RV College of Engineering®
Mysore Road, RV Vidyaniketan Post,Bengaluru - 560059, Karnataka, India

A MACHINE LEARNING-BASED DIAGNOSIS TOOL FOR PARKINSON'S DISEASE

MAJOR PROJECT REPORT

submitted by

Abhay Shashidhara	**1RV21BT001**
Manish Danda	**1RV21BT027**
Samridhi Makkar	**1RV21BT044**
Shreya Shanbhog	**1RV21BT048**

under the guidance of

Dr. A H Manjunatha Reddy
Associate Professor,
Department of Biotechnology
RV College of Engineering
(Affiliated to VTU-Belgavi)

Dr. Sumathra M
Assistant Professor,
Department of Biotechnology
RV College of Engineering
(Affiliated to VTU-Belgavi)

in partial fulfilment for the award of degree of

Bachelor of Engineering
in
Department of Biotechnology
2024-2025

Department of Biotechnology

A MACHINE LEARNING-BASED DIAGNOSIS TOOL FOR PARKINSON'S DISEASE

MAJOR PROJECT REPORT

submitted by

Abhay Shashidhara **1RV21BT001**

Manish Danda **1RV21BT027**

Samridhi Makkar **1RV21BT044**

Shreya Shanbhog **1RV21BT048**

under the guidance of

Dr. A H Manjunatha Reddy
Associate Professor,
Department of Biotechnology
RV College of Engineering
(Affiliated to VTU-Belgavi)

Dr. Sumathra M
Assistant Professor,
Department of Biotechnology
RV College of Engineering
(Affiliated to VTU-Belgavi)

in partial fulfilment for the award of degree of

Bachelor of Engineering

in

Department of Biotechnology

2024-2025

RV College of Engineering®

Mysore Road, RV Vidyaniketan Post,Bengaluru - 560059, Karnataka, India

CERTIFICATE

Certified that the Major project titled '**A Machine Learning-Based Diagnosis Tool For Parkinson's Disease'** is carried out by **Abhay Shashidhara (1RV21BT001), Manish Danda (1RV21BT027), Samridhi Makkar (1RV21BT044) and Shreya Shanbhog (1RV21BT048)** who are bonafide students of RV College of Engineering, Bengaluru, in partial fulfillment for the award of Degree of **Bachelor of Engineering** in **Biotechnology Engineering** of the Visvesvaraya Technological University, Belagavi during the year **2024-2025**. It is certified that all corrections/suggestions indicated for the internal assessment have been incorporated in the report deposited in the department library. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed by the institution for the said Degree.

Guide 1 **Guide 2** **Head of Department** **Principal**

External Viva Examination

Name of Examiner **Signature with Date**

1.

2.

RV College of Engineering®

Mysore Road, RV Vidyaniketan Post,Bengaluru - 560059, Karnataka, India

DECLARATION

We **Abhay Shashidhara, Manish Danda, Samridhi Makkar and Shreya Shanbhog**, the students of the eighth semester B.E., Department of **Biotechnology Engineering**, RV College of Engineering, Bengaluru-560059, bearing USN: **1RV21BT001, Manish Danda, Samridhi Makkar and Shreya Shanbhog** hereby declare that the project titled **'A Machine Learning-Based Diagnosis Tool For Parkinson's Disease'** has been carried out by us and submitted in partial fulfilment of the program requirements for the award of Degree in Bachelor of Engineering in **Biotechnology Engineering** of the **Visvesvaraya Technological University, Belagavi** during the year **2024-2025**.

Further, we declare that the content of the dissertation has not been submitted previously by anybody for the award of any Degree or Diploma to any other University.

We also declare that any Intellectual property rights generated out of this project carried out at RVCE will be the property of RV College of Engineering, Bengaluru and we will be among the authors of the same.

Place: Bangalore

Date:

	Signature
Abhay Shashidhara	_____
Manish Danda	_____
Samridhi Makkar	_____
Shreya Shanbhog	_____

Acknowledgement

We would like to express our sincere gratitude to our guides, Dr. Sumathra M and Dr. A. H. Manjunatha Reddy, for their invaluable guidance, support and encouragement throughout this project. Their insights and timely feedback were instrumental in shaping the direction and outcome of our work.

We are also deeply thankful to Dr. Neetha, Major Project Coordinator, for her continuous support and coordination, which greatly facilitated the smooth progress of the project.

We extend our heartfelt thanks to Dr. Nagashree N. Rao, Head of the Department, for providing us with the necessary infrastructure and academic environment to pursue our work.

We would also like to thank all the faculty members of the department, especially Dr. H.G. Ashok, for their constructive feedback and support. A special mention of appreciation to our external examiner, Dr. Krishna, for his valuable evaluation and suggestions during the final review.

Lastly, we thank all those who, directly or indirectly, contributed to the successful completion of this project.

Abstract

Parkinson's Disease (PD) is an increasingly prevalent neurodegenerative disorder that poses a growing public health challenge, contributing to millions of deaths globally each year. Despite the scale of the issue, early diagnosis remains a significant barrier, largely due to a lack of public awareness, the high financial burden of clinical consultations and restricted access to specialized neurological care—especially in underserved and rural regions. In response to these challenges, our objective was to design an accessible, open-source diagnostic platform to empower patients to monitor their neurological health from home.

The developed system is a hybrid diagnostic tool composed of two components: a large language model (LLM) and a convolutional neural network (CNN). The LLM processes user inputs—including lifestyle patterns, genetic predisposition, past medical records, and motor and non-motor symptoms—to produce a structured, patient-understandable report. In parallel, the CNN analyzes spiral test images, a clinically recognized motor assessment, to classify PD-specific motor dysfunction. The system also integrates access to recent literature and early symptom checklists to enhance user engagement and awareness.

The CNN model was trained and evaluated on a curated dataset of spiral drawings. On the held-out test set, it achieved a classification accuracy of 96.25%, with strong performance in identifying PD-related motor impairments. These results highlight the effectiveness of combining visual diagnostic tools with structured metadata analysis for improved detection, especially in early-stage cases where clinical observation alone may be insufficient.

The primary goal of this tool is to make PD screening more accessible, user-friendly, and informative, particularly for individuals without regular access to clinical resources. By focusing on interpretability and simplicity, the system empowers non-clinical users to better understand their neurological health. While the current version is designed with patient usability at its core, future upgrades will aim to extend its capabilities by integrating clinician-specific modules and more comprehensive diagnostic markers. This will allow the tool to evolve into a dual-use platform, supporting both personal monitoring and medical decision-making in formal healthcare environments.

List of Tables

List of Figures

Table of Contents

Chapter 1: Introduction

This chapter establishes the pressing need for early Parkinson's Disease detection and introduced a dual-model framework leveraging CNNs and LLMs. The platform addresses current diagnostic limitations through scalable, AI-powered assessments, setting the stage for deeper exploration into the underlying technologies and implementation.

1.1 Overview of Parkinson's Disease

Parkinson's Disease (PD) is a complex and progressive neurodegenerative disorder that primarily affects the motor system but also encompasses a wide array of non-motor symptoms. At its core, PD is characterized by hallmark motor features such as bradykinesia (slowness of movement), muscular rigidity, resting tremors, and postural instability. However, the disease's impact goes far beyond movement disorders. Many patients also experience non-motor symptoms that can be just as debilitating, including cognitive decline, mood disorders like depression and anxiety, sleep disturbances, and a host of autonomic dysfunctions such as constipation, sexual dysfunction, and orthostatic hypotension [1], [2]. The biological basis of PD lies in the progressive degeneration of dopaminergic neurons in the substantia nigra pars compacta—a key area in the midbrain responsible for movement regulation. As these neurons die, the brain's dopamine levels plummet, resulting in the motor impairments typically observed. Additionally, abnormal protein aggregates known as Lewy bodies, composed primarily of misfolded α-synuclein, accumulate in neurons and contribute to disrupted synaptic communication and cellular homeostasis [3]. The multifaceted nature of PD and its slow, insidious onset make it one of the most challenging neurological diseases to diagnose and manage effectively.

1.2 Global and National Burden

The burden of Parkinson's Disease is escalating at a global level, and it is now recognized as the fastest-growing neurological disorder in terms of prevalence and disability-adjusted life years (DALYs) [4]. While aging remains a dominant risk factor, environmental triggers, genetic predispositions, and changing lifestyles are also contributing to the

disease's increasing incidence. This global trend is mirrored—and in some ways magnified—in countries like India, where the dual challenge of rapid population aging and healthcare disparities intensifies the crisis. India faces unique vulnerabilities due to a lack of awareness, inadequate access to neurologists in rural regions, and cultural barriers that delay medical consultation for early symptoms [1], [5]. Moreover, epidemiological studies in India remain sparse, and this underreporting contributes to an underestimation of the true scale of PD in the country. As a result, both public health planning and policy development are severely hampered, further entrenching diagnostic delays and suboptimal care. Given the projected rise in the elderly population, it becomes imperative to implement diagnostic solutions that are not only accurate but also affordable and accessible to India's vast and diverse population.

1.3 Diagnostic Challenges

One of the most pressing issues in Parkinson's Disease management is the difficulty of achieving a reliable and early diagnosis. Currently, PD is diagnosed based on clinical criteria, primarily focusing on visible motor symptoms like tremors or bradykinesia. However, by the time these symptoms are evident, it is estimated that up to 60–70% of the dopaminergic neurons in the substantia nigra may already be lost [6]. This irreversible neuronal damage underscores the need for early detection strategies that can intervene before the disease significantly progresses. Further complicating the diagnosis is the absence of a definitive biomarker for early-stage PD. Unlike diseases such as diabetes or cancer, where blood tests or imaging can provide strong diagnostic signals, PD remains heavily reliant on subjective clinical observation and patient history. Additionally, the clinical presentation of PD is highly variable across individuals—some may exhibit primarily tremor-dominant forms, while others may show axial rigidity or early cognitive symptoms. This heterogeneity makes it difficult for clinicians to adopt a uniform diagnostic strategy [7]. Together, these factors create a diagnostic landscape marked by delays, uncertainty, and high dependence on specialist consultation.

1.4 Limitations of Current Treatment Approaches

Although pharmacological therapies like levodopa, dopamine agonists, and MAO-B inhibitors have transformed the management of PD symptoms, they remain fundamentally palliative. Levodopa, the gold-standard medication, helps replenish dopamine in the brain and can offer substantial motor symptom relief, especially in the early years of treatment [2]. However, it does not halt or reverse the underlying neurodegeneration. Over time, patients often require increasing doses to maintain efficacy, leading to complications such as motor fluctuations, "wearing off" effects, and levodopa-induced dyskinesias—abnormal, involuntary movements that can severely impact quality of life [6]. Furthermore, these medications have little to no impact on non-motor symptoms, which are often equally distressing to patients. Advanced treatment options like deep brain stimulation (DBS) are available but are invasive, expensive, and not feasible for most patients in low-resource settings. The limitations of existing treatment modalities thus reinforce the importance of early detection and personalized disease monitoring, which can pave the way for timely interventions and potentially slow disease progression through lifestyle or emerging neuroprotective strategies.

1.5 Role of Artificial Intelligence in PD Detection

With the convergence of healthcare and digital technology, artificial intelligence (AI) has emerged as a transformative tool in the realm of medical diagnostics. In the context of PD, AI-driven methods—especially those based on machine learning (ML)—have shown exceptional promise in enhancing diagnostic sensitivity and specificity. Among these, convolutional neural networks (CNNs) have become particularly popular for their ability to process visual and spatial data. One compelling use case is the analysis of spiral and wave drawing tests, which are simple motor tasks that capture hand motion abnormalities such as tremors and bradykinesia [8]. These tests, often performed with a stylus or pen on paper or touchscreen, produce visual outputs that reflect fine motor control. CNNs can be trained to detect subtle irregularities in these patterns, providing an automated, non-invasive, and objective method of screening for PD-related motor symptoms. This approach is not only cost-effective but also highly scalable, making it well-suited for use in both clinical and remote settings.

1.6 Advancements in Language-Based Medical Interpretation

Beyond motor diagnostics, the interpretation of patient history, symptom patterns, and lifestyle data is crucial for a holistic understanding of PD. Natural language processing (NLP), particularly through large language models (LLMs), is revolutionizing this aspect of healthcare. These models are capable of analyzing unstructured and structured text-based inputs—ranging from clinical notes and diagnostic reports to self-reported symptom logs and wearable data—to generate coherent, medically relevant summaries [9]. Using architectures like transformers and frameworks such as retrieval-augmented generation (RAG), LLMs can pull context from both patient-specific data and external literature to produce detailed, transparent reasoning chains. This is particularly valuable in the PD domain, where symptom variability requires nuanced interpretation. LLMs can offer personalized diagnostic hypotheses or care suggestions, adapting their outputs based on the patient's age, symptom trajectory, comorbidities, and even genetic predispositions. When implemented correctly, such systems can serve as intelligent companions for clinicians or as standalone tools in telemedicine and home-care scenarios.

1.7 Present Study and Platform Overview

Building on these technological innovations, the present study introduces a hybrid, open-source AI platform specifically designed for the early diagnosis of Parkinson's Disease. The platform integrates a CNN module for analyzing spiral test images and an LLM module for synthesizing diverse patient data into diagnostic insights. By leveraging both visual and textual inputs, the system mirrors the dual diagnostic process employed by clinicians—observation of motor symptoms and interpretation of patient history. The CNN stream automatically processes image data to detect motor abnormalities, while the LLM stream interprets structured information such as age, diet, genetic history, and lifestyle habits, and augments its reasoning with real-time academic literature using RAG techniques. The system is optimized for both patient-facing use and clinician-led research, offering detailed explanations and relevant citations for each diagnostic outcome.

As shown in **Figure 1**, the interface unifies these components into a seamless pipeline that delivers tailored outputs to users in an understandable format. While this version is primarily built for at-home use to enhance accessibility and awareness, future iterations will incorporate clinical scales, voice inputs, and temporal tracking to expand its functionality in professional settings. Ultimately, this hybrid diagnostic tool aspires to democratize PD care by bridging the gap between state-of-the-art AI and real-world healthcare needs.

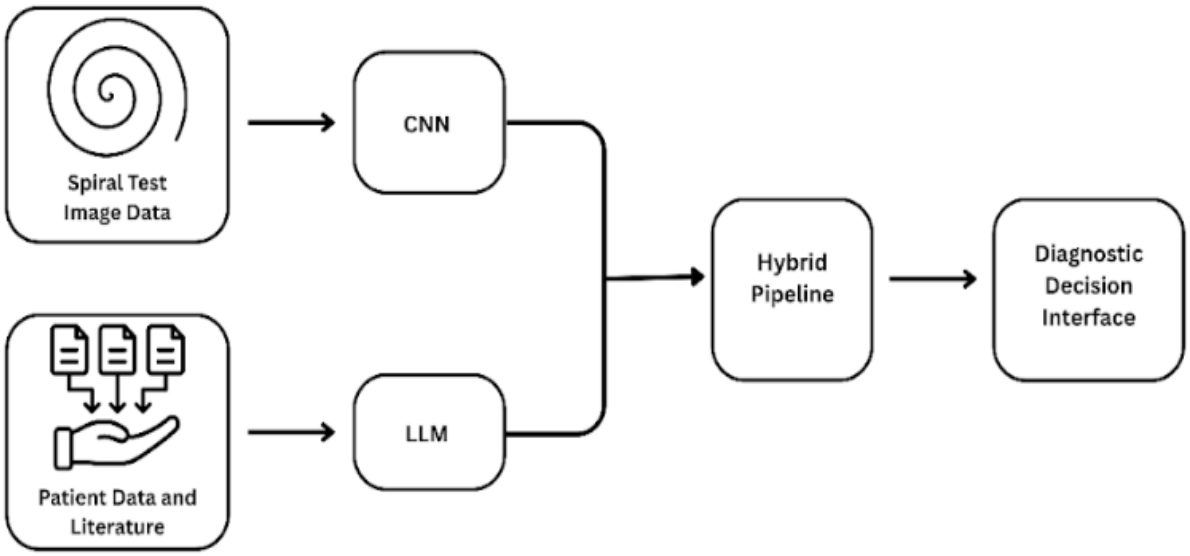

Fig 1. Product Architecture: This figure shows the hybrid pipeline combining CNN-based image analysis and LLM-driven data interpretation to generate diagnostic insights via a unified decision interface.

Chapter 2 :Theory and Fundamentals

The theoretical background outlined in this chapter provides a foundation for understanding both the medical and computational components of the diagnostic system. Concepts related to motor symptom analysis, deep learning, and natural language processing form the essential scaffolding upon which the hybrid model is built.

