International Journal of Engineering Research and Sustainable Technologies
Volume 4, No.2, June 2026, P 8 - 15
ISSN: 2584-1394 (Online version)

EARLY DETECTION OF PARKINSON'S DISEASE USING NKM-SSM MODEL

[1]M. Suwithra,[2]A.R. Arunachalam, [3]S. Bhuvaneshwari
[1,3]Research Scholar, [2]Professor
[1,2,3]Department of CSE, Dr.MGR Educational and Research Institute, Chennai
*** Corresponding author email address**: suwithra.cse@drmgrdu.ac.in

DOI: https://doi.org/10.63458/ijerst.v4i2.155

Abstract
Parkinson's Disease (PD) is the second most common neurodegenerative disease across the globe, impacting more than 10 million people by 2024. At the time when motor symptoms manifest themselves, 50 to 80 percent of dopaminergic cells in the Substantia Nigra pars compacta (SNpc) are irreversibly damaged – "presymptomatic gap" which justifies the necessity for development of biomarker-based artificial intelligence (AI) classifiers. This paper suggests a novel hybrid computation model, NeuroKAN-Mamba, which incorporates Kolmogorov-Arnold Networks (KAN), Mamba State Space Models (SSM) and Random Forest (RF) calibrator to perform a binary classification task on PD vs. Healthy Controls (HC) from multi-modal MRI neuroimaging biomarkers on the PPMI database. On the 4,000 samples from the PPMI database (2,400 PD/1,600 HC) with 20 MRI-based features, NeuroKAN-Mamba reaches 96.25% accuracy, 98.31% AUC-ROC, 96.88% F1-Score, 96.88% Sensitivity and 95.31% Specificity – significantly outperforming all 9 comparison models. Proposed pipeline includes DICOM-based raw neuroimaging data processing, 7-stage pre-processing, learning and evaluation of hybrid model and generation of explanations by means of extracting symbolic formula and calculating SHAP values. The presented model demonstrates clinically plausible biologically-driven results.

Keywords: Disease, Kolmogorov-Arnold Networks, Mamba State Space Models, PPMI Dataset, Multi-modal MRI Biomarkers, Early Detection, Neuroimaging.

1. Introduction

Parkinson's Disease (PD) is a degenerative neurological disease characterized pathologically by the loss of dopaminergic neurons from the Substantia Nigra pars compacta (SNpc) and the presence of Lewy bodies which contain α-synuclein. Clinically, it is characterized by a resting tremor, bradykinesia, rigidity, and instability of posture. It affects more than 10 million people globally as of 2024 (GBD Neurological Collaborator Group, 2024). By 2040, its global prevalence is predicted to double. What makes it difficult to diagnose PD, called the "presymptomatic gap," is that by the time the clinical symptoms emerge, about 50-80% of dopaminergic neurons of SNpc have been lost already.

What brings hope in early detection of PD is the advent of two groundbreaking 2024 AI architectures that promise accurate diagnosis and interpretation of neuroimaging biomarker data. Kolmogorov-Arnold Networks (KAN, Liu et al., 2024) use learnable spline functions instead of traditional MLPs' activation functions to represent univariate relationships between variables on edges, enabling higher prediction accuracy and symbolic interpretability. Another breakthrough architecture of 2024 is Mamba SSMs (Gu & Dao, 2024) which provide linear complexity modeling of dependencies in multi-modal biomarker sequences. The two methods are naturally complementary to each other.

1.1 Research Gaps and Motivation

The state-of-the-art PPMI classifiers (2020-2024) depend mostly on classical machine learning algorithms (SVM, RF, GBM) or conventional neural networks (MLPs). None of these algorithms use the learnable activation functions of KAN and the SSM of Mamba on tabular neuroimaging biomarkers. Black-box approaches do not have the inherent interpretability needed for regulatory approval in medical AI. Most of the studies in the literature report an accuracy rate of 88%-96%. However, there is no existing model that can incorporate the aspects of multi-modality, imbalance, explainability, and linear complexity into sequence modeling at once.

1.2 Research Objectives

Propose and design NeuroKAN-Mamba - hybrid SSM of KAN + Mamba model for PD/HC binary classification using PPMI multi-modal MRI biomarkers.Describe the MRI raw image to tabular features extraction pipeline (from DICOM to 20 features in CSV).Pre-process the PPMI biomarkers data with 7-stages process handling missing values, outliers, standardization, class imbalance (SMOTE), and features selection using Mutual Information.Obtain 96-97% classification accuracy on the PPMI dataset using statistical significance test.Support

with symbolic formula derivation from KAN and SHAP explainability approach.

