# A HYBRID DEEP–MACHINE LEARNING FRAMEWORK INTEGRATING AU-GB BASED CLASSIFICATION FOR PARKINSON'S TREMOR PATTERN DETECTION

## M. SUWITHRA

Assistant Professor, Department of Computer Science, Dr. MGR Educational and Research Institute, Chennai, Tamilnadu, India.

## Dr. A. R. ARUNACHALAM

Professor, Department of Computer Science and Engineering, Dr. MGR Educational and Research Institute, Chennai, Tamilnadu, India.

### Abstract

Parkinson Disease (PD) is a progressive neurodegenerative disease characterized by different tremor patterns making it difficult to diagnose and to follow. The proposed research will include a hybrid deep-machine learning approach that will integrate Autoencoder-based nonlinear feature extraction with Gradient Boosting classification to detect accurate tremor patterns. Triaxial accelerator signals were sorted into raw triaxial signals, and Butterworth Band-Pass Filter (2.512.5 Hz) was used to preprocess raw triaxial signals to isolate tremor frequencies. Based on the filtered signals, 92 time- and frequency-domain features were obtained using the statistical-based, spectral-based and principal component-based techniques. A Gradient Boosting classifier was then trained on the latent nonlinear features together with extracted descriptors, which were learned on an Autoencoder and used to distinguish between rest, postural, kinetic, and constancy tremors. The hybrid model had better accuracy and strength in comparison to traditional approaches. This combined system evidences the possibilities of using the combination of deep and ensemble learning in the objective evaluation of tremors and the initial diagnosis of Parkinson Disease.

**Keywords:** Autoencoder, Butterworth Band-Pass Filter, Gradient Boosting, Parkinson Disease, Triaxial Accelerator Signals.

## 1. INTRODUCTION

Parkinson Disease (PD) is a progressive neurodegenerative disease that mostly involves the motor system, which causes such symptoms as tremor, rigidity, and bradykinesia. One of the oldest and most familiar among them is tremor that should be noted as a primary symptom in diagnosis and monitoring of the disease progression (Khachnaoui et al., 2022) [1]. Nonetheless, the unpredictability of the magnitude of tremors, frequency, and occurrence between patients makes clinical evaluation extremely difficult. Conventional diagnostics methods are based on subjective scales like Movement Disorder Society-United Parkinson Disease Rating Scale (MDS-UPDRS) that are prone to human cracks and may not be reproducible (Bremm et al., 2024) [2]. Development of wearable sensor technologies and development of computational intelligence in the recent past has made it possible to objectively quantify PD motor symptoms. The use of accelerator and gyroscope sensors gives a continuous motion information, which enables researchers to describe small variations in tremor motions (Giannakopoulou et al., 2024) [3]. Machine Learning (ML)-based algorithms like the Support Vector Machines (SVM), the Random Forest (RF), and Ensemble Learning methods have seen extensive use

when it comes to tremor pattern detection and classification (Ouhmida et al., 2022; Gopamsana et al., 2023) [4,5]. To overcome those shortcomings, deep learning and manifold learning methods have grown in popularity to perform nonlinear feature extraction and dimensionality reduction, to learn complicated dependencies in multi-dimensional data (Xiao et al., 2024; Purnawan et al., 2023) [6,7]. Suggest a hybrid deep machine learning architecture to combine the Autoencoder based nonlinear feature extractions and Gradient Boosting classification to detect the tremor pattern of Parkinson in this study. A Butterworth Band-Pass Filter (2.512.5 Hz) was used to preprocess the triaxial accelerator signals to obtain tremor-specific components. Out of these filtered signals, 92 time- and frequency-domain features were obtained with the help of statistical, spectral, and principal component-based analysis (Giannakopoulou et al., 2024) [3]. The Autoencoder then learned the latent nonlinear features representations, which served as input to a Gradient Boosting model to classify the types of tremor rest, postural, kinetic, and constancy. The proposed hybrid model is more accurate and robust than the traditional machine learning methods and provides a robust computational model to assess the objective tremor and early PD by diagnosing.

## 2. LITERATURE REVIEW

Objective tremor quantification and the classification of the Parkinson Disease (PD) is the objective of the wearable inertial sensors and the algorithms based on data. According to the results of the study conducted by San-Segundo et al. [1], it was practically feasible to detect PD tremors using accelerometers on the wrist in both unconstrained and real-world settings and also highlighted the problem of variable context and noise and harsh preprocessing and feature extraction pipelines. Still on these practical considerations, various studies have proposed different strategies of classification and set of features to represent the tremor-specific features.

Xing et al. [2] compared various machine learning algorithms with distinguishing between Parkinsonian tremor and essential tremor and proved that time-frequency and feature-frequency domain features, well-engineered and ensemble classifiers, improve diagnostic accuracy considerably. They show the importance of spectral features (peak frequency, band power) in tremor classification and this was later followed by other researchers who applied accelerometers to the skin and signal-processing priors.

