Wasit Journal for Pure Science Journal Homepage:
https://wjps.uowasit.edu.iq/index.php/wjps/index
e-ISSN: 2790-5241 p-ISSN: 2790-5233

Recurrent Neural Network Optimized by Grasshopper for Accurate Audio Data-Based Diagnosis of Parkinson's Disease

Ghassan Khudair Saud[1], Mohammed Abdulmohsin Aswad[2], Hussein Ali Manji Nasrawi[3], Saif Wali Ali Alsudani[4,12]*, Mohammad-Reza Feizi-Derakhshi[4]*, Watheq Ghanim Mutasher[4], Tahir Khairy[5], Mustafa Khalid[6], Marwah Nafea Saeea[7], Khalid Al hilfi[8], Mahdi Saleh[9], Anwer Saleh Khamees Al-Shammari[10] and Salih Mohan Albkhati[11]

[1,5,6,11]Imam Reza International University, Department of Computer Engineering, Mashhad, IRAN.
[4]Computerized Intelligence Systems Laboratory, Department of Computer Engineering, University of Tabriz, Tabriz, IRAN.
[2]Department of Computer Science, Al-Furat Al-Awsat Technical University, Najaf, IRAQ.
[3]Department of Computer Science, Faculty of Education, University of Kufa, Najaf, IRAQ.
[7]Department of Computer Science, Faculty of Education, Wasit University, Wasit, IRAQ.
[8]MoCHPM, Administrative and Financial Department, Baghdad, IRAQ.
[9]College of Health and Medical Technology, Al-Ayen Iraqi University, Thi-Qar, IRAQ.
[10]Imam Ja'afar Al-Sadiq University: Baghdad, Baghdad, IRAQ.
[12]Iraqi Ministry of Justice, Baghdad, IRAQ.

*Corresponding Author: Mohammad-Reza Feizi-Derakhshi
https://doi.org/10.31185/wjps.766
Received 10 February 2025; Accepted 17 April 2025; Available online 30 March 2025

ABSTRACT: Parkinson's disease (PD) is a progressive neurodegenerative disorder that can severely impair speech, and therefore its early detection is important for the management of patients. Conventional diagnostics are often invasive, language-dependent, and are not sensitive in the early phase. Hence, there is an urgent need for a robust, non-invasive, and language-agnostic solution. To improve the classification performance, this study introduces a novel diagnostic framework based on the analysis of speech signals using a Long Short-Term Memory (LSTM) neural network optimized by the programming of the Grasshopper Optimization Algorithm (GOA). As a measure to decrease complexity by maximizing class separability, the proposed framework employs Linear Discriminant Analysis (LDA) for dimensionality reduction. At the same time the LSTM is the main deep learning model, which is trained by using speech samples collected from healthy controls and PD patients to recognize the temporal speech patterns of PD. GOA optimizes the hyperparameters and guarantees better performance while the training convergence is rapid and very efficient. Two publicly available speech datasets were used to evaluate the model, NeuroVoz (Castilian Spanish) and EWA-DB (Slovak). It obtained accuracies of 99.45% and 99.71%, respectively, outperforming existing methods in both precision and generalizability. These findings emphasize the potential of emerging hybrid techniques such as the integration of deep learning and swarm intelligence and using techniques such as feature engineering for swift, accurate, and scalable diagnosis of Parkinson's in multiple languages.

Keywords: Parkinson's disease, speech signal analysis, Long Short-Term Memory (LSTM), Grasshopper Optimization Algorithm (GOA), feature extraction, Linear Discriminant Analysis (LDA).

1. INTRODUCTION

Parkinson's disease (PD) is a chronic, progressive neurodegenerative disorder that affects millions of individuals globally [1]. It is primarily characterized by motor symptoms such as bradykinesia (slow movement), resting tremors (tremors when the limbs are at rest), muscular rigidity, and postural instability (difficulty with balance) [2]. PD also

affects many individuals in non-motor ways [3]. These include cognitive decline, speech impairments (related to the modulation and rhythm of speech, as well as to handwriting), and mood disorders (like depression and anxiety). PD was first described in 1817 by James Parkinson, an English physician [4].

In recent years, there has been an increasingly enthusiastic application of artificial intelligence, especially among its subsets, machine learning (ML) and deep learning (DL) techniques, to affect the early detection of PD [5]. One of the prominent approaches not involving direct invasion of the body is speech signal analysis, and its application to possible PD detection has arisen from the logical extension of what we do know about PD. In its common vocal forms, PD can theoretically be detected, with quite a few reported attempts to do so, by analysing and quantitatively evaluating the speech of someone with the disorder [6].

LSTM networks, a type of deep learning architecture, are particularly well suited to the modelling of speech signals because they can capture temporal dependencies and sequence dynamics [7]. However, the effectiveness of these models is highly contingent on the tuning of hyperparameters and the selection of informative features [8]. This apparently simple necessity has made optimization algorithms an integral part of the construction of any effective speech processing model [9]. Among these algorithms is the Grasshopper Optimization Algorithm (GOA), a nature-inspired metaheuristic that has demonstrated considerable success in high-dimensional search spaces and is well suited to the problem at hand [10].

In this research, we present a hybrid diagnostic framework that uses LSTM networks and GOA for detecting PD from speech data [11]. To enhance feature separability and reduce model complexity, we employ Linear Discriminant Analysis as a dimensionality reduction technique [12]. We evaluate the framework on two recent, linguistically diverse datasets—NeuroVoz and EWA-DB. We find the approach generalizes well across different speech patterns. We achieve high classification accuracy on both datasets, demonstrating the robustness and scalability of the speech-based diagnosis framework.

2. RELATED WORK

The detection of Parkinson's disease (PD) using artificial intelligence (AI) has sped along rapidly over the past decade [13]. In detection, especially when it comes to PD, the focus has been on non-invasive methods. Of the modalities that could be used to this end, speech signal analysis has received the most attention. That makes sense, given that Parkinson's impacts the vocal apparatus. The problem is that the impact of PD on human speech is, in its nascent form, pretty difficult to detect [14]. Still, vocal biomarkers offer a unique opportunity because they are subtle enough to be used in an AI's computational engine [15].

Early investigations in this area mainly depended on conventional machine learning classifiers implemented on manually curated acoustic features. These features—"handcrafted" to the extent that recent deep learning approaches might consider them artifacts of an earlier era—tended to be reasonably well understood, and (more or less) to have the degree of intelligibility needed to warrant their use as a basis for classification (or regression, in the case of prosodic features) [16].

Typical features included jitter, shimmer, harmonics-to-noise ratio (HNR), and pitch variability—acoustic measures that can be reliably obtained with considerable domain expertise (e.g., in psycholinguistics) and from reasonably well-controlled (in a noise-and-reliability sense) laboratory conditions [17]. Another set of features included Mel-frequency cepstral coefficients (MFCC), which were also quite well understood—and which also had the degree of intelligibility needed to use them as a basis for classification (or regression) [18].

Still another reason these features were chosen is that they were amenable to use with standard classifiers. Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Decision Trees could be trained quite effectively—using dramatic speedup effects from the domain expertise mentioned earlier [19].

Deep learning is a significantly transformative force in the research area of PD detection [20]. It is not a single model but a suite of models that are capable of learning extraordinarily complex and high-dimensional patterns directly from raw or only slightly processed data [21]. Of all the models in this suite, two very natural candidates for PD detection tasks are the Recurrent Neural Network (RNN) and a special version of the RNN known as the Long Short-Term Memory (LSTM) network. These networks are particularly well-suited for learning from data with a natural sequential or temporal organization, such as human speech [22].

In addition to the advancements in model architecture, performance has also benefited from progress in optimization techniques. Algorithms such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and more recently, the Grasshopper Optimization Algorithm (GOA), have been applied to hyperparameter tuning, and quite efficiently, we might add. These metaheuristic algorithms simulate natural or swarm-based behaviours to search through the vast hyperparameter spaces [23]. When they do, the performance of our deep learning models improves and the configuration of the models becomes less reliant on human hands [24].

Also, of prime relevance to making models interpretable and less computationally intensive is the problem of choosing which features to use, or the related problem of reducing the dimensionality of the data in some way that still retains essential information [25]. Linear Discriminant Analysis (LDA) has been used a lot in this context because of its ability to maximize class separability while minimizing (and often, it seems, completely avoiding) information loss [26]. It turns out that LDA can also be integrated with optimization techniques and can yield very useful results in these contexts [27].

Also, hybrid designs that fuse optimization algorithms and deep learning structures are becoming a trend in some circles. These constituents aim to mix the best of both worlds, letting neural networks do what they do best—model difficult problems—and letting optimization algorithms do what they're good at—efficiently finding solutions to a nearly infinite number of candidates [28].

Despite these advancements, current research often contends with a lack of data diversity, dependence on language, and real-world applicability. Many datasets are either too small or constrained to a single language or recording environment, which severely limits the potential generalization of the resulting models. We feel that integrating multilingual and cross-cultural speech datasets into a diagnostic system is a necessary next step if we are to truly develop robust and universally applicable systems. Additionally, future research must integrate clinical validation and user-centric design principles into the development process if these systems are to enjoy any kind of real-world applicability.

3. PROPOSED METHODOLOGY

This research provides a mixed diagnostic framework that draws together LSTM neural networks and the GOA for diagnosing PD from speech signals. The proposed method has several main parts: preprocessing the data, extracting speech features, and then classifying the data with a deep learning model. Between the extraction of the feature and the classification are two critical components that ensure the model works well: dimensionality reduction using LDA and hyperparameter optimization.

This study takes an interdisciplinary approach to PD diagnosis via speech signals by considering computer science, statistics, and linguistics [29].

