

Detecting Parkinson's Disease Through Handwriting Using a CNN Model

The University of Science and Technology of Hanoi
Group 6

Nguyen Tran Minh Quan 22BI13375
Bui Thanh Vinh 22BI13469
Pham Tien Dat 22BI13080
Nguyen Quoc Khach 22BI13210

Machine Learning in Medicine
Lecturer: Assoc. Prof. Tran Giang Son

1 Introduction

Parkinson's Disease is a progressive disorder of the nervous system that primarily affects movement. One of its early symptoms is hand tremor, which often alters a person's ability to write or draw smoothly. As these motor changes can be subtle, handwriting analysis has been studied as a possible tool for early detection. In this project, we explored the use of a deep learning model to classify whether an individual is affected by Parkinson's Disease based on images of spiral and meander drawings. Our objective was to build a working model that could learn to identify visual differences between drawings made by healthy individuals and those with Parkinson's, contributing to research in non-invasive screening methods.

2 Dataset

The dataset we used comes from Kaggle and includes spiral and meander drawings collected from both Parkinson's patients and healthy individuals. These drawings were divided into four groups:

- Spiral images from healthy individuals (SpiralControl)

- Spiral images from Parkinson's patients (SpiralPatients)
- Meander images from healthy individuals (MeanderControl)
- Meander images from patients (MeanderPatients)

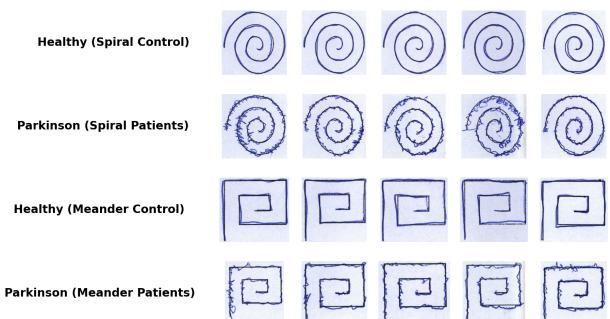


Figure 1: Display Sample

Here's a breakdown of the dataset:

Category	Number of Images
SpiralControl	73
SpiralPatients	297
MeanderControl	73
MeanderPatients	297

Here's the Distribution plot:

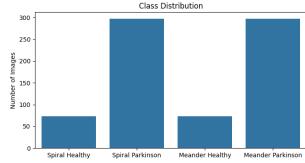


Figure 2: Class Distribution

As we can see, the dataset contains significantly more images from patients than from healthy individuals. This imbalance posed some challenges during model training and evaluation, particularly for the under-represented control classes.

3 Data Preprocessing

Before training our model, we performed several preprocessing steps:

- Verified the images and removed any corrupted files.
- Converted all images to grayscale to simplify the input and reduce computational load.
- Resized each image to 128x128 pixels for uniformity.
- Normalized the pixel values by scaling them to a 0–1 range.
- Assigned labels to each category (0 for healthy, 1 for Parkinson's) and encoded them accordingly.
- Split the dataset into training and test sets in an 80:20 ratio.

4 CNN Model Architecture

We built a simple Convolutional Neural Network (CNN) using TensorFlow and Keras. The architecture includes:

- Two convolutional layers with 32 and 64 filters, each followed by a ReLU activation and max pooling layer.
- A flattening layer and a fully connected dense layer with 128 units and a ReLU activation function.
- A dropout layer with a rate of 0.5 to prevent overfitting.
- An output layer using softmax activation to classify images into four categories.

We compiled the model using the Adam optimizer and trained it using the categorical cross-entropy loss function. The model was trained over 10 epochs with a batch size of 32.

5 Results and Evaluation

After training, the model achieved an accuracy of approximately 81.7% on the test set. Beyond accuracy, we evaluated performance using precision, recall, and F1-score:

Metric	Score
Accuracy	81.7%
Precision	82.0%
Recall	67.3%
F1-score	68.7%

In terms of class-wise performance, the model performed best on the meander images—both control and patient—with F1-scores above 0.80. The Spiral-Control category had the lowest performance, likely due to the small number of available training samples.

To visualize the results, we displayed a sample of test images along with their predicted labels, true labels, and the model's confidence levels. This helped

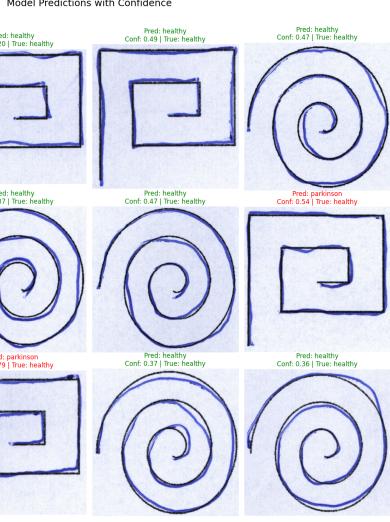


Figure 4: Model Prediction

us understand where the model was confident and accurate, and where it struggled.

