Chapter 1

Introduction

"The secret of getting ahead is getting started." - Mark Twain

The surge in the availability of affordable and powerful computing hardware has led to the integration of various computing devices into our daily lives, with a wide array of applications and interfaces developed for human interaction. These systems are more user-friendly when interactions mimic natural human behaviors like voice or gesture communication. Hand Gesture Recognition (HGR), a key component of Human-Computer Interaction (HCI), is focused on creating technologies that interpret human hand gestures as commands. These HGR systems are utilized in areas like prosthetics [3], sign language recognition [4, 5], rehabilitation [6], and device control, using data from popular input sources such as gloves, vision sensors, inertial measurement units (IMUs), and electromyography (EMG) sensors [4].

However, these commonly used data acquisition methods have limitations: gloves can be uncomfortable for extended use and are inaccessible for amputees [7], and vision sensors often struggle with noise. They are highly dependent on lighting conditions [5], while IMUs can suffer from calibration and drift issues [8]. In contrast, EMG signals are less affected by environmental lighting conditions, are cost-effective, and are easy to use due to their availability as wearable sensors. They can capture both the execution and intention of hand movements by measuring the electrical activity produced by muscles [9]. This makes EMG particularly suitable for amputees, as it allows for more intuitive and precise control of prosthetic devices, overcoming the limitations posed by other sensors [10]. Moreover, EMG sensors can directly detect the user's muscle activity, providing a more reliable and responsive interface for HGR systems.

1.0.1 The Role of Electromyography in Hand Gesture Recognition for Intuitive HCI

Electromyography (EMG) plays a pivotal role in Human-Computer Interaction (HCI), particularly in Hand Gesture Recognition (HGR). EMG technologies track and interpret the electrical activities of muscles during hand movements. This capability allows for the conversion of muscle movements into meaningful digital inputs, enabling a more natural and intuitive interaction with computers and devices. EMG-based HGR systems are crucial in developing user-friendly interfaces, especially in areas that require gesture-based controls, providing a bridge between human motion and digital response. These signals can be recorded or measured in two distinctive ways: surface EMG and intramuscular EMG [11, 12]. Intramuscular EMG is collected by inserting electrodes inside the muscles in an invasive manner and is useful in studying muscle activity having a smaller cross-sectional area or located deep inside the skin [13]. Surface EMG, on the other hand, is a noninvasive technique that is capable of recording and detecting electrical signals during muscles.This method involves placing electrodes on the skin's surface above the muscle of interest. These electrodes detect the electrical potentials generated by muscle fibers when they contract or relax [14].

1.0.2 surface Electromyography(sEMG)

Surface electromyography (sEMG) is a method that measures and records the electrical activity generated by skeletal muscles. The sEMG is a nonstationary microelectrical signal with an amplitude ranging from 0.01 to 10 mV and a frequency spectrum primarily between 20 and 500 Hz, particularly concentrated in the 50 to 150 Hz range [15]. sEMG can detect the electrical potentials produced when muscle cells are activated either electrically or neurologically. The recordings are made with electrodes placed on the skin surface above the muscles of interest. It captures two primary states of skeletal muscle activity: the resting state, where each muscle cell (fiber) maintains an electrical potential of around -80 milliVolts [16, 17]. The alternate state occurs during muscle contraction, which generates an electric potential in a motor unit (MU) consisting of both muscle fibers and a motor neuron. Variations in this electric potential are due to a motor neuron activating a neuromuscular junction, sending two opposing intracellular action potentials. These potentials are spread by the depolarization (activation) and repolarization (resetting) processes in each muscle fiber [18]. The combined intracellular action potentials within a motor unit's muscle fibers create a motor unit action potential (MUAP). EMG measures these MUAPs to analyze muscle contractions and activity. [16, 17].