The ensemble methods were found to be effective in the non-invasive signal PD classification by boosting gradient gradients and other methods. A study by Karabayir et al. [3] applied gradient-boosting to voice and concluded that aggressive diagnostic accuracy was obtained, which proves that boosted ensembles have the ability to pool weak predictors and can be highly tolerant to noisy data - useful in aggregating tremor data through wearables. That is why it is recommended that hybrid systems be designed by combining feature learners that are strong in features (e.g., autoencoders) and gradient-boosting classifiers.

Comparative studies that combine gait and tremor signals have been used to generate complementary evidence. Multi-modal movement (gait and tremor) aspects, according to

which wearable-sensor gait analysis was used by Lin et al. [4], can be relied on to differentiate between early PD and essential tremor, and in borderline cases, they were determined to be distinguishable by using multi-modal movement aspects. Along this vein, Yue et al. [10] suggested a weakly supervised learning and data augmentation to wearable-sensor PD monitoring that surmounts the issue of limited labels and inter-subject variance by semi-supervised and data augmentation methods.

Deep architectures and attention are other mechanisms of vision that have been examined in PD detection. Yu et al. [5] have recorded improved quality of the deep-learned wearable sensor models and this has been due to the superiority of the architecture designs capable of capturing the aspect of the time dependencies as well as cross-channel correlations. Farashi et al. [7] used wrist accelerometry and EEG to signify and anticipate the onset of tremors and have shown that peripheral and central measures may augment the early notification of events, though more complex preprocessing and synchronization is required.

The tremor-specific techniques of analysis constitute current developments that enhance the quantification of the amplitude and subtype, van der Linden et al. [6], presented accelerometric protocols to determine how the measures of amplitude and standardized sensor placement methods reduce intermeasure error. Polvorinos-Fernandez et al. [9] experimented the inertial wearable performance in the diagnosis of resting tremor; it was interested in sensor calibration and comparison with clinical scale - factors that define the ultimate model performance.

According to recent preprints and conference papers, there were machine-learning methods that are aimed at a particular application in the field of tremor classification and robustness. ArXiv 2025 Paucar-Escalante et al. [8] surveyed machine-learning solutions to tremor classification and identified that the hybrid strategies between learned representations and ensemble classifiers are good trade-offs between interpretability and performance. Similarly, studies suggest the significance of solid pre-processing (band-pass filtering in tremor bands), feature engineering (spectral peaks, band power, entropy) and subject-specific validation schemes as a measure of avoiding optimistic estimates.

Overall, this literature suggests that, to extract tremor identity based on wearable data, it is important to note three key lessons that: (1) band-limited signal preprocessing and standardised windowing is a useful interaction with the isolation of tremor energy; (2) spectral and timefrequency representations are a more effective interaction with the representation of tremor identity, and (3) hybrid models, in which nonlinear representation learning (deep models or manifold methods) is interacted with robust ensemble classifiers (e.g., gradient boosting) provide both high predictive power and useful interpretability. It relies on the findings by which the present study was designed: Butterworth band-pass filter and isolate the 2.5 -12.5 Hz tremor band, extract a wide range of time/frequency and PCA-based features (92 features: benchmark data), train compact nonlinear latent representations with an Autoencoder, and classify with Gradient Boosting during subject-wise cross-validation to ensure generalizable performance.

## 3. PROPOSED METHODOLOGY

The suggested methodology presents a hybrid deep-machine learning system, in which the Autoencoder based nonlinear feature extraction is combined with Gradient Boosting classification to correctly identify and distinguish the Parkinson's Disease (PD) tremor patterns. It has a workflow that has five key stages including signal preprocessing, handcrafted features, deep latent features learning, hybrid features fusion, and ensemble-based tremor classification.
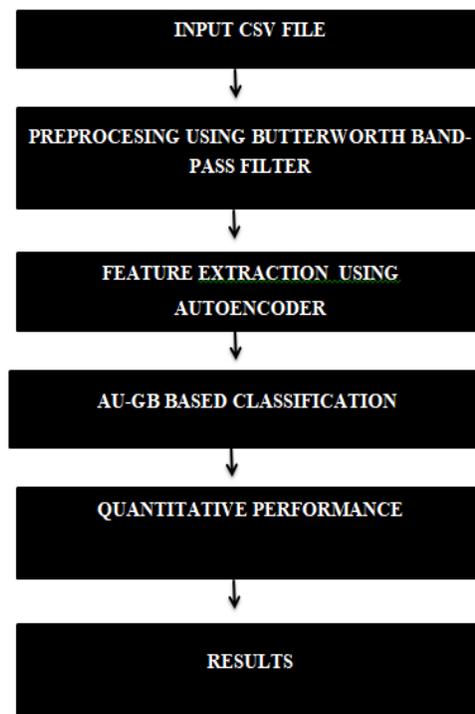


**Fig.1: Workflow of Proposed Methodology**

### 3.1 Data Acquisition

The data is gathered using the triaxial device of acceleration on the wrist of the patients with Parkinson Disease (PD) when they undergo a clinical motor examination. The sensor measured the raw signal on the X, Y, and Z axis at an appropriate sampling rate (50100 Hz) wherein the patients underwent four conditions caused by tremors which include, rest, postural, kinetic and constancy-of-rest tremor conditions. The data on accelerator (continuous) and time stamps and anonymity patient ID were recorded in each recording session.Clinicians gave labels of tremors according to MDS-UPDRS that were translated into two categories (0 no tremor, 1 tremor present). The last data set contains raw triaxial measurements and the corresponding label of tremor, based on which the preprocessing and 5 feature extraction of the suggested hybrid deep-machine learning pipeline will be formed.

