic generalization	Moderate accuracy
Basaia et al. [5]	Multi-center 3D CNN + clinical data	Multi-center T1	Prognosis; transfer learning	Low accuracy for mild PD
Hussain et al. [6]	Swin transformer + CNN fusion	MRI	High reported accuracy	Large model; data-hungry
Alrawis et al. [7]	U-Net, autoencoder, EfficientNet + attention	PPMI, OASIS, MIRIAD	Denoising and segmentation	Generic denoising; complex pipeline
Sar et al. [20]	MultiParkNet multimodal fusion	Speech, drawing, imaging, ECG	Multimodal; XAI	Requires many modalities

Table I summarizes the reviewed studies. Approaches range from morphometry with classical classifiers to large hybrid networks, and the most accurate single-cohort results contrast with the moderate multi-center performance. None of the reviewed works combines Rician-aware denoising with optimized local-phase texture, shape descriptors and compact hybrid networks, which is the gap addressed by SGLN-Net.

SGLN-Net: Proposed Methodology

SGLN-Net processes T1-weighted MRI in three stages, consistent with the research objective of hybrid deep learning on enhanced images: enhancement by ISGF, multimodal feature extraction (texture, shape and deep features), and hybrid classification. The overall pipeline is shown in Fig. 1.

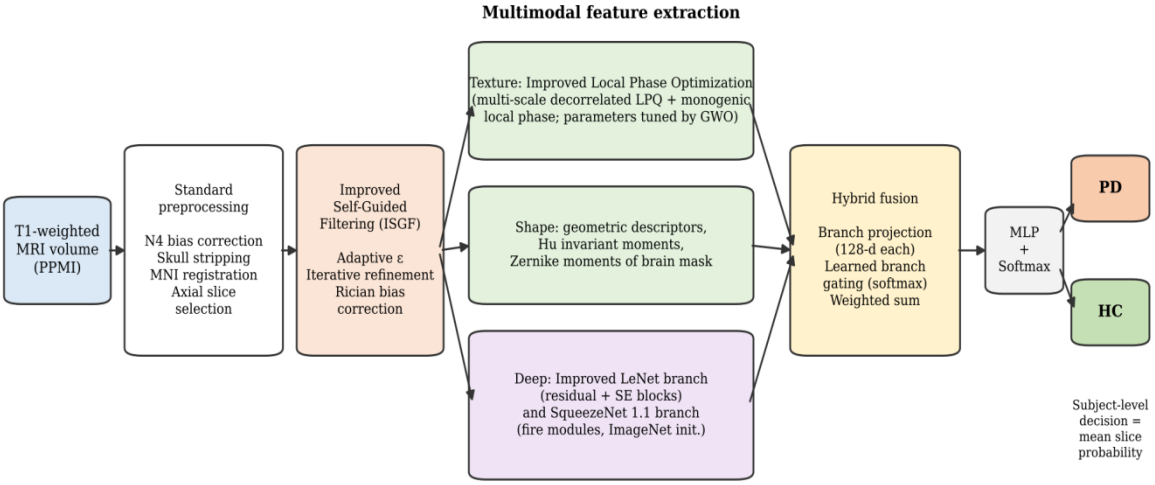


Fig. 1. Block Diagram of the Proposed SGLN-Net Methodology

Fig. 1. shows that each T1-weighted volume is bias-corrected, skull-stripped and registered, axial slices covering the basal ganglia and midbrain are extracted, and each slice is enhanced by ISGF. Texture, shape and deep features are then computed and combined by a gated fusion module, and slice probabilities are averaged to obtain a subject-level decision.

Data Collection and Standard Preprocessing

Baseline T1-weighted scans of PD patients and healthy controls are obtained from the Parkinson's Progression Markers Initiative (PPMI) [25]. Intensity inhomogeneity is corrected with N4 [26], non-brain tissue is removed with HD-BET [27], and volumes are affinely registered to MNI space. From each volume, 24 equally spaced axial slices between 38% and 58% of the superior–inferior extent, covering the basal ganglia and midbrain, are resampled to 128 × 128 pixels and normalized to the 1st–99th intensity percentiles of brain tissue.

Improved Self-Guided Filtering (ISGF)

Magnitude MRI is the modulus of a complex signal with independent Gaussian noise in both channels, so the observed intensity M follows a Rician distribution [9], as given in (1):

$$M = \sqrt{(A + n_r)^2 + n_i^2} \qquad (1)$$

where A is the true signal and nr, ni ~ N(0, σ²). In the guided filter [10], the output q is a local linear transform of a guidance image I within a window ωk of radius r, as in (2):

$$q_i = \bar{a}_i I_i + \bar{b}_i, \quad a_k = \frac{\mathrm{cov}_k(I,p)}{\sigma_k^2 + \varepsilon_k} \qquad (2)$$

where p is the input, bk = μp,k − akμI,k, and the bars denote averaging over all windows covering pixel i. In self-guided filtering, I = p. ISGF modifies this filter in three ways. First, the regularization is made edge-aware by scaling it with a normalized local-variance weight Γ, following weighted guided filtering [28], as given in (3):

$$\varepsilon_k = \frac{\varepsilon}{\Gamma(k)} \qquad (3)$$

$$\Gamma(k) = (\sigma_k^2 + \lambda)\frac{1}{N}\sum_j \frac{1}{\sigma_j^2 + \lambda} \qquad (4)$$