Muscle contractions can be categorized into two main types: static and dynamic [19, 20]. A static contraction keeps muscle fiber length unchanged and lacks joint movement, yet muscle fibers still contract, such as when holding a hand or making a peace sign. Static contraction occurs when muscle fibers contract without changing their length and without causing joint movement, such as when holding a hand still or making a peace sign. In contrast, a dynamic contraction involves changing the length of muscle fibers and causing joint movement, as seen when waving a hand in a greeting gesture [20].

EMG signals are generated during both types of muscle contractions. To build an efficient Hand Gesture Recognition (HGR) system, it is necessary to model the relationship between static and dynamic contractions and their corresponding EMG signals. Few existing mathematical models can be used for this purpose, which are as follows:

1.0.2.1 EMG Signals: Mathematical models based on contractions

EMG signals can be analyzed as outcomes of stochastic processes influenced by muscle contractions [21]. These contractions can be categorized into two types, which are:

- **Model of Static Contraction (MSC)**: For static contractions, EMG signals are represented as a stationary process, modeled by the equation:

$$\mathrm{EMG}_{\mathrm{static}}(t) = \sum_{i=1}^{M} s_i(t) * m_i(t), \tag{1.1}$$

  Where M denotes the number of active motor units (MUs), $s_i(t)$ the impulse train of each MU, $m_i(t)$ the motor unit action potentials (MUAPs), and $*$ the convolution operation. However, the model can be considered a non-stationary process when influenced by factors such as muscular fatigue and temperature, which affect the EMG signals. [22].

- **Model of Dynamic Contraction (MDC)**: For dynamic contractions, the signals are modeled as a non-stationary process, similar to amplitude modulation:

$$\mathrm{EMG}_{\mathrm{dynamic}}(t) = a(t) \cdot w(t) + n(t), \tag{1.2}$$

  with $a(t)$ indicating EMG signal intensity, $w(t)$ a unit-variance Gaussian process, and $n(t)$ representing noise and signal artifacts [20, 23].

1.0.2.2 Limitations of Contraction Models

These models, while theoretically robust, face several limitations [24]. Their complex parameter estimation makes practical implementation challenging. The

MDC model, in particular, struggles with non-stationary processes, which are time-varying and harder to model accurately [23]. Noise and signal artifacts further complicate accurate EMG signal interpretation. Additionally, real-world adaptability is limited by external factors like electrode placement and individual physiological differences. High computational intensity for processing and analysis is another drawback, especially for real-time applications. Moreover, the assumed linearity between EMG components and muscle contractions may not always hold true. Due to these limitations, machine learning methods are increasingly preferred for handling non-stationary data and overcoming the challenges of traditional EMG signal models.

Through this research, we focus on addressing several challenges in existing applications and developing a robust, accurate, and cost-effective sEMG-based gesture recognition system. Our work aims to assist the development of sEMG-based HGR applications that can benefit various fields, including smart classroom technologies and clinical tools for neurological diagnosis [25].

1.0.3 Motivation

The primary motivation for this research is to design and develop novel sEMG-based recognition techniques focusing on accurately recognizing static and dynamic hand gestures.

An efficient EMG-based recognition model capable of accurately classifying complex hand movements can enhance our understanding of the relationship between sEMG (muscle activation patterns) and corresponding hand movements. This understanding is valuable for various applications, including biometric authentication [26] [27], signature verification [28] [29], hand gesture recognition [30], and e-learning, where it can assist students in writing [31] [32] [33]. Additionally, analyzing handwriting movements can aid in diagnosing neurological diseases such as Parkinson's [34] and Alzheimer's [35]. The methodologies developed through this research can also serve as a foundation for creating prosthetic devices with higher degrees of freedom and precision control [3].

From the literature review, we pinpointed several areas for potential research, which are outlined next:

Advancing HCI for Educational Innovation: Traditional HCI approaches predominantly use mechanical tools like keyboards and mice, which can be cumbersome and non-instinctive for users. The shift to more organic forms of interaction, particularly hand gestures, is driven by the aim to establish more intuitive and accessible interfaces. This need is especially pressing in educational environments, such as smart classrooms, where digitizing handwritten characters is crucial. The adoption of EMG-based hand gesture recognition presents a pioneering solution, offering a natural and efficient way to convert handwritten content into digital formats, thereby enhancing interactivity and technological integration in modern educational settings.