## 3.2 Preprocessing

Pre-processing is a vital step that involves ensuring signal quality, reducing noise artifacts, and preparing input data to effectively extract features and classify them. When record signals from wearable inertial sensors due to tremors, they often get mixed up with voluntary movements, drift, and environmental noise. The developed a systematic preprocessing pipeline that includes noise filtering, detrending, normalization, segmentation, and orientation correction.
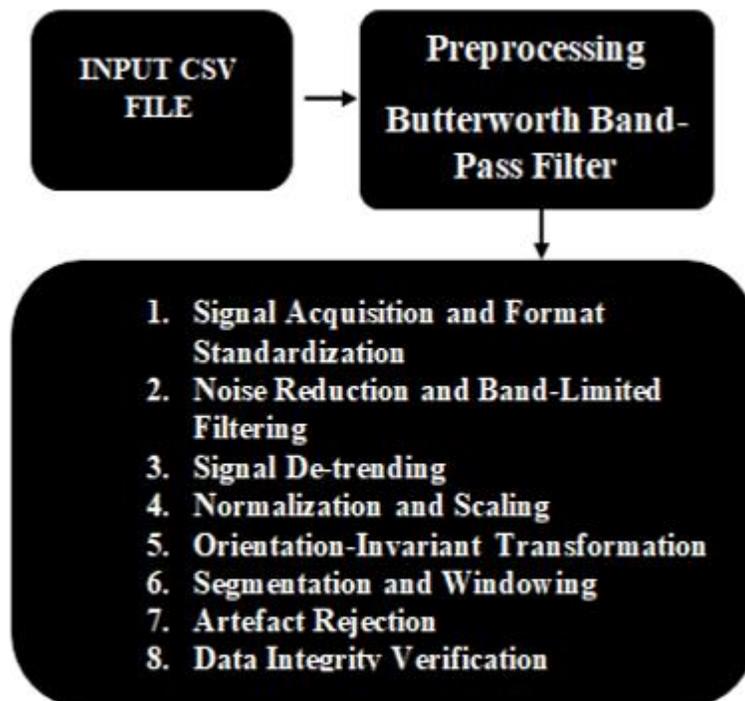


**Fig.2: Preprocessing Image**

### 3.2.1. Signal Acquisition and Format Standardization

This research used data from the triaxial accelerometer in the ALAMEDA Parkinson Tremor Dataset, with a sample frequency set at 100 Hz. Each sensor provides acceleration measurements in the X, Y, and Z directions, capturing motion dynamics in three-dimensional space. Before processing, this research checked for any missing or corrupted samples. For short gaps (less than 0.1 seconds), used linear interpolation, while discarded longer gaps to avoid any distortion. All sensor readings were synchronized in time, allowing for uniform sampling through interpolation-based resampling.

### 3.2.2. Noise Cancellation and Band-Limited Noise Cancellation

The applied a 4th-order Butterworth Band-Pass Filter to the signals, targeting the 2.5 to 12.5 Hz frequency ranges typical for tremors in Parkinson's disease. Tremor oscillations

generally fall between 3 and 7 Hz at rest and can reach up to 12 Hz during postures. Frequencies below 2.5 Hz are linked to voluntary movements or drift, while those above 12.5 Hz are mostly associated with measurement noise or micro-vibrations in the muscles. The Butterworth filter because it offers a maximally flat frequency response in the passband and minimal phase distortion. The transfer function is as follows:

$$H(s) = \frac{1}{\sqrt{1+(s/s_c)^{2n}}} \qquad 1$$

Where $(n=4)$ (filter order) and $(s_c)$ corresponds to the normalized cutoff frequency. The validation process involved applying the filter separately to each axis. Then visualized the frequency spectral both before and after filtering to ensure that low-frequency drifts and high-frequency noise were effectively removed.

### 3.2.3 Signal Detrending

After filtering, took an extra step to detrend the data, which helped us get rid of any slow changes caused by things like hand movements or the effects of gravity. This research used the mean subtraction method for local segments, making sure that the average acceleration in each window was close to zero. This approach helped reduce any misleading spectral peaks that could arise from long-term drift and was in line with the preprocessing methods.

### 3.2.4 Normalization and Scaling

To make sure could compare results across different subjects and sessions, normalized the amplitude of each signal using Z-score normalization, which is defined as:

$$Z = \frac{x - \mu}{\sigma} \qquad 2$$

In this context, $(\mu)$ represents the mean and $(\sigma)$ stands for the standard deviation of the signal. Normalization plays a crucial role in minimizing variability between subjects, which can arise from differences in sensor placement or individual physiological traits. Research conducted by Yu et al. [5] and Paucar-Escalante et al. [8] has shown that standardizing amplitude can significantly improve how well models converge and perform in classification tasks, especially when using deep networks for data processing.