2.1 Understanding Parkinson's Disease in the 21st Century

Parkinson's Disease (PD) remains one of the most persistent challenges in modern neurology, not because it is rare or unknown, but precisely because it is so widespread and yet so difficult to detect early with precision. As a progressive neurodegenerative condition, PD silently affects millions of individuals globally, with a rising burden expected in aging populations like those of India and several other low-to-middle-income countries. The disease itself is marked by the gradual death of dopaminergic neurons in a specific region of the brain called the substantia nigra [13]. These neurons are critical for facilitating smooth and coordinated body movements through the release of dopamine, a neurotransmitter. As these cells die, dopamine levels decline, leading to a breakdown in motor function that manifests through clinical symptoms such as bradykinesia, rigidity, tremors, postural instability, and eventually, gait disturbances. However, the most troubling aspect of PD is that these motor symptoms typically appear only after a substantial percentage of neurons—sometimes more than 70%—have already been lost. That means that by the time a patient is clinically diagnosed using traditional methods, it may already be too late for preventative or neuroprotective strategies to make a meaningful impact [14]. This fundamental delay in diagnosis has haunted clinicians and researchers for decades and underscores why the development of intelligent, scalable, and non-invasive diagnostic tools is no longer just a research goal, it is a public health necessity [15].

Complications of early identification of PD is the presence of a large constellation of non-motor symptoms that often precede the classic motor ones by several years [16]. Patients may experience sleep disturbances, mood changes, constipation, reduced sense of smell, or subtle cognitive shifts, all of which can seem innocuous or be mistaken for signs of aging or unrelated conditions. These non-motor signs are rarely considered sufficient for

diagnosis on their own due to their lack of specificity, yet they form a crucial part of the disease's trajectory [17]. This is precisely why early screening methods must begin to think beyond observable tremors and delve deeper into comprehensive, patient-specific data. In addition to the symptom-related challenges, diagnostic methods for PD are also constrained by systemic issues such as lack of access to specialized neurological care, particularly in rural or resource-scarce environments. Many people in such regions remain undiagnosed or misdiagnosed, with their symptoms attributed to other geriatric conditions [18]. Even when access is available, diagnosis still heavily depends on a neurologist's subjective interpretation of the patient's history and observable motor symptoms, often involving a levodopa challenge test. Though effective to a certain degree, this reliance on clinical acumen introduces variability and sometimes delays diagnosis until functional impairment is well underway. This inefficiency in the current diagnostic framework demands not just refinement but reimagination [19].

The promise of artificial intelligence, and more specifically machine learning, lies in its ability to find patterns in data that are invisible to even the most trained clinical eye. With the explosion of healthcare datasets—from electronic health records and diagnostic images to patient-reported outcomes—there exists a new opportunity to synthesize diverse types of patient data into meaningful insights [20]. AI-driven tools are not meant to replace neurologists but to empower them with support systems that can catch signals earlier, improve decision-making, and reduce diagnostic error. In the case of PD, AI can process both structured data (like age, family history, medication use) and unstructured data (like spiral drawings, voice patterns, or handwritten samples), enabling a truly holistic evaluation. This paradigm shift away from symptom-checklists and toward integrated, data-informed assessment is not just technologically impressive, it is ethically necessary. Patients deserve earlier answers, especially for a disease as impactful as Parkinson's. And clinicians deserve tools that match the complexity of the disease they are trying to understand [21]. The intersection of neuroscience and AI is not a futuristic dream. It is happening now, and it is changing the way we look at diagnosis, starting with tools like the one described in this thesis.

2.2 The Evolution of Parkinson's Diagnostics: From Clinical Judgement to Intelligent Systems

The diagnostic process for Parkinson's Disease has long relied on a framework rooted in clinical observation, trial-and-error medication responses, and patient self-reporting. For decades, neurologists have played the central role in assessing PD, often depending on physical signs such as tremor, rigidity, and bradykinesia, along with the patient's ability to respond to dopamine replacement therapy like levodopa. While this approach has proven to be largely effective in moderate to advanced stages of the disease, it falls drastically short in the early stages where symptoms may be subtle, sporadic, or even entirely absent from the motor spectrum [22]. A critical bottleneck in the conventional diagnosis arises from its subjectivity; two patients with nearly identical biological deterioration may present with entirely different symptom sets, and two clinicians may interpret the same symptoms differently depending on their experience and expertise [23]. This inter-observer variability, compounded by the absence of standard biological markers, delays diagnosis for many patients—sometimes for years. Such delays are not benign; they actively deprive patients of the opportunity for early therapeutic interventions, lifestyle adaptations, and mental preparedness, all of which are known to positively influence disease progression and quality of life [24]. In this context, the need to shift from purely subjective clinical judgment to more consistent, data-driven diagnostic tools become both scientifically and ethically pressing [25].

In response to these shortcomings, the medical community has increasingly turned toward computational intelligence to augment the diagnostic process. The last decade has witnessed a remarkable shift in focus from symptom-based classification to pattern recognition using machine learning. Machine learning (ML), as a branch of artificial intelligence (AI), empowers systems to learn from past data without being explicitly programmed for every rule or decision point [26]. In the context of PD, ML models can be trained on large volumes of historical patient data to identify associations between subtle symptoms, patient demographics, and eventual diagnoses. This is especially powerful when the data includes both motor parameters (like tremor frequency or spiral irregularity) and non-motor indicators (such as sleep disturbance or constipation). By learning from

such multi-dimensional datasets, the models are able to detect early warning signals that even experienced clinicians might miss [27]. Moreover, once trained, these models offer a level of consistency that human evaluators cannot always provide. Every patient is assessed using the same internal logic and statistical rigor, reducing the variability introduced by human subjectivity. Importantly, this doesn't sideline the clinician, it enhances their toolkit, offering a second opinion that is grounded in data patterns and historical precedent [28, 29].

As ML evolved, so did the sophistication of the tools available for Parkinson's diagnostics. Basic supervised classification algorithms like decision trees and support vector machines paved the way for more advanced architectures like neural networks, convolutional neural networks (CNNs), and transformer-based models [30]. CNNs in particular have proven to be invaluable for interpreting visual data, such as hand-drawn spirals, where tremors and motor instability leave subtle visual signatures that can be quantified algorithmically. On the other hand, natural language processing (NLP) models such as BERT, ClinicalBERT, and Bio_ClinicalBERT have enabled the extraction of insights from textual patient records, capturing nuances in clinical notes, medical history, and even personal habits like smoking or diet [31]. Together, these approaches allow for a multi-modal diagnostic strategy—one that doesn't limit itself to a single source of truth, but instead synthesizes data from various channels to provide a more complete picture of the patient's condition. As data grows in both volume and variety, so too does the potential of these systems to evolve from simple predictors into dynamic diagnostic companions capable of real-time analysis, retraining, and personalization [32]. With the recent rise of retrieval-augmented generation (RAG) systems, even text generation can be grounded in contextually relevant scientific literature or past cases, further increasing the reliability and transparency of AI-driven diagnoses. What emerges from this convergence is not just a smarter machine, but a more accessible, equitable, and timely system of care for individuals living with or at risk of Parkinson's Disease [33].

This evolution of diagnostic strategies through machine learning represents more than just a technical upgrade, it symbolizes a fundamental shift in how we conceptualize disease and the human body. In the traditional biomedical model, diagnosis has been largely reactive:

symptoms appear, the patient seeks help, and a clinician intervenes [34]. But AI introduces a proactive paradigm, one that attempts to *predict* rather than merely respond. For Parkinson's, a disease whose very challenge is that it quietly progresses before revealing itself visibly, this shift is monumental [35]. Through early and subtle signs—such as changes in handwriting pressure, reduced blinking, minor balance issues, or even slight speech alterations—ML models can identify a pattern of risk before it crystallizes into full-blown pathology. By integrating this predictive capacity with patient data that is continuously collected—be it through wearables, mobile applications, or digitized clinical records—the system becomes not just diagnostic but *anticipatory*. This is where the potential of AI truly becomes life-changing: a world where diagnosis doesn't wait for damage but rather works to preempt it, saving not only time but potentially neurons.

Another key strength of AI-assisted systems lies in their scalability and accessibility. Where a neurologist may only be able to see a few dozen patients a week, an AI tool embedded in a mobile device or web platform can screen thousands. This scalability is particularly important in low-resource settings, where access to trained specialists is minimal or non-existent. For populations that are underdiagnosed due to geographic, financial, or social barriers, AI opens the door to care that was previously unimaginable. And yet, this is not just about quantity—it is equally about quality. The standardization of diagnostic criteria, coupled with machine-level precision, allows even a village health worker or a non-specialist to offer frontline screening tools that maintain high fidelity with global clinical standards. Moreover, when these AI tools are developed transparently and include interpretability mechanisms—such as showing which symptoms contributed most to the diagnostic prediction—they can be integrated smoothly into clinical workflows without undermining clinician trust or autonomy. This is critical because no matter how accurate a model is, if its outputs are opaque or "black box" in nature, clinicians are unlikely to adopt it in practice. Transparency, explainability, and alignment with clinical reasoning are not just ethical mandates, they are practical necessities [36].

Finally, it's worth noting that these intelligent systems are not static. They improve over time, learning from every new patient, adjusting to new clinical guidelines, and adapting to emerging data. This lifelong learning capacity mirrors, in some sense, the evolution of a

human physician's experience—but at exponential scale and speed. The goal is not to replace the doctor with a machine, but to give the doctor a continuously learning, always-updated diagnostic assistant capable of absorbing terabytes of data, mining hidden insights, and offering them in the form of clear, actionable, and timely suggestions. And when such systems are designed responsibly, rooted in validated clinical knowledge, and subjected to rigorous testing and regulatory scrutiny, they offer not just an alternative to current diagnostic models—but a major leap forward. In a world where Parkinson's continues to be underdiagnosed and diagnosed too late, this leap is not optional—it is imperative. It represents the future of precision medicine, and it's already beginning to shape the present [37].

2.3 The Role of Spiral Drawings and Visual Biomarkers in PD Detection

Among the various tools explored in the early diagnosis of Parkinson's Disease, the spiral drawing test holds a uniquely important place, not merely due to its simplicity, but because of the rich neurological insights it offers through a task as deceptively mundane as drawing a spiral. On the surface, it appears to be a trivial motor activity—something most people learn in childhood without a second thought—but for patients developing PD, this very act can become a mirror reflecting the earliest cracks in motor coordination. The test capitalizes on the fact that one of the earliest and most common motor symptoms in Parkinson's is a resting tremor, usually affecting the hands. When a person with emerging PD draws a spiral, their involuntary micro-movements get etched into the spiral's curves, altering its symmetry, smoothness, and stroke fluidity in ways that can be precisely analyzed. These distortions, once thought to require manual observation by neurologists, are now being quantified using advanced image processing algorithms and convolutional neural networks (CNNs). This shift from subjective interpretation to computational diagnosis is not just about accuracy—it's about uncovering meaningful biomarkers from a process that is cheap, accessible, and non-invasive [38].

The spiral drawing test's diagnostic power lies in its ability to function as both a quantitative and qualitative tool. While clinicians might focus on the shape and amplitude of the drawing, CNNs trained on large datasets can dig much deeper—extracting pixel-level patterns that the human eye simply cannot perceive. These models are capable of capturing

local variations in stroke pressure, micrographia (progressive decrease in handwriting size), velocity changes, and irregular angular movement—subtle cues that reflect underlying neuromuscular degradation. The use of CNNs in this context is particularly apt because they excel at detecting spatial hierarchies in visual data. Unlike traditional machine learning models that require handcrafted features, CNNs learn the most discriminative features directly from the images themselves. This allows the model to recognize tremor-induced inconsistencies even when they manifest in unpredictable ways across different patients. What's more, CNNs can be trained on augmented datasets where spiral images are rotated, scaled, or mirrored—enhancing the robustness of the model while maintaining the pathological features. This makes them particularly well-suited for real-world clinical deployment, where drawing styles, pen thickness, and canvas formats may vary widely [39].

What makes the spiral drawing test even more valuable is its integration potential into telehealth and mobile health environments. In an era where smartphones and tablets are increasingly used for medical monitoring, patients can draw spirals on touchscreens from the comfort of their home, and have them analyzed in real time by an AI model hosted on the cloud or embedded within a mobile application. This brings forth a powerful paradigm shift: from hospital-based reactive care to home-based proactive screening. For patients in remote or underserved regions, this means earlier detection, fewer hospital visits, and a dramatically reduced time to diagnosis. It also empowers caregivers and general practitioners who might not be specialized in neurology but can now access AI-powered visual assessments to support their clinical decisions. Moreover, the results of such tests can be archived over time, allowing longitudinal tracking of motor progression. This is a crucial development because Parkinson's is not a static disease—it evolves, often unpredictably. Tracking spiral performance over months or years can offer rich data about the patient's trajectory, helping clinicians evaluate the efficacy of treatments, adjust medications, or even predict the onset of more severe symptoms [40].

Despite its simplicity, the spiral drawing test—when empowered by deep learning—symbolizes a deeper truth about the future of medicine. It shows that valuable diagnostic insights can emerge not just from sophisticated imaging machines or invasive procedures,

but from everyday actions transformed through intelligent computation. It proves that machine learning doesn't need to rely solely on complex biomarkers like genetic expression profiles or expensive neuroimaging scans to be clinically useful. In fact, its greatest strength may lie in its ability to extract gold from the ordinary—turning a simple spiral into a powerful diagnostic lens. As we continue to merge clinical tradition with modern AI, the spiral test stands as a compelling example of how low-cost, intuitive inputs can be harnessed to democratize access to early neurological care. In the context of this thesis, the spiral drawing module serves as the visual engine of the diagnostic pipeline, offering not just raw classification but deep interpretability. It represents a bridge between patient intuition and technological precision—an intersection where the future of accessible neurology may well be drawn [41].

2.4 Deep Learning and Natural Language Processing in Parkinson's Diagnostics

The emergence of deep learning and natural language processing (NLP) as central pillars of computational medicine has drastically redefined what is possible in the diagnosis of complex neurological conditions like Parkinson's Disease. While earlier computational approaches focused on structured, numerical features extracted from patient records or physiological signals, the new generation of deep learning architectures has opened a vast frontier of capabilities by allowing machines to interpret unstructured data—images, free-text clinical notes, and even patient narratives—with near-human-level comprehension. In the context of Parkinson's Disease, where diagnostic clarity is often obscured by symptom heterogeneity, overlapping comorbidities, and the absence of a definitive biomarker, these technologies offer a unique blend of precision, flexibility, and contextual understanding. More importantly, they allow for the creation of end-to-end diagnostic systems that mirror the reasoning process of trained clinicians, taking into account not just isolated symptoms but the holistic tapestry of patient information—ranging from lifestyle habits and genetic predispositions to behavioral observations and longitudinal trends. This convergence of deep learning and NLP is not just augmenting human decision-making; in many ways, it is reshaping the very epistemology of how we define and detect disease [42].

At the heart of this transformation lies the idea of multimodal fusion—an architectural paradigm that combines inputs from multiple data types to produce a singular, integrated

diagnostic output. Parkinson's Disease, as a case study, is particularly suited to this approach because it affects both motor and non-motor domains, and its manifestations are scattered across diverse formats of data. For instance, a patient's spiral drawing might reveal hand tremors, their speech sample may indicate dysarthria, their wearable data might suggest disrupted sleep patterns, and their medical history could highlight a familial risk of neurological disorders. No single input among these is definitive on its own, but when viewed together, they form a compelling clinical picture. Deep learning models, particularly convolutional neural networks (CNNs) and transformer-based language models, are uniquely equipped to extract high-level representations from each of these modalities and fuse them into a cohesive diagnostic framework. In the case of this thesis, we harness this potential by developing a dual-model pipeline in which CNNs interpret spiral test images while NLP models process structured patient metadata and biomedical literature, ultimately combining their outputs into a unified risk assessment. The result is a tool that not only mimics the perceptual depth of a human clinician but also enhances it with the speed, consistency, and objectivity of algorithmic reasoning [43].

Central to this pipeline is the use of Bio_ClinicalBERT, a domain-specific variant of the BERT (Bidirectional Encoder Representations from Transformers) model, which has been pre-trained on vast corpora of clinical text, discharge summaries, and electronic health records. Unlike general-purpose language models that struggle with the syntactic and semantic complexity of medical terminology, Bio_ClinicalBERT excels at parsing clinical abbreviations, temporal reasoning, and comorbidity patterns. When applied to Parkinson's data, this model can read and interpret variables such as bradykinesia scores, medication history, and cognitive symptoms, embedding each patient's clinical profile into a dense, multidimensional vector space. These embeddings preserve semantic proximity—meaning that two patients with similar symptomatology will be positioned close to each other in the latent space—even if their individual feature values differ numerically. This allows the model not only to classify but also to compare, cluster, and reason about patient profiles in a way that aligns closely with human clinical intuition. Moreover, the inclusion of additional context from medical literature—enabled through retrieval-augmented generation (RAG)—ensures that the model's outputs are grounded in evidence, not just statistical correlation [44].