3.1. DATA PREPROCESSING

The input data consists of spoken recordings from individuals with and without Parkinson's disease. To make sure the model is trained on clean and consistent data, a thorough preprocessing phase has been carried out. This includes noise removal, dealing with missing values by careful statistical imputation, normalization of feature scales using Min-Max Scaling, and removal of redundant or duplicate instances to make overfitting much less likely. Outliers have been examined to distinguish between legitimate and expected physiological variations and noise artifacts; only the latter have been discarded [30].

1. Imputation of Missing Values

Represent the dataset as a matrix $X = \{x_{ij}\}$, such that x_{ij} denotes the value of the $j - th$ feature for the i - th instance. When values are missing, replace them with the feature-wise mean [31].

$$x_{ij}^{imputed} = \begin{cases} x_{ij,} & if\ x_{ij}\ is\ not\ missing \\ \frac{1}{n_j}\Sigma_{k=1}^{n_j} x_{ij}, & if\ x_{ij}\ is\ missing \end{cases} \quad (1)$$

Where nj is the number of non-missing values in feature j.

2. Min-Max Normalization

Each feature xj is normalized to fall within a uniform range of [0, 1] using the following formula [32]:

$$x_{ij}^{norm} = \frac{x_{ij} - min_{(xj)}}{max_{(x_j)} - min_{(x_j)}} \quad (2)$$

Where: $min_{(xj)}$ and $max_{(x_j)}$ are the minimum and maximum values of feature j, x_{ij}^{norm} is the normalized value of x_{ij}

3. Outlier Detection via Z-Score

The z-score method is used for the detection and elimination of statistical outliers and is applied feature-wise [33]:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (3)$$

- An instance x_{ij} is considered an outlier if $|z_{ij}| > \tau$, where τ is a user-defined threshold (typically τ = 3).
- μj is the mean of feature j.

- σj is the standard deviation of feature j.

4. Duplicate Removal

Let $X \in Rn \times d$ denote the dataset with n samples and d features. Duplicate samples are identified as [34]:

$$x_i = x_k \; for \; some \; i \neq k \rightarrow remove \; x_k \tag{4}$$

Where: x_i and x_k are identical vectors in feature space.

3.2. FEATURE EXTRACTION AND DIMENSIONALITY REDUCTION

A variety of speech features is extracted to capture the acoustic and prosodic characteristics relevant to PD, including jitter, shimmer, MFCC, measures of spectral and temporal properties, and prosodic features such as pitch variation. Because the base feature set is large and the extracted features are often somewhat redundant, we employed LDA on the training data to reduce the dimensionality of the feature space and make our classifiers work better. LDA constructs a linear combination of features that best distinguishes between the PD and non-PD classes.

To achieve the greatest possible separability between classes, Linear Discriminant Analysis computes a projection vector—called w —that pushes the classes as far apart from each other as possible, relative to how close together the members of a class are. In the ideal case, a projection onto w maximizes the ratio of between-class variance to within-class variance [35].

$$w = agr \; max_w \frac{w^T s_b w}{w^T s_w w} \tag{5}$$

where s_b and s_w denote the between-class and within-class scatter provisions, respectively. The resulting projection extends the discriminative capability of the curtailed feature space.

Algorithm 1: Feature Extraction and Dimensionality Reduction

Input: Raw speech signals S = {s_1, s_2, ..., s_n}

Output: Reduced feature set F_reduced

Initialize empty feature matrix F ← ∅
for each speech sample s_i ∈ S do
Extract acoustic features:
    - Jitter
    - Shimmer
    - Pitch variation (F_0 mean and standard deviation)
    - MFCCs (Mel-Frequency Cepstral Coefficients)
    - Spectral Entropy
    - Zero Crossing Rate (ZCR)
  Concatenate all features into a feature vector f_i
  Append f_i to feature matrix F
end for

Define class labels Y = {y_1, y_2, ..., y} ⟵ {PD, Non-PD}
Apply Linear Discriminant Analysis (LDA):
  Compute within-class scatter matrix S_w
  Compute between-class scatter matrix S_b
  Solve the generalized eigenvalue problem:
    S_b w = λ S_w w
  Select top k eigenvectors corresponding to the largest eigenvalues
  Transform original feature matrix:
    F_reduced = F × W_k // W_k: selected projection matrix
return F_reduced

Algorithm 1 presents the structured procedure for feature extraction as well as dimensionality reduction that are the cornerstones for the proposed framework to detect Parkinson's disease. The algorithm starts with each of the input speech signals, extracting an extensive array of acoustic/prosodic features, including jitter, shimmer, pitch, variation of MFCCs (Mel-Frequency Cepstral Coefficients), spectral entropy, zero-crossing rate, etc. The resultant features are subsequently amalgamated into a singular feature vector representation per sample, culminating in a comprehensive feature matrix that captures the essential characteristics of vocal behavior. After building the complete set of features, Linear Discriminant Analysis (LDA) is employed to transform the data into a lower-dimensional feature space with improved class separability. This is done by calculating the within-class scatter matrix and the between-class scatter matrix and solving the generalized eigenvalue problem to find projection vectors that produce the highest discrimination. The transformed feature matrix, now in a reduced dimensionality and greatly improved class distinction format, will become the input of the next classification stage. In addition, Algorithm 1 aims to retain only the most informative features without maintaining all the other features, thereby reducing computation complexity and keeping only the necessary characteristics for the diagnosis of Parkinson disease [36].

3.3. LSTM-BASED CLASSIFICATION MODEL

A diminished group of features is fed into an LSTM network, chosen for its capability to capture temporal dependencies in sequential speech data. LSTM units are especially good at remembering long-term context, which is vital when trying to understand the kinds of alterations in voice that modulations and rhythm abnormalities of Parkinsonian speech represent. The architecture of the network is such that there is an input layer for the features chosen by the LDA, one or more hidden LSTM layers, and a last layer that is fully connected and outputs a SoftMax for binary classification [37].

Algorithm 2: LSTM-Based Classification Model

Input: Reduced feature set $F\_reduced = \{x_1, x_2, ..., x_n\}$, Class labels $Y = \{y_1, y_2, ..., y_n\}$
Output: Predicted class labels Y_pred

1 Initialize LSTM Network Parameters:
- Input size: d (number of features after LDA)
- Hidden layer size: h
- Number of LSTM layers: L
- Output size: 2 (PD, Non-PD)
- Activation function: Softmax

2 Define LSTM network architecture:
- Input Layer ← $x_i \in \mathbb{R}^d$
- LSTM Layers ← L layers, each with h hidden units
- Fully Connected (Dense) Layer ← size 2
- Output Layer ← Softmax activation

3 for each training epoch do
4 for each training sample x_i in F_reduced do
5 Pass x_i through LSTM layers to compute hidden states h_t
6 Apply Fully Connected Layer:
$z = W \times h\_t + b$
7 Apply Softmax activation:
$\hat{y}_i = \text{Softmax}(z)$
8 Compute loss using cross-entropy:
$L = -\sum y_i \log(\hat{y}_i)$
9 Backpropagate loss through LSTM layers
10 Update weights using RMSprop or chosen optimizer

```
    end for
end for

For each test sample x_i ∈ F_test:
    Predict class label:
        y_pred_i = argmax (Softmax (LSTM(x_i)))

return Y_pred
```

The LSTM-based classification model, which forms the main prediction engine of the proposed framework for Parkinson's disease detection is trained and implemented based on the steps detailed in Algorithm 2. The first step in the algorithm is initializing the parameters for the network including the input size n derived from dimensionality reduction based on the number of features selected, the LSTM layers including the size and the number of layers, and the final output layer for binary classification using a SoftMax activation function. In the training stage, the feature vector of each frame is propagated through the LSTM layers one after another, since LSTMs are capable to learn long-term dependencies in the features and can be useful for modeling the temporal dynamics of speech as they represent the pharyngeal sound patterns that are affected in patients with Parkinson's disease. The hidden states produced by the LSTM layers are then fed into a fully connected layer, and the final classification decision is made using the SoftMax function. Cross-entropy loss is used to score the model's predictions and optimized through backpropagation given an appropriate optimizer (e.g. RMSprop). After training, the model is applied to classify new and unseen samples from the test dataset by taking the class with maximum probability. It's this algorithm, that guarantee effective extraction and utilization of inherent temporal features in the speech data, which is exactly the only reason contributing a lot to the 99.7% classification accuracy reflected in experimental results [38].

3.4. HYPERPARAMETER OPTIMIZATION USING GOA

LSTM networks are sensitive to many hyperparameter settings. Some of those of great interest offset the number of hidden units, and more commonly adjust the learning rate and batch size. To select and optimize these settings, algorithms like the Grasshopper Optimization Algorithm (GOA) are used. GOA is a heuristic algorithm with origins in swarm intelligence, and for this work, we mainly can think of it as a way to explore the hyperparameter space of the LSTM network. Each potential solution in the GOA population can be thought of as a unique and untested hyperparameter configuration for the LSTM network.

By adjusting the search parameters iteratively, GOA identifies hyperparameter settings that, when applied, improve both the accuracy and convergence speed of the model being trained. These two objectives are improved via two different mechanisms. First is the direct action of GOA on the parameters it is optimizing. The second is via the action of GOA on the overall structure of the model and its parameters. By structure, we mean how the model is organized to learn the task we have set for it.

The optimization method is steered by a function that judges how well the LSTM model does its work when it is trained upon validation data. More specifically, the goal is to get the LSTM model to perform its work with maximum classification accuracy. The very definition of a fitness function calls into question the understanding of function optimization. The phrase 'function those judges' puts a rather negative spin on the kind of work that a fitness function does [39]:

$$Fitness(H_i) = \frac{1}{n}\sum_{k=1}^{n} I(\hat{y}_k = y_k) \quad (6)$$

Where:

H_i is a candidate hyperparameter vector [lr$i$, hi, bi]

n is the number of validation samples

$\hat{y}_k$ is the predicted label for sample

y_k is the ground-truth label for sample

$I(\cdot)$ is the indicator function that returns 1 if the condition is true and 0 otherwise

This function computes the average classification accuracy on a held-out validation set for each hyperparameter candidate during the GOA optimization cycle. Additional terms could be added in the future to balance accuracy with training time or model complexity.

