UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING

Alen Salkanović

A DEEP LEARNING MODEL FOR TOUCHSCREEN HANDWRITING RECOGNITION BASED ON A SENSOR FUSION APPROACH

DOCTORAL DISSERTATION

Rijeka, 2025.

UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING

Alen Salkanović

A DEEP LEARNING MODEL FOR TOUCHSCREEN HANDWRITING RECOGNITION BASED ON A SENSOR FUSION APPROACH

DOCTORAL DISSERTATION

Supervisor: Prof. Sandi Ljubić, PhD

Rijeka, 2025.

SVEUČILIŠTE U RIJECI
TEHNIČKI FAKULTET

Alen Salkanović

MODEL DUBOKOGA UČENJA ZA PREPOZNAVANJE RUKOPISA NA DODIRNIM ZASLONIMA ZASNOVAN NA PRISTUPU FUZIJE OSJETILA

DOKTORSKI RAD

Rijeka, 2025.

Supervisor: Prof. Sandi Ljubić, PhD (University of Rijeka, Faculty of Engineering)

Doctoral dissertation was defended on _____ at University of Rijeka, Faculty of Engineering, in front of the committee:

1. Prof. Ivo Ipšić, PhD - committee chair (University of Rijeka, Faculty of Engineering)

2. Prof. Kristijan Lenac, PhD - member (University of Rijeka, Faculty of Engineering)

3. Assist. Prof. Žiga Emeršič, PhD - member (University of Ljubljana – Faculty of Computer and Information Science, Slovenia)

ACKNOWLEDGEMENTS

ABSTRACT

The use of biometric data to recognize individuals is a frequently studied topic in the field of computer security. Biometric identification relies on analyzing and measuring human traits to determine a person's identity. Handwriting analysis represents a conventional biometric modality that captures the distinct writing patterns and characteristics unique to each individual. Traditional approaches primarily concentrate on offline handwriting verification, where static images of handwritten text are analyzed to distinguish between genuine and forged signatures. The emergence of touchscreen devices with a variety of built-in sensors has enabled a more comprehensive approach to online handwriting recognition, leveraging dynamic writing characteristics such as pressure, speed, the number and order of the strokes, pen tilt, and other relevant attributes.

Related studies in this field highlight the predominance of certain research methodologies. Specifically, in most instances, the analysis focuses solely on the signature, verification methods are more widely used compared to recognition methods, and the proposed approaches primarily depend on a specific device or sensor to gather handwriting biometric data. This doctoral dissertation introduces a novel approach to user authentication that extends beyond conventional online signature verification systems. The study aims to address the research gap by broadening the focus from verification to a wider range of handwriting identification tasks. This refers to a general process of determining the handwriting owner that involves multi-class classification. The proposed system utilizes sensor fusion, a technique that integrates and processes measurements collected by multiple sensors. Differing from the use of specialized hardware, such as custom-designed pens or complex sensor solutions, which are often limited and costly, this system employs readily available off-the-shelf components and common touchscreen smart devices. In contrast to prevalent research mainly centered on handwritten signatures, the investiga-

tion examines the gathering and analysis of handwriting in various forms, including short sentences, words, and individual letters. The analysis incorporates two distinct methods of handwriting input, utilizing both a stylus and a finger.

A total of 60 participants took part in a controlled experiment to form a newly curated handwriting biometrics dataset, consisting of measurements from different types of sensors. Custom CNN models were employed for feature extraction and classification tasks to identify participants' handwriting. The accuracy of the model was analyzed in relation to the impact of three factors: train set size, input modality, and handwriting form. The obtained results show a statistically significant effect of set size on the model's accuracy, with the highest accuracies observed with the largest train set size. The statistical analysis also revealed that higher accuracy was consistently achieved for recognizing stylus-based handwriting compared to finger-based writing.

An ablation study was conducted to analyze the impact of each sensor in the fusion-based setup. The acquired findings indicated that the proposed person recognition system yields the highest accuracy for signatures, while also showing notably high accuracy for sentences and words. Overall, these findings suggest that the most suitable sensor fusion subset is highly dependent on both the handwriting form and the input modality. When all sensors are included, the model's accuracy improves for recognizing sentences, words, and letters that are entered using a stylus. This highlights the benefit of utilizing a comprehensive multi-sensor setup. Hence, it can be concluded that including additional sensors can certainly enhance the accuracy of the model, thereby justifying their integration into the experiment apparatus.

PROŠIRENI SAŽETAK

Korištenje biometrijskih podataka za prepoznavanje osoba često je istraživana tema u području računalne sigurnosti. Biometrijska identifikacija oslanja se na analizu i mjerenje ljudskih karakteristika kako bi se utvrdio identitet osobe. Analiza rukopisa predstavlja standardni biometrijski modalitet koji bilježi jedinstvene obrasce i značajke rukopisa specifične za svakog pojedinca. Tradicionalni pristupi uglavnom su usmjereni na verifikaciju rukopisa, pri čemu se analiziraju slike rukopisa kako bi se razlikovali originalni i krivotvoreni potpisi. Pojava uređaja sa zaslonom osjetljivim na dodir i brojnim ugrađenim senzorima omogućila je sveobuhvatniji pristup prepoznavanju osoba pomoću rukopisa, koristeći dinamičke značajke pisanja poput razine pritiska, brzine pisanja, broja i redoslijeda poteza, nagiba olovke i drugih atributa od značaja.

Srodne studije u ovom području ističu određene metodologije koje prevladavaju u istraživanjima. Konkretno, u većini slučajeva analiza se usmjerava isključivo na potpis, metode verifikacije su zastupljenije od metoda identifikacije, a predložena rješenja uglavnom se oslanjaju na određeni uređaj ili specifični senzor za prikupljanje biometrijskih podataka. Ova doktorska disertacija uvodi novi pristup autentifikaciji korisnika koji nadilazi tradicionalne sustave verifikacije potpisa. Cilj istraživanja je prebaciti fokus s verifikacije na širi raspon zadataka povezanih s identifikacijom rukopisa. Ovo se odnosi na općeniti postupak određivanja vlasnika rukopisa koji uključuje višeklasnu klasifikaciju. Predloženi pristup implementira koncept fuzije senzora, tehniku koja objedinjuje i obrađuje podatke prikupljene s više senzora. Za razliku od korištenja specijaliziranog sklopovlja, poput posebno dizajniranih olovaka ili složenih senzorskih sustava, koji su često ograničeni i skupi, ovaj sustav koristi lako dostupne komponente i uobičajene pametne uređaje s dodirnim zaslonima. Za razliku od većine istraživanja koja su uglavnom usmjerena na potpise, ovo istraživanje ispituje prikupljanje i analizu rukopisa u različitim oblicima,

uključujući kratke rečenice, riječi i pojedinačna slova. Osim toga, analiza obuhvaća dva različita načina pisanja – korištenjem olovke (stylusa) i prsta.

Ukupno 60 sudionika sudjelovalo je u kontroliranom eksperimentu za formiranje novog skupa biometrijskih podataka, koji se sastoji od mjerenja s različitih tipova senzora. Kako bi se klasificirao rukopis sudionika, CNN modeli dubokoga učenja korišteni su za zadatke izdvajanja značajki i klasifikacije. Točnost modela analizirana je u odnosu na utjecaj tri različita čimbenika: veličine skupa za treniranje, modaliteta unosa i oblika rukopisa. Dobiveni rezultati pokazuju statistički značajan učinak veličine skupa na točnost modela, pri čemu je najveća točnost uočena kod najveće veličine skupa za treniranje. Statistička analiza također je pokazala da je postignuta viša točnost za prepoznavanje rukopisa olovkom u usporedbi s pisanjem pomoću prsta.

Provedena je ablacijska studija kako bi se analizirao utjecaj pojedinačnih senzora unutar sustava temeljenog na fuziji. Dobiveni rezultati pokazali su da predloženi sustav za prepoznavanje rukopisa postiže najvišu točnost za potpise, dok također pokazuje izrazito visoku točnost za rečenice i riječi. Općenito, rezultati ukazuju na to da najpogodniji podskup fuzije senzora uvelike ovisi o obliku rukopisa i modalitetu unosa. Kada su uključeni svi senzori, točnost modela se poboljšava za prepoznavanje rečenica, riječi i slova unesenih pomoću olovke, što dodatno naglašava prednost korištenja sustava s više senzora. Stoga se može zaključiti da uključivanje dodatnih senzora svakako može povećati točnost modela, opravdavajući njihovu implementaciju u eksperimentalni aparatus.

Ključne riječi: autentifikacija, biometrija, digitalni potpisi, prepoznavanje rukopisa, strojno učenje, fuzija senzora, rukopis na zaslonu osjetljivom na dodir

CONTENTS

Chapter 1

INTRODUCTION

Determining and verifying people's identities is a critical necessity throughout many domains of contemporary society. It prevents unauthorized access to sensitive resources and systems, thereby reducing the risks of fraud and forgery. Confirming identities holds significant importance across diverse sectors, spanning from banking and education to the legal system and healthcare. Financial organizations are obligated to ascertain only authorized persons can access accounts and conduct transactions, especially in the current context where all transactions can be done over the network [1]. Identification is essential in education for maintaining academic integrity and accurate records, while in the legal system it facilitates justice administration, security and proper documentation. Healthcare facilities increasingly adopting smart Internet of Things (IoT) systems need to protect patients against privacy breaches, tampering, and data falsification.

Conventional mechanisms including passwords or personal identification numbers (PINs) are prone to disadvantages like limited complexity, vulnerability to attacks, theft, and management difficulties [2]. Advances in information technology extensively transform traditional access controls by introducing novel and sophisticated methods, including multi-factor authentication (MFA) [3]. To mitigate vulnerabilities and improve the security of conventional password-based systems, MFA incorporates an additional layer of protection. Among the various types available, the most commonly used forms include one-time passwords (OTPs), smart cards, tokens, certificates, and biometric methods. While passwords, cards, and physical tokens are considered non-biometric methods to identify individuals, biometric identification relies on unique personal traits such as finger-

prints, gait or hand geometry. The utilization of a specific modality depends on security requirements and usage scenarios. Non-biometric methods are often used for everyday tasks such as accessing buildings or logging into computers and online accounts. In contrast, biometric identification is used in contexts where security demands are elevated, and require a high level of reliability [4]. This method utilizes unique and inherent human characteristics to verify identities, offering a reliable and alternative solution to conventional approaches.

1.1.   Biometric systems

A biometric systems measure one or more unique physical, behavioral patterns, or psychophysiological characteristics of a human body. The first pertains to the physical attributes of an individual's body, typically remaining consistent over time. These traits stem from unique biological attributes that vary significantly between individuals, making them challenging to replicate [5]. Examples include iris, hand geometry, fingerprints, facial features, ear shape, and vein patterns. In contrast, behavioral or affective traits pertain to learned behaviors and measurable patterns in human activities. Behavioral characteristics are inherently dynamic and subject to change over time, influenced by factors such as mood, age, health, and other variables. The modalities encompass gait, body posture, voice, keyboard typing patterns, and handwriting style [6].

Biometric systems are categorized in two distinct types: the unimodal biometric system (UBS) and the multimodal biometric system (MBS). Each type has its own benefits and drawbacks, and the choice typically relies on specific needs and constraints of the application context. In UBS, only one type of biometric trait is employed for user verification or identification. Data is collected and processed from a unique physiological or behavioral trait of an individual. These systems are easier to implement, but vulnerable to spoof attack, external factors and reduced accuracy if the biometric feature is degraded [7]. MBS integrates several biometric features into a multimodal approach, such as combining fingerprint, facial, and iris recognition. Expanding the range of identifying modalities increases security, enhances recognition accuracy and improves the reliability of person identification [8]. Multimodal biometric systems can be categorized into several different types, such as those involving multiple sensors, modalities, algorithms, samples,

and instances [9]. Compared to the unimodal approach, MBS offers the benefit of reducing the probability of several traits getting simultaneously compromised. However, these systems also exhibit certain limitations, especially when implementing on a large scale. There may be incompatibilities among different devices, acquiring and storing data from various sources can be challenging, and the system may encounter difficulties in maintaining high performance.

A biometric system can operate in two modalities for identity management purposes, specifically verification and identification [10]. Verification, alternatively referred to as authentication, denotes the procedure validating the identity claimed by a subject. In this instance, the provided biometric data is compared with a pre-existing template corresponding to a particular individual, serving as reference data for the claimed identity (a one-to-one match) [11]. When a match is detected based on a specific level of similarity, the system confirms the individual's identity. Otherwise, if the outcome does not satisfy the required threshold, the individual is rejected. Verification is frequently employed in scenarios where individuals need to confirm their identity to access a system, location, or service, such as using a fingerprint to gain access to a smartphone.

