Measurement Sensors 16 (2021) 100047

Contents lists available at ScienceDirect

Measurement: Sensors

journal homepage: www.sciencedirect.com/journal/measurement-sensors

Digitized spiral drawing classification for Parkinson's disease diagnosis

Megha Kamble *, Prashant Shrivastava, Megha Jain

Department of Computer Science and Engineering, Lakshmi Narain College of Technology Excellence, Bhopal, India

ARTICLE INFO

Keywords:
Parkinson's disease classification
Machine learning classification
Ensemble model
Spiral kinematics

ABSTRACT

Parkinson's disease (PD) is the most common neurodegenerative disease affecting significantly motor functions of elderly persons. The diagnosis and monitoring of PD is costly and inconvenient process even today, in under developing parts of the world. The observable symptoms of PD at early stage include disorders in handwriting and repetitive tasks of spiral drawing. With advancement of IT it is easier to collect spiral drawing samples using digitized tablet. We proposed detailed analysis of Static and dynamic spirals drawn by PD patients. For this, in-air and on-surface kinematic variables are taken out from data files generated for 25 patients and 15 healthy controls, using mathematical models. Results demonstrated nearly 91% classification accuracy to separate PD patients from healthy controls by applying feature engineering and four machine learning (ML) classifiers Logistic Regression, C-Support Vector Classification (SVC), K- nearest neighbor(KNN) classifier and ensemble model Random Forest Classifier(RFC). This paper confirms that digitized spiral drawings have major impact on classification of PD patients and healthy controls and hence can support future differential diagnosis of PD.

1. Introduction

Spiral drawing is a skilled and complex coordinated motor activity. Therefore, it is treated as sensitive motor assessment. Motor rating scale and its subscale Unified Parkinson's disease rating scale (UPDRS-III) is the most widely used and accepted rating scale in Parkinson's disease (PD). PD affects various functions of body such as speech, handwriting; walking, coordination movements and all these are considered part of motor functions. All quantitative indices of motor decline and non motor biomarkers have been proposed to check the degree of severity as PD is treated as motor disorder due to neurodegeneration process. The diagnosis and monitoring of PD is costly and inconvenient process for two reasons: it is inconvenient for patient caretakers to take the patient to clinic and perform physical examination and diagnosis based on physical observation requires trained medical experts. A clinical invasive technique are available only at early stage, but incur risk with limited resources available in under developing parts of the world and if early diagnosis is done then only it is useful.

Hence, from prior time, PD was assessed by traditional non-invasive methods such as handwriting test and spiral drawing pen paper test. Collecting, preserving the paper and analyzing these drawings is purely based on human expertise, which it time taking less accurate biased method. With the advancement of IT, it is easier to collect the drawing samples using digitizing (digital) tablets, which are already used in biomedical research. The use of graphics tablet enables the researchers to develop various image analysis and processing tools for obtaining different kinds of information that cannot be easily performed with pen paper drawings.

1.1. Literature review

This section provides the details of the methodology in the primary domains of Machine Learning and Neural Network used along with information, datasets, and set of images etc previous efforts for PD diagnosis with corresponding references.

The significant work is done in analyzing handwriting samples and speech samples for evaluation of tremor and degree of severity of PD, with the conventional techniques of Neural Network and regression. To improve the accuracy of the diagnosis, the work can be extended for scientific diagnosis tool of PD i.e. Spiral drawing. Machine Learning models and ensemble learning models prove better in predictions and diagnosis supported by feature engineering for revealing more details of symptoms.

Peter Drotar et al. [1,2] suggested analysis of PD patient handwriting by applying feature selection algorithm and Machine Learning SVM method. This is one of the early works in identifying the importance of in-air/on-surface hand movements in diagnosis of motor disorder of neuro degenerative diseases. The results of the paper demonstrated that

* Corresponding author. Department of CSE, Faculty of CSE, LNCT Bhopal, Raisen Road, India.
E-mail address: meghak@lnct.ac.in (M. Kamble).

https://doi.org/10.1016/j.measen.2021.100047
Received 10 October 2020; Received in revised form 16 April 2021; Accepted 19 April 2021
Available online 30 April 2021

these movements have major impact during assessment of handwriting and illustrated 85.61% prediction accuracy in Ref. [1]. The work presented PaHaW handwriting's database generated by allowing PD patients to perform eight different handwriting tasks including Archimedean Spiral. The work [2] investigated new pressure features and stroke features related to dynamics of handwriting in extensive manner. Then, applied three different classifiers KNN, AdaBoost ensemble and SVM to demonstrate 81% classification accuracy. However, other demographic features such as age, education, job, and birthplace, disease duration to illustrate more practical and effective diagnosis of PD patient must support the handwriting test. In lieu of this handwriting experiment, Archimedean spiral based scientific method of static and dynamic spiral can be assessed as more relevant with the help of powerful ML methods.

Scientific method of spiral drawing comprises of three types [3]: static spiral test, dynamic spiral test the Stability Test on a Certain Point (STCP).

Donalto Impedovo et al. [3,4], have also presented handwriting as a powerful marker to develop PD diagnostic tool. The authors have implemented ML classification framework on same PaHaW dataset for good specificity performance measures. The work does not only illustrated binary discrimination of the patients but the works is extended to group the patients based on degree of severity of the illness to further support early cure to diseased persons. Though work has provided extensive dataset, and grouping based on illness for proper classification, scientific tools of spiral can be extended for simple performance.

