On-Device Machine Learning for Diagnosis of Parkinson's Disease from Hand Drawn Artifacts

Sai Vaibhav Polisetti Venkata[1], Shubhankar Sabat[1], Chinmay Anand Deshpande[1],
Asiful Arefeen[2], Daniel Peterson[2], Hassan Ghasemzadeh[2]

***Abstract*—Diagnosis of neuro-degenerative diseases sometimes leads to inaccurate decisions and excessive medical costs. Deep learning models can help with diagnosis of such diseases like Parkinson's and aid in treating symptoms. This study introduces a novel system that integrates pervasive computing, mobile sensing, and machine learning to classify hand-drawn images and provide diagnostic insights for screening of Parkinson's disease patients. We carefully design a computational framework that combines data augmentation techniques with optimized convolutional network design for on-device and real-time image classification. We assess the performance of our system using images of Archimedean spirals drawn by hand and demonstrate that the proposed techniques achieve 76% accuracy and can run on an Android smartphone. Our study demonstrate that pervasive computing tools may offer an inexpensive and effective tool for diagnosis of Parkinson's disease.**

***Index Terms*—Machine learning, convolutional neural networks, Parkinson's disease, mobile sensing, pervasive computing.**

I. Introduction

Parkinson's disease (PD) is the second most prevalent neurodegenerative condition, surpassed only by Alzheimer's disease. Tremors are common, however, they are often accompanied by stiffness or slowed mobility. Depression, anxiety, and apathy are common in patients with Parkinson's disease, thus cognitive and behavioral problems may arise. However, PD indications and severity of symptoms are not consistent across different people and no two patients exhibit the same symptoms [1]. Therefore, early warning indicators could be subtle and go unnoticed by patients or clinicians. Early identification of symptoms and risk of PD could have important clinical implications. For example, tracking of subtle PD symptoms over time could assist clinicians with medication administration and prognosis, as well as clinical trials to quantify changes due to treatments.

A number of technologies have been developed to assist clinicians and scientists in identifying PD symptoms and tracking them in a highly quantitative way. Since PD affects different limbs or organs, recent researches have utilized a diverse set of modalities. For example, early works explored voice-based PD assessment methods [2]. Nakul et al. [3] compared various machine learning techniques for detecting Parkinson's disease and concentrated on patients' individual vocal traits since monitoring motor functions at an early stage is challenging. However, the sole focus was on the theoretical approach and not on deploying the solution on an embedded device to enable pervasive utilization of the technology. Gunduz et al. [4] employed a 9-layer convolutional neural network (CNN) and a parallel convolutional framework to classify PD patients. A predictive model was developed by Haq et al. using L1 Support Vector Machine to segregate PD patients from healthy ones [5]. While such approaches are powerful in their ability to assess many people remotely, voice-based PD assessment has limitations. For example, the sensitivity of this metric is modest, as there are many people who may have speech impairment and not necessarily PD [6]. Similarly, although inexpensive, variability in language may reduce the effectiveness of such voice-based assessments.

Other than voice-based PD assessment, emphasis has also been given on movement [7] and handwriting kinematics. For example, a common practice for PD diagnosis or quantization is through the use of wearable sensors. Inertial measurement unit (IMU) sensors for hand tremor quantization [8] and assessing movement abnormalities [9], ground reaction force (GRF) sensors for freezing gait analysis [10], surface EMG (sEMG) sensors for muscle movement assessment [11], and inertial sensors to quantify gait abnormalities [7] are frequently used in prior research. However these approaches also have limitations. For example, reliable data requires that 1) sensors are provided to participants, and 2) participants are able to don them effectively, and 3) algorithms are able to partition large amounts of data to identify the specific tasks of interest.

PD patients face abnormalities while pronouncing and writing as tremor interrupts and micrographia sets in. Hence, handwriting based PD assessment offers a pervasive solutions for PD assessment [2], [12]. This approach does not rely on deployment of expensive sensors, and is not impacted by language barriers and therefore partially circumvents some of the challenges noted above. Zham et al. [13] introduced Composite Index of Speed and Pen-pressure (CISP) of drawing as a feature for measuring the severity of Parkinson's disease. Participants drew Archimedean spirals and the association between drawing speed, pen-pressure, and CISP and disease severity of the subjects were calculated and examined. [14] prioritized hand writing movements to assist the detection of PD. The pressure exerted on the surface of the paper while writing contributed to the diagnosis of PD along with

This work was supported in part by the National Science Foundation, under grants CNS-2210133, CNS-2227002, IIS-1852163, and IIS-1954372.