In identification modality, the system determines the subject's identity by comparing all registered templates in the database without requiring the person to claim an identity. The biometric framework performs a comprehensive search of the complete database to find a correspondence for the submitted biometric data. In this context, identification may be conceptualized as a series of verification processes [12]. It proves beneficial in scenarios where entity must be identified without providing any credentials. Identification is crucial in contexts where an individual's identity must be established from a large biometric database, without prior knowledge of the claimed identity. This is common in law enforcement agencies and access control systems, such as border control.

The choice of biometric system is largely influenced by the nature of the application. Before selecting a system, it's important to consider existing security measures (passwords, smart cards, PINs, security tokes, etc.), as biometrics may complement rather than replace them. Depending on the application, either verification or identification functionalities may be required. Additionally, it is essential to determine which biometric trait will be measured and analyzed based on the system's intended use.

The handwritten signature is a commonly utilized biometric characteristic, remain-

ing prevalent for identity verification in administrative, financial, and legal fields [13]. Throughout centuries, traditional pen-and-paper signatures have attained social and regulatory acceptance globally, reflecting the distinct learned actions of individuals challenging to replicate. Despite initiatives aimed at digitizing signatures, handwritten versions remain prevalent in practice. This is primarily due to their widespread applicability and acceptance compared to electronic alternatives [14]. Signatures refer to a distinctive writing style of a person and are not regarded as simply a combination of letters or words. In the realm of biometrics, handwritten signatures are categorized under behavioral characteristics, reflecting the learned patterns and habits of an individual's handwriting. This contrasts with physiological biometrics, such as fingerprints or iris patterns, which are inherent physical attributes. The dynamic and personal nature of signature creation makes it a distinctive identifier, but it is susceptible to variations due to different influences. The features of signatures can change considerably due to factors like emotional state, physical condition, level of focus or attention, and age. This variability is further influenced by cultural practices, personal habits, psychological state, and physical condition [15].

1.2. Handwritten signature verification systems (HSVS)

Handwritten signature verification systems (HSVS) involve analyzing and processing handwritten signatures to confirm an individual's claimed identity. The verification process typically involves comparing the provided signature against stored reference signatures. Two distinct approaches to signature authentication refer to writer-dependent (WD) and writer-independent (WI) verification systems. Writer-dependent systems are developed to authenticate signatures by examining distinct characteristics inherent to the handwriting of an individual. These systems require the enrollment of multiple signature samples from a particular person during a registration process. Systems are individually trained for each unique user, and both the learning and verification processes are based on the signer. A limitation of this method is the necessity to generate a model for every additional writer undergoing verification [16]. Writer-independent systems employ a generic model to verify signatures from any writer, regardless of whether their samples

were included in the training dataset [17]. Such approaches prioritize identifying general patterns and characteristics that are typical across signatures, rather than focusing on individual variations. WI systems require a single global classifier that can be utilized with just one signature sample from each user. These systems are trained on a diverse dataset comprising signatures from numerous individuals, thereby eliminating the requirement for prior training on specific persons [18]. However, verifying signatures in a writer-independent manner is more difficult because of the considerable morphological variability among different writers. Another strategy that can be employed represents a hybrid WD-WI system, which combines the WD and WI methods.

Advanced HSVS employ machine learning techniques, such as deep learning, to extract distinctive features and improve accuracy in distinguishing genuine signatures from forgeries. Forgeries are generally classified into three main forms based on different levels of attempts to falsify or imitate signatures: random forgery, unskilled forgery (simple or basic forgery), and skilled forgery. Random forgery is characterized by an attempt to imitate the owner's signature in a random style. The attacker attempts to mimic the signature without having prior knowledge of its genuine form. This results in a forged representation that significantly differs in appearance and structure from the authentic one. Unskilled forgery happens when the forger is familiar with the victim's name but lacks knowledge of the exact details of how the signature is written. As a result, they try to imitate the signature in their own manner, even though they have no previous experience. In this case, the forgery closely resembles the genuine signature in shape and visual characteristics. In cases of skilled forgery, the attacker has thorough knowledge of the victim's name as well as their signature, demonstrating considerable expertise in accurately replicating the signature. Detecting this type of forgery is challenging, as visual inspection may fail due to its close resemblance to the genuine writing style [19].

The signatures to be authenticated by the HSVS can be obtained through two different approaches. Depending on the method of data gathering, the systems are classified as offline (static) and online (dynamic) [20]. Offline signatures denote physical signatures on paper that are verified after the writing process has concluded. Online signatures, in contrast, represent digital signatures obtained through electronic devices and are typically verified immediately or in real-time. Prominent advantages of signatures as a biometric determinant comprise their consistency, non-intrusive data collection methods, and

5

the potential to incorporate dynamic features [21]. With technological development, it becomes possible to acquire dynamic signature characteristics typically obscured on paper, including writing speed, size fluctuations, sequence of strokes, tilt of the pen, and pressure intensity. Both offline and online HSVS systems employ distinct techniques and technologies, which are further detailed in the subsections 1.2.1. and 1.2.2.

Signature verification commonly follows a process consisting of four main stages [22]. These steps are generally applicable across different implementations of signature verification and include:

– Data acquisition: offline systems use devices like cameras or scanners to convert physical signatures into digital images. Online systems utilize tablets, electronic pens, and smartphones for acquiring signatures in real-time. Additionally, publicly accessible datasets can also be utilized for evaluating system performance.

– Preprocessing: this entails refining and organizing the acquired data for subsequent analysis. For instance, digital images may undergo preprocessing like noise reduction, binarization, and deskewing. Raw readings from digital pens or tablets is refined by removing noise and normalizing the data.

– Feature extraction: features are derived from the photograph of the signature, encompassing shape descriptors, texture characteristics, and statistical metrics. Online signatures involve extracting dynamic components of signing, such as pen trajectory, velocity, acceleration, pressure variations, pen tilt, and timestamp.

– Classification: serves as the technique for evaluating the authenticity of a given signature. The extracted characteristics are matched against the database entries to determine whether they are genuine or forged.

1.2.1. Offline HSVS

The offline HSVS pertains to a system in which templates, primarily signatures, are derived from imaging devices such as optical scanners or cameras. The signature is displayed as a digital photo, with the inked mark created on paper utilizing a handheld tool like a pen or pencil. The system is tasked with determining if an image of a signature is authentic or a forged copy [23]. In the offline scenario, the signatures are obtained after

the paper-writing process has concluded. Therefore, only static data is extracted from the original writing procedure, including aspects such as shape, outline, length and height, slant, bounding box, and baseline. This indicates that the dynamic features involved in the creation of the signature, like the pen's movement and writing speed, are not obtained or stored. Due to this absence of information, verifying offline signatures is regarded as more difficult process compared to online signatures [24]. Additionally, the characteristics of a signature can be influenced by the type of writing instrument and distinct, highly stylized or unconventional writing styles. However, offline systems provide a notable benefit in that signatures can be acquired without the requiring specialized processing devices.

Verifying signatures offline typically relies on analyzing geometric characteristics of the signature. However, certain methods also attempt to extract dynamic elements from the signature image. Since dynamic information cannot be directly derived from static signature images, specific characteristics are used to partially represent these dynamic aspects, referred to as pseudo-dynamic information. The term "pseudo-dynamic" is used to differentiate real dynamic data, captured during the writing process, from information that can be inferred or reconstructed from the static image [25].

1.2.2. Online HSVS

For online signature verification, measurements are recorded in the form of a time series signal, typically collected in real-time as signatures are being executed. Since signatures from a single individual are not exactly identical, it is essential to capture multiple sample signatures from every individual to serve as training data [26]. Signatures are acquired using electronic input devices, including digital pens, pressure-sensitive tablets, or touchscreens on smartphones and handheld devices. Therefore, real-time dynamic properties regarding the signing procedure can be gathered during the act of handwriting, including spatial and temporal characteristics of a signature. Online signatures consist of a sequence of sampled points, each comprising various extracted features [27]. These factors, among others, encompass acceleration, speed, pressure exertion, stroke sequence, position, timing, and pen angle (including altitude and azimuth). They are recorded as function of time and saved continuously throughout the signature - collection procedure.

In comparison to offline systems, these available attributes enhance accuracy, as repli-

cating dynamic characteristics is notably challenging. The variability in features at each point of the signature enhance its uniqueness and increases the difficulty of forgery [28]. In online scenarios, identity verification benefits from requiring forgers to replicate not only the static signature image, but also the unique signing gesture. Accurately imitating this gesture poses a greater challenge than recreating the image alone. Moreover, although a signature follows a particular movement pattern, it consistently generates a signal that varies from one occurrence to another [29].

Some drawbacks in online signature verification include dealing with the fluctuations of handwritten signatures, as they may change in the long term. The characteristics of the digital medium, such as the resolution and size of the signing area on tablets or touchscreens, can influence how a person signs. The verification process may be affected by variations in hardware or software configurations, influencing accuracy and reliability. Certain individuals may consider signing on digital devices to be less intuitive or comfortable compared to the traditional method of signing on paper. Moreover, people vary in their coordination and consistency, which leads to the development and refinement of their signing motion and gesture over time. Consequently, these variations within the signatures of the same individual (intra-class differences) become an important consideration.

1.3.   Handwritten signature identification

In biometric systems, it's crucial to differentiate between verification and identification. Handwritten signature identification represents a distinct biometric approach compared to signature verification. It involves determining the identity of an individual by analyzing a given signature against a database of known signatures. In contrast to signature verification, which confirms whether a presented signature matches a claimed identity, identification involves searching through multiple signature samples to find the most probable match. This distinction is important, as verification follows in a one-to-one comparison, where a signature is checked against a specific reference sample. On the other hand, identification utilizes a one-to-many approach, requiring the classification of an unknown signature against a large database of multiple potential identities without prior claim information [30].

Signature identification presents distinct challenges when compared to verification. One of the main difficulties is addressing intra-person variability, which is affected by inconsistencies in factors like pen pressure, stroke dynamics, and writing speed. Additionally, different individuals may have signature styles that closely resemble each other, leading to potential misidentification. This issue contrasts with signature verification, where the primary concern is preventing forgeries rather than resolving identity ambiguities. Since the system must distinguish between multiple individuals, the system demands a robust approach to feature extraction and classification to minimize the risk of false identifications [31].

A variety of approaches have been suggested for signature identification, employing machine learning and deep learning techniques. Contemporary methods make use of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to extract dynamic characteristics of signatures. The focus of this dissertation is on using deep learning techniques for recognizing and identifying individuals through handwriting. Unlike related solutions, the proposed system is based on sensor fusion for person recognition, not just from signatures, but also from other forms of handwriting patterns like short sentences, words, and individual letters.

1.4.   Hypotheses and Contributions

Various approaches to user verification and identification through handwriting are presented in Chapter 2. In the context of online handwriting person recognition, signatures have been the most extensively studied as the primary biometric marker for identifying individuals. Furthermore, research focusing on the biometric characteristics of signatures typically relies on a limited number of sensors. There is also a noticeable gap in existing solutions that investigate the application of multi-sensor setups for user identification through signatures. Other forms of handwriting have been used to a far lesser extent.

As mentioned, this dissertation aims to address the research gap by investigating a multi-sensor approach to provide a more detailed biometric profile by incorporating additional data sources. Additionally, while signatures represent a well-practiced action that can be consistently replicated, they are not the only form of handwriting that can serve as a biometric indicator. This research expands the scope beyond signatures by

exploring other handwriting forms, including sentences, words, and individual letters, for verifying individuals. Unlike signatures, which may be easily forged or imitated, natural handwriting behaviors exhibited in longer text forms (e.g., sentences and words) offer unique biometric traits that vary both within and between individuals. Incorporating these diverse handwriting forms facilitates a more comprehensive analysis of handwriting patterns, making the identification process more adaptable to inter-personal and intra-personal variations. Based on the identified research gaps and limitations in existing studies, the research hypotheses are defined as follows:

1. Besides signature, other forms of handwriting can be used for person identification.

2. The dynamic features extracted from touchscreen handwriting samples will show intra-person consistency and inter-person variability.

3. Deep learning models based on CNNs can be developed for person recognition based on touchscreen handwriting (both for stylus and finger input).

4. The implementation of a sensor fusion approach could improve the accuracy of person recognition based on touchscreen handwriting.

From these hypotheses, the research's scientific contributions are outlined and presented in the following manner:

1. A novel experimental apparatus that combines sensory devices, peripherals, and software support, to capture dynamic touch-based handwriting features.

2. An original dataset of biometric handwriting data, based on sensor fusion and comprising dynamic characteristics of signatures and other handwriting forms (short sentences, words, and individual letters) from stylus and finger input.