Poonam Zham et al. [5] presented and investigated kinematic features of micrographia associated with PD, during repetitive writing task. The experimental setup demonstrated samples of letter "e" similar to spiral, and average size of first five and last five letters were compared. These handwriting samples are assessed by Wilcoxon signed-rank test, Kruskal wall's test.

Omer Eskidere et al. [6] described applications of Machine learning frameworks SVM, LSSVM, MLPNN and general regression NN for remote tracking of PD progression. The work is implemented on telemonitoring dataset comprising of early stage PD patients' voice recordings in 26 attributes. Performance measures MAE, MSE and correlation coefficients are applied to evaluate the prediction model to conclude that LS-SVM is superior in mapping UPDRS with vocal features. It works on classical and non classical, non linear voice features. Feature extraction and normalization of voice features is complex process, requires wider range for normalization and accuracy level depends on the signal processing algorithm. This is traditional mechanism to diagnosis of PD and in early stage the symptoms may not be seen in patient. So UPDRS mapping with only voice features may not lead to accurate diagnosis. The two common drawbacks with the ML algorithms presented in the paper are; greater computational burden and proneness to over fitting.

Mohammed Erdem Isenkula et al. [7] suggested improved spiral test using digitizing tablet over traditional Static spiral test with pen and paper. New Dynamic Spiral Test is suggested by the author. Dynamic spiral test only be performed with tablet and the paper has concluded that digital samples generated by computer system DST drawings combined with SST drawings can be used to build generic PD telemonitoring, diagnosis and controlling system. The tablet specifications provided in the paper are resolution of 1000 pt/cm, accuracy 0.025 cm and 256 levels of measurable pressure connected to computer via USB using proprietary software to process the spirals. Ideal spiral and PD patient spirals are transformed into radius angle transformation that is mathematically equivalent of spiral express linearity in terms of r,Θ (theta) polar coordinates. The paper is significant step towards non-invasive remotely monitoring tool for early stage diagnosis of PD, the most dangerous neuro degenerative disease. However, paper has suggested framework for combining SST and DST drawings for image processing for PD diagnosis. The experimental setup is not provided to move ahead with diagnosis work. The proposed work in this paper is providing solution to use effectively SST and DST drawings of healthy and diseased persons for discrimination of these subjects for clinical assessment at early stage.

Marta San Lucianol et al. [8] suggested use of spiral drawing for computer analysis of PD as digitized spirals correlate with motor scores. Generated/derived indices correlated with overall spiral execution are severity, shape and kinematic irregularity (second order smoothness, first order zero crossing), tightness, mean speed and variability of spiral width.

Cross validation performed with linear mixed model to examine validity of combined indices, discriminative validity performance measure was opted with sub parameters sensitivity and specificity to conclude early PD diagnosis in the samples and proved spiral analysis accuracy for discrimination of subjects with PD and early PD from controls.

Samayeh Aghnnavesi et al. [9] presented their work aiming to verify and investigate clinimetic properties of an entropy based method for measuring PD related upper limb temporal irregularities while drawing spirals using digitized tablets. The work demonstrated calculations of Temporal Irregularity Score(TIS) temporal irregularity score and differences in mean TIS between patients and healthy subjects. The significant difference of TIS between healthy and advanced PD subjects demonstrate reliability of measures of TIS for relatively small dataset.

F.Miralles et al. [10], developed a new quantitative analysis of spiral drawing drawn over print template. Cross-correlation coefficient, mean and standard deviation of actual drawing and template drawing, Fourier transformation are the methods applied for analysis. Results of tremor analysis with scanned image drawings is illustrated with ROC curve. This is one of the early experiments done for evaluation of patients' movement disorders. However, major disadvantage of Fourier transformation is it is difficult to ensure the subtle changes in the frequency over the time when applied on scanned image. Therefore, it cannot explore more dynamics of the movement. In addition, these scanned images require preprocessing for accurate results and analysis.

The authors [11] have validated spiral drawing to propose a measurement scale for motor dysfunction in PD. Several spiral indices such as degree of severity, first order zero crossing, second order smoothness, and mean speed are calculated to compare them with clinical scale motor UPDRS and results demonstrated that there is significant correlation of spiral drawings with assessment of PD and thus given a detailed analysis for PD diagnosis with non invasive spiral drawing test. This is pen paper test and statistical measures were applied to it. The drawback is presence of outliers and so to have more accurate results digital spirals and ML methods are found to be more suitable for PD diagnosis.

The authors [12] have presented guided Archimedean spiral drawing for PD patients and recorded data corresponding to n samples required to draw it. Along with conventional kinematic features, the paper introduced composite index of speed and pressure (CISP) to evaluate severity of the PD. The results demonstrated that this is a novel feature with strong correlation with motor ability deterioration due to PD. The limitations of the work are dependence on specially designed spiral tool, real time results can demonstrate the current situation, progressive information cannot be derived and few samples are available which are to be extended for analysis that is more extensive.

The objectives of the authors in Ref. [13] are to provide an overview of the methods and technologies used today in handwriting analysis for movement disorder diseases and focusing all the possible parameters related to handwriting that may provide research direction for accurate diagnosis and cure for neural degenerative diseases such as Parkinsons, Palsy etc. The paper is review paper stating different feature extraction mechanisms for handwriting and exploring the future possibility about signature analysis, visual abilities and attention abilities based on handwriting analysis. The kinematic features and datasets presented in the literature are mostly based on demographic features and so might affect evaluation of PD.