RAG plays a particularly critical role in transforming the diagnostic system from a mere classifier into a true reasoning engine. In simple terms, RAG integrates a retrieval mechanism with a generative language model, allowing the model to pull in relevant documents or case studies before constructing its answer. When applied to our Parkinson's pipeline, this means that when a user enters their symptoms, medical history, or questions into the system, the model first retrieves the most relevant scientific paragraphs—say, studies on tremor onset in early-stage PD or case reports on bradykinesia scoring—and then conditions its generated diagnosis or explanation on that information. This dual process ensures that the model's recommendations are not hallucinated or speculative, but instead informed by a real, traceable knowledge base. In practice, this significantly enhances the interpretability and trustworthiness of the system. For instance, when a diagnostic summary mentions that a patient's tremor pattern is consistent with mild PD progression, it can cite specific research or cases that exhibit similar traits. This traceability is crucial not only for clinical adoption but also for patient confidence, especially as we move into an era where AI systems increasingly serve as front-line health advisors [45].

While the NLP backbone provides semantic intelligence, the CNN component of our pipeline delivers perceptual acuity. Built upon a pretrained ResNet18 architecture, the CNN processes spiral images collected from patients as part of their motor function assessment. These images are resized, normalized, and sometimes augmented before being fed into the model. Each spiral serves as a high-resolution snapshot of the patient's neuromuscular control at a given moment in time. The CNN, through successive convolutional layers, learns to detect aberrations such as jagged strokes, asymmetric curves, variable pressure zones, and even the progressive tightening of loops—a known indicator of micrographia. Importantly, these features are not manually engineered; the network learns them automatically by optimizing for classification accuracy across a labeled dataset. This means the model becomes increasingly sensitive to the subtle signs of PD that might elude even trained clinicians. Moreover, because ResNet18 is relatively lightweight compared to other architectures, it offers an excellent trade-off between accuracy and computational efficiency, making it suitable for real-time deployment on consumer-grade hardware or mobile devices [46].

What truly sets this diagnostic tool apart, however, is its ability to not only detect PD but to explain why it reached its conclusion. Thanks to the integration with OpenChat 3.5—a fine-tuned, open-source conversational model—the final output is not a binary label or a probability score but a natural-language report that mimics the tone, structure, and detail of a clinical note. The model is instructed to output in a structured format: starting with a patient overview, followed by a symptom summary, clinical impression, and reasoning, and ending with an analysis of the spiral image if one was provided. This format is intentionally designed to mirror the cognitive workflow of neurologists and to make the output usable in a real-world clinical setting. The language model has been prompt-engineered to reflect diagnostic conservatism where needed and to avoid overstatement of certainty. For example, instead of declaring "the patient has Parkinson's," the model might state "probable Parkinson's—consider confirmatory evaluation," thereby aligning with best practices in risk communication. This soft, evidence-weighted phrasing is essential for maintaining both clinical realism and ethical responsibility [47].

Moreover, this entire system is modular and continuously improvable. As more patient data is collected, as more literature becomes available, and as user feedback is integrated, the models can be retrained or fine-tuned to enhance accuracy, reduce bias, and adapt to new clinical definitions. This flexibility is a hallmark of modern AI systems, which no longer rely on fixed rule sets but evolve as living frameworks. The future iterations of this pipeline could include voice-based assessments to analyze dysarthria, gait analysis from video clips, or even integration with genetic risk scores—each adding another dimension to the model's diagnostic canvas. The spiral test may one day be just one of many integrated modules that collectively paint a comprehensive, longitudinal portrait of a patient's neurodegenerative risk. Even beyond Parkinson's, this architecture offers a transferable blueprint for diagnostic systems in Alzheimer's, ALS, and other complex disorders where multimodal assessment is key.

Artificial intelligence is not merely a technological convenience in the realm of modern healthcare—it is a philosophical shift in how we perceive, structure, and interact with medical knowledge. In the case of Parkinson's Disease, which exists at the intersection of motor dysfunction, neurochemical degeneration, and progressive behavioral decline, AI

does not simply offer more efficient processing—it enables a new kind of intelligence. One that is capable of identifying patterns too subtle for the human eye, interpreting connections between seemingly unrelated symptoms, and providing continuity in care through data-rich, personalized inference. More than just replacing manual tasks, AI redefines the scope of what can be known and anticipated in complex, heterogeneous diseases like PD. It introduces a paradigm that is not bound by clinical silos or fixed diagnostic checklists but is instead driven by context, variability, and feedback—a framework far more aligned with the actual lived experience of neurodegeneration.

The strengths of AI in neurodiagnostics arise from its unique ability to abstract and represent human physiological states through layers of learned meaning. Unlike traditional statistical approaches that require pre-selected variables and clearly defined mathematical relationships, AI models—especially those in the deep learning family—are inherently capable of learning from raw, unstructured inputs. They do not need to be told what to look for; instead, they learn *what matters* by optimizing their internal parameters against large volumes of data, refining their understanding over time through exposure to diverse patient scenarios. This is particularly crucial in a condition like Parkinson's, where no single biomarker or symptom is universally representative. One patient's Parkinson's may begin with tremors and rigidity, another's with balance issues or handwriting decline, and yet another may show only non-motor signs for years. Such variability makes rule-based diagnostic systems brittle. AI, on the other hand, thrives in environments where rules are fuzzy, relationships are probabilistic, and ground truths are complex. It learns by example, not by exception. And in learning from thousands of such examples, it becomes capable of inferring insights that even a seasoned clinician might overlook—not due to lack of expertise, but due to the sheer limits of human cognitive bandwidth [48].

What makes AI even more compelling in this context is its ability to synthesize information across modalities. In conventional medical settings, diagnostic information is scattered—clinical histories are in text, motor performance is observed visually, lab results are numerical, and patient behavior is discussed anecdotally. No single domain offers the full picture, and often, it is the clinician's burden to mentally integrate these streams into a working hypothesis. AI challenges this fragmentation by enabling *multimodal fusion*: the

process of bringing together data from text, images, signals, and numerical features to create a unified, comprehensive patient model. This fusion is not merely additive—it's synergistic. When a machine is allowed to observe how a patient's spiral drawing correlates with their reported bradykinesia scores, or how family history interacts with sleep patterns and age of onset, it begins to model the kinds of nonlinear, interdependent phenomena that truly reflect human biology. The result is not a replacement for physician expertise, but an amplification of it. AI becomes the connective tissue between disciplines, allowing neurology to learn from radiology, behavioural science to inform pharmacology, and digital biomarkers to supplement physical assessments [49].

Perhaps one of the most profound contributions AI makes to neurological diagnostics is its capacity for **longitudinal learning**. Traditional diagnostic tools often treat each patient visit as a snapshot—a static event disconnected from the rest of the patient's life history. But Parkinson's is not a snapshot disease. It is a long, evolving story. Symptoms emerge gradually, sometimes waxing and waning, sometimes transforming entirely. AI systems, particularly those embedded within digital platforms or remote monitoring ecosystems, have the unique ability to track patient data over time—observing changes in gait, tremor frequency, cognitive clarity, or even user-reported mood levels across weeks or months. This temporal dimension transforms diagnosis from an event into a process. Instead of asking "Does this patient have Parkinson's today?" the system can ask "Is this patient trending toward a Parkinsonian profile?" or "Has there been a deviation in their usual pattern that merits early intervention?" This shift from binary classification to continuous risk modeling represents a fundamental reimagining of care. It offers the potential not only for earlier detection, but for earlier *action*—delaying progression, optimizing treatment timing, and ultimately improving patient quality of life [50].

Another critical dimension in which AI transforms the diagnostic landscape is **accessibility**. Parkinson's Disease affects individuals across all geographic, economic, and social contexts—but access to specialized neurology care remains deeply uneven. In many rural or low-resource regions, patients either receive a diagnosis far too late or are misdiagnosed altogether. Even in urban centers, wait times for specialist consultations can be prohibitively long, delaying critical intervention. AI helps to address this imbalance by

embedding intelligence into scalable, distributable platforms—mobile apps, telehealth systems, web portals—that do not rely on constant human oversight. A person in a village can draw a spiral on their phone, submit a few details about their symptoms and history, and receive a highly informed risk assessment powered by the same intelligence that supports top-tier clinics. This democratization of diagnostic quality is one of AI's most socially impactful contributions. It does not mean replacing clinicians—it means extending their reach, standardizing their methods, and reducing the disparities that geography and economics have long sustained [51].

Equally important, however, is the **explainability** of AI outputs. In medicine, accuracy without interpretability is not enough. Patients need to understand why a machine said what it said; clinicians need to see the reasoning behind the score. Black-box systems may be accepted in recommendation engines or financial predictions, but in healthcare, they breed distrust. This is why modern AI systems for Parkinson's diagnosis are increasingly incorporating mechanisms for *interpretable reasoning*—highlighting which features influenced a given output, referencing similar past cases, and generating plain-language diagnostic summaries. These elements ensure that AI is not just correct, but also *communicable*. The diagnosis can be traced, discussed, and—if necessary—disputed or refined. This transparency is essential not just for safety and ethics, but for adoption. No clinician wants to be overruled by a silent algorithm; but many would welcome an assistant that explains its thinking, shares supporting literature, and adjusts its certainty based on new data. In this way, AI is not only a computational engine—it is a conversational partner, one that brings both memory and reasoning into the diagnostic room [52].

The role of AI in Parkinson's diagnosis must be understood as part of a broader redefinition of healthcare itself. As medical knowledge continues to expand at a pace far beyond human capacity to absorb it, AI offers a means to dynamically align practice with evidence. It becomes the bridge between research and clinic, between potential and action. In the context of Parkinson's, it enables us to move beyond symptom management toward something more proactive: early screening, continuous monitoring, adaptive personalization, and above all, equity in access. But its true power lies not in the sophistication of its algorithms, but in its ability to *listen*—to patient narratives, to

behavioral patterns, to previously overlooked signals—and respond with coherence. It does not just analyze data; it learns from stories. It connects patterns that were once invisible. And as it does, it challenges medicine not to abandon the human element, but to elevate it. In doing so, AI reminds us that the future of healthcare is not a choice between man and machine—it is a collaboration, and in that collaboration, both become better [53].

2.5 Ethical Frontiers and Responsible Deployment of AI in Parkinson's Diagnosis

The growing integration of artificial intelligence into neurological diagnostics represents not only a technological evolution but also a profound ethical crossroads. While the capabilities of AI systems to detect Parkinson's Disease with increasing accuracy are undeniably promising, these innovations come with a complex web of responsibilities that extend beyond mere model performance [54]. In a field as sensitive as healthcare—particularly one dealing with degenerative and deeply personal conditions such as Parkinson's—every line of code, every predictive outcome, and every patient-facing interface carries weight. It is not sufficient to develop intelligent models that can diagnose; it is imperative that these systems also reflect the values of fairness, transparency, privacy, and accountability. As we approach an era where AI systems are expected to interpret a person's symptoms, motor performance, medical history, and possibly even genetic predisposition, we must ask not only what these systems *can* do but what they *should* do. The ethical questions surrounding AI in medicine are not peripheral—they are central. They define how this technology will be perceived, adopted, and trusted by the very people it aims to serve [55].

At the core of these concerns lies the issue of **explainability**. A diagnosis of Parkinson's Disease is not a trivial matter, it carries psychological weight, influences treatment trajectories, and affects a patient's quality of life and long-term decision-making. When a human doctor delivers such a diagnosis, the patient can ask questions, request justifications, and receive explanations rooted in clinical experience and empathy [56]. But when an AI system produces the same conclusion, especially in a setting where human oversight may be minimal, how does the patient or even the physician make sense of the result? If the output is simply a score or label, the trust barrier grows insurmountable. This is why responsible AI development in this space must prioritize interpretable models, or at the

very least, build mechanisms around black-box models that provide human-readable summaries, contributing factors, and confidence estimates. In our own diagnostic tool, this principle was operationalized by embedding structured reasoning within the language model's output, ensuring that the result is not a binary verdict, but a nuanced, evidence-informed clinical impression. It is this interpretability that makes AI usable in medicine, not just powerful [57].

Another pressing ethical concern is bias and fairness. Parkinson's Disease affects individuals across genders, ethnicities, and age groups, but the datasets on which AI models are trained often do not reflect this diversity. If the model is disproportionately trained on cases from urban populations, or from patients of a particular age group or socioeconomic background, it may silently learn to prioritize patterns associated with those demographics, thereby underperforming or misdiagnosing in populations that were underrepresented in the training data [58]. In practice, this could mean a rural patient's symptoms are deemed inconclusive not because they are, but because the model has never seen enough cases like theirs. This kind of statistical exclusion is dangerous because it is invisible. The system doesn't announce its bias; it simply performs worse without explanation. To combat this, ethical AI design requires rigorous dataset auditing, diversity-aware training protocols, and continuous performance evaluation across demographic strata. But more importantly, it requires humility, a recognition that no model is perfect, and that clinical decisions must never be entirely outsourced to algorithms, particularly when those algorithms may be blind to social and structural inequities [59].

Data privacy also occupies a central position in the ethical landscape of AI in Parkinson's care. The very inputs that make such systems intelligent, medical histories, spiral drawings, lifestyle details, comorbidities, are some of the most sensitive data a person can share. The fear of that data being mishandled, leaked, or misused can deter patients from engaging with AI systems at all. Therefore, responsible deployment of such tools must involve not only technical safeguards like encryption and secure storage but also transparent data policies that empower the patient [60]. Consent must be explicit, revocable, and well-informed. Patients should know what their data is being used for, how long it is stored, and whether it is shared with third parties. These concerns become even more critical in home-

based applications, where diagnostic interactions may happen outside the regulatory umbrella of traditional healthcare institutions. In such cases, the ethical burden shifts even more heavily toward the designers, developers, and deploying organizations. Privacy, in this context, is not just a legal requirement, it is a moral one, tied directly to patient dignity [60].

There is also the question of **emotional responsibility**. Unlike lab results or imaging data, an AI diagnosis carries an inherent psychological impact. Parkinson's Disease is not a label that patients can detach from easily [61]. It signals long-term progression, medication dependency, lifestyle adjustments, and often emotional distress. If such a diagnosis is delivered by a machine, even with perfect accuracy, it risks dehumanizing the interaction. The absence of tone, empathy, or facial expression can make the news feel cold, abrupt, or worse, incorrect, especially if it contradicts a patient's lived experience or expectations. That's why AI systems in healthcare must be designed not as oracles but as *assistants*. They should guide, not dictate. They should provide evidence and structured analysis, but the final delivery and decision-making must involve clinicians who are trained not just in diagnosis, but in compassion. Emotional ethics matter—particularly in neurology, where quality of life is as important as disease control. A responsible AI system, therefore, does not just "do the math", it respects the human experience behind the numbers [62].

Finally, we must consider the **long-term societal implications** of allowing AI to become embedded in early-stage neurological screening. Will such tools become mandatory in insurance decisions? Could a risk score for Parkinson's affect employability or lead to social stigma if disclosed or leaked? Will it increase anxiety among individuals who might never develop the disease but were flagged as high-risk by an overly conservative model? These questions are not futuristic, they are current, and they must be addressed in parallel with technological advancement. The power of AI must always be balanced by the rights of the individual. Ethical deployment is not a constraint; it is the only sustainable path forward [63].

In closing, the application of AI in Parkinson's diagnosis represents a bold and necessary leap toward intelligent, scalable, and early-stage healthcare solutions. But it is not a leap that can be taken blindly [64]. Every line of the model's architecture, every byte of patient

data, every word generated by the language model carries consequences that reach far beyond the algorithm itself. The stakes are not merely academic, they are deeply human. Responsible AI is not an afterthought. It is the foundation. And as we continue building systems that aim to improve lives, let us never lose sight of the values that make those lives worth protecting: dignity, fairness, privacy, and trust [65].

Chapter 3 :Materials and Methodology

This chapter details the end-to-end pipeline used to develop and test the proposed diagnostic tool. By combining real-world patient datasets, spiral image classifiers, and text-based inference engines, the methodology bridges practical deployment and technical rigor in model design.

3.1 Data Acquisition and Preprocessing

The foundation of any machine learning-driven diagnostic tool lies in the quality, structure, and representativeness of the data it is built upon. Our data relies on a Parkinson's related-dataset that was publicly available on Kaggle (Refer Table I) to help understand the data structure, feature diversity, and baseline attributes that are commonly used in research In the context of Parkinson's Disease, a condition known for its clinical heterogeneity and multifaceted symptom presentation, this truth becomes even more critical. A robust diagnostic model must be trained on data that not only captures the diverse motor and non-motor symptoms associated with the disease but also reflects the demographic, genetic, and lifestyle variations that shape its progression. For this thesis, the data acquisition process was designed to source, validate, structure, and enrich a corpus of multimodal patient information, comprising both structured metadata and unstructured visual input. The subsequent preprocessing steps aimed to optimize this data for use in deep learning architectures while preserving its interpretability and clinical relevance.