so that regularization is weak at anatomical edges (high local variance) and strong in homogeneous tissue. Second, the filter is applied iteratively, with the current estimate q(t) used as the guide for iteration t + 1, which progressively sharpens the guide while the input remains the original image. Third, the Rician bias is removed from the final estimate using the second-moment relation E[M²] = A² + 2σ², as in (5):

$$\hat{A} = \sqrt{\max\left(q^{(T)^2} - 2\hat{\sigma}^2, 0\right)} \qquad (5)$$

where $\hat{\sigma}$ is estimated from the signal-free background: background magnitudes follow a Rayleigh distribution whose mode equals σ, so $\hat{\sigma}$ is taken as the mode of the intensity histogram below the Otsu threshold. A window radius of 3, ε = 0.01 and T = 3 iterations are used.

Multimodal Feature Extraction

Texture: Improved Local Phase Optimization (ILPO)

LPQ [11] computes the short-term Fourier transform of an w × w neighbourhood at four low frequencies, as in (6):

$$F(u,x) = \sum_{y \in N_x} f(x-y) e^{-j2\pi u^T y} \qquad (6)$$

with u ∈ {[a,0], [0,a], [a,a], [a,−a]} and a = 1/w. The real and imaginary parts form an 8-dimensional vector that is decorrelated by whitening under a correlation model with coefficient ρ, and the signs of the whitened coefficients are binary-coded into 256 labels whose histogram describes texture. To capture structure at several scales, ILPO concatenates weighted LPQ histograms at three window sizes and adds statistics of the monogenic signal [29], whose local phase and orientation are computed with log-Gabor filters at two wavelengths, as in (7):

$$\varphi = \arctan\left(\frac{\sqrt{h_1^2 + h_2^2}}{f}\right) \qquad (7)$$

where f is the band-pass filtered image and h1, h2 are its Riesz transforms. Local phase is invariant to contrast and complements LPQ in the same way that phase congruency captures features independently of intensity [30]. The ILPO vector is given in (8):

$$v_{ILPO} = \left(\alpha_1 h_{w_1}, \alpha_2 h_{w_2}, \alpha_3 h_{w_3}, s_{\lambda_1}, s_{\lambda_2}\right) \qquad (8)$$

where hw is the LPQ histogram at window w and sλ contains the 8-bin phase histogram, the amplitude-weighted 8-bin orientation histogram and the log mean amplitude at wavelength λ. The nine parameters θ = (w1, w2, w3, α1, α2, α3, ρ, λ1, λ2) are optimized by the Grey Wolf Optimizer (GWO) [31], which moves each candidate towards the three best solutions α, β and δ, as in (9):

$$X^{(t+1)} = \frac{1}{3}\sum_{l \in \{\alpha,\beta,\delta\}} \left(X_l - A_l \left| C_l X_l - X^{(t)} \right|\right) \qquad (9)$$

where Al and Cl are random coefficient vectors whose range decreases linearly over iterations. The fitness is the mean Fisher score of the ILPO features, computed only on training slices of each fold, to which random Gaussian blur and Rician noise are applied so that scales that are stable under degradation are favoured (Section IV-B), as in (10):

$$J(\theta) = \frac{1}{D}\sum_{d=1}^{D} \frac{\sum_c n_c (\mu_{c,d} - \mu_d)^2}{\sum_c n_c \sigma_{c,d}^2} \qquad (10)$$

Eight wolves and 15 iterations are used, with window sizes constrained to odd values in [3, 15], ρ ∈ [0.70, 0.99] and λ ∈ [4, 32] pixels.

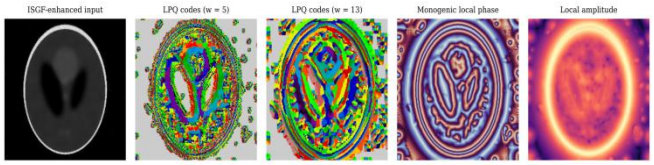


Fig. 2 ILPO Representations Computed on an ISGF-Enhanced Phantom Slice

Fig. 2. illustrates the ILPO representations on an enhanced phantom slice. Small LPQ windows respond to fine edges and noise-scale structure, larger windows encode coarser regional texture, and the monogenic phase and

amplitude highlight boundaries independently of local contrast, showing why the scales carry complementary information.

Shape Features

A brain-tissue mask is obtained by thresholding the smoothed slice. From its largest connected component, eight geometric descriptors (area, perimeter, eccentricity, solidity, extent, axis ratio, circularity and Euler number), the seven Hu invariant moments [32] in log scale, and 25 Zernike moment magnitudes up to order 8 [33] are computed. The Zernike moment of order n and repetition m is defined in (11):

$$Z_{nm} = \frac{n+1}{\pi} \sum_{x,y} f(x,y) R_{nm}(\rho) e^{-jm\theta} \quad (11)$$

where Rnm is the radial polynomial and (ρ, θ) are polar coordinates normalized to the unit disk. Shape features summarize global brain morphology in the selected slices, including ventricular and cortical changes that accompany neurodegeneration.

Deep Features

Deep features are learned by the two convolutional branches of the hybrid classifier described next, so that their representations are optimized jointly with the handcrafted branch.

