A Multiple Feature Selection based Parkinson's Disease Diagnostic System Using Deep Learning Neural Network Classifier

Dr.N. Suresh Kumar[1*], Dr.M. Senthamil Selvi[2], D. Gayathri[3]

Abstract

Parkinson's disease is one among the most common neurological disorder that is increasing day by day. Parkinson's disease could become more dangerous in not treated well in time. Hence, early and automated detection of Parkinson's disease is greatly essential. In current research, five classifiers such as DT, NB, NN, RF and SVM are exploited for PD detection. Nonetheless, more time is greatly necessitated for using further classifiers for PD detection. Enhanced Convolutional Neural Network (ECNN) method is greatly suggested for mitigating this issue in PD detection. Primarily, input PD dataset processing is done as preprocessing steps for validating dataset quality to accomplish the process. On the basis of Multiple Feature Evaluation Approach (MFEA), feature assessment processing is done which comprises quite a lot of feature assessment and ranking algorithms to weight features worth and features set extraction. Furthermore, a feature selection model by principal component analysis is suggested for classifier accuracy improvement. To conclude, Enhanced Convolutional Neural Network (ECNN) training is done over these certain features for PD detection, besides its performance evaluation is done through various metrics.

Key Words: Neurological Disorder, Brain Nerve Cells, Principal Component Analysis, Feature Subset Selection, and Deep Neural Network.

DOI Number: 10.14704/nq.2021.19.7.NQ21104 **NeuroQuantology 2021; 19(7):209-220**

Introduction

Parkinson's disease (PD) is regarded as enduring degenerative syndrome in which motor system is usually affected and various symptoms of this disease are shaking, trouble with drinking, talking, thinking and walking. Dopamine, a chemical existing in brain nerve cells helps in sending signals to the brain region controlling muscles movement and this deficiency leads to this disease. There is estimation that about 7 to 10 million people are affected due to this disease. In some cases, mortality might happen due to this disease and symptoms may be worsen as time progress besides more server illness.

Consequently, initial recognition and treatment is greatly necessitated for PD patients. Unfortunately, not even a single perfect diagnostic test does exist presently and PD detection intensely relies on information that patient affords (Jain & Shetty 2016); (Grover et al., 2018). As a result, developing an automatic system for PD disorder detection is greatly required. The PD affected people speech may comprise certain subtle irregularities that cannot be perceived through listeners nevertheless can be useful for PD detection via using acoustical examination. In recent times, several techniques have been developed for speech analysis and people classification suffering from PD along with healthy people.

Corresponding author: Dr.N. Suresh Kumar
Address: [1*]Assistant Professor (Sl.Gr), Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India; [2]Professor, Department of IT, Sri Ramakrishna Engineering College, Coimbatore, India; [3]Assistant Professor, Department of Electronics and Communication Engineering, PSNA College of Engineering and Technology, Dindigul
E-mail: [1*]nsuresh2@gmail.com; [2]senthamilselvi@srec.ac.in; [3]deivagayathri@gmail.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Received: 29 March 2021 **Accepted:** 30 June 2021

The people speech recording is being done besides certain explicit features extraction is one for discriminating people with PD from healthy subjects using diverse techniques in this research. Recently, new features were suggested while others have utilized diverse feature selection approaches for eliminating replicated features as well as feature vector size reduction4 is done for classification purpose (Aich et al., 2017); (Chakraborty et al., 2020); (Aich et al., 2017).
Various inspired algorithms relative research representation is attained for different evolutional techniques performance assessment in optimal features selection necessitated for PD analysis. Several machine learning and classification approaches examining is done for analysis purpose. Sadikov et al., (2015) utilized spirography features for motor function assessment in Parkinson's disease. In (Devarajan & Ravi 2019); (Castaño-Pino et al., 2020) sleep Behavior Disorder (RBD) non-motor features and Cerebrospinal fluid (CSF) measurements are exploited for early PD detection and classification using five classifiers such as DT, NB, NN, RF and SVM.
Over conventional techniques, for PWP speech, potentially superior classification performance is offered by Deep neural networks (DNNs). In DNNs, auto encoders (AEs) are used for minimizing feature dimension and softmax layer is used for sample classification. But this is not the case on conventional techniques. In different medical applications, DNNs are used successfully (Badem et al., 2016). For classification problems, they are mostly preferred due to the recent enhancements in deep neural networks field (Badem et al., 2016). For complex classification problems which are not resolvable effectively using other classification methods, deep neural network's application has given a new area.
Typically, deep learning methods are divided into four classes: CNN-based methods, restricted Boltzmann machines (RBMs), auto encoders, and sparse coding. Recently, the CNN-based methods have attracted more and more attention around the world, which have achieved promising results in literature. By considering this the Enhanced Convolutional Neural Network (ECNN) is used in this research.
The major contribution of the work in which the PD detection using enormous classifier is considered as time consuming process which is mitigated by presenting Enhanced Convolutional Neural Network (ECNN) in this research. Primarily, input PD dataset processing is done as preprocessing steps for validating dataset quality to accomplish the process. On the basis of Multiple Feature Evaluation Approach (MFEA), feature assessment processing is done which comprises quite a lot of feature assessment and ranking algorithms for weighing features worth and features set extraction. Furthermore, a feature selection model by principal component analysis is suggested for classifier accuracy improvement. At last, Enhanced Convolutional Neural Network (ECNN) training is done over these selected features for PD detection.

Related Works

The reviews about different PD detection methods are given in detail.
Fayyazifar & Samadiani (2017) suggested an approach for prompt PD detection using Genetic algorithm for optimal feature set selection by which feature vector dimension reduction is achieved from 22 to 6 features. Furthermore, detection rate of 96.55% using AdaBoost and 98.28% using Bagging algorithms is attained for classification process, respectively.
Ali et al., (2019) suggested two-dimensional data selection technique for selecting features and samples. Here, feature ranking is done through chi-square statistical model, explores optimal ranked features subset and iteratively chooses samples. Experimental outcomes reveal that, on multiple voice data types, suggested technique outclasses traditional methods pertaining to PD detection accuracy.
Upadhya et al., (2018) utilized Thomson Multitaper (TMT) windowing and Single Taper Smooth (STS) windowing technique by suggesting MFCC and PLP voice in combination with neural network classifier. By which, healthy people classification from early stage Parkinson diseased patients is achieved. Both technique's performance evaluation described. Short time speech spectra weighted average helps in achieving variance reduced power spectrum which might be found through TMT windows set. PLP and MFCC features computation is attained by this spectrum. The outcomes reveal that, in contrast with STS window, a supreme enhancement (around 6.6%) in classification accuracy is achieved for nine tapers, adaptive weights by MFCC as features.
Li et al., (2017) utilized all speech segments features of every subject organized for suggesting hybrid feature learning algorithm. In this manner, innovative and highly efficient features deprived of

feature transformation are attained. Initially, hybrid features construction is done via segments and features merging followed by high proficient hybrid feature selection were directed by several parameters. Then, selected hybrid features are used in PD classification. Different assessment criteria are presented for feature selection process. Experimental outcomes infer that this suggested algorithm attains new features (hybrid feature) with reasonable classification accuracy and very steady as well as expressive features are selected.