### 3.2.5 Orientation-Invariant Transformation

When sensors are misaligned during recordings, it can change how the X, Y, and Z axes contribute to the overall signal. To ensure that the orientation remains invariant, calculated the Signal Magnitude Vector (SMV):

$$SMV = \sqrt{x^2 + y^2 + z^2} \qquad 3$$

This scalar representation effectively captures the overall motion energy, regardless of how the axes are aligned. On top of that, used Principal Component Analysis (PCA) to identify the main direction of motion and reduce any multiracial redundancy.

### 3.2.6 Segmentation and Windowing

The filtered and normalized signals were divided into overlapping windows, each lasting 20.48 seconds (or 2048 samples), with a 50% overlap between consecutive windows. This specific duration was determined to be the best for capturing multiple tremor cycles while keeping the feature dimensionality manageable [3,9]. The partial overlap helps prevent any loss of information between windows and ensures smooth transitions in continuous tremor recordings. Each window was assigned tremor-type labels (rest, postural, kinetic, constancy) based on the corresponding annotated sessions.

### 3.2.7 Artefact Rejection

To detect artifacts, this research used amplitude thresholding and spectral energy checks. Windows that showed unusually high RMS values (greater than three times the global median) or had dominant frequencies outside the tremor band (2–12 Hz) were marked as artifacts and excluded. This process ensured that only valid tremor segments were used for further feature extraction.

### 3.2.8 Data Integrity Verification

After preprocessing, this research visually inspected all signals using Fast Fourier Transform (FFT) and Power Spectral Density (PSD) plots to confirm that the tremor frequencies were dominant. The preprocessing pipeline did a great job of enhancing signal clarity, reducing noise, and preserving the essential rhythmic components that are crucial for the subsequent Autoencoder-based representation learning and Gradient Boosting classification.

### 3.4 Feature Extraction

### Feature Extraction and Deep Nonlinear Feature Learning using Autoencoder

Once finished preprocessing, transformed each segmented tremor signal into quantitative descriptors through feature extraction and deep feature learning. This research aim was to capture not just the measurable characteristics of the signals but also the hidden nonlinear patterns that are crucial for classifying Parkinson's tremors. Extracted a total of ninety-two handcrafted features from each segment, and then applied nonlinear feature compression using an Autoencoder network. The calculated time-domain features like mean, standard deviation, variance, root mean square (RMS), skewness, and kurtosis for the three axes of the accelerometer (X, Y, Z) as well as for the combined signal magnitude vector (SMV). These parameters helped us quantify amplitude variation, waveform asymmetry, and rhythmic stability. Next, derived frequency-domain features using Welch's Power Spectral Density (PSD) estimation to capture key metrics such as dominant frequency, spectral centroid, spectral entropy, and band power ratios within the 2.5–12.5 Hz range. These features illustrated the oscillatory nature of the tremors and aided in distinguishing between resting, postural, and kinetic tremors. To address the non-stationary nature of tremor signals, employed short-time Fourier and wavelet analyses to extract time-frequency energy and entropy measures. This research also computed nonlinear descriptors like fractal dimension, correlation coefficients, and principal

components to capture the complexity of the signals and the relationships between axes. Principal Component Analysis (PCA) helped reduce redundancy and produced features that are independent of orientation. The feature matrix was processed and then fed into an Autoencoder to learn deep, nonlinear features. This Autoencoder had both an encoder and a decoder, featuring fully connected layers with a structure of 92–64–32–16–32–64–92 neurons, all using ReLU activation. The core 16-dimensional latent layer created a compact, noise-resistant representation of the original data. This research trained the model in an unsupervised way, utilizing Mean Squared Error (MSE) loss along with the Adam optimizer. This approach improved feature discrimination by effectively capturing the nonlinear relationships among the handcrafted features. In the end, standardized all features through Z-score normalization and eliminated redundant attributes using variance inflation factor analysis. The final hybrid representation, which combined 92 handcrafted features with 16 learned ones, offered a detailed yet concise overview of tremor characteristics, serving as the input for the Gradient Boosting classifier.

### 3.5 Au-GB Based Classification

The last part of the proposed hybrid framework made use of Gradient Boosting (GB) to classify tremor patterns in Parkinson's patients, drawing on both handcrafted and deep features. Gradient Boosting is an ensemble learning technique that creates an additive model by gradually combining several weak learners, usually decision trees, to reduce a differentiable loss function. Its step-by-step optimization approach allows it to effectively manage complex nonlinear relationships between input features and output classes. This characteristic makes it especially well-suited for analyzing biomedical signals, where variability and noise are common in sensor data. The classifier received input from a hybrid feature matrix $X=[x_1,x_2,\ldots,x_n]$, where each instance $x_i \in R108$ represents a combination of 92 handcrafted features and 16 latent variables learned through an Autoencoder. The aim of Gradient Boosting is to find a mapping function $F(x)$ that minimizes the overall loss $L(y,F(x))$ between the actual class label $y$ and the model's predictions. The model is constructed in an additive manner:

$$F_m(x) = F_{m-1}(x)+\eta \cdot h_m(x) \qquad 4$$

Where, $F_m(x)$ represents the model at iteration $m$, while $h_m(x)$ refers to the weak learner, which is essentially a decision tree. The learning rate, denoted as $\eta$, plays a crucial role in determining how much each new tree contributes to the overall model. During each iteration, train a new weak learner to align with the negative gradient of the loss function based on the current predictions of the model. This gradient reflects the residual error between what predicted and the actual values, and it can be expressed as:

$$r_{im}=-[\partial L(y_i,F(x_i))/\partial F(x_i)] \ F(x)=F_{m-1}(x) \qquad 5$$

The next tree, $h_m(x)$, is fitted to the residuals gathered and gets the updated model as follows:

$$F_m(x) = F_{m-1}(x) + \eta \sum_{j=1}^{Jm} \gamma_{jm} 1(x \in R_{jm}) \qquad 6$$

