Early detection of Parkinson’s Disease from Spiral and Wave Drawings using 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, Professor
Electronics 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 a common 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 thes drawings, HOG feature descriptor extracts the features of input images and machine learning technique is used to classify these images and predict which one is drawn by the person with PD disease with a high accuracy rate.

Keywords — Machine learning, Histogram Oriented Gradients.

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 drop in the quantity of the amino acid dopamine. Although the exact etiology 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 developing Parkinson'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](image-url)
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 table below.

<table>
<thead>
<tr>
<th>Image Type</th>
<th>No. of Images in the Training set</th>
<th>No. of Images in the Test set</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave</td>
<td>72</td>
<td>30</td>
<td>102</td>
</tr>
<tr>
<td>Spiral</td>
<td>72</td>
<td>30</td>
<td>102</td>
</tr>
</tbody>
</table>

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 techniques like SVM, RANDOM FOREST CLASSIFIER, XG BOOST, KNN as they give accurate results compared with other techniques.

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, it employs 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. Both hand-drawn spiral and wave patterns can be found in our dataset. Here, we train the model while taking spiral patterns into account. The data was divided into training and test groups. The model is trained using the training data, and the results are predicted using the 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.

\[
\text{Accuracy (\%)} = \frac{TP + TN}{TP + FP + TN + FN} \times 100
\]

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

\[
\text{Sensitivity (\%)} = \frac{TP}{TP + FN} \times 100
\]

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

\[
\text{Specificity (\%)} = \frac{TN}{FP + TN} \times 100
\]

III. RESULTS AND DISCUSSION
Spiral drawing outputs:

Wave drawing outputs:

Comparison of results with base paper:

[Figure 3: Output for spiral dataset]

[Figure 4: Output for wave dataset]
### Table 2: Results comparison with base paper

#### Accuracies comparison for different machine learning techniques:

<table>
<thead>
<tr>
<th>MODEL</th>
<th>Wave</th>
<th>Spiral</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base paper</td>
<td>Obtained output</td>
</tr>
<tr>
<td></td>
<td>accuracy</td>
<td>sensitivity</td>
</tr>
<tr>
<td>RANDOM FOREST CLASSIFIER</td>
<td>71. %</td>
<td>88. %</td>
</tr>
<tr>
<td>SVM</td>
<td>73. %</td>
<td>66. %</td>
</tr>
<tr>
<td>XGBOOST</td>
<td>86. %</td>
<td>93. %</td>
</tr>
<tr>
<td>KNN</td>
<td>86. %</td>
<td>89. %</td>
</tr>
</tbody>
</table>

### 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 determine whether 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 accuracy level of 90% and is used for individual testing.
OUTPUT FOR HEALTHY PERSON:

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 the highest accuracy 90%.
The dataset taken up for consideration is examined and were considered for utilization as per the requirements. The Random Forest algorithm which works 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 aimed to cover a broader space of imaging and machine learning Technologies. Moreover, we emphasize the 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 at every stage of our main project at college. He instilled a spirit of understanding and support in us, which helped us carry out our project successfully. 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 to carry out our project 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