Early detection of Parkinson's Disease from Spiral and Wave Drawingsusing Image Processing and Machine Learning techniques

A.Sree Sai , M .Bhuvaneswari , Ch .Joshetha , K.Padmanjali , D.V.A.N. Ravi Kumar

Student, Student, Student,
Student, ProfessorElectronics and
Communication Engineering
Gayatri Vidya Parishad College of Engineering for Women, Visakhapatnam, India

Abstract -This paper develops a model that can recognize Parkinson's disease (PD) at its early stage. PD is acommon neurodegenerative disorder of the central nervous system that presents with progressive slow movement, tremor, limb rigidity, and gait alterations, including stooped posture, shuffling steps, freezing gait, and falling. There are many methods to detect the presence of pd disease. The proposed predictive analytics framework is a combination of histogram-oriented gradient (HOG) feature descriptor and machine learning technique which is used to gain insights from patients. Our proposed system provides accurate results by integrating spiral and wave drawing inputs of normal and Parkinson's affected patients. From thesedrawings, HOG feature descriptor extracts the features of input images and machine learning technique is used toclassify these images and predict which one is drawn by the person with pd disease with a high accuracy rate.

Keywords— Machine learning,HistogramOrientedGradients.

I. INTRODUCTION

Parkinson's disease (PD), or simply Parkinson's, is a chronic, degenerative condition that primarily affects the motor system of the central nervous system. Typically, symptoms appear gradually, and as the condition progresses, non-motor symptoms increase in frequency. Tremor, rigidity, slowness of movement, and trouble walking are the most noticeable early signs. Parkinson's disease can lead to delayed motor movements, anxiety, depression, problems sleeping, changes in behavior, and more. Among the main causes of Parkinson's disease are environmental influences and genetic inheritance. The loss of neurons that generate the chemical messenger dopamine in the brain causes Parkinson's disease. Parkinson's disease is caused by aberrant brain activity, which is brought on by a dropin the quantity of the amino acid dopamine.
Although the exact Aetiology of Parkinson's disease is still unknown, a number of factors, including genes, environmental factors, and triggers, seem to be involved.People experience the symptoms of this illness for many years prior to receiving a diagnosis. According to the estimated findings, Parkinson's disease affects 7–10 million people worldwide. The risk of developingParkinson's disease is highest in people over the age of 50. Our main objective is to use the HOG approach to assess the aesthetic appeal of these drawings created by the individuals and then train a machine learning model to categorize these drawings.

II. METHODOLOGY

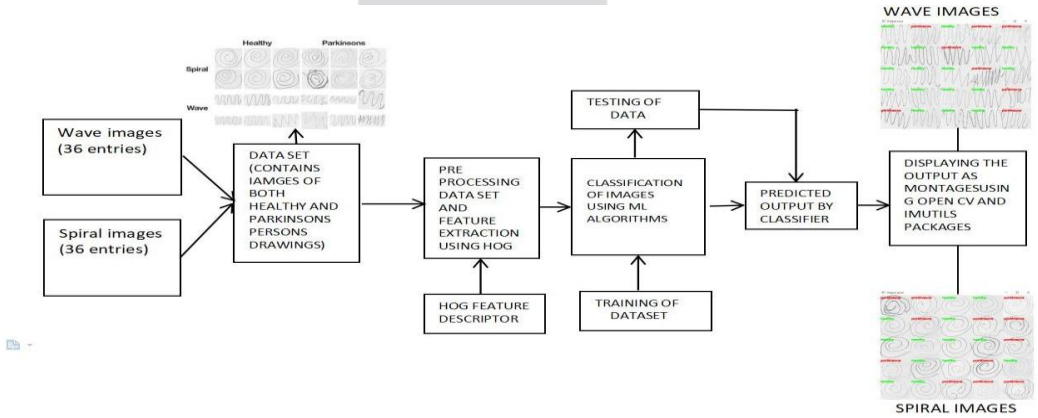


Figure 1: PROPOSED SYSTEM

The above figure shows the methodology we used, below is the explanation of the methodology.