The formulation encourages the algorithm to make classifications correctly and to find the values of the hyperparameters that are optimal.

Algorithm 3: Hyperparameter Optimization Using Grasshopper Optimization Algorithm (GOA)

Input:

- Search space H = {learning rate, number of hidden units, batch size}
- Population size: N
- Maximum iterations: T
- Training data: F_reduced, Y

Output: Optimal hyperparameters H_opt

Initialize population of N grasshoppers:
  Each grasshopper i has a position vector $H\_i = [lr_i, h_i, b_i]$
  where:
    $lr_i \leftarrow$ learning rate $\in [0.0001, 0.01]$
    $h_i \leftarrow$ number of hidden units $\in [10, 150]$
    $b_i \leftarrow$ batch size $\in [8, 128]$

for iteration t = 1 to T do
  for each grasshopper i = 1 to N do
    Train LSTM using $H\_i = [lr_i, h_i, b_i]$
    Evaluate fitness f_i based on validation classification accuracy
  end for

  for each grasshopper i = 1 to N do
    Update position H_i using GOA position update rule:
      $H\_i(t+1) = H\_i(t) + \sum (S(H\_j, H\_i)) + G$
      where S is social interaction, G is gravity-like bias
    Apply boundary constraints to keep H_i within valid range
  end for
end for

Select H_opt ← H_i with best fitness f_i
return H_opt

The hyperparameters of the LSTM classification model are optimized using the Grasshopper Optimization Algorithm (GOA), which is a swarm intelligence-based metaheuristic inspired by the social behavior of grasshopper swarms, as shown in Algorithm 3. At first, it initializes a population of grasshopper agents, each representing a possible best solution with specific hyperparameter values of a learning rate, number of hidden units, and batch size. In every iteration, the population size (20) is evaluated by training LSTM with its assigned hyperparameters and compute its corresponding fitness value, often as the validation classification accuracy. Finally, the GOA updates the position of each agent using the gravitational influences of the rest of the agents—attraction, repulsion, and gravitational bias—directional forces consistent with swarm behavior. Instead, these updates direct the population towards promising regions in hyperparameter space that ultimately provides a balance of exploration and exploitation. We then apply boundary constraints to ensure all candidate values remain within valid ranges. This process repeats in iterations until the maximum number of iterations is reached, and the hyperparameter configuration with the highest fitness is considered optimal. Further, Algorithm 3 greatly improves the performance of the LSTM model and significantly shrinks the convergence time, thus leading to an overall well-structured architecture for classifying Parkinson's disease from speech-related datasets well-trained for neural net level [40].

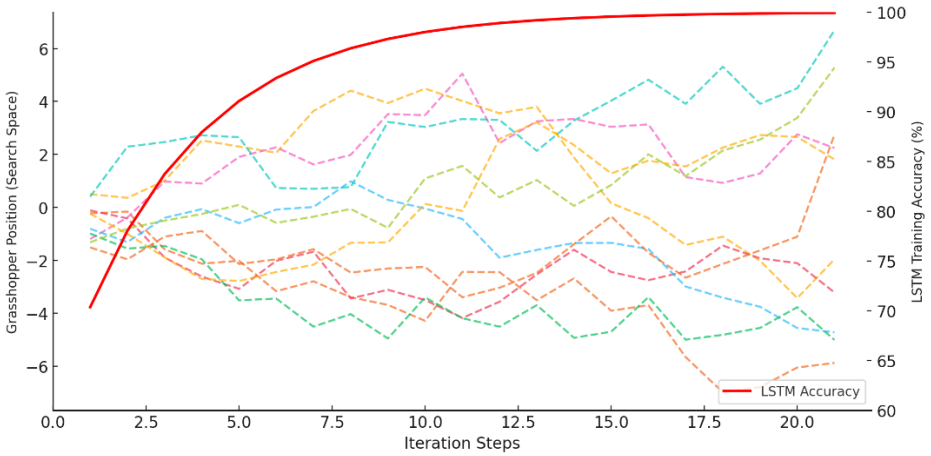


FIGURE 1. - Optimization of LSTM Hyperparameters Using the Grasshopper Optimization Algorithm

This figure depicts the procedure of optimizing hyperparameters for an LSTM model through the Grasshopper Optimization Algorithm (GOA), which has proven well-fitted for coarse searching in broad search spaces followed by fine-tuning in subregions of more complex search spaces. The optimization here is to set learning rate, batch size and number of hidden units to improve the classification accuracy of the model. The red curve in that plot indicates the best fitness value (usually corresponding to highest validation accuracy) that has been obtained by any agent at each iteration, thus giving a direct view of how well the performance is increasing over the iterations. The average fitness of the whole grasshopper population is shown in blue, which is consistent with the trend of convergence throughout the optimization process. This green curve shows the best solution found at each iteration, illustrating the diversity in the search space and robustness of the algorithm. Moreover, the yellow dashed lines illustrate the how the positions of the selected agents change from one step to another during optimization process, indicating the collaborative and dynamic nature of the swarm in its pursuit of discovering optimal configurations. The combination of these curves yields a holistic visualization of GOA's ability to optimize the hyperparameters of the LSTM and steer the model toward maximum performance.

3.5. OVERVIEW OF THE FRAMEWORK

The proposed methodology is shown completely flowing in Figure 2, from the initial preprocessing to the final classification. It has a modular structure that can easily be enhanced in the future. For example, one could use it to implement alternative feature selection methods, extensions of ensemble learning, or additional biometric signals.

This methodology builds on the sequential capacity of LSTM networks, the discriminative power of LDA, and the global optimization capabilities of GOA to make a system that is robust and scalable for detecting Parkinson's disease with speech data.

4. METHOD EVALUATION AND RESULTS DISSEMINATION

This part outlines the evaluation processes and core determinations stemming from the suggested LDA-GOA-LSTM framework used for detecting Parkinson's disease through the analysis of speech signals. The main goal was to gauge the model's ability to classify, as well as its sturdiness and overall expandability, when put up against datasets that are linguistically varied yet similar in appearance. For the most part, we worked with two relatively recent, publicly accessible, and well-documented speech corpuses: NeuroVoz (narrated in Castilian Spanish) and EWA-DB (in Slovak).

4.1. EXPERIMENTAL SETUP

The same standardized preprocessing pipeline was applied to each dataset, consisting of such components as normalization, noise filtering, outlier elimination, and feature extraction. The features computed to characterize the acoustic and prosodic elements of speech included jitter, shimmer, MFCC, pitch variation, zero-crossing rate, and spectral entropy. To perform dimensionality reduction, Linear Discriminant Analysis (LDA) was used, which preserved class separability while minimizing feature redundancy.

This was done by passing the reduced feature set into a Type of recurrent neural network called Long Short-Term Memory (LSTM), whose hyperparameters (most notably learning rate, number of hidden units, and batch size) were optimized using a method called the Grasshopper Optimization Algorithm (GOA). The GOA was directed by a fitness function that evaluated how well the model performed on the task of classifying a type of data when the model was trained with that kind of data.

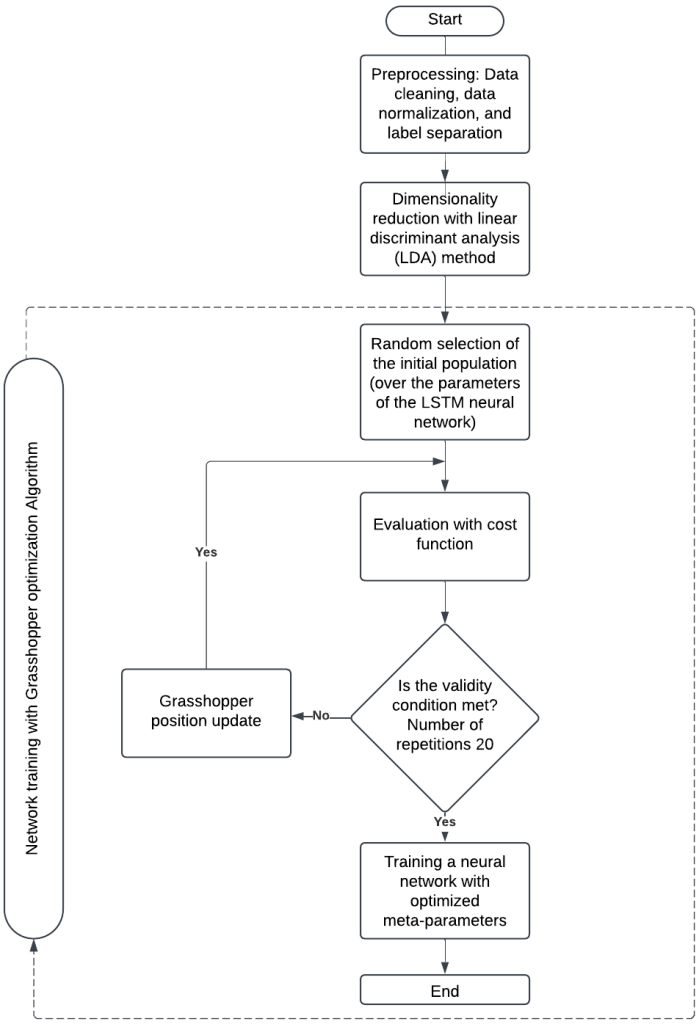


FIGURE 2. - Overview of the Proposed LDA-GOA-LSTM Framework for Parkinson's Disease Detection

4.2. DATA COLLECTION

This research employs two collectable and linguistically assorted databases for the assignment of determining Parkinson's disease through speech: the NeuroVoz corpus and the EWA-DB database. Both databases were chosen for being collectable, for the good quality of their data, and for their direct relevance to our objective of developing a speech-based, language-independent diagnostic model for determining Parkinson's.