Hybrid Improved LeNet–SqueezeNet Classifier

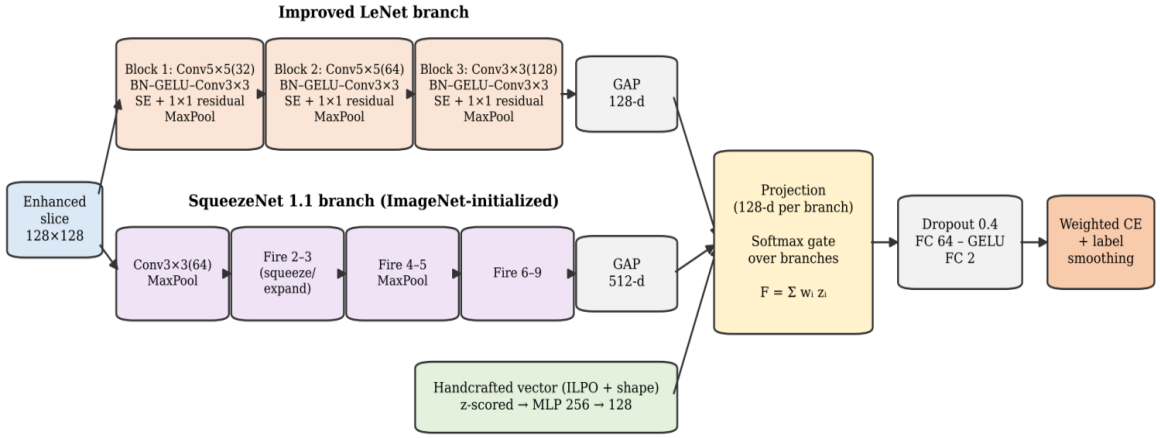


Fig. 3. Architecture of the Hybrid Improved LeNet–SqueezeNet Classifier

Fig. 3. shows the classifier. The original LeNet-5 [13] uses two small convolutional layers with average pooling and tanh activations, which is insufficient for 128 × 128 MRI slices. The Improved LeNet branch keeps LeNet's compact, three-stage structure but uses wider layers (32, 64 and 128 channels), batch normalization, GELU activations, a 1 × 1 residual shortcut in each block and squeeze-and-excitation (SE) channel attention [34]. For a block input x, the output is given in (12):

$$y = \mathrm{MaxPool}\left(\mathrm{GELU}\left(\mathrm{SE}\left(\mathrm{BN}\left(W_2 * \mathrm{GELU}\left(\mathrm{BN}(W_1 * x)\right)\right)\right) + W_s * x\right)\right) \quad (12)$$

where SE rescales each channel by a sigmoid gate computed from its global average. The SqueezeNet 1.1 branch [12] uses fire modules, in which a 1 × 1 squeeze layer reduces channels before parallel 1 × 1 and 3 × 3 expand layers, giving a strong ImageNet-initialized feature extractor with only about 0.7 million parameters. The handcrafted ILPO and shape vector is z-scored using training statistics and mapped by a two-layer perceptron. Each branch output is projected to a 128-dimensional vector zi, and the branches are fused by learned gating, as in (13):

$$w = \mathrm{softmax}\left(W_g[z_1; z_2; z_3]\right), \quad F = \sum_i w_i z_i \quad (13)$$

The fused vector passes through a dropout-regularized two-layer head. Training minimizes class-weighted cross-entropy with label smoothing 0.05, and the subject-level probability is the mean of the slice probabilities, as in (14):

$$\hat{y}_s = \frac{1}{|S_s|}\sum_{k\in S_s} p_k \qquad (14)$$

Algorithm 1: SGLN-Net PD Detection

Input: T1-weighted volumes V with labels; folds K = 5

Output: subject-level PD probabilities

Step 1: For each V: N4 correction, skull stripping, MNI registration; extract 24 axial slices

Step 2: For each slice: estimate $\hat{\sigma}$ (Rayleigh mode); apply ISGF (T = 3 iterations, adaptive ε, Rician correction)

Step 3: For each outer fold k (split by subject):

Run GWO on training slices to obtain ILPO parameters θ*

Compute ILPO(θ*) and shape features for all slices; z-score with training statistics

Train hybrid network on training subjects; early stopping on inner validation subjects

Predict test slices; average probabilities per subject

Step 4: Report subject-level metrics over folds and seeds

End Algorithm

Algorithm 1 summarizes the procedure. GWO tuning, feature standardization and early stopping use training or validation subjects only, so no information from test subjects influences any stage of the pipeline.

Results and Discussion

This section first reports quantitative validations of the ISGF enhancement stage on a phantom with known ground truth and of the ILPO texture descriptor on natural textures. It then defines the experimental protocol, hypotheses and success criteria for PD classification on PPMI data, for which experiments are in progress. All code was implemented in Python 3.11 with NumPy, SciPy, scikit-image and PyTorch 2.3.

Validation of the Enhancement Stage

Because clean reference images are not available for patient MRI, ISGF was evaluated on a 256 × 256 Shepp–Logan phantom normalized to [0, 1] and corrupted with Rician noise generated according to (1) at σ = 0.03, 0.05 and 0.08. Each condition was repeated with 10 independent noise realizations, and peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) with respect to the clean phantom were computed. ISGF used the estimated $\hat{\sigma}$; non-local means [21] used $h = 0.8\hat{\sigma}$ with 5 × 5 patches.