Sharma et al., (2019) suggested an innovative model entitled Modified Grey Wolf Optimization (MGWO) on basis of modern Grey Wolf Optimizer (GWO) which functions as a search stratagem for feature selection. GWO is regarded as meta-heuristic algorithm inspired through wolves hunt down behavior. k-nearest neighbor, random forest classifier, decision tree are greatly utilized for feature selection. The suggested system evaluation is achieved through different voice datasets types, handwriting (meander and spiral) and speech. Parkinson disease prediction has attained 94.83% accuracy, 98.28% detection rate, 16.03% false alarm rate and helps individuals receiving functional treatment at an initial stage. Suggested bio-inspired algorithm is steady adequate for determining optimal features subset. Finally, outcomes obtained from suggested algorithm assessment on datasets associated with Optimized Cuttlefish Algorithm (OCFA) outcomes are utilized for comparison. Thereby suggested algorithm has been validated by experimental outcomes in terms of accuracy and features chosen.

Nilashi et al., (2018) designed a system UPDRS prediction by utilizing Incremental support vector machine for Motor-UPDRS and Total-UPDRS prediction. For minimizing dimensionality, non-linear iterative partial least squares are deployed and for clustering self-organizing map is used. Approach is assessed by conducting numerous experimentations with PD dataset. Outcomes are compared with existing research. Prediction accuracies estimated through MAE for Motor-UPDRS is 0.4967 and Total-UPDRS is 0.4656. Experimental outcomes are substantiated by suggested technique demonstration for UPDRS prediction. In healthcare, for PD prediction, this can also be implemented as an intelligent system.

Gupta et al., (2018) suggested an enhanced optimized version of crow search algorithm (OCSA). This methodology is greatly utilized in Parkinson's disease prediction with 100% accuracy which serves individual for proper treatment at primary stage. 20 benchmark datasets is exploited to assess OCSA performance and outcomes comparison is done with original chaotic crow search algorithm (CCSA). It also infers that this suggested algorithm can attain optimal features subset with maximum classification accuracy and selected features are very steady along with minimized number of features selected.

Machine learning approach based automatic Parkinson disease detection using person's voice/speech is proposed by Indira, & Can (2013). To make a discrimination among Parkinson disease affected people and healthy people, a technique based on pattern recognition and fuzzy C-means clustering is proposed. Around 45.83% accuracy, 75.34% sensitivity and 68.04% accuracy are achieved using this.

For predicting Parkinson's disease severity, a deep neural network is investigated by Grover et al., (2018). When compared with other available methods, better accuracy is achieved using proposed DNN model. Better performance is achieved using motor UPDRS score based classification when compared with total UPDRS score classification. Severity prediction is done in a better manner using this. A dataset with 5875 instances is used in experimentation. Using a larger dataset, proposed technique's accuracy is enhanced further.

Using a wrapper feature selection technique, five various classification paradigms are proposed by Armañanzas et al., (2013) which are able to predict every class variables with accuracy of 72-92%. In addition, dichotomy problems are formulated by converting three major classification classes namely severe, moderate and mild. In this, better performance is shown by binary classifier and various non-motor symptoms subsets are selected.

The PD patient's remote Tele-monitoring is done using k-Nearest Neighbour (k-NN) and SVM as proposed by Sakara & Kursunb (2014), where, at regular interval, PD patients voice recording are taken. Gender, age, voice recordings taken at baseline, three months interval and six months interval are used as features. In patients UPDRS score, in significant deterioration detection, better performance is shown by Support Vector Machine.

For discriminating PD and healthy controls, support vector machine, random forest and feature selection are proposed by Tsanas et al. (2012). Only ten dysphonia features are used for achieving a classification accuracy of around 99%.

A nonlinear signal approach is proposed in Tsanas et al., (2011). Large dataset which is formed using speech/voice records without requiring clinical physician presence is used. Various speech signal algorithms are applied. Using a classification algorithm and nonlinear regression, performed this paper and visibility of accurate, cost-effective, remote, frequent self-administered speech tests based UPDRS telemonitoring is supported by this.
A support vector machine, pattern recognition and artificial neural network are proposed by Sharma & Giri (2014). In Parkinson disease diagnosis, experts are assisted using these methods. Healthy people's biomedical voice signals are used for creating this dataset and around 85.294% accuracy is obtained for Parkinson disease detection.
A Multi-Layer Perceptron (MLP) with Back-Propagation learning algorithm is proposed by Khemphila & Boonjing (2012). Effective Parkinson's disease (PD) diagnosis is done using this algorithm. Doctor's experience and expertise are used in medical diagnosis. But, there exist a wrong treatment and diagnosis. For diagnosis, various tests needs to be taken by patients. In most cases, majority of these tests are not contributing to effective diagnosis. Patients diagnosis are sorted using artificial neural networks. With 91.45% accuracy, training dataset is predicted and with 80.77%, validation dataset is predicted using this technique.
For finding Parkinson disease patient's freezing, a Bayesian Belief Network (BBN) is proposed by Saad et al., (2013). An video dataset which us extracted from real Parkinson disease patients is used in experimentation and this dataset is available in online. A matrix is there in every file which has three sensors measurement data in x, y and z directions. Occurrence of Weather Freezing of Gait (FoG) is monitored. Data synchronization is used for labelling these annotations via video which records every patients run and appeared results while testing the models with Bayesian Naïve Classifier (BNC) classifier.
For Parkinson's disease diagnosis, to propose a decision support system, Artificial Neural Networks (ANN), Support Vector Machines (SVM), Classification and Regression Trees (CART), Regression and Bayesian Networks are proposed by Hadjahmadi & Askari (2012). Movement freezing or slowing are is generally produced by this Parkinson's disease disorder. Using regression tree and classification, around 93.7% accuracy is achieved by proposed system.