3. A deep machine learning model based on the early sensor approach for touchscreen handwriting recognition.

4. A method for the evaluation of the impact of individual sensor or sensor subset on the accuracy of handwriting-based person recognition.

1.5. Research Methodology

The research was carried out in three distinct phases, each directly contributing to achieving the scientific objectives of the study.

- The initial phase of the study involved a comprehensive review of the relevant research within the subject domain, encompassing both offline and online approaches. The review first covered studies that used sensors for verification purposes, where most research in the field of handwriting biometrics concentrated on differentiating authentic and forged signatures. The focus then shifted to works concerning handwriting-based person identification, which involved multi-class classification to determine the writer, with particular emphasis on online approaches that captured and analyzed dynamic handwriting features. This phase has revealed a research gap, which is further discussed in Section 2.3. of Chapter 2.

- The second phase of the study involved designing the empirical research methodology, configuring the experimental setup, and acquiring sensory data. An experimental apparatus for collecting dynamic touch-based handwriting features was developed based on the sensor fusion approach. This phase included testing individual sensors, both embedded in smart devices and external sensors, to determine their suitability for integration into the sensor fusion setup. The required software support was developed, and additional 3D-printed accessories were designed and created. Furthermore, an experiment was organized and conducted in a controlled setting, following the experimental practices and methods commonly used in the domain of Human-Computer Interaction (HCI). The collected biometric handwriting data were then categorized for further use. Permission for conducting the experiment and collecting data was requested and obtained from the Ethics Committee of the University of Rijeka, Faculty of Engineering. The first and second scientific contributions were achieved during this phase.

- In the third phase, the data obtained were subjected to suitable pre-processing. All sensory data were structured into separate vectors, with each vector containing specific measurements originating from a particular sensor. Further pre-processing

11

steps included the application of re-sampling and normalization techniques. The next step involved developing a deep learning model for feature extraction and classification tasks. To address the inherently nondeterministic nature of neural networks and their training process, both the feature extractor and classifier were trained and evaluated multiple times. Appropriate statistical tests were used to investigate the main effects of within-subject factors: input modality (finger, stylus), train set size, and handwriting form (signatures, sentences, words, and letters). An ablation study was conducted to analyze the impact of individual sensors or sensor subsets within the fusion-based setup. This phase culminated in the review and discussion of the methods proposed, employed, and assessed throughout the study. The last two scientific contributions were achieved through the activities conducted during this stage.

The research in this dissertation consists of distinct yet interconnected segments, some of which are based on the author's previous work. Certain sections of the study originate from the author's earlier research, specifically "A Sensor-Fusion-Based Experimental Apparatus for Collecting Touchscreen Handwriting Biometric Features" [32], "Beyond Signatures: Leveraging Sensor Fusion for Contextual Handwriting Recognition" [33], and "Augmenting Around-Device Interaction by Geomagnetic Field Built-in Sensor Utilization" [34]. These articles are publicly available and distributed under the Creative Commons Attribution (CC BY) license (`https://creativecommons.org/licenses/by/4.0/`), which permits the reproduction and modification of the content with proper attribution.

1.6.   Structure of the Doctoral Dissertation

This dissertation is organized into six chapters. The initial chapter, Chapter 1, provides an introduction to the field of biometric systems, narrowing the focus to the specific area of handwritten signature verification systems. It delves into both offline and online signature verification methods, outlining their specific applications. In addition to verification, this chapter also addresses handwritten signature identification, highlighting its potential for person recognition and its distinction from verification tasks. This chapter, along with the introductory contextualization, outlines the hypotheses and contributions of the dissertation and provides an overview of the employed methodological approach.

Chapter 2 presents a review of existing studies on handwriting-based biometric systems, focusing on their application in verification and identification methods. It reviews prior work on the methodologies, achievements, and limitations of verification systems, which aim to confirm the authenticity of a writer, as well as identification systems designed to determine the identity of an individual. The chapter concludes by identifying the research gap in person recognition via different handwriting forms, which forms the basis for the contributions of this dissertation.

Chapter 3 provides a comprehensive overview of the experimental setup and the collected data. It begins by outlining the sensor fusion technique applied in the study, explaining the rationale behind combining multiple data sources for handwriting-based person recognition. This is followed by a thorough description of the apparatus, describing the equipment and setup used to obtain dynamic handwriting data. Information about the participants is then presented, along with a review of the experimental design, methodology, and the procedural steps taken during data collection. The chapter concludes with an overview of the recorded sensor data, summarizing its structure and characteristics.

Chapter 4 focuses on the development and implementation of a person recognition system using a convolutional neural network (CNN) deep learning model. An introduction to CNNs is provided, explaining their architecture and relevance to the task of person recognition via handwriting. The acquired dataset used for training and testing is described, followed by details on data preprocessing techniques applied to prepare the dataset for analysis. In addition, the chapter discusses the utilized Dynamic Time Warping (DTW) method, along with the proposed Classification Potential (CP) metric. Feature extraction methods and the classification process are detailed, presenting the approach taken to recognize individuals based on the extracted features from their handwriting.

The overall results of the study are presented in Chapter 5, accompanied by a detailed discussion of the findings. The outcomes of the DTW analysis are shown, along with graphs that visually support the findings. Furthermore, the chapter presents the outcomes of investigating the main effects with regard to the influence of three different factors (handwriting form, input modality, and train set size) on the system's accuracy. Finally, findings of an ablation study are provided, analyzing the impact and contribution of individual sensors within the fusion-based setup.

Chapter 6 provides a summary of the dissertation, highlighting the main findings and contributions of the research. It also addresses potential limitations and suggests directions for future work in this field, drawing on the results of the research.

Chapter 2

REVIEW OF RELATED STUDIES

This chapter presents a thorough review of the existing studies related to handwriting biometric identification systems. The purpose of this review is to show a detailed survey of machine learning-based signature verification and identification systems across different aspects: datasets, preprocessing methodologies applied, techniques for feature extraction, models used for machine learning-driven verification and identification, along with the measures used to assess performance of the proposed systems.

The review of existing literature is structured into sections to categorize studies related to various approaches in this field. Both sections discuss relevant studies, with highlighted findings concentrating on single biometric modalities, specifically signatures or handwriting. This is closely aligned with the research in this thesis, particularly in relation to sensor fusion techniques for recognizing individuals through handwriting. It is important to emphasize that the review of related studies pertains to both offline and online verification or identification methods, although the research in this dissertation relates to the online approach.

The first section provides an overview of handwriting verification systems in this domain, with most works focusing on the various sensors used for verification purposes, aiming to differentiate authentic from forged handwriting samples. These systems address a binary classification task, determining whether a sample is authentic or forged. While this is a critical aspect of biometric security, many studies tend to overlook the broader context of user recognition across different handwriting forms.

The second section covers related work on handwriting identification systems, high-

15

lighting the various types of sensors, collected data and biometric features used for person recognition. While most existing research has primarily focused on handwriting verification, a notable research gap remains in the area of person recognition via handwriting. Instead of focusing solely on verification, this section encompasses studies related to handwriting identification tasks. These tasks refers to the general process of identifying the handwriting owner, which involves multi-class classification. Additionally, this section includes studies on analyzing handwriting features for individual identification across different handwriting forms, extending beyond just signatures. In comparison to signatures, this area has also been addressed in a fewer number of studies overall.

2.1. Review of handwriting verification systems

This section provides a review of works that exclusively focus on handwriting verification. It pertains to proposed online systems for collecting dynamic biometric data during writing, used to distinguish between genuine and forged samples. The paragraphs include a description of the devices and sensors employed for capturing handwriting characteristics. Furthermore, the discussion extends to common public datasets used in the research, outlining their characteristics and specifics. Moreover, the machine learning methods proposed for feature extraction and classification are described and detailed. The section also highlights the results achieved by various studies, presenting the advancements and findings of other researchers in this domain.

2.1.1. Radar and acoustics-based solutions

The literature presents a range of innovative solutions that utilize distinct sensor technologies to capture and analyze signature characteristics. For instance, mmSign [35] employs millimeter wave (mmWave) technology to detect hand movements during signing using the commercial mmWave radar. Utilizing advanced signal processing algorithms, mmSign derives time-velocity feature maps from recorded mmWave signals and employs a transformer-based model for authentication. The authors utilized new meta-learning strategy with task generation and data augmentation to improve learning from limited samples, enabling quick adaptation to new users. It achieves high accuracy with a small

number of labeled signatures while effectively resisting common forgeries. mmHSV [36] is an alternative approach that leverages mmWave radar to provide in-air handwritten signature verification system. The proposed method combines hand shape characteristics and writing process attributes to create electronic signatures. It models the biometrics of the handwritten signature process by extracting phase-dependent biometrics and behavioral attributes from the mmWave radar signals. A few-sample learning network is employed to integrate these features and verify users. Evaluations with commercial mmWave devices show that mmHSV achieves high accuracy and efficiency, maintaining a low false acceptance rate (FAR) at a fixed threshold and performing well with untrained users.

In addition to mmWave-based approaches, a variety of acoustic-based solutions have also been developed for signature verification. These systems leverage the unique characteristics of sound and vibration produced throughout the signing to capture subtle variations in signature dynamics. For instance, some methods analyze the frequency patterns and amplitude of pen strokes, while others use time-domain features of recorded audio to distinguish genuine signatures from forgeries.

The ASSV [37] is an acoustic-based solution designed to identify changes in signals generated by the movement of a user's hand and pen while signing. It employs a smartphone to capture these signal variations by emitting and receiving inaudible acoustic signals while the user writes on a standard piece of paper. A binary classification model employing a deep CNN is used to differentiate authentic signatures from forged ones, ensuring that no retraining is needed when adding a new user. The results demonstrate that ASSV achieves high performance with low latency. Although additional accessories are not required, the authors note that variations in the placement of the built-in speaker and microphone across smartphone brands present a limitation.

SilentSign [38] is a comparable solution that utilizes acoustic sensors (microphone and speaker) integrated in common smart devices to track the changes in distance of the pen tip while signing. It addresses certain constraints found in the proposed ASSV system, particularly concerning device orientation and the required distance between the smartphone and the writing surface. While employing a similar acoustic sensing method, it provides a larger signing area and eliminates the requirement for the smartphone to be aligned parallel to the signature surface. The authors of SilentSign do not employ neural networks in their system's architecture. Instead, they utilize traditional machine learning

classifiers: logistic regression (LR), naive Bayes (NB), random forest (RF), and support vector machine (SVM). After evaluating the performance of these classifiers, the authors determine that SVM is the most effective, demonstrating a high area under the curve (AUC) and a low equal error rate (EER).

SonarSign [39], similar to ASSV, also uses inaudible acoustic signals but introduces a more advanced verification system by incorporating an attentional multi-modal Siamese network. This system utilizes both the smartphone's speakers and microphones to send and receive a custom training sequence. It then records the resulting echoes to estimate the channel impulse response (CIR) based on the dynamic hand movements during the signing process. SonarSign demonstrates high effectiveness for unseen users, showing strong performance in real-world scenarios. SonarSign also addresses device variability limitations through an advanced spatio-temporal feature extraction technique, enhancing robustness compared to ASSV and SilentSign.

Sadak et al. [40] proposed an approach that is different from others, as it relies on the natural sound of the pen interacting with paper. It captures the friction sounds using only the internal microphone of mobile devices, without emitting any inaudible signals, providing a unique biometric data source. This method distinguishes genuine signatures from forgeries by employing dynamic time warping (DTW) to calculate similarity distances between audio signals. The authors indicate that EER values can vary considerably based on specific parameters of the signing process, such as the type of pen, paper, and recording device used. The proposed solution faces issues with background noise since the subtle pen-paper friction sounds are easily influenced by environmental factors. Additionally, using a single microphone without extra sensors makes the system sensitive to changes in writing style and device positioning.

2.1.2. Motion-based solutions

Regarding other types of sensors, related studies in this domain typically rely on accelerometers, either standalone or in combination with a gyroscopes. Most research focuses on the utilization of off-the-shelf devices with built-in accelerometers and gyroscopes, such as smartwatches or smartphones.

In their study, Li and Sato [41] investigated the use of commercially available smart-

watches for biometric verification through signature analysis. The authors employed the accelerometer and gyroscope of the smartwatch to capture motion data while participants sign on a tablet, aiming to explore conventional contact-based signature verification. The research use custom database, consisting of both authentic and forged signatures from 20 participants, with a total of 400 signatures recorded and aligned into 2900 pairs for analysis. Siamese recurrent neural networks (RNNs) are utilized to process sequential data, avoiding the necessity for manual feature extraction. The architecture transforms temporal motion data into fixed-length vectors, enabling efficient comparison and showcasing high performance.