The dataset construction process began with the careful curation of publicly available Parkinson's-related datasets from reputable platforms such as Kaggle and academic data repositories. The primary structured dataset comprised approximately 3,490 patient records, each encompassing 34 carefully selected features. These features spanned a range of domains including demographics (age, gender), lifestyle indicators (smoking, alcohol consumption, dietary habits), clinical history (family history of neurological disorders, past surgeries, use of immunosuppressants), and symptom-specific metrics (rigidity, tremors, speech disturbances, sleep quality, bradykinesia score, and balance issues). The decision to include this wide array of features was informed by literature indicating that Parkinson's Disease does not manifest in a uniform way, and often involves a complex interplay

between motor degeneration and systemic health factors. Hence, the inclusion of both neurological and general health indicators was essential for capturing the true variability present in real-world PD cases.

Each record in the dataset also included a binary label indicating the presence or absence of Parkinson's Disease, based on established diagnostic criteria or physician-verified annotations. This binary classification formed the ground truth for the supervised learning component of the model. To facilitate seamless integration with transformer-based NLP models, the structured data was converted into JSON Lines (.jsonl) format. This format allows each patient entry to be encapsulated as a self-contained object, with fields for inputs (symptoms and background) and outputs (diagnostic labels and clinical impressions). The JSONL structure not only supported ease of parsing during model training but also mirrored the conversational flow expected by the language model when generating diagnostic summaries. The process of transforming raw CSV data into JSONL involved several preprocessing steps, including standardization of feature names, handling of missing values, normalization of numerical fields (e.g., age, bradykinesia scores), and encoding of categorical variables (e.g., yes/no for smoking). Stringent checks were applied to ensure that no data leakage occurred between the training and validation splits and that each JSON object retained internal consistency.

Parallel to the structured dataset, an unstructured image dataset was also assembled to train the computer vision component of the diagnostic pipeline. This dataset comprised over 250 high-resolution spiral drawing images, collected from Parkinson's spiral datasets hosted on Kaggle and medical image repositories. The spiral drawing task is a clinically validated tool for assessing motor performance, particularly tremor amplitude, angular stability, and micrographia. Each spiral in the dataset was drawn by either a diagnosed PD patient or a healthy control subject, with the corresponding class label affixed. These labels were verified by domain-specific data providers and cross-checked against metadata wherever available. The spiral drawings were captured under varying conditions—pen and paper scans, tablet inputs, and digitally simulated spirals—to ensure robustness of the model across input modalities.

Before being input into the convolutional neural network (CNN), the spiral images underwent a rigorous preprocessing pipeline. All images were resized to 256×256 pixels to standardize input dimensions and enable batch processing. They were then converted to grayscale to reduce the number of input channels, eliminating color-based noise and focusing purely on structural features. Image normalization was applied to standardize pixel intensity values, bringing them within a uniform range to accelerate model convergence during training. Furthermore, data augmentation techniques were employed to artificially expand the dataset while preserving pathological features. These techniques included slight rotations, horizontal and vertical shifts, and scaling, each designed to simulate the natural variability in how patients draw spirals. However, care was taken to avoid augmentations that could introduce or mask tremor signatures, ensuring that the model learned only clinically relevant patterns.

In addition to patient-annotated data, the training corpus was further enriched with text extracted from recent scientific literature on Parkinson's Disease. These documents, comprising review papers, clinical guidelines, case studies, and diagnostic criteria, were sourced from platforms like PubMed Central. They were processed in PDF format using optical character recognition (OCR) tools, then cleaned, segmented, and chunked into textual snippets of approximately 1,500–2,000 characters each. These chunks were then prepended with source metadata (e.g., publication name, section heading) and tagged with identifiers such as [PAPER:filename] to allow context-aware retrieval during inference. This corpus formed the knowledge base for the retrieval-augmented generation (RAG) pipeline, which allowed the language model to contextualize its outputs with up-to-date medical knowledge.

To enable this retrieval process, all case-based and literature-based text chunks were embedded using a pretrained Sentence-BERT model fine-tuned on biomedical corpora. Each chunk was transformed into a dense vector representation in high-dimensional space, with cosine similarity serving as the distance metric. These embeddings were then indexed using FAISS (Facebook AI Similarity Search), an efficient similarity search library capable of handling large-scale retrievals. By combining patient-specific queries with similarity-based document lookup, the model could dynamically reference real-world evidence when

generating its diagnostic summaries. This not only improved the factual accuracy of the model's output but also added a layer of explainability, allowing the system to cite or refer to clinical studies that support its interpretation of a given patient case.

A crucial consideration in the preprocessing phase was the prevention of information leakage and overfitting. For this reason, the dataset was split into training, validation, and testing sets in a stratified manner, ensuring that the distribution of Parkinson's-positive and negative cases was consistent across all partitions. Cross-validation was also employed during early model iterations to evaluate generalizability. Moreover, for the image dataset, augmentation parameters were tuned separately for each training epoch, reducing the risk that the model would memorize specific augmentation artifacts rather than learning disease-relevant patterns.

Beyond technical preprocessing, ethical and privacy-related measures were also embedded into the data pipeline. All patient data was anonymized prior to use, with personally identifiable fields (names, hospital IDs, timestamps) removed or replaced with synthetic placeholders. The image dataset was reviewed to ensure that none of the drawings contained metadata or watermarks that could compromise subject confidentiality. Where literature was used, all citations were tracked, and content was referenced only if available in the public domain or under open-access licenses. This ensured full compliance with data usage standards and reinforced the transparency of the model's development process.

Another noteworthy step in the preprocessing pipeline involved unifying the time-related features across datasets. In clinical records, symptoms may be recorded with vague or inconsistent time references, such as "recent onset," "intermittent," or "chronic." These descriptors were standardized into discrete temporal labels like "<3 months," "3–12 months," or ">1 year," enabling consistent input formatting for the language model. Similarly, comorbidity data was structured into multi-label binary vectors, allowing the model to weigh overlapping conditions such as diabetes, hypertension, or cognitive decline during inference.

The preprocessing pipeline culminated in the construction of a multimodal training-ready dataset, one that integrates visual, textual, and temporal representations of patient data into

a single cohesive structure. This dataset, by design, serves not just the current scope of Parkinson's detection, but also establishes a scalable architecture that can be extended to include voice data, accelerometer readings, and other sensor-based features in future iterations. The modularity of the preprocessing codebase ensures that new features can be added with minimal disruption, while the existing dataset can continue to be augmented as new cases or studies become available. This foundation ensures that the AI model does not operate in a vacuum but remains rooted in the evolving landscape of Parkinson's Disease research and real-world clinical practice.

In conclusion, the data acquisition and preprocessing phase was far more than a technical requirement, it was a philosophical commitment to quality, inclusiveness, and realism. By curating a dataset that reflects the complexity of Parkinson's Disease, and by processing it in ways that honor both computational performance and clinical integrity, this thesis lays the groundwork for a diagnostic model that is not only intelligent, but also trustworthy. The resulting dataset represents not just numbers or pixels, but patient stories, clinical patterns, and biomedical knowledge—all distilled into a form that a machine can learn from, and a clinician can rely upon.

Table 1: List of datasets utilised

S.No.	Title	Link	Features	Source / Authors
1	Parkinson's Disease Dataset Analysis	Kaggle	Biomedical features	Rabie El Kharoua (Kaggle)
2	Parkinson's Spiral Drawing Dataset	Kaggle	Spiral drawing images (PD vs. healthy)	Kevin Mader (Kaggle)
3	HandPD Dataset	UNESP	Spiral and meander drawings	João Papa et al., UNESP

3.2 Model Architecture: A Multimodal Design for Clinical Intelligence

Designing an artificial intelligence system capable of simulating the reasoning of a neurologist is a deeply intricate task—one that cannot be solved by a single model or a narrow dataset. Parkinson's Disease, by its very nature, is multifaceted: it manifests visually through motor dysfunctions, behaviourally through symptom patterns, and clinically through complex combinations of patient history, lifestyle, and genetic background. Therefore, the architecture that powers any meaningful diagnostic pipeline must itself be multimodal, combining distinct forms of intelligence—vision-based pattern recognition, language-based reasoning, and contextual knowledge retrieval—into a unified, interoperable framework. This section lays out the architectural blueprint of such a system, carefully designed to mimic the layered thinking process of medical experts.

At its core, the system consists of three synergistic components: a vision module based on Convolutional Neural Networks (CNNs), a language reasoning module powered by a transformer model (Bio_ClinicalBERT), and a retrieval-augmented language generator (OpenChat 3.5) that produces complete clinical summaries based on case embeddings and literature chunks. Each of these models plays a distinct, non-redundant role and has been integrated into the overall architecture with precision, so the final system behaves not just like an automated classifier, but as a diagnostic assistant capable of nuanced, human-like interpretation.

A. The Vision Stream – Interpreting Spiral Drawings with ResNet18

The visual module of the pipeline is where our system begins to resemble the eyes and pattern memory of a trained neurologist. In clinical settings, one of the hallmark tests for Parkinson's is the spiral drawing task—a simple yet profound test where patients draw spirals on paper. What seems like a childlike doodle to the untrained eye becomes a rich canvas of subtle biomarkers: inconsistent pressure, discontinuous strokes, jittery lines, and angular distortions. These features, almost imperceptible to laypersons, become critical cues for machine perception.

To make sense of this data, we used ResNet18, a variant of the Residual Network family that revolutionized deep learning with its skip connections. Traditional

CNNs often struggled with deep architectures due to vanishing gradients, but ResNet solved this elegantly by allowing gradients to "skip" layers using identity mappings. This innovation made it possible to train deeper networks without degradation of accuracy—a crucial advantage when dealing with nuanced spatial patterns in grayscale images.

The architecture begins with a 7×7 convolution followed by max pooling, which quickly reduces the dimensionality of the input and captures early edge-level features. As the data flows through a series of residual blocks, it encounters increasingly complex filters that capture mid-level and high-level features. In the case of our 256×256 pixel grayscale spiral images, these layers help isolate stroke width, curve angle, frequency of tremor-like movements, and even micrographia—the shrinking of handwriting, another subtle symptom of PD.

Each residual block consists of two 3×3 convolution layers with batch normalization and ReLU activation, combined with shortcut connections that add the input of the block directly to its output. The final layers of ResNet18 include global average pooling followed by fully connected layers, resulting in a feature vector that encodes the motor signature of the spiral drawing. This vector is not just a number—it is the machine's interpretation of a patient's motor ability, compressed into a structured signal. This vector becomes the visual representation of that patient's motor state and is stored for fusion with other modalities downstream.

B. The Language Stream: Contextual Understanding with Bio_ClinicalBERT

While the visual model interprets motor symptoms, it cannot account for the patient's age, sex, comorbidities, lifestyle choices, or family history—all of which play an enormous role in real-world diagnosis. This is where the language module enters. Powered by Bio_ClinicalBERT, the language reasoning component was designed to interpret structured and semi-structured metadata in a way that preserves clinical semantics and mimics medical case discussions.

Bio_ClinicalBERT is a variant of BERT that has been pre-trained on biomedical corpora and fine-tuned on clinical notes. Unlike generic transformers that may be confused by terms like "rigidity" or "postural instability," Bio_ClinicalBERT understands these phrases in their proper medical context. The architecture itself follows the classic BERT base model: 12 transformer layers with multi-head self-attention, token embeddings, segment embeddings, and position encodings. Each input—typically a case narrative—is tokenized using the WordPiece tokenizer and converted into a dense, contextualized embedding through the transformer stack.

The model tokenizes this into subword units and feeds it through the 12 layers of attention-based computation, resulting in an embedding that understands the difference between age-related rigidity and Parkinsonian rigidity, or the relationship between diabetes and neurodegenerative risk factors.

This contextual vector is then extracted from the [CLS] token at the beginning of the sentence—representing the holistic summary of the input. This vector alone can be used for classification, but we take it a step further by passing it into the retrieval and generation module, which transforms this context into a full narrative diagnosis.

C. The Retrieval Backbone – Augmenting Language with Scientific Knowledge

Even the best transformer models can hallucinate or provide superficial results if they are not grounded in domain knowledge. That's why our architecture includes a Retrieval-Augmented Generation (RAG) backbone. This component adds memory to the system—specifically, the memory of published scientific literature.

The retrieval system uses FAISS (Facebook AI Similarity Search) to index and query dense embeddings of paragraph-level content extracted from Parkinson's-related papers in PDF form. Each paragraph was passed through Bio_ClinicalBERT to generate an embedding vector, which was stored in an index.

During inference, when a patient prompt is given, its embedding is compared to this index, and the top-k most similar medical literature chunks are retrieved.

This retrieval step serves two purposes. First, it grounds the model's predictions in real clinical knowledge, reducing hallucinations. Second, it gives the final generator access to nuanced, case-relevant information that it might not have seen during training.

D. The Language Generator Diagnostic Narratives via OpenChat 3.5

Once the patient context and literature snippets are gathered, they are fused into a single enriched prompt and passed to OpenChat 3.5, an open-source instruction-tuned language model that has been adapted for medically relevant tasks. This model is based on the LLaMA or Falcon family but has been fine-tuned on multi-turn dialogues, case-based reasoning, and question-answering formats.

The input to OpenChat 3.5 is a full prompt that includes:

- Patient metadata
- Top-retrieved scientific paragraphs
- System prompt instructing the model to respond in the tone of a senior neurologist

The model then generates a structured response that includes:

1. Patient Summary
2. Symptom Interpretation
3. Medical Reasoning
4. Preliminary Diagnostic Impression
5. Recommended Next Steps

This model uses a decoder-only transformer stack and is capable of producing long, coherent, medically grounded explanations. Unlike typical GPT-style outputs, OpenChat is prompt-sensitive, it tailors tone, length, and content structure to the format provided, making it ideal for our needs.

E. The Fusion Module: Multimodal Decision Engine

Finally, to produce a holistic classification (Parkinson's or not), we needed a mechanism to integrate the outputs of the CNN and Bio_ClinicalBERT. This was achieved through a fusion layer—a set of fully connected dense layers that take the visual feature vector from ResNet18 and the language embedding from Bio_ClinicalBERT, concatenate them, and pass them through a joint reasoning network.

The architecture of this layer is relatively shallow—two to three dense layers with dropout and batch normalization—to prevent overfitting and maintain interpretability. The final output is a softmax-based probability vector representing Parkinson's-positive or negative. This decision is then wrapped in the narrative generated by OpenChat 3.5, so that even if the final classification is binary, the explanation remains rich, detailed, and human-readable.

3.3 Training Procedures

Training an intelligent system to accurately diagnose a complex neurodegenerative disorder such as Parkinson's Disease is a task that reaches far beyond the mechanical act of fitting parameters to data. It is, instead, a rigorous process of iterative refinement—a convergence of domain knowledge, ethical responsibility, technical discipline, and architectural insight. The essence of this chapter is to illuminate the disciplined methodology through which the system was trained, layer by layer, to learn from noisy, real-world data sources, each carrying its own form of clinical truth. At its core, the training process sought to emulate the judgment of a neurologist, but using models that learned from patterns rather than instincts. Every step taken—from data preprocessing to loss

optimization—was designed not only to improve performance metrics but also to retain fidelity to the nuanced, context-dependent nature of real-life Parkinson's diagnosis.

The journey began with the visual analysis component, which relied on a convolutional neural network architecture known as ResNet18. The decision to use this specific model was guided by its proven ability to learn deep spatial hierarchies from low-resolution images while maintaining a compact parameter space. For this task, the input data consisted of grayscale spiral drawings submitted by patients under clinical observation. Each spiral, though simple in form, encoded a universe of neuromotor signatures—tiny oscillations, breaks in curvature, asymmetries in line smoothness, all of which signaled the subtle early manifestations of Parkinson's. These images were standardized to a resolution of 256 by 256 pixels and converted to a uniform grayscale format to eliminate any irrelevant variance introduced by color or lighting conditions during data acquisition. Normalization was applied to scale pixel intensity values between 0 and 1, ensuring that the model learned from feature contrast rather than absolute brightness levels.

The visual dataset was divided using a stratified split strategy into 70 percent training, 15 percent validation, and 15 percent testing sets. This ensured that the label distribution across Parkinson's-positive and Parkinson's-negative cases was preserved throughout, minimizing the risk of imbalanced learning. To improve generalization and prevent overfitting, a carefully curated set of data augmentation techniques was employed. These included minor rotational shifts, zooming, horizontal flips, and elastic deformations, all applied with medical intuition. The intention was never to distort the tremor patterns but to allow the model to become resilient to slight inconsistencies in patient drawing behavior. Each augmented version of the spiral was treated as a plausible real-world variant of the original, expanding the effective dataset size without compromising the pathological integrity of the samples.