Table 1. Comparison of the Existing Approaches

Authors name	Methods	Dataset Description	Result performance
Indira & Can (2013)	fuzzy C-means	Speech signal	Accuracy- 68.04% sensitivity- 75.34%, specificity- 45.83%
Indira & Can (2013)	ANN	Speech signal	Recognition rate- 92 %
Grover et al., (2018)	deep neural network	Speech dataset as high or low	Accuracy of training dataset- 83.367% Accuracy of test dataset 81.6667% respectively
Armañanzas et al., (2013)	Wrapper feature selection	non-motor symptoms	Accuracy range- 72% to 92%
Sakara, Batalu & Kursunb (2014)	SVM	Gender, Age, voice recording	Accuracy - 76% Sensitivity- 34%
Tsanas et al., (2012)	SVM	Speech signal	Accuracy - 98.6%
Tsanas et al., (2011)	Regression & Classification	Speech signal	Percentile- 5–95
Sharma & Giri (2014)	SVM	Speech signal	Accuracy- 85.29%
Khemphila & Boonjing (2012)	ANN	Speech signal	Accuracy- 83.33%Accuracy
Saad et al., (2013)	Bayesian Naive Classifier (BNC)	Video recorded	Accuracy- 74.31%
Hadjahmadi & Askari (2012)	classification and regression tree	Recorded voice signals	Accuracy- 93.7%

From literature, over other feature vector classification based classifier approaches, better performance is shown by DNN classifier as said (Sharma et al., 2019); (Wu et al., 2018). For PD diagnosis, for addressing aforementioned classification problem. A DNN classifier is proposed in this paper. Using parents generated speech signal, PD is diagnosed accurately in this proposed DNN classifier. Using AE, features are learnt in proposed DNN and using softmax layer, a robust classifier is designed.

Enhanced Convolutional Neural Network (ECNN) Based Parkinson Diseases Detection

Enhanced Convolutional Neural Network (ECNN) Based Parkinson Diseases Detection is elaborated in this section. The suggested system comprises of four main stages, Parkinson's dataset preparation via preprocessing steps is the first step followed by feature extraction process using MFEA model. Subsequently, feature selection through principal component analysis for classifier accuracy improvement is the third step. Fourth one is Enhanced Convolutional Neural Network (ECNN) training over these features for PD detection. The suggested PD detection overall process is presented in figure 1.

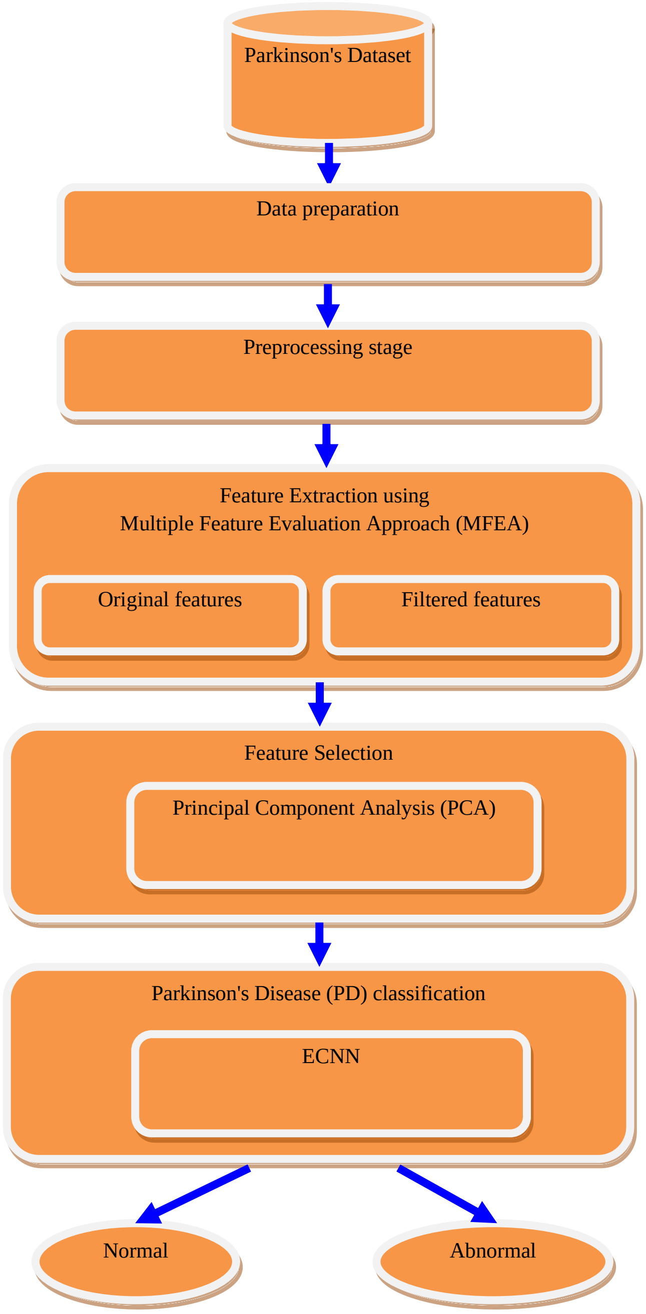

Figure 1. General Process of the Proposed PD Detection

1) Data Preparation

Many researches are being carried out for Parkinson's dataset extraction. Once physical signals are obtained signals are analyzed and filtered to 5923 signals (Ramayya et al., 2014); (Zhang & Tian 2016); (Shi et al., 2018). Main research objective is density function updating and thereby data forcing is done to develop more typically distributed along with stable UPDRS features for regression arrangement. Many linear signals processing approaches like classification and regression tree (CART) in vowel phonation (dysphonia) measures are deployed for attaining signal characteristics and these features characteristics are regarded as Parkinson's dataset. For comparing measures, non-parametric statistical tests are utilized and UPDRS comprising Gaussian kernels and Spearman correlation coefficient for ensuring accurate evaluation. Later, extracted features mapping is done through regression approaches for obtaining final dataset.

Unified Parkinson's disease Rating Scale (UPDRS): The statistical process is mainly involved in regression analysis for substantiating features dependency and estimating relationships amid features. By using Bayesian information and Akaike information criterion class model construction is done in feature selections process. A lot of features can be handled by these techniques with dissimilar sequence besides best fit of the model is estimated. The input features are rejected and passed which are not associated with target output for ensuing test nodes (Lecturio, Pathology, Stages and Life Expectancy, Parkinson's Disease- Symptoms 2018). At the meantime, work of (Little et al., 2007) exploited Least Absolute Shrinkage and Selection Operator (LASSO) as feature selection technique. Nonetheless, both techniques shared identical results affecting features number and dimensions. Parkinson's dataset outcomes is said to have 23 numerical real numbers features. First two features are recording number and patient name, which are excluded in classification process.

As a final point, cross-validation statistical technique is mainly deployed for model generalization. In model's parameters setting, it is achieved by using unused data. Dataset division is done into two training along with testing subsets for classification outcome's Mean Absolute Error (MAE) computation. The overall tests outcomes reveal acceptable prediction accuracy. Henceforth when Parkinson's dataset is analyzed, it reveals that, (1) all features are real numbers numerical (2) it is extensive and no noisy or non-numeric data does exist, and (3) there is no observable patterns. Nonetheless, Parkinson's dataset unbalancing classification process complexity and diminutions in classification outcomes robustness.