Table 2. Denoising Performance on the Rician-Corrupted Phantom (Mean Over 10 Realizations)

Method	σ = 0.03 PSNR / SSIM	σ = 0.05 PSNR / SSIM	σ = 0.08 PSNR / SSIM
Noisy input	28.57 / 0.362	24.14 / 0.258	20.10 / 0.190

Method	σ = 0.03 PSNR / SSIM	σ = 0.05 PSNR / SSIM	σ = 0.08 PSNR / SSIM
Gaussian (σ = 1)	24.86 / 0.481	23.12 / 0.429	20.57 / 0.370
Non-local means	30.62 / 0.523	26.17 / 0.488	21.96 / 0.455
Self-guided filter	29.64 / 0.505	25.63 / 0.445	21.58 / 0.317
ISGF (proposed)	35.64 / 0.905	33.92 / 0.884	30.47 / 0.946

Table II shows that ISGF achieved the highest PSNR and SSIM at all noise levels, improving PSNR over the noisy input by 7.07, 9.78 and 10.37 dB and over the best competing filter (non-local means) by 5.02, 7.75 and 8.51 dB. The standard deviation of ISGF over realizations was below 0.9 dB in PSNR; its SSIM varied more at σ = 0.03 and 0.05 (SD 0.079 and 0.108) because small errors in $\hat{\sigma}$ leave a faint residual background floor in some realizations.

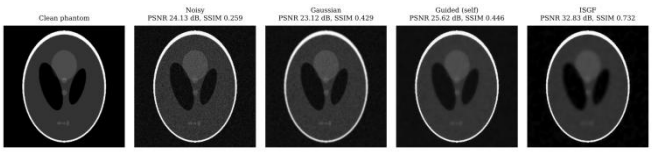


Fig. 4. Visual Comparison of Filters on the Rician-Corrupted Phantom (σ = 0.05)

Fig. 4. compares the filters on one realization. Gaussian smoothing blurs the small ellipses, and the self-guided filter preserves edges but leaves the noise floor in the background and dark regions. ISGF preserves boundaries and restores near-zero intensity in signal-free regions, which is the effect of the Rician correction.

Table 3. Ablation of ISGF Components (σ = 0.05, 10 Realizations)

Configuration	PSNR (dB)	SSIM
Self-guided filter (1 iteration, fixed ε)	25.63 ± 0.01	0.445 ± 0.001
+ Edge-aware adaptive ε	25.73 ± 0.02	0.448 ± 0.001
+ 3 iterations	25.74 ± 0.02	0.473 ± 0.000
+ Rician correction, true σ	34.26 ± 0.05	0.943 ± 0.001
+ Rician correction, estimated $\hat{\sigma}$ (ISGF)	34.02 ± 0.84	0.904 ± 0.107

Table III attributes the improvement to individual components. Adaptive regularization and iteration produced only small gains in PSNR (0.11 dB) and a moderate gain in SSIM (0.028), whereas Rician bias correction accounted for most of the improvement. Using the estimated $\hat{\sigma}$ instead of the true σ cost 0.24 dB on average, confirming that the background-mode estimator is adequate. Two caveats apply: the phantom is piecewise constant with a large signal-free background, which favours bias correction and edge-preserving filters, and Rician correction could also be added to other filters. The benefit of ISGF for PD classification will therefore be judged by the downstream ablation described below rather than by PSNR alone.

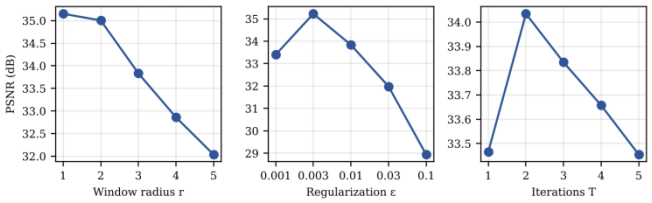


Fig. 5. Sensitivity of ISGF to Its Parameters (σ = 0.05, Mean PSNR Over 5 Realizations)

Fig. 5. reports a one-at-a-time sensitivity analysis around the default setting (r = 3, ε = 0.01, T = 3). On the phantom, smaller windows (r = 1–2, 35.0–35.2 dB) and ε = 0.003 (35.22 dB) performed better than the defaults (33.83 dB), and two iterations were slightly better than three (34.03 dB). Larger ε caused over-smoothing, reducing PSNR to 28.94 dB at ε = 0.1. The defaults were fixed a priori for MRI, whose structures are less piecewise constant than the phantom; on patient data, r, ε and T will be selected on validation subjects.

Table 4. Accuracy of the Background-Mode Noise Estimator

True σ	Estimated $\hat{\sigma}$ (mean ± SD)	Relative error
0.02	0.0207 ± 0.0010	+3.4%
0.03	0.0308 ± 0.0018	+2.7%
0.05	0.0507 ± 0.0055	+1.4%
0.08	0.0886 ± 0.0093	+10.7%
0.10	0.1110 ± 0.0085	+11.0%

Table IV evaluates the noise estimator over 10 realizations per level. The relative error was below 4% for σ ≤ 0.05, the range typical of clinical T1-weighted MRI, and rose to about 11% at higher noise, where partial overlap between background and low-intensity tissue in the histogram biases the mode upwards. An overestimated $\hat{\sigma}$ slightly over-corrects dark regions, which is consistent with the stable PSNR of ISGF at σ = 0.08 in Table II.