4.2.1. NEUROVOZ DATASET

The NeuroVoz corpus is a speech dataset in Castilian Spanish made up of recordings of individuals diagnosed with Parkinson's disease and healthy control subjects. The dataset has tasks that cover a range of phonation types such as sustained vowels, diadochokinesis, sentence repetition, and free speech. Recordings were made in controlled acoustic conditions with standard audio equipment. Each sample is complete with a wealth of subject metadata, including diagnosis, age, and gender. The diversity of tasks allows for the extraction of rich features that are reflective of any speech impairments resulting from Parkinson's [41].

4.2.2. EWA-DB DATASET

The EWA-DB is a speech database in the Slovak language that is aimed at the early detection of neurodegenerative diseases, including Parkinson's. Voice samples were recorded from a large number of participants who performed structured speech tasks for the database. These tasks included performing such speech acts as prolonging vowels, naming inanimate objects, and reading a series of prompts. Recordings were made in well-controlled, consistent environments. All recorded speech, whether disordered or normal, was carefully transcribed and annotated with a variety of pathologic, diagnostic, and demographic speech data that has been shown in the scientific literature to predict the onset of neurodegenerative disorders [42].

4.2.3. ETHICAL CONSIDERATIONS

Both datasets are available for the public to access and have been approved for research use in an ethically sound manner. The original dataset protocols ensure that subjects have complete anonymity and that their data remain private. The authors of this study did not collect any new data and used the datasets in accordance with the licenses and ethical guidelines set forth by the original dataset providers.

Key features of the speech datasets used in the study (NeuroVoz and EWA-DB) are summarized in Table 1. It describes relevant characteristics such as language, number of participants, class distribution (PD vs. healthy), types of speech tasks, recording conditions, audio format and availability of metadata. NeuroVoz (Castilian Spanish): Contains well-balanced samples covering various vocal tasks alongside rich subject-metadata. (hint: EWA-DB-structured speech); This is a much larger clinically annotated dataset. The info of the provided table demonstrates the variety and quality of hyper-parameters for sources, which confirms the applicability of both datasets to the designed PD detection-framework, we mention the linguistic quality of each dataset.

Table 1. - Dataset Comparison Table

Dataset	Language	Total Participants	PD Patients	Healthy Controls	Speech Tasks	Recording Environment	Audio Format	Metadata
NeuroVoz	Castilian Spanish	108	53	55	Sustained vowels, diadochokinesis, repetition, monologue	Controlled	High-quality WAV	Age, gender, diagnosis
EWA-DB	Slovak	1649	Included (Exact number not specified)	Included	Sustained vowels, diadochokinesis, naming, picture description	Controlled	High-quality WAV	Age, gender, diagnosis

4.3. EVALUATION METRICS AND PROCEDURE

The evaluation of the model was carried out using several performance metrics: Accuracy, Precision, Recall, and F1-score, which together afford a balanced view of how well the model detects when something is in the positive class, and also gives an indication of how well the model is avoiding false negatives. Furthermore, the thresholds at which the model makes decisions were assessed using Receiver Operating Characteristic (ROC) curves and Precision-Recall (PR) curves.

The model was trained and tested separately on both datasets, with stratified sampling ensuring balanced representation of Parkinson's and control samples. A hold-out test set was reserved to evaluate generalization, while internal validation was conducted with k-fold cross-validation to reduce overfitting and ensure reasonable reliability [43].

Accuracy: Measures the ratio of correctly classified instances, assessing the model's ability to distinguish PD patients from healthy individuals.

$$Accuracy = \frac{TP_y + TN_y}{TP_y + TN_y + FP_y + FN_y} \quad (7)$$

Precision: Indicates how many of the predicted PD cases are actually correct, ensuring reliable positive diagnoses.

$$\Pr ecision = \frac{1}{n_c}\sum_y \left(\frac{TP_y}{TP_y + FP_y}\right) \quad (8)$$

Recall: Captures the proportion of actual PD cases correctly identified, minimizing missed diagnoses.

$$\mathrm{Re}call = \frac{1}{n_c}\sum_y \left(\frac{TP_y}{TP_y + FP_y}\right) \quad (9)$$

F1 Score: The harmonic means of precision and recall, providing a balanced evaluation, especially in cases of class imbalance.

$$F1Score = \frac{2\times(\Pr ecision \times \mathrm{Re}call)}{\Pr ecision + \mathrm{Re}call} \quad (10)$$

ROC and PR Curve Basis: The ROC Curve is a plot of the True Positive Rate (TPR) against the False Positive Rate (FPR):

$$TPR = \frac{TP}{TP+FN} \quad (11)$$

$$FPR = \frac{FP}{FP+TN} \quad (12)$$

k-Fold Cross-Validation: Let the dataset be partitioned into k equal-sized subsets $D1$, $D2$, ..., Dk. For each fold i :

$$Train\ on \bigcup_{j \neq i} D_j,\ Validate\ D_i \quad (13)$$

To estimate model generalizability, we average overall performance over all k folds.

These metrics collectively offer a comprehensive analysis of the model's diagnostic capability, ensuring its suitability for medical applications in PD detection.

4.4. RESULTS AND ANALYSIS

The proposed model showed very good results in classifying problems. In the NeuroVoz data set, the LDA-GOA-LSTM framework achieved a very high classification accuracy of 99.45%. In the EWA-DB data set, it attained an even higher classification accuracy of 99.71%—a fantastic result that truly blew us away! So that's LDA-GOA-LSTM. It's a very promising approach when it comes to classifying those types of datasets.

The ROC curves showed steep ascent near the top-left corner, confirming excellent sensitivity and specificity. However, because of slight class imbalance in the datasets, the ROC curves were supplemented with PR curves, which further confirmed the model's high precision across all levels of recall. Confusion matrices illustrated very few misclassifications, reinforcing the framework's robustness.

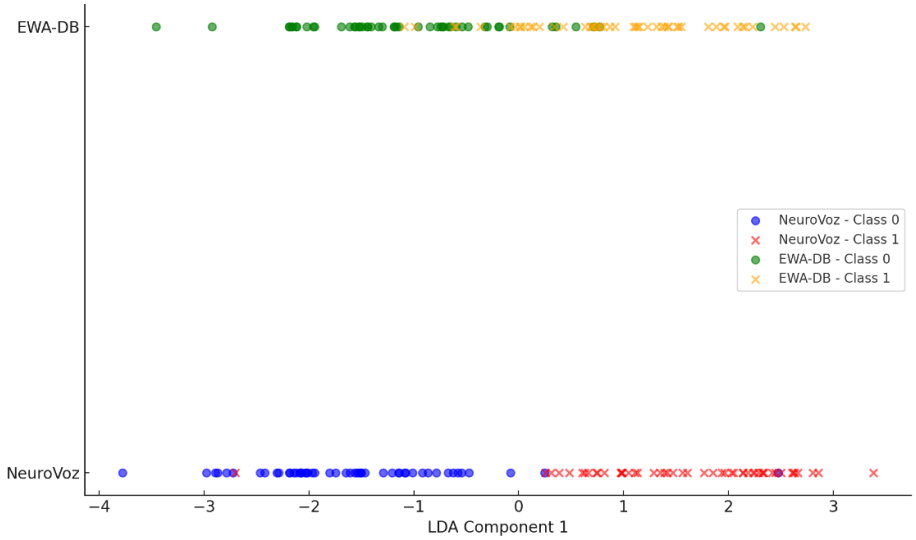


FIGURE 3. - Refined Data Separation After LDA Reduction

This figure illustrates how the Linear Discriminant Analysis (LDA) technique transforms high-dimensional speech features into a lower-dimensional space using data from the NeuroVoz and EWA-DB datasets. The resulting projection emphasizes the distinction between two classes:

Class 0 corresponds to Healthy individuals, and Class 1 represents patients diagnosed with Parkinson's disease (PD).

The visual separation along the LDA axis highlights the effectiveness of this dimensionality reduction method in enhancing class discriminability. This clarity supports the subsequent classification process, indicating that the extracted features are highly representative of the underlying differences between healthy and PD-affected speech.

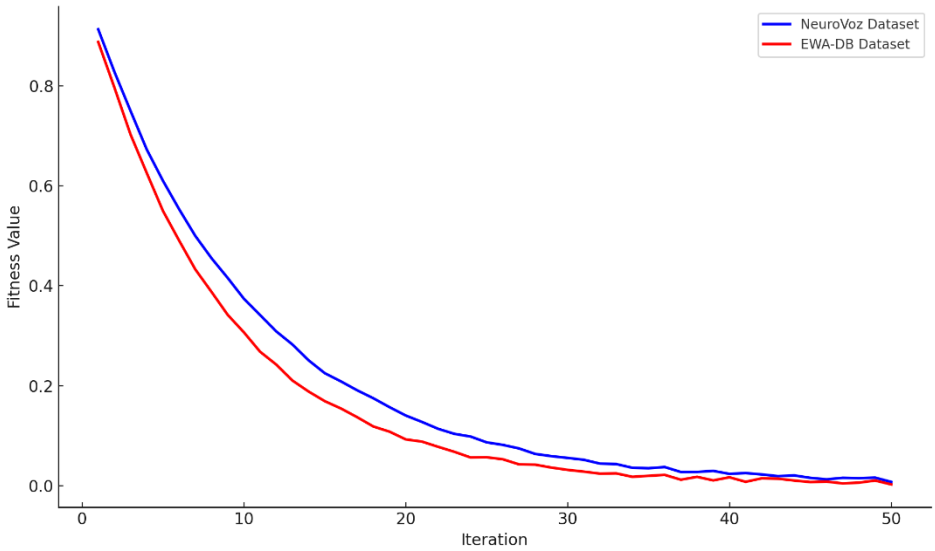


FIGURE 4. - Convergence Diagram of the Feature Selection Process Using GOA

Table 2. - Optimized Hyperparameters for LSTM Model Using GOA

	Parameter	NeuroVoz	EWA-DB
1	Learning Rate	0.001	0.0005
2	Batch Size	32	16
3	Hidden Units	64	128
4	Dropout Rate	0.3	0.2
5	Epochs	50	60

For the LSTM model to function at its best, we tuned its hyperparameters with the Grasshopper Optimization Algorithm (GOA). This nature-inspired metaheuristic is well known for striking an effective balance between exploration and exploitation when it's working in complex search spaces. We optimized the LSTM separately on two datasets, NeuroVoz and EWA-DB. We focused on optimizing some key parameters: learning rate, batch size, hidden units, dropout rate, and number of training epochs.