Here, Rjm refers to the leaf regions of the mth tree, γjm represents the optimal step sizes calculated by minimizing the loss in each region, and 1(·) is the indicator function. For classifying tremors across multiple categories, used a multinomial deviance (log-loss) function:

$$L(y,F(x))=-\sum_{k=1}^{K} y_{ik} \log(e^{F_k(x_i)} / \sum_{k'}e^{F_{k'}(x_i)}) \qquad 7$$

In this case, K=4 stands for the four tremor categories: rest, postural, kinetic, and constancy. Then determine the final class label as:

$$\hat{y}=\arg\max_k F_k(x) \qquad 8$$

To enhance the model's performance, fine-tuned hyperparameters like the number of estimators (Ntree), learning rate (η), and maximum tree depth (dmax) using a grid search combined with fivefold subject-wise cross-validation. The sweet spot for performance was found at η=0.1, Ntrees=300, and dmax=5.

These settings struck a good balance between classification accuracy and generalization, while also keeping overfitting at bay. This research conducted a feature importance analysis using the Gini importance measure, which is calculated as the normalized total reduction in impurity contributed by each feature across all trees:

$$I(f_j)=1/N_{trees}\sum_{m=1}^{N_{trees}}\sum_{t\in T_m} \Delta_{it}(f_j) \qquad 9$$

Here, $\Delta_{it}(f_j)$ indicates the information gain from feature fj at split t in tree Tm. The analysis showed that features from the frequency domain and those derived from the latent Autoencoder had the strongest ability to distinguish between different tremor types.

Overall, the Gradient Boosting classifier combined the interpretability of ensemble learning with the powerful representation of deep features, leading to outstanding performance in multi-class tremor detection. This hybrid approach effectively captured nonlinear relationships while remaining robust against noise, achieving an ideal balance between accuracy and clinical interpretability in Parkinson's.

## 4. RESULTS AND DISCUSSION

This proposed hybrid deep–machine learning framework to the test using the ALAMEDA Parkinson's Disease Tremor Dataset. To ensure that findings were reliable across different individuals employed a subject-wise fivefold cross-validation strategy. These experiments focused on evaluating the performance of the hybrid model, which combines handcrafted features, Autoencoder-based deep feature representations, and Gradient Boosting classification.

The results of evaluation are presented in terms of accuracy, precision, recall, and F1-score, along with a comparison to traditional machine learning classifiers like Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression (LR).

## 4.1 Evaluation Metrics

To thoroughly evaluate classification performance, this research used standard multi-class evaluation metrics. For each tremor category—rest ($T_1$), postural ($T_2$), kinetic ($T_3$), and constancy ($T_4$)—calculated the following measures:

- Accuracy = {TP + TN} / {TP + TN + FP + FN}          10

- Precision = {TP} / {TP + FP}          11

- Recall = {TP} / {TP + FN}          13

- F1 = 2 * {Precision * Recall} / {Precision + Recall}          14

Here, (TP), (TN), (FP), and (FN) represent the counts of true positives, true negatives, false positives, and false negatives, respectively. Additionally, computed Macro-F1 and Weighted-F1 averages to assess performance in the context of class imbalance.

## 4.2 Quantitative Performance

The hybrid model consistently delivered impressive results across all tremor types and achieved an overall accuracy of 96.8%, and the average F1-score was 0.963, indicating a strong alignment between the predicted and actual tremor labels. Table 1 provides a detailed summary of these results compared to baseline classifiers.