Data Collection

The dataset used in here is the Parkinson's Drawing Dataset from Kaggle . The dataset includes Spiral and Waves drawings created by healthy and Parkinson's disease patients. The Train Set and Test Set are already included in the dataset as shown in the table1 below.

Image Type	No. of Images in the Training set	No. of Images in the Test set	Total
Wave	72	30	102
Spiral	72	30	102

Table 1: Data set

Data Pre-processing

Preprocessing pictures makes them better than they were originally by enhancing their quality. The aim of picture acquisition is to gather images that have less noise than HD images. The main benefit of this module is the higher clarity, reduced noise, and reduced distortion of the images. Segmentation's goal is to simplify or make an image representation easier to analyze.

Feature Extraction

In Our Project, we have used Histogram Oriented Gradient for Feature Extraction from Spiral and Wave Images. HOG works in five stages as:

Stage 1. Standardizing the image preceding explanation.

Stage 2. Figuring gradients in equally the vertical and horizontal direction.Stage 3. Getting weighted polls in spatial and orientation cells.

Stage 4. Difference between normalizing covering spatial cells.

Stage 5. Gathering all histograms of oriented gradients to shape the last component vector.

Machine learning Algorithms Description

For the prediction of Parkinson's disease there are various Machine learning techniques. But in our project, we have used four techniqueslike **SVM, RANDOM FOREST CLASSIFIER, XG BOOST, KNN** as they give accurate results compared with other techniques.

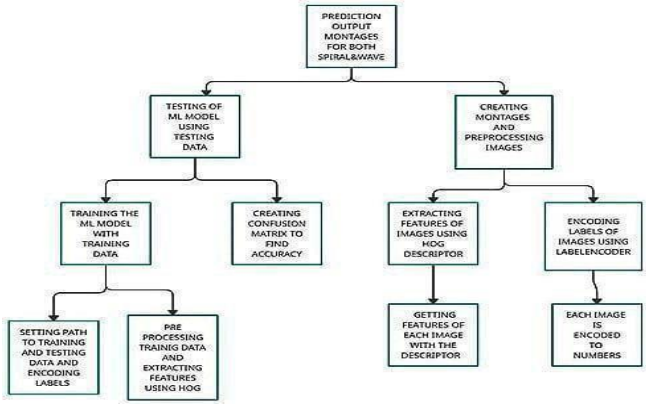


Figure 2: Content diagram of the project

Support vector machine(svm)

SVMs are utilized in web pages, intrusion detection, face identification, email categorization, gene classification, and handwriting recognition, among other applications. We utilize SVMs in machine learning for a number of reasons, including this. Both classification and regression on linear and non-linear data are supported.

Random forest classifier

In random forests, which are an ensemble version of Multiple choice bushes, each tree will narrow its emphasis to one particular capability while preserving a broad perspective of all other options. Every tree in the random area will do its own data training. To reduce the effects of characteristic connected with response, every tree will split its nodes.

XG Boost

A well-liked supervised-learning algorithm for regression and classification on sizable datasets is called XG Boost (extreme Gradient Boosting). To get reliable results, itemploys shallow decision trees that are formed sequentially and a highly scalable training technique that prevents overfitting**.**

K-Nearest Neighbour(KNN)

One of the simplest supervised machine learning techniques for classification is K-Nearest Neighbors. A data point is categorized based on the classifications of its neighbors. It archives all cases in its database and groups fresh cases according to characteristics in common.

Train and Test Data

The following step is to specify the training path and testing path after importing the necessary libraries. Bothhand-drawn spiral and wave patterns can be found in our dataset. Here, we train the model while taking spiralpatterns into account. The data was divided into training and test groups. The model is trained using the training data, and the results are predicted usingthe testing data as shown in the fig.2.