Validation of the ILPO Texture Descriptor

To examine whether multi-scale local phase and GWO tuning improve the robustness of texture description, a three-class texture task was constructed from the brick, grass and gravel images (512 × 512) distributed with scikit-image. Forty 48 × 48 patches per class were drawn from the left half of each image for training and forty from the right half for testing, so that training and test regions did not overlap. Features were standardized and classified by L2-regularized logistic regression trained on clean patches; test patches were evaluated clean and after Gaussian blur (σ = 1.5 pixels), Rician noise (σ = 0.05) or both, with three noise realizations. GWO used six wolves and eight iterations on half of the training patches, with the Fisher-score fitness computed either on clean patches or on randomly degraded copies.

Table 5. Texture Classification Accuracy (%) Under Degradation

Descriptor	Clean	Blur	Noise	Blur + noise	Mean (degraded)
LPQ (w = 7)	95.83	91.67	83.33	67.78	80.93
ILPO, default θ	100.00	86.67	91.39	74.44	84.17

Descriptor	Clean	Blur	Noise	Blur + noise	Mean (degraded)
ILPO, GWO (clean fitness)	100.00	66.67	63.06	43.89	57.87
ILPO, GWO (augmented fitness)	98.33	95.00	85.83	71.94	84.26

Table V shows that multi-scale ILPO with default parameters was more accurate than single-scale LPQ on clean patches (100% vs. 95.83%) and under noise (91.39% vs. 83.33%), but less accurate under blur alone.

When GWO maximized the Fisher score on clean patches, it selected a 3 × 3 window that separated the clean training textures perfectly but generalized poorly to degraded patches, reducing the mean degraded accuracy to 57.87%.

Computing the fitness on degraded training copies selected larger windows (7, 9 and 11), gave the best accuracy under blur (95.00%) and the highest mean degraded accuracy (84.26%), although the default parameters remained better under noise.

This experiment motivated the augmented fitness used in SGLN-Net, and it shows that optimization of texture parameters must target robustness rather than separability on clean data.

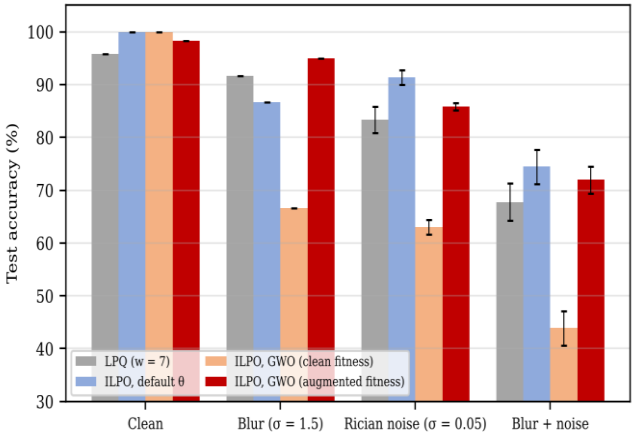


Fig. 6. Texture Classification Accuracy of LPQ and ILPO Variants Under Degradation

Fig. 6. visualizes these results; error bars denote the standard deviation over noise realizations. The comparison is limited to natural textures and a small number of patches, so it validates the behaviour of the descriptor rather than its diagnostic value for PD, which is addressed by hypothesis H2.

Classification Protocol

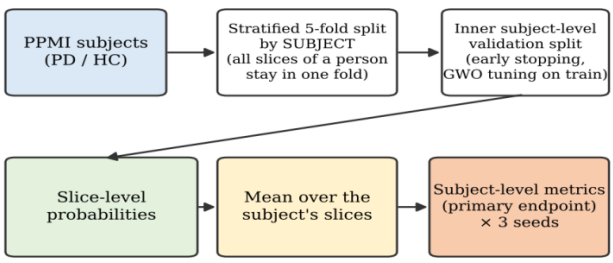


Fig. 7. Subject-Level Evaluation Protocol

The protocol is illustrated in Fig. 7. The subjects, as opposed to the slices, were split into five stratified folds; therefore, all of the slices from one individual were either in the training set or in the testing set. This prevented any "leakage" that can lead to artificially inflated performance in 2D MRI classification systems [8]. For each training partition, an inner split of the subject's data provided validation data for early stopping, and GWO parameter tuning used only slices from the training partition. The main outcome variable was the subject-level performance derived from averaging the slice-level probabilities; slice-level performance variables were presented as secondary outcomes. All of the different configurations were run using three random seed values.

Table 6. Simulation Parameters

Parameter	Value
Dataset	PPMI baseline T1-weighted MRI (PD, HC)
Preprocessing	N4, HD-BET, affine MNI registration
Slices per subject	24 axial (38–58% of S–I extent), 128 × 128
ISGF	r = 3, ε = 0.01, T = 3, Rician correction
ILPO / GWO	3 LPQ scales, 2 log-Gabor scales; 8 wolves, 15 iter.
Shape features	8 geometric + 7 Hu + 25 Zernike
Optimizer / learning rate	AdamW, 3×10^{-4}, cosine annealing
Batch size / epochs	32 / max. 60, early stopping (patience 10)
Loss	Class-weighted CE, label smoothing 0.05
Validation	Subject-level stratified 5-fold CV, 3 seeds

Table VI lists the configuration. The slice range covers the basal ganglia and midbrain, which are most relevant to PD, while keeping the number of slices per subject moderate so that the subject-level average is not dominated by uninformative slices.

