FEATURE RELEVANCE ANALYSIS FOR HANDWRITING BASED IDENTIFICATION OF PARKINSON'S DISEASE

SAMAN KHAWAR

01-249192-014

A thesis submitted in fulfilment of the requirements for the award of degree of Masters of Science (Data Science)

Department of Computer Science

BAHRIA UNIVERSITY ISLAMABAD

SEPTEMBER 2021

Thesis Completion Certificate

Scholar Name: SAMAN KHAWAR

Registration Number: 63096

Enrollment: 01-249192-014

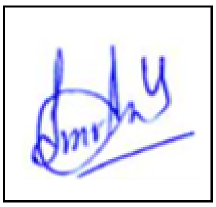
Program of Study: MS(DS)

Thesis Title: FEATURE RELEVANCE ANALYSIS FOR HANDWRITING BASED IDENTIFICATION OF PARKINSON'S DISEASE

It is to certify that the above scholar's thesis has been completed to my satisfaction and, to my belief, its standard is appropriate for submission for examination. I have also conducted plagiarism test of this thesis using HEC prescribed software and found similarity index **17%**. that is within the permissible limit set by the HEC for the MS/M.Phil degree thesis. I have also found the thesis in a format recognized by the BU for the MS/M.Phil thesis.

Principal Supervisor Name: Dr. Imran Siddiqi

Principal Supervisor Signature:



Date: 02 September 2021

Author's Declaration

I, **Saman Khawar** hereby state that my MS/M.Phil thesis titled is my own work and has not been submitted previously by me for taking any degree from Bahria University or anywhere else in the country/world. At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw/cancel my MS/M.Phil degree.

Name of Scholar: Saman Khawar

Date: 2 September 2021

Plagiarism Undertaking

I, solemnly declare that research work presented in the thesis titled **"Feature Relevance Analysis for Handwriting Based Identification of Parkinson's Disease"** is solely my research work with no significant contribution from any other person. Small contribution / help wherever taken has been duly acknowledged and that complete thesis has been written by me. I understand the zero tolerance policy of the HEC and Bahria University towards plagiarism. Therefore I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred / cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS/M.Phil degree, the university reserves the right to withdraw / revoke my MS/M.Phil degree and that HEC and the University has the right to publish my name on the HEC / University website on which names of scholars are placed who submitted plagiarized thesis.

Name of Scholar: Saman Khawar

Date: 2 September 2021

Dedication

I dedicate my thesis work to my supervisor,my respected teachers and my mother.

Acknowledgements

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main thesis supervisor, Professor Dr. Imran Siddiqi, for encouragement, guidance, critics and friendship.I am also very thankful to my co-supervisors Mam. Momina Moetesum for her guidance, advices and motivation. Without their continued support and interest, this thesis would not have been the same as presented here.

Librarians at Bahria University also deserve special thanks for their assistance in supplying the relevant literature. My fellow postgraduate students should also be recognised for their support. My sincere appreciation also extends to all my colleagues and others who have provided assistance at various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family members.

Abstract

Parkinson's disease is a severe neurodegenerative disorder that impairs the motor system over time, causing the slowness of speech and movements, as well as abnormal writing abilities due to tremors. Parkinson's patients are not suitable for all types of PD diagnosis tests due to their physical problems. As a result, a handwriting test can be used to construct an automated diagnostic tool as a potential marker. While traditional techniques focused on the effectiveness of online and offline or combining both features of handwriting from established templates characterizing the presence and absence of PD. In this study, we use the PaHaW dataset to carry out a comprehensive study to assess the optimal set of features that are more informative as a function of the templates from which they are extracted. For this purpose, We extract online and offline features subjects,combined extracted features and employed a feature selection mechanism , such as a genetic algorithm and correlation, to find the most relevant features that describe the presence and absence of PD by achieving an overall accuracy of 77.46% .

TABLE OF CONTENTS

LIST OF TABLE

LIST OF FIGURE

List of Acronym

PD	Parkinson's Disease
CNN	Convolution Neural Network
CGP	Cartesian Genetic Programming
GA	Genetic Algorithm
RNN	Recurrent Neural Network
SVM	Support Vector Machine

CHAPTER 1

INTRODUCTION

Parkinson's disease is the second most common neurological disorder after Alzheimer's [4]. Parkinson's disease (PD) affects around 10 million persons globally [5].PD affects 1–2 per 1,000 of the population [6, 7].PD affects 1% of the population over the age of 60, but is uncommon in people under the age of 50 [7, 8]. The prevalence of Parkinson's disease rises with age, reaching around 4% in the oldest age groups [8, 9]. These prevalence rates are expected to any increase because of the population[10]. Parkinson Disease (PD) is characterized by motor symptoms and non-motor symptoms including akinesia, bradykinesia, rigidity, and tremor, postural imbalance and vocal disabilities [11, 12]. Traditional diagnostic procedures for the diagnosis of PD include neuroimaging strategies such as SPECT and CT scans, shown in Figure 1.1, which shows vital potential within the determination of PD however needs expensive instrumentality. Furthermore, these strategies are compelling only when the disease has progressed to the final stage, further highlighting the complexities of PD analysis.[13].According to clinicopathological research [14, 10],up to 25% of PD patients are misdiagnosed within last stages of their illness. As a result, there's a lot of work being done to develop accurate systems for detecting and diagnosing Parkinson's disease in its early stages.

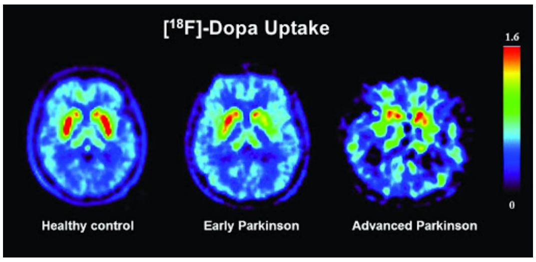


Figure 1.1: SPECT scanning of PD patient.[1]

With the advancement of technology, researchers are able to propose many solutions and decision support systems to identify the early stage of PD patients. Some of the studies [15] used signal acquisition through wearable sensors monitoring free muscular movements to predict PD, while other studies [16] used breath or voice analysis [17, 18, 19, 20] to predict PD. Voice processing for diagnosis of PD offered very promising results by achieving 98% overall classification accuracy [21] Likewise, Bradykinesia(slowness of movement), in the literature is directly related to handwriting. Some of the recent studies [22, 23] recommended that handwriting is often used as a good tool for early diagnosing of PD and a few preliminary pieces of knowledge suggest that handwriting would possibly function as a diagnostic marker for PD diagnosis by identifying micrographia. The idea is illustrated in Figure 1.2 where a PD subject attempts to write a sentence, over the period of time, handwriting starts deteriorating. Initially, the size of the letters and horizontal alignment are fine. However, it becomes hard for the PD subject to maintain the size and alignment of words, and the words at the end are almost impossible to read.

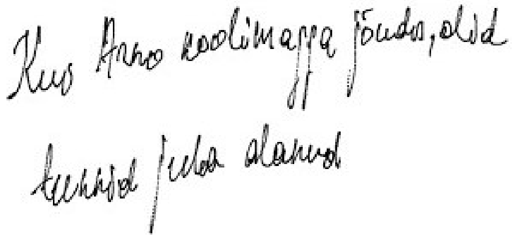


Figure 1.2: Handwriting of patient suffering with Micrographia.

Tremors damage handwriting because the involuntary oscillating movement of one or more body parts of the patient, as depicted in Figure.1.3, causes the hands or fingers to twitch slightly while the patient writes or draws something.

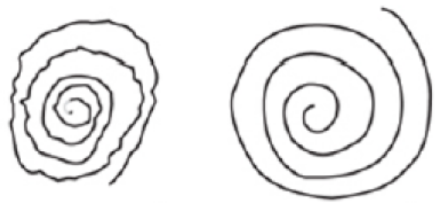


Figure 1.3: Subject with Parkinson's disease.

Another symptom known as Bradykinesia is in which the patient's handwriting speed is slow and the graphomotor task takes longer time than usual. Some researchers collected data using gadgets (digitizers or tablets), while others used hand-drawing shapes to come up with solutions and hypotheses for their study. Several preliminary studies have suggested that handwriting can be used as an effective non-invasive tool for the early diagnosis of PD. So, based on these considerations we have attempted to develop a system that is specifically designed for early detection of Parkinson's disease. As a result, we'll use Drotar et al dataset that is intended for Parkinson's disease patients.

1.1 Problem Statement

Identification of Parkinson's disease through modalities like handwriting or speech has been thoroughly investigated in the literature. The correlation between PD and changes in writing patterns has also been established in a number of studies [24, 13]. From the perspective of handwriting analysis, a number of static (offline) [13] and dynamic (online) [24] features have been identified that can serve as effective indicators of PD. Combining online and offline features are also known to improve the identification performance [24].

In most cases, the identified attributes are mapped to computational features which are extracted from established templates and are fed to a classifier to determine the presence or absence of PD. An important factor in choosing the type of features is the drawing or handwriting template under study. While the previous studies primarily target combining features or decisions on multiple templates to enhance the overall performance, to the best of our knowledge, no investigations have been carried out to study the relationship between the template under study and the extracted features.

The proposed research is aimed at feature relevance analysis for the identification of PD through handwriting. More specifically, we intend to carry out a comprehensive study using different feature selection techniques to assess the optimal set of features for this problem. Furthermore, some features may be more appropriate with specific templates hence we also aim to study which template performance is better =on optimal set of features.

1.2 Research Objectives

The objectives of Research include the following.

- To combine the online and offline attributes of writing and study the system performance
- To study the relevance of both static and dynamic features of handwriting in identification of PD.
- To investigate the performance of template and identify the relevant set of features for a given template.

1.3 Research Contributions

The research carried out in this study has resulted in the design and development of a a system that predicts Parkinson disease through computerized analysis of handwriting The main purpose of the proposed study is to design and development a system that can predict Parkinson's disease by use of computerized handwriting analysis. The key contribution of the research is the manipulation of the offline and online features to identify a relevant set of features that can predict the absence and presence of PD. In case of no availability of specialized hardware devices to directly capture online handwriting, offline attributes can be useful Features extracted from control subjects and PD patients are fed to feature selection techniques to assess the optimal set of features, and that optimal set of features is then fed to a learning algorithm to learn to discriminate between the two classes. Support Vector Machine (SVM) classifier is investigated for this purpose. Experiments on a benchmark dataset report promising classification rates.

1.4 Thesis Organization

This document is organized as follows. Chapter 2 presents a discussion on the work related to prediction of Parkinson disease from handwriting Chapter 3 describes the method that we have adopted in order to achieve the objectives along with the key concepts behind the approaches. Chapter 4 outlines the metrics used to test our methods, describes the experiments, presents the findings we obtained and their interpretation. Chapter 5 incorporates the concluding remarks and recommendations for future work.

CHAPTER 2

LITERATURE REVIEW

Parkinson's disease is caused by the loss of pigmented neurons in the mid-brain region's substantia nigra, which control muscle movements. Dopamine, a neurotransmitter involved in the control and regulation of body movements, is reduced when these neurons are lost. This causes tremors, sluggish movements, hypertonia, and balance issues. [25] These symptoms have an effect on the individual's hand-wrist movements, which have a negative impact on his or her handwriting. Computer-aided handwriting analysis allows for the identification of prospective patterns that may be useful in the detection and classification of Parkinson's disease.Several studies[22, 26] have been published that indicate handwriting the analysis is an effective tool for PD diagnosis. Many handwriting features were proposed in the in the writing for the identification of PD[13, 26, 27, 28]. Based on their technique of knowledge acquisition, extracted features can be classified into two types: Static and Dynamic. Static features will be taken from offline handwriting samples, whilst dynamic features will be derived from online handwriting samples. These studies used a variety of machine learning techniques to examine the static and dynamic features' ability to discriminate PD. In this chapter, we will discuss related work on handwriting analysis and potential strategies used for early Parkinson's prediction.

2.1 Online Features Analysis

Handwriting requires the participation of various body parts such as fingers, arms and also includes our motor neurons, a healthy person manages the participation of all parts for the writing task, however when we perform a writing task to the patient, the motor neurons do not function properly.A number of solutions for detecting Parkinson's disease and other similar disorders have been developed in recent years, one of which is wearable sensors that are attached to the patient's body. In 2011,[29], they integrated their device with smart gloves, which detected the level of motor dysfunction in PD using smart gloves and assessed the movement of fingers while writing, making non-invasive approaches more effective and less expensive.

In 2013, Dortar et al. [26] created a handwriting based dataset, which acquires handwritten signals (on-surface and in-air) using a digitizing tablet Intuos 4M and presented a template consisting of seven completely different handwriting tasks with an addition Archimedean spiral drawing task .In this study, they evaluated three types of features, i.e supported in-air movement, primarily based on-surface movement and combination of both groups of features to effectively diagnose PD. 75 samples in which 38 patients and 37 healthy subjects were employed. They applied classification using SVM (Support vector machine) and achieved a classification accuracy of 80% using 16 features selected from in-air movement. In 2014, [2] Drotar et al., combined various online in-air and on-surface features by using feature selection techniques and a support vector machine learning classifier to discriminate PD patients from healthy controls, attaining an accuracy of 85%. In a subsequent study [10], The authors achieved 88.13% accuracy using the SVM classifier with radial kernel for automated diagnosis, working with kinematic and spatio-temporal handwriting measures as well as handwriting measures including entropy, signal energy, and empirical mode decomposition. In 2015, authors extended the similar work [30], by using a combination of dynamic features and achieving 89% area under the ROC curve (AUC) for PD classification. In 2016, authors [31], used online kinematic and pressure features of handwriting to train different classifiers and achieving 81.3% accuracy with SVM, 78.9% with AdaBoost classifier, and 71% with KNN, respectively. In the sequence of experiments, authors additionally prompt that performance of identification of PD depends on the selection of template used. After a year Pereira at al [32], have acquired the NewHandPd dataset, which includes both off-line images and on-line signals (extracted from the smart pen). Each person fills the structure by composing on paper with a digitising pen and drawing four spirals and four meanders. Diverse machine learning algorithms, such as Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Optimal Path Finder (OPF), Random Forest (RF), and Restricted Boltzman Machines (RBM), were evaluated, revealing that when on-line data is used, the CNN ImageNet design could achieve precision of 87.14 % in its best setup, whereas the SVM achieved the highest performance on off-line data, with an accuracy of 66.72%

Researchers have not only used handwriting features to classify PD, but they have also used posture features. In 2014, Graca et al. [3] used mobile devices to predict PD by completing a different tasks (spiral analysis, gait analysis, tip analysis) in which 35 samples were collected from drawing Archimedes spiral. They extracted various features (Spatio-temporal , pressure features and gait features) and fed these features to the different classifier such as C4.5, RipperK and Bayesian network for classification. The reported accuracy with mentioned classifiers were 86.67%, 80.83% and 87.50% respectively. In 2018, Impedovo et al. [11] used PaHaW [26] database to investigate those specific dynamic features of the handwriting that can help to idenify people suffering from PD . They worked on online kinematic features were employed six classifiers including SVM (RBF, linear), KNN, LDA (Linear Discriminant Analysis), NB (Gaussian Naiïve Bayes), RF (Random Forest), ADA (AdaBoost) reporting accuracy's of 71%, 68%, 67.90%, 66%, 57%, 73%, 61% on 72 subjects, respectively.

In another valuable research by Angelillo et al. [33] in 2019, the researcher retrieved features from the raw data of different tasks using the PaHaW dataset, which comprises many tasks done by similar subjects, by utilising the dynamics of the handwriting process. Techniques such as Shannon and Renyi entropy, signal-to-noise ratio, and empirical mode decomposition (EMD) were used to figure both on-surface and in-air horizontal and vertical parts of handwriting.After extracting features, the prescient capability of every task is assessed exclusively and the best tasks, i.e. those with the most noteworthy forecast, are fed into a group of classifiers(SVM, AdaBoost, Logistic Regression, Linear Discriminant), whose predictions are obtained via majority voting and its achieved highest classification accuracy of 88.33%.In 2019 [34] Cartesian Genetic Programming is a technique for detecting Parkinson's disease (PD) by analysing the handwriting of PD patients and healthy controls. The adoption of such an approach is particularly intriguing because it allows for the inference of explicit classification models while also allowing for the automatic identification of an appropriate subset of features relevant for a correct diagnosis. The approach was tested using characteristics collected from handwriting examples in the PaHaW dataset, which is freely available. In 2020,Ammour et al. [35] worked on the Arabic Handwriting dataset and extracted the number of features of different categories like Kinematics on surface In-air, Mechanical, Inclination, Pen Up features and used the semi-supervised approach for classification(Clustering and PCA) obtaining 97.3% of classification accuracy. In same year Amina Naseer et al. [20] worked on PaHaW dataset and performed features extracted via CNN- Alexnet pre-trained model.The selected features were fed to SVM classifier for PD identification and obtained 98.28% of accuracy.In the same year, another research by Ismail Canturk [36] used the Fuzzy recurrence plot (FRP) approach to convert time-series signals into grayscale surface graphics. For attribute extraction, these FRPS were fed into two pre-trained deep learning algorithms (AlexNet and GoogleNet). These collected attributes were passed into k-NN and SVM classifiers, yielding a 94% promising outcome.