Ramachandra et al. [42] developed a smartwatch-based user verification system that captures accelerometer data from two different smartwatch models. Similar to the approach in this dissertation, other forms of handwriting were utilized in addition to the signature. Specifically, the data was collected while 30 participants performed three tasks: signing their name, writing a predefined common sentence (*"you are smart"*), and writing a self-chosen sentence. Handwriting samples were captured from both paper and a tablet device to compare the effects of different writing mediums. The study utilized continuous wavelet transform (CWT) to extract time-frequency features and employed the pre-trained ResNet50 for deep feature extraction. These features were then classified using an ensemble of classifiers, with decisions fused using the SUM rule to improve verification accuracy. The researchers found that while signature-based verification was effective, writing a self-chosen sentence often led to improved results, particularly across different devices. However, there was some decline in performance.

Levy et al. [43] further used the accelerometer and gyroscope sensors to obtain dynamic motion data as participants signed physical documents. This included signing checks, credit card receipts, and vote-by-mail forms while wearing a smartwatch. Biometric data was collected from 66 participants, including both genuine and forged signatures, producing a dataset of 1980 recordings. The authors applied a machine learning model leveraging DTW and discrete cosine transform (DCT) for feature extraction and classification tasks, demonstrating notable results.

Using the same sensors (accelerometer and gyroscope), certain studies focused on in-air signature verification instead of standard writing methods on paper or tablet. In this approach, users perform signatures without physical contact, capturing dynamic data

from smartwatch or smartphone motion sensors.

In their follow-up studies, Li et al. [44, 45] explored in-air signature verification, allowing users to sign without direct interaction with paper or touchscreen. They utilized accelerometer and gyroscope sensors, alongside device attitude readings, to capture motion data as participants signed names using their fingers in the air. The research used a custom database with 22 participants, each contributing 10 authentic and 10 forged signatures, resulting in 440 signatures. RNNs are employed to process variable-length sequential data collected during signing. The recorded motion data is converted into fixed-length vectors to compare genuine and forged signatures. The system highlighted its effectiveness in authenticating users through in-air gestures and distinguishing genuine signatures from skilled forgeries.

Guo and Sato [46] utilized the same custom dataset for in-air signature classification using multiple CNNs. The researchers first applied the DTW algorithm to validate the dataset and then employed four CNN models (LeNet, AlexNet, VGG, and ResNet) to classify the signatures. The novelty of their approach involved transforming the time-sequential in-air signature signals into three-dimensional static representations for CNN input. Additionally, the study evaluated the performance of the CNNs with that of a transformer model, which was used as a baseline for time-sequential signal classification. Within the models tested, ResNet reached the highest accuracy when fine-tuned using the Adagrad optimizer. The research highlights that CNN models outperform traditional time-sequential models in handling this medium-sized in-air signature dataset, which showed signs of overfitting.

Shao et al. [47] introduced AirSign, an authentication system that combines built-in acoustic and motion sensors in smartphones for biometric verification through in-air signature gestures. The authors used the earpiece speaker to transmit inaudible signals and records echoes with integrated microphones to track hand geometry and signature motion, while the accelerometer and gyroscope capture hand movement during signing. The research involved thirty participants, each providing registered, genuine, and forged signatures during the data collection phase. AirSign utilizes multidimensional DTW for feature comparison, computing similarities between registered and input signatures. In addition to signature features, the system also integrates a hand geometry classifier and a motion classifier, combining the results to authenticate users. The results showed high

accuracy, validating its effectiveness without requiring additional hardware.

Buriro et al. [48] introduced an in-air signature authentication system that uses natural arm movements to authenticate users. Participants wore a smartwatch with integrated accelerometer and gyroscope sensors, which capture 3D motion data as they wrote their names in the air. The system processes this data utilizing DTW and a one-class multilayer perceptron (OMLP) classifiers. The authors explored two fusion levels, sensor and feature fusion. Testing with several participants revealed that the OMLP classifier with feature fusion outperformed the DTW-based approaches with a limited number of training samples.

Realme [49] is another solution for verifying handwritten signatures utilizing a wrist-worn MetaWear MetaMotionR smart device. It represents a commercially available rechargeable wearable device. The device features an integrated gyroscope and accelerometer to measure three-axis motion. These sensors record wrist movement while users perform genuine and fake signatures on plain sheets of paper, without requiring additional accessories such as a stylus. In the experiment, 10 participants provided 12 genuine and forged signature samples after a one-week gap. Several machine learning classifiers were evaluated to distinguish authentic signatures from forged ones, including SVM, decision tree (DT), RF, and the BayesNet algorithm. Among these, BayesNet proved to be the most effective, showcasing high classification accuracy and a low EER.

Several studies use Leap Motion sensor for contactless biometric authentication by tracking hand and finger movements in 3D space. These studies focus on signature recognition and verification, showing improved accuracy compared to traditional 2D methods by utilizing 3D spatial data and classifiers like k-NN, hidden Markov model (HMM), and least squares SVM (LS-SVM) [50, 51, 52]. Deep learning approaches, such as CNNs, have been utilized with Leap Motion and demonstrated high accuracy and low error rates in distinguishing genuine from forged signatures [53].

2.1.3. Smart pen solutions

Another approach to online signature verification involves using smart pens or styluses equipped with various sensors. Proposed solutions typically utilize off-the-shelf devices or custom pen-like systems featuring sensors such as pressure sensors, accelerometers, and

gyroscopes to capture detailed movement or pressure data while writing.

Lech and Czyżewski [54] proposed handwritten signature verification system including a wireless biometric pen and a $7''$ resistive touch screen integrated into an LCD display. Alongside a surface pressure sensor and two touch pressure sensors, the biometric pen features a three-axis accelerometer and gyroscope. The pen mimics the form of a ballpoint pen tip but does not dispense ink onto the paper. The signature verification process relies on six dynamic measures derived from the findings of the DTW method.

Subedi et al. [55] proposed a sensorized writing stylus with two off-the-shelf accelerometers to capture hand movements during writing. The system, using a Raspberry Pi for data collection, extracts spatial and temporal features like acceleration and stroke direction. Tested with three participants, including genuine and forged signature samples, the system achieved high classification accuracy using an MLP with one hidden layer. In follow-up research [56], the authors integrated two triaxial accelerometers within the stylus, eliminating external mounts. Using a larger dataset of ten participants, the system employed a more advanced MLP with three hidden layers. Despite the increased complexity and dataset size, the system's accuracy was slightly lower, attributed to the variability in the signatures resulting from the more diverse dataset.

Kurowski et al. [57] developed an automated system for biometric handwritten signature verification. The system uses a specialized electronic pen equipped with sensors such as pressure sensor, a 6-axis gyroscope-accelerometer, a 3-axis inclinometer, 2 MEMS (micro-electromechanical systems) microphones, speaker, and additional sensors. The pen is compatible with both computer and mobile phone screens, capturing both dynamic and static signature features. To extract relevant signature features, the authors used a convolutional neural network (CNN) trained with the triplet loss method, which generates fixed-length embeddings that are compared using Euclidean distance. The system was evaluated on a large dataset of signatures and demonstrated low error rates in detecting both random and skilled forgery attempts.

Zhou et al. [58] presented a solution for verifying handwritten signatures that combines both offline and online verification techniques. Specifically, during the signature-writing process, both the offline image and online data of the signature are collected in real-time. A smart pen, equipped with a camera and pressure sensor, allows for the simultaneous collection of both the physical signature and digital data. The static and dynamic features

of the signature are extracted and verified using SVM and DTW, respectively.

2.1.4.   Existing datasets

Several publicly available datasets have been widely utilized in handwriting biometric research, serving as benchmarks for evaluating signature verification models. Among the most frequently used are SigComp 2009 [59], SVC-2004 [60], MCYT-100 [61], SUSIG [62], and BiosecurID [63]. A common trend among researchers leveraging these datasets is the utilization of CNN architectures for developing verification systems [64, 65, 66]. Data augmentation techniques are often employed, addressing the challenge of limited sample sizes in certain datasets [67, 68]. In addition to conventional CNN-based approaches, some studies focus on lightweight CNN architectures specifically designed for mobile-based signature verification, enhancing performance for resource-constrained devices [69, 70]. Furthermore, other research investigates hybrid models that combine CNNs with RNNs or transformers to improve feature extraction and classification accuracy [71, 72].

The related studies described so far focus on verification systems. The following chapter provides an overview of research related to handwriting identification systems.

2.2.   Review of handwriting identification systems

Given the focus on person recognition via handwriting in the dissertation, a comprehensive review of relevant literature in this domain is provided. It is important to highlight that, as mentioned earlier, there are notably fewer studies addressing person identification through handwriting in comparison to works related to verification. The reviewed works rely on the use of one or more distinct sensors for person identification purposes, examining various technologies and approaches used to capture and analyze handwriting dynamics. While the review covers studies using various types of neural networks, the emphasis is placed on those employing CNN architecture. The review supports the use of CNN in this thesis, while also considering other relevant neural networks that contribute to this field. The aim is to contextualize the current advancements and techniques in person identification via handwriting.

2.2.1. Machine learning-based approaches

Machine learning-based approaches to signature identification rely on explicit feature extraction techniques prior to classification. Once these features are selected and extracted, a standard set of classifiers is applied, utilizing well-established models.

One of the common feature extraction techniques in offline signature recognition is the use of texture-based feature extraction methods or image texture descriptors, such as histogram of oriented gradients (HOG), local binary patterns (LBP), and binary statistical image features (BSIF) which capture texture and gradient variations within the signature images [73, 74, 75]. Feature extraction techniques such as scale-invariant feature transform (SIFT) and speeded-up robust features (SURF) have also been investigated to extract distinctive keypoints from signatures [76].

While offline methods primarily rely on images, online signature identification benefits from dynamic handwriting features. Research using sensor-based devices like Wacom tablets has explored handwriting kinematics by extracting time-series representations of pen movements. Across multiple studies, a variety of machine learning classifiers have been used to distinguish individuals based on their handwriting characteristics. The most frequently employed classifiers include SVM, RF, and LR [77, 78]. Additionally, other classifiers like KNN and DT are commonly used, with some studies further incorporating ensemble methods such as the extra tree classifier (ETC) to improve accuracy [79].

Numerous studies use publicly available datasets like CEDAR and MCYT, enabling comparisons across various research works. Additionally, some studies introduce custom datasets tailored to specific writing styles or languages. For example, certain research focuses on non-Latin scripts, such as Devanagari and Kirgiz [80, 81]. Other studies combine both public and private datasets, merging well-established benchmarks with newly developed multilingual collections that include Chinese, Uyghur, and Kazakh [82, 83]. Additionally, some studies focus on Arabic handwriting, such as the AHWDB1 and AHWDB2 datasets used for writer identification [84].

The authors in [85] tested both ML and DL models for handwriting-based person identification. They evaluated 14 ML algorithms, such as SVM and Random Forest, alongside 7 DL architectures, including CNNs and BiLSTM. The results showed that both ML and DL methods were effective for handwriting-based person identification.

2.2.2.   Deep Learning-Based Approaches

Deep learning models have become a widely utilized and advanced technique for signature-based person recognition and handwriting identification, particularly leveraging CNN models and hybrid architectures.

Multiple studies employ pre-trained deep learning models, benefiting from transfer learning by fine-tuning well-established architectures like ResNet, EfficientNet, and GoogleNet on custom collected datasets [86, 87]. Some research focuses on custom CNN architectures designed for signature classification. LS2Net [88] and CNN-GC [89] represent novel CNN architectures developed for large-scale signature identification and trained on publicly available datasets. CNN models with modified input representations have been employed to accommodate 3D air signatures captured via Leap Motion sensors [90]. Similarly, multi-scale CNNs have been designed for in-air gesture signatures, using parallel convolutional layers to enhance feature extraction and achieving higher accuracy than standard pre-trained models [91]. Research has explored handwriting classification beyond signatures, for instance, employing LeNet to classify the handwriting of different individuals based on a proprietary dataset [92].

Beyond individual CNN models, multi-branch architectures have been proposed to improve feature extraction at different scales. One approach combines ResNet-18 with a feature pyramid network (FPN), enhancing classification accuracy by capturing features at various scales [93]. Similarly, another study employs CNNs as a baseline model alongside capsule networks, achieving notable accuracy in both verification and identification tasks by retaining spatial relationships within the image data [94].

Regarding hybrid approaches, a hybrid CNN-BiLSTM model has been introduced for Bengali handwriting identification [95]. The CNN component extracts spatial features, while the bidirectional long short-term memory (BiLSTM) captures sequential dependencies in handwriting data collected using a stylus. The hybrid architecture notably improved classification accuracy over standalone CNN or LSTM models.