Utilizing the Potential of sEMG Signals in Human-Machine Interaction: Surface Electromyography (sEMG) provides a rich data source about muscle activity, which can be harnessed for interpreting hand gestures. The motivation here lies in effectively translating this biological data into meaningful digital commands, making Human-Machine interactions more seamless and intuitive. Leveraging sEMG signals in Human-Machine Interaction, especially in robotics, enhances the ability to interpret hand gestures accurately and translate them into commands, enhancing control and interaction with robotic systems.

Advancing Prosthetics and Rehabilitation: sEMG's applications extend beyond HCI into areas like prosthetics and rehabilitation. Developing better gesture recognition systems can significantly improve the quality of life for amputees and those undergoing muscle rehabilitation by providing more natural and responsive control over prosthetic limbs and rehabilitation devices.

Addressing Challenges in sEMG Processing: Despite its potential, sEMG signal processing is fraught with challenges like signal interference and variability. The motivation includes tackling these issues using advanced machine learning frameworks, thereby enhancing the reliability and effectiveness of gesture recognition systems.

In summary, this thesis aims to push the boundaries of HCI by leveraging the power of EMG signals and machine learning. This research aims to develop more natural, intuitive, and accessible ways for humans to interact with technology, making significant contributions to fields such as sign language recognition, handwritten character recognition for smart classrooms, medical rehabilitation (Alzheimer's detection using handwriting dynamics), and the development of intuitive interfaces for computing, robotics, and prosthetic devices utilizing EMG signals.

1.0.4 Problem Statement

This thesis delves into the challenge of advancing the field of surface electromyography (sEMG) by designing and developing innovative recognition techniques. The focus is on both static and dynamic hand gestures, encompassing the realms of American Sign Language (ASL), handwritten characters, and hand grasp gesture recognition. The complexity of accurately capturing and interpreting the sEMG signals associated with these diverse gestures necessitates the use of sophisticated data-driven methods, with a primary emphasis on machine learning paradigms. The key challenge lies in harnessing these signals to create reliable and efficient recognition systems that can be practically applied in various real-world scenarios.

To frame the solution, the following assumptions were made throughout the thesis chapters, as all chapters deal with sEMG-based recognition pipelines for static or dynamic gestures:

- **High-Fidelity EMG Data:** Throughout the experiments, it is assumed that the EMG sensors employed provide accurate, high-fidelity data with minimal signal degradation.
- **Context Awareness:** The system is assumed to have an awareness of the context in which gestures are performed, thereby enabling the filtering out of irrelevant movements and focusing on intentional gestures.
- **Handling Sensor Drift:** It is assumed that the pipeline can handle sensor drift, where sensor characteristics may change over time. Techniques such as periodic recalibration or adaptive algorithms are presumed to be utilized to address this issue.
- **Stationarity of EMG Signals**: The EMG signals are assumed to be stationary, implying that their statistical properties remain constant over time.

1.1 Thesis Objectives

This thesis aims to develop sEMG-based hand gesture recognition models that overcome current limitations, improve accuracy and performance, and exhibit substantial real-world applicability. The primary objectives are as follows:

1.1.1 Mapping Electromyographic (EMG) Signals to Hand Gestures (Static and Dynamic)

This objective focuses on developing effective machine-learning pipelines to classify hand gestures by analyzing the structural and functional patterns in Electromyography (EMG) signals. Specific models are designed to classify static hand gestures and dynamic hand gestures separately. This involves constructing dedicated classification models with adaptive classifiers, which are trained and evaluated on test data. The process includes identifying and analyzing the best signal preprocessing methods, feature extraction techniques, and classification algorithms to design a reliable and accurate sEMG-based hand gesture recognition framework. The detailed steps are provided below:

1.1.1.1 Collecting sEMG Signals and Signal Preprocessing

The experimental framework leverages both newly collected and publicly available datasets comprising raw surface electromyography (sEMG) signals.These datasets encompass a range of static and dynamic hand gestures relevant to American Sign Language, handwritten character recognition, and hand grasping activities. The acquired raw EMG signals undergo preprocessing using various filtering and feature extraction techniques to select the most suitable ones. After conducting multiple inspections, we use Butterworth band-pass filtering methods to improve signal quality and reduce the effects of noise and artifacts.