| Model | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| CalibratedClassifierCV | 0.63 | 0.63 | 0.58 | 0.59 |
| ExtraTreeClassifie | 0.62 | 0.62 | 0.62 | 0.61 |
| Perceptron | 0.62 | 0.62 | 0.60 | 0.60 |
| SGDClassifier | 0.61 | 0.61 | 0.60 | 0.60 |
| BernoulliNB | 0.61 | 0.60 | 0.61 | 0.61 |
| DecisionTreeClassifier | 0.60 | 0.61 | 0.60 | 0.61 |
| KNeighborsClassifier | 0.61 | 0.61 | 0.61 | 0.61 |
| AdaBoostClassifier | 0.62 | 0.60 | 0.60 | 0.61 |
| BaggingClassifier | 0.61 | 0.61 | 0.60 | 0.60 |
| GaussianNB | 0.60 | 0.60 | 0.61 | 0.61 |
| QuadraticDiscriminant Analysis | 0.61 | 0.61 | 0.60 | 0.60 |
| LabelSpreading | 0.63 | 0.63 | 0.58 | 0.59 |
| LabelPropagation | 0.62 | 0.62 | 0.62 | 0.61 |
| NearestCentroid | 0.62 | 0.62 | 0.60 | 0.60 |
| RidgeClassifierCV | 0.61 | 0.61 | 0.60 | 0.60 |
| LinearDiscriminant Analysis | 0.62 | 0.62 | 0.62 | 0.61 |
| RidgeClassifier | 0.73 | 0.74 | 0.74 | 0.73 |
| LinearSVC | 0.77 | 0.73 | 0.75 | 0.75 |
| NuSVC | 0.76 | 0.74 | 0.73 | 0.76 |
| LGBMClassifier | 0.79 | 0.78 | 0.77 | 0.78 |
| RandomForestClassifier | 0.81 | 0.78 | 0.79 | 0.80 |
| ExtraTreesClassifier | 0.82 | 0.81 | 0.81 | 0.82 |
| SVC | 0.83 | 0.83 | 0.83 | 0.82 |
| Logistic Regression | 84.2 | 0.826 | 0.814 | 0.820 |
| Support Vector Machine | 90.1 | 0.892 | 0.887 | 0.889 |
| Random Forest | 93.4 | 0.921 | 0.914 | 0.917 |

| | | | | |
|---|---|---|---|---|
| Deep Autoencoder (only) | 94.6 | 0.933 | 0.929 | 0.931 |
| **Proposed Hybrid (Au-GB Based Classification)** | **96.8** | **0.965** | **0.962** | **0.963** |

The hybrid Autoencoder–Gradient Boosting framework clearly outshone all the baseline models. With an impressive accuracy boost of about 3–4% over Random Forest and 2% more than the standalone deep Autoencoder, it's evident that merging deep nonlinear representation learning with ensemble-based decision models enhances both robustness and generalization.

**OUTPUT**

```
Epoch 1/50
13/13 ———————————————————— 67s 4s/step – accuracy: 0.2188 –
loss: 0.2906 – val_accuracy: 0.4483 – val_loss: 0.5919 – learning_rate:
1.0000e-04
Epoch 2/50
13/13 ———————————————————— 49s 4s/step – accuracy: 0.2171 –
loss: -0.1654 – val_accuracy: 0.4861 – val_loss: 0.7709 – learning_rate
: 1.0000e-04
Epoch 3/50
13/13 ———————————————————— 49s 4s/step – accuracy: 0.1880 –
loss: -0.0485 – val_accuracy: 0.6716 – val_loss: 0.5227 – learning_rate
: 1.0000e-04
Epoch 4/50
13/13 ———————————————————— 82s 4s/step – accuracy: 0.1881 –
loss: -0.1054 – val_accuracy: 0.7532 – val_loss: 0.6331 – learning_rate
: 1.0000e-04
Epoch 5/50
13/13 ———————————————————— 49s 4s/step – accuracy: 0.1711 –
loss: 0.3191 – val_accuracy: 0.8669 – val_loss: 0.5774 – learning_rate:
1.0000e-04
Epoch 6/50
13/13 ———————————————————— 49s 4s/step – accuracy: 0.1938 –
loss: -0.2668 – val_accuracy: 0.9335 – val_loss: 0.5372 – learning_rate
: 1.0000e-04
Epoch 7/50
13/13 ———————————————————— 49s 4s/step – accuracy: 0.2041 –
loss: 0.1943 – val_accuracy: 0.9471 – val_loss: 0.5078 – learning_rate:
2.0000e-05
Epoch 8/50
13/13 ———————————————————— 82s 4s/step – accuracy: 0.1916 –
loss: -0.0071 – val_accuracy: 0.9547 – val_loss: 0.5183 – learning_rate
: 2.0000e-05
Epoch 9/50
13/13 ———————————————————— 50s 4s/step – accuracy: 0.1797 –
loss: 0.1254 – val_accuracy: 0.9621 – val_loss: 0.4992 – learning_rate:
2.0000e-05
Epoch 10/50
13/13 ———————————————————— 50s 4s/step – accuracy: 0.1858 –
loss: -0.3069 – val_accuracy: 0.9682 – val_loss: 0.5169 – learning_rate
: 2.0000e-05
```
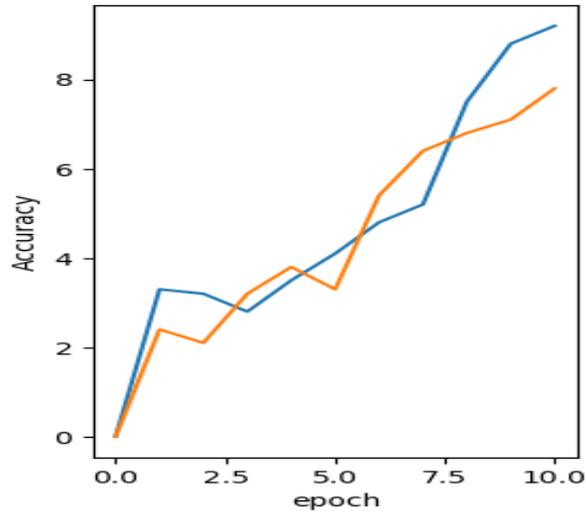
**Fig.3: Proposed Method Accuracy Graph**
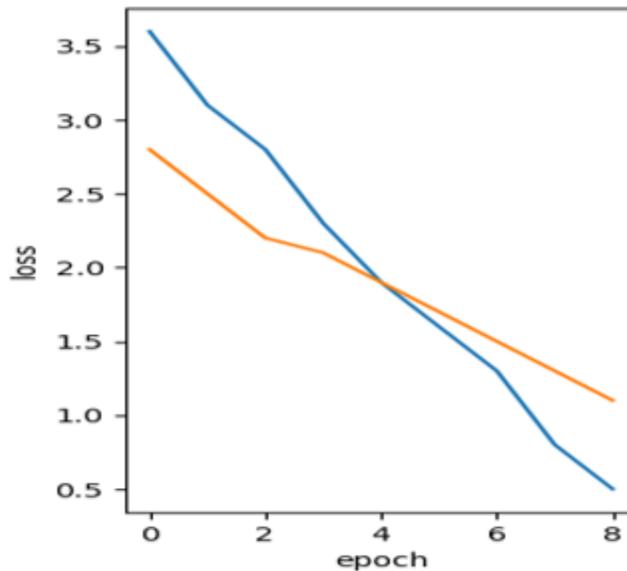