An alternative approach integrates a CNN with a spatio-temporal siamese neural network (ST-SNN), using a dual-branch architecture to extract both spatial and sequential patterns from air-written signatures [96]. The approach proved highly effective for 3D motion-based handwriting analysis, utilizing a Leap Motion sensor to track hand move-

ments in three-dimensional space. The custom-designed architecture outperformed traditional machine learning techniques such as SVMs, HMMs, DTW, and RF classifiers.

The studies presented in this subsection are summarized in Table 5.13 in Section 5.6., highlighting the results and comparing them with the author's findings in this dissertation.

2.3. Research gap

This dissertation addresses several notable gaps in handwriting biometric systems research. The primary areas of inquiry that this study aims to explore and contribute to include:

- Focus on verification over identification:

  - The majority of the reviewed studies in the field of handwriting biometrics are centered on verification systems. These approaches involve distinguishing between genuine and forged handwriting samples, representing a two-class classification problem where a decision is made about whether a sample is authentic or forged. In contrast, this study addresses the research gap by switching the focus from verification to handwriting recognition, which involve the general process of detecting the handwriting owner through multi-class classification.

- Narrow scope of handwriting forms:

  - Related studies often miss the wider context of user recognition across different handwriting forms. The emphasis tends to be primarily on gathering user signatures, overlooking the potential for identifying users through other forms of handwriting, such as sentences, words and individual letters. Accordingly, the study investigates the feasibility of recognizing a person through various forms of handwriting, without relying exclusively on signatures, expanding the analysis to general handwriting.

- Lack of sensor fusion approaches:

  - This study incorporates the sensor-fusion concept, setting it apart from similar studies that primarily rely on a single type of sensor or device. The objective was to introduce a broader range of sensor readings into the analysis of

handwriting dynamics. The measurements were mainly sourced from standard smart devices and integrated sensors, rather than specialized hardware. Additionally, data on handwriting dynamics is collected not only for stylus input but also for finger input, which remains the primary method of interacting with touchscreen devices. This broadens the study to include an analysis of the proposed solution regarding the input modality.

Chapter 3

EXPERIMENT AND DATA DESCRIPTION

This part of the dissertation outlines a detailed overview of the experimental setup. The proposed apparatus is described, with each of the devices, sensors, and accessories used in the configuration thoroughly detailed. Additionally, their specific functions and roles within the system are explained. Following that, the procedures employed in the conducted experiment are described. Finally, the data collection process is presented, including the obtained data types and technologies utilized for their acquisition.

3.1. Sensor fusion setup

Research within the domain of handwriting verification or identification primarily focus on employing a single device or sensor. In contrast, this study emphasizes leveraging sensor fusion capabilities to facilitate user identification based on their handwriting. Sensor fusion involves combining readings from multiple sensors, providing advantages compared to using a single data source [97]. This approach can address the constraints of individual sensors, such as limited coverage or susceptibility to noise.

A diverse array of sensors was integrated into the proposed apparatus, considering all devices at disposal. External sensor modules were added to complement the integrated sensors on these devices. The objective was to evaluate how sensors, whether used individually or in groups, affect person recognition accuracy through different handwritten

29

Table 3.1: An overview of the sensors used in the proposed apparatus.

Sensor	Device	Data
Accelerometer	Smartwatch	Acceleration in three orthogonal directions (m/s^2)
Camera	Smartphone	Handwriting video for ArUco marker tracking, 720p resolution and 30 FPS
Gyroscope	Smartwatch	The rotation rate around the x, y, and z axes, in degrees per second (°/s)
Magnetometer	Tablet computer	The magnetic field for all three physical axes (μT)
Piezoelectric sensor 1, model PZT-5	External, connected to the tablet using a USB sound card and adapter	Electrical output in response to mechanical strain (V)
Piezoelectric sensor 2, model FT-50T-3A1	External, connected to the smartphone using 3.5 mm audio connector	Electrical output in response to mechanical strain (V)
Touchscreen	Tablet computer	Touch position, size and orientation of contact area, pressure, interaction type

forms. This facilitates determining the most effective set of sensors for user identification, using not only signatures but also handwriting samples of sentences, words, and individual letters. Employing sensor fusion aimed to investigate the potential for achieving enhanced accuracy, robustness, stability, reliability, and other beneficial characteristics in comparison to single-sensor systems [98]. A comprehensive list of the devices used, along with both internal and external sensors, and the data gathered by the sensors, is provided in the Table 3.1.

In the subsequent chapter, a thorough overview of individual devices and sensors is provided, outlining their characteristics and roles within the research context.

3.2. Apparatus description

This subsection presents a description of the devices, sensors, and additional accessories utilized in the apparatus. The presented details include device models, sensor specifications, technical characteristics, and the utilized software tools and platforms. The proposed system depicted in Fig. 3.1 encompasses a tablet, a smartphone, a smartwatch, and two external piezoelectric sensors. The absence of a single device integrating all essential sensors was the inherent limitation requiring the utilization of multiple devices.

Figure 3.1: The experimental setup depicting all devices and sensors, including a tablet, a smartphone, a smartwatch, and two external piezoelectric sensors.

The specific devices employed in research were the only equipment available to the author and are interchangeable with others of similar function. Additionally, certain software restrictions imposed by the device manufacturers, as described later in this subsection, could have simplified the apparatus implementation.

All devices are interconnected and communicate through wireless networking (Wi-Fi) and Bluetooth communication protocols. Custom applications for the Android and Tizen operating systems have been created for synchronizing the devices, thereby enabling reliable data gathering. The system offers a high level of adaptability, enabling the selection of various settings throughout the experiment.

A perforated plate with dimensions of 120 by 60 centimeters functions as the mounting platform for individual components of the apparatus. The panel itself is securely attached

to the desk to prevent any potential movement while the participant is engaging in handwriting activities. In addition to using screws, a variety of metal brackets, holders, and plates were utilized to securely position the devices in their assigned locations. The reason for securing the equipment at a fixed position is to ensure consistency, reliability, and control over the experimental setup. This enables subjects interacting with the apparatus under the same conditions, thereby mitigating variation within the experimental setting. The strategy reduces the probability of accidental interruptions during the experiment and enables accurate data to be obtained.

3.2.1. Tablet sensors

The primary component employed in the proposed apparatus was the tablet computer. It was manufactured by Samsung Electronics and pertains to the specific model released in 2019 named Galaxy Tab S6. The unit utilized in the study features a Qualcomm SM8150 Snapdragon 855 system-on-chip and 6 gigabytes of RAM. The device is initially distributed with the Android Pie operating system, utilizing the One UI 1.5 user interface. At the time of writing, the most recent software version available for update is Android $12L$, in conjunction with One UI 4.1.1. Given the specific design of the applications for the initial versions, the upgrade was delayed to retain the original system version.

The device is equipped with a large 10.5-inch touchscreen AMOLED display, providing sufficient screen area for a variety of tasks. A screen exhibits a resolution of 1600 by 2560 pixels at ~ 287 pixels per inch (PPI) density and $16 : 10$ aspect ratio. These characteristics make it particularly suitable for handwriting input, seamlessly accommodating both stylus and finger interaction.

Aside from touchscreen sensor, the device incorporates a multitude of integrated sensors. Information about these components was obtained using three specific applications: Sensors Toolbox, Sensors Multitool, and Device Info: System & CPU Info. Information from the applications was utilized to find detailed functionalities and capabilities of the available sensors for collecting measurement data.

The magnetometer embedded into the tablet device itself was of particular significance for the research. While primarily devised for orientation detection or navigation tasks, the author's study [34] showcased its adaptability for drawing or writing with a strong

Table 3.2: Technical specifications of the AK09918 integrated circuit (IC), a 3-axis electronic compass featuring advanced high-sensitivity Hall sensor technology. This IC is integrated into the Samsung Galaxy S6 Tab tablet device utilized within the apparatus.

Model	AK09918C
Sensitivity $[\mu T/LSB]$	0.15
Data output	16-bit data out for each 3-axis magnetic component
Analog power supply voltage (min.) $[V]$	1.65
Analog power supply voltage (typ.) $[V]$	1.8
Analog power supply voltage (max.) $[V]$	1.95
Current consumption (typ.)	
• power-down mode $[\mu A]$	1
• measurement mode $[mA]$	avg. current consumption at $100Hz$ repetition rate: 1.1
Serial interface	I^2C
Operating temperature $[°C]$	-30 to $+85$
Device package	WL-CSP
Number of pins	4
Footprint $[mm]$	$0.76 \times 0.76 \times 0.5$

permanent magnet. Intentionally disrupting the magnetic field with powerful magnets while writing using a stylus or finger results in a unique pattern for each user. According to the official datasheet available on the Asahi Kasei Microdevices Corporation website, the Galaxy Tab S6 model incorporates the AK09918 3-axis electronic compass [99]. The integrated circuit (IC) generates a 16-bit data output for each of the 3-axis magnetic components. It features a 0.15 μT/LSB measuring sensitivity, with maximum measurement range of 4192 μT. Detailed technical specifications of the magnetometer sensor are presented in the Table 3.2.

3.2.2.   Stylus and 3D-printed ring

In addition to the smart devices, the proposed apparatus include a variety of supplementary accessories. A pair of neodymium magnets were employed to induce a disturbance in the magnetic field. Each magnet possesses a residual magnetization (remanence) of 1.19 Tesla and belongs specifically to the N-38SH grade. The magnet's dimensions are expressed as follows: length, width, and height measuring $2 \times 10 \times 6$ mm, respectively. For enhanced magnetometer readings in the presence of magnets, they are arranged in a vertically stacked configuration, one atop another.

To ensure the magnets were positioned correctly on the stylus and the user's finger during writing, custom-made holders were employed. The holders were designed utilizing specialized software and then fabricated using 3D printing technology. For the stylus, the holders were constructed as housings that precisely correspond to the thickness of both the magnets and the pen to prevent any movement during writing. Alongside the magnet holders, specialized mounts were also printed to securely affix ArUco markers of the required sizes. The first two items in the upper row from the left in the Fig. 3.2 are specifically designed 3D-printed holders for ArUco markers, intended for attachment to tip of the stylus (S Pen). In the same figure, the holder for two stacked magnets attached to the bottom of the stylus is shown as the first on the right.

A ring-shaped holder was employed for writing with the finger, specifically using the index finger. Different sizes of rings were produced based on available online templates, providing a personalized fit for each person. Each of the holders is labeled with a unique
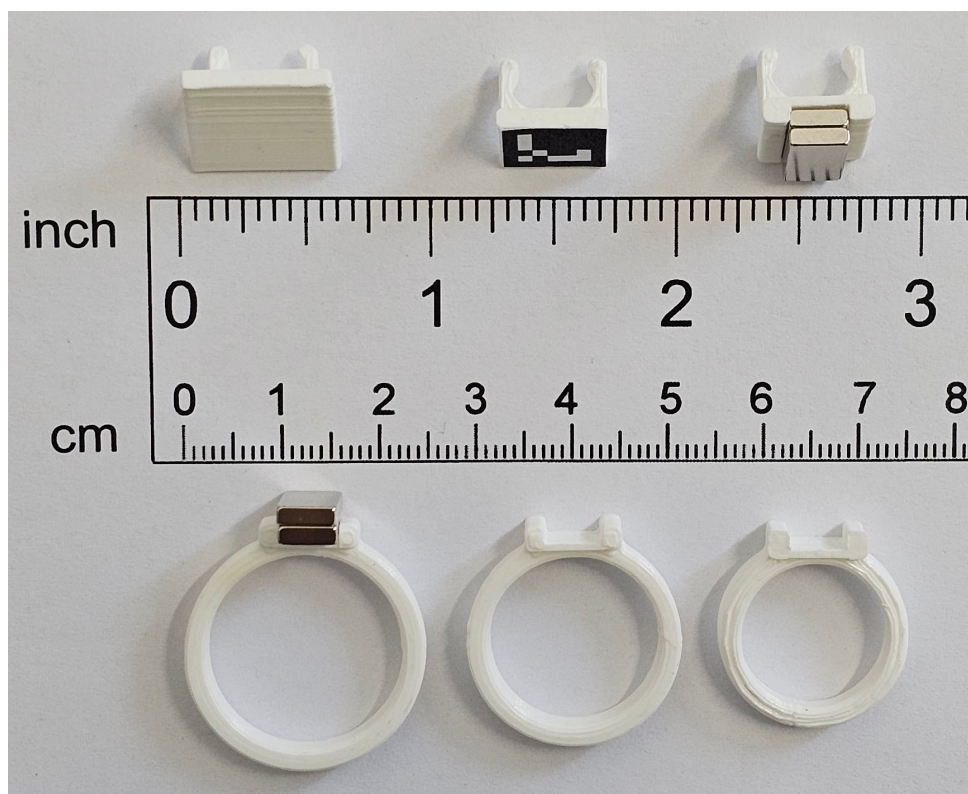
Figure 3.2: Upper row: the first two on the left are custom designed 3D printed mounts for ArUco markers of different sizes for the stylus (S Pen). The holder for two stacked magnets that attaches to the bottom of the stylus is shown as the first on the right. Bottom row: ring-shaped holders for two stacked neodymium magnets of various diameters utilized for finger writing activities.

label to ensure the proper orientation and alignment of the magnets. The described circular holders for two stacked neodymium magnets, varying in diameter and designed for finger writing, are displayed in the bottom row of Fig. 3.2.