[1]Sai Vaibhav Polisetti Venkata, Shubhankar Sabat, and Chinmay Anand Deshpande are Computer Engineering graduate students, Arizona State University, Tempe, AZ 85287, USA email: spolise6@asu.edu

[2]Asiful Arefeen, Daniel Peterson, and Hassan Ghasemzadeh are with the College of Health Solutions, Arizona State University, Phoenix, AZ 85004, USA

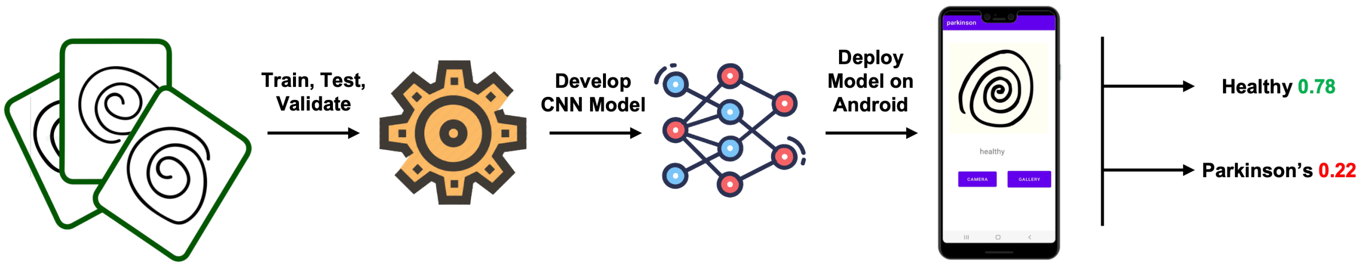


Fig. 1: Workflow of our proposed methodology. The system starts with preprocessing the hand drawn spirals and proceeds by training the CNN model. Next, the developed model is deployed on Android platform for doing analysis.

classifiers like Support Vector Machine, AdaBoost and K-Nearest Neighbours (kNN) and provided comparative analysis using a database with few samples. [15] proposed a framework with gradient boosting and kNN to detect PD patients from their hand drawn waves and spirals with high sensitivity. However, the feasibility of these techniques for real-time and on-device symptom assessment has not been demonstrated to date. Furthermore, we sense lack of effort to make the resources public and allow reproducibility for future aspirants.

In this paper, we propose a novel framework for reliable, robust, yet inexpensive approach for image analysis based PD diagnosis. All the experiments and results presented in this paper are fully reproducible with the code and data made publicly available. Although PD can be assessed in multiple ways, our focus is patients with hand tremor who face some abnormalities while drawing spirals. Our approach depends only on guided hand drawn spirals and guarantees accessibility from user end. Our approach is unique in a sense that it is not limited to theoretical aspects but implemented on smartphone for ease of usage and practicability.

II. System Design

A. System Overview

Once a hand drawn spiral dataset is available, as shown in Figure 4, our proposed method begins with preprocessing the data. Then we trained a CNN model with the preprocessed data to classify PD and healthy patients. We converted the model into TensorFlow Lite version to use it on smartphone device. Finally, we developed a smartphone application to get an interface and make it easily usable.

B. Preprocessing Image Data

1) Resizing: Input images can be of much higher resolution than needed for PD assessment. We propose to resize the images to a small uniform size (i.e., 180x180 pixels). This reduces the number of input features fed to the neural network. This also reduces the training time and computational complexity. Resizing causes some of the features to be diminished. However, this is not an issue in our case as the hand drawn spirals do not contain fine grained information.

2) Normalization: The pixel values of an R-G-B image range from 0 to 255. Different combinations of the three pixel values result in different colours, all three 0's being black and all three 255's being white. Scaling or normalizing the pixel values is a necessary step in preprocessing. The pixel values are scaled between 0 and 1 making sure equal contribution from each image towards the total loss of the model. This also ensures uniform learning rate for all images, faster convergence during training phase and the fact that every pixel has a similar distribution.

3) Data Augmentation: We have a neural network at the core of our system and training it with small number of image samples is difficult. This is the case in many real-world applications where collecting high quality and labeled training data remains an issue. To address this challenge, we perform data augmentation to increase our sample size. Data augmentation is used to increase the amount of data by slightly modifying the existing images and creating more samples. This helps reduce over-fitting when training and generalize better. As part of augmentation, we do zooming in/out, manipulate brightness and contrast, flip the images horizontally and vertically, and tilt the images by some degrees to expand our input dataset.

C. Neural Network Architecture

We used convolution neural networks (CNN or ConvNet) as it is the state of the art algorithm to classify and analyze visual data. CNN generally consists of an input layer, multiple hidden layers and an output layer. The hidden layers perform the actual Convolution by performing dot product of the kernel and the input image matrix. CNN also contains pooling layers which reduce the dimensions of data by combining outputs of neuron clusters into a single neuron in the next layer. The Softmax function is an activation function generally used in the final layer of the neural network to normalize the output of the network over predicted class values. It applies the standard exponential function to each element of the input vector and normalizes them.