2) Feature Evaluation

Multi-agent system is operated through Multiple Feature Evaluation Approach (MFEA). This system has intelligent agents set possessing an interaction with each other surrounded by an environment. These agents serve jointly for mitigating the problems which are not elucidated independently. The agents' application in a system benefits for system flexibility improvement by its functionality segregation and facilitating interaction with its modules in the course of runtime.

MFEA comprises five agents in which every agent functions as specific feature assessment technique. The features are assessed along with ranking by agents and feature vectors subsets are obtained by search algorithm implementation. The processes take account of weighting every feature and feature ordering subsets rendering to their distinct assessment outcomes. The agents on basis of features subsets implement feature filtering process which yields various feature vectors preliminary copies. At that time collaboration takes place with each other for generating optimized feature vector. MFEA overview is illustrated in figure 2.

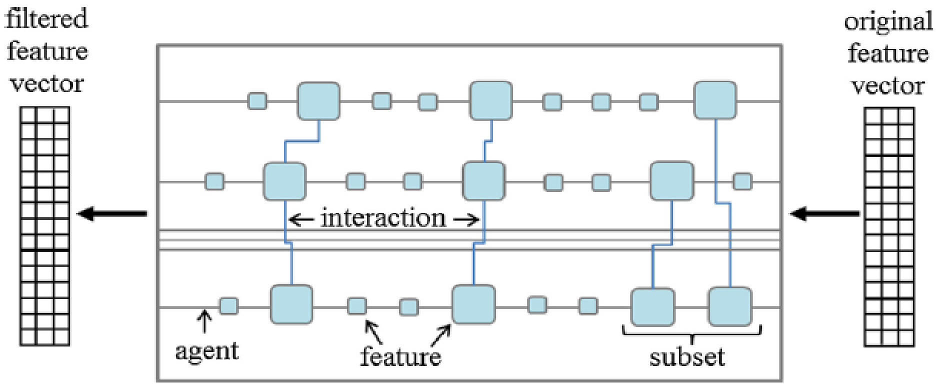

Figure 2. An Overview to the MFEA

The following factors define feature evaluation operators based on five agent:

- $\propto_1$ is the first operator, through which an Autocorrelation feature evaluator, ∂_1 is operated. Pinch period of a subset of features is determined by ∂_1 for evaluating feature subset's worth. If the correlation coefficient scores of a feature is lower than 0.95, it is considered being efficient, or else it is rejected.
- $\propto_2$ is the second agent that helps operating a CFS feature evaluator, ∂_2. Each feature's predictive capability and redundancy level within each feature are measured by ∂_2, through which feature subset's worth is evaluated.
- $\propto_3$ is the third agent, by which a Gain Ratio feature evaluator, ∂_3 is operated. Gain ratio as regards the respective class is measured by ∂_3 for evaluating feature subset's worth.
- $\propto_4$ is a fourth agent, through which an Info Gain feature evaluator, ∂_4 is operated. Information obtained as regards the respective class is measured by ∂_4 for evaluating feature subset's worth.
- $\propto_5$ is a fourth agent that helps operating a SVM evaluator, ∂_5. Using classifier, feature subset's worth is evaluated by ∂_5.

The feature's appearance frequency and features rank's mean score value are measured through collaboration of agents. In the selected features, a maximum appearance frequency and lesser mean rank values are involved. Subsequently, subsets of the chosen features are combined for making the filtered feature vector. Through this method, optimal features accompanied by maximum ranks are selected, meanwhile the features are prevented from over-reduction or overweighting. As such, the nominal features are ignored/terminated for generating a feature vector which is compatible with various kinds of classifiers. Since the features may differently be ranked by each of the features, the weak features are collectively ignored by them. Thus, the MFEA tends to enhance the classification task and ignore classifiers' global minimum learning, which can be derived through over fitting some of the classified features.

Consider that feature vector's m instance's length, V is possessed by initial dataset D. Then, define V $=\{X,Y\}$, where it has input parameters set, X $= \{x_1, x_2, \dots\dots, x_n\}$ and output parameters Y=$\{y_1, y_2, \dots\dots, y_m\}$. Finally, the D is formulated as follows,

$$D=\begin{bmatrix} x_1 & x_2.. & x_n \\ x_1 & x_2.. & x_n \\ x_1 & x_2.. & x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ . \\ y_m \end{bmatrix} \quad (1)$$

The features of each agents are evaluated by the corresponding agents by using $\propto_1: X \to x_1 » \ldots » x_n$, where the permutation process of X is signified by ». In multi-agent system's collective framework, filter function is included, through which the included feature, $\text{in}x_{k1}{}^{in}$ is determined from the excluded feature, x^{ex} of received X.

$$\text{Filter}(X) = \begin{cases} x_{k1}^{in} \, f_i \geq \frac{1}{n}\sum_{i=1}^{n} f_i \wedge rank\left(x_i, \frac{1}{f_i}\sum_{j=1}^{f_i} r_{i,j}\right) \leq t, \\ x_{k2}^{ex} otherwise \end{cases} \quad (2)$$

In which, feature is denoted by x_i; the total number of features of the X is signified by n. Besides, consider that i= $\{1,2 \ldots . n\}$, where x^{in}+ x^{ex}=X. The indexes for the x^{in} is denoted by k1, and indexes for x^{ex} is signified by k2. The appearance frequency of x_i is represented by f_i, and $r_{i;j}$ is the rank of x_i by a ∂_j in which, j is corresponding feature evaluator agent index. Here, a function is referred as rank, by which feature's referenced rank is returned as an integer value, and required number of $x^{in \text{ is}}$ threshold is signified by t.

In accordance with the outcomes, the final quantity of selected features is determined by t, which can be either assigned by a human or estimated by the agents autonomously. By applying the following equation (3), a specific length of feature vector of D is estimated.

$$t = \frac{\frac{1}{k0}\sum_{k0=1}^{d} s_k * n}{m} \; t<n \quad (3)$$

Here, D's size is represented by m; D's original features count is signified by n. In terms of a test attempt, the training diminutions are denoted by d; an index of the attempt is denoted by k_0. In each attempt, the training dataset size is indicated by s. In the algorithm given below, multi-agent system's MFEA is depicted, where the filtered features are denoted by $\tilde{X}$.