The tablet device employed in the apparatus is equipped with a stylus accessory designed specifically for use with Samsung Galaxy tablets and smartphones. The stylus known as the S Pen offers advanced functions such as pressure sensitivity, air gestures, and tilt detection. It enables users with precise input capabilities, facilitating activities such as writing, drawing, navigating, and interacting with the device interface. Notably, the S Pen features a sensitivity range of 4096 discrete levels, which varies based on the applied pressure. Nevertheless, the stylus represents a specific limitation that is relevant to this research.

The constraint particularly pertains to the S Pen Remote Software Development Kit (SDK). This SDK enables the utilization of advanced functionalities of the S Pen, including button interactions, movement tracking, and gesture recognition. However, based on a reply from the official Samsung Developers forum [100], the S Pen Remote SDK does not provide developers the capability to collect pressure levels and tilt data. Regardless of this limitation, it remains possible to utilize the MotionEvent object [101] provided by the Android framework to obtain data regarding stylus pressure levels. It represents an event detailing a user interaction with the screen, utilizing both the stylus and the finger. The object contains details concerning the event, including the type of action performed, event coordinates, pressure, size, orientation, and additional properties determined by the event type. Although certain data can be obtained during finger interaction, pressure information for finger writing was unavailable with the utilized tablet.

To address the SDK constraint and enable the data collection regarding the tilt of the stylus, an ArUco marker was utilized. It represents a square marker featuring a specific black and white pattern that cameras detect and track, providing precise 3D pose and orientation information [102]. The function of the marker in this research is to acquire data on the inclination of the S Pen during writing. Although feasible, ArUco markers were not utilized during finger writing in this study. The reason for this is the challenges posed by finger writing, notably marker occlusion, occurring when the marker is obstructed or partially obscured by the finger. To attach the marker to the stylus during writing, custom mounts were created using 3D printing technology. Holders of various sizes were
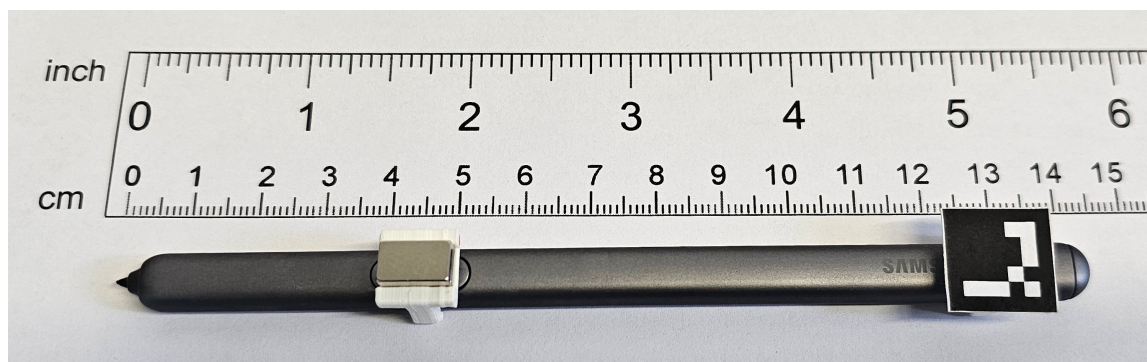
Figure 3.3: The S Pen stylus equipped with both a magnet and an ArUco marker, securely held in place by custom-designed holders created using 3D printing technology.

designed and created to accommodate different dimensions of ArUco markers. Multiple designs were tested to determine the most appropriate option, resulting in the selection of a 15×15 mm variant utilizing the original ArUco dictionary. The marker was printed on a piece of paper with the appropriate dimensions, adhered to the holder and then secured to the top of the stylus. The S Pen stylus equipped with the described magnet holders and a mount for the ArUco marker is shown in Fig. 3.3.

3.2.3. Piezoelectric sensors

Regarding external sensory components, the proposed apparatus includes two piezoelectric sensors. These devices function as audio capture instruments, similar to conventional microphones. However, in contrast to microphones, these sensors are not affected by distant sounds and are suitable for capturing noises while writing. Their research purpose is to gather sounds from the tablet surface corresponding to the momentary pressure exerted to the screen during writing. Additionally, it is feasible to detect the vibration patterns and directional changes arising from the stylus or finger moving across the display. The recorded audio signal therefore includes data pertaining to the gestures performed on the touchscreen, such as gliding, swiping, tapping, and pressure application. For the purposes of this study, the piezoelectric sensors are capable of delivering the required bandwidth and dynamic range for the audio signal. The analog-to-digital conversion capabilities of both devices were presumed identical, and thus, no additional analysis or examination was conducted.

Sensors operate based on the piezoelectric effect, in which particular materials, like

quartz or specific ceramics, produce an electric charge. An electrical output is generated when mechanical force or stress is applied to the piezoelectric crystal. Different pressure levels on the screen during writing produce signal variations, enabling recognition of unique user patterns.

The generated signal is obtained using the tablet integrated microphone input. The limitation of the device is the absence of a physical 3.5 mm audio connector. To overcome this constraint, an external USB 2.0 sound card dongle manufactured by Delock was utilized for sound acquisition. The adapter is equipped with a USB 2.0 Type-A male connector and offers connectivity options for both microphone input via a 3.5 mm mono jack and speaker output through a 3.5 mm stereo jack. It incorporates the CMedia CM119 chipset, ensuring low power consumption and reliable audio processing. The device features a signal-to-noise ratio (SNR) of approximately 90 dB and supports a sample rate of 16-bit / 48 kHz. The sound card proved suitable for research purposes as it does not require an external power source.

A USB-C to USB-A adapter was used to connect the sound card to the tablet. Since the apparatus includes another audio recording device, a smartphone was also used to incorporate an additional piezoelectric sensor. The device model utilized includes a 3.5 mm audio input interface, hence an adapter was unnecessary in this case.

The Fig. 3.4 depicts two distinct models of piezoelectric sensors utilized within the proposed apparatus. Both sensors are positioned beneath the tablet, with one situated in the bottom left corner and the other in the bottom right corner. Two distinct types were selected, varying in shape, dimensions, and technical specifications. The devices differ in their output signal strengths, frequency spectra, dynamic ranges, and sensitivity levels. In addition to the two mentioned models, a variety of other piezoelectric sensor types have been tested during the research. These types involved different shapes and dimensions, including those in the form of a rod as depicted in Fig. 3.5. Moreover, there are more expensive sensor variants available on the market, featuring larger surface areas, higher sensitivity, and increased capacity. However, considering the specific objectives of this study, the selected piezoelectric devices nonetheless provide acceptable results.

Both piezoelectric sensors employed in the apparatus are described in detail below. The first model utilized is rectangular in shape, featuring a metal plate made of elastic steel and larger ceramic dimensions measuring 50×30 mm. The second model is circular,
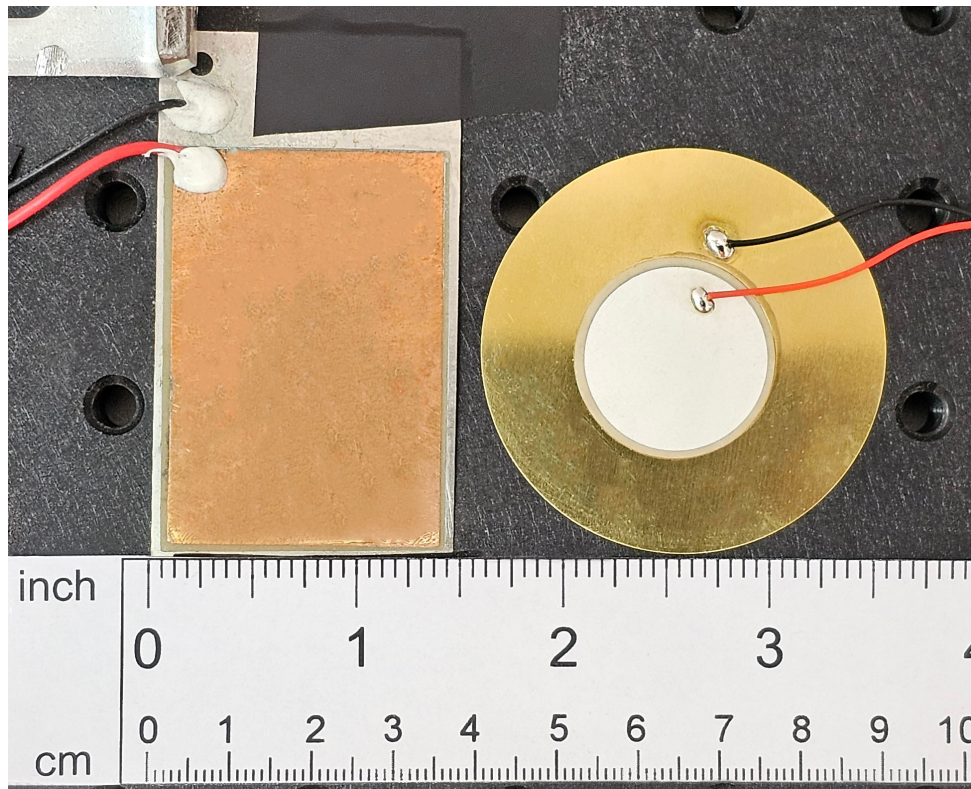
Figure 3.4: The square PZT-5 and the round FT-50T-3A1 piezoelectric sensors models in the apparatus, positioned beneath the lower left and right edges of the tablet, respectively.
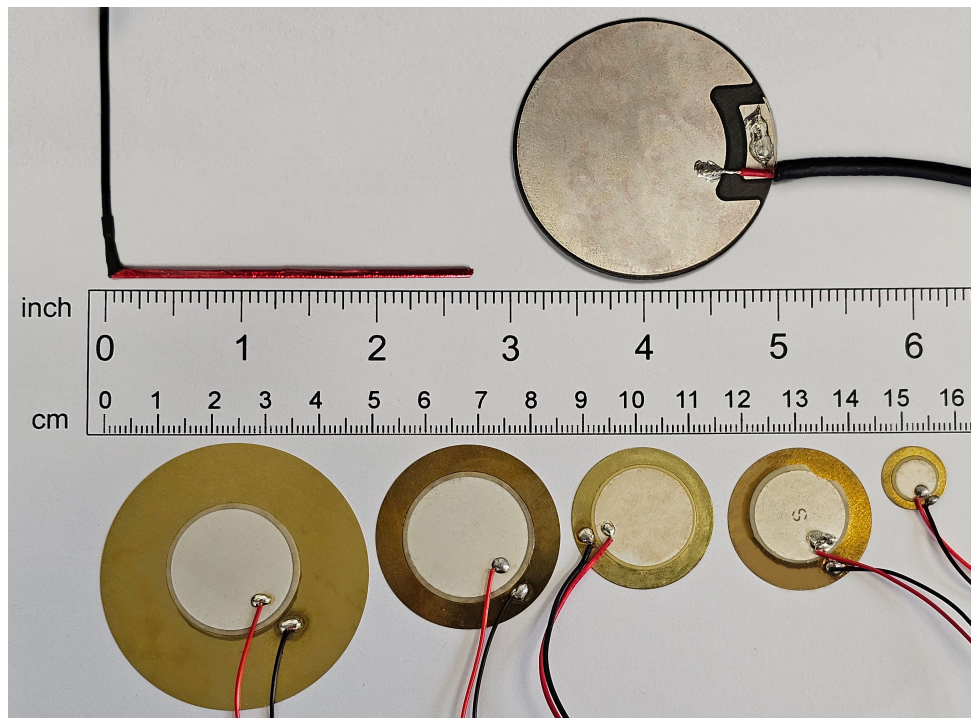

Figure 3.5: Comparison of different piezoelectric sensor models with varying sizes and shapes, tested in the apparatus during the research.

with a metal plate made of brass and a ceramic diameter of 25 mm.

According to the provided data, the first PZT-5 square model is characterized by a larger attached piezo crystal and a higher current rating. As a result, it is expected to provide a stronger signal, a hypothesis that has been experimentally confirmed. This device features a higher resonant frequency, making it more sensitive to higher frequencies. In contrast, the FT-50T-3A1 model demonstrates a lower resonant frequency and is more responsive to lower frequencies. A more detailed overview and comparison of the characteristics of the piezoelectric sensors are provided in the Table 3.3.