Evaluation

The proposed system's effectiveness is assessed using the following performance measures: classification accuracy, sensitivity, and specificity, denoted by the terms TP as true positive, FP as false positive, TN as true negative, and FN as false negative.

$$\text{Accuracy (\%)} = \frac{TP + TN}{TP + FP + TN + FN} x100$$

The percentage of correctly identified PD subjects is determined by the sensitivity or true positive rate.

$$\text{Sensitivity (\%)} = \frac{TP}{TP + FN} x100$$

The percentage of correctly identified healthy or non-PD subjects is determined by the specificity.

$$\text{Specificity (\%)} = \frac{TN}{FP + TN} x100$$

III. RESULTS AND DISCUSSION

Spiral drawing outputs:

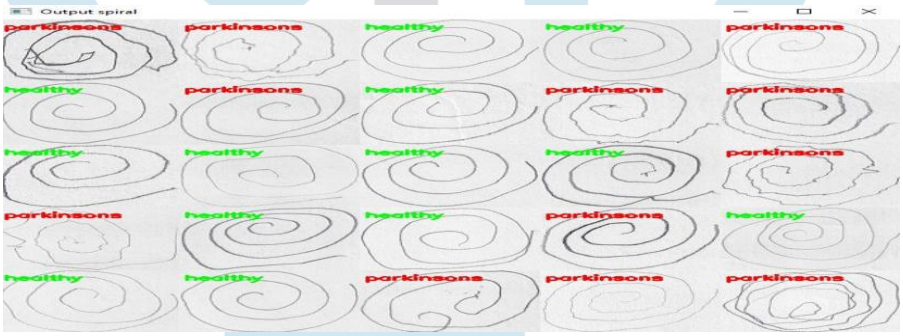
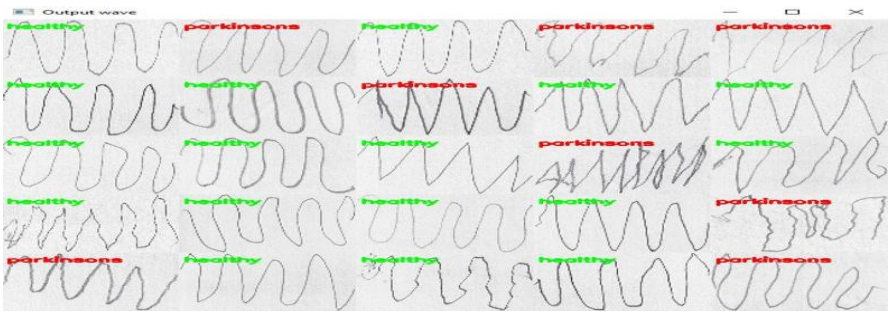


Figure3: Output for spiral dataset

Wave drawing outputs:



Comparision of results with base paper:

Figure 4: Output for wave dataset

MODEL	Wave						Spiral					
	Base paper			Obtained output			Base paper			Obtained output		
	accuracy	sensitivity	specificity	accuracy	sensitivity	specificity	accuracy	sensitivity	specificity	accuracy	sensitivity	specificity
RANDOM FOREST CLASSIFIER	71.33%	68.33%	73.33%	75.00%	83.33%	66.67%	83.33%	75.67%	90.33%	90.00%	86.67%	93.33%
SVM	73.33%	66.67%	80.00%	73.33%	73.33%	73.33%	70.00%	80.00%	60.00%	86.67%	80.00%	93.33%
XGBOOST	86.67%	93.33%	80.33%	70.00%	73.33%	66.67%	86.33%	65.00%	96.33%	73.33%	73.33%	73.33%
KNN	86.33%	89.67%	76.67%	76.67%	86.67%	66.67%	89.33%	91.67%	80.33%	80.00%	86.67%	73.33%

Table 2: Results comparison with base paper

Accuracies comparison for different machine learning techniques:

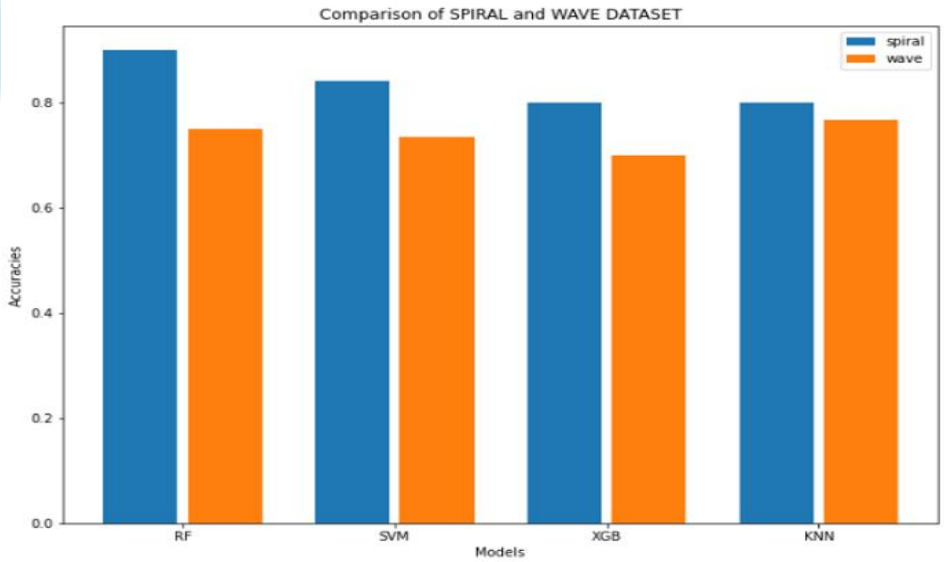


Table 2: Results comparison with base paper

Conclusion after comparing accuracies:

From the Table it is clear that Random Forest classifier has given the highest accuracy amongst all other machine learning algorithms.

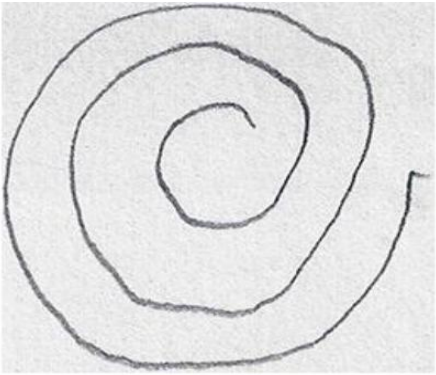
From figure it is clear that when compared to wave drawings, spiral drawings dataset has given more accurate results.

Hence, we can conclude by collecting spiral drawings from the patient and applying RandomForest classifier model we can determinewhether the person is affected by Parkinson's or not.

IV.INDIVIDUAL TESTING :

Based on the previous results the random forest classifier model is fixed at the obtained accuracylevel of 90% and is used for individual testing.

OUTPUT FOR HEALTHY PERSON:



```
clf = RandomForestClassifier(n_estimators=35,random_state=8) # Your pre-trained

# Load the color image
clf.fit(trainX, trainY)

image_path2="C:/Users/appa rao/Desktop/V01HE03.png"
try:
    image = cv2.imread(image_path2)
    if image is None:
        raise Exception("Error loading image")
except Exception as e:
    print("Error: " + str(e))
    exit()
# Convert the color image to grayscale
gray_image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
resized = cv2.resize(gray_image, (200,200))

image = cv2.threshold(resized, 0, 255,cv2.THRESH_BINARY_INV | cv2.THRESH_OTSU)[1
# Extract HOG features from the grayscale image
features = feature.hog(image, orientations=9, pixels_per_cell=(8, 8), cells_per_

prediction = clf.predict([features])

# Output the prediction result
if prediction == [1]:
    print("Positive for Parkinson's disease")
else:
    print("Negative for Parkinson's disease")
```

```
Negative for Parkinson's disease
```

V. CONCLUSION

Our paper is aimed to cover a broader space of imaging and machine learning technologies. Moreover, we emphasize the early detection of Parkinson's disease using hand drawn of spiral and wave images by using a Simple Algorithms with low cost, such that treatment and support can be provided to patients as soon as possible. We have implemented this using HOG and Random Forest Classifier, Support Vector Machine, XG Booster and K Nearest Neighbor. But the conclusion is that spiral images dataset with random forest classifier has given thehighest accuracy 90%.