Algorithm 1: The MFEA

1. begin;
2. initial parameters;
3. Perform for every active a_j
4. $X \leftarrow prepare\left(a_j ; D\right)$
5. a_j do i until n
6. $f_{i;j} \leftarrow$ evaluate (a_j, x_i);
7. $r_{i;j} \leftarrow rank\ (a_j, x_i)$
8. $x_{i;j} \leftarrow search\ (a_j, x_i)$
9. end-do
10. $X^{\frown} \leftarrow select\ (a_j, x_{1,2,\ldots,j})$
11. end-for
12. $\tilde{X}$ $collaborate$ (
$t\ update(a_{1,2,\ldots\ldots,t})$,
$\{x^{in}, x^{ex}\} filter(a_{1,2;\ldots} \tilde{X}_{1,2;\ldots}\}$
);
13. Stop;

As represented in Algorithm 1, consider that feature vector's m instance's length is comprised in initial dataset D. The features of each agents are evaluated by the corresponding agents. In a collective framework of the multi-agent system, a filter function is included, by which the X is received and the included feature x^{in} is determined from the excluded feature x^{ex} of received X. In accordance with the outcomes, the final quantity of selected features is determined by t.

3) Feature Selection Using Principle Component Analysis (PCA)

To perform the feature selection task, the Principal Component Analysis (PCA) is significantly preferred in this study. PCA is an efficient dimensionality reduction model that is capable of diminishing the wavelet coefficients size from speech signals. Consider that, a dataset C that accompanies the size of *N*, and dimension of *d*. Then, the sample mean m_j of j^{th} feature is computed as follows, 215

$$m_j = \frac{1}{N}\sum_{i=1}^{N} C(i,j) \quad (4)$$

Then, the zero-mean dataset B is estimated through,

$$B = C - em^T \quad (5)$$

In which, an $N \times 1$ vector of each one is defined by e. Then, the generation of $d \times d$ covariance matrix Z is formulated as follows,

$$z = \frac{B * B}{N-1} \quad (6)$$

As a fourth step, the covariance matrix Z has an eigen decomposition expression as

$$z = XYX^{-1} \quad (7)$$

Here, the Eigen vector matrix is represented by X, and the Eigen value matrix (also known as a diagonal matrix) is signified by Y.

$$Y = \begin{bmatrix} Y(1,1) & & \\ & Y(2,2) \ddots & \\ & & Y(d,d) \end{bmatrix} \quad (8)$$

Then, X and Y need to rearranged. Hence, the eigen value is in a decreasing way

$$Y(1,1) \geq Y(2,2) \geq \cdots .. \geq Y(d,d) \quad (9)$$

Subsequently, for each eigen vector, the cumulative variance is computed as given by,

$$G(K) = \sum_{I=1}^{K} Y(i,i) \quad (10)$$

Accordingly, a vector can be generated in the form of

$$G = [G(1)G(2) \ldots \ldots . G(d)] \quad (11)$$

Then, let the threshold be T. In this way, choose L^* that converges

$$\left\{L \left| \frac{G(L)}{G(d)} \geq T \right.\right\} \quad (12)$$

Ultimately, output L^* most important principal components (Pan et al., 2020); (Jenelius & Koutsopoulos 2017). Consequently, the optimal features can be selected from this process, and it can be fed into subsequent process as an input.

4) Parkinson Diseases Detection Using Enhanced Convolutional Neural Network

Further, Enhanced Convolutional Neural Network (ECNN) is fed with selected features as an input. Being specific among robust deep networks, (Gao et al., 2018); (Liu et al., 2018) CNN is highly capable of involving multiple hidden layers with the ability to convolution and subsampling, through which a low level to high level of features can be extracted from the input data. Generally, there are three layers involved in such kind of networks, namely convolution, sub sampling/pooling and full connection layers. The selected feature is fed into the network as an input. As represented in Figure 5, the input layer involved in the network takes the input (features); from the output layer, the system procures the trained output; and the layers exist in middle of input and out layers are termed as hidden layers. In order to obtain the accurate outcomes, the weight values of the features are optimized in this proposed ECCNN model.

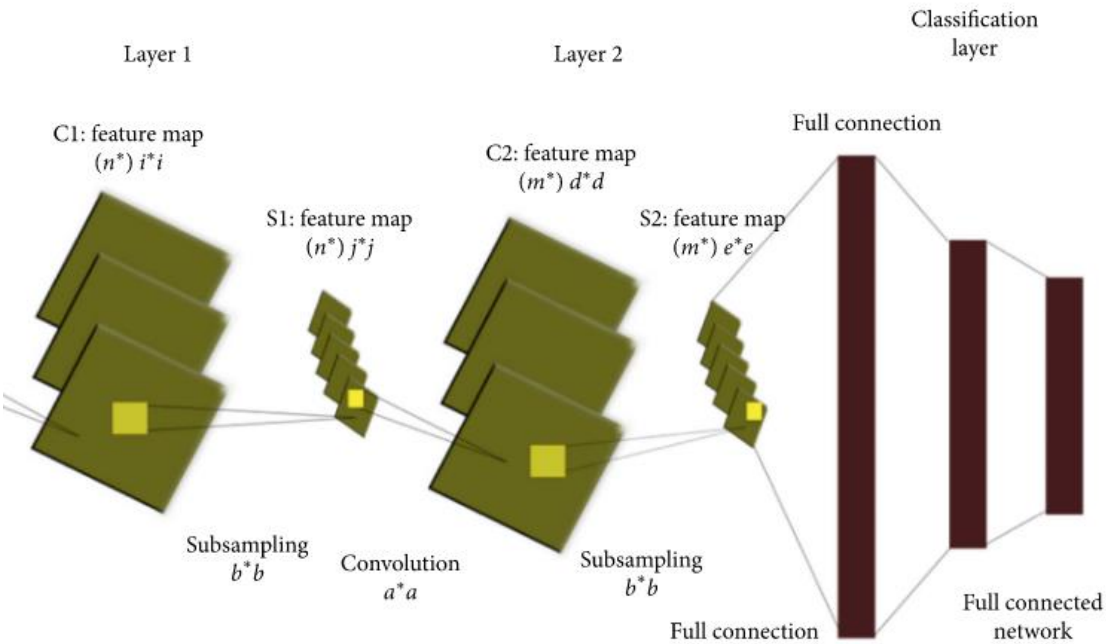

Figure 3. Convolutional Neural Network Structure

Convolution Layer

During the process of this layer, input data with size R∗C is convolved with size a∗a sized filter or kernel. In input matrix, every individual block is independently convolved with filter, through which a pixel is generated in output. Besides, n output features are generated using the convolution result of the input data and kernel. In general, convolution matrix kernel is termed as filter, whereas output data features (which is attained through convolving kernel), and input data are collectively termed as feature maps with i∗i size. In CNN, several convolutional layers can be included, besides subsequent convolutional layers outputs and inputs are called as feature vector. In every convolution layer, numerous n filters involved that are convolved with the input, in which filters count applied in convolution process and generated feature maps (n∗) depth are equal. Remember that, at a specific location of input data, every filter map is taken as specific feature (Zhang et al., 2019); (Prabhakar et al., 2017); (Bhagat et al., 2020).