2.2 Offline Features Analysis

In other studies, certain authors did not used any dataset collection system, they use hand-drawing samples and shapes.In 2015, Pereira et al. [27] have collected the HandPd is a dataset composed of images extracted from handwriting exams of 92 people divided into 18 healthy people (Healthy Group) and 74 patients (Patients Group). They worked on automatic Parkinson's disease diagnosis using spirals and meanders in forms as shown in Figure 3.1, that are then compared with the template for feature extraction, which was assessed employing three methods: Nave Bayes (NB), Optimum-Path Forest (OPF), and Support Vector Machines with Radial Basis Function (SVM-RBF), with the best results on the NB classifier that gave around 79% order accuracy. This study additionally indicated that meander samples play a very important role, resulting in higher accuracy than spiral samples.

In 2017, Loconsole at el. [37] used a limited number of features extracted from EMG (ElectroMyoG raphy) signals obtained at the arm level (time feature) and scans of traditional paper sheets (vision-based features) by utilising computer vision and applied an Artificial Neural Network-based classifier employing a Multi-Objective Genetic Algorithm (MOGA) achieving 95% accuracy. In 2018, Khatamino etal. [38] used HandWritten datasets that comprise of the Static Spiral Test (SST), the Dynamic Spiral Test (DST) and Stability Test on Certain Point (STCP) of 57 patients and 15 control healthy individuals [39].Author used a CNN-based deep learning approach and accomplished a precision of 88%. In the same year, Momina et al. [13], utilized Convolutional

Neural Networks(using the Alex-Net pre-prepared model)to extract visual features from numerous representations of different graphomotor tests delivered of 72 subjects (Patient and Health Group) subjects.These features are fed to a Support Vector Machine (SVM) classifier accomplishing accuracy of 83%. In 2019, Diaz et al. [40] worked with PAHAW offline data(images) that extracted features from CNN using a pre-trained VGG16 network. To reduce overfitting, the authors applied feature selection algorithm before classification.they applied different classifier (SVM, Random forest) achieved accuracy of 86.76% and also the examined which handwriting task performed better than other .Another study by Ribeiro et al. [41] used same dataset for the classification of PD and used Recurrent Neural Network(RNN) achieving 85% accuracy at the spiral and 89% on the meander. In 2019, Gupta et al.[42] used PaHaW off-line hand-drawn Archimedean spiral data and presented a novel distance based features PD prediction by extracting Fourier Transform based distance features,Tremor Estimation feature and combined distance-based features and fed these extracted features to the SVM classifier for classification and the reported accuracy of 81.66%.

In 2019, another author Rosa et al. [34] proposed an evolutionary approach to discriminate PD using hand shape analysis .they applied Cartesian Genetic Programming on a set of static features on HandPD dataset to show which handwriting template performed better. The results of the experiments indicated that the features derived by spirals are less informative than those derived by meanders, and that the global accuracy reached by meander analysis outperforms that of other studies.Their study also showed that, in its best configuration, the CGP performs better than state-of-the-art techniques for PD diagnosis proposed in the literature.

2.3 Summary of related work

Author	Year	Dataset	Handwriting Task	Features	Analysis	Results
Drotar et al.[26]	2013˜	Parkinson's	Letters, Words, Sentences and Archimedean Spiral	Online in-air Surface Features	SVM	80.09%
Drotar et al.[2]	2014	PaHaW	Letters, Words, Sentences and Archimedean Spiral	Online in-air and on-surface Features	SVM	85%
Drotar et al.[10]	2014	PaHaW	Letters, Words, Sentences and Archimedean Spiral	Online Spatial, Temporal Kinematic, Entropy, Signal Energy,	SVM	88%
Graca˜ et al. [3]	2014	Graca's Dataset	Archimedean Spiral	Online Spatial-Temporal and Pressure Feature	C4.5, RipperK , Bayesian Networks	86.67% 80.83% 87.50%
Dortar et al. [30]	2015	PaHaW	Letters, Words, Sentences and Archimedean Spiral	Online Spatial, Temporal Kinematic, Entropy, EMD and Pressure	SVM	˜89.09%
Pereira et al.[43]	2015	HandPD	Archimedean Spiral	Offline Mean Relative ˜Tremor(MRT) and Spatial Features	Naïve Bayes (NB), Optimum-Path Forest (OPF), SVM	78.90% 77.10% 75.80%
Dortar et al. [31]	2016	PaHaW	Letters, Words, Sentences and Archimedean Spiral	Online˜Kinematic, Pressure Features	SVM ADABOOST K-NN	81.3% 78.9% 71 %
Pereira et al.[27]	2016	NewHandPD	Archimedean Spiral and Meander	˜Pen-based Features	CNN OPF	87.1% on Meander Tasks
Laconsole et al.[37]	2017	Laconsole Dataset	Sentence, repetitive loops	Online and Offline features	ANN	95%
Impedovo et al.[11]	2018	PaHaW	Letters, Words, Sentences and Archimedean Spiral	Spatial, Temporal Kinematic, Entropy, Signal Energy, EMD, Pressure	SVM (RBF, Linear), KNN, LDA NB , RF ,AdaBoost	71% 68% 67.90% 66% 57%% 61%

Table 2.1: Summary of related works on Handwriting based Parkinson Prediction 2.1

Author	Year	Dataset	Handwriting Task	Features	Analysis	Results
Khatamino et al.[38]	2018˜	HW dataset	Archimedean Spiral	Dynamic and Visual features	CNN	88%
Angelillo et al.[44]	2019˜	PaHaW	Letters, Words, Sentences and Archimedean Spiral	Online Spatio-Temporal Kinematic	SVM, AdaBoost, Logistic Regression, Linear Discriminant	88.33%
Ribeiro et al.[41]	2019	New HandPD	Archimedean Spiral and Meander	Kinematic , Spatio-Temporal	RNN	85% (Spiral) 89%(Meander)
Diaz et al.[40]	2019	PaHaW	Letters, Words, Sentences and Archimedean Spiral	CNN based Visual Features	SVM	86.76%
Parziale et al.[45]	2019	HandPD	Archimedean Spiral and Meander	Offline Mean Relative Tremor and Spatial Features	SVM Decision Tree Random Forest	73.63%
Gupta et al.[42]	2019	PahaW	Archimedean Spiral	Spatial Features	SVM	81.66%
Alae et al.[35]	2020	Arabic Dataset	Arabic Text	Online Kinematics on surface/In-air Mechanical, Inclination, Pen Up.	Clustering and Principal Component Analysis(PCA)	97.3%
Amina et al.[20]	2020	PaHaW	Letters, Words, Sentences and Archimedean Spiral	CNN(Alex-net) based Visual Features	SVM	98.28%
Ismail et al.[36]	2020	HW dataset	Archimedean Spiral	CNN based Visual Features using AlexNet or GoogleNet	SVM KNN	98.28%

Table 2.2: Summary of related works on Handwriting based Parkinson Prediction 2.2

2.4 Benchmarking Datasets

In any research domain, the availability of datasets is one of the key requirements for the analysis of neurological disease.Collection of datasets is a very difficult activity in medical field since it presents a particular problem for selecting participants, choosing a acquisition device, and finding the most suitable handwriting tasks. The number of dataset use for the prediction of PD are discussed below. In this section, the datasets that are used by in pervious techniques to evaluate their approaches have been reviewed.

- **PAHAW Dataset:** This dataset consists multiple handwriting samples from 37 people with Parkinson's disease (19 men/18 women) and 38 healthy people (20 men/18 women). The samples came from the Movement Disorders Center at Masaryk University's First Department of Neurology and St. Anne's University Hospital in Brno, Czech Republic. Each participant was given eight handwriting assignments to complete at their own pace. The signals were captured using a Wacom Intuos 4M digitising tablet with a sampling rate of 150 Hz. Each individual perform eight tasks according to template illustrated in Figure 2.1

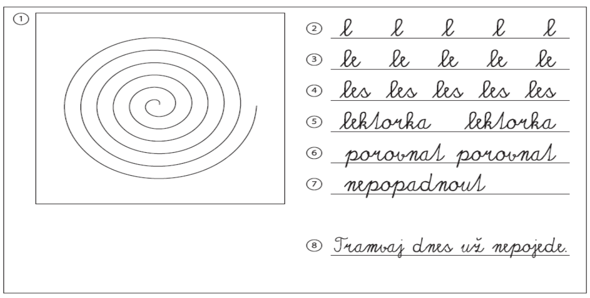


Figure 2.1: Template proposed by [2]

- **HandPD dataset:** This dataset contains 92 individuals, 18 of Healthy Group and 74 of Patients Group, the latter being composed of people suffering from Parkinson's Disease (PD). Botucatu Medical School, So Paulo State University - Brazil, gathered the handwritten exams. The main task includes filling out a form that consists of four spirals and four meanders.[27] as shown in Figure 2.2

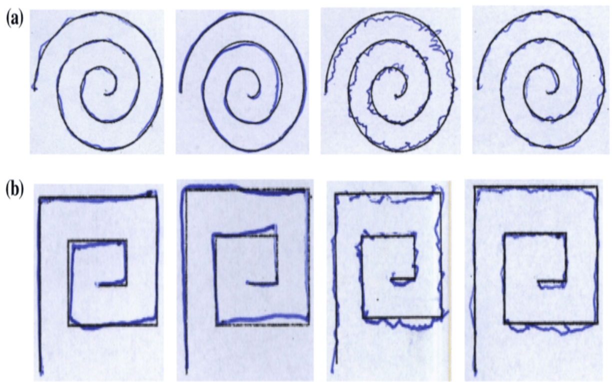


Figure 2.2: HandPD Dataset tasks

- **NewHandPD Dataset:** This dataset is composed of 66 individuals that are obtained from 35 Parkinson's patients (21 males and 10 females) and 31 healthy subjects (18 males and 17 females).Every individual was approached to draw 12 exams, 4 spirals, 4 meanders, 2 circled movements (one circle in the air and another on the paper). Some handwritten dynamics features were likewise recorded utilizing an advanced pen, having images from 4 spirals,4 meanders ,4 circles and signals for all 12 exams.So every individual sample includes 9 images and 12 signals.[46]

- **Graca et al. Dataset** In 2014, Graca et al. [3] used mobile devices to predict PD by completing a different tasks (spiral analysis, gait analysis, tip analysis) in which 35 samples were collected from drawing Archimedes spiral as shown in Figure 2.3.They allow the user to perform the task on a mobile screen because it's easy to manage.

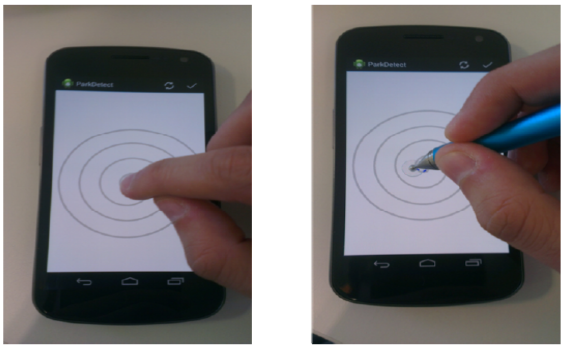


Figure 2.3: Image sample of grace et al.[3] dataset

- **Arabic Handwriting Dataset:** Arabic Handwriting dataset used for PD prediction with 28 Parkinson's patients and 28 healthy subject. This data set completed with three tasks shown in Figure 2.4.

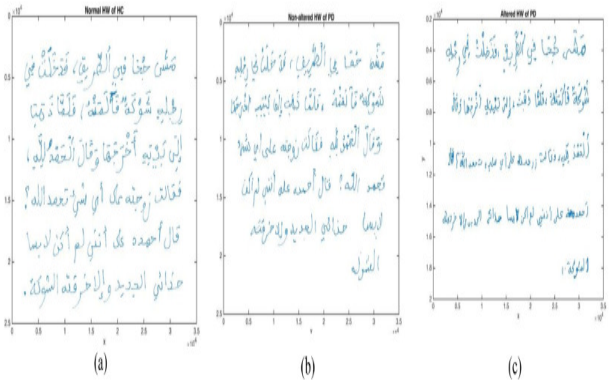
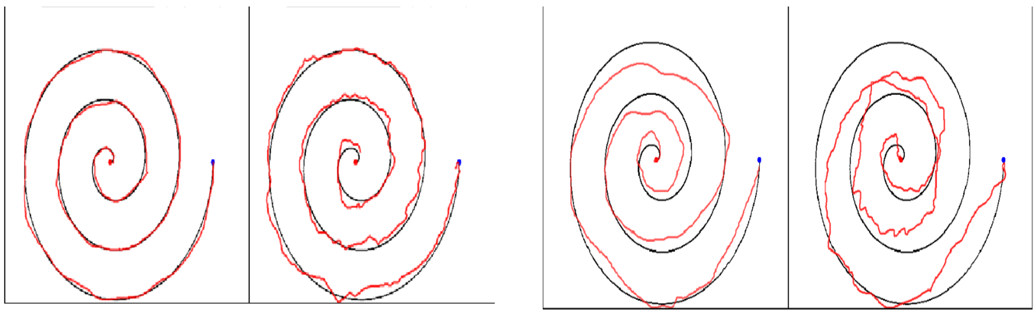


Figure 2.4: Arabic Handwriting dataset

- **The Hand written Dataset:** The Hand-Written (HW) dataset was gathered at Istanbul University's Cerrahpasa Faculty of Medicine's Department of Neurology[26, 47]. This dataset contains time-series data from handwriting spiral exams of individuals in two groups: healthy people and Parkinson's disease. The dataset contains 72 individuals, 57 of whom are patients and 15 of whom are healthy controls. Exams are provided to everyone in the same way (recommend to draw inward to outward). As illustrated in Figure 2.5, it consists of three types of handwriting tests: the Static Spiral Test (SST), the Dynamic Spiral Test (DST), and the Stability Test on Certain Point (STCP).



(a) SST Test

(b) DST Test

Figure 2.5: Hand Written dataset

- **Mirjana et al dataset** In this study[48],There were 43 participants in total: 33 patients with Parkinson's disease and ten healthy controls (HC). This dataset consists of 4 tasks,

  1. Writing a sentence between two lines(a distance of 1 cm when looking at the laptop)
  2. Typing a sentence between two lines(a distance of 1 cm, with the monitor out of sight)
  3. Writing a paragraph without space restriction while looking at the screen

4. Writing a sentence without space restriction, with the monitor out of sight

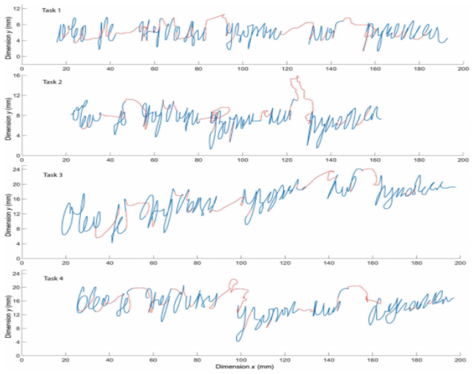


Figure 2.6: Signal Tasks.

2.5 Summary

This chapter presented an overview of the techniques presented for the identification of PD using handwritten analysis.Recent studies primarily target both static and dynamic features or selected features to enhance the over all performance, to the best of our knowledge, no investigations have been carried out to study the relationship between the template under study and the computed features.Our study will combined both A summary of related works (2013-present) can can be seen in Table

CHAPTER 3

METHODOLOGY

In the preceding chapter, significant contributions in the field of Parkinson disease identification using handwriting analysis was discussed. In most situations, the discovered attributes are mapped to computational features derived from known templates and input into a classifier to identify whether or not PD exists. The drawing or handwriting template under study is a key aspect in determining the type of features to use. While past research has focused on combining features or decisions from various templates to improve overall performance. In this chapter, the methodology used for feature relevance analysis for the identification of PD through handwriting is explained in length. Workflow for proposed approach is shown in Figure 3.1. The goal of our proposed system is to assess the optimal set of features and to study which features are better suited to specific templates.

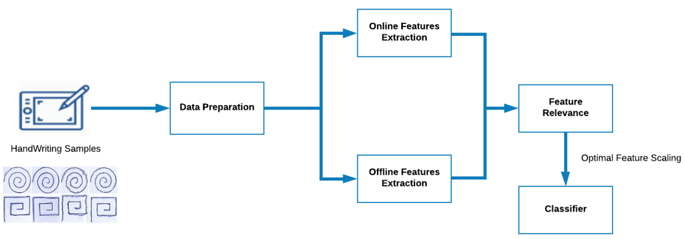


Figure 3.1: Proposed Approach.