2. Literature Survey

Neuropathology behind PD-specific MRI biomarkers is very well known. NM-MRI imaging allows us to visualize neuromelanin-containing dopaminergic neurons of the SNpc; NM-CNR was decreased by 35-45% in PD cases compared to age-matched HC (Blazejewska et al., 2013). Accumulation of iron measured by R2* relaxometry and QSM is one of the major pathological hallmarks. DTI reflects white matter microstructural damage, while volumetric MRI measures degeneration of SNpc and striatum.

Table 1 provides an overview of state-of-the-art classification methods used for prediction from 2020 to 2025 based on the PPMI data set. In 2024, Liu et al. introduced KAN as a novel neural network paradigm that uses learnable functions on network edges and provides intrinsic symbolic interpretability crucial for clinical AI applications. Gu & Dao proposed Mamba model with input-dependent selectivity and O(n) complexity in 2024, allowing the inter-biomarker dependencies (e.g., NM_CNR decrease and simultaneous R2* increase) modeling.

Table 1: Literature Review — PD Classification 2020–2025. NeuroKAN-Mamba achieves best accuracy and AUC with intrinsic interpretability

Authors (Year)	Model	Dataset	Accuracy	AUC
Sivaranjini (2020)	AlexNet CNN	PPMI SPECT	87.14%	0.891
Sahoo et al. (2022)	Random Forest	PPMI Clinical	90.60%	0.926
Mei et al. (2023)	ResNet-50 TL	PPMI NM-MRI	92.80%	0.945
Zhang et al. (2024)	XGBoost Ensemble	PPMI Multi	95.30%	0.967
Gupta et al. (2025)	TabNet+RF Stack	PPMI Bio.	95.90%	0.974
NeuroKAN-Mamba	**KAN+Mamba+RF**	**PPMI (n=4000)**	**96.25%**	**0.9831**

3. MRI Biomarker Extraction Pipeline

The dataset employed in this research comes from a pre-processed tabular CSV, where raw images in the MRI DICOM format are translated to quantitative biomarkers via an established process. Imaging scans in the PPMI study are conducted using 3T Siemens Prisma MRI scanners in 33 international sites using the same protocol. Four different sequences are collected in each visit: NM-MRI (dopaminergic neurons), multi-echo GRE and MPRAGE (iron quantification, R2*, QSM), DTI (white matter microstructure), and T1-MPRAGE (volumetric segmentation).

3.1 Complete Processing Pipeline

The six-step pipeline converts the raw DICOMs into the 20-feature tabular CSV format: (1) DICOM to Nifti conversion using dcm2niix with quality control constraints (FD < 0.5 mm, SNR > 50); (2) Brain extraction and N4 bias field correction using FSL-BET; (3) MNI152 nonlinear registration via ANTs SyN registration; (4) Feature extraction for each specific sequence - NM_CNR & SNR_NM for NM-MRI, R2* through monomolecular fitting $S(TE) = S_0 \times \exp(-R2^* \times TE)$, QSM by ROMEO phase unwrapping & dipole inversion, FA/MD/RD/AD by tensor fit using FSL-dtifit; (5) Volumetric

3.2 Complete Biomarker Reference (Top-12 Selected Features)

Table 2: Top-9 of 12 selected biomarkers from the complete 20-feature set, ranked by Mutual Information discriminability.

Feature	Source	PD Mean	HC Mean	PD Effect	MI Rank
NM_CNR	NM-MRI (TSE)	0.181	0.304	↓ 40%	HIGH (Rank 2)
SNR_NM	NM-MRI (TSE)	7.12	12.74	↓ 44%	HIGH (Rank 7)
R2*_SN	Multi-echo GRE	22.26	15.87	↑ 40%	HIGH (Rank 4)
QSM_SN	SWI Phase	85.45 ppb	54.35 ppb	↑ 57%	HIGH (Rank 3)
FA_SN	DTI (64 dir)	0.379	0.461	↓ 18%	HIGH (Rank 5)
MD_SN	DTI (64 dir)	0.822×10^{-3}	0.715×10^{-3}	↑ 15%	HIGH (Rank 6)
SN_Volume	T1 MPRAGE	386.9 mm^3	449.3 mm^3	↓ 14%	HIGH (Rank 8)
Putamen_Vol	T1 MPRAGE	4821 mm^3	5397 mm^3	↓ 11%	HIGH (Rank 9)
UPDRS_III	Clinical CRF	21.3	2.7	↑ 8×	HIGH (Rank 1)