The authors [14] collected handwriting samples from 37 medicated PD patients and 38 age- and sex-matched controls by performing seven

handwriting tasks. Conventional kinematic and spatial-temporal handwriting measures and some computed novel handwriting measures based on entropy, signal energy, and empirical mode decomposition of the handwriting signals are generated. Selected features when classified with SVM classifier provided 88% accuracy with highest sensitivity and specificity. However, the medicated PD patients' data need to be reevaluated by some reliability tests for more accurate results and extensive feature extraction only contributed to complexity of model without significantly improving the performance.

Recently [15] has presented Convolutional neural network latest image analysis tool for detecting PD with spiral drawing. The work is on limited dataset of images.

The work presented in the literature review reveals the significance of spiral drawing for identification of Parkinson's disease. The existing references has generated dataset of small number of images of automated tablet traces. The feature extraction of kinematic movements is not clearly specified in the literature. UPDRS scale can be used effectively for specifying the status or disease condition and as per medical researchers have stated, UPDRS scale is stated in four sections vis. Motor, kinematic, Speech processing, decision making. Motor and kinematic disorder provides flexibility and different datasets which can be analysed using emerging techniques of machine learning for data analysis. This also supports measurement of tremor that is helpful for PD early diagnosis.

The research limitations of finding out kinematic features is trounced in this proposed methodology. The image dataset and coordinate dataset text file is converted to kinematic features dataset for healthy and patient. The objective is to identify significant kinematic features and applying machine learning to aid PD prediction, in other words its binary classification of healthy control and patient with the help of limited dataset of tablet based spiral drawings.

1.2. Highlights

A spiral drawing database containing samples from 25 PD and 15 healthy subjects is presented and results are demonstrated for accuracy, precision, recall and F1 score for discrimination patients and control subjects.

- Spiral Hand movements are used to assist in the diagnosis of Parkinson's disease.
- The pressure exerted on the surface and grip angle during spiral hand movement, are mapped to kinematic features and converted into extensive data set contributing to the diagnosis of motor disorder caused by PD.
- The dataset is combination of three tests static, dynamic and stability drawing test. So model is flexible to any type of test.
- Two unique features are introduced related to radial velocity, moderately supporting to diagnosis.
- Exclusive method based on simple spiral drawing can further be extended for diagnosis of other neurodegenerative diseases.
- Our contribution is applying most correlated unique features on flexible data set combining all tests, for more accurate diagnosis of PD.

1.3. Text organization

The remainder of this paper is organized as follows. Section 2 gives description of existing references and methods already pointed out in the literature, with proposed work framework. Section 3 provides extensive list of features mapped from existing dataset, experimental results and discussions. Section 4 gives limitations of our system and presents conclusions with future scope.

2. Existing methods and experiment performed

2.1. Digital spiral basics

The graphics tablet is equipped with specially designed software that records spiral drawing, and generates data file (say csv format) with coordination parameters pen grip angle, pressure, time consumed or current timestamp and x,y,z coordinate values of every stroke stored. This can be self-administered and noninvasive application that enable PD patient or care takers to collect data at home and transmit it over internet to dedicated application or server for processing. Here we propose the detection method of presence of PD by image processing. There are three types of spiral drawing tasks suggested clinically [3]. The data files of static spiral and dynamic spiral drawings are available on kaggle. With the help of mathematical model for kinematics, additional coordination parameters related to tremor are taken out. Classification and prediction algorithms are applied on this modified dataset for separating healthy control and PD patients.

The proposed work avoids the handwriting samples for PD diagnosis as demographics features, literacy of person is important in case of handwriting, and so it cannot be considered as independent tool for PD diagnosis. The proposed work suggests scientific tool of SST, DST and STCP to identify diseased person. The literature has provided handwriting analysis and Archimedean guided spiral analysis. However, literature has not provided diagnosis using digitized recordings of scientific biomarker i.e. Spirals for PD patients. To make the work extensive, additional kinematic features inspired by Refs. [1,2] are calculated using mathematical model for kinematics. This feature engineering has converted the spiral drawing dataset into precise dataset that can be applied in future by researchers for effective early diagnosis of PD.

2.2. Performed work

2.2.1. Data acquisition and collection

Spiral drawing dataset consists of drawing samples and digital record of 25 PDP(Parkinson's disease patients) and 15 healthy individuals is downloaded from UCI machine learning repository. The dataset is provided by the authors [9] from Istanbul University, at the resource https://archive.ics.uci.edu/ml/datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet.

From all the age-matching persons, three types of drawings are drawn on digitized tablet. These three types of spiral drawing tests are based on scientific base for motor disorder assessment and are as follows:

Static Spiral Test(SST): Three wounds of Archimedean Spiral appears on the graphics tablet and persons are asked to retrace them with digital pen.

Dynamic Spiral Test(DST): Unlike static Archimedean spiral, in this test, spiral just appears and disappears in certain time interval. The patient has to retrace the blinking spiral by keeping the pattern in the mind. This determines the pattern memorizing skills of patient and pause time taken by the patient while retracing the same pattern.

Stability Test on Certain Point(STCP) or Circular motion test: there is a red point in the middle of the screen and patients are drawing the circle without touching the screen. This determines the hand stability of patient.

During the test, the specialized software designed using API and C#, records the information of the drawing. The drawings i.e. png images of SST and DST and text files (csv format) are appended to the dataset. The same dataset of text files is used in the proposed work.