Table 7. Approximate Parameter Count of the Classifier Branches

Component	Parameters
Improved LeNet branch (3 residual SE blocks)	≈ 0.34 M
SqueezeNet 1.1 feature extractor	≈ 0.72 M
Handcrafted MLP (842 → 256 → 128)	≈ 0.25 M
Projections, gate and head	≈ 0.09 M
SGLN-Net total	≈ 1.40 M
ResNet-50 backbone (reference)	≈ 23.5 M

Component	Parameters
VGG16 convolutional layers (reference)	≈ 14.7 M

Table VII, computed analytically from the layer configuration, shows that SGLN-Net is about 17 times smaller than a ResNet-50 backbone. This compactness is intended to reduce over-fitting on the few hundred subjects typically available in PPMI-based studies and to allow deployment on modest hardware.

Table 8. Measured Runtime per 128 × 128 Slice (Single CPU Core)

Operation	Time (ms)
Self-guided filter (1 iteration)	0.86
ISGF (3 iterations, $\hat{\sigma}$ estimation, Rician correction)	3.65
ILPO features (3 LPQ scales, 2 monogenic scales)	18.39
Shape features (geometric, Hu, Zernike)	8.05

Table VIII reports runtimes measured on one core of a 2.3 GHz Intel Xeon processor, averaged over 20 runs after warm-up. ISGF is about four times slower than a single self-guided filter pass but still processes a slice in under 4 ms, and the complete handcrafted feature extraction takes about 30 ms per slice, i.e. under one second for the 24 slices of a subject. GWO tuning, which evaluates the Fisher score for 8 wolves over 15 iterations, is performed once per fold on a training subset.

Evaluation Metrics

With PD as the positive class, accuracy, precision, recall, specificity and F1-score are computed from the subject-level confusion matrix, as in (15) and (16):

$$\mathrm{Acc} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}},\ \mathrm{Spec} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} \qquad (15)$$

$$\mathrm{F1} = \frac{2\cdot\mathrm{Prec}\cdot\mathrm{Rec}}{\mathrm{Prec}+\mathrm{Rec}},\ \mathrm{Prec} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}},\ \mathrm{Rec} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \qquad (16)$$

The area under the ROC curve (AUC) and the Matthews correlation coefficient are also reported. Differences between SGLN-Net and the strongest competing configuration will be assessed with DeLong's test on subject-level AUC and McNemar's test on paired predictions at $\alpha = 0.05$.

Hypotheses and Expected Outcomes

Table 9. Hypotheses, Experiments and Success Criteria

Hypothesis	Experiment	Success criterion
H1: ISGF improves classification	No filter, Gaussian, self-guided filter vs. ISGF, full model	Higher subject-level AUC with ISGF than with any other filter
H2: ILPO outperforms single-scale LPQ	Texture: none vs. LPQ (w = 7) vs. ILPO	ILPO gives the highest AUC among texture options

Hypothesis	Experiment	Success criterion
H3: Handcrafted features complement deep features	Hybrid with vs. without handcrafted branch	AUC gain with handcrafted branch
H4: The hybrid exceeds its branches and reference CNNs	LeNet, Improved LeNet, SqueezeNet, ResNet-50, VGG16 vs. hybrid	Hybrid best; significant vs. best reference (DeLong $p < 0.05$)
H5: Clinically useful discrimination	Full SGLN-Net, subject level	Subject-level AUC ≥ 0.85 and sensitivity $\geq 80\%$

Table IX states the hypotheses. The absolute target in H5 is deliberately set below the 95–97% accuracies reported in single-cohort studies [6], [7], [14], because subject-level evaluation removes identity leakage and multi-center studies indicate that rigorous T1-only evaluation yields substantially lower figures [4], [5]. The primary criterion is therefore relative: SGLN-Net should outperform its own components and re-implemented reference networks under an identical protocol.

Planned Ablation Study

Table 10. Ablation Configurations

Configuration	Purpose
Filter: none / Gaussian / self-guided / ISGF	Effect of enhancement on classification (H1)
Texture: none / LPQ / ILPO	Effect of multi-scale optimized phase texture (H2)
Shape features off / on	Contribution of morphology
Deep only vs. deep + handcrafted	Complementarity of feature modalities (H3)
LeNet-5 / Improved LeNet / SqueezeNet / hybrid	Contribution of each deep branch (H4)
ResNet-50, VGG16 (same protocol)	Reference networks (H4)

Table X lists the ablation configurations; all share the same subject folds, seeds and training budget, so differences can be attributed to the varied component. The learned branch weights w_i of (13) will additionally be reported to show how the model balances deep and handcrafted evidence.

Comparison with Existing Studies

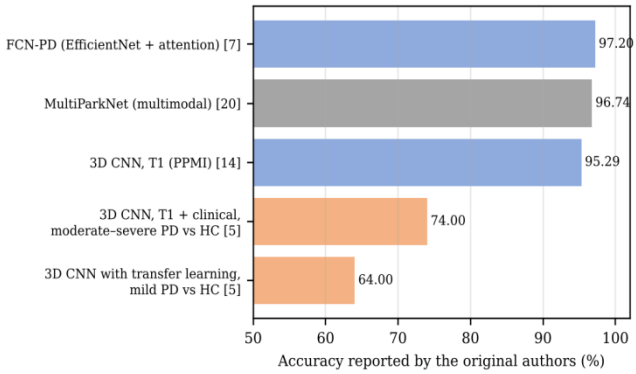


Fig. 8. Reported Accuracies of Recent MRI-Based PD Detection Studies

Fig. 8. summarizes the accuracies reported by recent studies. Single-cohort results cluster around 95–97% [7], [14], [20], whereas the multi-center study of Basaia et al. [5] reported 74% for moderate-to-severe and 64% for mild PD from imaging. The figure values are taken from the original publications and were obtained with different datasets, splits and task definitions, so they are not directly comparable with each other or with the subject-level results that SGLN-Net will report.