A coaxial mono audio cable was employed to facilitate the connection between the sensors and the tablet or smartphone. These types of cables are commonly utilized to connect devices such as microphones, musical instruments, or other audio sources to amplifiers, mixers, or sound systems. The single center conductor of the coaxial cable was soldered to the center of the piezo contact surface to facilitate signal transmission. The shielding ground conductor was subsequently soldered to the metal contact substrate of the piezo sensor. Furthermore, a standard 3.5 mm mono audio plug is soldered to the

Table 3.3: Characteristics comparison of the two distinct piezoelectric models employed as sound vibration and pressure input sensors.

| Model	FT-50T-3A1	PZT-5
Shape	Round	Square
Metal plate material	Brass	Elastic Steel
Contact material	Silver	Copper
Metal plate dimensions $[mm]$	$\Phi 50$	65×37
Metal plate thickness $[mm]$	0.18	0.25
Metal plate area $[mm^2]$	1963	2405
Metal plate volume $[mm^3]$	353.34	601.25
Ceramic dimensions $[mm]$	$\Phi 25$	50×30
Ceramic thickness $[mm]$	0.27	0.2
Ceramic area $[mm^2]$	490.9	1500
Ceramic volume $[mm^3]$	132.54	300
Total thickness $[mm]$	0.45	0.45
Operating temperature range $[°C]$	-20 to 70	-20 to 70
Resonant frequency $[kHz]$	3.2	28
Maximum voltage $[V]$	15	12
Resonant impedance $[\Omega]$	≤ 200	1002
Static capacity $[nF]$	4.2	102
Maximum current $[mA]$	3	8
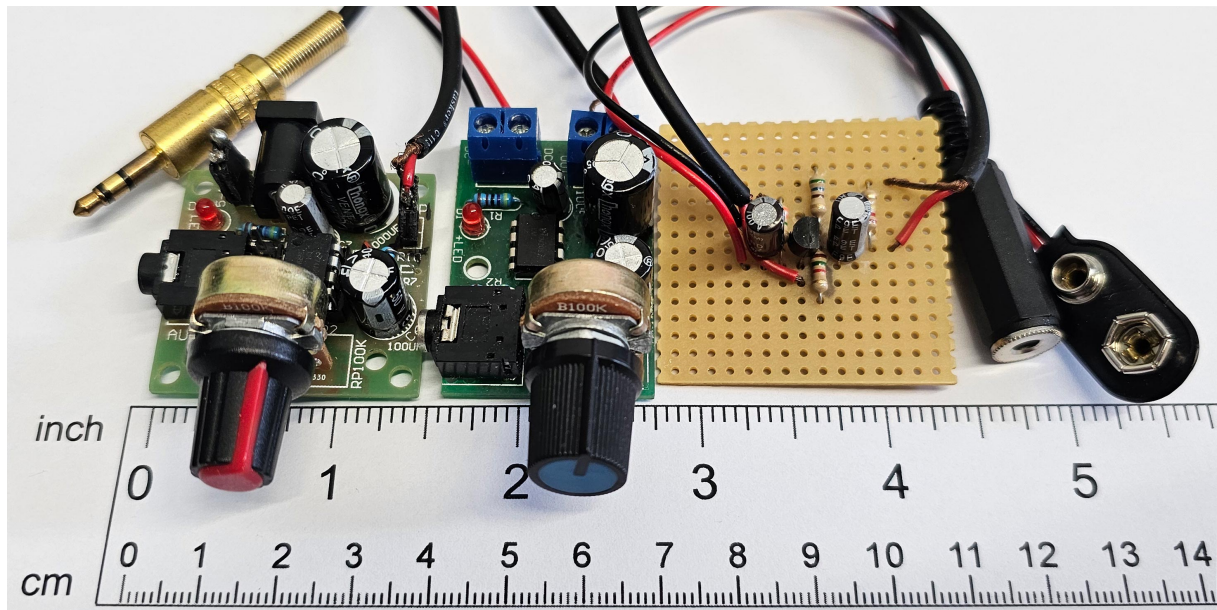
Figure 3.6: Various audio amplifiers were tested during the research in an attempt to amplify the signals obtained from the piezoelectric sensors.

opposite end of the coaxial cable for device connectivity. In this case, the inner conductor of the cable was soldered to the tip terminal, and the ground conductor to the sleeve terminal of the audio connector.

During the research, three different amplifiers were tested in an attempt to amplify the obtained signal from both piezoelectric sensors. The objective was to investigate how various amplifier models affect signal enhancement. This was particularly relevant in scenarios where the acquired signal exhibits lower intensity, such as during finger-writing tasks. In addition to signal amplification, it was required to minimize the levels of noise introduced into the signal. The Fig. 3.6 shows several models of audio signal amplifiers that were examined in this study. All amplifier models operate on a DC power supply ranging from 3V to 12V, capable of delivering output power ranging from 0.5W to 12W. Audio input is accepted through a standard 3.5 mm audio jack for connecting sensor.

For testing amplifier performance, sensor was used in various scenarios: without amplifiers and in three different cases using distinct amplifier models. The Fig. 3.7 showcases the data obtained from the FT-50T-3A1 piezoelectric sensor model described earlier. It is divided into four segments, each displaying a specific waveform corresponding to a particular use case. The x-axis represents time in seconds, while the y-axis indicates the amplitude of the sensor readings. The amplitude of signals varies significantly across the

(a) Raw piezo sensor signal, no amplification.

(b) Low amplification level

(c) Moderately amplified signal.

(d) Highest amplification level.
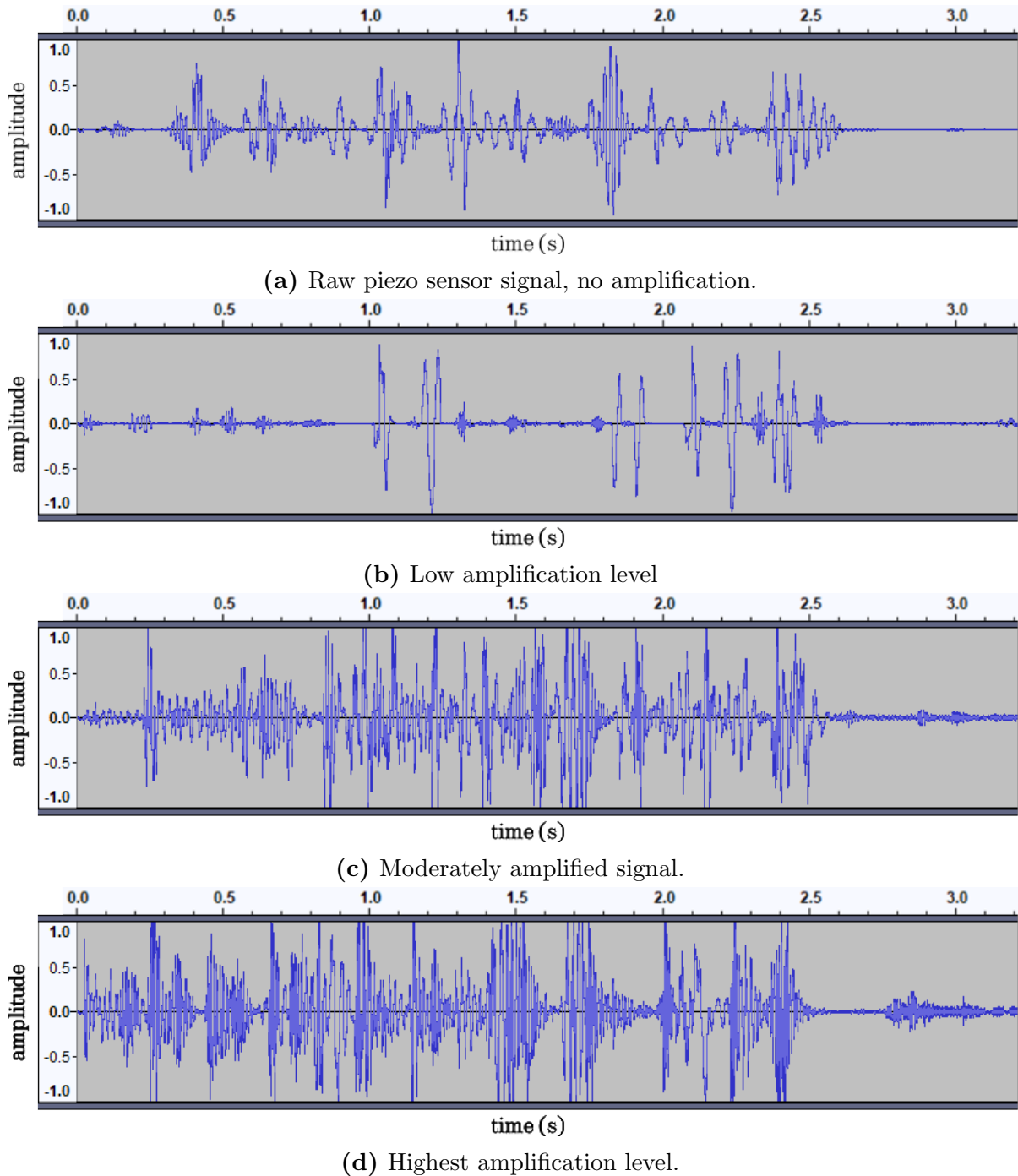
Figure 3.7: Waveforms of piezoelectric sensor signals at different amplification levels, illustrating variations in amplitude, signal detail, and noise. The amplitudes have been normalized, and the signals are displayed using Audacity software.

different segments, along with the level of noise present in the signals. The readings pertain to the vibrations and changes in pressure detected on the screen during writing the same word using a stylus.

The very top segment (Fig. 3.7a) displays the raw readings obtained from the sensor

without any amplification. The acquired signal lacks the pronounced peaks observed in the amplified measurements. When compared to the third and fourth segments (Fig. 3.7c and 3.7d), it exhibits signal of lower amplitude and sparser distribution with occasional spikes. Nevertheless, the waveform demonstrates a notably cleaner profile maintaining a relatively low noise level. The second segment (Fig. 3.7b) shows a less detailed waveform with lower amplitude and fewer smaller spikes compared to the first segment. The signal appears less cluttered with fewer high-frequency components, featuring smoother transitions between peaks and troughs. This suggest that the amplification level may not be able to adequately capture finer details of stylus or finger writing motions. Here, an amplifier operates akin to a trigger system activated by a strong spike, whereas in other instances, the signal remains unamplified. There is a certain level of noise present, although it is less prominent compared to the third and fourth segments. Finally, the third and fourth segments (Fig. 3.7c and 3.7d) demonstrate more prominent signal features, indicating stronger amplification. With increased amplitude, the signal reveals finer variations that were previously too weak to detect. However, this also amplifies background noise, making it more noticeable and introducing small oscillations that may not be part of the original signal. The fourth segment, distinguished by its highest amplification, displays the greatest amount of noise. However, it also reveals the most detailed signal characteristics. Both waveforms exhibit dense and high-amplitude signals with prominent peaks. In this context, the noise is more pronounced, manifesting as increased fluctuations and irregularities in the signal. This suggests that excessive amplification may lead to distortion, making it harder to distinguish meaningful signal features from noise.

Considering all the points discussed, it has been concluded that utilizing the raw piezoelectric signal is an adequate and appropriate option for research purposes. The measurements obtained during writing provides sufficient details regarding the actions of gliding, tapping, and applying pressure, whether on the touchscreen or the tablet itself. The sensor readings, without the use of amplifiers, were suitable to capture these actions while avoiding the noise introduced by amplifiers. Despite the absence of amplifiers, the apparatus incorporates two different models of piezoelectric sensors, each characterized by unique technical specifications. This enables improved tracking of pressure and vibration dynamics. Moreover, it extends the sensitivity range, facilitating the detection of a broader spectrum of signal intensities and variations.

3.2.4. Smartwatch sensors

The Samsung Electronics smartwatch has been incorporated into the apparatus to facilitate the acquisition of additional data from its integrated sensors. The available device model used in the research was the Galaxy S3 Frontier, which was introduced in 2016. It operates on version 4.0.0.7 of the wearable platform based on the Tizen operating system, along with One UI Watch version 1.5. The gadget is equipped with a 1.3-inch circular AMOLED display, offering a resolution of 360×360 pixels, offering a pixel density of ~ 278 pixels per inch (ppi). The device employs the Exynos 7 Dual 7270 chipset and is equipped with 768 MB of RAM. It is outfitted with various sensors, comprising a heart rate monitor, barometer, ambient light sensor, accelerometer, and gyroscope. The primary function of the smartwatch in the proposed apparatus is to collect gyroscope and accelerometer data while users actively engage in writing on the tablet screen.