1.1.1.2 Feature Engineering: Feature ensemble and metaheuristic feature selection

In this process, the preprocessed EMG signal is represented by a set of parameters (features). Feature normalization is performed to ensure consistent scaling, mitigating potential issues caused by inadequately scaled features. This step is crucial for enhancing the system's ability to distinguish between different classes, eventually influencing the accuracy of the final classification.

Two advanced feature engineering approaches are employed and explored: Feature ensemble and meta-heuristics feature selection. Meta-heuristic feature selection approaches are designed to efficiently search the feature space, balancing exploration (searching new areas) and exploitation (refining known good areas). By doing so, they can improve the accuracy and efficiency of feature selection, potentially leading to better model performance. Additionally, these approaches can be cost-effective because they aim to find an optimal subset of features more quickly and with fewer resources compared to exhaustive search methods. Meanwhile, feature ensemble involves combining multiple feature selection methods to create a comprehensive set of features that captures various aspects of the data, thereby enhancing the model's performance.

By integrating metaheuristic and feature ensemble selection, the goal is to capture the most relevant and informative aspects of the EMG signals relevant to various gestures. This comprehensive approach targets enhancing the system's ability to accurately classify static and dynamic hand gestures, making it well-suited for real-world applications.

1.1.1.3 Hand Gesture Recognition: Classification Using Machine Learning Methods

To achieve accurate classification results, we utilized a range of supervised methods. Our experiments incorporated both baseline machine learning algorithms and advanced tree-based techniques, including Support Vector Machine (SVM), Naïve Bayes (NB), Decision Tree (DT), eXtreme Gradient Boosting (XGB), and CatBoost. These algorithms were employed to detect static and dynamic hand gestures from sEMG patterns, and their effectiveness was evaluated based on the highest classification accuracy achieved.

Our primary concern was identifying the most representative feature set, ensuring that even baseline machine learning classifiers (often used as a starting point for building more complex models) could achieve high accuracy. This approach is aimed at developing robust real-world applications that perform effectively with simpler models, as simpler models are less computationally intensive.

1.1.2 Dynamic Hand Gesture Recognition: sEMG-Based approach for Handwritten Character Recognition (HCR)

Another objective of this thesis is to develop an sEMG-based dynamic hand gesture recognition system. This system leverages deep representation learning [36] to construct an effective HCR pipeline. The goal is to enhance real-time human-computer interaction (HCI) applications, with a specific emphasis on smart classroom environments and clinical handwriting dynamics analysis.

1.1.3 Multi-Modal Handwritten Character Recognition

This thesis proposes the development of a multi-modal handwritten character recognition (HCR) pipeline that integrates physiological sensors (EMG) and inertial sensors (IMU). The goal is to enable the model to learn both within-modality and inter-modality correlations, thereby enhancing data representation. The approach involves early feature fusion [37] and the implementation of a modified deep auto-encoder architecture to learn an enriched combined data representation. Additionally, the thesis aims to assess the impact of fusing EMG and IMU sensors on the recognition of dynamic hand gestures. This comprehensive approach seeks to improve the accuracy and robustness of HCR systems, with potential applications in smart classroom settings using whiteboards and clinical handwriting dynamics analysis for conditions such as Alzheimer's disease.

1.1.4 Efficient hand grasp recognition using a minimal number of sEMG signals

The objective is to develop an sEMG-based hand grasp recognition (HGR) pipeline to differentiate between various hand grasps, which are essential for applications in prosthetics and robotics. The proposed approach leverages a game theory-based feature selection method to identify the most representative features, thereby ensuring high recognition accuracy. The framework is designed and optimized to minimize the number of sEMG sensors required, enhancing efficiency and practicality. The proposed pipeline is validated using a publicly available dataset, ensuring rigorous evaluation and reliability of results.