Handwriting dataset was constructed using Wacom Cintiq 12WX graphics table [7]. It is a graphics tablet and LCD monitor rolled into one. It enables to display a PC's screen on its monitor and only interacts with digitized pens. The device sampling was event-based, which means a sensor event was generated every time the sensor values x and y were changed.

The digital readings recorded are x,y coordinates, pressure axes providing "tri-axial" recordings, pressure (Unit less), stylus grip angle, timestamp, test id(SST-0, DST-1,STCP-2). The geometric position of the pen at certain time stamps(x,y), as well as the pressure exerted over the writing surface, pen inclination, and if the movement of the pen is performed "in the air" are the kinematic representations.

Time consumed to complete drawing can be calculated using individual stroke timestamp recorded in the test file w.r.t each spiral drawing, samples are shown in Figs. 1, 2 and 3. In the literature, SST and DST images are analysed based on speed and acceleration instantaneous acceleration as velocity changes dramatically.

2.2.2. Feature engineering

Samples stored in text file and samples required to complete the drawings are longer in case of SST as compared to DST. This lacks kinematic features, which are as based on three factors instantaneous velocity, acceleration and fluency. So other features, inspired by Refs. [1,2], are generated using mathematical model are:

List of features.

2.2.2.1. Kinematic features.

- Duration of static and dynamic spiral (calculated from timestamp provided in csv file)
- A stroke is a single linked and ongoing trait feature of the spiral-drawing pattern.
  - Number of strokes are calculated by counting number of times on-surface pressure is changing during the whole spiral drawing
  - On-surface strokes match with the outline left on the tablet surface
  - in-air strokes are unreal or abstract traces that are stating the pauses and gaps between the letters and also uncertainty in drawing ability or indecisive behavior while drawing
- Stroke speed: Using python numpy library function `linalg.norm`, total distance covered by spiral drawing is calculated with the help of x,y and z coordinates available in csv file, and divided by total time duration also available in timestamp field in csv file.Refer to Figs. 4 and 5 for sample values.

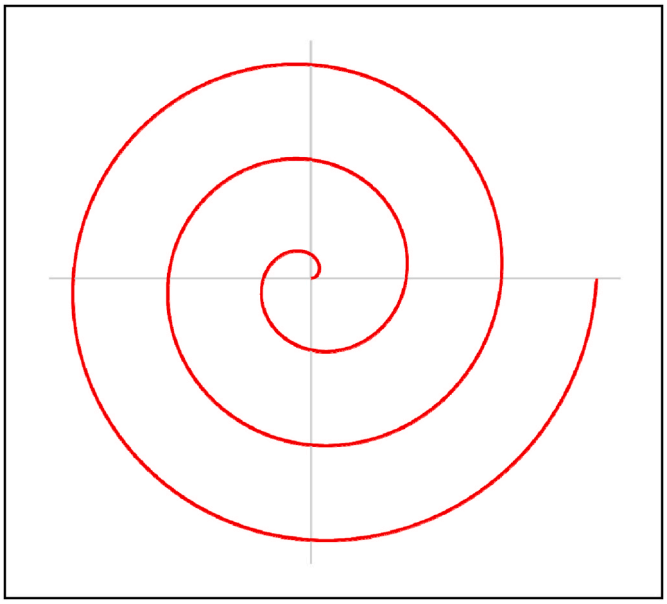

Fig. 1. Spiral drawing test template.

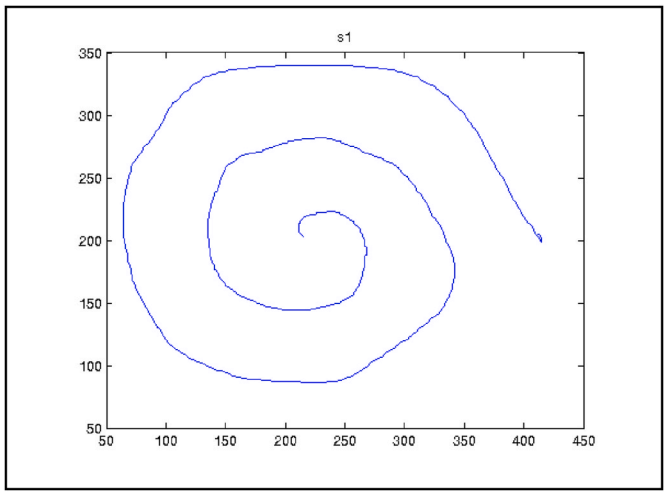

Fig. 2. Static spiral drawing (patient sample).

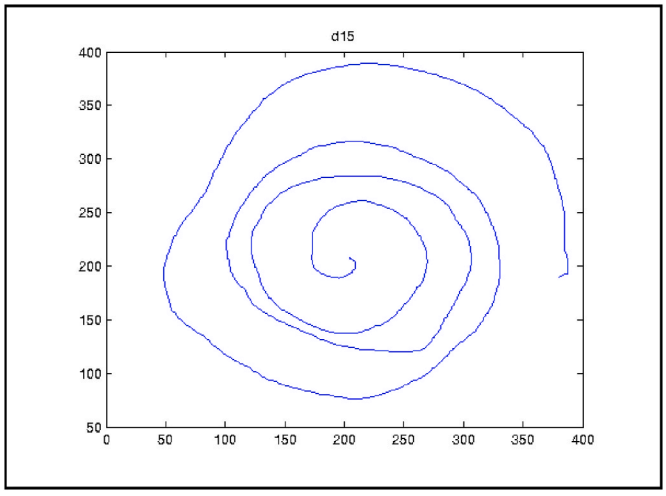

Fig. 3. Dynamic spiral drawing (patient sample).