3.1 DataSet

Data acquisition is a very difficult task when it comes to medical problems. In our research, we used the Parkinson's Disease Handwriting Database (PaHaW). This database consists of samples obtained from 37 Parkinson's patients (19 males and 18 females) and 38 healthy subjects (20 males and 18 females). All members involved in PD diagnosis enlist from the movement disorder center at The Department of Neurology, Masaryk University, and St. Annes Hospital in Smo, Czech Republic. And all samples write in the native language of the participants and the participant completed all tasks according to the template. All tasks are the following:

1. Drawing an Archimedes spiral

2. Writing in cursive the letter l

3. The bigram le

4. The trigram les

5. Writing in cursive the word lektorka ("female teacher" in Czech)

6. porovnat ("to compare")

7. nepopadnout ("to not catch")

8. Writing in cursive the sentence Tramvaj dnes uz nepojede ("The tram won't go today")

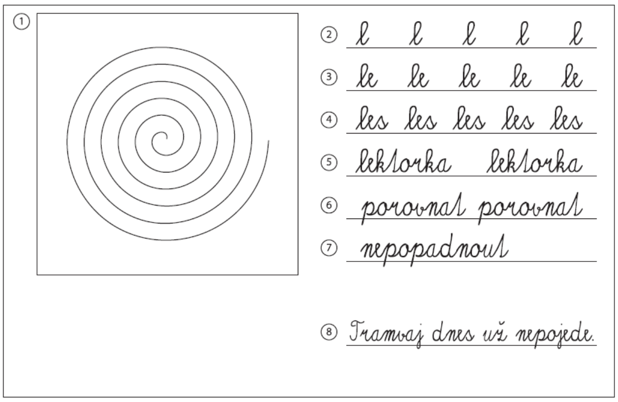


Figure 3.2: PAHAW template.

3.2 Acquisition Device

The number of devices used for data acquisition, in our problem the authors used Wacom Intuos 4 M digitizer or digital pen for data collection shown in Figure 3.3, some patients may be unfamiliar with the device, so patient writing on paper, the paper is set on the digitizer. By this device, some useful features have been acquired: (x-y) the coordinate of the pen moves in a different direction, the time stamps, the pen Orientation (azimuth and altitude) and the pressure, and if the pen moves in-air the button state is 0 otherwise (on the surface) the button state is 1. All the features reported are the numeric values shown in Figure 3.3

3.3 Data Prepossessing

The Acquisition Device discusses in the previous section use for collecting the pen-based data. These included all of the functional attributes that could be used for the derived kinematics features. All of these features are sequentially measured within the same time intervals. In Literature, most authors work with these sequential features and measure only mean values, and then feed them to the model[2]. But when they transform the sequential value

(a) Digitized smart pen

4597	4821	777779	1	3225	519	845
4597	4821	777787	1	3225	519	897
4597	4821	777794	1	3225	519	937
4598	4821	777802	1	3225	519	955
4598	4821	777809	1	3225	519	957
4599	4821	777817	1	3225	519	955
4601	4823	777824	1	3225	519	961
4604	4825	777832	1	3216	518	971
4607	4828	777839	1	3216	518	991
4611	4832	777847	1	3216	518	997
4617	4836	777854	1	3216	518	1007
4623	4841	777862	1	3216	518	1021
4630	4847	777870	1	3216	518	1011

(b) Features Acquire by device

Figure 3.3: Feature acquired using digitized pen

into mean values, the beneficial information is lost. Therefore, in this study, we use all sequential values and give the classifier as it is an emphasis on the output of sequential data analysis.

3.4 Features Extraction

In this section extracted features employed, offline and online features of handwriting discuss in a later section.

3.4.1 Online Features

Online features provide valuable information for the diagnosis of PD. In our research, we are working on online features that have been calculated from raw data in the PaHaW database. There are a variety of features: (x-y) coordinate, azimuth, altitude, pressure, timestamp, button status. By using these features, new features derived such as velocity, acceleration, distance and also calculated the mean, standard deviation, median, 10th, and 90th percentile of the whole signal acquired by the device give it to the predictor classifier.

3.4.2 Offline Features

In offline features, these features are derived by handwriting and hand-drawing. Some researchers used offline functionality instead of online attributes[27]. In 2015, Pereira used the hand-drawing shape to extract the features. In our case study, we used the dortar et al. dataset.[2] The original dataset contains online features extracted from the device (X, Y coordinates, button states, pressure) so that we can convert all online features to images, online features in the form of numeric data, and by plotting x, y coordinates into images shown in Figure 3.5

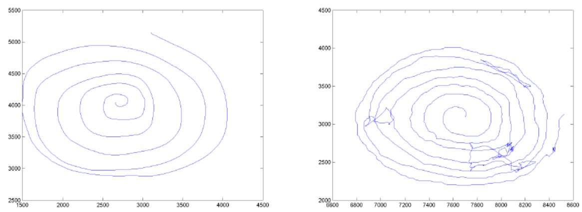


Figure 3.4: Offline image of the Archimedean spiral (a): Healthy Subject (b): Parkinson Patient

All attributes or features of the dataset includes information about PD and healthy subject, by using visualization techniques and apply different filters on images for features extraction and enhanced the dataset because original image not enough to train the dataset. We feed the images to pre-trained model VGG16 and extracted features from the last FC layers which contain 4096 features.

- **VGG16**

  Visual Geometry Group at Oxford introduces Vgg16. This is a 16-layer convolutional neural network. The model uses a set of pre-trained weights from ImageNet. In ImageNet, a dataset of over 14 million images belonging to 1000 classes, the model achieves 92.7 % accuracy. The default RGB image input size for the VGG16 model is 224 x 224 pixels with three channels.The used architecture of VGG16 is summarized in Figure 3.5.

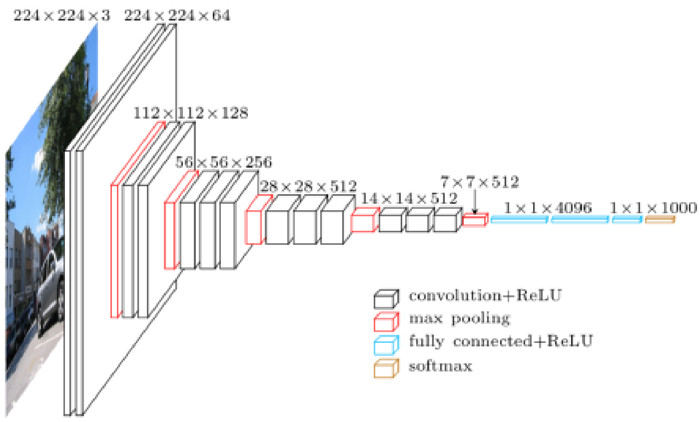


Figure 3.5: VGG16 Architecture.

3.5 Feature Selection

Feature selection is the process of selecting relevant and informative features with the motivation of data/feature set reduction, performance improvement, and data understanding[49]. The primary goal of a feature selection procedure would be to find the features (or feature components) that are useful in identifying the presence and absence of PD. Filter, wrapper, and embedding approaches are the three basic kinds of feature selection algorithms[50]. Generation, evaluation, stop criterion, and validation are the four key processes of a feature selection approach. A search strategy is used in the generation process to obtain a subset of features (usually utilising forward selection, backward removal, bidirectional, and other methods). The efficiency of the resulting subset is then evaluated using an evaluation criterion, which might be independent (filter) or dependent (measurement) (wrapper). After each iteration, a stopping condition is examined to decide when the selection process should be terminated. Typical criteria involve achievement of optimal subset or bounds on a number of features or iterations etc. Once the stopping condition is met, the resultant subset of features can be confirmed [51].

For our problem, we employed a genetic algorithm (GA), a wrapper approach a Correlation, a filter approach for feature selection.

3.5.1 Genetic Algorithms

Genetic algorithm is one of the most advanced feature selection algorithms. It is a stochastic function optimization method based on natural genetics and biological evolution mechanics.In nature, organisms' genes tend to evolve through generations to improve their ability to adapt to their surroundings. It acts on a population of individuals to better approximations over time. A state diagram for the feature selection process with the genetic algorithm is shown in Figure 3.6 .

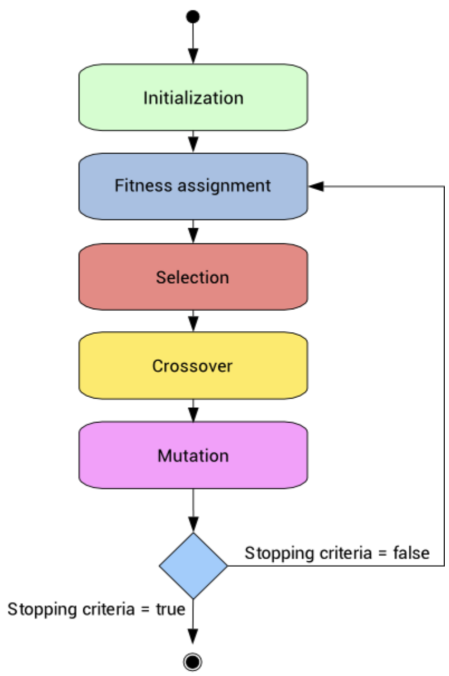


Figure 3.6: Genetic Algorithm.

As with natural adaptation, this process results in the evolution of populations that are better suited to their environment than the individuals from which they were formed. This technique has an advantage over others in that it permits the best answer to emerge from the best of previous solutions. In this study, we used the basic application of genetic algorithms as the objective of our system is to select an optimal set of features that provides better performance than all features. We will first analyze the relevance of the features i.e 97 online and 4096 offline features. The GA is used to generate individuals of length (97 & 4096) and the set bits are used to select the respective features. We executed the GA ten times and extracted features that are almost selected every time we runs GA.We used the following parameters for GA:

- Population Size: 50,
- Crossover Rate: 0.5,
- Mutation Rate: 0.2,
- Selection Rule: logistic regression
- Number of Generations: 10.

The initial population is generated at random. The fitness function is used to evaluate the chromosomes in each generation, with the present population's fitness values being utilized to find the offspring of the next generation. When the specified number of generations has been evaluated, the procedure comes to an end. The best individual of the final generation determines the selected feature subset. The division of online and offline features according to their relevance is explained in the next Chapter 4.

3.5.2 Correlation

Another approach for feature selection is Correlation. It's a metric for determining the degree of linear correlation between an input feature and an output feature. It has a range of +1 to -1, with 1 denoting total positive correlation and -1 denoting total negative correlation. As high correlation features are more linearly dependent, they have roughly the same impact on the dependent variable. When there is a strong correlation between two features, one of them may be dropped. The correlation mathematical formula is shown in 3.5.2.

$$P_{X,Y} = \frac{COV(X,Y)}{\sigma X \sigma Y}$$

3.6 Classification

Classification is an important part of research because evaluating the quality of the literature, we are providing the best results on this disease. In the literature survey, many techniques used for classification the most commonly used help vector machine, Random forest, Naive Bayes, neural networks, etc. Some researchers used a combination of classifiers and often used several neural networks to improve overall accuracy. In our implementation, we use three classifiers support vector machines.We applied this classifier on online and offline features data extracted by the feature relevance method.

3.6.1 SVM

SVM (Support Vector Machine) is a supervised machine learning model for binary and regression problem classification. Each data item is represented as a point in n-dimensional space (where n is the number of features), with the value of each feature being the value of a certain position in the SVM algorithm[52]. In general, SVM is divided into two types: linear and non-linear. A linear SVM computes a linear decision boundary using a linear kernel. Figure 3.7 shows a two-dimensional data example of a linear SVM.

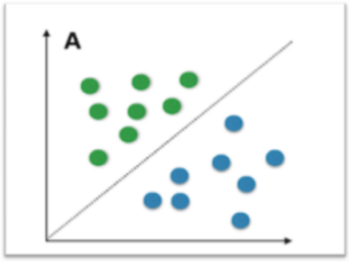


Figure 3.7: SVM classifier.

For higher dimensions, planes or hyper-planes are computed. A Non-linear SVM (Figure 3.8) uses a non-linear kernel.

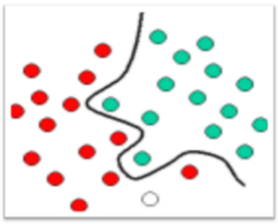


Figure 3.8: SVM classifier.

There are various classification algorithms used in machine learning, but we only utilised SVM because it is superior than most of them because it has a lower computational complexity and delivers faster predictions with higher accuracy. Different classifiers, such as Navies Bayes, Decision Trees, and others, have been employed in studies, but SVM has shown to be more promising than the others..

3.7 Summary

In this chapter discuss the detail of the PaHaW dataset used in our study and present the online and offline extracted features. A brief overview of feature selection and SVM classifier examined our work is also present. In the next chapter, we will discuss the results and different experiments.

CHAPTER 4

ANALYSIS & RESULTS

This chapter describes the specifics of all experiments, computes the efficiency of the different models, and examines the effectiveness of extracted features in each task to identify the presence and absence of PD. All experiments performed by Drotar et al. dataset(PaHaW) are presented in section 2.4. In this chapter, we first discussed training and testing data and then show results using Feature selection techniques and classification models.

4.1 Training Test Datasets

The PaHaW dataset contains 75 sample files, all of which are used for the experiment (37 Parkinson's patients and 38 healthy subjects). When we extracted the features from 75 samples, we used folding techniques for rotating 75 samples, In our scenario, we divide 75 sample data into fivefold. By using this technique, we are able to estimate the skill of our model on unseen data.

4.2 Performance Metrics

The functionality of the proposed system is evaluated by using standard measures of accuracy. Each of these is briefly described and figure out the accuracy of each task with the optimal set of features using SVM Classifier. We discussed the Standard measures are as follows.

- **True Positive** Data instance belongs to a specific class and is correctly classified by the algorithm that data belongs to the same class. In the case of PD identification, a PD subject is correctly classified as PD.

- **False Negative** The algorithm detects that the data does not pertain to a specific class, however it belongs to that class. In other words, a PD subject is wrongly classified as Healthy.
- **False Positive** The data instance does not belong to a specific class, but it is incorrectly identified by the algorithm as belonging to that class. In this case, a healthy subject is wrongly classified as a patient.
- **True Negative** The algorithm identifies that data does not belong to a specific class; however, the data actually belongs to another class. In this case, a healthy subject is correctly classified as healthy.

4.2.1 Accuracy

The accuracy calculate the ability of overall system to precisely classify the PD patient and healthy subject.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

4.3 Results and Discussion

In this chapter, we will explain the performance of the system,we present the accuracy of every single task with all features and relevance features set and then fed it to the classifier discussed in chapter 3.6. There are different sets of experiments used for classification with online data and offline data.

4.3.1 Results on all features (Online and Offline)

In this experiment, we simple classification model was applied on every single task for classification and the result mentioned in the below table 4.1

Tasks	**Online Features (97)**	**Offline Features (4096)**
Archimedean Spiral	57%	57%
Repetitive(l)	60%	43%
Repetitive(le)	62%	57%
Repetitive(les)	60%	64%
Word(leplorka)	50%	64%
Word(porovnal)	53%	50%
Word(nepopadnoul)	47%	64%
Sentence	67%	64%
Overall Accuracy	57%	58%

Table 4.1: Task-Wise Accuracy on All Features

4.3.2 Results Using GA Technique (Online and Offline)

In this experiment, we applied the Genetic Algorithm (GA) Feature selection technique on each task features set and extracted the optimal feature subset. We performed 10 iterations to determine specific features that were selected almost every time we runs the Genetic Algorithm. We fed these selected features data to a machine learning classifier i.e. SVM for PD classification and the obtained result after Genetic feature selection techniques are mentioned in the below table 4.2,4.3

Tasks	**No of Features Selected**	**SVM**
Archimedean Spiral	6	78%
Repetitive(l)	8	74%
Repetitive(le)	5	75%
Repetitive(les)	8	80%
Word(leplorka)	5	60%
Word(porovnal)	5	80%
Word(nepopadnoul)	5	74%
Sentence	6	73%
Overall Accuracy		74%

Table 4.2: GA Online Features Results

Tasks	**No of Features Selected**	**SVM**
Archimedean Spiral	65	86%
Repetitive(l)	21	71%
Repetitive(le)	38	64%
Repetitive(les)	33	78%
Word(leplorka)	114	71%
Word(porovnal)	45	71%
Word(nepopadnoul)	39	71%
Sentence	129	72%
Overall Accuracy		79%

Table 4.3: GA Offline Features Results

4.3.3 Results using Correlation technique(Online and Offline)

In this experiment, we build a correlation matrix, which examines the correlation of all features (for all possible feature combinations). We selected highly correlated features by setting a threshold of 0.5, It removed the first feature that is correlated with anything else without any other insight. We removed these highly correlated features and selected the features which are not highly correlated columns and have an absolute correlation smaller than 0.5. We fed these selected features data to a machine learning classifier i.e. SVM for PD classification. The obtained result after Correlation feature selection techniques are mentioned in the below table 4.4,4.5

Tasks	**No of Features Selected**	**SVM**
Archimedean Spiral	14	64%
Repetitive(l)	17	61%
Repetitive(le)	21	70%
Repetitive(les)	17	61%
Word(leplorka)	16	61%
Word(porovnal)	20	70%
Word(nepopadnoul)	19	53%
Sentence	20	78%
Overall Accuracy		64.75%

Table 4.4: Correlation Online Feature Results

Tasks	No of Features Selected	SVM
Archimedean Spiral	2172	79%
Repetitive(l)	1874	86%
Repetitive(le)	1187	87%
Repetitive(les)	1862	71%
Word(leplorka)	1725	89%
Word(porovnal)	1931	87%
Word(nepopadnoul)	1801	71%
Sentence	1759	64%
Overall Accuracy		79%

Table 4.5: Correlation Offline Features Results

4.3.4 Combined features result analysis

In this experiment, firstly we employed task-wise classification by combining all online and offline features. We also combined both online and offline features after feature selection and performed task-wise classification. and the obtained results are mentioned in the table below 4.6