1.1.5 Enhancing Interpretability with Explainable AI (XAI)

Moreover, the thesis places a strong emphasis on enhancing the interpretability of the proposed pipelines for both static and dynamic hand gesture recognition by utilizing Explainable AI (XAI) methods, with a particular focus on SHAP (SHapley Additive exPlanations). These techniques ensure that the decision-making processes of the models are transparent and understandable, making the systems more reliable and trustworthy if used in practical applications.

This involves identifying influential features in all the mentioned classification tasks, significantly contributing to the fields of biomedical signal processing, electromyography, and human-computer interaction. Enhanced interpretability is crucial for applications such as American Sign Language (ASL) interpretation, smart classroom technology, and clinical analysis.

However, we acknowledge that the XAI component in the thesis was not positioned as a central contribution, but rather as a supportive tool to inform and interpret model behavior—particularly in relation to feature relevance and parameter space analysis.

The summary of the overall research goal can be listed as follows.

- *"To design and validate classification models that can efficiently recognize hand gestures(Static/Dynamic) using sEMG signals, exploring and employing appropriate signal pre-processing, feature extraction, and classification algorithms.*

- *"To construct an sEMG-based dynamic hand gesture recognition system that utilizes deep representation learning, thereby improving Handwritten Character Recognition(HCR) pipelines for real-time applications."*
- *"To develop a multi-modal HCR pipeline that integrates EMG and IMU sensors, enhancing data representation and analyzing the impact of this fusion on gesture recognition accuracy."*
- *"To develop an efficient hand grasp recognition system that prioritizes the use of minimal number of sEMG sensors, this approach leverages game theory-based feature selection to achieve higher accuracy and practical applicability."*
- *"To improve the interpretability of sEMG-based recognition systems by deploying Explainable AI methods, identifying influential features in classification tasks and ensuring transparency and trust in the models' decision-making processes."*

1.2 Contribution of thesis

This thesis investigates innovative methodologies to enhance the recognition of static and dynamic hand gestures through surface electromyography (sEMG). Our research encompasses a progression from solving less complex problems (static gestures) to tackling more intricate (dynamic gestures)challenges within sEMG-based hand gesture recognition. To address the problems, we developed four new datasets—ASL-10, ASL-24, HCR_{emg}, and HCR_{MM}—covering a wide range of hand gestures for various sEMG applications. These datasets include various American Sign Language (ASL) and handwritten character gestures (both on paper and on a whiteboard). These datasets were utilized for both static and dynamic hand gesture recognition, forming the foundation for creating robust frameworks. Detailed descriptions of these datasets are provided in Section 2.3.3.

For static hand gesture recognition, we designed efficient pipelines aimed at improving accuracy for building real-time applications. Our ASL gesture recognition pipeline, employing a novel ensemble feature selection method, achieved an average classification accuracy of 99.91% on the ASL-24 dataset using eight sEMG channels. We validated our approach using the publicly available Ninapro benchmark dataset [1]. (Details discussed in Chapter 3)

The thesis also introduced an efficient pipeline for distinguishing various static hand grasps. Hand grasps refer to the various ways we use our hands to hold and manipulate objects. We employ a less explored game theory-based feature selection approach. This method achieved up to 98.2% recognition accuracy on a publicly available dataset. By focusing on using a minimal number of sensors, we ensured that these pipelines are not only accurate and reliable but also cost-effective, making them suitable for real-time applications such as prosthetics. (Detailed discussed in Chapter 4)

Building on these successes, we tackled the more complex problem of dynamic hand gesture recognition. We proposed a sEMG sensor-based handwritten character recognition pipeline that leverages deep feature representation learning [36] for improved performance. A deep autoencoder variant was designed for feature extraction, achieving up to 98.72% accuracy with baseline machine learning classifiers. Furthermore, we explored the feasibility and analyzed the impact of multimodal sensor fusion by integrating sEMG with Inertial Measurement Unit (IMU) data, resulting in 99.01% accuracy for handwritten characters on a whiteboard. (Detailed discussed in Chapter 5 and Chapter 6).