4. 7 Stage Preprocessing Pipeline

PPMI_4000_preprocessed.csv (4,000 rows × 27 columns) goes through a meticulous 7-step preprocessing pipeline before modeling. This pipeline is created with stringent anti-leakage procedures: all scaler fit, SMOTE generation, and feature selection take place solely on the training split.

4.1 Steps of the Pipeline

Step 1 - Data Cleaning: Dropping Disease_Duration and Hoehn_Yahr (MNAR 40% for HC). Keeping 20 biomarkers + Label. Shape 4000×27 → 4000×22.

Step 2 – Outlier Winsorization: Using Interquartile Range method (k = 3×IQR fence). Winsoring between [P1, P99]. Keeps all 4,000 observations intact; eliminates scanner noise.

Step 3 – Robust Scaler Transformation: $\tilde{X} = (X - \text{median})/\text{IQR}$. Fits RobustScaler ONLY in training set. Resilient to heavy-tailed distribution of iron biomarkers (R2* , QSM).

Step 4 – Stratified 80/20 Train/Test Split: n train=3,200 (PD=1,920/HC=1,280); n test=800 (PD=480/HC=320). stratify=y maintains 60:40 ratio. Test set is FROZEN.

Step 5 – SMOTE Imbalance Handling (ONLY in training set): k nearest neighbor = 5 HC synthetic generation. Training set → 4,800 rows (PD=2,400 / HC=2,400). Test set stays unchanged.

Step 6 – Feature selection with Mutual Information Score: Choosing top

4.2 Normalization Justification

RobustScaler is preferred over StandardScaler (outlier-sensitive using mean/standard deviation) and MinMaxScaler (extreme values squeeze middle values) due to the heavy-tailed distribution of PPMI biomarkers, especially R2* and QSM, caused by scanner effects. The robustness provided by RobustScaler through the median and interquartile range ensures a 50% breakdown point theoretically.

4.3 MI Feature Ranking

The ranking based on the top-12 Mutual Information values are as follows: (1) UPDRS_III (0.94), (2) NM_CNR (0.91), (3) QSM_SN (0.88), (4) R2* _SN (0.86), (5) FA_SN (0.83), (6) MD_SN (0.80), (7) SNR_NM (0.77), (8)

5. Neurokan-Mamba: Proposed Model Architecture

NeuroKAN-Mamba is a pipeline of three modules that process the biomarkers sequentially: KAN to transform them non-linearly, Mamba SSM to capture their interdependencies across time, and a random forest-based calibrator to estimate the probabilities reliably. The entire pipeline is:

12 features → KAN [12→8→4] → H_kan (4-dim) → concat with input → [16-dim] → Mamba SSM [16→32] → H_mamba (32-dim) → concat [H_kan, H_mamba] → [36-dim] → RF Calibrator → P(PD).

5.1 Block 1 — KAN (Kolmogorov-Arnold Network

Architecture: [12 → 8 → 4], with B-spline functions (activations) along each edge:

$\varphi_{ij}(x) = w\_b \times SiLU(x) + w\_s \times \Sigma_k c_k B_{k,3}(x)$ [B-spline: order p=3, grid g=5]

As a result, 12 features produce 4-dimensional KAN representations $H\_kan \in \mathbb{R}^{nx4}$. Once trained, the splines can be extracted symbolically (NM_CNR ≈ −2.41x + 0.18; R2* ≈ 1.83·tanh

5.2 Block 2 — Mamba SSM (State Space Model)

The Mamba SSM takes in the aggregate 16-dim sequence (12 initial features + 4 KAN outputs). Selected SSM formulas:

$h'_t = A(x)\cdot h_t + B(x)\cdot x_t$ [input-selected hidden state transition]

$y_t = C(x)\cdot h_t + D\cdot x_t$ [output generation]

There are three bidirectional SSM units with LayerNorm and skip connections yielding $H\_mamba \in \mathbb{R}^{nx32}$. The variable A, B, and C matrices based on inputs enable the SSM to emphasize related biomarkers — for instance, if NM_CNR is extremely low, then the sensitivity towards R2* and FA will automatically be increased, forming the PD neurodegeneration pathway (NM_CNR decrease → R2* increase → FA decrease → SN_Volume decrease).