The architecture of our model is explained below-

- In our solution, we used a Convolution Neural Network with 3 2D Convolution layers containing 32, 64 and 64 neurons respectively.
- We applied same padding with kernel size of 3x3 for convolution layers.
- ReLU activation function was used in all hidden layers.
- The final layer had 128 neurons.

- The output layer was of 2 neurons as it is a binary classification problem and used Softmax activation function with Adam optimizer.

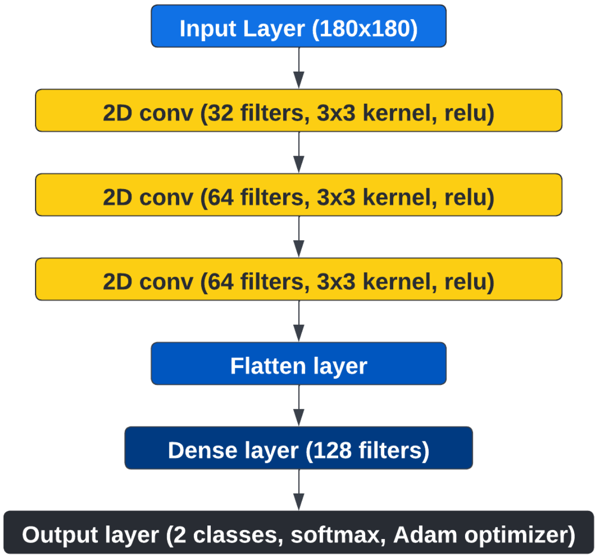


Fig. 2: CNN model architecture used in this study.

D. Android Application

The unique aspect of our project is designing an offline Android application to determine the subject's condition. The application takes image inputs (spirals drawn by the user) from both camera and gallery of the user. The machine learning model was converted into TensorFlow Lite format and deployed onto android platform. Our application was designed using Android Studio and Java programming language.

III. Validation

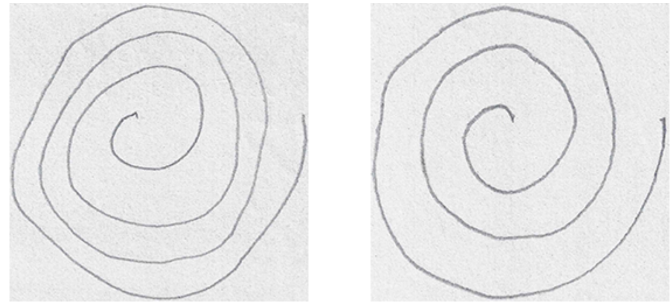
A. Dataset

We used publicly available dataset from [13]. It contains drawings of spirals and waves from 51 healthy people and 51 PD patients and available as 36 training and 15 testing samples under each class. This data set was originally collected with the measure of speed of hand and pressure applied on the surface while drawing. However, for our study, we used the hand drawn spirals, a subset of the entire dataset.

We used the PIL (Python Image Library) or "Pillow" [16] to perform image resizing. Through data augmentation, we expanded our data set to 718 samples from the original size of 102 samples. The data-set was split into 503 samples of training data and 215 samples of testing data.

B. Performance

We demonstrate our results with different models and under different setups. For better generalization and handling the small dataset, we performed image augmentation and as shown in Table I, accuracy increased by almost 21% when we used CNN with the augmented images. Although augmentation helped improve the performance of kNN model as well, the accuracy was capped at 68.3%.



(a) Parkinson's patient (b) Healthy

Fig. 3: Sample images from the dataset

Confusion matrix for the CNN model has been depicted in Figure 4 highlighting the high rates of true positives and true negatives. Also, note that F1 scores are around 0.9 which implies that the model generalizes well for both cases and controls.

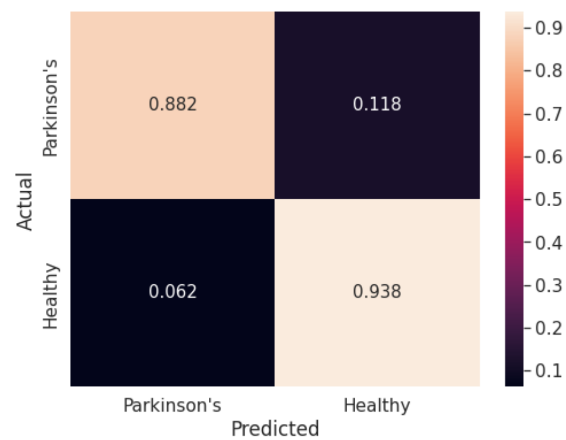


Fig. 4: Confusion matrix for PD screening using hand drawn spirals. **F1 scores** for Parkinson's and Healthy categories are **0.92** and **0.89** respectively.