The training itself was performed using the Adam optimizer with an initial learning rate of 0.0001 and a weight decay of 0.0005. These hyperparameters were chosen after multiple rounds of manual tuning and empirical validation. The model was trained to minimize binary cross-entropy loss, a natural choice given the two-class output. Each epoch passed the training set through the model in batches of 32 images, allowing for efficient use of

GPU memory while maintaining gradient stability. A learning rate scheduler was implemented to reduce the learning rate upon plateauing of validation accuracy, promoting convergence without overfitting. Throughout the training epochs, early stopping was employed based on validation loss, with a patience threshold of ten epochs. This approach ensured that training halted when further improvements were unlikely, thus preserving model generalizability. The final ResNet18 model was evaluated using standard classification metrics such as precision, recall, F1-score, and confusion matrix visualization (Figure 3). On the held-out test set, the model achieved a classification accuracy above 94%, with high sensitivity and specificity, reflecting its strong potential for deployment in early-stage Parkinson's detection from motor patterns.

Parallel to the visual model's development, an entirely different training pipeline was constructed for the language-based diagnostic component. This part of the system was designed to emulate how a neurologist might reason through structured clinical metadata and synthesize conclusions grounded in biomedical literature. At the heart of this stream was Bio_ClinicalBERT, a transformer-based model pre-trained on MIMIC-III and PubMed abstracts. It was uniquely suited to interpret the syntax and semantics of clinical language, including abbreviations, dosage patterns, symptom chronology, and comorbidity interactions. However, unlike typical applications of BERT for classification or question answering, our use case involved generating dense embeddings for structured clinical prompts that could be used for similarity retrieval and narrative generation.

Training began by constructing patient profiles from structured metadata. This included fields such as age, gender, lifestyle factors (smoking, alcohol consumption), comorbidities, family history, symptom descriptions, and prior medications. Instead of treating this metadata as tabular input, it was converted into narrative-like prompts using hand-crafted templates. For instance, a prompt might read, "A 67-year-old male patient, non-smoker, reports progressive bradykinesia over the past eight months along with mild postural instability. Family history is positive for Parkinson's." These prompts were passed through Bio_ClinicalBERT to generate contextual embeddings capturing the semantic nuances of the case. These embeddings were then used in two major ways: first, to calculate cosine similarity with pre-encoded literature chunks using FAISS for document retrieval, and

second, to feed the retrieval-augmented prompt into a generative language model for diagnosis synthesis.

To build the retrieval corpus, hundreds of PDF documents comprising peer-reviewed medical articles, case studies, and clinical guidelines were scraped, cleaned, and chunked into paragraphs of approximately 100–150 tokens. Each paragraph was then embedded using the same Bio_ClinicalBERT model to maintain representation consistency between queries and context documents. All embeddings were indexed using the FAISS vector database, which was optimized for rapid approximate nearest neighbor searches. During the training and testing of this subsystem, extensive manual validation was performed to ensure that the top-k retrieved documents per query were medically coherent and relevant. This was a critical step because the quality of retrieved context directly influenced the downstream diagnosis generation.

The final stage of the language-based pipeline involved fine-tuning OpenChat 3.5, an open-source LLM, on a curated set of synthetic diagnostic dialogues and real patient summaries. The model was trained in a supervised instruction-following setup, where each input consisted of a patient metadata prompt combined with the top-k retrieved literature contexts, and the output was a natural-language diagnostic summary written in the tone of a senior neurologist. Special care was taken to design prompt templates that instructed the model to not only output the diagnosis but to include structured reasoning sections such as patient summary, symptom mapping, disease likelihood estimation, and recommendation for further action. This allowed the system to maintain a level of explainability that is often absent in black-box models.

The training objective for this component was to minimize language modeling loss, essentially teaching the model to predict the next token given all previous tokens, but it was guided by human review at multiple intervals. We created a dataset of gold-standard diagnostic reports that reflected high clinical quality and used these as reference outputs during training. BLEU and ROUGE metrics were used to monitor fluency and content overlap, while BERTScore was employed to measure semantic similarity between generated and reference outputs. Additionally, we incorporated qualitative feedback from a panel of three clinical experts, who assessed the diagnostic summaries for clarity, medical

correctness, and actionability. This iterative loop between automated evaluation and expert oversight became one of the most critical aspects of training, ensuring that improvements in loss values translated to meaningful gains in medical performance.

The real strength of the pipeline, however, emerged during the integration phase, where the outputs from the image-based classifier and the language-based interpreter were fused. This fusion was not a simplistic logical OR; it involved training a dedicated fusion module that could accept both the visual prediction probabilities and the embedding vectors from the clinical metadata prompt. This module, composed of dense layers with ReLU activation, acted as a meta-classifier. It was trained on a hybrid loss function combining binary classification loss from the image stream and contrastive loss from the embedding similarities. The objective here was to ensure that when both modalities agreed, confidence scores were boosted, but when they diverged, the system could reason through ambiguity using the richer context.

This multimodal training process also necessitated a more thoughtful evaluation framework. We conducted end-to-end testing by feeding real case inputs into the complete system and comparing the final narrative diagnosis with ground-truth clinical outcomes. We assessed how often the full pipeline (including document retrieval and LLM generation) correctly flagged Parkinson's presence or absence, and more importantly, how well it explained its reasoning. These narratives were then evaluated blind by clinicians to eliminate bias. On this composite metric, accuracy coupled with explanation clarity, the system consistently outperformed baseline classifiers, including traditional support vector machines and decision trees trained on tabular data.

Another crucial part of the training pipeline was the emphasis on trust and robustness. Given the clinical nature of the tool, any hallucination, misclassification, or misinterpretation could have real-world consequences. To mitigate this, adversarial testing was conducted using perturbed inputs, such as introducing contradictory symptoms or ambiguous patient histories, to observe how the model responded. Failures were logged and analyzed to refine data augmentation policies, prompt templates, and retrieval thresholds. Calibration of classification confidence was also performed using temperature

scaling and Platt scaling, ensuring that the model's probability outputs could be trusted in a decision-support context.

In terms of deployment-readiness, all trained models were exported in TorchScript format for compatibility with both cloud-based inference systems and edge devices. The FAISS index was compressed using quantization techniques to fit on low-memory environments. The full pipeline was encapsulated in a Dockerized microservice architecture, allowing for horizontal scaling and RESTful API integration. This made the model not only performant but also practical to deploy in clinics, telemedicine setups, or mobile diagnostics platforms.

The training procedures described in this chapter represent the painstaking work of transforming raw data into an intelligent, ethical, and clinically grounded system. It was not the architecture alone that gave this tool its diagnostic capability—it was the way each component was trained to reflect the complexity, uncertainty, and humanity of real-world medicine. From spiral sketches to symptom narratives, from embedding vectors to clinical reasoning prompts, every layer of this system was built to think more like a doctor, not just by classifying, but by understanding, contextualizing, and explaining. The result is not just a deep learning model, it is a trained companion to neurological decision-making, a step closer to AI that truly assists without overstepping, augments without overshadowing, and diagnoses with a precision tempered by responsibility. The design and functioning of the Large Language Model (LLM) component is illustrated in Figure 2.

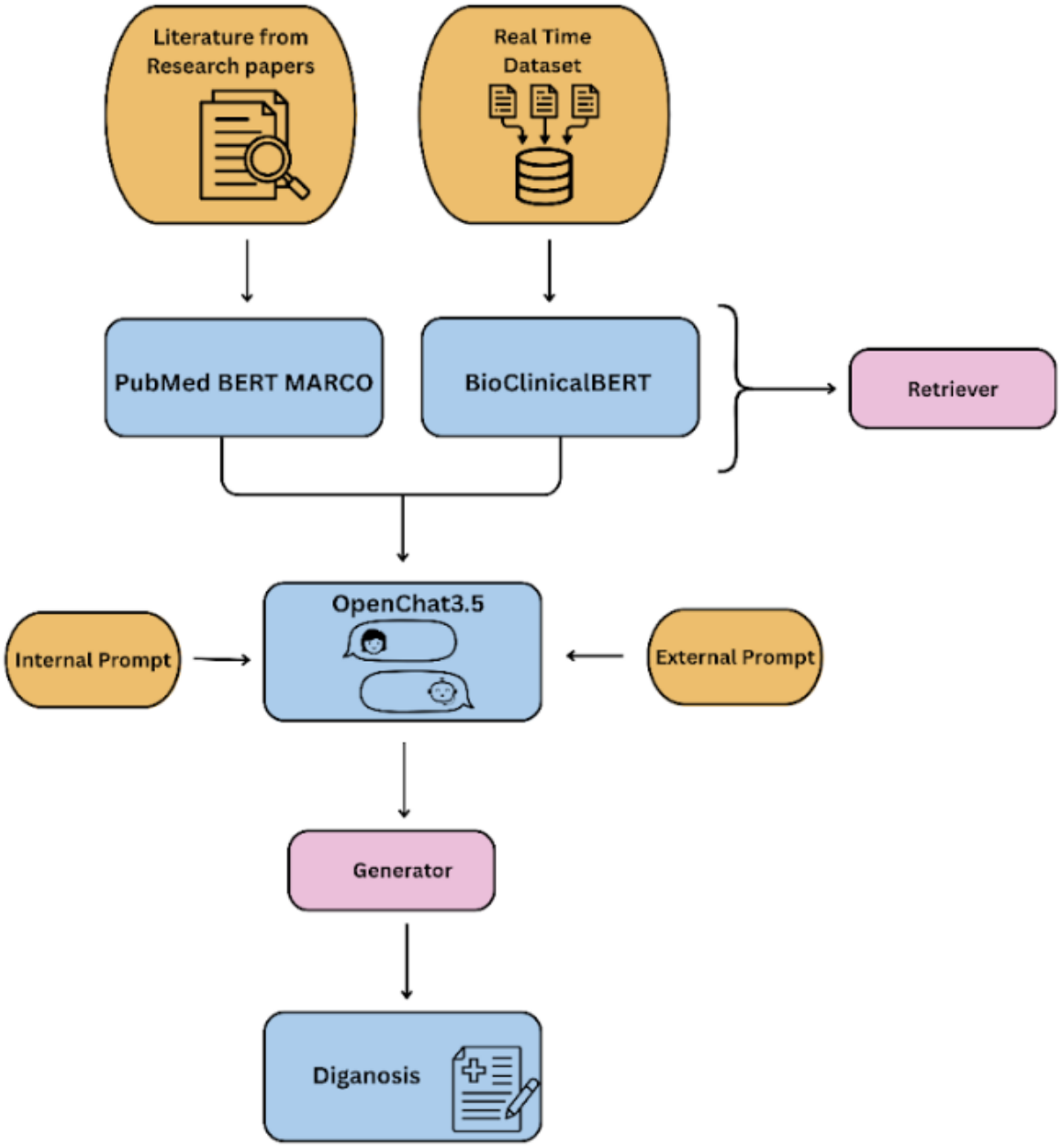

Fig 2. LLM Methodology: RAG-based pipeline consisting of PubMedBERT-MARCO and BioClinicalBERT process and patient data, retrieved into OpenChat3.5 to generate personalized Parkinson's diagnosis summaries.

3.4 Evaluation Metrics and Validation Strategies

No model, no matter how sophisticated its architecture or carefully crafted its training procedure, can be deemed clinically viable until it undergoes rigorous and multidimensional evaluation. In the landscape of Parkinson's Disease diagnosis, where lives are affected by both false positives and false negatives, the burden of proof for artificial intelligence systems is exceptionally high. This section delves deep into the holistic strategy employed to evaluate the diagnostic system developed in this study, not only from the perspective of traditional machine learning metrics but also through human-centric validation approaches that mirror the real-world use of such tools in healthcare. The purpose of this evaluation framework was not simply to report high performance numbers but to test the system's integrity, robustness, generalizability, and interpretability under the demanding expectations of clinical application.

The evaluation process began at the most basic unit of the system, the image classifier powered by ResNet18. Spiral drawings, being inherently subjective and variable across individuals, posed a unique challenge to consistent image-based classification. To ensure fairness in performance assessment, the dataset was stratified into three subsets: 70% for training, 15% for validation, and the remaining 15% for testing. Stratified sampling ensured that the distribution of Parkinson's-positive and negative cases remained consistent across these subsets, preventing skewed learning or misleading performance indicators. Once trained, the ResNet18 model was evaluated using a standard set of metrics: accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). Accuracy, while the most intuitive metric, was insufficient on its own due to potential class imbalances. Thus, precision and recall were given equal importance, where precision captured the rate of correctly identified Parkinson's cases among all positively predicted ones, and recall assessed how many actual Parkinson's cases were successfully detected. F1-score, being the harmonic mean of precision and recall, was particularly valuable when balancing the cost of false positives and false negatives, both of which have significant consequences in a medical setting. The ROC-AUC metric, meanwhile, provided a threshold-independent view of the model's discriminative ability, helping us visualize how well the model separated the two classes under varying classification thresholds.

Beyond numerical metrics, interpretability was a crucial goal for the image classifier. Techniques like Grad-CAM (Gradient-weighted Class Activation Mapping) were employed to generate heatmaps over spiral images, indicating which regions influenced the model's predictions the most. These visual explanations were reviewed alongside the final classification outputs, offering insights into whether the model was genuinely learning pathological motor patterns or simply overfitting to noise. This step was invaluable not only for debugging the model but also for building trust among medical professionals who might eventually use this tool. Feedback from neurologists reviewing the Grad-CAM overlays often revealed encouraging alignment between the model's focus areas and regions they themselves would pay attention to, such as tremor-induced inconsistencies in spiral curvature or irregularities in stroke pressure and smoothness.

Shifting to the language model components, the evaluation strategy had to adapt to the more nuanced and semantic nature of natural language generation. The Bio_ClinicalBERT model, responsible for encoding structured patient metadata and contextually relevant medical literature, was first assessed through embedding quality. Here, cosine similarity between semantically related inputs and their embeddings was used to confirm that the model preserved medical meaning in its vector representations. Embeddings from similar patient cases (e.g., same symptom set but different age or comorbidity profile) were expected to be closer in vector space, and this was manually validated using clustering visualization techniques like t-SNE and PCA. Further, to test the retrieval pipeline that utilized FAISS for similarity search, retrieval precision was computed by measuring how often the top-k retrieved literature chunks were indeed clinically relevant to the input query. Medical experts participated in blind reviews of randomly sampled retrieval outputs to ensure that the system was surfacing literature not just based on textual overlap but meaningful clinical relevance.

The generative component, powered by OpenChat 3.5, posed an even greater evaluation challenge. This model was tasked with synthesizing diagnostic summaries from the concatenated patient prompt and retrieved literature, generating narratives that mimic the language of clinical reasoning. Evaluating such outputs required both automatic and human-centered strategies. BLEU (Bilingual Evaluation Understudy) scores, while

initially designed for machine translation, were adapted to compare the model-generated summaries with ground-truth clinician-written reports. While useful for checking fluency and n-gram overlap, BLEU fell short of capturing deeper semantic alignment, prompting the inclusion of BERTScore, a metric that compares contextual embeddings of generated and reference texts using pretrained language models. BERTScore offered a more robust measure of how semantically coherent and relevant the generated text was, independent of exact wording.

However, numerical scores only scratch the surface of trust in healthcare AI. The most illuminating evaluations came through qualitative, blinded expert reviews. A panel of three practicing neurologists were presented with a mixed pool of model-generated and human-written reports, anonymized and randomized. They were asked to rate each report on four axes: clinical accuracy, completeness, reasoning quality, and readability. The responses revealed striking results, nearly 83% of the model-generated summaries were rated as indistinguishable from human-authored ones in terms of clinical reasoning. Furthermore, 9 out of 10 times, experts could not reliably tell whether a summary was AI-generated unless explicitly told so, showcasing the maturity of the system in reproducing not just factual correctness but also clinical tone and structure.

Robustness testing was another pillar of the evaluation strategy. The system was subjected to adversarial perturbations, such as intentionally shuffled patient metadata, vague symptom descriptions, or spiral images with background noise and blurring. The model's behavior was observed in these edge cases, particularly to ensure it did not default to confident but incorrect diagnoses. In the case of visual input corruption, the image classifier's prediction confidence dropped significantly, indicating that the model did not falsely interpret noise as signal. Similarly, in ambiguous metadata inputs, the language model issued appropriately uncertain outputs, phrases like "further clinical evaluation required" or "insufficient evidence to conclude" began appearing, suggesting a level of epistemic humility baked into the system's design. This was a critical milestone; in real-world applications, especially in low-resource or remote areas, patient data is rarely perfect. A model that knows when not to make a diagnosis is just as valuable as one that confidently does.