Task	**All Features Accuracy**	**Selected Features Accuracy (GA)**	**Selected Features Accuracy(Correlation)**
Archimedean Spiral	59%	81.67%	72%
Repetitive(l)	52%	78.32%	74%
Repetitive(le)	61%	74.32%	78%
Repetitive(les)	62%	75.24%	66%
Word(leplorka)	57%	76.35%	75%
Word(porovnal)	54%	74.11%	78%
Word(nepopadnoul)	55%	76.08%	63%
Sentence	66%	83.57%	71%

Table 4.6: Combined Features Result Analysis

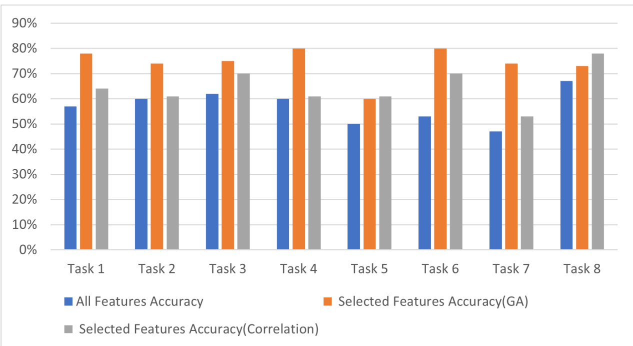


Figure 4.1: Comparison between all and selected features accuracy on online features

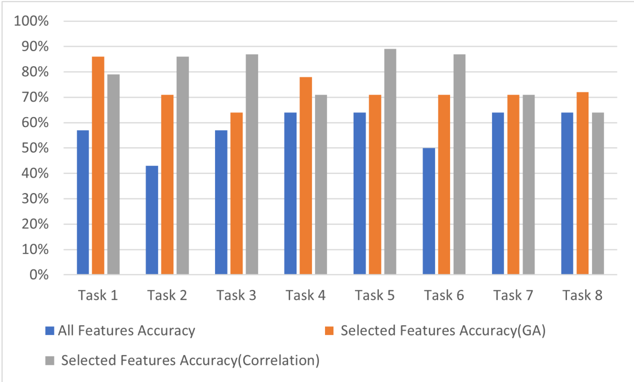


Figure 4.2: Comparison between all and selected features accuracy on offline features

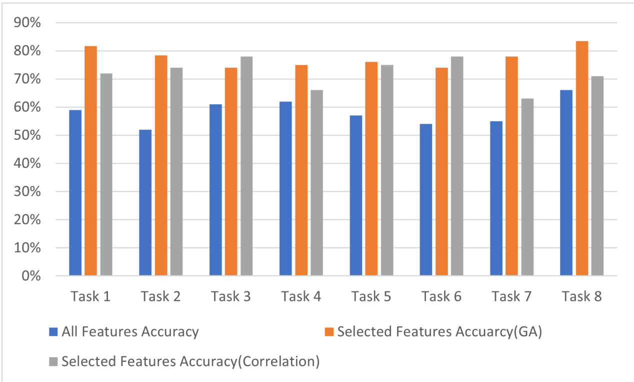


Figure 4.3: Comparison between all and selected features accuracy on combined features

To further, establish the viability of the proposed model we compare with the studies on task-based selection in table 4.8. All of the studies have utilized the PaHaW dataset for the assessment of the proposed method. The authors in [30, 13] have employed several static and dynamic features extraction techniques to predict PD using PaHaW dataset. We use combined features of handwriting to demonstrate their use in identifying the presence or absence of Parkinson's disease.The accuracies reported in different experiments are comparable to those reported in the literature.

Task	Impedevo et al [11]	Angelilo et al [44]	Diaz et al [40]	Momina et al [53]	Proposed Technique
Archimedean Spiral	54.67%	53.75%	75.00%	89.64%	81.67%
Repetitive(l)	61.80%	67.08%	64.16%	75.00%	78.32%
Repetitive(le)	72.28%	62.50%	58.33%	73.75%	74.32%
Repetitive(les)	55.28%	57.91%	71.67%	72.32%	75.24%
Word(leplorka)	59.80%	54.58%	75.41%	79.46%	76.35%
Word(porovnal)	63.71%	56.75%	63.75%	74.46%	74.11%
Word(nepopadnoul)	60.98%	61.67%	70.00%	79.28%	76.08%
Sentence	71.95%	70.40%	67.08%	81.42%	83.57%

Table 4.8: Comparison with existing studies

We discussed our result according to the most effective features set on the specific task in PaHaW dataset. We observed features extracted from GA give a better classification performance (77.46% global accuracy) using an SVM classifier than features extracted from correlation selection techniques. Nevertheless, considering the very limited amount of text (drawing) available in each task, the realized accuracies are indeed promising. Comparing the performance of different tasks, it can be observed that after feature selection **Sentence** task has the highest accuracy in overall experiments. The Archimedean spiral tasks achieved second highest accuracy in overall experiments. The word-based tasks "porovonal, nepopadnoul,leplorka," task provides remarkably similar results. On the other hand,repetitive letter tasks "l" and "les" obtained the highest accuracy "le" have the lowest accuracy. "le" tasks were comparatively obtained less effectively than other tasks in our proposed system.

4.4 Summary

This chapter presents the details of all experiments carried out to prove our thoughts to bring up in this research. We used feature relevance methodologies to evaluate the performance of features online and offline individually, then integrated both types of features and applied SVM for classification. Task-wise accuracy was used in these investigations, and the system was compared against others in the literature.

CHAPTER 5

CONCLUSION & FUTURE WORK

The potential for handwriting features to indicate Parkinson's disease is discussed in this study. The literature has looked at both online and offline features, however, in our study, we used a combination of online and offline features and extracted a set of features that performed better on a specific task template This study does not deny previous research on online and offline features; rather, it enhances the author's expertise and demonstrated the utility of both online and offline features. we explore the feature relevance techniques to detect offline and online features. Evaluation on a standard data set (PAHAW), our proposed system reports overall accuracy of 77.46% when we combined both features and applied feature selection. Another an important aspect of our studies is that we explained which templates perform better after feature selection.

In conclusion, we have demonstrated that such a method could be beneficial to clinicians in the diagnosis of Parkinson's disease since it allows them to select the most important features of the disease and, as a result, develop a set of guidelines for defining fresh testing protocols. This work will be used to guide future work and research in this field because of the performance metric across numerous datasets.

REFERENCES

[1] T. Booth, M. Nathan, A. Waldman, A.-M. Quigley, A. Schapira, and J. Buscombe, "The role of functional dopamine-transporter spect imaging in parkinsonian syndromes, part 1," *American Journal of Neuroradiology*, vol. 36, no. 2, pp. 229–235, 2015.

[2] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Analysis of in-air movement in handwriting: A novel marker for parkinson's disease," *Computer methods and programs in biomedicine*, vol. 117, no. 3, pp. 405–411, 2014.

[3] R. Graça, R. S. e Castro, and J. Cevada, "Parkdetect: Early diagnosing parkinson's disease," in *2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1–6, IEEE, 2014.

[4] C. Tucker, Y. Han, H. Black Nembhard, W.-C. Lee, M. Lewis, N. Sterling, and X. Huang, "A data mining methodology for predicting early stage parkinson's disease using non-invasive, high-dimensional gait sensor data," *IIE transactions on healthcare systems engineering*, vol. 5, no. 4, pp. 238–254, 2015.

[5] D. Hirtz, D. Thurman, K. Gwinn-Hardy, M. Mohamed, A. Chaudhuri, and R. Zalutsky, "How common are the "common" neurologic disorders?," *Neurology*, vol. 68, no. 5, pp. 326–337, 2007.

[6] O.-B. Tysnes and A. Storstein, "Epidemiology of parkinson's disease," *Journal of Neural Transmission*, vol. 124, no. 8, pp. 901–905, 2017.

[7] S. von Campenhausen, B. Bornschein, R. Wick, K. Bötzel, C. Sampaio, W. Poewe, W. Oertel, U. Siebert, K. Berger, and R. Dodel, "Prevalence and incidence of parkinson's disease in europe," *European Neuropsychopharmacology*, vol. 15, no. 4, pp. 473–490, 2005.

[8] L. M. De Lau and M. M. Breteler, "Epidemiology of parkinson's disease," *The Lancet Neurology*, vol. 5, no. 6, pp. 525–535, 2006.

[9] Y. Zou, J. Tan, N. Li, J. Yang, B. Yu, J. Yu, W. Du, W. Zhang, L. Cui, Q. Wang, *et al.*, "The prevalence of parkinson's disease continues to rise after 80 years of age: a cross-sectional study of chinese veterans," *Eur Rev Med Pharmacol Sci*, vol. 18, no. 24, pp. 3908–3915, 2014.

[10] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Decision support framework for parkinson's disease based on novel handwriting markers," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 3, pp. 508–516, 2014.

[11] D. Impedovo, G. Pirlo, and G. Vessio, "Dynamic handwriting analysis for supporting earlier parkinson's disease diagnosis," *Information*, vol. 9, no. 10, p. 247, 2018.

[12] M. Alviano, G. Greco, and F. Scarcello, *AI* IA 2019–Advances in Artificial Intelligence: XVIIIth International Conference of the Italian Association for Artificial Intelligence, Rende, Italy, November 19–22, 2019, Proceedings*, vol. 11946. Springer Nature, 2019.

[13] M. Moetesum, I. Siddiqi, N. Vincent, and F. Cloppet, "Assessing visual attributes of handwriting for prediction of neurological disorders—a case study on parkinson's disease," *Pattern Recognition Letters*, vol. 121, pp. 19–27, 2019.

[14] A. M. Vlaar, A. E. Bouwmans, M. J. Van Kroonenburgh, W. H. Mess, S. C. Tromp, P. G. Wuisman, A. G. Kessels, A. Winogrodzka, and W. E. Weber, "Protocol of a prospective study on the diagnostic value of transcranial duplex scanning of the substantia nigra in patients with parkinsonian symptoms," *BMC neurology*, vol. 7, no. 1, pp. 1–6, 2007.

[15] A. Salarian, H. Russmann, C. Wider, P. R. Burkhard, F. J. Vingerhoets, and K. Aminian, "Quantification of tremor and bradykinesia in parkinson's disease using a novel ambulatory monitoring system," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 2, pp. 313–322, 2007.

[16] U. Tisch, I. Schlesinger, R. Ionescu, M. Nassar, N. Axelrod, D. Robertman, Y. Tessler, F. Azar, A. Marmur, J. Aharon-Peretz, *et al.*, "Detection of alzheimer's and parkinson's disease from exhaled breath using nanomaterial-based sensors," *Nanomedicine*, vol. 8, no. 1, pp. 43–56, 2013.

[17] J. Mekyska, Z. Smekal, M. Kostalova, M. Mrackova, S. Skutilova, and I. Rektorova, "Motor aspects of speech imparment in parkinson's disease and their assessment," *Ceska A Slovenska Neurologie A Neurochirurgie*, vol. 74, no. 6, pp. 662–668, 2011.

[18] A. Tsanas, M. A. Little, C. Fox, and L. O. Ramig, "Objective automatic assessment of rehabilitative speech treatment in parkinson's disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 181–190, 2013.

[19] J. Rusz, R. Čmejla, H. Ržičková, J. Klempíř, V. Majerová, J. Picmausová, J. Roth, and E. Ržička, "Acoustic assessment of voice and speech disorders in parkinson's disease through quick vocal test," *Movement Disorders*, vol. 26, no. 10, pp. 1951–1952, 2011.

[20] A. Naseer, M. Rani, S. Naz, M. I. Razzak, M. Imran, and G. Xu, "Refining parkinson's neurological disorder identification through deep transfer

learning," *Neural Computing and Applications*, vol. 32, no. 3, pp. 839–854, 2020.

[21] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, "Novel speech signal processing algorithms for high-accuracy classification of parkinson's disease," *IEEE transactions on biomedical engineering*, vol. 59, no. 5, pp. 1264–1271, 2012.

[22] S. Rosenblum, M. Samuel, S. Zlotnik, I. Erikh, and I. Schlesinger, "Handwriting as an objective tool for parkinson's disease diagnosis," *Journal of neurology*, vol. 260, no. 9, pp. 2357–2361, 2013.

[23] J. Mucha, J. Mekyska, Z. Galaz, M. Faundez-Zanuy, K. Lopez-de Ipina, V. Zvoncak, T. Kiska, Z. Smekal, L. Brabenec, and I. Rektorova, "Identification and monitoring of parkinson's disease dysgraphia based on fractional-order derivatives of online handwriting," *Applied Sciences*, vol. 8, no. 12, p. 2566, 2018.

[24] M. Moetesum, I. Siddiqi, F. Javed, and U. Masroor, "Dynamic handwriting analysis for parkinson's disease identification using c-bigru model," in *2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 115–120, 2020.

[25] J. H. Man, L. Groenink, and M. Caiazzo, "Cell reprogramming approaches in gene-and cell-based therapies for parkinson's disease," *Journal of controlled release*, vol. 286, pp. 114–124, 2018.

[26] M. Isenkul, B. Sakar, and O. Kursun, "Improved spiral test using digitized graphics tablet for monitoring parkinson's disease," 05 2014.

[27] C. R. Pereira, D. R. Pereira, F. A. Silva, J. P. Masieiro, S. A. T. Weber, C. Hook, and J. P. Papa, "A new computer vision-based approach to aid

the diagnosis of parkinson's disease," *Computer Methods and Programs in Biomedicine*, vol. 136, pp. 79–88, 2016.

[28] L. Palmerini, L. Rocchi, S. Mellone, F. Valzania, and L. Chiari, "Feature selection for accelerometer-based posture analysis in parkinson's disease," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 3, pp. 481–490, 2011.

[29] K. Niazmand, K. Tonn, A. Kalaras, U. M. Fietzek, J.-H. Mehrkens, and T. C. Lueth, "Quantitative evaluation of parkinson's disease using sensor based smart glove," in *2011 24th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 1–8, IEEE, 2011.

[30] P. Drotár, J. Mekyska, Z. Smékal, I. Rektorová, L. Masarová, and M. Faundez-Zanuy, "Contribution of different handwriting modalities to differential diagnosis of parkinson's disease," in *2015 IEEE international symposium on medical measurements and applications (MeMeA) proceedings*, pp. 344–348, IEEE, 2015.

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

[32] C. R. Pereira, S. A. Weber, C. Hook, G. H. Rosa, and J. P. Papa, "Deep learning-aided parkinson's disease diagnosis from handwritten dynamics," in *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp. 340–346, Ieee, 2016.

[33] M. Angelillo, D. Impedovo, G. Pirlo, and G. Vessio, *Performance-Driven Handwriting Task Selection for Parkinson's Disease Classification*, pp. 281–293. 11 2019.

[34] R. Senatore, A. Della Cioppa, and A. Marcelli, "Automatic diagnosis of neurodegenerative diseases: An evolutionary approach for facing the interpretability problem," *Information*, vol. 10, no. 1, p. 30, 2019.

[35] A. Ammour, I. Aouraghe, G. Khaissidi, M. Mrabti, G. Aboulem, and F. Belahsen, "A new semi-supervised approach for characterizing the arabic on-line handwriting of parkinson's disease patients," *Computer methods and programs in biomedicine*, vol. 183, p. 104979, 2020.

[36] I. Canturk, "Fuzzy recurrence plot-based analysis of dynamic and static spiral tests of parkinson's disease patients," *NEURAL COMPUTING & APPLICATIONS*, 2020.

[37] C. Loconsole, G. F. Trotta, A. Brunetti, J. Trotta, A. Schiavone, S. I. Tatò, G. Losavio, and V. Bevilacqua, "Computer vision and EMG-based handwriting analysis for classification in Parkinson's disease," in *International Conference on Intelligent Computing*, pp. 493–503, Springer, 2017.

[38] P. Khatamino, İ. Cantürk, and L. Özyılmaz, "A deep learning-cnn based system for medical diagnosis: An application on parkinson's disease handwriting drawings," in *2018 6th International Conference on Control Engineering & Information Technology (CEIT)*, pp. 1–6, IEEE, 2018.

[39] B. E. Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgen, S. Delil, H. Apaydin, and O. Kursun, "Collection and analysis of a parkinson speech dataset with multiple types of sound recordings," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 4, pp. 828–834, 2013.

[40] M. Diaz, M. A. Ferrer, D. Impedovo, G. Pirlo, and G. Vessio, "Dynamically enhanced static handwriting representation for parkinson's disease detection," *Pattern Recognition Letters*, vol. 128, pp. 204–210, 2019.

[41] L. C. Ribeiro, L. C. Afonso, and J. P. Papa, "Bag of samplings for computer-assisted parkinson's disease diagnosis based on recurrent neural networks," *Computers in biology and medicine*, vol. 115, p. 103477, 2019.

[42] J. D. Gupta and B. Chanda, "Novel features for diagnosis of parkinson's disease from off-line archimedean spiral images," in *2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST)*, pp. 1–6, IEEE, 2019.