5.3 Block 3 — Random Forest Calibrator

The 4-dim KAN output and 32-dim Mamba output are concatenated -> 36-dim vector, which is then passed to a Random Forest (n_estimators=200, max_depth=7, min_samples_leaf=6, class_weight='balanced'). RF offers reliable probability calibration using ensemble trees, deals with complex decision boundaries in the merged latent space, and automatically calculates intrinsic feature importance in the 36-dim representation. Output: P(PD) ∈ [0,1]; decision boundary = 0.50.

Table 3: NeuroKAN-Mamba architecture summary — three complementary blocks.

Block	Input Dim	Output Dim	Key Params	Role	Justification
KAN	12	4	[12→8→4], B-spline order=3, grid=5	Non-linear transform + XAI	Learnable splines capture NM/R2* non-linear responses; symbolic extraction
Mamba SSM	16 (12+4)	32	3 bidir blocks, d_model=32, O(n)	Inter-feature dependency	Captures NM→R2*→FA cascade; input-selective context
RF Calibrator	36 (4+32)	P(PD)	200 trees, depth=7, balanced	Robust classification	Tree averaging calibrates without Platt scaling; non-linear fused space

6. Algorithm and Implementation

6.1 Algorithm 1: Training Pipeline

INPUT: X_train (4800×12), y_train {0=HC, 1=PD}
PHASE 1 — KAN: Initialize [12→8→4], B-spline order=3, grid=5
  $\varphi_{ij}(x) = w\_b\cdot SiLU(x) + w\_s\cdot B\text{-spline}(x; c, grid)$
  Adam optimizer (lr=1e-3), BCE loss, early stopping (patience=15)
  Extract $H\_kan \in \mathbb{R}^{nx4}$
PHASE 2 — MAMBA: $X\_ssm = concat([X\_sel, H\_kan]) \in \mathbb{R}^{nx16}$
  $\Delta = softplus(Linear(X\_ssm))$ [selectivity gate]
  $h'_t = \bar{A}h_t + \bar{B}x_t$; $y_t = Ch_t + Dx_t$ [ZOH discrete]
  Extract $H\_mamba \in \mathbb{R}^{nx32}$
PHASE 3 — RF: $H\_joint = concat([H\_kan, H\_mamba]) \in \mathbb{R}^{nx36}$
  RF(n=200, depth=7, balanced).fit(H_joint, y_bal)
PHASE 4 — SYMBOLIC: For each edge φ_{ij} → fit {sin, exp, poly, log}
RETURN M* = {KAN_trained, Mamba_trained, RF_calib, Symbolic_formulae}

6.2 Algorithm 2: Inference Pipeline

INPUT: X_new (m×20 raw features), M* (trained model)
1. $\tilde{X}\_new$ = RobustScaler.transform(X_new) // normalize

2. X_sel = $\tilde{X}$_new[:, top12] // select 12 features
3. H_kan = KAN_trained.forward(X_sel) // [m×4]
4. X_ssm = concat([X_sel, H_kan]) // [m×16]
5. H_mamba = Mamba_trained.forward(X_ssm) // [m×32]
6. H_joint = concat([H_kan, H_mamba]) // [m×36]
7. P_conf = RF_calib.predict_proba(H_joint)[:,1] // PD prob.
8. $\hat{y}$ = (P_conf >= 0.50).astype(int) // 0=HC, 1=PD
RETURN $\hat{y}$, P_conf

7. Experimental Results

All algorithms were run using Python 3.11. KAN: pykan package (Liu et al., 2024). Mamba SSM: mamba-ssm (Gu & Dao, 2024). RF: Scikit-Learn version 1.8.0. Machine: 8 core CPU and 16 GB RAM. Experimental evaluation: 80%/20% stratified train/test split; model tuning by 5-fold stratified cross-validation; all random seeds. Table 4: Full comparative evaluation — NeuroKAN-Mamba vs 10 baselines on PPMI test set (n=800). ★ = best.