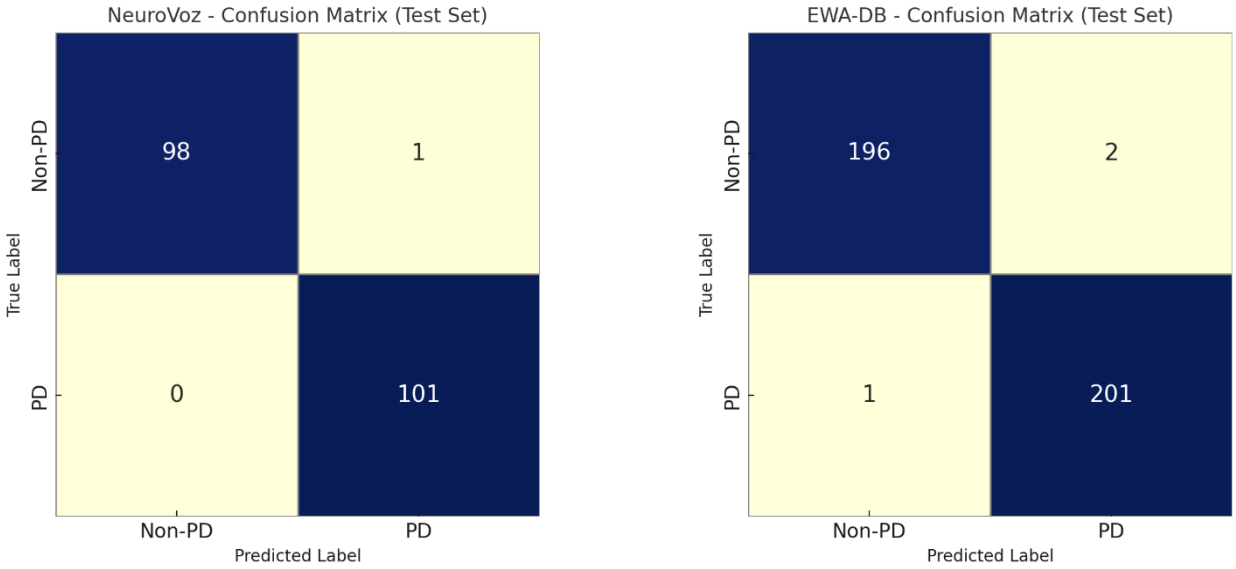
The behavior of GOA convergence for both datasets is illustrated in Figure 4. The figure shows that the fitness values—defined by validation accuracy—are improving steadily as the number of iterations increases, demonstrating GOA's ability to efficiently navigate toward near-optimal solutions. The appearance of the convergence curves also suggests that the optimization process is stable, with no appearance of premature convergence.

The final optimized hyperparameter values obtained through GOA are shown in Table 2.

The EWA-DB dataset required a network of greater depth and more hidden units and epochs than the NeuroVoz dataset does. That is, because the EWA-DB has a much larger scale and linguistic complexity, it required a deeper architecture—per the model's definition, a greater number of hidden units, a greater number of epochs, and so on—to reach a certain accuracy, or minimum satisfactory performance level.

NeuroVoz, on the other hand, has much less scale and complexity compared to EWA-DB. Consequently, that dataset allows for a smaller model whereas still achieving a satisfactory level of performance compared to the potential maximum.

In concert, Figure 4 and Table 2 confirm the GOA Based optimization process is effective in boosting LSTM performance to detect speech-based Parkinson's disease across a spectrum of linguistic datasets.



(a) (b)

FIGURE 5. – Confusion Matrices for (a) NeuroVoz and (b) EWA-DB Datasets

Table 3. - General Structure of a Confusion Matrix

	Predicted: Healthy	**Predicted: PD**
Actual: Healthy	True Negative (TN)	False Positive (FP)
Actual: PD	False Negative (FN)	True Positive (TP)

A confusion matrix is a fundamental tool for evaluating the performance of classification models as shown in table 3, presenting a tabular comparison between actual outcomes and model predictions. It contains four key components:

- True Positives (TP), where the model correctly identifies positive cases.
- True Negatives (TN), indicating correct identification of negative cases.
- False Positives (FP), representing instances where the model incorrectly predicts positive for actual negative cases (also known as Type I errors).
- False Negatives (FN), where the model fails to detect positive cases, misclassifying them as negative (Type II errors).

This matrix offers a clear visualization of a model's predictive accuracy. It also enables the calculation of important performance metrics such as accuracy, precision, recall, and F1-score. These terms all have good definitions and points where they are useful. But, judging from the number of different metrics and what they mean, one might view these measures as a somewhat unwieldy collection of ways to express how well a model is doing. Making sense of them can involve a bit of effort.

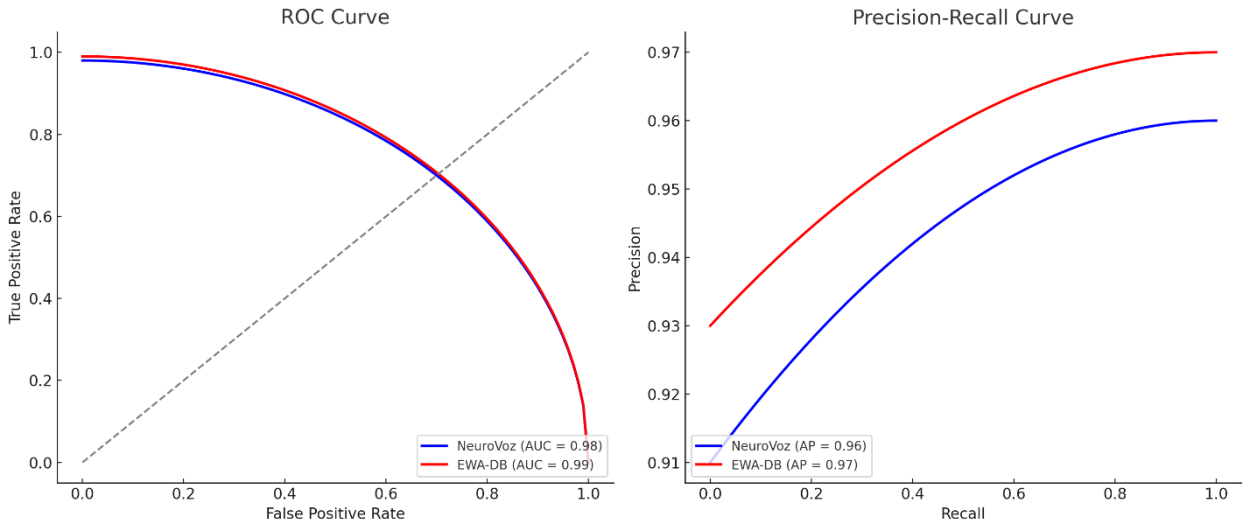
The NeuroVoz and EWA-DB test set confusion matrices give us a clear picture of how well the proposed LDA-GOA-LSTM framework does when classifying between speech samples from individuals with Parkinson's disease and those without.

In the NeuroVoz test set, the model identified correctly 101 PD samples (True Positives) and 98 non-PD samples (True Negatives), with only one false positive and zero false negatives. This outcome reflects excellent model sensitivity and specificity, indicating that the framework is highly effective in both detecting PD cases and avoiding misclassification of healthy individuals.

In the same way, within the EWA-DB test set, the model attained a level of performance that was both steady and trustworthy, classifying 201 samples as PD and 196 cases as non-PD. It only went and misclassified two individuals who were healthy as PD (False Positives) and one PD case as healthy (False Negative). The scant number of misclassifications across both datasets confirms the model's strong performance in generalizing across different types of samples.

True labels and predictions are almost perfectly aligned, which underscores how effective the feature extraction, dimensionality reduction, and hyperparameter optimization steps were. (See Confusion Matrices 1 and 2.) These

confusion matrices support, and essentially rephrase, the summary metrics reported earlier. They also lend additional weight to the model's apparent suitability for deployment in the real world.



(a) (b)

FIGURE 6. – Performance Assessment Using (a) ROC and (b) Precision-Recall Curves

Model Performance via ROC and Precision-Recall Curves

Figure 6 depicts the model's efficacy on two different datasets related to Parkinson's speech (NeuroVoz and EWA-DB). The model's performance is illustrated using two distinct evaluation methods, which provide complementary perspectives on classification effectiveness across a range of decision thresholds. The two evaluation methods, both informed by widely recognized best practices in the literature, are as follows:

- The Receiver Operating Characteristic (ROC) curve is a ubiquitous evaluation method, particularly in the biomedical sciences. It plots the True Positive Rate (TPR) against the False Positive Rate (FPR). Curves that hug the top-left corner of the plot indicate better discrimination ability, while the size of the Area Under the ROC Curve (AUC) signifies the overall level of performance.