TABLE I: Accuracy comparison for our approach with and without image augmentation. Accuracy also shown for kNN based model for comparison.

Model type		Without Augmentation	With Augmentation
CNN	Training Accuracy	74.10%	90.57%
	Testing Accuracy	55.17%	76.10%
kNN	Testing Accuracy	61.29%	68.3%

IV. Conclusions

We proposed a framework to screen PD patients who have difficulty in drawing spirals due to hand tremor. Eventually, we built a CNN model, a smartphone application and deployed them on Android platform. We assessed our system using accuracy and F1 score. Due to constraints in the dataset, we could not estimate the severity of PD which is a limitation of our approach. Furthermore, we could have tried different techniques for model optimization and develop lighter model for low latency application without compromising the performance. All these improvements are left to explore in future.

Acknowledgment

References

[1] "Understanding Parkinson's Disease," https://www.ohsu.edu/brain-institute/understanding-parkinsons-disease, 2022, [Online; accessed 6-April-2022].

[2] H. Zhang, A. Wang, D. Li, and W. Xu, "Deepvoice: A voiceprint-based mobile health framework for parkinson's disease identification," *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pp. 214–217, 2018.

[3] Y. Nakul, A. Gupta, and H. Sachdeva, "Parkinson disease detection using machine learning algorithms," 2021.

[4] H. Gunduz, "Deep learning-based parkinson's disease classification using vocal feature sets," *IEEE Access*, vol. 7, pp. 115 540–115 551, 2019.

[5] A. U. Haq, J. P. Li, M. H. Memon, J. Khan, A. Malik, T. Ahmad, A. Ali, S. Nazir, I. AHAD, and M. Shahid, "Feature selection based on l1-norm support vector machine and effective recognition system for parkinson's disease using voice recordings," *IEEE Access*, vol. 7, pp. 37 718–37 734, 2019.

[6] C. Moreau and S. Pinto, "Misconceptions about speech impairment in parkinson's disease," *Movement Disorders*, vol. 34, 2019.

[7] I. E. Maâchi, G.-A. Bilodeau, and W. Bouachir, "Deep 1d-convnet for accurate parkinson disease detection and severity prediction from gait," *Expert Syst. Appl.*, vol. 143, 2020.

[8] H. S. Jeon, W. Lee, H. Park, H. J. Lee, S. K. Kim, H. Kim, B. S. Jeon, and K. S. Park, "Automatic classification of tremor severity in parkinson's disease using a wearable device," *Sensors (Basel, Switzerland)*, vol. 17, 2017.

[9] N. M. Rad, T. van Laarhoven, C. Furlanello, and E. Marchiori, "Novelty detection using deep normative modeling for imu-based abnormal movement monitoring in parkinson's disease and autism spectrum disorders," *Sensors (Basel, Switzerland)*, vol. 18, 2018.

[10] A. Marcante, R. D. Marco, G. Gentile, C. Pellicano, F. Assogna, F. E. Pontieri, G. Spalletta, L. Macchiusi, D. Gatsios, A. Giannakis, M. Chondrogiorgi, S. Konitsiotis, D. I. Fotiadis, and A. Antonini, "Foot pressure wearable sensors for freezing of gait detection in parkinson's disease," *Sensors (Basel, Switzerland)*, vol. 21, 2021.

[11] A. M. Baraka, H. A. Shaban, M. A. El-Nasr, and O. Attallah, "Wearable accelerometer and semg-based upper limb bsn for tele-rehabilitation," *Applied Sciences*, 2019.

[12] N. Zhi, B. K. Jaeger, A. Gouldstone, R. Sipahi, and S. Frank, "Toward monitoring parkinson's through analysis of static handwriting samples: A quantitative analytical framework," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, pp. 488–495, 2017.

[13] P. Zham, D. K. Kumar, P. Dabnichki, S. P. Arjunan, and S. Raghav, "Distinguishing different stages of parkinson's disease using composite index of speed and pen-pressure of sketching a spiral," *Frontiers in Neurology*, vol. 8, 2017.

[14] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smekal, and M. Faúndez-Zanuy, "Evaluation of handwriting kinematics and pressure for differential diagnosis of parkinson's disease," *Artificial intelligence in medicine*, vol. 67, pp. 39–46, 2016.

[15] 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)*, pp. 1–6, 2020.

[16] A. Clark, "Pillow (pil fork) documentation," 2015. [Online]. Available: https://buildmedia.readthedocs.org/media/pdf/pillow/latest/pillow.pdf