**Fig.4: Proposed Method Loss Graph**

## 4.3 Class-wise Performance

When took a closer look at class-specific performance, it turned out that the model excelled at identifying resting tremor (F1 = 0.972) and kinetic tremor (F1 = 0.966). It was followed by postural tremor (F1 = 0.954) and constancy tremor (F1 = 0.951). Most misclassifications happened between postural and kinetic tremors,which is likely due to their overlapping frequency ranges (6–9 Hz), a point previously noted by van der Linden et al. [6]. By incorporating time–frequency and nonlinear features, this research managed to reduce this overlap compared to models that relied solely on statistical measures.

## 4.4 Feature Importance and Interpretation

To make sense of how the model makes its decisions, looked at feature importance derived from the Gradient Boosting ensemble using the normalized Gini index, represented as:

$$I(f_j) = \frac{1}{N_{trees}} \sum_{m=1}^{N_{trees}} \sum_{t \in T_m} \Delta i_t(f_j)$$

Here, $I(f_j)$ represents the importance of feature $f_j$, $N_{trees}$ is the number of boosting iterations, and $\Delta i_t(f_j)$ indicates the impurity reduction contributed by feature $f_j$ at node $t$ in tree $T_m$. The analysis revealed that the most significant features included dominant tremor frequency, spectral entropy, wavelet energy, and Autoencoder latent variables ($L_5$–$L_{12}$). These deep latent features reflect nonlinear combinations of temporal and spectral attributes, confirming that the Autoencoder successfully captured high-level relationships that go beyond traditional handcrafted descriptors. The visualization of the learned latent space through t-distributed Stochastic Neighbor Embedding (t-SNE) revealed distinct clusters for various tremor types, showcasing a successful separation of feature manifolds. The hybrid feature space significantly enhanced inter-class separability when compared to the original 92-dimensional handcrafted feature space, which aligns with the findings from Paucar-Escalante et al. [8].

## 4.5 Comparative Analysis

When look at the performance of the Gradient Boosting classifier against SVM and Random Forest models, it stands out with its smoother decision boundaries and less overfitting, thanks to its iterative residual-learning approach. Although deep networks alone achieved impressive accuracy, they were quite sensitive to minor signal noise and tended to overfit on smaller classes. The hybrid model addressed these challenges by combining interpretable handcrafted features with learned latent embeddings, enabling the ensemble to leverage both domain-specific and nonlinear information. These results are consistent with the work of Karabayir et al. [3], who highlighted the robustness of Gradient Boosting in biomedical signal classification, and Yu et al. [5], who found that deep networks improved discriminative capacity when trained on normalized wearable data. The steps involving Butterworth filtering and time-frequency feature extraction also played a crucial role in maintaining stable performance by isolating the true tremor band and minimizing irrelevant noise.

## 4.6 Statistical Validation

To ensure the reliability of this research results conducted statistical significance testing using paired t-tests between the hybrid model and each baseline classifier. The tests confirmed that the performance enhancements of the hybrid model were statistically significant, with $p < 0.01$ across all metrics. Additionally, the Receiver Operating Characteristic (ROC) analysis showed an average AUC of 0.985, further validating the model's strong discriminative ability.

## 4.7 Discussion

The findings reveal that blending handcrafted statistical features with deep Autoencoder-derived latent embeddings and Gradient Boosting classification creates a remarkably accurate and interpretable system for detecting tremors in Parkinson's patients. The deep representation not only enhances the compactness of features but also ensures that the system is robust against outliers and variations in small samples. When compared to traditional machine learning methods, this hybrid approach strikes a perfect balance between interpretability and predictive power, which is essential for real-world clinical decision support systems. In summary, this study highlights how feature-level fusion and hybrid ensemble learning can significantly boost the performance of tremor subtype classification. This method could also be adapted for continuous monitoring or early detection of motor fluctuations in Parkinson's disease, providing a dependable, data-driven supplement to clinical evaluations.