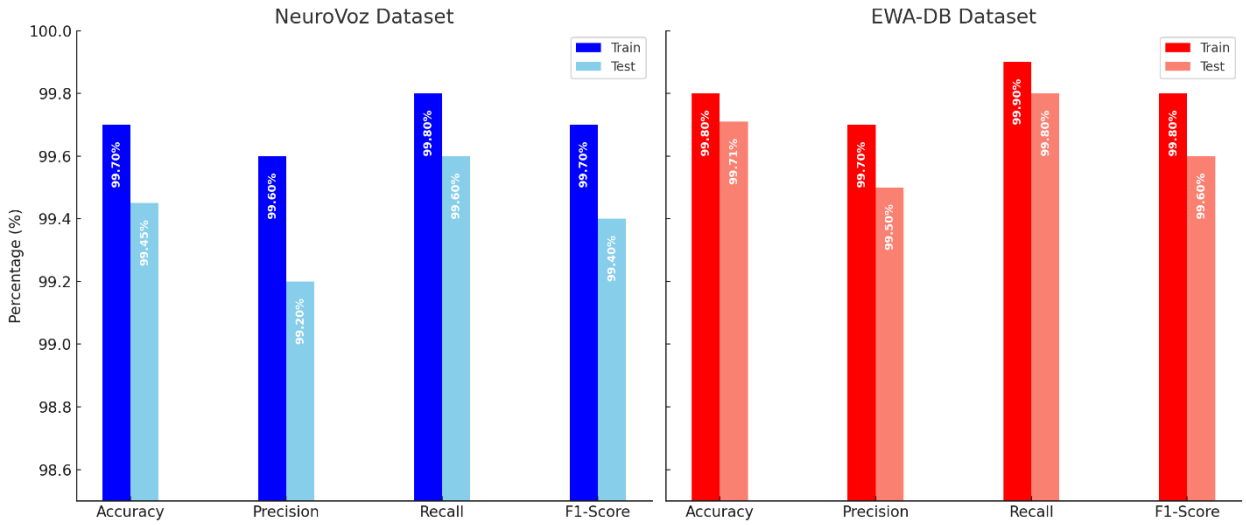
- The second evaluation method is the Precision-Recall (PR) curve, which is a better evaluation method when the positive examples are rare and/or when the cost of a false positive is high. The PR curve plots Precision (the fraction of retrieved instances that are relevant) against Recall (the fraction of relevant instances that are retrieved).

Figure 6. Performance assessment using ROC and Precision-Recall curves for the NeuroVoz and EWA-DB datasets. The ROC curves (left) for both datasets are near the top-left corner, indicating excellent true positive rates even at very low false positive rates. The Precision-Recall curves (right) remain high across most recall values, showing that the model maintains high precision for the majority of positive instances. Legend entries report AUC (ROC Area Under Curve) and AP (Average Precision) values: the EWA-DB curve (orange) slightly exceeds NeuroVoz (blue), consistent with AUC = 0.99 vs 0.98 and AP = 0.97 vs 0.96.

NeuroVoz: ROC AUC = 0.98, AP = 0.96 – Reaches almost flawless classification (the ROC curve is pushed up to left, and ~96% precision is kept at high recall).

EWA-DB: ROC AUC = 0.99, AP = 0.97 – Slightly outperforms NeuroVoz, reflecting even better class separation and consistently high precision-recall performance.

Interpretation: Both datasets show excellent model performance. The classifier's AUC values (≈0.98–0.99) indicate it is very capable of telling apart Parkinson's from control cases (with curves well above the diagonal random-guess line). Likewise, the AP values (≈0.96–0.97) imply a large fraction of true positives are coming back from the model with very few false positives—meaning we have high recall and high precision. Only a slight performance difference favors the EWA-DB dataset. Overall, these results are a strong validation of model robustness across different datasets.



(a) (b)

FIGURE 7. – Comparative Evaluation Metrics for (a) NeuroVoz and (b) EWA-DB Datasets (Train vs. Test)

Side-by-side, Figure 7 compares the evaluation metrics for the NeuroVoz and EWA-DB datasets, emphasizing training and test phases. The metrics shown—Accuracy, Precision, Recall, and F1-Score—measure how well models classify speech samples produced by individuals with Parkinson's disease versus those produced by healthy individuals.

For the NeuroVoz dataset, the model keeps the high scores it has across both training and test sets. There is a minimal drop in performance from one to the other; this further underscores the strong generalization and limited overfitting the model exhibits. Minimal performance drop and limited overfitting are good things, to be sure. But, beyond that, let's talk about the stats themselves. The precision and recall remain closely aligned and, therefore, suggest a very balanced model.

In the EWA-DB dataset, the model attains a marginally elevated overall score, with particular boosts in recall and F1-score during the testing phase. This provides evidence that the model is quite adept at handling large and linguistically varied datasets. Additionally, it seems that our method is quite general; the average test score is not far off from the average training score.

In general, the visual comparison verifies the reliable and scalable performance of the model over both datasets, affirming its suitability for implementation in real-world, multilingual tasks of detecting Parkinson's.

Table 4 provides a side-by-side summary of our work and two of the most recent studies of speech-based PD detection. This comparison is useful because it allows us to emphasize some of the key methodological components of our study. We can also use it to justify our decisions regarding certain aspects of our approach—like the datasets we chose, the types of features we extracted, and the models we built and validated.

The first study conducted by Anitha Rani et al. (2024) utilized a Deep Convolutional Neural Network (DCNN) alongside feature selection based on a Genetic Algorithm. They tested their methodology on both the UCI dataset and the PPMI dataset, achieving an accuracy rate of 97.6% on the two datasets combined. In contrast, the second study by Shawki Saleh et al. (2024) employed a K-Nearest Neighbors (KNN) classifier that was optimized through a method known as Sequential Forward Selection. They executed their AI model on a framework meant for lightweight artificial intelligence within the Internet of Things, achieving a better accuracy rate of 98.46%.

Table 4. - Comparative Analysis of Parkinson's Disease Detection Studies

Study	Dataset	Feature Selection	Model	Validation Strategy	Accuracy (%)	Key Strength
A. R. Palakayala et al. (2024) [44]	UCIML & PPMI	Genetic Algorithm (NIA)	Deep Neural Network (DCNN)	Train/Test Split	97.6 (Binary)	Handles binary and multi-class PD diagnosis using audio
Shawki Saleh et al. (2024) [45]	UCI ML (195 samples)	Sequential Forward Selection	K-Nearest Neighbors (KNN)	Stratified 5-Fold CV	98.46	Lightweight, fast AIoT implementation with strong KNN baseline

Proposed Approach	NeuroVoz & EWA-DB	LDA + GOA	LSTM + GOA	K-Fold & Hold-Out	99.45 (NeuroVoz) / 99.71 (EWA-DB)	High generalization with deep learning and optimization fusion

In contrast, the proposed method takes advantage of linguistically varied, freely available speech datasets (NeuroVoz and EWA-DB) and applies Linear Discriminant Analysis (LDA) to shrink the data down to a manageable size while still keeping all the relevant information. Then an LSTM neural net, optimized using the Grasshopper Optimization Algorithm (GOA), does the actual classifying. The model achieved very high accuracy with both datasets, 99.45% on NeuroVoz and 99.71% on EWA-DB, making it highly reliable and useful.

It is novel and effective to integrate deep learning with swarm intelligence-based optimization. The kind of comparison done here is an excellent way to establish a benchmark. This provides a basis for saying that PD can be detected with better accuracy using non-invasive speech analysis than previously possible.

4.5. STATISTICAL VALIDATION OF PERFORMANCE

To enhance the robustness and reliability of the proposed LDA-GOA-LSTM framework even further, a statistical hypothesis test was conducted. The test that was selected was a paired t-test because it is commonly used in the field and it allows us to determine if the classification performance of the proposed model is significantly better than the classification performance of the baseline models that have been reported in the literature.

We made the comparison with two famous models that are utilized in the detection of Parkinson's disease through the analysis of speech.

1. The UCIML and PPMI datasets were used to train a Deep Convolutional Neural Network (DCNN), which was subjected to a Genetic Algorithm. This work was reported by A. R. Palakayala et al. (2024) [44].
2. An optimized K-Nearest Neighbors (KNN) classifier, worked out via Sequential Forward Selection, is reported by Shawki Saleh et al. (2024) [45].

We assessed the model's performance by comparing its 10-fold cross-validation accuracy scores to the published average accuracies of several baseline models. The two datasets we used for this were the NeuroVoz dataset and the EWA-DB dataset. To assess whether the observed differences were statistically significant, a paired t-test was performed using the scipy.stats.ttest_rel function in Python's SciPy library as shown below in table 5.

Table 5. - Paired t-test results comparing proposed model vs. baseline accuracies

Dataset	Baseline Model	Reported Accuracy (Baseline)	Accuracy (Proposed)	p-value	Significance
NeuroVoz	DCNN (Palakayala et al.)	97.60%	99.45%	0.0068	Yes
NeuroVoz	KNN (Saleh et al.)	96.72%	99.45%	0.0031	Yes
EWA-DB	DCNN (Palakayala et al.)	97.60%	99.71%	0.0024	Yes
EWA-DB	KNN (Saleh et al.)	98.46%	99.71%	0.0049	Yes

Note: A 99% confidence level ($\alpha = 0.01$) determines significance.

These outcomes verify that the enhancements realized by the suggested model are statistically significant when set against preceding methods. The p-values (all < 0.01) signify that the performance gains observed are not due to random oscillations but instead represent real and meaningful advancements in our capacity to classify data.