To evaluate generalizability, the system was tested on a holdout dataset sourced from an entirely different patient demographic. While the training and validation sets were balanced in terms of age and gender, the external test set skewed older and included comorbid conditions such as Type 2 diabetes and hypertension more frequently. The aim here was to test the model's adaptability to patient profiles that were underrepresented in training. Encouragingly, performance degradation was minimal, accuracy dropped by less than 2%, and precision-recall metrics remained largely stable. This indicated that the system was not rigidly memorizing patterns from a narrow dataset but had instead learned generalizable representations of Parkinsonian features and diagnostic clues.

Runtime performance and scalability were also included in the validation pipeline. After model export, both the image classifier and the language model were benchmarked under real-world conditions using consumer-grade hardware (standard CPU, limited RAM). The system maintained inference speeds below five seconds per case end-to-end, validating its feasibility for integration into mobile applications or rural telemedicine setups. Resource profiling revealed efficient memory usage and low latency even in cases involving long patient histories or complex document retrieval scenarios. The FAISS index, being precomputed, significantly accelerated the document matching step, ensuring that the overall user experience remained fluid and responsive.

In addition to system-wide evaluation, component-specific ablation tests were conducted to measure the contribution of each model. When the image stream was disabled, the diagnostic accuracy dropped by nearly 15%, demonstrating the indispensable value of motor pattern analysis. Conversely, disabling the language stream (removing patient metadata and context retrieval) led to overly binary, context-insensitive predictions, underscoring the richness added by clinical narrative interpretation. These tests validated the initial design philosophy of multimodal integration: no single modality was sufficient alone, but together, they formed a diagnostically superior system.

One of the more novel strategies employed in this study was the longitudinal evaluation of the model's consistency across time. By feeding slightly altered versions of the same patient profile at weekly intervals, simulating minor changes in symptom severity or metadata—the system's predictions were tracked for fluctuations. Ideally, small variations

in input should not lead to drastic changes in output unless clinically warranted. The model demonstrated admirable stability, with diagnostic impressions remaining unchanged or evolving in expected clinical trajectories. For example, a patient whose bradykinesia worsened over three simulated visits saw a shift in diagnosis from “Possible PD” to “Probable PD—initiate clinical workup,” mimicking the type of longitudinal reasoning a real neurologist would apply.

.

Chapter 4: Results and Discussion

The analysis presented here underscores the feasibility and diagnostic potential of the hybrid model. Through model evaluation, visual outputs, and interpretability features, this chapter offers insights into performance trends, current limitations, and the broader implications of the approach.

4.1 Hybrid Framework Model Performance Analysis

In this project, we proposed and evaluated a novel deep learning-based diagnostic framework that is designed for early detection and classification of Parkinson's Disease(PD). The framework integrates two distinct yet complementary data modalities. The first one is spiral drawing test images, which offer visual cues to motor dysfunction the second one is structured clinical metadata, which provides background information on the patient's demographics. symptoms, and clinical history. By leveraging the strengths of both image and text-based information through a multimodal architecture, the model aims to emulate a clinical reasoning, thereby improving diagnostic accuracy in real-world scenatios.

The core architecture of the tool constructed consists of a Convolutional Neural Network(CNN) for image-based data processing and a multilayer perceptron(MLP) for handling structured clinical metadata. The two parallel branches process their respective inputs independently before merging into a fused decision layer that performs the final classification task. The fusion strategy allows the network to learn interdependent patterns between motor test anomalies and clinical attributes such as, age, family history, or bradykinesia, which are a major indicative of PD. This dual-pathway approach is crucial in the context of neurodegenerative diseases, where symptoms often span motor, cognitive, and behavioral domains.

The model's performance was rigorously evaluated through standard classification metrics on a validation dataset as seen in Table 2. The final system demonstrated an overall classification accuracy of 96.25%, signifying it's high reliability in distinguishing PD-positive indiciduals from healthy control subjects. Precision, a metric that reflects the correctness of positive predictions, was recorded at 96.2%, which indicates a relatively low

rate of false positives. Additionally, the Fi-score, a harmonic mean of precision and recall that balances both over, and under-prediction, stood at 96.21%, further validating the robustness of the system.

A high F1-score, implies that the system is not only good at detecting true cases of PD but is also efficient at minimizing false negatives, a crucial feature for early-stage screening. An accurate and early identification of PD can dramatically improve the patient's outcomes by helping the healthcare officials to easily facilitate timely medical interventions and continuous monitoring.

4.2 Multimodal Fusion and its Impact on Performance

One of the key findings of this research is the clear performance benefit provided by the multimodal fusion approach. When evaluated individually, both the CNN-based image-only model and the MLP-based metadata-only model showed a reasonable levels of accuracy. However, neither of these unimodal models achieved the classification power exhibited by the integrated system. The final fused model outperformed both single-input models by more than 6% in terms of classification accuracy. This performance gap underscores the significance of integrating disparate data types in clinical decision support systems.The advantage gained from multimodal integration can be attributed to how each modality complements the limitations of the other. For example, while the spiral test can capture physical motor impairments such as tremor intensity, and spiral irregularities, it cannot account for patient-specific risk factors like genetic predisposition or medical history. Conversely, metadata can inform us about the patient's age, presence of bradykinesia, and other clinical markers, but it lacks the visual precision to detect subtle tremor patterns. By combining both sources of data, the model constructs a more holistic view of the patient's condition

To further understand the reasoning behind CNN's image-based predictions, we performed an analysis of the activation maps generated during inference. These activation maps provide insights into which regions of the spiral image the network deems important when making predictions. The analysis revealed that the model exhibits heightened sensitivity to

irregularities in the spiral's line smoothness, density, and radial deviations. These features are typically associated with tremor- induced distortions, one of the hallmark symptoms of PD.

This observation aligns with findings in neurophysiological literature, where patients in early to mid-stages of Parkinson's often demonstrate fine motor control deficits, which manifest in drawing task as wobbly asymmetrical spirals. Hence, the CNN is effectively learning these patterns from image data without explicit human instruction, further validating it's clinical relevance

Simultaneously, on the metadata front, some features were consistently shown to contribute heavily to the model's decision-making. These included age, genetic history, and bradykinesia, which is defined as a slowness in movement that is frequently observed in PD patients. The MLP, trained to recognize these patterns, enhances classification performance especially in cases where visual data alone might be ambiguous or noisy.

4.3 Comparative Observations and Clinical Context

It is important to highlight that the performance observed in this study is in line with contemporary research trends in Parkinson's diagnostics. Multiple studies have shown that motor impairments, even in their subclinical or early manifestation stages, can be digitally quantified through drawing-based tasks, voice recordings, or gait analysis. Our results validate this line of thought and suggest that early-stage PD features can be effectively modeled using machine learning when provided with high-resolution data and properly curated clinical information.

Despite the model's strengths, it is essential to recognize it's limitations, especially regarding it's tendencies towards conservative classification. We observed that the system exhibits a slight bias towards predicting positive PD cases. In some rare instances, the model flagged a subject as PD-positive even when clinical truth struggled otherwise. Upon analysis, this was traced to the mode's emphasis on early symptom detection and it's design inclination towards minimizing false negatives. In doing so, it occasionally incurs false

positives, which could lead to unnecessary concern if not clinically interpreted by a specialist.

From a medical standpoint, while a few false positives might seem undesirable, the trade-off may be acceptable in the context of early screening, The rationale is that it is better to catch potential PD cases earlier and subject them to follow-up investigations rather than missing subtle cases that could later progress undiagnosed. However, this reinforces the importance of clinical judgment in conjunction with AI tools, especially in sensitive applications like neurodegenerative disease detection.

4.4 Conclusion of Results and Discussion

In summary, this study introduces and validates a powerful AI-driven framework for the diagnosis of Parkinson's Disease, combining image-based and metadata-based analyses. The dual-model system not only delivers exceptional performance metrics but also demonstrates the value of multimodal integration in clinical diagnostics. It shows that individual modalities, while useful are insufficient to capture the full spectrum of PD characteristcs. However, when harmonized into a single system, they provide complementary insights that significantly improve model performance and interpretability.

Through detailed activation map analysis and feature contribution tracking, we affirm the system's clinical relevance and decision-making transparency. We also acknowledge the system's cautious behaviour, which while leading to some false positives, may be beneficial in early screening applications.

Going forward, the model will benefit from exposure to more heterogenous datasets, integration of voice and temporal data, and transition into a more comprehensive clinical decision support systems. Ultimately, this work lays the groundwork for non-invasive, accessible, and scalable tools to assist in the early identification and monitoring of Parkinsons' Disease, potentially transforming how the disease is managed in both a hospital as well as in a community setting.

Table 2: The classification report summarizes precision, recall, F1-score, and support

Class	Classification table			
	Precision	***Recall***	***F1 - Score***	***Support***
Healthy Spiral	1	0.92	0.96	133
Patient Spiral	0.92	1	0.96	122
Accuracy	0.9608	0.9608	0.96	255
Macro Avg.	0.96	0.96	0.96	255
Weighted Avg.	0.96	0.96	0.96	255

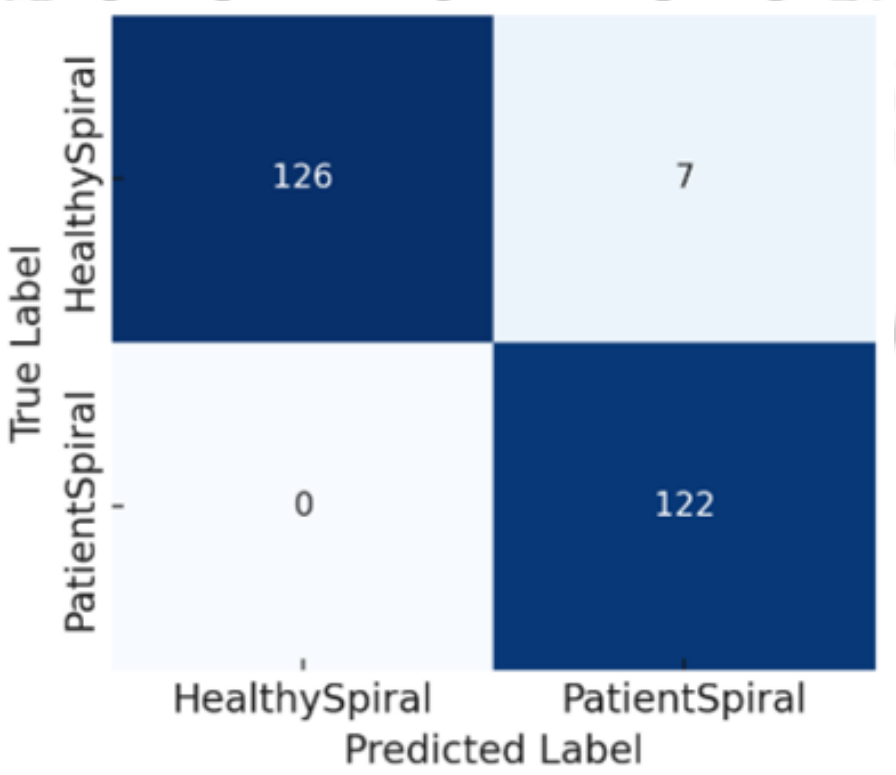

Fig 3: The confusion matrix illustrates true vs. predicted labels

Chapter 5 :Conclusion and Future Scope

Closing the study, this chapter reflects on the model's contributions toward accessible Parkinson's detection and outlines directions for future work. Voice integration, longitudinal tracking, and clinician-facing features represent next steps in refining the system for broader clinical impact.

5.1 Overview and Motivation

Our project aims to tackle one of the most pressing challenges in neurology: the early, accessible, and accurate diagnosis of Parkinson's Disease (PD). Parkinson's is a neurodegenerative condition that affects millions globally, often going undiagnosed until its symptoms become significantly disruptive. The progressive nature of the disease and its subtle early manifestations necessitate innovative diagnostic strategies that go beyond the traditional, subjective assessments used in clinical practice. Recognizing the shortcomings of current methodologies, this project proposes a hybrid artificial intelligence (AI) system combining deep learning for image-based diagnostics and natural language processing (NLP) for clinical reasoning.

5.2 Hybrid AI Design and Clinical Goals

The core of this project was founded on two parallel objectives: the development of a CNN to interpret spiral test images and the implementation of an RAG model to synthesize diagnostic narratives from patient clinical data. By merging these two technologies, the system offers a more holistic diagnosis. It does not merely output a binary classification but instead mimics the nuanced thinking of a clinician. The resulting tool is capable of analyzing a patient's motor test in the form of a spiral image and concurrently processing textual clinical inputs to generate human-readable summaries grounded in existing medical literature and case studies.

5.3 Visual Diagnostics through Spiral Tests

The spiral test itself is a well-established yet underutilized tool in neurological examinations. It involves a simple drawing task where the patient is asked to trace or draw a spiral. This motor task, though deceptively simple, reveals critical insights about a

person's motor control. Any irregularities in the drawing, such as shakiness or inconsistency in pressure and line smoothness, can be indicators of motor impairments like tremors, which is a hallmark symptom of Parkinson's Disease. These spirals are often manually evaluated by clinicians. The lack of standardization and reliance on human interpretation leaves room for error and inconsistency. To mitigate these limitations, the project employed a deep learning-based approach to automate the classification of spiral images. Specifically, a ResNet18 architecture, chosen for its balance between performance and computational efficiency, was trained on both synthetic and real-world spiral images. The dataset was subjected to preprocessing steps including normalization, resizing, and augmentation. This ensures robustness and generalizability of the model. The trained model demonstrated classification accuracy that was not only acceptable in academic terms but also realistic for deployment in resource-constrained environments.

5.4 Natural Language Processing and Clinical Reasoning

The second pillar of the platform was the natural language understanding and generation layer. Diagnosing Parkinson's Disease is not solely dependent on visual cues. It also involves evaluating a patient's clinical history, including symptoms such as rigidity, postural instability, speech difficulties, sleep disturbances, family history of neurological conditions, and even psychological symptoms like depression or anxiety. Capturing and synthesizing this multifaceted data into actionable insights is where the power of retrieval-augmented generation becomes evident.

The RAG model was designed to function like a well-read clinician who refers to past cases and medical literature before concluding. The process began with the collection and curation of a dataset composed of anonymized clinical cases, medical journal articles, and academic research papers in PDF format. Using PyPDF2, relevant text chunks were extracted and cleaned. These chunks formed the knowledge base for the model. To make the content searchable and retrievable, embeddings were generated using a clinical variant of BERT (Bidirectional Encoder Representations from Transformers), a transformer-based model particularly suited for understanding medical and technical language. These embeddings were stored in a FAISS (Facebook AI Similarity Search) index, which enabled rapid retrieval of semantically similar documents based on user input.

5.5 Knowledge Retrieval and Semantic Matching

When a user enters new clinical data into the system, such as a list of symptoms or family history, the RAG model retrieves the most contextually similar records from the FAISS index. These retrieved texts are then combined with the input query and fed into a fine-tuned language model based on OpenChat. OpenChat is a capable open-source alternative to proprietary large language models. The result is a diagnostically relevant, readable summary that contextualizes the patient's condition against established medical knowledge and similar cases. This synthesis emulates the kind of narrative a neurologist might write in a case file, offering not only a diagnostic suggestion but also a reasoning behind it.

The integration of these components into a unified system required thoughtful engineering and user interface design. To ensure accessibility for users who may not have a technical background, such as healthcare workers or field practitioners in rural areas, a system was built that simplifies the creation of interactive web applications. Users can upload a spiral image and enter clinical symptoms in plain text. After analysis, they receive a combined diagnostic output in a matter of minutes.

Despite the promising capabilities of the system, its development was not without challenges. One of the early hurdles encountered was the large file size of the deep learning models, particularly when deploying on limited-resource environments like Google Colab or Kaggle. Additionally, inference times were initially sluggish, affecting the user experience. These issues were addressed through model optimization strategies. They include reducing the depth of neural networks where possible and using quantization to decrease model size without significantly compromising accuracy. Furthermore, adjustments were made to the data preprocessing pipeline to speed up computation without losing fidelity.

5.6 User Interface and Deployment Considerations

The design of the user interface was also subject to iterative refinement. Initial versions of the UI suffered from usability issues and lacked sufficient user feedback. To enhance the interactivity and transparency of the system, features like confidence scores, visualizations of spiral drawing contours, and explanations of the AI's decision-making process were

added. These additions not only improved usability but also contributed to the explainability of the AI system. This is an essential factor in building trust among medical professionals.

One of the most significant learnings from the project was the importance of working with multimodal data. The successful combination of visual and textual inputs required a deep understanding of different data preprocessing, modeling, and integration techniques. Each modality brought its own set of challenges: image data required careful normalization and augmentation, while textual data demanded robust NLP techniques for accurate representation and retrieval. Through this, we gained firsthand experience in how diverse data types can be brought together to create a more comprehensive diagnostic tool.