eISSN 1303-5150

$C_j^{(l)}$ signifies the l-th convolution layer output that comprises the feature maps. It is estimated as follows,

$$C_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i,j}^{(l-1)} * C_j^{(l)} \quad (13)$$

Here, the bias matrix is denoted by $B_i^{(l)}$; convolution kernel/filter of size a∗a is signified by $K_{i,j}^{(l-1)}$, through which j-th feature map in layer (l − 1) is linked with i-th feature map in same layer. Feature maps are included in output $C_i^{(l)}$ layer. In (14), the first convolutional layer $C_i^{(l-1)}$ is input space, denoted by $C_i^{(0)} = X_i$.
Feature map is generated by the kernel. Then, convolutional layer output's nonlinear transformation is applied with activation function next to convolution layer.

$$Y_i^{(l)} = Y(C_i^{(l)}) \quad (14)$$

Here, the activation function output is denoted by $Y_i^{(l)}$; and the input received by the activation function is represented by $C_i^{(l)}$.
During the process, sigmoid, tanh, and Rectified Linear Units (ReLUs) are deployed as an activation functions. Accordingly, $Y_i^{(l)} = \max(0, Y_i^{(l)})$ is applied as ReLUs in this work. Since this function is capable of effectively reducing the impacts of interaction and nonlinear effects, it is widely utilized in deep learning methods. If negative input is received by ReLU, it converts output to 0; if receives positive input, same input value will be returned. Due to the error derivative, this activation function getting faster over other functions. Because it turns out to be very lesser in saturating area. Consequently, weights update nearly vanish, termed as vanishing gradient problem.

Sub Sampling Layer

This layer tends to spatially decrease features maps dimensionality that is derived from preceding convolution layer. As such, in order to perform the sub sampling operation within the mask and the feature maps, a size b∗b mask is chosen. Remember that, amid the input data of convolution layer, the rotation and translation is tolerated with the help of a sub sampling layer. On the basis of mean of the weights of features, optimal weights are updated in this study.

$$\text{Weighted mean } w_H = \frac{N}{\sum_{i=1}^{N} w x_i} \quad (15)$$

Here, N denotes Number of features; w indicates Weight value of the feature; x_i signifies Features.

Full Connection

The following Softmax activation function is applied on the output

$$Y_i^{(l)} = f(z_i^{(l)}), \text{ where } z_i^{(l)} = \sum_{i=1}^{m_i^{(l-1)}} w_H y_i^{(l-1)} \quad (16)$$

In which, the weighted Harmonic mean of the features is denoted by w_H, which needs to be adjusted by the fully connected layer for organizing the depiction of every layer; the transfer function is signified by f that refers to the nonlinearity. Finally, the classification selection process is accomplished through the classifier by connecting the fully connected and output layer.

Results and Discussion

In this section, proposed Enhanced Convolutional Neural Network (ECNN) framework is evaluated on the basis of empirical results obtained through MATLAB 2013 a. Accordingly, the existing approaches, like KNN and FCLAM are considered to compare with the proposed method in order to assess its efficiency. During that, Precision, Recall and F-Measure have been taken as performance parameters, and the implementation was performed on the Parkinsons Dataset, which was gathered from Machine Learning repository belongs to University of California-Irvine (UCI). In this dataset, there are 252 individuals, in which 3 voice records have included for each individual. Besides, from the Neurology Department in Cerrahpasa Faculty of Medicine, Istanbul University, dataset was collected, in which the details of 188 PD patients (male-107; female-81) as well as 64 healthy persons (male-23; female-41) have included. The PD patients' age range from 33 to 87, whereas the age of healthy objects fall between 41 to 82. Microphone's frequency response is fixed at 44.1 KHz during data collection. Post-reviewed by doctor, repeated repetition of the vowel /a/ letter with three replicates was gathered for each person.

Table 1. Software Components

SRL Num	Software Component	Component Description
1	MATLAB 2013a	Intel core Processor
2	Training Phase	80% of Samples
3	Testing Phase	20% of Samples
4	Parkinsons Data Set	

Table 2. Performance Comparison Results

Metrics	Methods		
	KNN	FCLAM	ECNN
Precision	70	85	91
Recall	72	83	90
F-measure	75	81	91
Accuracy	72	83	92

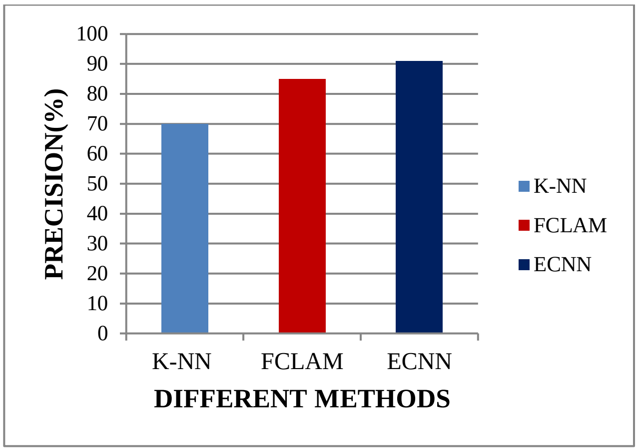

Figure 4. Precision Results Comparison

In figure 4, Precision results individually obtained by the existing K-NN and FCLAM, and the proposed ECNN technique are compared. Throughout this work, for maximizing values as well as Precision rate's quality, the input is gone through preprocessing task with the help of ranking system. In the figure, implemented methods lie on X-axis; and Y-axis stands for their Precision rates. The outcomes prove that the novel ECNN approach efficiently secures 91% Precision rate, and it is capable of outperforming the existing K-NN and FCLAM, since they solely obtain 70%, and 85% precision rates, respectively.

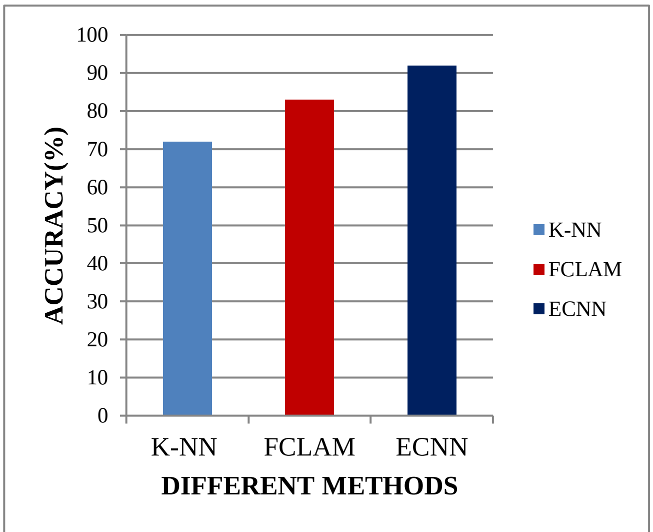

Figure 5. Comparison Result of Accuracy

Figure 5 compares individual Accuracy rates of the existing K-NN and FCLAM, and the proposed ECNN technique. Concerning important features, frequency band division system is utilized by the proposed PCA, through which the ECNN accuracy is significantly enhanced. In the figure, implemented methods lie on X-axis; and Y-axis stands for their Precision rates. The outcomes depict the proficiency of the novel ECNN approach to procure 92% Accuracy rate, which is superior to the existing K-NN and FCLAM, since they are able to solely acquire 72%, and 83% accuracy rates, respectively.