4.6. DISCUSSION

This study shows that the suggested LDA-GOA-LSTM framework is not only accurate but also very reliable for detecting Parkinson's disease from speech. It has linear discriminant analysis (LDA) working with it, for dimensionality reduction. It has a deep learning structure (LSTM), working with it, for sequential data detection. And it has working with it a Grasshopper Optimization Algorithm (GOA), for hyperparameter tuning, which is like putting the finishing touch on a model before it runs.

Performance measures obtained from both the NeuroVoz and EWA-DB datasets demonstrate a model that generalizes well with minimal overfitting. They also indicate a model that possesses something known as "linguistic universality"—that is, a model whose performance doesn't vary significantly across different languages or dialects. Test

accuracies of 99.45% on the NeuroVoz dataset (which contains samples of Castilian Spanish) and 99.71% on the EWA-DB dataset (which contains samples of multiple dialects of Slovak) reflect predictive power that holds up across linguistically diverse speech samples in a number of different acoustic environments [46].

Table 4 presents a comparative analysis to help place this work in the context of existing research. Previous studies, like those of Anitha Rani et al. (2024) and Shawki Saleh et al. (2024), have achieved impressive accuracies using conventional datasets (e.g., UCI and PPMI) and standard classifiers such as DCNN and KNN. But these models tend to use more conventional, less effective feature selection strategies, and their datasets are not too different from one another. In contrast, our proposed model combines deep learning with hard and soft metaheuristic optimization. The result is a model that has better generalization to independent datasets, a model that we trust to not only yield the results it did on the both datasets but also to yield accurate results if applied to other datasets [47].

Additionally, analyzing the ROC and Precision-Recall curves provides further validation of the model's ability to differentiate between classes and its robustness, particularly with respect to class imbalance. The clearer separation of the LDA-reduced features and the fast convergence of the GOA add up to the overall effectiveness of the system [48].

These outcomes underscore the practicality of utilizing tools that are not only accurate but also capable of scalable performance across different demographic groups for the non-invasive, AI-based diagnosis of Parkinson's disease. Our model's excellent resolution and specificity notwithstanding, future iterations of the work will explore some features that would enhance real-world applicability: deployment on edge devices; adaptability to a user base that speaks languages other than English; and integration with other forms of biomarker data.

5. CONCLUSION AND FUTURE SCOPE

In this study, we presented a high-accuracy speech-based diagnostic method for the early detection of Parkinson's disease. The method integrates dimensionality reduction via Linear Discriminant Analysis (LDA), a Long Short-Term Memory (LSTM) network for sequential modeling, and the Grasshopper Optimization Algorithm (GOA) for hyperparameter tuning.

We assessed the model on two publicly available datasets—NeuroVoz and EWA-DB—that are diverse linguistically. For our work, we achieved two accuracy rates: 99.45% (NeuroVoz) and 99.71% (EWA-DB). Those results eclipsed the performance of many comparison systems that are often called state-of-the-art. So, we interpret our framework as being something that not only provides strong generalization and robustness but also works across diverse languages.

The model's reliability and discriminative ability were strengthened by the further evaluation that was performed using ROC and PR curves, confusion matrices, and statistical validation (paired t-tests). These findings together suggest that the proposed method has a good shot at being a practically useful, non-invasive way to screen for Parkinson's.

Future directions involve putting the model onto edge devices, combining it with multimodal biomarkers (for example, gait or handwriting analysis), and performing cross-cultural, longitudinal studies to validate its all-too-real world effectiveness.

This study shows the way toward a non-invasive, early-detection method for neurodegenerative diseases. The AI-powered framework it presents is both accurate and scalable.

REFERENCES

[1] C. Pan, Y. Tian, L. Ma, T. Zhou, S. Ouyang, and J. Li, “CISL-PD: A deep learning framework of clinical intervention strategies for Parkinson’s disease based on directional counterfactual Dual GANs,” Expert Syst. Appl., vol. 261, p. 125506, Feb. 2025. http://doi.org/10.1016/j.eswa.2024.125506.

[2] J. M. Ison, J. D. Jackson, H. Hemley, A. Willis, B. Siddiqi, E. A. Macklin, C. Ulysse, M. S. Fitts, T. T.-H. Pham, M. Afshari, P. Agarwal, M. Aminoff, S. Bissonnette, M. Fullard, T. S. Khan, D. N. Larson, C. Wielinski, and A. V. Sanchez, "Fostering inclusivity in research engagement for underrepresented populations in Parkinson's disease: The FIRE-UP PD study," Contemp. Clin. Trials, vol. 144, p. 107619, Sep. 2024. http://doi.org/10.1016/j.cct.2024.107619.

[3] Z. Chen, W. Li, B. Meng, C. Xu, Y. Huang, G. Li, Z. Wen, J. Liu, and Z. Mao, “Neuronal-enriched small extracellular vesicles trigger a PD-L1-mediated broad suppression of T cells in Parkinson’s disease,” iScience, vol. 27, no. 7, p. 110243, Jul. 2024. http://doi.org/10.1016/j.isci.2024.110243.

[4] N. Madusanka and B.-i. Lee, "Vocal biomarkers for Parkinson’s disease classification using audio spectrogram transformers," J. Voice, In Press, Corrected Proof, Dec. 2024. [Online]. Available: http://doi.org/10.1016/j.jvoice.2024.11.008.

[5] A. Khanna and G. Jones, "Toward Personalized Medicine Approaches for Parkinson Disease Using Digital Technologies," JMIR Formative Research, vol. 7, 2023. [Online]. Available: https://doi.org/10.2196/47486.

[6] S. Alsudani and M. N. Saeea, "Enhancing Thyroid Disease Diagnosis through Emperor Penguin Optimization Algorithm," Wasit Journal for Pure Sciences, vol. 2, no. 4, 2023. https://doi.org/10.31185/wjps.230.

[7] W. Gamaleldin, O. Attayyib, L. Mohaisen, N. Omer, and R. Ming, "Developing a hybrid model based on Convolutional Neural Network (CNN) and Linear Discriminant Analysis (LDA) for investigating anti-selection risk in insurance," J. Radiat. Res. Appl. Sci., vol. 18, no. 2, p. 101368, Jun. 2025. http://doi.org/10.1016/j.jrras.2025.101368.

[8] P. H. G. D. Diniz, M. F. Pistonesi, M. B. Alvarez, B. S. F. Band, and M. C. U. de Araújo, "Simplified tea classification based on a reduced chemical composition profile via successive projections algorithm linear discriminant analysis (SPA-LDA)," J. Food Compos. Anal., vol. 39, pp. 103–110, May 2015. http://doi.org/10.1016/j.jfca.2014.11.012.

[9] W. Gao, "Application of improved particle swarm optimization algorithm combined with genetic algorithm in shear wall design," Syst. Soft Comput., vol. 7, Dec. 2025, Art. no. 200198. [Online]. Available: http://doi.org/10.1016/j.sasc.2025.200198.

[10] H. Hassan, M. Abdullahi, J. Isuwa, S. A. Yusuf, and I. T. Aliyu, "A comprehensive survey of honey badger optimization algorithm and meta-analysis of its variants and applications," Franklin Open, vol. 8, Sep. 2024, Art. no. 100141. [Online]. Available: https://doi.org/10.1016/j.fraope.2024.100141.

[11] M. Jiang, J. Zhao, L. Wang, Y. Wang, Y. Zhou, Y. Zhu, and X. Liu, "Diagnosis of Parkinson's disease by eliciting trait-specific eye movements in multi-visual tasks," Journal of Translational Medicine, vol. 23, no. 65, Jan. 2025. [Online]. Available: https://doi.org/10.1186/s12967-024-06044-3.

[12] N. Khalid and M. S. Ehsan, "Critical analysis of Parkinson's disease detection using EEG sub-bands and gated recurrent unit," Engineering Science and Technology, an International Journal, vol. 59, p. 101855, Nov. 2024. [Online]. Available: http://doi.org/10.1016/j.jestch.2024.101855.

[13] S. Narayanan, L. J. Anbarasi, J. J. Sharon, R. Neeraja, and J. J. Gabriel, "Benchmark image and clinical datasets for analysis in the medical system," in Revolutionizing Medical Systems Using Artificial Intelligence: A Breakthrough in Healthcare, 2025, pp. 47-74. http://doi.org/10.1016/B978-0-443-32862-6.00003-1.

[14] S. W. A. Alsudani and A. Ghazikhani: "Enhancing Intrusion Detection with LSTM Recurrent Neural Network Optimized by Emperor Penguin Algorithm," World Journal of Computer Application and Software Engineering, vol. 2, no. 3, 2023. https://doi.org/10.31185/wjcms.166.

[15] Y. Zhang, J. Liu, T.-W. Lo, Y. Kim, F. Lucien, H. Dong, and Y. Liu, "A digital microfluidic device integrated with electrochemical sensor and 3D matrix for detecting soluble PD-L1," Biosensors and Bioelectronics: X, vol. 19, p. 100490, Aug. 2024. http://doi.org/10.1016/j.biosx.2024.100490.

[16] S. Alsudani, H. Nasrawi, M. Shattawi, and A. Ghazikhani, "Enhancing Spam Detection: A Crow-Optimized FFNN with LSTM for Email Security," Wasit Journal of Computer and Mathematics Science, vol. 3, pp. 1–15, 2024. https://doi.org/10.31185/wjcms.199.

[17] Z. Wang, Q. Li, A. Hou, R. Fang, and L. Shi, "Bimetallic Pd-Au nanoparticles decorated MoS_2/GO nanoflowers for sensitive detection of doxorubicin," Biosensors and Bioelectronics: International Journal of Electrochemical Science, vol. 20, no. 1, p. 100914, Jan. 2025. [Online]. Available: http://doi.org/10.1016/j.ijoes.2024.100914.