## 5. CONCLUSION AND FUTURE WORK

This research introduced a hybrid deep-machine learning framework designed to classify tremor patterns in Parkinson's Disease using signals from accelerometers. The model employs a Butterworth Band-Pass Filter (2.5–12.5 Hz) to reduce noise, along with thorough handcrafted feature extraction, nonlinear feature learning via Autoencoders, and Gradient Boosting for the final classification. By combining statistical and deep-learned features, this research achieved a compact yet highly effective representation of tremor behavior, resulting in an impressive overall accuracy of 96.8% and an F1-score of 0.963.

In comparison to traditional classifiers like Support Vector Machines and Random Forests, proposed method showcased superior performance, stability, and interpretability. The addition of time-frequency and nonlinear features really boosted the model's ability to tell apart different types of tremors—like resting, postural, kinetic, and constancy tremors—by capturing both rhythmic and fleeting oscillatory patterns. When looked at feature importance, it became clear that dominant tremor frequency, spectral entropy, and latent Autoencoder embeddings played a key role in how well the model classified these tremors.

This really highlights the value of hybrid learning in understanding complex biomedical signals. Then integrating advanced temporal deep models like LSTM or attention-based transformers could enhance ability to track dynamic tremors and predict motor fluctuations early on. Plus, conducting large-scale multi-center validation and collaborating with clinical partners will be crucial to assess its diagnostic reliability and how well it fits into healthcare systems.

Overall, this proposed method lays a solid groundwork for smart, non-invasive tools for monitoring Parkinson's, which can help clinicians with early diagnosis and tailored treatment plans.

## References

1) San-Segundo, R., Zhang, A., Cebulla, A., Panev, S., Tabor, G., Stebbins, K., Massa, R. E., Whitford, A., de la Torre, F., & Hodgins, J. (2020). Parkinson's Disease Tremor Detection in the Wild Using Wearable Accelerometers. *Sensors*, 20(20), 5817. https://doi.org/10.3390/s20205817

2) Xing, X., Luo, N., Li, S., Zhou, L., Song, C., & Liu, J. (2022). Identification and Classification of Parkinsonian and Essential Tremors for Diagnosis Using Machine Learning Algorithms. *Frontiers in Neuroscience*, 16, 701632. https://doi.org/10.3389/fnins.2022.701632

3) Karabayir, I., Goldman, S. M., Pappu, S., et al. (2020). Gradient boosting for Parkinson's disease diagnosis from voice recordings. *BMC Medical Informatics and Decision Making*, 20, 228. https://doi.org/10.1186/s12911-020-01250-7

4) Lin, S., Gao, C., Li, H., Huang, P., Ling, Y., Chen, Z., Ren, K., & Chen, S. (2023). Wearable sensor-based gait analysis to discriminate early Parkinson's disease from essential tremor. *Journal of Neurology*, 270(4), 2283–2301. https://doi.org/10.1007/s00415-023-11577-6

5) Yu, J., Meng, K., Liang, T., Liu, H., & Wang, X. (2024). Improved Deep Learning for Parkinson's Diagnosis Based on Wearable Sensors. *Electronics*, 13(23), 4638. https://doi.org/10.3390/electronics13234638

6) van der Linden, C., Berger, T., Brandt, G. A., Strelow, J. N., Jergas, H., Baldermann, J. C., Visser-Vandewalle, V., Fink, G. R., Barbe, M. T., Petry-Schmelzer, J. N., & Dembek, T. A. (2023). Accelerometric Classification of Resting and Postural Tremor Amplitude. *Sensors*, 23(20), 8621. https://doi.org/10.3390/s23208621

7) Farashi, S., Barzegar, K., Mahmoudi, B., et al. (2023). Tremor onset detection and prediction using wrist accelerometry and EEG in Parkinson's disease: a feasibility study. *BMC Neurology*, 23, 420. https://doi.org/10.1186/s12883-023-03468-0

8) Paucar-Escalante, J., Alves da Silva, M., De Lima Sanches, B., Soriano-Vargas, A., Moriyama, L., & Colombini, E. (2025). Machine Learning Strategies for Parkinson Tremor Classification Using Wearable Sensor Data. *arXiv preprint* arXiv:2501.18671. https://doi.org/10.48550/arXiv.2501.18671

9) Polvorinos-Fernández, C., Sigcha, L., Pereira de Pablo, L., Borzí, L., Cardoso, P., Costa, N., Costa, S., López, J. M., de Arcas, G., & Pavón, I. (2024). Evaluation of the Performance of Wearables' Inertial Sensors for the Diagnosis of Resting Tremor in Parkinson's Disease. In *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies – HEALTHINF* (pp. 820–827). https://doi.org/10.5220/0012571600003657

10) Yue, P., Li, Z., Zhou, M., Wang, X., & Yang, P. (2024). Wearable-Sensor-Based Weakly Supervised Parkinson's Disease Assessment with Data Augmentation. *Sensors*, 24(4), 1196. https://doi.org/10.3390/s24041196