[43] C. R. Pereira, D. R. Pereira, F. A. Da Silva, C. Hook, S. A. Weber, L. A. Pereira, and J. P. Papa, "A step towards the automated diagnosis of parkinson's disease: Analyzing handwriting movements," in *2015 IEEE 28th international symposium on computer-based medical systems*, pp. 171–176, IEEE, 2015.

[44] M. T. Angelillo, D. Impedovo, G. Pirlo, and G. Vessio, "Performance-driven handwriting task selection for parkinson's disease classification," in *International Conference of the Italian Association for Artificial Intelligence*, pp. 281–293, Springer, 2019.

[45] A. Parziale, A. Della Cioppa, R. Senatore, and A. Marcelli, "A decision tree for automatic diagnosis of parkinson's disease from offline drawing samples: experiments and findings," in *International Conference on Image Analysis and Processing*, pp. 196–206, Springer, 2019.

[46] C. R. Pereira, S. A. T. Weber, C. Hook, G. H. Rosa, and p. . . y. . . J. P. Papa title = Deep Learning-aided Parkinson's Disease Diagnosis from Handwritten Dynamics, booktitle = Proceedings of the SIBGRAPI 2016 - Conference on Graphics, Patterns and Images

[47] A. Lozano and A. Lang, "Pallidotomy for parkinson's disease," *Advances in neurology*, vol. 86, pp. 413–20, 02 2001.

[48] V. M. Jerkovic, V. Kojic, N. D. Miskovic, T. Djukic, V. S. Kostic, and M. B. Popovic, "Analysis of on-surface and in-air movement in handwriting of subjects with parkinson's disease and atypical parkinsonism," *Biomedical Engineering/Biomedizinische Technik*, vol. 64, no. 2, pp. 187–194, 2019.

[49] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, *Feature extraction: foundations and applications*, vol. 207. Springer, 2008.

[50] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.

[51] I. Siddiqi, K. Khurshid, and N. Vincent, "Feature relevance analysis for writer identification," in *Document Recognition and Retrieval XVIII*, vol. 7874, p. 78740F, International Society for Optics and Photonics, 2011.

[52] M. Maliha, A. Tareque, and S. S. Roy, *Diabetic retinopathy detection using machine learning.* PhD thesis, BRAC University, 2018.

[53] M. Moetesum, I. Siddiqi, F. Javed, and U. Masroor, "Dynamic handwriting analysis for parkinson's disease identification using c-bigru model," in *2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pp. 115–120, IEEE, 2020.

Document Viewer

Turnitin Originality Report

Processed on: 25-Oct-2021 18:45 PKT
ID: 1683596448
Word Count: 9466
Submitted: 1

123 By 123 123

Similarity Index	Similarity by Source	
17%	Internet Sources:	9%
	Publications:	15%
	Student Papers:	3%

include quoted include bibliography exclude small matches mode: quickview (classic) report Change mode print refresh download

2% match (publications)
Siddiqi, Imran, Khurram Khurshid, Nicole Vincent, and Christian Viard-Gaudin. "<title>Feature relevance analysis for writer identification</title>", Document Recognition and Retrieval XVIII, 2011.

1% match (Internet from 14-May-2021)
https://dokumen.pub/image-analysis-and-processing-iciap-2019-20th-international-conference-trento-italy-september-913-2019-proceedings-part-i-1st-ed-2019-978-3-030-30641-0-978-3-030-30642-7.html

1% match (publications)
Momina Moetesum, Imran Siddiqi, Nicole Vincent, Florence Cloppet. "Assessing visual attributes of handwriting for prediction of neurological disorders—A case study on Parkinson's disease", Pattern Recognition Letters, 2019

1% match (publications)
Pedram Khatamino, Ismail Canturk, Lale Ozyilmaz. "A Deep Learning-CNN Based System for Medical Diagnosis: An Application on Parkinson's Disease Handwriting Drawings", 2018 6th International Conference on Control Engineering & Information Technology (CEIT), 2018

1% match (publications)
Maria Teresa Angelillo, Donato Impedovo, Giuseppe Pirlo, Gennaro Vessio. "Chapter 20 Performance-Driven Handwriting Task Selection for Parkinson's Disease Classification", Springer Science and Business Media LLC, 2019

1% match (student papers from 27-Aug-2021)
Submitted to Coventry University on 2021-08-27

1% match (Internet from 27-Nov-2020)
https://www.degruyter.com/view/journals/bmte/64/2/article-p187.xml

1% match (publications)
Momina Moetesum, Imran Siddiqi, Farah Javed, Uzma Masroor. "Dynamic Handwriting Analysis for Parkinson's Disease Identification using C-BiGRU Model", 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2020

1% match (publications)
Rosa Senatore, Antonio Della Cioppa, Angelo Marcelli. "Automatic Diagnosis of Parkinson Disease through Handwriting Analysis: A Cartesian Genetic Programming Approach", 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), 2019

<1% match (Internet from 20-May-2021)
https://dokumen.pub/international-conference-on-innovative-computing-and-communications-proceedings-of-icicc-2020-volume-2-1st-ed-9789811551475-9789811551482.html

<1% match (Internet from 01-Jun-2021)
https://dokumen.pub/deep-learning-for-medical-decision-support-systems-1st-ed-9789811563249-9789811563256.html

<1% match (Internet from 27-Nov-2020)
https://www.mdpi.com/2076-3417/9/21/4666/htm

<1% match (Internet from 14-Mar-2020)
https://www.mdpi.com/2078-2489/9/10/247/html

<1% match (Internet from 22-Jan-2020)
https://link.springer.com/article/10.1007%2Fs00521-020-04735-8

<1% match (Internet from 22-Oct-2020)
https://link.springer.com/chapter/10.1007%2F978-3-319-63312-1_43

<1% match (Internet from 12-Mar-2020)
https://link.springer.com/article/10.1007%2Fs00034-019-01246-3

<1% match (Internet from 12-Apr-2019)
https://link.springer.com/content/pdf/10.1007%2F978-1-59745-462-9.pdf

<1% match (publications)
"Abstracts", Movement Disorders, 2019

<1% match (student papers from 20-Sep-2021)
Submitted to Liverpool John Moores University on 2021-09-20

<1% match (publications)
Drotár, Peter, Jiří Mekyska, Irena Rektorová, Lucia Masarová, Zdenek Smékal, and Marcos Faundez-Zanuy. "Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease", Computer Methods and Programs in Biomedicine, 2014.

<1% match (Internet from 14-Nov-2019)
https://epdf.pub/new-trends-in-applied-artificial-intelligence-20-conf.html

<1% match (Internet from 24-Aug-2021)
https://www.ijimai.org/journal/sites/default/files/2021-08/ijimai6_7_2.pdf

<1% match (student papers from 16-Oct-2020)
Submitted to Federation University on 2020-10-16

<1% match (publications)
Peter Drotar, Jiri Mekyska, Zdenek Smekal, Irena Rektorova, Lucia Masarova, Marcos Faundez-Zanuy. "Prediction potential of different handwriting tasks for diagnosis of Parkinson's", 2013 E-Health and Bioengineering Conference (EHB), 2013

<1% match (publications)
F. Cavaliere, A. Della Cioppa, A. Marcelli, A. Parziale, R. Senatore. "Parkinson's Disease Diagnosis: Towards Grammar-based Explainable Artificial Intelligence", 2020 IEEE Symposium on Computers and Communications (ISCC), 2020

<1% match (publications)
Carlos Alonso-Martinez, Marcos Faundez-Zanuy, Jiri Mekyska. "A Comparative Study of In-Air Trajectories at Short and Long Distances in Online Handwriting", Cognitive Computation, 2017

<1% match (student papers from 04-Jun-2012)
Submitted to University of Ulster on 2012-06-04

<1% match (Internet from 23-Jul-2018)
https://repositorio.ufscar.br/bitstream/handle/ufscar/9299/TeseCRP.pdf?isAllowed=y&sequence=1

<1% match (publications)
Donato Impedovo, Giuseppe Pirlo. "Dynamic Handwriting Analysis for the Assessment of Neurodegenerative Diseases: A Pattern Recognition Perspective", IEEE Reviews in Biomedical Engineering, 2019

<1% match (publications)
Moustafa, Ahmed A., Srinivasa Chakravarthy, Joseph R. Phillips, Ankur Gupta, Szabolcs Keri, Bertalan Polner, Michael J. Frank, and Marjan Jahanshahi. "Motor symptoms in Parkinson's disease: A unified framework", Neuroscience & Biobehavioral Reviews, 2016.

<1% match (student papers from 03-Apr-2018)
Submitted to Vels University on 2018-04-03

<1% match (publications)
"Intelligent Computing Theories and Application", Springer Science and Business Media LLC, 2018

<1% match (Internet from 05-Sep-2021)
http://sersc.org

<1% match (publications)
Moises Diaz, Miguel Angel Ferrer, Donato Impedovo, Giuseppe Pirlo, Gennaro Vessio. "Dynamically enhanced static handwriting representation for Parkinson's disease detection", Pattern Recognition Letters, 2019

<1% match (publications)
Ali Mohammad Alqudah, Hiam Alquraan, Isam Abu-Qasmieh, Alaa Al-Badarneh. "Employing Image Processing Techniques and Artificial Intelligence for Automated Eye Diagnosis Using Digital Eye Fundus Images", Journal of Biomimetics, Biomaterials and Biomedical Engineering, 2018

<1% match (publications)
Iqra Kamran, Saeeda Naz, Imran Razzak, Muhammad Imran. "Handwriting dynamics assessment using deep neural network for early identification of Parkinson's disease", Future Generation Computer Systems, 2021

<1% match (Internet from 16-Jul-2020)
https://elab.engineering.asu.edu/wp-content/uploads/2019/05/Ranadeep_MS_Thesis.pdf

<1% match (Internet from 13-Aug-2019)
https://onlinelibrary.wiley.com/doi/full/10.1002/mds.27087

<1% match (Internet from 02-Oct-2021)
https://www.science.gov/topicpages/a/arc+weld+process

<1% match (student papers from 09-Jul-2021)
Submitted to Aston University on 2021-07-09

<1% match (publications)
Peter Drotar, Jiri Mekyska, Zdenek Smekal, Irena Rektorova, Lucia Masarova, Marcos Faundez-Zanuy. "Contribution of different handwriting modalities to differential diagnosis of Parkinson's Disease", 2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings, 2015

<1% match (Internet from 30-Aug-2018)
https://www.ismrm.org/17/TraditionalPoster.pdf

<1% match (publications)
"Poster Presentations", Movement Disorders, 2013.

<1% match (publications)
Alae Ammour, Ibtissame Aouraghe, Ghizlane Khaissidi, Mostafa Mrabti, Ghita Aboulem, Faouzi Belahsen. "A new semi-supervised approach for characterizing the Arabic on-line handwriting of Parkinson's disease patients", Computer Methods and Programs in Biomedicine, 2020

<1% match (publications)
Clayton R. Pereira, Danilo R. Pereira, Francisco A. Silva, João P. Masieiro, Silke A.T. Weber, Christian Hook, João P. Papa. "A new computer vision-based approach to aid the diagnosis of Parkinson's disease", Computer Methods and Programs in Biomedicine, 2016

<1% match (Internet from 24-Oct-2021)
https://ebin.pub/advances-in-artificial-intelligence-and-data-engineering-select-proceedings-of-aide-2019-1st-ed-9789811535130-9789811535147.html

<1% match (Internet from 23-Jan-2020)
https://escholarship.org/content/qt8f18p61p/qt8f18p61p.pdf?t=pk18xk

<1% match (Internet from 01-Feb-2020)
http://export.arxiv.org

<1% match (publications)
"VIII Latin American Conference on Biomedical Engineering and XLII National Conference on Biomedical Engineering", Springer Science and Business Media LLC, 2020

<1% match (publications)
Chawki Djeddi, Imran Siddiqi, Abdeljalil Gattal, Somaya Al-Maadeed, Abdellatif Ennaji. "Influence of codebook patterns on writer recognition: An experimental study", Expert Systems, 2021

<1% match (publications)
Khatib, S., J.P.M. Finberg, F. Artoul, Y. Lavner, S. Mahmood, U. Tisch, H. Haick, Y. Aluf, and J. Vaya. "Analysis of volatile organic compounds in rats with dopaminergic lesion: Possible application for early detection of Parkinson’s disease", Neurochemistry International, 2014.

<1% match (publications)
Kinjal Chaudhari, Ankit Thakkar. "Survey on handwriting-based personality trait identification", Expert Systems with Applications, 2019

<1% match (publications)
Liaqat Ali, Ce Zhu, Noorbakhsh Amiri Golilarz, Ashir Javeed, Mingyi Zhou, Yipeng Liu. "Reliable Parkinson’s Disease Detection by Analyzing Handwritten Drawings: Construction of an Unbiased Cascaded Learning System Based on Feature Selection and Adaptive Boosting Model", IEEE Access, 2019

<1% match (publications)
Maria Teresa Angelillo, Fabrizio Balducci, Donato Impedovo, Giuseppe Pirlo, Gennaro Vessio. "Attentional Pattern Classification for Automatic Dementia Detection", IEEE Access, 2019

<1% match (publications)
Shail Raval, Rahil Balar, Vibha Patel. "A Comparative Study of Early Detection of Parkinson's Disease using Machine Learning Techniques", 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), 2020

<1% match (publications)
Yuanyuan Liu, Nelly Penttila, Tiina Ihalainen, Juulia Lintula, Rachel Convey, Okko Rasanen. "Language-Independent Approach for Automatic Computation of Vowel Articulation Features in Dysarthric Speech Assessment", IEEE/ACM Transactions on Audio, Speech and Language Processing, 2021

<1% match (publications)
Zhaohan Zhang, Mu Li, Katharine Flores, Rohan Mishra. "Machine learning formation enthalpies of intermetallics", Journal of Applied Physics, 2020

<1% match (Internet from 02-Mar-2020)
https://res.mdpi.com/bookfiles/book/1394/Selected_Papers_from_the_218_41st_International_Conference_on_Telecommunications_and_Signal_P v=1583137276

<1% match (publications)
Claudio Loconsole, Giacomo Donato Cascarano, Antonio Lattarulo, Antonio Brunetti et al. "A comparison between ANN and SVM classifiers for Parkinson’s disease by using a model-free computer-assisted handwriting analysis based on biometric signals", 2018 International Joint Conference on Neural Networks (IJCNN), 2018

<1% match (publications)
Gennaro Vessio. "Dynamic Handwriting Analysis for Neurodegenerative Disease Assessment: A Literary Review", Applied Sciences, 2019

<1% match (publications)
"Intelligent Computing Theories and Application", Springer Science and Business Media LLC, 2017

<1% match (publications)
João Paulo Folador, Maria Cecilia Souza Santos, Luiza Maire David Luiz, Luciane Aparecida Pascucci Sande de Souza et al. "On the use of histograms of oriented gradients for tremor detection from sinusoidal and spiral handwritten drawings of people with Parkinson’s disease", Medical & Biological Engineering & Computing, 2021

<1% match (publications)
Mahmood Saleh Alzubaidi, Uzair Shah, Haider Dhia Zubaydi, Khalid Dolaat, Alaa A. Abd-Alrazaq, Arfan Ahmed, Mowafa Househ. "The Role of Neural Network for the Detection of Parkinson’s Disease: A Scoping Review", Healthcare, 2021