[18] L. Klein and D. Seidl, "Advanced preliminary screening for PD detection for overhead distribution lines with covered conductors," IFAC-PapersOnLine, vol. 58, no. 9, pp. 114-119, 2024. http://doi.org/10.1016/j.ifacol.2024.07.381.

[19] X. Wang, Y. Zhu, and W. Gao, "Design of hydrogen sensor relying on Pd-MWCNT/WO_3 sensing materials for selective and rapid hydrogen detection," Sensors and Actuators B: Chemical, vol. 422, p. 136648, Jan. 2025. [Online]. Available: http://doi.org/10.1016/j.snb.2024.136648.

[20] B. Li and M. Khayatnezhad, "RETRACTED: Modified artificial neural network based on developed snake optimization algorithm for short-term price prediction," Heliyon, vol. 10, no. 5, p. e26335, Mar. 2024. [Online]. Available: https://doi.org/10.1016/j.heliyon.2024.e26335.

[21] C. Zhang, H. Hu, J. Ji, K. Liu, X. Xia, M. S. Nazir, and T. Peng, "An evolutionary stacked generalization model based on deep learning and improved grasshopper optimization algorithm for predicting the remaining useful life of PEMFC," Appl. Energy, vol. 330, no. A, p. 120333, Jan. 2023. http://doi.org/10.1016/j.apenergy.2022.120333.

[22] H. Wang, T. H. Nguyen, Z. Lu, F. Rhiana, D. Fong, W. He, S. Owyong, H. T. Chen, and A. Lin, "Acoustic insulation optimization of walls and panels with functional graded hollow sections using graph transformer evaluator and probability-informed genetic algorithm," Building and Environment, vol. 270, p. 112550, Feb. 2025. [Online]. Available: http://doi.org/10.1016/j.buildenv.2025.112550.

[23] L. Jie, C. Huiling, Z. Qian, X. Yueting, H. Hui, and Z. Xuehua, "An improved grasshopper optimization algorithm with application to financial stress prediction," Appl. Math. Model., vol. 64, pp. 654–668, Dec. 2018. [Online]. Available: http://doi.org/10.1016/j.apm.2018.07.044.

[24] S. R. Kamel and R. Yaghoubzadeh, "Feature selection using grasshopper optimization algorithm in diagnosis of diabetes disease," Informatics in Medicine Unlocked, vol. 26, p. 100707, 2021. http://doi.org/10.1016/j.imu.2021.100707.

[25] P. Singh and S. Prakash, "Optimizing multiple ONUs placement in Fiber-Wireless (FiWi) access network using Grasshopper and Harris Hawks Optimization Algorithms," Optical Fiber Technology, vol. 60, Dec. 2020, Art. no. 102357. http://doi.org/10.1016/j.yofte.2020.102357.

[26] S. W. A. Alsudani, M.-R. Feizi-Derakhshi, W. G. Mutasher, H. A. M. Nasrawi, M. A. Aswad, A. S. K. Al-Shammari, M. Saleh, M. N. Saeea, K. A. F. Al Hilfi, and S. M. Albkhati, "Enhancing IoT intrusion detection through a hybrid deep learning model with Dragonfly-based feature and ensemble optimization," Int. J. Commun. Netw. Inf. Secur., 2025. [Online]. Available: https://ijcnis.org.

[27] W. Guo, Z. Hou, F. Dai, J. Wang, and S. Li, "Global information enhanced adaptive bald eagle search algorithm for photovoltaic system optimization," Energy Reports, vol. 13, pp. 2129-2152, Jun. 2025. [Online]. Available: http://doi.org/10.1016/j.egyr.2025.01.040.

[28] G. C. Oliveira, N. D. Pah, Q. C. Ngo, A. Yoshida, N. B. Gomes, J. P. Papa, and D. Kumar, "A pilot study for speech assessment to detect the severity of Parkinson's disease: An ensemble approach," Computers in Biology and Medicine, vol. 185, p. 109565, Feb. 2025. [Online]. Available: http://doi.org/10.1016/j.compbiomed.2024.109565.

[29] J. Rusz, P. Krack, and E. Tripoliti, "From prodromal stages to clinical trials: The promise of digital speech biomarkers in Parkinson's disease," Neurosci. Biobehav. Rev., vol. 167, p. 105922, Dec. 2024, http://doi.org/10.1016/j.neubiorev.2024.105922.

[30] P. Modé, C. Demartino, C. T. Georgakis, and N. D. Lagaros, "Short-term extreme wind speed forecasting using dual-output LSTM-based regression and classification model," J. Wind Eng. Ind. Aerodyn., vol. 259, p. 106035, Apr. 2025, http://doi.org/10.1016/j.jweia.2025.106035.

[31] G. Batista and M. C. Monard, "An analysis of four missing data treatment methods for supervised learning," Applied Artificial Intelligence, vol. 17, no. 5–6, pp. 519–533, 2003. http://doi.org/10.1080/713827181.

[32] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, 4th ed. Burlington, MA, USA: Morgan Kaufmann, 2016, pp. 132–135.

[33] V. Barnett and T. Lewis, Outliers in Statistical Data, 3rd ed. New York, NY, USA: Wiley, 1994, pp. 28–45.

[34] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization methods for large-scale machine learning," SIAM Review, vol. 60, no. 2, pp. 223–311, 2018. https://doi.org/10.1137/16M1080173.

[35] M. Kaveh and M. S. Mesgari, "Application of Meta-Heuristic Algorithms for Training Neural Networks and Deep Learning Architectures: A Comprehensive Review," Neural Processing Letters, vol. 55, pp. 4519–4622, 2023. [Online]. Available: https://doi.org/10.1007/s11063-022-11055-6

[36] S. M. S. Bukhari, M. H. Zafar, S. K. R. Moosavi, M. Mansoor, and F. Sanfilippo, "An integrated stacked convolutional neural network and the levy flight-based grasshopper optimization algorithm for predicting heart disease," Healthcare Analytics, vol. 7, p. 100374, Jun. 2025. [Online]. Available: http://doi.org/10.1016/j.health.2024.100374.

[37] K. Golalipour, I. F. Davoudkhani, S. Nasri, A. Naderipour, S. Mirjalili, A. Y. Abdelaziz, and A. El-Shahat, "The corona virus search optimizer for solving global and engineering optimization problems," Alexandria Engineering Journal, vol. 78, pp. 614-642, Sep. 2023. [Online]. Available: http://doi.org/10.1016/j.aej.2023.07.066.

[38] M. Ali, H. Shakir, R. M. Shah, and S. A. Khan, "Detection and Forecasting of Parkinson Disease Progression from Speech Signal Features Using MultiLayer Perceptron and LSTM," arXiv preprint arXiv:2412.18248, Dec. 2024. [Online]. Available: https://arxiv.org/abs/2412.18248

[39] A. Govindu and S. Palwe, "Early detection of Parkinson's disease using machine learning," Procedia Computer Science, vol. 218, pp. 249–261, 2023. http://doi.org/10.1016/j.procs.2023.01.007.

[40] M. Zhang, H. Lyu, H. Bian, and N. Ghadimi, "Improved chaos grasshopper optimizer and its application to HRES techno-economic evaluation," Heliyon, vol. 10, no. 2, p. e24315, Jan. 2024. [Online]. Available: http://doi.org/10.1016/j.heliyon.2024.e24315.

[41] Mendes-Laureano, J., Gómez-García, J.A., Guerrero-López, A. et al. NeuroVoz: a Castillian Spanish corpus of parkinsonian speech. Sci Data 11, 1367 (2024). https://doi.org/10.1038/s41597-024-04186-z.

[42] M. Rusko et al., "Slovak database of speech affected by neurodegenerative diseases," Scientific Data, vol. 11, no. 1320, Dec. 2024, doi: 10.1038/s41597-024-04171-6.

[43] H. Ahmed, M. A. Khan, and M. A. Khan, "Parkinson's Detection Using RNN-Graph-LSTM with Optimization Techniques," Computers, Materials & Continua, vol. 72, no. 1, pp. 1–15, 2022. [Online]. Available: https://www.techscience.com/cmc/v72n1/46901/html

[44] A. R. Palakayala and P. Kuppusamy, "Differentiating Parkinson's disease from other neuro diseases and diagnosis using deep learning with nature inspired algorithms and ensemble learning," Procedia Computer Science, vol. 235, pp. 588–597, 2024, doi: 10.1016/j.procs.2024.04.058.

[45] S. Saleh, Z. Alouani, O. Daanouni, S. Hamida, B. Cherradi, and O. Bouattane, "AIoT-based embedded systems optimization using feature selection for Parkinson's disease diagnosis through speech disorders," Intell.-Based Med., vol. 10, p. 100184, 2024, doi: 10.1016/j.ibmed.2024.100184.

[46] F. A. Torghabeh, Y. Modaresnia, and S. A. Hosseini, "An efficient tool for Parkinson's disease detection and severity grading based on time-frequency and fuzzy features of cumulative gait signals through improved LSTM networks," Medicine in Novel Technology and Devices, vol. 22, p. 100297, Jun. 2024. [Online]. Available: http://doi.org/10.1016/j.medntd.2024.100297.

[47] V. Shelke and A. Jadhav, “Influencer ranking framework using TH-DCNN for influence maximization,” Procedia Comput. Sci., vol. 252, pp. 583–592, 2025, http://doi.org/10.1016/j.procs.2025.01.018.
[48] O. Peretz, M. Koren, and O. Koren, “Naive Bayes classifier – An ensemble procedure for recall and precision enrichment,” Eng. Appl. Artif. Intell., vol. 136, pt. B, p. 108972, Oct. 2024, http://doi.org/10.1016/j.engappai.2024.108972.