Implementation and Reproducibility

All stages are implemented as scripts that run end to end from a subject list. ISGF, ILPO, GWO and the shape descriptors are implemented in NumPy and SciPy, and the classifier in PyTorch. The feature standardization statistics, GWO parameters and early stopping decisions are computed from the training or validation subject of each fold, and saved along with the results, including per subject probabilities, ROC curves and learned branch weights, such that all tables and figures of the classification study can be reproduced by running the code without manual intervention. Both the phantom and texture experiments reported in sections IV-A and IV-B are completely reproducible using fixed random seeds.

Discussion

The phantom experiments demonstrate the effectiveness of ISGF as a subsequent stage for enhancing Rician corrupted images. The major benefit of ISGF over traditional filter-based methods is its explicit bias correction capability, whereas traditional filters do not include such explicit bias correction. In addition, the parameter analysis indicated that the default settings for the ISGF were not ideal for the phantom. Thus, the analysis supports the idea that enhancement parameter tuning should be done with respect to the specific type of image being processed rather than using fixed default settings. One of the main questions related to hypothesis H1 is whether the enhancement results in improved performance in PD classification. Since the phantom does not represent the complex nature of real T1-weighted images including bias fields, partial volume effects, and anatomical variation, this is the central issue of hypothesis H1. A cautious finding emerged from the texture experiment; namely, selecting descriptor parameters for class separability based upon clean data resulted in selecting delicate fine-scale descriptors, whereas selecting them based upon degraded copies resulted in robust fine-scale descriptors. This concept extends to MRI, since scanner variability acts as a degradation mechanism, and this principle is embedded within the GWO fitness of SGLN-Net. The ILPO and shape descriptors were developed so that they could contribute information that may not be learned from small datasets by compact CNNs, and the gated fusion provides a means by which their contribution can be measured via the branch weights.

The limitations that need to be addressed are several: The validation of the enhancement was performed with a synthetic phantom and texture robustness was demonstrated using natural images, not with MRI. Also, there are no results from the PPMI study at this time. PPMI is a research cohort with standardized acquisition, so external validation on independent scanners and clinical populations will be required. The 2D slice-based design is computationally efficient but discards some volumetric context, and the fixed slice range assumes accurate registration. Finally, the framework uses a single imaging modality; integration with gait, handwriting and speech is the subject of the next stage of this research programme.

Conclusion and Future Work

This study proposed SGLN-Net, a framework for PD detection from T1-weighted MRI that combines an Improved Self-Guided Filter with multimodal feature learning and a hybrid Improved LeNet–SqueezeNet classifier. ISGF extends self-guided filtering with edge-aware adaptive regularization, iterative refinement and Rician bias correction, and on a Rician-corrupted phantom it improved PSNR by 7.07–10.37 dB and SSIM by 0.54–0.76 over the noisy input, outperforming Gaussian, non-local means and standard self-guided filtering. Texture is described by ILPO, which optimizes multi-scale local phase descriptors with a Grey Wolf Optimizer, shape by geometric, Hu and Zernike descriptors, and deep features by two complementary lightweight branches fused through learned gating. A subject-level, leakage-free protocol with explicit hypotheses was defined for the PPMI classification experiments.

Future work will complete the PPMI experiments and ablations, validate the framework on independent multi-center cohorts, and extend it to 3D processing and early-stage and differential diagnosis. The image-enhancement and hybrid feature components will also be integrated with gait, handwriting and speech in the multimodal fusion stage, and with explainable AI methods in the final stage of this research programme.

References

[1] E. R. Dorsey et al., “Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016,” Lancet Neurology, vol. 17, no. 11, pp. 939–953, 2018.

[2] J. Jankovic, “Parkinson's disease: clinical features and diagnosis,” Journal of Neurology, Neurosurgery & Psychiatry, vol. 79, no. 4, pp. 368–376, 2008.

[3] C. G. Goetz et al., “Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results,” Movement Disorders, vol. 23, no. 15, pp. 2129–2170, 2008.

[4] M. Camacho, M. Wilms, P. Mouches, H. Almgren, R. Souza, R. Camicioli, Z. Ismail, O. Monchi, and N. D. Forkert, “Explainable classification of Parkinson's disease using deep learning trained on a large multi-center database of T1-weighted MRI datasets,” NeuroImage: Clinical, vol. 38, p. 103405, 2023.

[5] S. Basaia et al., “Multi-center 3D CNN for Parkinson's disease diagnosis and prognosis using clinical and T1-weighted MRI data,” NeuroImage: Clinical, vol. 48, 2025.

[6] S. S. Hussain, P. M. Shah, H. Dawood, X. Degang, A. Alshamayleh, M. A. Khan, and T. M. Ghazal, “A swin transformer and CNN fusion framework for accurate Parkinson disease classification in MRI,” Scientific Reports, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-93671-5.