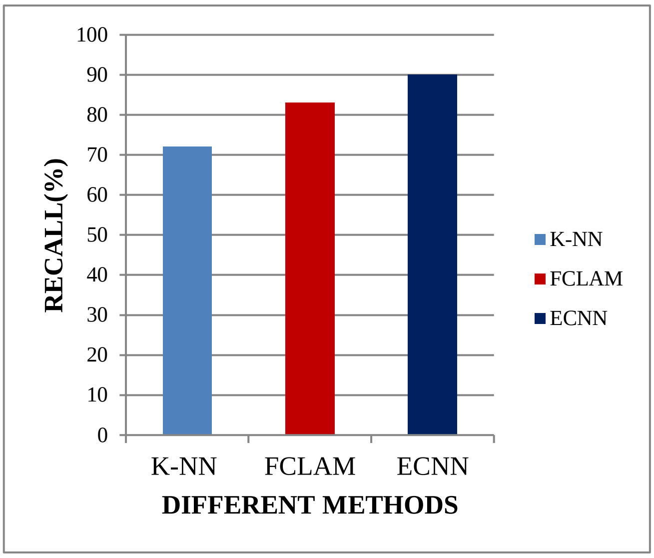

Figure 6. Comparison Result of Recall

In Figure 6, the individual Recall rates of the existing K-NN and FCLAM, and the proposed ECNN technique are compared. In the figure, implemented methods lie on X-axis; and Y-axis stands for their Recall rates, where the novel ECNN approach proves its efficiently by securing 90% Recall. Besides, the proposed approach shows its capability to outperform the existing K-NN and FCLAM, as they are able to obtain only 72%, and 83% Recall rates, respectively.

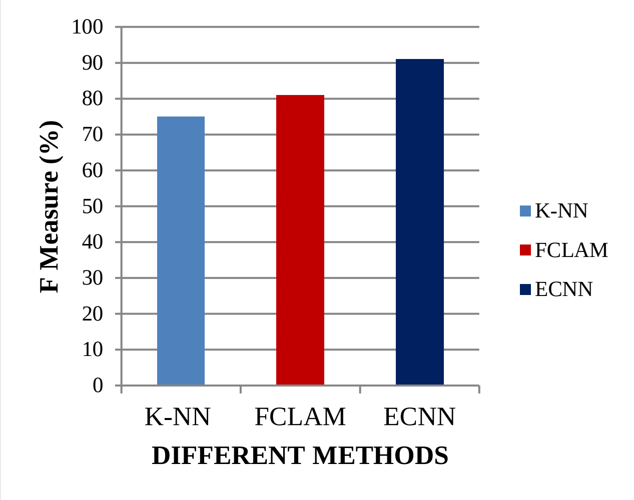

Figure 7. Comparison Result of F -Measure

Figure 7 compares individual F-measure values of the existing K-NN and FCLAM, and the proposed ECNN technique. With regard to learn features, numerous convolution functions are utilized by the proposed method, through which the F-measure rate of ECNN is significantly enhanced. In the figure, X-axis represents implemented methods; and Y-axis stands for their F-measure rates. The outcomes depict the efficiency of the proposed novel ECNN approach to obtain 91% F-measure rate, which is higher than the existing K-NN and FCLAM, since they are able to solely acquire 75%, and 81% F-measure values, respectively.

Conclusion and Future Work

An enhanced framework for Parkinson diseases detection is mainly suggested in this research. Mostly, input PD dataset processing is done as preprocessing steps for validating dataset quality to accomplish the process. On the basis of Multiple Feature Evaluation Approach (MFEA), feature assessment processing is done which comprises quite a lot of feature assessment and ranking algorithms to weight features worth and features set extraction. Furthermore, a feature selection model by principal component analysis is suggested for classifier accuracy improvement. To conclude, Enhanced Convolutional Neural Network (ECNN) training is done over these certain features for PD detection; besides its performance evaluation is done through various metrics. It is thereby validated that suggested methodology outperforms in a better way pertaining to accuracy as well time consumption is reduced through single deep learning model rather than deploying more classifiers along with feature selection model. From the results, the proficiency of the novel ECNN approach to procure 92% Accuracy rate, which is superior to the existing K-NN and FCLAM, since they are able to solely acquire 72%, and 83% accuracy rates, respectively. Conversely, deep learning approach yields additional computational complexities and hence certain other hybrid methodologies are necessitated in future for PD detection. At last, because of PD's small vocal datasets, proposed method is generalized to be much larger datasets in the future.

References

Jain S, Shetty S. Improving accuracy in noninvasive telemonitoring of progression of Parkinson's Disease using two-step predictive model. *Third International Conference on Electrical, Electronics, Computer Engineering and their Applications (EECEA)* 2016: 104-109.

Grover S, Bhartia S, Yadav A, Seeja KR. Predicting severity of Parkinson's disease using deep learning. *Procedia computer science* 2018; 132: 1788-1794.

Aich S, Sain M, Park J, Choi KW, Kim HC. A text mining approach to identify the relationship between gait-Parkinson's disease (PD) from PD based research articles. *International conference on inventive computing and informatics (ICICI)* 2017: 481-485.

Chakraborty S, Aich S, Han E, Park J, Kim HC. Parkinson's Disease Detection from Spiral and Wave Drawings using Convolutional Neural Networks: A Multistage Classifier Approach. *International Conference on Advanced Communication Technology (ICACT)* 2020: 298-303.

Aich S, Sain M, Park J, Choi KW, Kim HC. A mixed classification approach for the prediction of Parkinson's disease using nonlinear feature selection technique based on the voice recording. *In International Conference on Inventive Computing and Informatics (ICICI)* 2017: 959-962.

Devarajan M, Ravi L. Intelligent cyber-physical system for an efficient detection of Parkinson disease using fog computing. *Multimedia Tools and Applications* 2019; 78(23): 32695-32719.