The dataset taken up for consideration is examined and Were considered for utilization as per the requirements. The Random Forest algorithm whichworks on the classification technique can be used through the steps described to find Parkinson's disease. This can be a great help for the society to protect the patients in remote areas by offering the treatment in proper time, thus preventing mortality. The problem can be tackled in the best way possible. Finally, the fundamental idea of the work is to protect the people as best as possible and help the local governing bodies and hence, the society because a life without health is as good as a river without water and keeping our body healthy is our duty.

All the existing systems have focused only on a particular imaging modality such as MRI or PET, or on one specific type of dementia only such as AD, and can only detect Parkinson's disease at the later stages usually costs much. This paper aimedto cover a broader space of imaging and machine learning Technologies. Moreover, we emphasizethe early detection and prediction of Parkinson's disease using hand drawings by using simple algorithms with low cost, such that treatment and support can be provided to patients as soon as possible.

VI. ACKNOWLEDGEMENT:

We sincerely thank our project guide, Dr.D.V.A.N RAVI KUMAR Associate Professor,for his guidance and constant encouragement atevery stage of our main project at college. Heinstilled a spirit of understanding and support inus, which helped us carry out our projectsuccessfully. We would like to express our sincere thanks to Dr. P. M. K. Prasad, Head of the Department of Electronics and Communication Engineering, for his valuable suggestions and constant motivation that greatly aided us in completing our seminar work successfully. We express our sincere gratitude to our Vice Principal, professor Dr. G. Sudheer, for his encouragement and cooperation in the completion of our project. We wish to express our deep sense of gratitude to our Principal, Professor Dr. R.K. Goswami, for allowing us tocarry out our project work successfully.
We would also like to express our gratitude towards our parents and members of Gayatri Vidya Parishad College of Engineering for Women for their kind cooperation and encouragement, which helped us complete our project.

REFERENCES

[1]. Zhou, Yuan, Sule Tinaz, and Hemant D. Tagare. "Robust Bayesian analysis of early-stage Parkinson's disease progression using DaTscan images." IEEE transactions on medical imaging 40.2 (2020): 549-561.
[2]. A. Bhan, S. Kapoor, M. Gulati and A. Goyal, "Early Diagnosis of Parkinson's Disease in brain MRI using Deep Learning Algorithm," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV),2021, pp.1467- 1470, Doi: 10.1109/ICICV50876.2021.9388571.
[3]. Anitha r, Nandhini t, Sathish raj s, Nikitha v, "Early Detection of Parkinson's Disease Using Machine Learning" in International Journal of Advance Research and Innovative Ideas in Education (IJARIIE), e-ISSN: 2395-4396 Volume/Issue: Volume6 Issue2 2020 DUI:16.0415/IJARIIE-1159.
[4]. J. Rusz, M. Novotný, J. Hlavinka, T. Tykalová and E. Růžička, "High-Accuracy Voice-Based Classification Between Patients with Parkinson's Disease and Other Neurological Diseases May Be an Easy Task with Inappropriate Experimental Design," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 8, pp. 1319-1321, Aug. 2017, Doi: 10.1109/TNSRE.2016.2621885.
[5]. Ferdib-Al-Islam and L. Akter, "Early Identification of Parkinson's Disease from Hand-drawn Images using Histogram of Oriented Gradients and Machine Learning Techniques," 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE), 2020, pp. 1-6, Doi: 10.1109/ETCCE51779.2020.9350870.