CHAPTER 1 INTRODUCTION Parkinson’s disease is the second most common neurologi- cal disorder after Alzheimer’s [1]. Parkinson’s disease (PD) affects around 10 million persons globally [2].PD affects 1–2 per 1,000 of the population [3, 4]. PD affects 1% of the population over the age of 60, but is uncommon in people under the age of 50 [4, 5]. The prevalence of Parkinson’s disease rises with age, reaching around 4% in the oldest age groups [5, 6]. These prevalence rates are ex- pected to any increase because of the population[7]. Parkinson Disease (PD) is characterized by motor symptoms and non-motor symptoms including akinesia, bradykinesia, rigidity, and tremor, postural imbalance and vocal disabilities [8, 9]. Traditional di- agnostic procedures for the diagnosis of PD include neuroimaging strategies such as SPECT and CT scans, shown in Figure ??, which shows vital potential within the determination of PD however needs expensive instrumentality. Furthermore, these strategies are com- pelling only when the disease has progressed to the final stage, further highlighting the complexities of PD analysis.[10].According to clinicopathological research [11, 7],up to 25% of PD patients are misdiagnosed within last stages of their illness. As a result, there's a lot of work being done to develop accurate systems for detecting and diagnosing Parkinson's disease in its early stages. Figure 1.1: SPECT scanning of PD patient. With the advancement of technology, researchers are able to propose many solutions and decision support systems to identify the early stage of PD patients. Some of the studies [12] used sig- nal acquisition through wearable sensors monitoring free muscular movements to predict PD, while other studies [13] used breath or voice analysis [14, 15, 16, 17] to predict PD. Voice processing for diagnosis of PD offered very promising results by achieving 98% overall classification accuracy [18] Likewise, Bradykinesia(slowness of movement), in the literature is directly related to handwriting. Some of the recent studies [19, 20] recommended that handwriting is often used as a good tool for early diagnosing of PD and a few preliminary pieces of knowledge suggest that handwriting would possibly function as a diagnostic marker for PD diagnosis by iden- tifying micrographia. The idea is illustrated in Figure 1.2 where a PD subject attempts to write a sentence, over the period of time, handwriting starts deteriorating. Initially, the size of the letters and horizontal alignment are fine. However, it becomes hard for the PD subject to maintain the size and alignment of words, and the words at the end are almost impossible to read. Figure 1.2: Handwriting of patient suffering with Micrographia. Tremors damage handwriting because the involuntary oscil- lating movement of one or more body parts of the patient, as de- picted in Figure.1.3, causes the hands or fingers to twitch slightly while the patient writes or draws something. Figure 1.3: Subject with Parkinson's disease (b) healthy subject. Another symptom known as Bradykinesia is in which the pa- tient's handwriting speed is slow and the graphomotor task takes longer time than usual. Some researchers collected data using gad- gets (digitizers or tablets), while others used hand-drawing shapes to come up with solutions and hypotheses for their study. Several preliminary studies have suggested that handwriting can be used as an effective non-invasive tool for the early diagnosis of PD. So, based on these considerations we have attempted to develop a sys- tem that is specifically designed for early detection of Parkinson's disease. As a result, we'll use Drotar et al dataset that is intended for Parkinson's disease patients. 1.1 Problem Statement Identification of Parkinson's disease through modalities like handwriting or speech has been thoroughly investigated in the lit- erature. The correlation between PD and changes in writing pat- terns has also been established in a number of studies [21, 10]. From the perspective of handwriting analysis, a number of static (offline) [10] and dynamic (online) [21] features have been identified that can serve as effective indicators of PD. Combining online and offline features are also known to improve the identification perfor- mance [21]. In most cases, the identified attributes are mapped to computational features which are extracted from established tem- plates and are fed to a classifier to determine the presence or ab- sence of PD. An important factor in choosing the type of features is the drawing or handwriting template under study. While the previous studies primarily target combining features or decisions on multiple templates to enhance the overall performance, to the best of our knowledge, no investigations have been carried out to study the relationship between the template under study and the computed features. Dynamic information could be more useful for certain templates while for others, static or visual information can provide useful clues. The proposed research is aimed at feature relevance analysis for the identification of PD through handwriting. More specifically, we intend to carry out a comprehensive study using different feature selection techniques to assess the optimal set of features for this problem. Furthermore, some features may be more appropriate with specific templates hence we also aim to study which features are more informative as a function of the template from which they are extracted. 1.2 Research Objectives The objectives of Research include the following. • To combine the online and offline attributes of writing and study the system performance • To study the relevance of both static and dynamic features of handwriting in identification of PD. • To investigate the performance of different features as a func- tion of acquisition template and identify the relevant set of features for a given template. 1.3 Research Contributions The research carried out in this study has resulted in the de- sign and development of a a system that predicts Parkinson disease through computerized analysis of handwriting The main purpose of the proposed study is to design and development a system that can predict Parkinson's disease by use of computerized handwriting analysis. The key contribution of the research is the manipulation of the offline and online features to identify a relevant set of fea- tures that can predict the absence and presence of PD. In case of no availability of specialized hardware devices to directly cap- ture online handwriting, offline attributes can be useful Features extracted from control subjects and PD patients are fed to feature selection techniques to assess the optimal set of features, and that optimal set of features is then fed to a learning algorithm to learn to discriminate between the two classes. Artificial Neural Network (ANN) and Support Vector Machine (SVM) classifiers are investi- gated for this purpose. Experiments on a benchmark dataset report promising classification rates. 1.4 Thesis Organization This document is organized as follows. Chapter 2 presents a discussion on the work related to prediction of Parkinson disease from handwriting Chapter 3 describes the method that we have adopted in order to achieve the objectives along with the key con- cepts behind the approaches. Chapter 4 outlines the metrics used to test our methods, describes the experiments, presents the find- ings we obtained and their interpretation. Chapter 5 incorporates the concluding remarks and recommendations for future work. CHAPTER 2 LITERATURE REVIEW Parkinson's disease is caused by the loss of pigmented neu- rons in the midbrain region's substantia nigra, which control muscle movements. Dopamine, a neurotransmitter involved in the control and regulation of body movements, is reduced when these neurons are lost. This causes tremors, sluggish movements, hypertonia, and balance issues. [22] These symptoms have an effect on the individ- ual's hand-wrist movements, which have a negative impact on his or her handwriting. Computer-aided handwriting analysis allows for the identification of prospective patterns that may be useful in the detection and classification of Parkinson's disease.Several studies[19, 23] have been published that indicate handwriting the analysis is an effective tool for PD diagnosis. Many handwriting features were proposed in the in the writing for the identification of PD[10, 23, 24, 25]. Based on their technique of knowledge acquisi- tion, extracted features can be classified into two types: Static and Dynamic. Static features will be taken from offline handwriting samples, whilst dynamic features will be derived from online hand- writing samples. These studies used a variety of machine learning techniques to examine the static and dynamic features' ability to discriminate PD. In this chapter, we will discuss related work on handwriting analysis and potential strategies used for early Parkin- son's prediction. 2.1 Online Features Analysis Handwriting requires the participation of various body parts such as fingers, arms and also includes our motor neurons, a healthy person manages the participation of all parts for the writing task, however when we perform a writing task to the patient, the motor neurons do not function properly.A number of solutions for de- tecting Parkinson's disease and other similar disorders have been developed in recent years, one of which is wearable sensors that are attached to the patient's body. In 2011,[26], they integrated their device with smart gloves, which detected the level of motor dysfunc- tion in PD using smart gloves and assessed the movement of fingers while writing, making non-invasive approaches more effective and less expensive. In 2013, Dortar et al. [23] created a handwriting based dataset, which acquires handwritten signals (on-surface and in-air) using a digitizing tablet Intuos 4M and presented a template consisting of seven completely different handwriting tasks with an addition Archimedean spiral drawing task .In this study, they evaluated three types of features, i.e supported in-air movement, primarily based on-surface movement and combination of both groups of fea- tures to effectively diagnose PD. 75 samples in which 38 patients and 37 healthy subjects were employed. They applied classification using SVM (Support vector machine) and achieved a classification accuracy of 80% using 16 features selected from in-air movement. In 2014, [27] Drotar et al., combined various online in-air and on- surface features by using feature selection techniques and a sup- port vector machine learning classifier to discriminate PD patients from healthy controls, attaining an accuracy of 85%. In a subse- quent study [7], The authors achieved 88.13% accuracy using the SVM classifier with radial kernel for automated diagnosis, work- ing with kinematic and spatio-temporal handwriting measures as well as handwriting measures including entropy, signal energy, and empirical mode decomposition. In 2015, authors extended the sim- ilar work [28], by using a combination of dynamic features and achieving 89% area under the ROC curve (AUC) for PD classifi- cation. In 2016, authors [29], used online kinematic and pressure features of handwriting to train different classifiers and achieving 81.3% accuracy with SVM, 78.9% with AdaBoost classifier, and 71% with KNN, respectively. In the sequence of experiments, au- thors additionally prompt that performance of identification of PD depends on the selection of template used. After a year Pereira at al [30], have acquired the NewHandPd dataset, which includes both off-line images and on-line signals (extracted from the smart pen). Each person fills the structure by composing on paper with a digitising pen and drawing four spirals and four meanders. Di- verse machine learning algorithms, such as Convolutional Neural Networks (CNN), Support Vector Machines (SVM), Optimal Path Finder (OPF), Random Forest (RF), and Restricted Boltzman Ma- chines (RBM), were evaluated, revealing that when on-line data is used, the CNN ImageNet design could achieve precision of 87.14 % in its best setup, whereas the SVM achieved the highest perfor- mance on off-line data, with an accuracy of 66.72% Researchers have not only used handwriting features to clas- sify PD, but they have also used posture features. In 2014, Graca et al. [31] used mobile devices to predict PD by completing a different tasks (spiral analysis, gait analysis, tip analysis) in which 35 samples were collected from drawing Archimedes spiral. They extracted various features (Spatio-temporal , pressure features and gait features) and fed these features to the different classifier such as C4.5, RipperK and Bayesian network for classification. The re- ported accuracy with mentioned classifiers were 86.67%, 80.83% and 87.50% respectively. In 2018, Impedovo et al. [8] used PaHaW [23] database to investigate those specific dynamic features of the handwriting that can help to idenify people suffering from PD . They worked on online kinematic features were employed six clas- sifiers including SVM (RBF, linear), KNN, LDA (Linear Discrimi- nant Analysis), NB (Gaussian Naiïve Bayes), RF (Random Forest), ADA (AdaBoost) reporting accuracy's of 71%, 68%, 67.90%, 66%, 57%, 73%, 61% on 72 subjects, respectively. In another valuable research by Angelillo et al. [32] in 2019, the researcher retrieved features from the raw data of different tasks using the PaHaW dataset, which comprises many tasks done by similar subjects, by utilising the dynamics of the handwriting pro- cess. Techniques such as Shannon and Renyi entropy, signal-to- noise ratio, and empirical mode decomposition (EMD) were used to figure both on-surface and in-air horizontal and vertical parts of handwriting.After extracting features, the prescient capability of every task is assessed exclusively and the best tasks, i.e. those with the most noteworthy forecast, are fed into a group of clas- sifiers(SVM, AdaBoost, Logistic Regression, Linear Discriminant), whose predictions are obtained via majority voting and its achieved highest classification accuracy of 88.33%.In 2019 [33] Cartesian Ge- netic Programming is a technique for detecting Parkinson's disease (PD) by analysing the handwriting of PD patients and healthy con- trols. The adoption of such an approach is particularly intriguing because it allows for the inference of explicit classification models while also allowing for the automatic identification of an appropri- ate subset of features relevant for a correct diagnosis. The approach was tested using characteristics collected from handwriting exam- ples in the PaHaW dataset, which is freely available. In 2020,Am- mour et al. [34] worked on the Arabic Handwriting dataset and extracted the number of features of different categories like Kine- matics on surface In-air, Mechanical, Inclination, Pen Up features and used the semi-supervised approach for classification(Clustering and PCA) obtaining 97.3% of classification accuracy. In same year Amina Naseer et al. [17] worked on PaHaW dataset and per- formed features extracted via CNN- Alexnet pre-trained model.The selected features were fed to SVM classifier for PD identification and obtained 98.28% of accuracy.In the same year, another research by 2.2 Offline Features Analysis In other studies, certain authors did not used any dataset col- lection system, they use hand-drawing samples and shapes.In 2015, Pereira et al. [24] have collected the HandPd is a dataset composed of images extracted from handwriting exams of 92 people divided into 18 healthy people (Healthy Group) and 74 patients (Patients Group). They worked on automatic Parkinson's disease diagno- sis using spirals and meanders in forms as shown in Figure 3.1, that are then compared with the template for feature extraction, which was assessed employing three methods: Nave Bayes (NB), Optimum-Path Forest (OPF), and Support Vector Machines with Radial Basis Function (SVM-RBF), with the best results on the NB classifier that gave around 79% order accuracy. This study ad- ditionally indicated that meander samples play a very important role, resulting in higher accuracy than spiral samples. In 2017, Loconsole at el. [35] used a limited number of fea- tures extracted from EMG (ElectroMyoG raphy) signals obtained at the arm level (time feature) and scans of traditional paper sheets (vision-based features) by utilising computer vision and applied an Artificial Neural Network-based classifier employing a Multi- Objective Genetic Algorithm (MOGA) achieving 95% accuracy. In 2018, Khatamino etal. [36] used HandWritten datasets that comprise of the Static Spiral Test (SST), the Dynamic Spiral Test (DST) and Stability Test on Certain Point (STCP) of 57 patients and 15 control healthy individuals [37].Author used a CNN-based deep learning approach and accomplished a precision of 88%. In the same year, Momina et al. [10], utilized Convolutional Neural Networks(using the Alex-Net pre-prepared model)to extract visual features from numerous representations of different graphomotor tests delivered of 72 subjects (Patient and Health Group) sub- jects.These features are fed to a Support Vector Machine (SVM) classifier accomplishing accuracy of 83% In 2019, Diaz et al. [38] worked with PAHAW offline data(images) that extracted features from CNN using a pre-trained VGG16 net- work. To reduce overfitting, the authors applied feature selection al- gorithm before classification.they applied different classifier (SVM, Random forest) achieved accuracy of 86.76% and also the exam- ined which handwriting task performed better than other .Another study by Ribeiro et al. [39] used same dataset for the classification of PD and used Recurrent Neural Network(RNN) achieving 85% accuracy at the spiral and 89% on the meander. In 2019, Gupta et al.[40] used PaHaW off-line hand-drawn Archimedean spiral data and presented a novel distance based features PD prediction by extracting Fourier Transform based distance features,Tremor Esti- mation feature and combined distance-based features and fed these extracted features to the SVM classifier for classification and the reported accuracy of 81.66%. In 2019, another author Rosa et al. [33] proposed an evo- lutionary approach to discriminate PD using hand shape analysis .they applied Cartesian Genetic Programming on a set of static features on HandPD dataset to show which handwriting template performed better. The results of the experiments indicated that the features derived by spirals are less informative than those derived by meanders, and that the global accuracy reached by meander analysis outperforms that of other studies.Their study also showed that, in its best configuration, the CGP performs better than state- of-the-art techniques for PD diagnosis proposed in the literature. Author Year Dataset Handwriting Task Features Analysis Results Drotar et al. 2013˜ Parkinson's Letters, Words, Sentences and Archimedean Spiral Online in-air Surface Features SVM 80.09% Drotar et al. 2014 PaHaW Letters, Words, Sentences and Archimedean Spiral Online in-air and on-surface Features SVM 85% Drotar et al. 2014 PaHaW Letters, Words, Sentences and Archimedean Spiral Online Spatial, Temporal Kinematic, Entropy, Signal Energy, SVM 88% Graca˜ et al. 2014 Graca's Dataset Archimedean Spiral Online Spatial-Temporal and Pressure Feature C4.5, RipperK , Bayesian Networks 86.67% 80.83% 87.50% Dortar et al. 2015 PaHaW Letters, Words, Sentences and Archimedean Spiral Online Spatial, Temporal Kinematic, Entropy, EMD and Pressure SVM ˜89.09% Pereira et al. 2015 HandPD Archimedean Spiral Offline Mean Relative ˜Tremor(MRT) and Spatial Features Naïve Bayes (NB), Optimum-Path Forest (OPF), SVM 78.90% 77.10% 75.80% Dortar et al. 2016 PaHaW Letters, Words, Sentences and Archimedean Spiral Online˜Kinematic, Pressure Features SVM ADABOOST K-NN 81.3% 78.9% 71 % Pereira et al. 2016 NewHandPD Archimedean Spiral and Meander ˜Pen-based Features CNN OPF 87.1% on Meander Tasks Laconsole et al. 2017 Laconsole Dataset Sentence, repetitive loops Online and Offline features ANN 95% Impedovo et al. 2018 PaHaW Letters, Words, Sentences and Archimedean Spiral Spatial, Temporal Kinematic, Entropy, Signal Energy, EMD, Pressure SVM (RBF, Linear), KNN, LDA NB , RF ,AdaBoost 71% 68% 67.90% 66% 57%% 61% Table 2.1: Summary of related works on Handwriting based Parkin- son Prediction 2.3 Benchmarking Datasets In any research domain, the availability of datasets is one of the key requirements for the analysis of neurological disease.Collection of datasets is a very difficult activity in medical field since it presents a particular problem for selecting participants, choosing a acquisi- tion device, and finding the most suitable handwriting tasks. The number of dataset use for the prediction of PD are discussed below. In this section, the datasets that are used by in pervious techniques to evaluate their approaches have been reviewed. • PAHAW Dataset: This dataset consists multiple hand- writing samples from 37 people with Parkinson's disease (19 men/18 women) and 38 healthy people (20 men/18 women). Author Year Dataset Handwriting Task Features Analysis Results Khatamino et al. 2018˜ HW dataset Archimedean Spiral Dynamic and Visual features CNN 88% Angelillo et al. 2019˜ PaHaW Letters, Words, Sentences and