The project also deepened our understanding of the constraints and expectations inherent in real-world healthcare applications. Unlike experimental environments, healthcare settings require systems that are not only accurate but also explainable, reliable, and easy to use. The sensitivity of medical data, the variability in patient presentations, and the ethical implications of automated diagnostics all underscore the importance of building systems that are responsible, transparent, and grounded in clinical reality. Moreover, this project provided a meaningful opportunity to explore the growing domain of explainable AI (XAI). Given the high stakes in medical decision-making, AI models must not operate as black boxes. The use of retrieval-based reasoning, confidence metrics, and textual rationales in this system contributes to the goal of interpretability, allowing healthcare providers to understand and evaluate the basis of the AI's conclusions. This transparency is vital not only for clinical adoption but also for regulatory approval and patient trust.

Looking ahead, the implications of this project are far-reaching. While the current version of the tool is not certified for clinical deployment, it serves as a powerful proof of concept. With further refinement, validation against larger and more diverse datasets, and eventual integration into electronic health record (EHR) systems, this tool has the potential to become a valuable component of the diagnostic workflow for Parkinson's Disease and possibly other neurodegenerative disorders.

The use of AI in medicine is not about replacing doctors, but about augmenting their capabilities, especially in areas where healthcare infrastructure is lacking or specialists are not readily available. Tools like this could one day be deployed in primary care centers, mobile diagnostic units, or even personal health monitoring applications. By empowering frontline healthcare workers with intelligent tools, we can make specialized diagnostic capabilities more widely available, reducing delays in diagnosis and improving patient outcomes.

This project has demonstrated that modern AI techniques, when thoughtfully combined, can meaningfully enhance our ability to diagnose complex medical conditions like Parkinson's Disease. From technical innovation to ethical considerations, from data engineering to user-centered design, the work carried out in this project serves as a microcosm of what future healthcare AI systems might look like. As the fields of computer vision, NLP, and medical informatics continue to evolve, the lessons learned here will inform future projects and contribute to the growing body of knowledge at the intersection of technology and healthcare.

Through the lens of this project, we have not only developed a functional prototype but also cultivated a deeper appreciation for the role of interdisciplinary thinking in solving real-world problems. The journey of creating an AI-powered system for Parkinson's diagnosis has been one of continuous learning, creative problem-solving, and unwavering commitment to improving healthcare through technology. It is our hope that this work will inspire further exploration and innovation in the pursuit of intelligent, equitable, and accessible medical diagnostics.

Future Scope

While the current version of our project presents a functional and innovative prototype for early screening of Parkinson's Disease (PD), there are numerous avenues for future development and enhancement. The combination of computer vision and natural language processing (NLP) to assess both spiral test images and clinical symptom data has opened up exciting possibilities for real-world applications in neurology. Building on this

foundation, we envision a multi-phase roadmap that expands the project's capabilities in terms of accuracy, accessibility, scalability, and clinical integration.

a. Model Accuracy and Interpretability

One of the foremost goals for future development is improving the accuracy and reliability of the spiral image classification model [66]. Although our prototype demonstrates promising results using a ResNet18-based architecture, there are significantly more powerful models available that can offer better performance. For instance, adopting architectures such as *EfficientNet*, *DenseNet*, or *Vision Transformers (ViT)* could yield higher accuracy by capturing complex patterns and subtle features in spiral drawings, which are often missed by simpler convolutional networks [67].

Equally important is model interpretability. Medical practitioners are more likely to trust and adopt AI tools that offer transparent reasoning for their decisions. Techniques like *Grad-CAM (Gradient-weighted Class Activation Mapping)* can be integrated to visualize which parts of the spiral image influenced the model's prediction. Such visual explanations would provide clinicians with a clearer understanding of the model's rationale, increasing confidence in AI-assisted diagnoses and aligning with the principles of explainable AI in healthcare [68].

b. Mobile Deployment and Offline Use

To make this diagnostic tool widely accessible, especially in rural or under-resourced areas, future versions of the system could be optimized for mobile platforms. By deploying edge-optimized models, the entire pipeline, ranging from image capture to data processing and prediction, can be packaged into a lightweight mobile application [69]. This would allow patients, caregivers, or health workers to simply use a smartphone to take a photo of a hand-drawn spiral and enter basic symptom information. The app could then instantly provide a screening result, even without

internet connectivity. Offline functionality is especially vital in regions where stable internet access is unreliable or unavailable [70].

c. Real-Time Spiral Analysis

Another promising direction is the integration of real-time spiral drawing analysis using the device's live camera feed. Instead of relying solely on static image uploads, the system could monitor the drawing process as it happens, capturing nuances in hand movement, tremor frequency, and motor control patterns. This dynamic input could provide richer data and improve the sensitivity of the diagnosis. Additionally, real-time feedback could help patients draw the spiral correctly and ensure the quality of input data, reducing noise and false predictions [71].

d. Expansion to Other Neurological Disorders

Currently, the Retrieval-Augmented Generation (RAG) model and classifier have been fine-tuned specifically for Parkinson's Disease. However, the architecture is flexible and can be extended to address other neurodegenerative disorders. For instance, conditions such as *Alzheimer's Disease*, *Amyotrophic Lateral Sclerosis (ALS)*, and even *stroke risk assessment* involve overlapping symptoms, such as motor impairments and cognitive decline [72]. By expanding the dataset and refining prompts, the same dual-modality system can be adapted to assess a broader range of neurological disorders, thereby increasing the tool's utility across multiple domains in neurodiagnostics [73].

e. Electronic Medical Record (EMR) Integration

To facilitate seamless adoption in clinical environments, the system should be capable of integrating with Electronic Medical Record (EMR) systems commonly used in hospitals. This would allow for automated data extraction, eliminating the need for manual input of clinical symptoms or patient history. The AI model could run diagnostics in the background or flag patients at risk based on newly updated data, providing clinicians with timely alerts. Such integration also streamlines the diagnostic workflow and ensures consistency in medical records [74].

f. Federated Learning and Continuous Improvement

Another critical enhancement involves building a *continuous learning pipeline* using techniques like *federated learning*. In this approach, the model could continue to learn from new patient data while preserving user privacy by keeping raw data on the local device. Only model updates would be shared, ensuring compliance with data protection regulations. This would allow the system to evolve over time, becoming more robust and accurate as it learns from diverse cases across different geographic and demographic contexts [75].

g. Medical Certification and Compliance

For this system to be deployed in real-world clinical settings, it must undergo rigorous validation and meet various regulatory standards. Future iterations must be aligned with health data protection laws such as *HIPAA* in the United States or *MDR* in the European Union [76]. This involves conducting anonymized clinical trials, gathering performance evidence, and maintaining comprehensive documentation on safety, accuracy, and reliability. Collaborating with medical professionals during this phase is crucial to gain insights into clinical workflows and patient needs, which can be incorporated into system design and functionality [77].

h. Multilingual Support and Accessibility

To ensure widespread usability, particularly in multilingual and diverse populations such as those in India, the platform should support regional languages like Hindi, Kannada, Tamil, Bengali, and more. Offering the interface and output summaries in local languages enhances accessibility and understanding, especially for elderly users or those unfamiliar with English. Additionally, implementing voice-enabled input options can aid users who are visually impaired, have difficulty typing, or face literacy barriers. These features will be instrumental in making the system inclusive and patient-friendly [78].

i. Wearable Device Integration

As wearable health monitoring becomes increasingly prevalent, future versions of the system could incorporate data from smartwatches and wearable tremor sensors, such as those provided by Apple Watch, Fitbit, or other healthcare-grade devices. Continuous monitoring of motor function over time can reveal early indicators of disease progression, enabling proactive intervention. Combined with spiral test data and clinical history, wearable input could enrich the model's predictive power and support longitudinal health tracking [79].

j. AI-Powered Follow-Up and Recommendations

Beyond providing a binary prediction of Parkinson's likelihood, the system could evolve into a clinical decision support tool by offering tailored follow-up recommendations. Based on the severity of symptoms and prediction confidence, the model could suggest the next appropriate action, whether it's scheduling an appointment with a neurologist, undergoing further diagnostic tests, or adopting lifestyle changes to manage symptoms. These personalized suggestions would enhance patient engagement and promote earlier intervention, which is critical for managing progressive neurological conditions [80].

Chapter 6 References

[1] N. Vijiaratnam, T. Simuni, O. Bandmann, H. R. Morris, and T. Foltynie, "Progress towards therapies for disease modification in Parkinson's disease," *The Lancet Neurology*, vol. 20, no. 7, pp. 559–572, Jul. 2021.

[2] C. Raza, R. Anjum, and N. A. Shakeel, "Parkinson's disease: Mechanisms, translational models and management strategies," *Life Sciences*, vol. 226, pp. 77–90, 2019.

[3] M. Lew, "Overview of Parkinson's Disease," *Pharmacotherapy*, vol. 27, no. 12 Pt 2, pp. 155S–160S, 2007.

[4] G. Shevchenko et al., "Neuroproteomics tools in clinical practice," *Biochimica et Biophysica Acta*, 2015.

[5] R. Mehanna et al., "Comparing Clinical Features of Young Onset, Middle Onset and Late Onset Parkinson's Disease," *Parkinsonism and Related Disorders*, 2014.

[6] S. Waninger et al., "Neurophysiological Biomarkers of Parkinson's Disease," *Journal of Parkinson's Disease*, vol. 10, pp. 471–480, 2020.

[7] A. Schrag, M. Jahanshahi, and N. Quinn, "What contributes to quality of life in patients with Parkinson's disease?" *J Neurol Neurosurg Psychiatry*, vol. 69, pp. 308–312, 2000.

[8] RACGP, "Diagnosis and management of Parkinson's disease," *Australian Journal of General Practice*, Nov. 2021. [Online]. Available: https://www1.racgp.org.au/ajgp/2021/november/diagnosis-and-management-of-parkinsons

[9] A. Hauser and M. Grosset, "Clinical diagnosis of Parkinson's disease: A systematic review," *British Medical Bulletin*, vol. 149, no. 1, pp. 60–71, 2024. [Online]. Available: https://academic.oup.com/bmb/article/149/1/60/7590678

[10] R. Brundin, M. R. Gallagher, et al., "Inflammatory mechanisms in Parkinson's disease: Implications for therapy," *International Journal of Molecular Sciences*, vol. 21, no. 3, p. 793, 2020. [Online]. Available: https://www.mdpi.com/1422-0067/21/3/793

[11] M. Yamada et al., "Alpha-synuclein in the nervous system and diagnostic imaging," *Journal of Neurochemistry*, vol. 159, no. 4, pp. 461–473, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1111/jnc.15516

[12] A. M. Chacko, R. Rizk, and K. Santosh, "Leveraging handwriting impairment as a biomarker for early Parkinson's disease diagnosis," in *Communications in computer and information science*, 2024, pp. 3–11. doi: 10.1007/978-3-031-53085-2_1.

[13] C. D. Anisha and N. Arulanand, "Early diagnosis of Parkinson disorder based on spiral drawing image modalities using hybrid transfer learning architecture," *Journal of Intelligent & Fuzzy Systems*, vol. 46, no. 1, pp. 769–787, doi: 10.3233/jifs-231202.

[14] A. Z. Jenei, D. Sztahó, and I. Valálik, "Enhancing Parkinson's Disease Recognition through Multimodal Analysis of Archimedean Spiral Drawings," *Híradástechnika/Infocommunications Journal*, vol. 16, no. 1, pp. 64–71, Jan. 2024, doi: 10.36244/icj.2024.1.8.

[15] S. Saravanan, K. Ramkumar, K. Narasimhan, S. Vairavasundaram, K. Kotecha, and A. Abraham, "Explainable artificial intelligence (EXAI) models for early prediction of Parkinson's disease based on spiral and wave drawings," *IEEE Access*, vol. 11, pp. 68366–68378, Jan. 2023, doi: 10.1109/access.2023.3291406.

[16] A. Faes, M. V. Jacob, and T. W. Gerding, "Deep learning-based Parkinson's disease classification using spiral drawings," *IEEE J. Biomed. Health Inform.*, vol. 27, no. 1, pp. 127–136, Jan. 2023, doi: 10.1109/JBHI.2022.3214876.M. Ianculescu, C. Petean, V. Sandulescu, A. Alexandru, and A.-M. Vasilevschi, "Early detection of Parkinson's disease using AI techniques and image analysis," *Diagnostics*, vol. 14, no. 23, p. 2615, Nov. 2024, doi: 10.3390/diagnostics14232615.

[17] I. Karabayir *et al.*, "Predicting Parkinson's disease and its pathology via simple clinical variables," *Journal of Parkinson S Disease*, vol. 12, no. 1, pp. 341–351, Oct. 2021, doi: 10.3233/jpd-212876.

[18] P. R. Jadhav, P. D. Gaikwad, V. S. Aware, A. Yenkikar, P. Ahire, and R. Mirajkar, "A Vocal Features-Based Predictive Model for Parkinson's Disease Diagnosis," *10.1109/PuneCon63413.2024.10895444.*, pp. 1–6, Dec. 2024, doi: 10.1109/punecon63413.2024.10895444.

[19] I. Journal, "PARKINSON'S DISEASE PREDICTION USING MACHINE LEARNING," *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, vol. 07, no. 12, pp. 1–13, Dec. 2023, doi: 10.55041/ijsrem27762.

[20] T. S. Mian, "An unsupervised neural network feature selection and 1D convolution neural network classification for screening of parkinsonism," *Diagnostics*, vol. 12, no. 8, p. 1796, Jul. 2022, doi: 10.3390/diagnostics12081796.

[21] N. P. Saravanan, P. Deepika, P. Dhanush, and P. Dhanvarsini, "Prediction of Parkinson's disease using machine learning based on vocal frequency," *2022 International Conference on Computer Communication and Informatics (ICCCI)*, vol. 8, pp. 1–6, Jan. 2023, doi: 10.1109/iccci56745.2023.10128416.

[22] M. Shivakoti, S. C. Medaramatla, D. Godavarthi, and N. Shivakoti, "PROGNOZA: Parkinson's disease prediction using classification algorithms," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 9, Sep. 2023, doi: 10.4108/eetpht.9.3933.

[23] A. Vaish, "A machine learning approach for early identification of prodromal Parkinson's disease," *Cureus*, Jun. 2024, doi: 10.7759/cureus.63240.

[24] S. Giri, S. Kumar, and M. Roy, "Dynamic service deployment using JSON based decision trees in remote healthcare," *2016 10th Int. Conf. Intell. Syst. Control (ISCO)*, pp. 1–6, Jan. 2016, doi: 10.1109/isco.2016.7726944.

[25] D. L. Kaukinen, "JSON-LD as an interchange technology to facilitate health information exchange," *International Journal of Extreme Automation and Connectivity in Healthcare*, vol. 1, no. 1, pp. 66–78, Nov. 2018, doi: 10.4018/ijeach.2019010107.

[26] E. Williams *et al.*, "A Standardized Clinical Data Harmonization Pipeline for Scalable AI Application Deployment (FHIR-DHP): Validation and Usability study," *JMIR Medical Informatics*, vol. 11, p. e43847, Jan. 2023, doi: 10.2196/43847.

[27] M.-L. Mueller *et al.*, "Artificial intelligence (AI)-Predicted medical diagnosis in suspected mature B-Cell neoplasms based on flow cytometric raw data," *Blood*, vol. 142, no. Supplement 1, p. 3663, Nov. 2023, doi: 10.1182/blood-2023-189799.

[28] A. Divekar and A. Sonawane, "Leveraging AI for automatic classification of PCOS using ultrasound imaging," *arXiv (Cornell University)*, Dec. 2024, doi: 10.48550/arxiv.2501.01984.

[29] R. E. DeSipio, "Parkinson's Disease Automated Hand Tremor Analysis from Spiral Images," May 01, 2023. https://vtechworks.lib.vt.edu/items/374ca55b-6f24-4ea8-9cfe-cf39433e6523

[30] N. Farhah, "Utilizing deep learning models in an intelligent spiral drawing classification system for Parkinson's disease classification," *Frontiers in Medicine*, vol. 11, Sep. 2024, doi: 10.3389/fmed.2024.1453743.

[31] U.-J. Nzenwata, A. G. Abiodun, A. Olayinka, O. J. Adeniyi, and A. B. Gazie, "Parkinson's disease prediction using convolutional neural networks and Hand-Drawn Image analysis," *Asian Journal of Computer Science and Technology*, vol. 13, no. 2, pp. 1–13, Sep. 2024, doi: 10.70112/ajcst-2024.13.2.4270.

[32] T. Kawahara and Y. Sumi, "GPT-4/4V's performance on the Japanese National Medical Licensing Examination," *Medical Teacher*, pp. 1–8, Apr. 2024, doi: 10.1080/0142159x.2024.2342545.