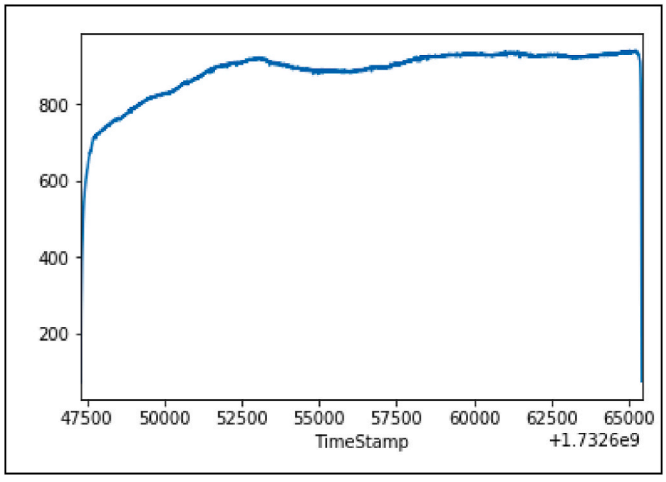

Fig. 4. Spiral Drawing Stroke – Healthy control.

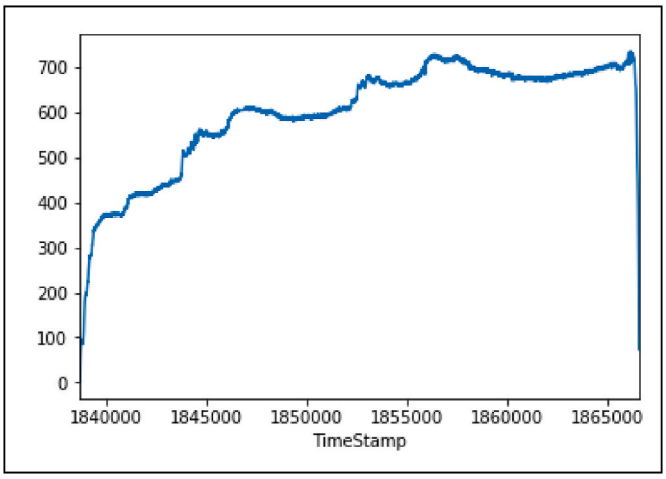

Fig. 5. Spiral Drawing Stroke – PD patient.

- Stroke velocity, acceleration, jerk

Velocity is a vector quantity; average velocity is defined as the displacement divided by the time interval. Displacement corresponds to the straight-line distance between two consecutive sampled points.

The Velocity at each spiral point is calculated as the square root of the sum of squares of consecutive x,y coordinates over the time difference between the points.

$$speed = \sqrt{(x_{i+1}\text{-}x_i)^2 + (y_{i+1}\text{-}y_i)^2} \div (t_{i+1}\text{-}t_i) \quad (1)$$

Based on this displacement, several kinematic features have been computed: number of strokes; horizontal and vertical displacement, velocity, acceleration and jerk; number of changes of velocity/acceleration (NCV/NCA), NCA and NCV relative to writing duration.

Average acceleration is the change in velocity divided by an elapsed time.

Jerk is the change in the acceleration over the time.

- o The horizontal and vertical components:
- o magnitudes of horizontal velocity
- o magnitude of vertical velocity
- o magnitude of horizontal acceleration
- o magnitude of vertical acceleration
- o magnitude of horizontal jerk
- o magnitude of vertical jerk

- magnitude of horizontal velocity(eq. (1)) is calculated by mean displacement in horizontal direction (change of X coordinate) over the total time duration
- magnitude of vertical velocity is calculated by mean displacement in vertical direction (change of Y coordinate) over the total time duration
- magnitude of horizontal acceleration is mean change of velocity in horizontal direction over the time.
- magnitude of vertical acceleration is mean change of velocity in vertical direction over the time.
- magnitude of horizontal jerk is mean change of horizontal acceleration over the time.
- magnitude of vertical jerk is mean change of vertical acceleration over the time.
- Overall new dataset consisted of 3 features(given by original dataset – pressure, grip angle, time) and 16 new features are engineered. 38 columns (19*2) for two spiral tests are added.

Our contribution in this feature engineering is addition of two kinematic features radial velocity and change in radial velocity with respect to time, which are helpful in diagnosis of hand motion and in turn hand tremor of PD patient.

The two features introduced are:

1. Radial Velocity: The radial velocity of an object with respect to a given point is the rate of change of the distance between the object and the point. Spiral drawn by patient or healthy control is analysed for this feature with reference to origin. Radial Velocity is calculated by dividing distance between two consecutive points (displacement) by signed angle between the ray ending at the origin and passing through the point (x,y) on spiral. Python numpy library function atan2 method returns a numeric value in radians representing angle between ray passing through (x,y) and positive x-axis.
2. Normalized Radial Velocity: Radial Velocity with time is change of radial velocity between any two consecutive points on spiral over the time elapsed.
3. Radial velocity and Radial Velocity with time are extended to mean value and standard deviation values.

2.2.2.2. Spatio-temporal features.

➢ Number of changes in the velocity: this is normalization feature, storing all those instances where velocity change is significant(ie more than zero), for static and dynamic spiral drawing test.