A significant contribution of this research is the development of the Modified Salp Swarm Algorithm (MSSA) for feature selection. This enhanced algorithm improves upon the original Salp Swarm Algorithm and has been rigorously validated against 14 benchmark datasets. The MSSA is specifically designed to select essential features from both sEMG and IMU sensor data, with the aim of enhancing handwritten character recognition. By effectively selecting relevant features, MSSA not only significantly boosts model accuracy and operational efficiency but also results in a smaller feature vector, thereby enhancing computational efficiency. This contribution underscores the utility of MSSA as an efficient feature selection tool for sEMG and other applications. (Detailed discussed in Chapter 7)

In an additional contribution, we focused on enhancing the interpretability and transparency of our model for classifying Alzheimer's disease using handwriting dynamics. By incorporating explainable AI techniques, specifically SHAP (SHapley Additive exPlanations), we gained detailed insights into the decision-making process of our ensemble model. This approach allowed us to identify the most influential features for predicting Alzheimer's disease based on handwriting movements. We introduced an effective stacking ensemble model designed to classify Alzheimer's disease, which was tested against a contemporary benchmark dataset and demonstrated superior precision. Utilizing SHAP, we effectively highlighted pivotal features, significantly enhancing the model's interpretability and transparency. (Detailed discussion in Chapter 8)

These contributions collectively advance the field of sEMG-based gesture recognition, offering robust solutions for static and dynamic gestures. The thesis also contributes to the field by analyzing the feasibility and improvement of existing techniques, such as ensemble feature selection for sEMG-based classification tasks. It also highlights the efficiency of several less-explored areas, such as deep feature representation learning and cooperative game-based feature selection for sEMG-based gesture recognition. Additionally, they extend the applicability to critical clinical applications, such as Alzheimer's disease classification using handwriting dynamics.

However, the proposed models still exhibit a few aspects that are less robust, highlighting opportunities for further research and improvement. **The detailed**