Archimedean Spiral Online Spatio-Temporal Kinematic SVM, AdaBoost, Logistic Regression, Linear Discriminant 88.33% Ribeiro et al. 2019 New HandPD Archimedean Spiral and Meander Kinematic , Spatio-Temporal RNN 85% (Spiral) 89% (Meander) Diaz et al. 2019 PaHaW Letters, Words, Sentences and Archimedean Spiral CNN based Visual Features SVM 86.76% Parziale et al. ˜ ˜ ˜ ˜ ˜ ˜ ˜ ˜ ˜ 2019˜ ˜ ˜˜ HandPD Archimedean Spiral and Meander Offline Mean Relative Tremor and Spatial Features SVM Decision Tree Random Forest 73.63% Gupta et al. 2019 PahaW Archimedean Spiral Spatial Features SVM 81.66% Alae et al. 2020 Arabic Dataset Arabic Text Online Kinematics on surface/In-air Mechanical, Inclination, Pen Up. Clustering and Principal Component Analysis(PCA) 97.3% Amina et al. 2020 PaHaW Letters, Words, Sentences and Archimedean Spiral CNN(Alex-net) based Visual Features SVM 98.28% Ismail et al. 2020 HW dataset Archimedean Spiral CNN based Visual Features using AlexNet or GoogleNet SVM KNN 98.28% Table 2.2: Summary of related works on Handwriting based Parkin- son Prediction The samples came from the Movement Disorders Center at Masaryk University's First Department of Neurology and St. Anne's University Hospital in Brno, Czech Republic. Each participant was given eight handwriting assignments to com- plete at their own pace. The signals were captured using a Wacom Intuos 4M digitising tablet with a sampling rate of 150 Hz. Each individual perform eight tasks according to template illustrated in Figure 2.1 Figure 2.1: Template proposed by [27] • HandPD dataset: This dataset contains 92 individuals, 18 of Healthy Group and 74 of Patients Group, the latter be- ing composed of people suffering from Parkinson's Disease (PD). Botucatu Medical School, So Paulo State University - Brazil, gathered the handwritten exams. The main task in- cludes filling out a form that consists of four spirals and four meanders.[24] as shown in Figure 2.2 Figure 2.2: HandPD Dataset tasks • NewHandPD Dataset: This dataset is composed of 66 in- dividuals that are obtained from 35 Parkinson's patients (21 males and 10 females) and 31 healthy subjects (18 males and 17 females).Every individual was approached to draw 12 ex- ams, 4 spirals, 4 meanders, 2 circled movements (one circle in the air and another on the paper). Some handwritten dy- namics features were likewise recorded utilizing an advanced pen, having images from 4 spirals,4 meanders ,4 circles and signals for all 12 exams.So every individual sample includes 9 images and 12 signals.[41] • Graca et al. Dataset In 2014, Graca et al. [31] used mobile devices to predict PD by completing a different tasks (spiral analysis, gait analysis, tip analysis) in which 35 sam- ples were collected from drawing Archimedes spiral as shown in Figure 2.3.They allow the user to perform the task on a mobile screen because it's easy to manage. Figure 2.3: Image sample of grace et al.[31] dataset • Arabic Handwriting Dataset: Arabic Handwriting dataset used for PD prediction with 28 Parkinson's patients and 28 healthy subjects. This data set completed with three tasks shown in Figure ??. Figure 2.4: Arabic Handwriting dataset • The Hand written Dataset: The Hand-Written (HW) dataset was gathered at Istanbul University's Cerrahpasa Fac- ulty of Medicine's Department of Neurology[23, 42]. This dataset contains time-series data from handwriting spiral ex- ams of individuals in two groups: healthy people and Parkin- son's disease. The dataset contains 72 individuals, 57 of whom are patients and 15 of whom are healthy controls. Exams are provided to everyone in the same way (recommend to draw inward to outward). As illustrated in Figure 2.5, it consists of three types of handwriting tests: the Static Spiral Test (SST), the Dynamic Spiral Test (DST), and the Stability Test on Certain Point (STCP). • Mirjana et al dataset In this study[43],There were 43 par- ticipants in total: 33 patients with Parkinson's disease and ten healthy controls (HC). This dataset consists of 4 tasks, 1. Writing a sentence between two lines(a distance of 1 cm when looking at the laptop) 2. Typing a sentence between two lines(a distance of 1 cm, with the monitor out of sight) (a) SST Test (b) DST Test Figure 2.5: Hand Written dataset 3. Writing a paragraph without space restriction while look- ing at the screen 4. Writing a sentence without space restriction, with the monitor out of sight Figure 2.6: Signal Tasks. 2.4 Summary This chapter presented an overview of the techniques pre- sented for the identification of PD using handwritten analysis.Recent studies primarily target both static and dynamic features or se- lected features to enhance the over all performance, to the best of our knowledge, no investigations have been carried out to study the relationship between the template under study and the computed features.Our study will combined both A summary of related works (2013-present) can can be seen in Table CHAPTER 3 METHODOLOGY In the preceding chapter, significant contributions in the field of Parkinson disease identification using handwriting analysis was discussed. In most situations, the discovered attributes are mapped to computational features derived from known templates and input into a classifier to identify whether or not PD exists. The drawing or handwriting template under study is a key aspect in determin- ing the type of features to use. While past research has focused on combining features or decisions from various templates to improve overall performance. In this chapter, the methodology used for fea- ture relevance analysis for the identification of PD through hand- writing is explained in length. Workflow for proposed approach is shown in Figure 3.1. The goal of our proposed system is to assess the optimal set of features and to study which features are better suited to specific templates. Figure 3.1: Proposed Approach. 3.1 DataSet Data acquisition is a very difficult task when it comes to med- ical problems. In our research, we used the Parkinson's Disease Handwriting Database (PaHaW). This database consists of samples obtained from 37 Parkinson's patients (19 males and 18 females) and 38 healthy subjects (20 males and 18 females). All members in- volved in PD diagnosis enlist from the movement disorder center at The Department of Neurology, Masaryk University, and St. Annes Hospital in Smo, Czech Republic. And all samples write in the native language of the participants and the participant completed all tasks according to the template. All tasks are the following: 1. Drawing an Archimedes spiral 2. Writing in cursive the letter l 3. The bigram le 4. The trigram les 5. Writing in cursive the word lektorka ("female teacher" in Czech) 6. porovnat ("to compare") 7. nepopadnout ("to not catch") 8. Writing in cursive the sentence Tramvaj dnes uz nepojede ("The tram won't go today") Figure 3.2: PAHAW template. 3.2 Acquisition Device The number of devices used for data acquisition, in our prob- lem the authors used Wacom Intuos 4 M digitizer or digital pen for data collection shown in Figure 3.3, some patients may be un- familiar with the device, so patient writing on paper, the paper is set on the digitizer. By this device, some useful features have been acquired: (x-y) the coordinate of the pen moves in a different direc- tion, the time stamps, the pen Orientation (azimuth and altitude) and the pressure, and if the pen moves in-air the button state is 0 otherwise (on the surface) the button state is 1. All the features reported are the numeric values shown in Figure 3.3 3.3 Data Prepossessing The Acquisition Device discusses in the previous section use for collecting the pen-based data. These included all of the func- tional attributes that could be used for the derived kinematics fea- tures. All of these features are sequentially measured within the same time intervals. In Literature, most authors work with these sequential features and measure only mean values, and then feed (a) Digitized smart pen (b) Features Acquire by device Figure 3.3: Feature acquired using digitized pen them to the model[2]. But when they transform the sequential value into mean values, the beneficial information is lost. Therefore, in this study, we use all sequential values and give the classifier as it is an emphasis on the output of sequential data analysis. 3.4 Features Extraction In this section extracted features employed, offline and online features of handwriting discuss in a later section. 3.4.1 Online Features Online features provide valuable information for the diagnosis of PD. In our research, we are working on online features that have been calculated from raw data in the PaHaW database. There are a variety of features: (x-y) coordinate, azimuth, altitude, pressure, timestamp, button status. By using these features, new features derived such as velocity, acceleration, distance and also calculated the mean, standard deviation, median, 10th, and 100th percentile of the whole signal acquired by the device give it to the predictor classifier. 3.4.2 Offline Features In offline features, these features are derived by handwrit- ing and hand-drawing. Some researchers used offline functionality instead of online attributes[24]. In 2015, Pereira used the hand- drawing shape to extract the features. In our case study, we used the dortar et al. dataset. [27] The original dataset contains online features extracted from the device (X, Y coordinates, button states, pressure) so that we can convert all online features to images, on- line features in the form of numeric data, and by plotting x, y coordinates into images shown in Figure 3.5 Figure 3.4: Offline image of the Archimedean spiral (a): Healthy Subject (b): Parkinson Patient All attributes or features of the dataset includes information about PD and healthy subject, by using visualization techniques and apply different filters on images for features extraction and enhanced the dataset because original image not enough to train the dataset. We feed the images to pre-trained model VGG16. We trained the model on our data and extracted features from the last FC layers which contain 4096 features. • VGG16 Visual Geometry Group at Oxford introduces Vgg16. This is a 16-layer convolutional neural network. The model uses a set of pre-trained weights from ImageNet. In ImageNet, a dataset of over 14 million images belonging to 1000 classes, the model achieves 92.7 % accuracy.