Castaño-Pino YJ, González MC, Quintana-Peña V, Valderrama J, Muñoz B, Orozco J, Navarro A. Automatic Gait Phases Detection in Parkinson Disease: A Comparative Study. *In 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 2020: 798-802.

Fayyazifar N, Samadiani N. Parkinson's disease detection using ensemble techniques and genetic algorithm. *In Artificial Intelligence and Signal Processing Conference (AISP)* 2017: 162-165.

Ali L, Zhu C, Zhou M, Liu Y. Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection. *Expert Systems with Applications* 2019; 137: 22-28

Upadhya SS, Cheeran AN, Nirmal JH. Thomson Multitaper MFCC and PLP voice features for early detection of Parkinson disease. *Biomedical Signal Processing and Control* 2018; 46: 293-301.

Li Y, Zhang C, Jia Y, Wang P, Zhang X, Xie T. Simultaneous learning of speech feature and segment for classification of

Parkinson disease. *International Conference on e-Health Networking, Applications and Services (Healthcom)* 2017: 1-6.

Sharma P, Sundaram S, Sharma M, Sharma A, Gupta D. Diagnosis of Parkinson's disease using modified grey wolf optimization. *Cognitive Systems Research* 2019; 54: 100-115.

Nilashi M, Ibrahim O, Ahmadi H, Shahmoradi L, Farahmand M. A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques. *Biocybernetics and Biomedical Engineering* 2018; 38(1): 1-15.

Gupta D, Sundaram S, Khanna A, Hassanien AE, De Albuquerque VHC. Improved diagnosis of Parkinson's disease using optimized crow search algorithm. *Computers & Electrical Engineering* 2018; 68: 412-424.

Ramayya AG, Misra A, Baltuch GH, Kahana MJ. Microstimulation of the human substantia nigra alters reinforcement learning. *Journal of Neuroscience* 2014; 34(20); 6887–6895.

Zhang J, Tian GY. UHF RFID tag antenna-based sensing for corrosion detection & characterization using principal component analysis. *IEEE Transactions on Antennas and Propagation* 2016; 64(10): 4405-4414.

Shi X, Nie F, Lai Z, Guo Z. Robust principal component analysis via optimal mean by joint ℓ2, 1 and Schatten p-norms minimization. *Neurocomputing* 2018; 283: 205-213.

Wu SX, Wai HT, Li L, Scaglione A. A review of distributed algorithms for principal component analysis. *Proceedings of the IEEE* 2018; 106(8): 1321-1340.

Pan F, Zhang ZX, Liu BD, Xie JJ. Class-Specific Sparse Principal Component Analysis for Visual Classification. *IEEE Access* 2020; 8: 110033-110047.

Jenelius E, Koutsopoulos HN. Urban network travel time prediction based on a probabilistic principal component analysis model of probe data. *IEEE Transactions on Intelligent Transportation Systems* 2017; 19(2): 436-445.

Gao H, Cheng B, Wang J, Li K, Zhao J, Li D. Object classification using CNN-based fusion of vision and LIDAR in autonomous vehicle environment. *IEEE Transactions on Industrial Informatics* 2018; 14(9): 4224-4231.

Liu Y, Suen CY, Liu Y, Ding L. Scene classification using hierarchical Wasserstein CNN. *IEEE Transactions on Geoscience and Remote Sensing* 2018; 57(5): 2494-2509.

Zhang Z, Wang C, Gan C, Sun S, Wang M. Automatic modulation classification using convolutional neural network with features fusion of SPWVD and BJD. *IEEE Transactions on Signal and Information Processing over Networks* 2019; 5(3): 469-478.

Prabhakar G, Kailath B, Natarajan S, Kumar R. Obstacle detection and classification using deep learning for tracking in high-speed autonomous driving. *IEEE region 10 symposium (TENSYMP)* 2017: 1-6.

Bhagat M, Kumar D, Mahmood R, Pati B, Kumar M. Bell Pepper Leaf Disease Classification Using CNN. *In 2nd International Conference on Data, Engineering and Applications (IDEA)* 2020: 1-5.

Little M, McSharry P, Roberts S, Costello D, Moroz I. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *Nature Preceding's* 2007: 1-35.

Sadikov A, Žabkar J, Možina M, Groznik V, Nyholm D, Memedi M. Feasibility of spirography features for objective assessment of motor symptoms in parkinson's disease. *In Conference on Artificial Intelligence in Medicine in Europe* 2015: 267-276.

Badem H, Caliskan A, Basturk A, Yuksel ME. Classification and diagnosis of the parkinson disease by stacked autoencoder. *National Conference on Electrical, Electronics and Biomedical Engineering (ELECO)* 2016: 499-502.

Indira, & Can M. Diagnosis of Parkinson's Disease using Fuzzy C-Means Clustering and Pattern Recognition. *South East Europe Journal of Soft Computing* 2013; 2(1): 42-49.

Grover S, Bhartia S, Yadav A, Seeja KR. Predicting severity of Parkinson's disease using deep learning. *Procedia computer science* 2018; 132: 1788-1794.

Armañanzas R, Bielza C, Chaudhuri KR, Martinez-Martin P, Larrañaga P. Unveiling relevant non-motor Parkinson's disease severity symptoms using a machine learning approach. *Artificial intelligence in medicine* 2013; 58(3): 195-202.

Sakara BE, Kursunb, Telemonitoring of changes of unified Parkinson's disease rating scale using severity of voice symptoms. *In Proc. 2nd International Conference on E-Health and Tele Medicine, Istanbul* 2014: 114-119.

Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO. Novel speech signal processing algorithms for high accuracy classification of Parkinson's disease. *IEEE Transactions on Biomedical Engineering* 2012; 59(5): 1264-1271.

Tsanas A, Little MA, McSharry PE, Ramig LO. Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *Journal of the Royal Society Interface* 2011; 8(59): 842-855.

Sharma A, Giri RN. Automatic recognition of Parkinson's Disease via artificial neural network and support vector machine. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* 2014; 4(3): 2278-3075.

Khemphila A, Boonjing V. Parkinsons disease classification using neural network and feature selection. *World Academy of Science, Engineering and Technology* 2012; 64: 15-18.

Saad A, Zaarour I, Zeinedine A, Ayache M, Bejjani P, Guerin F, Havre-France L. A preliminary study of the causality of freezing of gait for Parkinson's disease patients: Bayesian belief network approach. *International Journal of Computer Science Issues* 2013; 10(3): 88-95.

Hadjahmadi AH, Askari TJ. A decision support system for Parkinson's disease diagnosis using classification and regression tree. *The Journal of Mathematics and Computer Science* 2012; 4(2): 257-263.

Shao J. Western culture acquisition of young cadres based on cognitive linguistics. *NeuroQuantology* 2018; 16(6): 423-429.