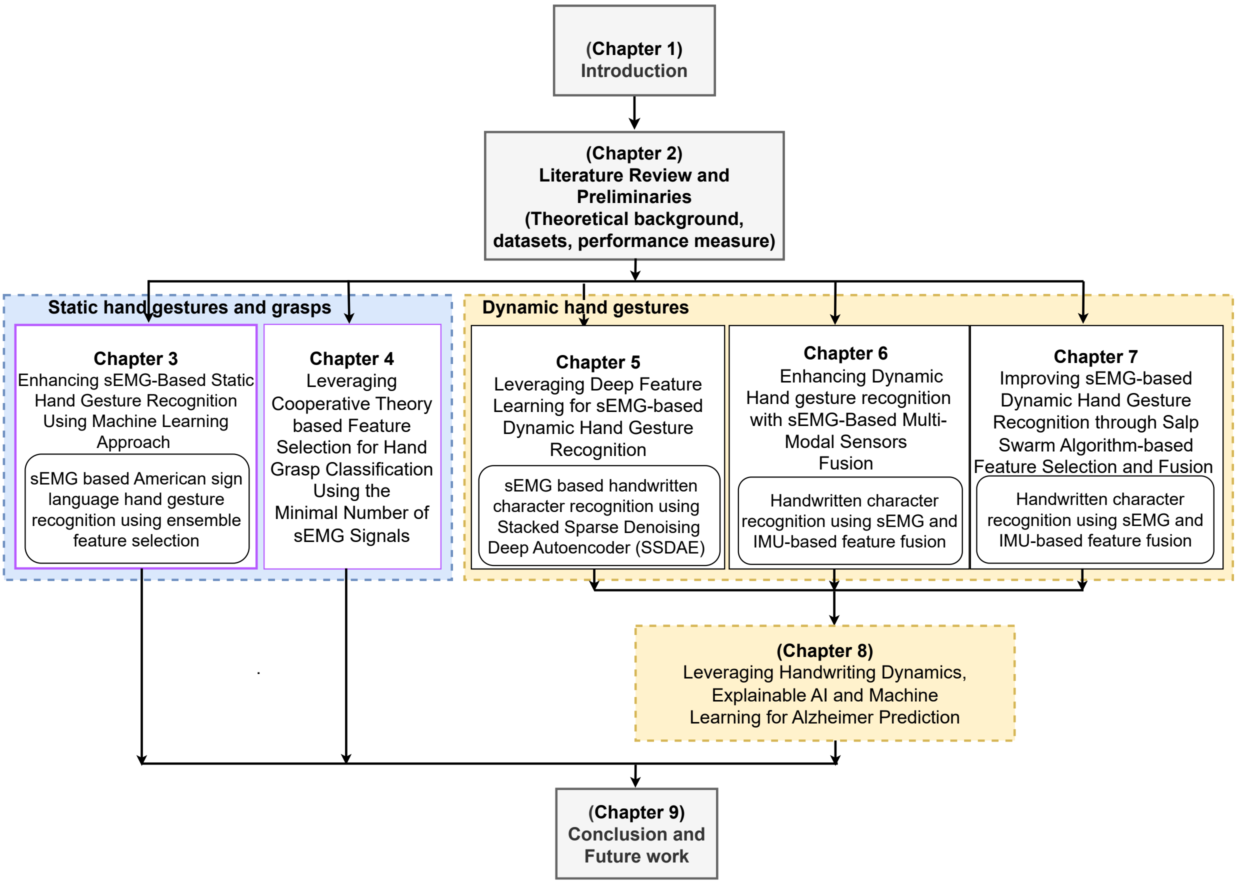

FIGURE 1.1: Thesis Structure

discussion of each contribution is provided in the respective chapters as separate sections. The organization of the rest of the thesis is as follows:

Chapter 2 provides a comprehensive theoretical background and literature review on traditional and contemporary approaches to EMG-based hand gesture recognition. It details popular datasets utilized in this research domain and discusses the tools and frameworks employed for implementing various recognition techniques presented in the thesis.

Chapter 3 investigates the potential of surface electromyography (sEMG) for static hand gesture recognition using machine learning methodologies. It introduces and evaluates an ensemble feature selection approach designed to enhance the performance of American Sign Language (ASL) gesture recognition.

Chapter 4 describes a cooperative game theory-based approach tailored for sEMG-based hand grasp recognition. This chapter focuses on using a minimal number of sensors while maintaining high recognition accuracy.

Chapter 5 explores deep representation learning for developing an efficient and robust handwritten character recognition (HCR) pipeline. It employs a stacked sparse denoising autoencoder network to obtain effective deep feature representations, which are then utilized for the recognition tasks.

Chapter 6 examines the impact of the multi-modal feature fusion of sEMG sensors with other sensors on dynamic gesture recognition. It discusses feature fusion approaches using deep feature learning and autoencoder techniques.

Chapter 7 aims to enhance sEMG-based dynamic hand gesture recognition by leveraging an improved Salp Swarm Algorithm (SSA) for feature selection and fusion. The chapter details the modifications made to the original algorithm and evaluates its effectiveness.

In Chapter 8 we examined the potential of dynamic hand gestures in predicting neurological conditions such as Alzheimer's disease by focusing on handwriting dynamics. Utilizing explainable AI, this study further investigates these dynamics using the "DARWIN" dataset to analyze motion patterns related to disease detection. The goal is to enrich our sEMG-based research, constructing a robust framework that supports early diagnosis and deepens our understanding of Alzheimer's through the study of fine motor skills.

Chapter 9 concludes the thesis by summarizing the major findings and addressing the research questions posed in each chapter. It also discusses the significance of the thesis outcomes within the context of current research and outlines potential directions for future work.