The default RGB image input size for the VGG16 model is 224 x 224 pixels with three channels.The used architecture of VGG16 is summarized in Figure 3.5. Figure 3.5: VGG16 Architecture. 3.5 Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">Feature Selection Feature selection is the process of Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">selecting relevant and in- formative Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">features with the motivation of data/feature set reduc- tion, Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">performance improvement, and data understanding[44]. The Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">primary goal of a feature selection procedure would be to find Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">the features (or feature components) that are useful Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">in identify- ing Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">the presence and absence Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">of PD. Filter, wrapper, and em- bedding approaches are the three basic kinds of feature selection algorithms[45]. Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">Generation, evaluation, stop criterion, and valida- tion are Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">the four key processes of a feature selection approach. A search strategy is used in the generation process to obtain a Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">subset of features (usually utilising Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">forward selection, backward removal, Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">bidirectional, and other methods). Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">The efficiency Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">of the resulting Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">subset is then evaluated using Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">an evaluation criterion, which might Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">be independent (filter) or dependent (measurement) (Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">wrapper). Af- ter Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">each iteration, a stopping condition is examined Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">to decide Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">when the selection process should be terminated. Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">Typical criteria involve achievement of optimal subset or bounds on a Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">number of features or iterations etc. Once Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">the stopping condition is met, the resultant subset of features can be confirmed [46]. Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">For our problem, we employed Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">a genetic algorithm (GA), a wrapper approach a Correlation, a filter approach for feature selec- tion. 3.5.1 Genetic Algorithms Genetic algorithm is one of the most advanced feature selec- tion algorithms. It is a stochastic function optimization method based on natural genetics and biological evolution mechanics.In nature, organisms' genes tend to evolve through generations to im- prove their ability to adapt to their surroundings. It acts on a population of individuals to better approximations over time. A state diagram for the feature selection process with the genetic al- gorithm is shown in Figure 3.6 . Figure 3.6: Genetic Algorithm. As with natural adaptation, this process results in the evo- lution of populations that are better suited to their environment than the individuals from which they were formed. This technique has an advantage over others in that it permits the best answer to emerge from the best of previous solutions. In this study, we used the basic application of genetic algorithms as the objective of our system is to select an optimal set of features that provides better performance than all Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">features. We will first analyze the relevance of the features i.e 97 online and 4096 offline features. The Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">GA is used to generate individuals of length (97 & 4096) Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">and the set bits are used to select the respective features. We executed Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">the ten times and extracted features that are almost selected every time we runs GA.We used the following parameters for GA: • Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">Population Size: 50, • Crossover Rate: 0.5, • Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">Mutation Rate: 0.2, • Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">Selection Rule: logistic regression • Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">Number of Generations: 10. Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">The initial population is generated at random. Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">The fitness function is used to evaluate the chromosomes in each generation, with the present population's fitness values being utilized to find the offspring of the next generation. Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">When the specified Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">number of generations has Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">been evaluated, the procedure comes to an end. The Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">best individual of the final generation determines the selected feature subset. The division Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">of online and offline Feature relevance analysis for writer identification", Document Recognition and Retrieval XVIII, 2011.">features according to their relevance is explained in the next Chapter 4 3.5.2 Correlation Another approach for feature selection is Correlation. It's a metric for determining the degree of linear correlation between an input feature and an output feature. It has a range of +1 to - 1, with 1 denoting total positive correlation and -1 denoting total negative correlation. As high correlation features are more linearly dependent, they have roughly the same impact on the dependent variable. When there is a strong correlation between two features, one of them may be dropped. The correlation mathematical for- mula is shown in COV (X, Y) PX,Y = σXσY 3.6 Classification Classification is an important part of research because evalu- ating the quality of the literature, we are providing the best results on this disease. In the literature survey, many techniques used for classification the most commonly used help vector machine, Ran- dom forest, Naive Bayes, neural networks, etc. Some researchers used a combination of classifiers and often used several neural net- works to improve overall accuracy. In our implementation, we use three classifiers support vector machines.We applied this classifier on online and offline features data extracted by the feature rele- vance method. 3.6.1 SVM SVM (Support Vector Machine) is a supervised machine learn- ing model for binary and regression problem classification. Each data item is represented as a point in n-dimensional space (where n is the number of features), with the value of each feature being the value of a certain position in the SVM algorithm[47]. In general, SVM is divided into two types: linear and non-linear. A linear SVM computes a linear decision boundary using a linear kernel. Figure 3.7 shows a two-dimensional data example of a linear SVM. Figure 3.7: SVM classifier. For higher dimensions, planes or hyper-planes are computed. A Non-linear SVM (Figure 3.8) uses a non-linear kernel. The ad- vantage of using a non-linear kernel is that it can capture and cal- culate much more complex and complicated relationships between the data points without performing extremely complicated and dif- ficult transformations on its own. But, naturally, non-linear kernels are more complex and time-consuming. Figure 3.8: SVM classifier. 3.7 Summary In this chapter discuss the detail of the PaHaW dataset used in our study and present the online and offline extracted features. A brief overview of feature selection and SVM classifier examined our work also present. In the next chapter, we will discuss the results and different experiments. CHAPTER 4 ANALYSIS & RESULTS This chapter describes the specifics of all experiments, com- putes the efficiency of the different models, and examines the effec- tiveness of extracted features in each task to identify the presence and absence of PD. All experiments performed by Drotar et al. dataset(PaHaW) are presented in section 2.3. In this chapter, we first discussed training and testing data and then show results using Feature selection techniques and classification models. 4.1 Training Test Datasets The PaHaW dataset contains 75 sample files, all of which are used for the experiment (37 Parkinson's patients and 38 healthy subjects). When we extracted the features from 75 samples, we used folding techniques for rotating 75 samples, In our scenario, we divide 75 sample data into fivefold. By using this technique, we are able to estimate the skill of our model on unseen data. 4.2 Performance Metrics The functionality of the proposed system is evaluated by using standard measures of accuracy. Each of these is briefly described and figure out the accuracy of each task with the optimal set of features using SVM Classifier. We discussed the Standard measures are as follows. • True Positive Data instance belongs to a specific class and is correctly classified by the algorithm that data belongs to the same class. In the case of PD identification, a PD subject is correctly classified as PD. • False Negative The algorithm detects that the data does not pertain to a specific class, however it belongs to that class. In other words, a PD subject is wrongly classified as Healthy. • False Positive The data instance does not belong to a spe- cific class, but it is incorrectly identified by the algorithm as belonging to that class. In this case, a healthy subject is wrongly classified as a patient. • True Negative The algorithm identifies that data does not belong to a specific class; however, the data actually belongs to another class. In this case, a healthy subject is correctly classified as healthy. 4.2.1 Accuracy The accuracy calculate the ability of overall system to pre- cisely classify the PD patient and healthy subject. Accuracy = TP + TN TP + TN + FP + FN 4.3 Results and Discussion In this chapter, we will explain the performance of the sys- tem,we present the accuracy of every single task with all features and relevance features set and then fed it to the classifier discussed in chapter 3.6. There are different sets of experiments used for classification with online data and offline data. 4.3.1 Results on all features (Online and Offline) In this experiment, we simple classification model was applied on every single task for classification and the result mentioned in the below table 4.1 Tasks Online Features (97) Offline Features (4096) Archimedean Spiral 57% 57% Repetitive(l) 60% 43% Repetitive(le) 62% 57% Repetitive(les) 60% 64% Word(leplorka) 50% 64% Word(porovnal) 53% 50% Word(nepopadnoul) 47% 64% Sentence 67% 64% Overall Accuracy 57% 58% Table 4.1: Task-Wise Accuracy on All features 4.3.2 Results Using GA technique (Online and Offline) In this experiment, we applied the Genetic Algorithm Fea- ture selection technique on each task features set and extracted the optimal feature subset. We performed 10 iterations to deter- mine specific features that were selected almost every time we runs the Genetic Algorithm. We fed these selected features data to a machine learning classifier i.e. SVM for PD classification and the obtained result after Genetic feature selection techniques are men- tioned in the below table 4.2,4.3 Tasks No of Features Selected SVM Archimedean Spiral 6 78% Repetitive(l) 8 74% Repetitive(le) 5 75% Repetitive(les) 8 80% Word(leplorka) 5 60% Word(porovnal) 5 80% Word(nepopadnoul) 5 74% Sentence 6 73% Overall Accuracy 74% Table 4.2: GA Online features results Tasks No of Features Selected SVM Archimedean Spiral 65 86% Repetitive(l) 21 71% Repetitive(le) 38 64% Repetitive(les) 33 78% Word(leplorka) 114 71% Word(porovnal) 45 71% Word(nepopadnoul) 39 71% Sentence 129 72% Overall Accuracy 79% Table 4.3: GA Offline features results 4.3.3 Results using Correlation technique(Online and Of- fline) In this experiment, we build a correlation matrix, which exam- ines the correlation of all features (for all possible feature combina- tions). We selected highly correlated features by setting a threshold of 0.5, It removed the first feature that is correlated with anything else without any other insight. We removed these highly correlated features and selected the features which are not highly correlated columns and have an absolute correlation smaller than 0.5. We fed these selected features data to a machine learning classifier i.e. SVM for PD classification. The obtained result after Correlation feature selection techniques are mentioned in the below table 4.4,4.5 Tasks No of Features Selected SVM Archimedean Spiral 14 64% Repetitive(l) 17 61% Repetitive(le) 21 70% Repetitive(les) 17 61% Word(leplorka) 16 61% Word(porovnal) 20 70% Word(nepopadnoul) 19 53% Sentence 20 78% Overall Accuracy 64.75% Table 4.4: Correlation Online Feature Results Tasks No of Features Selected SVM Archimedean Spiral 2172 79% Repetitive(l) 1874 86% Repetitive(le) 1187 87% Repetitive(les) 1862 71% Word(leplorka) 1725 89% Word(porovnal) 1931 87% Word(nepopadnoul) 1801 71% Sentence 1759 64% Overall Accuracy 79% Table 4.5: Correlation Offline Features Results 4.3.4 Combined features result analysis In this experiment, firstly we employed task-wise classification by combining all online and offline features. We also combined both online and offline features after feature selection and performed task-wise classification. and the obtained results are mentioned in the table below 4.6 Task All Features Accuracy Selected Features Accuracy (GA) Selected Features Accuracy(Correlation) Archimedean Spiral 59% 81.67% 72% Repetitive(l) 52% 78.32% 74% Repetitive(le) 61% 74.32% 78% Repetitive(les) 62% 75.24% 66% Word(leplorka) 57% 76.35% 75% Word(porovnal) 54% 74.11% 78% Word(nepopadnoul) 55% 76.08% 63% Sentence 66% 83.57% 71% Table 4.6: Combined features Result Analysis Figure 4.1: Comparison between all and selected features accuracy on online features Figure 4.2: Comparison between all and selected features accuracy on offline features Figure 4.3: Comparison between all and selected features accuracy on combined features To further, establish the viability of the proposed model we compare with the studies on task-based selection in table 4.7. All of the studies have utilized the PaHaW dataset for the assessment of the proposed method. The authors in [28, 10] have employed sev- eral static and dynamic features extraction techniques to predict PD using PaHaW dataset. We use combined features of handwrit- ing to demonstrate their use in identifying the presence or absence of Parkinson's disease.The accuracies reported in different experi- ments are comparable to those reported in the literature. Task Impedevo et al [8] Angelilo et al [48] Diaz et al [38] Momina et al [49] Proposed Technique Archimedean Spiral 54.67% 53.75% 75.00% 89.64% 81.67% Repetitive(l) 61.80% 67.08% 64.16% 75.00% 78.32% Repetitive(le) 72.28% 62.50% 58.33% 73.75% 74.32% Repetitive(les) 55.28% 57.91% 71.67% 72.32% 75.24% Word(leplorka) 59.80% 54.58% 75.41% 79.46% 76.35% Word(porovnal) 63.71% 56.75% 63.75% 74.46% 74.11% Word(nepopadnoul) 60.98% 61.67% 70.00% 79.28% 76.08% Sentence 71.95% 70.40% 67.08% 81.42% 83.57% We discussed our result according to the most effective fea- tures set on the specific task in PaHaW dataset. We observed features extracted from GA give a better classification performance (77.46% global accuracy) using an SVM classifier than features ex- tracted from correlation selection techniques. Nevertheless, consid- ering the very limited amount of text (drawing) available in each task, the realized accuracies are indeed promising. Comparing the performance of different tasks, it can be observed that after feature selection Sentence task has the highest accuracy in overall ex- periments. The Archimedean spiral tasks achieved second highest accuracy in overall experiments. The word-based tasks "porovonal, nepopadnoul,leplorka," task provides remarkably similar results. On the other hand,repetitive letter tasks "l" and "les" obtained the highest accuracy "le" have the lowest accuracy. "le" tasks were comparatively obtained less effectively than other tasks in our pro- posed system. 4.4 Summary This chapter presents the details of all experiments carried out to prove our thoughts to bring up in this research. We used feature relevance methodologies to evaluate the performance of features on- line and offline individually, then integrated both types of features and applied SVM for classification. Task-wise accuracy was used in these investigations, and the system was compared against others in the literature. CHAPTER 5 CONCLUSION & FUTURE WORK The potential for handwriting features to indicate Parkinson's disease is discussed in this study. The literature has looked at both online and offline features, however, in our study, we used a combi- nation of online and offline features and extracted a set of features that performed better on a specific task template This study does not deny previous research on online and offline features; rather, it enhances the author's expertise and demonstrated the utility of both online and offline features. we explore the feature relevance techniques to detect offline and online features. Evaluation on a standard data set (PAHAW), our proposed system reports overall accuracy of 77.46% when we combined both features. Another an important aspect of our studies is that we explained the importance of features for specific templates In conclusion, we have demonstrated that such a method could be beneficial to clinicians in the diagnosis of Parkinson's dis- ease since it allows them to select the most important features of the disease and, as a result, develop a set of guidelines for defining fresh testing protocols. This work will be used to guide future work and research in this field because of the performance metric across numerous datasets. REFERENCES [1] C. Tucker, Y. Han, H. Black Nembhard, W.-C. Lee, M. Lewis, N. Sterling, and X. Huang, "A data mining methodology for predicting early stage parkinson's disease using non-invasive, high-dimensional gait sensor data," IIE transactions on health- care systems engineering, vol. 5, no. 4, pp. 238–254, 2015. [2] D. Hirtz, D. Thurman, K. Gwinn-Hardy, M. Mohamed, A. Chaudhuri, and R. Zalutsky, "How common are the "com- mon" neurologic disorders?," Neurology, vol. 68, no. 5, pp. 326– 337, 2007. [3] O.-B. Tysnes and A. Storstein, "Epidemiology of parkinson's disease," Journal of Neural Transmission, vol. 124, no. 8, pp. 901–905, 2017. [4] S. von Campenhausen, B. Bornschein, R. Wick, K. Bötzel, C. Sampaio, W. Poewe, W. Oertel, U. Siebert, K. Berger, and R. Dodel, "Prevalence and incidence of parkinson's disease in europe," European Neuropsychopharmacology, vol. 15, no. 4, pp. 473–490, 2005. [5] L. M. De Lau and M. M. Breteler, "Epidemiology of parkin- son's disease," The Lancet Neurology, vol. 5, no. 6, pp. 525–535, 2006. [6] Y. Zou, J. Tan, N. Li, J. Yang, B. Yu, J. Yu, W. Du, W. Zhang, L. Cui, Q. Wang, et al., "The prevalence of parkinson's disease continues to rise after 80 years of age: a cross-sectional study of chinese veterans," Eur Rev Med Pharmacol Sci, vol. 18, no. 24, pp. 3908–3915, 2014. [7] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Decision support framework for parkinson's disease based on novel handwriting markers," IEEE Transactions on Neural Systems and Rehabilitation En- gineering, vol. 23, no. 3, pp. 508–516, 2014. [8] D. Impedovo, G. Pirlo, and G. Vessio, "Dynamic handwriting analysis for supporting earlier parkinson's disease diagnosis," Information, vol. 9, no. 10, p. 247, 2018. [9] M. Alviano, G. Greco, and F. Scarcello, AI* IA 2019–Advances in Artificial Intelligence: XVIIIth International Conference of the Italian Association for Artificial Intelligence, Rende, Italy, November 19–22, 2019, Proceedings, vol. 11946. Springer Na- ture, 2019. [10] M. Moetesum, I. Siddiqi, N. Vincent, and F. Cloppet, "Assess- ing visual attributes of handwriting for prediction of neurolog- ical disorders—a case study on parkinson's disease," Pattern Recognition Letters, vol. 121, pp. 19–27, 2019. [11] A. M. Vlaar, A. E. Bouwmans, M. J. Van Kroonenburgh, W. H. Mess, S. C. Tromp, P. G. Wuisman, A. G. Kessels, A. Winogrodzka, and W. E. Weber, "Protocol of a prospective study on the diagnostic value of transcranial duplex scanning of the substantia nigra in patients with parkinsonian symp- toms," BMC neurology, vol. 7, no. 1, pp. 1–6, 2007. [12] A. Salarian, H. Russmann, C. Wider, P. R. Burkhard, F. J. Vingerhoets, and K. Aminian, "Quantification of tremor and bradykinesia in parkinson's disease using a novel ambulatory monitoring system," IEEE Transactions on Biomedical Engi- neering, vol. 54, no. 2, pp. 313–322, 2007. [13] U. Tisch, I. Schlesinger, R. Ionescu, M. Nassar, N. Axelrod, D. Robertman, Y. Tessler, F. Azar, A. Marmur, J. Aharon- Peretz, et al., "Detection of alzheimer's and parkinson's dis- ease from exhaled breath using nanomaterial-based sensors," Nanomedicine, vol. 8, no. 1, pp. 43–56, 2013. [14] J. Mekyska, Z. Smekal, M. Kostalova, M. Mrackova, S. Sku- tilova, and I. Rektorova, "Motor aspects of speech imparment in parkinson's disease and their assessment," Ceska A Sloven- ska Neurologie A Neurochirurgie, vol. 74, no. 6, pp. 662–668, 2011. [15] A. Tsanas, M. A. Little, C. Fox, and L. O. Ramig, "Objec- tive automatic assessment of rehabilitative speech treatment in parkinson's disease," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 22, no. 1, pp. 181–190, 2013. [16] J. Rusz, R. Čmejla, H. Ržičková, J. Klempíř, V. Majerová, J. Picmausová, J. Roth, and E. Ržička, "Acoustic assess- ment of voice and speech disorders in parkinson's disease through quick vocal test," Movement Disorders, vol. 26, no. 10, pp. 1951–1952, 2011. [17] A. Naseer, M. Rani, S. Naz, M. I. Razzak, M. Imran, and G. Xu, "Refining parkinson's neurological disorder identifica- tion through deep transfer learning," Neural Computing and Applications, vol. 32, no. 3, pp. 839–854, 2020. [18] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, "Novel speech signal processing algorithms for high- accuracy classification of parkinson's disease," IEEE transac- tions on biomedical engineering, vol. 59, no. 5, pp. 1264–1271, 2012. [19] S. Rosenblum, M. Samuel, S. Zlotnik, I. Erikh, and I. Schlesinger, "Handwriting as an objective tool for parkin- son's disease diagnosis," Journal of neurology, vol. 260, no. 9, pp. 2357–2361, 2013. [20] J. Mucha, J. Mekyska, Z. Galaz, M. Faundez-Zanuy, K. Lopez- de Ipina, V. Zvoncak, T. Kiska, Z. Smekal, L. Brabenec, and I. Rektorova, "Identification and monitoring of parkinson's dis- ease dysgraphia based on fractional-order derivatives of online handwriting," Applied Sciences, vol. 8, no. 12, p. 2566, 2018. [21] M. Moetesum, I. Siddiqi, F. Javed, and U. Masroor, "Dynamic handwriting analysis for parkinson's disease identification us- ing c-bigru model," in 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 115–120, 2020. [22] J. H. Man, L. Groenink, and M. Caiazzo, "Cell reprogramming approaches in gene-and cell-based therapies for parkinson's dis- ease," Journal of controlled release, vol. 286, pp. 114–124, 2018. [23] M. Isenkul, B. Sakar, and O. Kursun, "Improved spiral test using digitized graphics tablet for monitoring parkinson's dis- ease," 05 2014. [24] C. R. Pereira, D. R. Pereira, F. A. Silva, J. P. Masieiro, S. A. T. Weber, C. Hook, and J. P. Papa, "A new computer vision- based approach to aid the diagnosis of parkinson's disease," Computer Methods and Programs in Biomedicine, vol. 136, pp. 79–88, 2016. [25] L. Palmerini, L. Rocchi, S. Mellone, F. Valzania, and L. Chiari, "Feature selection for accelerometer-based posture analysis in parkinson's disease," IEEE Transactions on Information Tech- nology in Biomedicine, vol. 15, no. 3, pp. 481–490, 2011. [26] K. Niazmand, K. Tonn, A. Kalaras, U. M. Fietzek, J.- H. Mehrkens, and T. C. Lueth, "Quantitative evaluation of parkinson's disease using sensor based smart glove," in 2011 24th International Symposium on Computer-Based Medical Systems (CBMS), pp. 1–8, IEEE, 2011. [27] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Analysis of in-air movement in hand- writing: A novel marker for parkinson's disease," Computer methods and programs in biomedicine, vol. 117, no. 3, pp. 405– 411, 2014. [28] P. Drotár, J. Mekyska, Z. Smékal, I. Rektorová, L. Masarová, and M. Faundez-Zanuy, "Contribution of different handwriting modalities to differential diagnosis of parkinson's disease," in 2015 IEEE international symposium on medical measurements and applications (MeMeA) proceedings, pp. 344–348, IEEE, 2015. [29] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Evaluation of handwriting kinematics and pressure for differential diagnosis of parkinson's disease," Artificial intelligence in Medicine, vol. 67, pp. 39–46, 2016. [30] C. R. Pereira, S. A. Weber, C. Hook, G. H. Rosa, and J. P. Papa, "Deep learning-aided parkinson's disease diagnosis from handwritten dynamics," in 2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 340–346, Ieee, 2016. [31] R. Graça, R. S. e Castro, and J. Cevada, "Parkdetect: Early diagnosing parkinson's disease," in 2014 IEEE Interna- tional Symposium on Medical Measurements and Applications (MeMeA), pp. 1–6, IEEE, 2014. [32] M. Angelillo, D. Impedovo, G. Pirlo, and G. Vessio, Performance-Driven Handwriting Task Selection for Parkin- son's Disease Classification, pp. 281–293. 11 2019. [33] R. Senatore, A. Della Cioppa, and A. Marcelli, "Automatic diagnosis of neurodegenerative diseases: An evolutionary ap- proach for facing the interpretability problem," Information, vol. 10, no. 1, p. 30, 2019. [34] A. Ammour, I. Aouraghe, G. Khaissidi, M. Mrabti, G. Aboulem, and F. Belahsen, "A new semi-supervised ap- proach for characterizing the arabic on-line handwriting of parkinson's disease patients," Computer methods and programs in biomedicine, vol. 183, p. 104979, 2020. [35] C. Loconsole, G. F. Trotta, A. Brunetti, J. Trotta, A. Schi- avone, S. I. Tatò, G. Losavio, and V. Bevilacqua, "Computer vision and emg-based handwriting analysis for classification in parkinson's disease," in International Conference on Intelli- gent Computing, pp. 493–503, Springer, 2017. [36] P. Khatamino, İ. Cantürk, and L. Özyılmaz, "A deep learning- cnn based system for medical diagnosis: An application on parkinson's disease handwriting drawings," in 2018 6th In- ternational Conference on Control Engineering & Information Technology (CEIT), pp. 1–6, IEEE, 2018. [37] B. E. Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgen, S. Delil, H. Apaydin, and O. Kursun, "Collection and analy- sis of a parkinson speech dataset with multiple types of sound recordings," IEEE Journal of Biomedical and Health Informat- ics, vol. 17, no. 4, pp. 828–834, 2013. [38] M. Diaz, M. A. Ferrer, D. Impedovo, G. Pirlo, and G. Ves- sio, "Dynamically enhanced static handwriting representation for parkinson's disease detection," Pattern Recognition Letters, vol. 128, pp. 204–210, 2019. [39] L. C. Ribeiro, L. C. Afonso, and J. P. Papa, "Bag of samplings for computer-assisted parkinson's disease diagnosis based on recurrent neural networks," Computers in biology and medicine, vol. 115, p. 103477, 2019. [40] J. D. Gupta and B. Chanda, "Novel features for diagnosis of parkinson's disease from off-line archimedean spiral images," in 2019 IEEE 10th International Conference on Awareness Sci- ence and Technology (iCAST), pp. 1–6, IEEE, 2019. [41] C. R. Pereira, S. A. T. Weber, C. Hook, G. H. Rosa, and p. . . y. . . J. P. Papa title = Deep Learning-aided Parkinson's Disease Diagnosis from Handwritten Dynamics, booktitle = Proceed- ings of the SIBGRAPI 2016 - Conference on Graphics, Patterns and Images [42] A. Lozano and A. Lang, "Pallidotomy for parkinson's disease," Advances in neurology, vol. 86, pp. 413–20, 02 2001. [43] V. M. Jerkovic, V. Kojic, N. D. Miskovic, T. Djukic, V. S. Kostic, and M. B. Popovic, "Analysis of on-surface and in- air movement in handwriting of subjects with parkinson's disease and atypical parkinsonism,” Biomedical Engineer- ing/Biomedizinische Technik, vol. 64, no. 2, pp. 187–194, 2019. [44] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Feature extraction: foundations and applications, vol. 207. Springer, 2008. [45] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” Journal of machine learning research, vol. 3, no. Mar, pp. 1157–1182, 2003. [46] I. Siddiqi, K. Khurshid, and N. Vincent, “Feature relevance analysis for writer identification,” in Document Recognition and Retrieval XVIII, vol. 7874, p. 78740F, International So- ciety for Optics and Photonics, 2011. [47] M. Maliha, A. Tareque, and S. S. Roy, Diabetic retinopathy de- tection using machine learning. PhD thesis, BRAC University, 2018. [48] M. T. Angelillo, D. Impedovo, G. Pirlo, and G. Vessio, “Performance-driven handwriting task selection for parkin- son’s disease classification,” in International Conference of the Italian Association for Artificial Intelligence, pp. 281–293, Springer, 2019. [49] M. Moetesum, I. Siddiqi, F. Javed, and U. Masroor, “Dynamic handwriting analysis for parkinson’s disease identification us- ing c-bigru model,” in 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 115–120, IEEE, 2020. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52