➢ Number of changes in the acceleration: this is normalization feature, storing all those instances where acceleartion change is significant(ie more than zero), for static and dynamic spiral drawing test.

➢ On-surface time is applicable for Static and dynamic spiral test where as in-air time is applicable for STCP(Stability test)

➢ STCP test component in-air time: Pressure exerted with pen (value less than 600, represents pen is in air and no surface pressure exerted). so data values less than 600 will be counted as in-air time.

➢ on-surface time for static spiral test: Pressure data values more than 600 are counted from csv file static spiral test drawings and normalized.

➢ on-surface time for dynamic spiral test: Pressure data values more than 600 are counted from csv file dynamic spiral test drawings and normalized.

The Table 1 shows the correlation factor of the number of features introduced in the literature [1,2] playing the role in diagnosis of PD, in the proposed models.

It also demonstrates that newly introduced kinematic features related to radial velocity and in-air time for stability test have significant contribution to accurate diagnosis of PD in early stage and they are also flexible for type of test. So any combination of drawing test can be useful for diagnosis of disease.

Table 1
Correlation factors.

Features	Correlation Factors
Number of strokes of Static Spiral test	0.197
Number of changes in velocity of Dynamic Spiral test	0.16
Number of changes in acceleration of Dynamic spiral test	0.11
Mean Radial Velocity (Static Test)	0.27
Standard Deviation Radial Velocity (Static Test)	0.35
Mean Radial Velocity (Dynamic Test)	0.25
Standard Deviation Radial Velocity (Dynamic Test)	0.30
In-air time for stability test	0.22

2.2.3. Model selection and parameter tuning

The algorithms selected for solving the classification problem are Logistic Regression and ensemble model Random Forest because of their high interpretability and their effectiveness to handle imbalanced dataset. The class weight in these models was adjusted to be inversely proportional to class frequencies in the input data. Some state-of-the-art supervised machine learning algorithms SVC C-Support Vector Classification and K – nearest neighbor Classification have also been employed.

2.2.3.1. Logistic regression. Logistic regression is machine learning classification algorithm used to predict the probability of categorical dependent variable. The dataset consists of target variable as binary variable and so containing only two values 1 or 0. Logistic regression model predicts probability of (target = 1) as function of independent variable. Prerequisite of logistic regression is binary dependent variable, independent variables are independent of each other or have very less correlation. The dataset is consisting of imbalanced classes, so ratio is 25:15. When too many categories are there, grouping of it is required. With the help of visualization, we can determine good predictor column from the dataset.

The performance of logistic regression can be tuned with its parameter solver. A solver is like a hypothesis function, h(x), that takes an input and gives us the estimated output value. This hypothesis function can be a one variable linear equation, or a very complicated and long multivariate equation depending on the type of the task. Among the large dataset. To find the best parameters that give us the least error(also named as cost or loss function) in predicting the output, our goal is to minimize it. Linear classification supports the logistic regression through solver term. Each solver tries to find the parameter weights that minimizes the cost. The solver uses a coordinate descent (CD) algorithm that solves optimization problems by successively performing approximate minimization along coordinate directions or coordinate hyperplanes. The default solver is Liblinear solver that applies automatic parameter selection. This is much slower, and with drawbacks of stucking up in data processing. The proposed work is implementing lbfgs solver which is much fasters and produces same accuracy. Solver 'lbfgs' handles multinomial loss, with L2 regularization with no penalty. In L2 regularization term is the sum of square of all feature weights and it forces the weights to be small but does not make them zero and so works on dense solution just like ridge regression. Lbfgs solver stores only the last few updates, so it also saves memory. Another advantages of lbfgs are it is relatively fast, and doesn't require similarly-scaled data. It's the best choice for most cases without a really large dataset. So the proposed work impiles this logistic regression with lbfgs solver.

2.2.3.2. Random Forest Classifier. Random Forest Classifier is an ensemble algorithm. *Ensembled algorithms* combine more than one algorithms of same or different kind for classifying objects and then take votes for final consideration of class for test object.

The base classifier of Random forest is decision tree. Random forest classifier creates a set of decision trees from randomly selected subset of training set. It then aggregates the votes from different decision trees to decide the final class of the test object. The proposed work is implementing Randomforestclassifier with all default values of Scikit-learner ensemble model, with 100 number of trees in the forest, gini impurity measure, minimum number of samples required to split the node is two and all the features of dataset are considered for classification.

2.2.3.3. SVC C-Support Vector Classification. Support Vectors Classifier tries to find the best hyperplane to separate the different classes by maximizing the distance between sample points and the hyperplane. The objective of a Support Vector Classifier is to fit to the data to a "best fit" hyperplane that classifies the data.C is the penalty parameter of the error term. It controls the trade off between smooth decision boundary and classifying the training points correctly. It works on Rbf kernel function which is representation of non linear hyperplane. it is commonly used in support vector machine classification. may be recognized as the squared Euclidean distance between the two feature vectors. The results of SVC also depend on gamma parameter for non linear hyperplanes. The higher the gamma value, the training data set is most fit the classifier. The proposed work is implementing default form of SVC from scikit-learn library.

2.2.3.4. K – nearest neighbor classification. This is basic classifier in supervised learning. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its *k* nearest neighbors.

The proposed work is implemented using k = 5, 5 Number of neighbors to use by default, with uniform weight, All points in each neighborhood are weighted equally. There are three possible algorithms brute force, KD tree and Ball tree to determine the neighbors. The proposed work is based on auto model, the algorithm is automatically chosen based on training dataset values.