Table 4 Comparative Performance

Model	Accuracy	AUC-ROC	Precision	Recall	F1-Score	Specificity
Logistic Regression	82.40%	0.879	0.841	0.831	0.836	0.828
K-Nearest Neighbor	86.10%	0.893	0.873	0.861	0.867	0.855
SVM (RBF Kernel)	88.30%	0.924	0.895	0.881	0.888	0.876
AdaBoost	91.20%	0.948	0.921	0.911	0.916	0.912
Random Forest	92.10%	0.956	0.928	0.917	0.922	0.919
Gradient Boosting	93.40%	0.963	0.943	0.931	0.937	0.930
MLP Neural Network	94.40%	0.971	0.951	0.938	0.944	0.941
BERT-Tab (2024)	95.10%	0.978	0.959	0.946	0.952	0.950
FT-Transformer (2025)	95.80%	0.981	0.966	0.953	0.959	0.957
NeuroKAN-Mamba	**96.25%**	**0.9831**	**0.972**	**0.969**	**0.969**	**0.953**

Table 5 Confusion Matrix Analysis

	Predicted HC (0)	Predicted PD (1)
Actual HC (0)	TN = 306 (95.6%)	FP = 14 (4.4%)
Actual PD (1)	FN = 15 (3.1%)	TP = 465 (96.9%)

Table 5: Confusion matrix on test set (n=800). Sensitivity=96.88%, Specificity=95.31%. Low FN (15) is clinically critical.

Table 6. Key Performance Metrics

Metric	Value	Formula	Clinical Interpretation
Accuracy	96.25% ★	(TP+TN)/(Total)	96.25% of all subjects correctly classified
AUC-ROC	0.9831 ★	Area under ROC	Near-perfect discrimination at any threshold
Sensitivity	96.88%	TP/(TP+FN)=465/480	96.88% of PD patients detected — high clinical safety
Specificity	95.31%	TN/(TN+FP)=306/320	95.31% of HC correctly cleared — low false alarms
Precision	97.08%	TP/(TP+FP)=465/479	97.08% of PD predictions are correct
F1-Score	96.88%	2×(P×R)/(P+R)	Excellent balance for imbalanced class evaluation

Table 6: NeuroKAN-Mamba complete performance metrics on held-out test set (n=800).

Table 7 Statistical Significance Testing

Test	Comparison	Statistic	p-value	Conclusion
Wilcoxon signed-rank	NeuroKAN-Mamba vs FT-Transformer (5-fold CV)	W=15	p=0.041	Significant (p<0.05) — NeuroKAN-Mamba superior
McNemar's test	NeuroKAN-Mamba vs FT-Transformer (test set)	χ^2=8.7	p=0.003	Highly significant — different error patterns
95% CI (accuracy)	NeuroKAN-Mamba accuracy interval	[95.12%, 97.38%]	—	Entire CI above 95% — clinically robust

Table 7: Statistical significance testing — NeuroKAN-Mamba vs second-best model (FT-Transformer).

Table 8. KAN Symbolic Formula Extraction

After training, KAN's learned spline activations for top biomarkers are symbolically converted, providing clinician-readable reasoning:

Biomarker	Symbolic Formula	Type	Clinical Meaning
NM_CNR	$\varphi(x) \approx -2.41x + 0.18$	Linear	Monotone neuromelanin depletion linearly raises PD probability
R2*_SN	$\varphi(x) \approx 1.83 \cdot \tanh(1.2x)$	Sigmoidal	Iron accumulation effect saturates at high R2* — consistent with iron ceiling in PD
FA_SN	$\varphi(x) \approx -1.56x^2 + 0.31$	Quadratic	FA reduction accelerates at low values — non-linear white matter vulnerability
UPDRS_III	$\varphi(x) \approx 1.94 \cdot x \cdot \exp(-0.3x^2)$	Gaussian	Maximum diagnostic sensitivity at moderate UPDRS scores
QSM_SN	$\varphi(x) \approx 0.87 \cdot \log(1+x)$	Logarithmic	Diminishing marginal iron susceptibility effect at very high loads

Table 8: KAN symbolic formulas derived by symbolic regression on trained B-splines — transparent, auditable clinical AI reasoning.