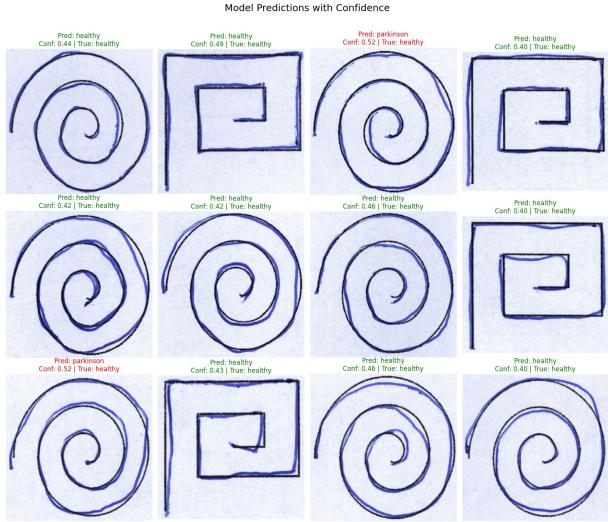


Figure 3: Confusion Matrix

Comparison with ResNet50V2

To further evaluate the effectiveness of our classification approach, we compared the performance of our simple CNN model with a more complex architecture—**ResNet50V2**. Both models were trained using the same image categories (spiral and meander drawings) obtained from Parkinson’s patients and healthy individuals. The classification report for ResNet50V2 is shown in Table 1.

Table 1: Classification Report of ResNet50V2

Class	Precision	Recall	F1-score	Support
Healthy	0.83	0.59	0.69	41
Parkinson	0.91	0.97	0.94	174
Accuracy				0.90 (215 samples)
Macro avg	0.87	0.78	0.81	215
Weighted avg	0.89	0.90	0.89	215

While the ResNet50V2 model achieved higher overall accuracy (90%) compared to the simple CNN (81.7%), a closer inspection reveals important caveats. Figure 5 shows a sample of the model’s predictions and confidence scores. Although many predictions appear confident, a number of healthy samples were misclassified as Parkinson’s with high confidence. This suggests that the model may be overfitting or failing to generalize subtle features distinguishing healthy patterns.

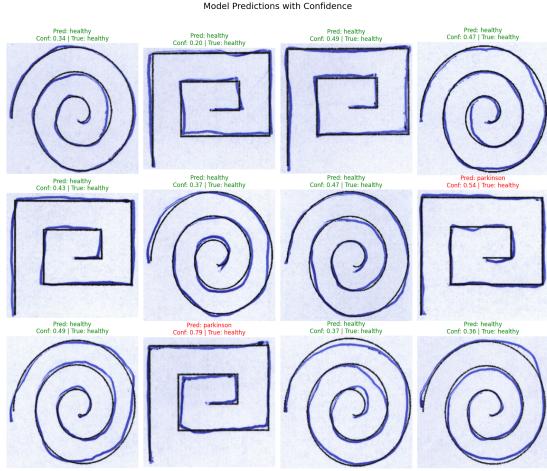


Figure 5: ResNet50V2 output predictions. Some healthy samples were misclassified with high confidence.

One possible explanation for this issue lies in the dataset handling process. The simple CNN model was trained on both spiral and meander datasets simultaneously, while keeping them as separate inputs. In contrast, ResNet50V2 was trained on a merged dataset, potentially losing class-specific spatial features. This merging may have reduced the model’s ability to distinguish between healthy and Parkinsonian patterns based on drawing type.

In conclusion, despite higher metrics, the ResNet50V2 model may not generalize as effectively in practice compared to the simpler CNN model, emphasizing the need for both qualitative and quantitative evaluations.

6 Discussion

The results show that our CNN model can identify Parkinson’s Disease from handwriting patterns with a decent level of accuracy. It worked especially well with meander drawings, which may contain more distinguishable features compared to spirals. However, the relatively small number of control samples—particularly spiral control drawings—limited the model’s ability to generalize well across all classes.

There’s room for improvement. In future work, we could balance the dataset by applying data augmentation techniques like rotation, scaling, and shifting. We also believe that using a more advanced CNN architecture or a pre-trained model (such as VGG or MobileNet) might lead to better performance. To increase transparency, it would also be useful to implement explainability techniques such as Grad-CAM to visualize which parts of the images the model focuses on when making decisions.

7 Conclusion

This project explored the use of convolutional neural networks to classify handwriting drawings from individuals with and without Parkinson’s Disease. We trained a simple CNN model on a public dataset of spiral and meander images and achieved a test accuracy of over 80%. While the results are promising, especially for meander drawings, the model’s performance can be further improved through data augmentation, transfer learning, and explainability tools. Overall, this work shows the potential of using handwriting analysis as a supporting tool in the early detection of Parkinson’s Disease.