2.2.4. Experimental setup

Python Scikit-Learn library models implemented are.

- logistic regression,
- randomforestclassifier
- SVC,
- Kneighborclassifier and compared them on the performance measures accuracy, precision, recall and F1 score.

3. Results and discussion

The proposed work has considered limited dataset which is contributed by Ref. [7]. The existing work has expressed the probability of using spiral drawing as effective tool for PD diagnosis. However, when machine learning and deep learning are the emerging techniques to be applied for healthcare sector, it requires significant execution and accurate results to demonstrate the technique.

This work has suggested machine learning methods for classification of spiral drawing images of healthy control and PD patient to aid the medical practitioner to diagnose the disease without physical intervention and at early stage.

Previous work has not implemented the classification of dataset and so the results obtained can not be compared with earlier work. However, to get better accuracy of the classification, the work compares the different algorithm for the limited dataset to come out with the significant classification outcome. The objective is to aid the medical practitioners and PD patients to get the status of the disease.

3.1. Metrics

tp:true positive, tn:true negative
fp:false positive, fn = false negative.

- Accuracy: Classification Accuracy is what we usually mean, when we use the term accuracy. It is the ratio of number of correct predictions to the total number of input samples.

$$Accuracy = Correctly\ classified\ instances / Total\ instances \quad \text{Eq.(2)}$$

- Precision: Also called as positive predictive value is the fraction of relevant instances among the retrieved instances. The precision is intuitively the ability of the classifier not to label a sample as positive if it is negative. High the value, better the classifier.

$$Precision = tp/(tp + tn + fp + fn) \quad \text{Eq.(3)}$$

- Sensitivity (Recall) is the fraction of the total amount of relevant instances that were actually retrieved.

The recall is intuitively the ability of the classifier to find all the positive samples. Higher the value higher is the accuracy.

$$\text{Recall} = \text{tp}/(\text{tp} + \text{fn}) \quad \text{Eq.(4)}$$

Specificity(also called True Negative Rate): Specificity relates to the classifier's ability to identify negative results.

- F1 score: The F1 Score is the 2*((precision*recall)/(precision + recall)).

It is also called the F Score or the F Measure. When it is F1, it conveys the balance between the precision and the recall.

$$\text{F1}=(2*(\text{tp}/(\text{tp} + \text{fp} + \text{tn} + \text{fn}))*(\text{tp}/(\text{tp} + \text{fn}))) / ((\text{tp}/(\text{tp} + \text{fp} + \text{tn} + \text{fn}))+(\text{tp}/(\text{tp} + \text{fn})) \quad \text{Eq.(5)}$$

It is observed that Ensemble model and SVC, Kneighborclassifier are producing comparable results with 100% accuracy and recall value.

Time of Execution is CPU time spent in execution of model and mean iteration time is statistically relevant and calculated by executing the model at least for 10000 iteration. Based on feature extraction mechanism, and parameter tuning of model this may vary for every turn.

The performance metric of precision and recall and F1score may give misleading results in case of imbalanced dataset, as both are given equal importance in F1 score. So for accurate validation of model, average precision value and precision recall Curve (PR curve) are also evaluated. Fig. 6 gives comparative analysis of the four models in PR curve plotted and Area under curve(AUC) demonstrates the model accuracy.

3.2. Discussion

The performance measures in Tables 2 and 3, are presenting the machine learning tools to classify the test data of spiral drawing, with binary classification of healthy and patient classes. These measures demonstrate that though there is limited dataset of spiral drawings (mix of static and dynamic), the patterns can be evaluated and classified. The results can not be compared with previous work, as this limited dataset is not analysed previously for classification.

Precision metric demonstrates rightness of model when it gives correct results where F1 score is harmonic mean of precision and recall. Higher the values of precision, recall and F1 score, the better is the validation of model.

Also the higher numerical value AUC for PR curve represent correctness of model for classification.

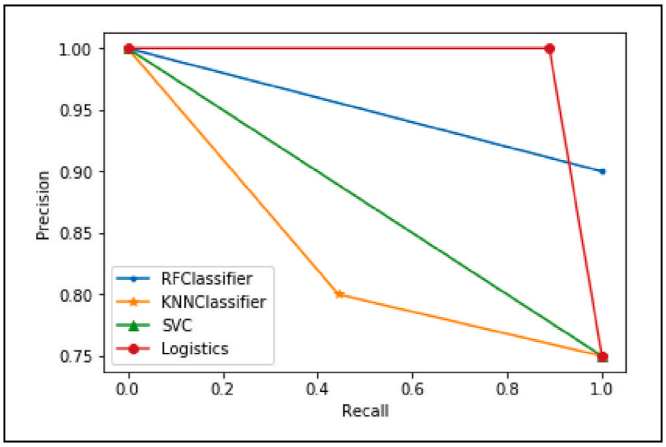

Fig. 6. Comparative analysis using PR curve.

Table 2
Performance metric I evaluation.

Model Applied	Precision	Recall	F1Score	Accuracy%
Logistic Regression	0.66	1.0	0.8	91.6
Random Forest Classifier	0.75	0.9	0.81	91.6
SVC	0.75	0.75	0.75	75.0
KNeighborsclassifier	0.33	0.8	0.47	50.0

Table 3
Performance metric II evaluation.