[7] M. Alrawis, F. Mohammad, S. Al-Ahmadi, and J. Al-Muhtadi, “FCN-PD: an advanced deep learning framework for Parkinson's disease diagnosis using MRI data,” Diagnostics, vol. 15, no. 8, p. 992, 2025.

[8] E. Yagis, S. W. Atnafu, A. García Seco de Herrera, C. Marzi, R. Scheda, M. Giannelli, C. Tessa, L. Citi, and S. Diciotti, “Effect of data leakage in brain MRI classification using 2D convolutional neural networks,” Scientific Reports, vol. 11, p. 22544, 2021.

[9] H. Gudbjartsson and S. Patz, “The Rician distribution of noisy MRI data,” Magnetic Resonance in Medicine, vol. 34, no. 6, pp. 910–914, 1995.

[10] K. He, J. Sun, and X. Tang, “Guided image filtering,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 6, pp. 1397–1409, 2013.

[11] V. Ojansivu and J. Heikkilä, “Blur insensitive texture classification using local phase quantization,” in Image and Signal Processing (ICISP 2008), Lecture Notes in Computer Science, vol. 5099, 2008, pp. 236–243.

[12] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,” arXiv:1602.07360, 2016.

[13] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[14] S. Chakraborty, S. Aich, and H.-C. Kim, “Detection of Parkinson's disease from 3T T1 weighted MRI scans using 3D convolutional neural network,” Diagnostics, vol. 10, no. 6, p. 402, 2020.

[15] G. Solana-Lavalle and R. Rosas-Romero, “Classification of PPMI MRI scans with voxel-based morphometry and machine learning to assist in the diagnosis of Parkinson's disease,” Computer Methods and Programs in Biomedicine, vol. 198, p. 105793, 2021.

[16] Ç. B. Erdaş and E. Sümer, “A fully automated approach involving neuroimaging and deep learning for Parkinson's disease detection and severity prediction,” PeerJ Computer Science, vol. 9, p. e1485, 2023.

[17] N. Islam, M. S. A. Turza, S. I. Fahim, and R. M. Rahman, “Advanced Parkinson's disease detection: a comprehensive artificial intelligence approach utilizing clinical assessment and neuroimaging samples,” International Journal of Cognitive Computing in Engineering, vol. 5, pp. 199–220, 2024.

[18] S. Desai, H. Chhinkaniwala, S. Shah, and P. Gajjar, “Enhancing Parkinson's disease diagnosis through deep learning-based classification of 3D MRI images,” Procedia Computer Science, vol. 235, pp. 201–213, 2024.
[19] V. Dentamaro, D. Impedovo, L. Musti, G. Pirlo, and P. Taurisano, “Enhancing early Parkinson's disease detection through multimodal deep learning and explainable AI: insights from the PPMI database,” Scientific Reports, vol. 14, no. 1, p. 20941, 2024.
[20] A. Sar, P. S. Puri, H. Naz, S. Aich, T. Choudhury, and L. A. Gabralla, “Multi-modal deep learning framework for early detection of Parkinson's disease using neurological and physiological data for high-fidelity diagnosis,” Scientific Reports, vol. 15, no. 1, p. 34835, 2025.
[21] A. Buades, B. Coll, and J.-M. Morel, “A non-local algorithm for image denoising,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, 2005, pp. 60–65.
[22] P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, “An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images,” IEEE Transactions on Medical Imaging, vol. 27, no. 4, pp. 425–441, 2008.
[23] J. V. Manjón, P. Coupé, L. Martí-Bonmatí, D. L. Collins, and M. Robles, “Adaptive non-local means denoising of MR images with spatially varying noise levels,” Journal of Magnetic Resonance Imaging, vol. 31, no. 1, pp. 192–203, 2010.
[24] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971–987, 2002.
[25] K. Marek et al., “The Parkinson Progression Marker Initiative (PPMI),” Progress in Neurobiology, vol. 95, no. 4, pp. 629–635, 2011.
[26] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, “N4ITK: improved N3 bias correction,” IEEE Transactions on Medical Imaging, vol. 29, no. 6, pp. 1310–1320, 2010.
[27] F. Isensee et al., “Automated brain extraction of multisequence MRI using artificial neural networks,” Human Brain Mapping, vol. 40, no. 17, pp. 4952–4964, 2019.
[28] Z. Li, J. Zheng, Z. Zhu, W. Yao, and S. Wu, “Weighted guided image filtering,” IEEE Transactions on Image Processing, vol. 24, no. 1, pp. 120–129, 2015.
[29] M. Felsberg and G. Sommer, “The monogenic signal,” IEEE Transactions on Signal Processing, vol. 49, no. 12, pp. 3136–3144, 2001.
[30] P. Kovesi, “Image features from phase congruency,” Videre: Journal of Computer Vision Research, vol. 1, no. 3, pp. 1–26, 1999.
[31] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey wolf optimizer,” Advances in Engineering Software, vol. 69, pp. 46–61, 2014.
[32] M.-K. Hu, “Visual pattern recognition by moment invariants,” IRE Transactions on Information Theory, vol. 8, no. 2, pp. 179–187, 1962.
[33] A. Khotanzad and Y. H. Hong, “Invariant image recognition by Zernike moments,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 5, pp. 489–497, 1990.
[34] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 7132–7141.