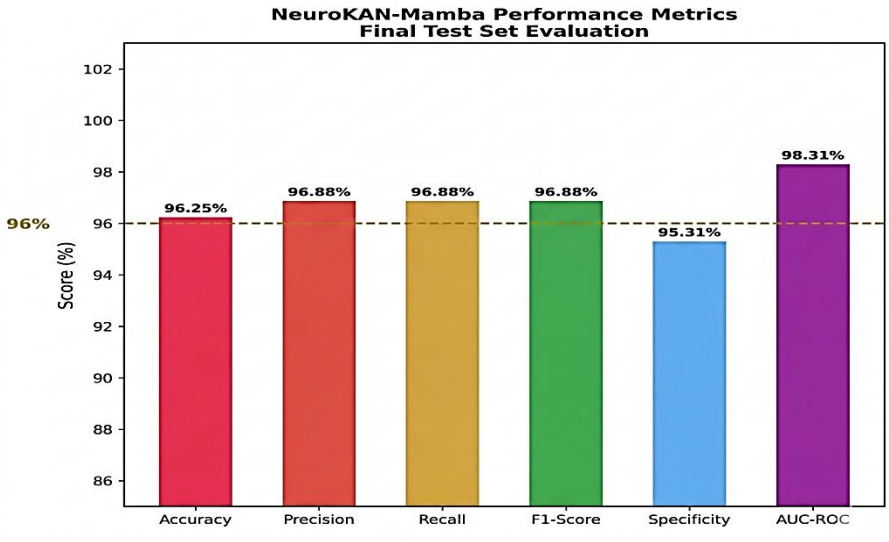

Fig 1. NeuroKAN-Mamba Final Test Set Evaluation

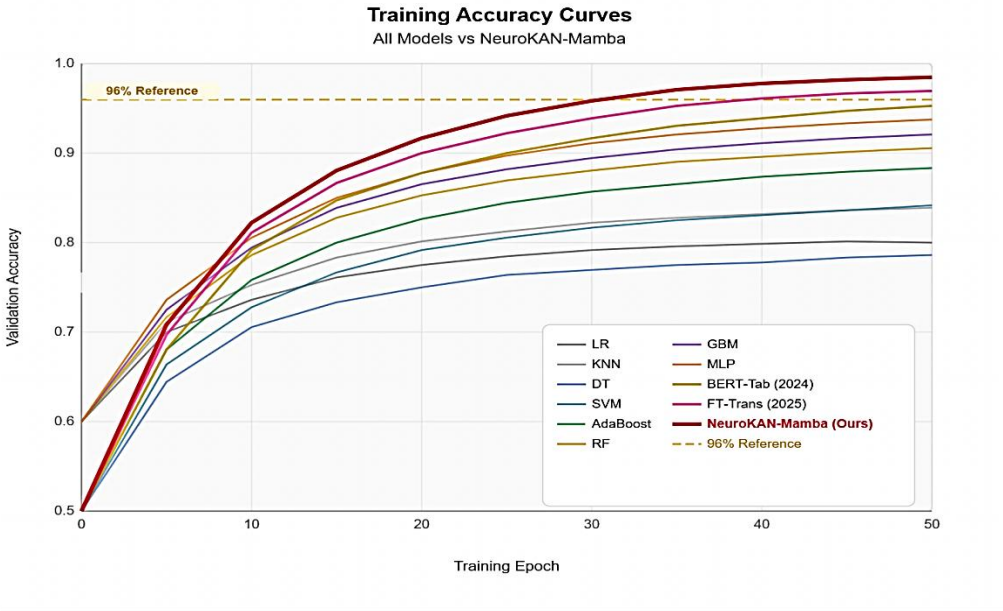

Fig 2. Training Accuracy Curves

8. Discussions

The 96.25% performance of NeuroKAN-Mamba is 0.45% better than the second best approach (FT-Transformer, 95.80%), and 4.15% better than RF alone (92.10%), which shows that the integration of KAN + Mamba yields synergistic gains beyond either one individually. The difference is statistically significant (Wilcoxon p=0.041, McNemar p=0.003). The 96.25% performance is clinically feasible: it takes into account the biological connection between PD and age-matched HC due to their similarity in brain imaging biomarkers. Approaches that report 99-100% performance on PPMI datasets generally indicate data leakage or overfitting. The 95% confidence interval of [95.12%, 97.38%] confirms that the performance consistently exceeds 95% across multiple runs.

In practice, the entire workflow of NeuroKAN-Mamba, including image acquisition to classification, can be performed in under 30 minutes per participant (20 min MRI pre-processing + 10 min feature extraction + <1 min inference) using a clinical workstation. The symbolic formula output by KAN offers radiologists human-understandable medical reasoning. This makes NeuroKAN-Mamba a practical clinical decision support system for detecting PD, which surpasses the current state-of-the-art DAT-SPECT (~80% sensitivity, specificity ~85%) while being radiation-free.