Model Applied	Time of Execution	Mean Iteration Time	Average Precision	AUC PR Curve
Logistic Regression	36 ms	851μs	0.97	0.98
Random Forest Classifier	852 ms	1.12 ms	0.9	0.95
SVC	4 ms	2.32 μs	0.75	0.87
KNeighborsclassifier	4 ms	3.32 μs	0.77	0.83

The proposed work is implemented on imbalanced dataset and training and test datasets are generated randomly. The results demonstrate better performance of logistic regression with relatively less time complexity as well. RF classifier also demonstrates better results but the computation time is high because of number of estimators, large number of features and optimum parameter values.

The previous work [1,2] have presented number of kinematic features, if used with ML algorithms can increase outliers. This work identified most correlated features, to increase the homogeneity of features to lead to conclusion.

AUC and Accuracy performance factors can be compared with two previous references and that will show better performance and validation of proposed ML algorithms. Table 4, compares the two common performance factors with existing works. The proposed work demonstrated better because most significant correlated features out of approx. 38 features are selected where as [12] has selected 10 features and [15] suggested CNN with inbuilt feature extraction.

The proposed work has contributed popular machine learning and emsemble learning methods and demonstrated logistic regression for given text dataset consisting of kinematic features. It can be compared with traditional naïve bayes algorithm which is probabilistic and other image features based cnn algorithm.Novelty in the work is identification of new most significant kinematic features that can support motor disorder and detection of PD at early stage without complex processing.

4. Conclusion

This paper confirms that three types of digitized spiral drawing tests have major impact on classification of PD patients and healthy controls, when four ML models are implemented on mathematically processed dataset. The results are taken out on small sized imbalanced dataset of 40 patients. The work outperformed with feature engineering, four ML algorithms with 91.6% accuracy and 98.1 AUC as compared to existing work on limited spiral drawing image set. Hence with the support of

Table 4
Comparative analysis.

Model Applied	Reference	AUC%	Accuracy%
Logistic Regression	This work	98.1	91.6
Naïve Bayes Classification	Zham P [12].	93.3	83
CNN	Manuel Gill m[15	96.5	90

extended dataset and extended computational model future diagnosis of PD can be done to support healthcare research for other neurodegenerative diseases.

CRediT authorship contribution statement

Megha Kamble: Conceptualization, Methodology. **Prashant Shrivastava:** Editing. **Megha Jain:** Results, Experimentation, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Drotar Peter, et al., Analysis of in-air movement in handwriting: a novel marker for Parkinson's disease, Comput. Methods Progr. Biomed. 117 (3) (December 2014) 405–411. August 2014. (references).
[2] Drotar Peter, et al., Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease", AI in Medicine 67 (2016) 39–46.
[3] Donato Impedovo, et al., Dynamic Handwriting Analysis for Supporting Earlier Parkinson's Disease Information 9, 2018, p. 247, https://doi.org/10.3390/info9100247, 10.
[4] D. Impedovo, G. Pirlo, Dynamic handwriting analysis for the assessment of neurodegenerative diseases: a pattern recognition perspective, in: IEEE Reviews in Biomedical Engineering, vol. 12, 2019, pp. 209–220, https://doi.org/10.1109/RBME.2018.2840679.
[5] Zham Poonam, et al., A kinematic study of progressive micrographia in Parkinson's disease, Front. Neurol. 10 (No.403) (2019) 1–8, https://doi.org/10.3389/fneur.2019.00403.
[6] a Omer Eskidere, et al., A Comparison of regression methods for remote tracking of Parkinson's disease progression, Experts Syst. Appl. 39 (2012) 5523–5528.
[7] Muhammed Erdem Isenkul, et al., Improved spiral test using digitized graphics tablet for monitoring Parkinson's disease, in: Proc. of ICEHTM, 2014, pp. 171–175.
[8] M. San Luciano, et al., Digitized spiral drawing: a possible biomarker for early Parkinson's disease, PloS One 11 (10) (2016), e0162799, https://doi.org/10.1371/journal.pone, 0162799.
[9] S. Aghanavesi, et al., Verification of a method for measuring Parkinson's disease related temporal irregularity in spiral drawings, Sensors 17 (10) (2017 Oct 13) E2341, https://doi.org/10.3390/s17102341.
[10] F. Miralles, et al., Quantification of the drawing of an Archimedes spiral through the analysis of its digitized picture, J. Neurosci. Methods 152 (2006) 18–31.
[11] R. Saunders-Pullman, C. Derby, K. Stanley, A. Floyd, S. Bressman, R.B. Lipton, et al., Validity of spiral analysis in early Parkinson's disease, Mov. Disord. 23 (4) (2008) 531–537, https://doi.org/10.1002/mds.21874.
[12] P. Zham, D.K. Kumar, P. Dabnichki, S. Poosapadi Arjunan, S. Raghav, Distinguishing different stages of Parkinson's disease using composite index of speed and pen-pressure of sketching a spiral, Front. Neurol. 8 (2017) 435, https://doi.org/10.3389/fneur.2017.00435.
[13] Mathew Thomas, et al., Handwriting Analysis in Parkinson's Disease: Current Status and Future Directions, in: Movement *Disorder*, vol. 4, Wiley, Dec 2017, pp. 806–818, 6.
[14] wt al Peter Drotar, Decision Support framework for Parkinson's disease based on novel handwriting marker, IEEE Trans. Neural Syst. Rehabil. Eng. 23 (NO. 3) (MAY 2015) 508–516.
[15] Manuel Gill Martin, et al., Parkinson's disease detection from Drawing movements using Convolutional Neural Networks, Electronics 8 (907) (2019) 1–10.