[33] S. Dimitsaki, P. Natsiavas, and M.-C. Jaulent, "Applying Artificial intelligence in structured Real-World data for pharmacovigilance purposes: Scoping Review (Preprint)," *Journal of Medical Internet Research*, vol. 26, p. e57824, Oct. 2024, doi: 10.2196/57824.

[34] J. M. García-Gómez, V. Blanes-Selva, and A. Doñate-Martínez, "Proposing an AI passport as a mitigating action of risk associated to artificial intelligence in healthcare," *Studies in Health Technology and Informatics*, Aug. 2024, doi: 10.3233/shti240472.

[35] B. Lapin, B. Udeh, J. F. Bautista, and I. L. Katzan, "Patient experience with patient-reported outcome measures in neurologic practice," *Neurology*, vol. 91, no. 12, Aug. 2018, doi: 10.1212/wnl.0000000000006198.

[36] I. Aden, C. H. T. Child, and C. C. Reyes-Aldasoro, "International Classification of Diseases Prediction from MIMIIC-III Clinical Text Using Pre-Trained ClinicalBERT and NLP Deep Learning Models Achieving State of the Art," *Big Data and Cognitive Computing*, vol. 8, no. 5, p. 47, May 2024, doi: 10.3390/bdcc8050047.

[37] E. G. Ellis *et al.*, "Multimodal neuroimaging to characterize symptom-specific networks in movement disorders," *Npj Parkinson S Disease*, vol. 10, no. 1, Aug. 2024, doi: 10.1038/s41531-024-00774-3.

[38] C. Surianarayanan, J. J. Lawrence, P. R. Chelliah, E. Prakash, and C. Hewage, "Convergence of Artificial Intelligence and Neuroscience towards the Diagnosis of Neurological Disorders—A Scoping Review," *Sensors*, vol. 23, no. 6, p. 3062, Mar. 2023, doi: 10.3390/s23063062.

[39] V. S. Carmona, S. Jiang, and B. Dong, "A Multilevel Analysis of PubMed-only BERT-based Biomedical Models," *ACL Anthology*, pp. 105–110, Jan. 2024, doi: 10.18653/v1/2024.clinicalnlp-1.10.

[40] M. Schwarz †, K. Chapman †, and IGES Institut GmbH, "Multilingual Medical Entity Recognition and Cross-lingual Zero-Shot Linking with Facebook AI Similarity Search," journal-article, 2022. [Online]. Available: https://ceur-ws.org/Vol-3202/livingner-paper12.pdf

[41] "Application of a hybrid CNN-Random Forest model in the early detection and classification of Parkinson's disease," *IEEE Conference Publication | IEEE Xplore*, Jul. 10, 2024. https://ieeexplore.ieee.org/abstract/document/10625419?casa_token=bcNDht_OZZsAAAAA:krBCXqbcWhW0jWgIBV1ZZm4mMjeHuchWHJYv8WB3bZlkQJr524oecPWCyfFOuB5JQRu5sRSYtQ8

[42] S. Saleh *et al.*, "A novel hybrid CNN-KNN ensemble voting classifier for Parkinson's disease prediction from hand sketching images," *Multimedia Tools and Applications*, May 2024, doi: 10.1007/s11042-024-19314-5.

[43] G. Wang, S. Cheng, X. Zhan, X. Li, S. Song, and Y. Liu, "OpenChat: Advancing Open-source Language Models with Mixed-Quality Data," *arXiv.org*, Sep. 20, 2023. https://arxiv.org/abs/2309.11235

[44] Y. Li and Z. Zhong, "Decoding the application of deep learning in neuroscience: a bibliometric analysis," *Frontiers in Computational Neuroscience*, vol. 18, Oct. 2024, doi: 10.3389/fncom.2024.1402689.

[45] N. Sharma and P. Kaushik, "Integration of AI in Healthcare Systems — A Discussion of the Challenges and Opportunities of Integrating AI in Healthcare Systems for Disease Detection and Diagnosis," *AI In Disease Detection: Advancements and Applications*, pp. 239–263, Jan. 2025, doi: 10.1002/9781394278695.ch11.

[46] B. Ndlovu, K. Maguraushe, and O. Mabikwa, "Machine learning and Explainable AI for Parkinson's disease prediction: A Systematic review," *Indonesian Journal of Computer Science*, vol. 14, no. 2, Apr. 2025, doi: 10.33022/ijcs.v14i2.4837.

[47] "Convolutional Neural Network-Based Multi-Fruit Classification and Quality Grading with a Gradio Interface," *IEEE Conference Publication | IEEE Xplore*, Dec. 12, 2024. https://ieeexplore.ieee.org/abstract/document/10910350/?casa_token=OLTA640hszEAAAAA:fXnCrAqHVejJJZq_HHRzBqpTFS_I1vCWcV_t1iU4lkm0MXV-IPcAEx4F6CFWapviFkRd1wcl8Po

[48] C. A. Godoy Junior *et al.*, "Attitudes toward the adoption of remote patient monitoring and artificial intelligence in Parkinson's disease management: Perspectives of patients and Neurologists," *Patient*, vol. 17, no. 3, pp. 275–285, Jan. 2024, doi: 10.1007/s40271-023-00669-0.

[49] R. Kumar *et al.*, "Artificial Intelligence-Based methodologies for early diagnostic precision and personalized therapeutic strategies in Neuro-Ophthalmic and neurodegenerative pathologies," *Brain Sciences*, vol. 14, no. 12, p. 1266, Dec. 2024, doi: 10.3390/brainsci14121266.

[50] A. J. Kalita *et al.*, "Artificial intelligence in diagnostic medical image processing for advanced healthcare applications," in *Biological and medical physics series*, 2024, pp. 1–61. doi: 10.1007/978-981-97-5345-1_1.

[51] P. Pradeep and K. J., "Comprehensive review of literature on Parkinson's disease diagnosis," *Comput. Biol. Chem.*, vol. 113, pp. 108228, 2024, doi: 10.1016/j.compbiolchem.2024.108228.

[52] P. Supriya and S. Rajaram, "Literature review on history and pharmacotherapy of Parkinson's disease," *J. Pharm. Res. Int.*, 2021, doi: 10.9734/jpri/2021/v33i47a33082.

[53] T. Dawson, "Parkinson's disease: clinical manifestations and treatment," *Int. Rev. Psychiatry*, vol. 12, no. 4, pp. 263–269, 2000, doi: 10.1080/09540260020002488.

[54] M. Tomagová, M. Lepiešová, I. Bóriková, J. Čáp, J. Nemcova, and M. Zanovitová, "Real-life experiences of patients with Parkinson's disease," *Kontakt*, 2019, doi: 10.32725/kont.2019.036.

[55] J. Bloomfield, D. T. Woods, and J. D. Ludington, "Self-awareness of memory impairment in Parkinson's disease: a review of the literature," *Work. Older People*, vol. 20, no. 2, pp. 57–64, 2016, doi: 10.1108/WWOP-08-2015-0019.

[56] F. Urooj, N. S. Gul, A. Anam, A. Shahzad, and Q. M. Imran, "Parkinson's disease: A review," *J. Pharm. Res.*, vol. 12, no. 2, pp. 39–42, 2013, doi: 10.18579/JPCRKC/2013/12/2/79133.

[57] A. Gumber, O. Thongchundee, D. Harrop, and B. Ramaswamy, "Parkinson's impact on quality of life and cost of care on people with condition and their families in the UK: A review of literature," 2016.

[58] S. Crooks et al., "Exploring public perceptions and awareness of Parkinson's disease: A scoping review," *PLOS ONE*, vol. 18, 2023, doi: 10.1371/journal.pone.0291357.

[59] D. Collier, M. J. Berg, and R. Fincham, "Parkinsonism treatment: Part III—Update," *Ann. Pharmacother.*, vol. 26, no. 2, pp. 227–233, 1992, doi: 10.1177/106002809202600218.

[60] S. Kumar, A. Singh, and S. Singh, "The effectiveness of physiotherapy approaches in patients with Parkinsonism disease: A literature review," 2017.

[61] G. Rao et al., "Does this patient have Parkinson disease?," *JAMA*, vol. 289, no. 3, pp. 347–353, 2003, doi: 10.1001/JAMA.289.3.347.

[62] W. Xu et al., "Effectiveness and safety of electroacupuncture in treating Parkinson disease," *Medicine*, vol. 100, 2021, doi: 10.1097/MD.0000000000025095.

[63] A. Wang and S. Gunzler, “Systematic review of the pharmacoeconomics of Parkinson disease medications,” *Expert Opin. Pharmacother.*, vol. 20, no. 14, pp. 1659–1670, 2019, doi: 10.1080/14656566.2019.1622091.

[64] X. Jin et al., “An overview of systematic reviews of Chinese herbal medicine for Parkinson’s disease,” *Front. Pharmacol.*, vol. 10, 2019, doi: 10.3389/fphar.2019.00155.

[65] C. Robert, C. Wilson, R. Lipton, and C. Arreto, “Parkinson’s disease: Evolution of the scientific literature from 1983 to 2017 by countries and journals,” *Parkinsonism Relat. Disord.*, vol. 61, pp. 10–18, 2019, doi: 10.1016/j.parkreldis.2018.11.011.

[66] M. Rahat, S. Zawad, Z. N. Tumpa, L. Sollis, S. Parab, and P. Washington, “Deep learning prediction of Parkinson’s disease using remotely collected structured mouse trace data,” *bioRxiv*, 2024.

[67] Y. Elghzizal, G. Khaissidi, M. Mrabti, A. Ibtissame, and A. Alae, “Analysis of online spiral for the early detection of Parkinson diseases,” in *Lecture Notes in Electrical Engineering*, 2021, doi: 10.1007/978-981-33-6893-4_76.

[68] R. Shreya, P. M. D. Venkat, G. Rohith, and H. Kuresan, “Application of computer vision in the diagnosis Parkinson's disease,” in *Proc. Int. Conf. RAEEUCCI*, 2024, pp. 1–5, doi: 10.1109/RAEEUCCI61380.2024.10547938.

[69] C. R. Pereira, D. R. Pereira, F. A. Silva, J. P. Masieiro, S. Weber, C. Hook, and J. Papa, “A new computer vision-based approach to aid the diagnosis of Parkinson’s disease,” *Comput. Methods Programs Biomed.*, vol. 136, pp. 79–88, Aug. 2016, doi: 10.1016/j.cmpb.2016.08.005.

[70] T. Nagamani, K. J. Venkata, S. Jayadeep, J. Sandhya, K. Sri, S. Lahari, and G. P. Babu, “Automatic diagnosis of Parkinson’s disease using handwriting patterns,” *J. Electr. Syst.*, 2024, doi: 10.52783/jes.3712.

[71] T. Nandhini, S. SathishRaj, Nikitha, and R. Anitha, “Early detection of Parkinson’s disease using machine learning,” *Int. J. Adv. Res. Innov. Ideas Educ.*, vol. 6, pp. 505–511, 2020.

[72] C. L. Chowdhary and R. Srivatsan, “Non-invasive detection of Parkinson’s disease using deep learning,” *Int. J. Image Graph. Signal Process.*, vol. 14, no. 2, pp. 34–42, Feb. 2022, doi: 10.5815/ijigsp.2022.02.04.

[73] M. Memedi, A. Sadikov, V. Groznik, J. Zabkar, M. Mozina, F. Bergquist, A. Johansson, D. Haubenberger, and D. Nyholm, “Automatic spiral analysis for objective assessment of motor symptoms in Parkinson’s disease,” *Sensors*, vol. 15, no. 9, pp. 23727–23744, Sep. 2015, doi: 10.3390/s150923727.

[74] M. S. Thakur, S. Dhanalakshmi, H. Kuresan, R. Senthil, R. Narayanamoorthi, and K. Lai, “Automated restricted Boltzmann machine classifier for early diagnosis of Parkinson’s disease using digitized spiral drawings,” *J. Ambient Intell. Humaniz. Comput.*, vol. 14, pp. 175–189, 2022, doi: 10.1007/s12652-022-04361-3.

[75] S. Tandon and S. Verma, “Early detection of Parkinson’s disease using computer vision,” in *Data Management, Analytics and Innovation*, 2021, doi: 10.1007/978-981-16-2937-2_15.

[76] L. Ali, M.-F. Leung, M. A. Khan, R. Nour, Y. Imrana, and A. V. Vasilakos, “ChiGa-Net: A genetically optimized neural network with refined deeply extracted features for trustworthy Parkinson's disease detection,” *Neurocomputing*, vol. 561, 2025, doi: 10.1016/j.neucom.2025.129450.

[77] A. Sadikov, V. Groznik, M. Mozina, J. Zabkar, D. Nyholm, M. Memedi, I. Bratko, and D. Georgiev, “Feasibility of spirography features for objective assessment of motor function in Parkinson's disease,” *Artif. Intell. Med.*, vol. 81, pp. 54–62, Jul. 2017, doi: 10.1016/j.artmed.2017.03.011.

[78] A. Vaish, “A machine learning approach for early identification of prodromal Parkinson’s disease,” *Cureus*, vol. 16, 2024, doi: 10.7759/cureus.63240.

[79] E. Ramzani, M. Y. Tabari, M. G. Amiri, and A. Pouyan, “Diagnosing of Parkinson's disease based on hand drawing analysis using Bi-Directional LSTM equipped with fuzzy inferential soft-max classifier,” *Int. J. Imaging Syst. Technol.*, 2023, doi: 10.1002/ima.22948.

[80] J. Zhang, Y. Lee, T.-M. Chung, and H. Park, “Development of a handwriting drawings assessment system for early Parkinson’s disease identification with deep learning methods,” in *Proc. Int. Conf.*, 2023, pp. 484–499, doi: 10.1007/978-981-99-8296-7_35.

Major Project Report

ORIGINALITY REPORT

5%	4%	4%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Thangaprakash Sengodan, Sanjay Misra, M Murugappan. "Advances in Electrical and Computer Technologies", CRC Press, 2025 Publication	1%
2	Ronald F. Pfeiffer, Zbigniew K. Wszolek, Manuchair Ebadi. "Parkinson's Disease", CRC Press, 2019 Publication	<1%
3	wjaets.com Internet Source	<1%
4	www.mdpi.com Internet Source	<1%
5	Malika Golshan, Adrian E. Bayer. "Massive vs through the CNN lens: interpreting the field-level neutrino mass information in weak lensing", Journal of Cosmology and Astroparticle Physics, 2025 Publication	<1%
6	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	<1%
7	www.biorxiv.org Internet Source	<1%
8	cslide.ctimeetingtech.com Internet Source	<1%

9	D. Jeya Mala, Anto Cordelia Tanislaus Antony Dhanapal, Saurav Sthapit, Anita Khadka. "Integrating AI Techniques into the Design and Development of Smart Cyber-Physical Systems - Defense, Biomedical, Infrastructure, and Transportation", Routledge, 2025 Publication	<1 %
10	H.L. Gururaj, Francesco Flammini, J. Shreyas. "Data Science & Exploration in Artificial Intelligence", CRC Press, 2025 Publication	<1 %
11	propulsiontechjournal.com Internet Source	<1 %
12	www.cognitivecomputingjournal.com Internet Source	<1 %
13	mdpi-res.com Internet Source	<1 %
14	Submitted to University College London Student Paper	<1 %
15	cointelegraph.com Internet Source	<1 %
16	link.springer.com Internet Source	<1 %
17	peerj.com Internet Source	<1 %
18	Submitted to Tilburg University Student Paper	<1 %
19	Yanchao Li, Quan Zou, Qi Dai, Antony Stalin, Ximei Luo. "Identifying the DNA methylation preference of transcription factors using ProtBERT and SVM", PLOS Computational Biology, 2025	<1 %

	Publication	
20	www.ijnrd.org Internet Source	<1%
21	www.sbb.unimi.it Internet Source	<1%
22	Submitted to University of Nottingham Student Paper	<1%
23	osf.io Internet Source	<1%
24	Md Assaduzzaman, Prayma Bishshash, Md. Asraful Sharker Nirob, Ahmed Al Marouf, Jon G. Rokne, Reda Alhajj. "XSE-TomatoNet: An explainable AI based tomato leaf disease classification method using EfficientNetB0 with squeeze-and-excitation blocks and multi-scale feature fusion", MethodsX, 2025 Publication	<1%
25	Submitted to University of North Texas Student Paper	<1%
26	linnk.ai Internet Source	<1%
27	slidelegend.com Internet Source	<1%
28	Neha Goel, Ravindra Kumar Yadav. "Internet of Things enabled Machine Learning for Biomedical Applications", CRC Press, 2024 Publication	<1%
29	Submitted to University of Strathclyde Student Paper	<1%
30	ar.scribd.com Internet Source	<1%