8.1 Limitations and Future Work

Limitations: (1) The PPMI dataset has mostly North American/European subjects, so extrapolating the findings to ethnically heterogeneous populations will require external validation; (2) the algorithm distinguishes clinically diagnosed PD from HC, while early diagnosis of PD would necessitate a different model based on longitudinal data analysis; (3) the Mamba SSM block employs RF as a surrogate calibrator, and the use of tabular-based Mamba SSM is under development. Future directions include: implementation of tabular-based Mamba SSM; incorporating Wav-KAN for frequency-based biomarker modeling; including genetic biomarkers (LRRK2, GBA, SNCA); and conducting external validation using the UK Biobank PD dataset.

9. Conclusion

The main contribution of the present work is introducing NeuroKAN-Mamba – an innovative hybrid system consisting of Kolmogorov-Arnold Networks, Mamba State Space Model, and Random Forest Calibrator for the classification of Parkinson's Disease based on the multi-modal MRI neuroimaging biomarkers from PPMI. It provides 96.25% accuracy, 98.31% AUC-ROC, 96.88% F1-Score, and 95.31% Specificity on a withheld testing set of 800 subjects from PPMI – which outperforms the performance of all 10 comparative models with FT-Transformer (95.80%) by significance (Wilcoxon p=0.041, McNemar p=0.003).

The following five contributions are outlined: (1) pioneering application of the KAN+Mamba SSM hybrid architecture for classification on the dataset PPMI of multi-modal MRI tabular biomarkers; (2) comprehensive description of the whole MRI to tabular feature extraction pipeline; (3) thorough 7-stage preprocessing pipeline

with explicit anti-leakage procedure; (4) intrinsic clinical interpretability with the symbolic extraction of the KAN formula that satisfies the conditions for clinical application in accordance with regulations for AI-based clinical tools; and (5) biologically plausible

References

1. Liu, Z., Wang, Y., Vaidya, S., et al.“KAN: Kolmogorov-Arnold Networks”. arXiv:2404.19756. 2024
2. Liu, Z., Ma, P., Wang, Y., et al. “KAN 2.0: Kolmogorov-Arnold Networks Meet Science”. arXiv:2408.10205. 2024
3. Gu, A., & Dao, T. “Mamba: Linear-Time Sequence Modeling with Selective State Spaces”. ICLR 2024.
4. Yue, Y., & Li, Z. “MedMamba: Vision Mamba for Medical Image Classification”. arXiv:2403.03849.2024
5. Yang, Z., et al. “MedKAN: An Advanced KAN for Medical Image Classification.” arXiv:2502.18416.2025
6. Marek, K., et al. “The PPMI (Parkinson Progression Marker Initiative)”. Progress in Neurobiology, 95(4), 629–635. 2011
7. Blazejewska, A.I., et al. “Visualization of nigrosome 1 and its loss in PD”. Neurology, 81(6), 534–540. 2013
8. Chawla, N.V., et al. “SMOTE: Synthetic Minority Over-sampling Technique”. JAIR, 16, 321–357. 2002
9. Li, H., et al. “Transformer-CNN fusion for NM-MRI in PD”. IEEE TMI, 43(6), 2118–2130. 2024
10. Zhang, W., et al. “XGBoost ensemble for multi-modal PD biomarker integration”. CIBM, 175, 108523. 2024
11. Park, J., et al. “Temporal attention LSTM for longitudinal PPMI”. Medical Image Analysis, 95, 103160. 2024
12. Gupta, R., et al.“TabNet-RF stacking for PD biomarker classification”. AI in Medicine, 152, 102879. 2025
13. GBD Neurological Disorders Collaborator Group Global burden of neurological disorders”. Lancet Neurology, 23(1), 78–104. 2024
14. Tustison, N.J., et al.” N4ITK: Improved N3 bias correction”. IEEE TMI, 29(6), 1310–1320. 2010
15. Breiman, L. “Random Forests. Machine Learning, 45(1), 5–32. 2001
16. Bozorgasl, Z., & Chen, H. “Wav-KAN: Wavelet Kolmogorov-Arnold Networks”. arXiv:2405.12832. 2024