The S3 Frontier incorporates an integrated LSM6DSL iNEMO 6DoF inertial measurement unit (IMU) [103], produced by STMicroelectronics. The system-in-package (SiP) encompasses both a 3D digital accelerometer and a 3D digital gyroscope. This combination facilitates precise orientation detection and movement tracking. Featuring low power consumption and high accuracy, the IMU is suitable for smartphones, battery-operated IoT devices, wearables, household devices, and gaming peripherals. Its characteristics and technical specifications are detailed in the Table 3.4.

To acquire accelerometer and gyroscope readings from the smartwatch, the official Galaxy Wearable application was employed. This proprietary software by Samsung Electronics enables the connectivity and administration of Samsung wearable devices with compatible smartphones or tablets. It functions as an integrated framework for managing devices, configuring options, personalization, and installing updates.

The reason for the smartwatch not being directly connected to the tablet is due to restrictions imposed by the manufacturer. In particular, establishing a connection between the Gear S3 Frontier smartwatch and the Galaxy S6 Tab tablet device proved unattainable. Due to several constraints, including battery life, the sensors were unable to stay operational during the entire experimental process. Since the smartwatch and tablet models used were not compatible, communication was mediated through a smartphone. In particular, the instructions to commence and halt sensor data acquisition were origi-

Table 3.4: The characteristics of the LSM6DSL system-in-package, which includes a 3D digital accelerometer and a 3D digital gyroscope. The SiP is integrated into the Galaxy S3 Frontier smartwatch utilized within the proposed apparatus.

Model	LSM6DSL
Power consumption $[mA]$	0.4 (combo normal mode) 0.65 (combo high-performance mode)
Analog supply voltage $[V]$	1.71 to 3.6
Independent IOs supply $[V]$	1.62
External magnetic sensor corrections	hard and soft ironing
Serial interface	I^2C & SPI with main processor data synchronization feature
Smart Embedded Functions	pedometer, step detector, step counter, significant motion, and tilt
Full-scale acceleration range $[g]$	$\pm2, \pm4, \pm8, \pm16$
Full-scale angular rate range $[dps]$	$\pm125, \pm245, \pm500, \pm1000, \pm2000$
Temperature range $[°C]$	-40 to $+85$
Smart FIFO	up to 4kbyte based on features set
Auxiliary SPI	for OIS data output for both gyroscope and accelerometer
Android compliance	Android M compliant
Footprint $[mm]$	$2.5 \times 3 \times 0.83$
Standard interrupts	free-fall, wakeup, 6D/4D orientation, click and double-click

nally sent from the tablet to the smartphone. The smartphone subsequently forwarded directives to the smartwatch through a successfully established connection.

A specialized desktop application for Windows OS has been developed for handling the acquired data from the smartwatch. The program was developed using the Python programming language, featuring an intuitive user interface designed to simplify user interaction. To facilitate communication between the watch and the computer, the Smart Development Bridge (SDB) was employed. It represents a command line tool that manages multiple connections with target devices and sends commands to specific devices. The SDB provides features like file transfer, port forwarding, remote shell commands, and target log output control. This client-server utility was utilized to bundle all files containing accelerometer and gyroscope sensor measurements. The approach facilitates the management of sensor data files, including tasks such as transferring, moving, and deleting files from the device's memory.

3.2.5. Smartphone sensors

The accessible smartphone device included in experimental setup was the Samsung Galaxy S9+ model. Released in 2018, it operates on the Android 9 (Pie) operating system, incorporating the One UI version 1.0 user interface. The device features a AMOLED display with HDR10 (high dynamic range) standard support. The display measures 6.2 inches diagonally and exhibits a screen-to-body ratio of approximately 84.2%. With a resolution of 1440×2960 pixels, it has a pixel density of approximately 529 pixels per inch. The specific model is based on the Exynos 9810 chipset, designed for distribution in the EMEA (Europe, Middle East, and Africa) market. It features an internal storage capacity of 64GB and a RAM capacity of 6GB.

In addition to enabling connectivity with a smartwatch, the primary function of the smartphone in this study was to utilize the built-in camera sensor. The sensor is capable of capturing videos at 4K resolution with 30/60 frames per second (FPS), 1080p resolution with 30/60/240 FPS, and 720p resolution with 960 FPS. It was employed to document the complete experimental procedure of handwriting with a stylus. As previously indicated, acquiring data regarding the inclination of the stylus was unattainable due to constraints within the S Pen Remote SDK. Consequently, the ArUco marker was positioned atop the S Pen stylus to address this limitation. Detecting and analyzing the positions of the ArUco markers from the video files facilitates the determination of the angle at which a participant holds the stylus. The custom developed Android application on the tablet device logs the start and end times of each task performed by the user. Since the camera recording is continuous, these time markers were utilized by a custom Python script to automatically trim the video based on the stored timestamps. The approach yields separate video segments corresponding to individual records of signatures, sentences, words, and letters. This allows for the precise identification of stylus inclination values that correspond to the particular writing task of the user.

The advantage of the Galaxy S9+ model is the presence of a physical 3.5 mm audio connector, a component absent in most modern smartphones. Therefore, it was suitable to employ this feature for interfacing with an external piezoelectric sensor. The headphone jack is additionally used for capturing supplementary touch screen pressure during the writing process. In this context, a secondary piezo sensor with a smaller surface area and

a spherical shape is utilized. This sensor is similarly positioned underneath the tablet, opposite the one attached to it. As previously noted, Fig. 3.4 shows both piezoelectric sensors: one is attached to the tablet using a sound card, while the other is directly connected to the smartphone using a 3.5 mm audio connector.

3.3. Experiment participants

In a controlled environment, a total of 60 subjects participated in the experiment. Out of the entire group of participants, 12 were female and 48 were male. The majority consisted of young students, enrolled at the Faculty of Engineering, University of Rijeka. A subset of the respondents in the experiment consisted of students who were currently enrolled in a Human-Computer Interaction (HCI) course. This proved beneficial as they observed the process and actively contributed to the execution of an HCI experiment in a real-world scenario. The remaining participants were colleagues and acquaintances of the author of this dissertation. The demographic characteristics of the participants are shown in Figure 3.8, which presents pie charts for age distribution, handedness (right-handed vs. left-handed), and gender (male vs. female).

Only nine persons identified themselves as predominantly left-handed. All subjects confirmed regular use of touchscreen devices and indicated expertise in their utilization techniques. The majority of users reported previous experience of writing on a touch-sensitive screen, either with an active or capacitive stylus. A smaller number wrote with their index finger, as they primarily used fingers for drawing activities.

The experiment took place in a controlled setting, where participants followed a predetermined schedule. An adequate time interval was maintained between their arrivals to allow for necessary recharging of devices, which was one of the limiting factors in the study. The testing period for each individual ranged from 70 to 90 minutes, including breaks. The test subjects were granted the flexibility to take pauses at any point during the experiment and for any duration they preferred. Simultaneous execution was considered unfeasible due to the absence of requisite equipment.

Before the experiment commenced, applicants were asked to provide their personal details, specifically their full name. An identification number was randomly assigned to facilitate differentiation between users. Each participant was provided with a document
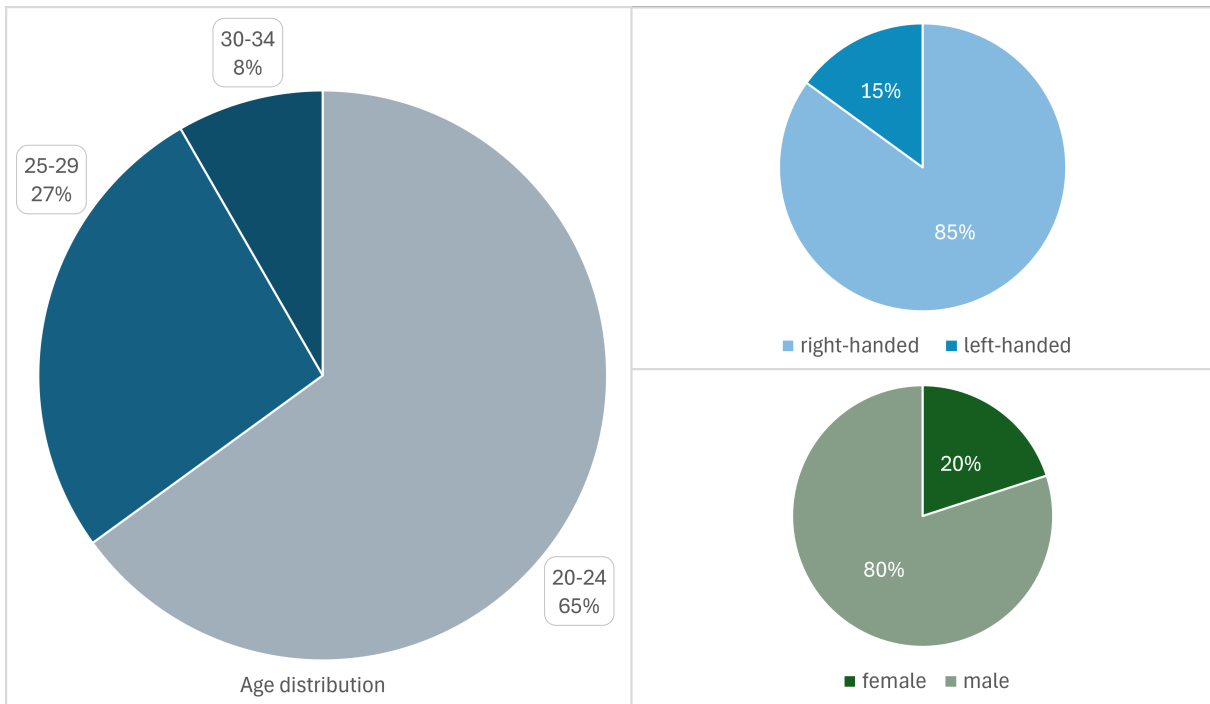
Figure 3.8: Visualization of demographic data, including age distribution, handedness, and gender breakdown of the participants.

that outlined a brief overview of the experiment, including the aims, the data collection procedures and research objectives. They were given explanations regarding the apparatus, detailing the devices used and their respective purposes. Additionally to written instructions, users received verbal guidance for further clarification. The experiment supervisor, who is also the author of this dissertation, demonstrated several trials showcasing the usage of the proposed system. If necessary, participants were advise to seek additional guidance at any point during the experiment.

A handwritten signature was requested as consent to participate in the conduct of the experiment and the gathering and examination of obtained data. The consent form can be found in the dissertation Appendix A. Each of the signed documents was properly stored and archived for future reference. Subjects were clearly informed that the origin of the data would be completely anonymized to ensure the protection of the privacy. The study protocol, the conduct of the experiment and the collection of data received approval by the Ethics Committee of the University of Rijeka, Faculty of Engineering (approval no. 640-08/23-01/3). The approval can be found in the dissertation Appendix B.

3.4.   Experimental procedure

The participants were seated at a table where the proposed apparatus was positioned at a predefined location. They were instructed to secure the smartwatch on their dominant wrist. The device was worn while engaging in writing activities using both a stylus and a finger. In a specific instance of finger-based writing, individuals utilized a custom 3D-printed ring that served as a holder for robust neodymium magnets. Another type of input modality involves the use of a specially equipped stylus, which includes 3D-printed holders for magnets and an AruCo marker. The objective was to examine how different input methods, such as using a stylus or a finger, affect the identification of individuals.

Before the experiment commenced, participants were given time to interact with the setup. In this manner, they were able to attain a detailed understanding of the system's functionality. This preliminary training stage could take place prior to the actual conduct of the experimental activities. Subjects were instructed to practice writing activities on the tablet screen, initially utilizing a stylus and subsequently using an index finger. During this phase, all sensor operations were suspended and consequently no data was collected or stored. Once the user had attained a thorough understanding of how the system functions, the process of data collection could be initiated.

To assure the precise and consistent sensor readings, the magnetometer was individually calibrated for every user at the onset of the experiment. Reference values were determined by computing the mean readings of the magnetometer along the x, y, and z axes derived from 1000 recorded measurements. While involved in the experiment, subjects were mandated to perform designated writing tasks. Given that signatures represent a pre-learned action for each individual, alternative methods of user identification were investigated. Apart from producing handwritten signatures, tasks included writing a pre-determined number of sentences, words, and letters. The research examines differences in the process of writing sentences, words, and individual letters among various individuals. Furthermore, the study investigates the intrinsic, individual characteristics inherent to these writing processes for each person.

To initiate data gathering, users were instructed to click on a dedicated button labeled as "Start" within the developed Android application. At this stage, the data acquisition process was initiated by the activation of all sensors. After concluding the writing task,
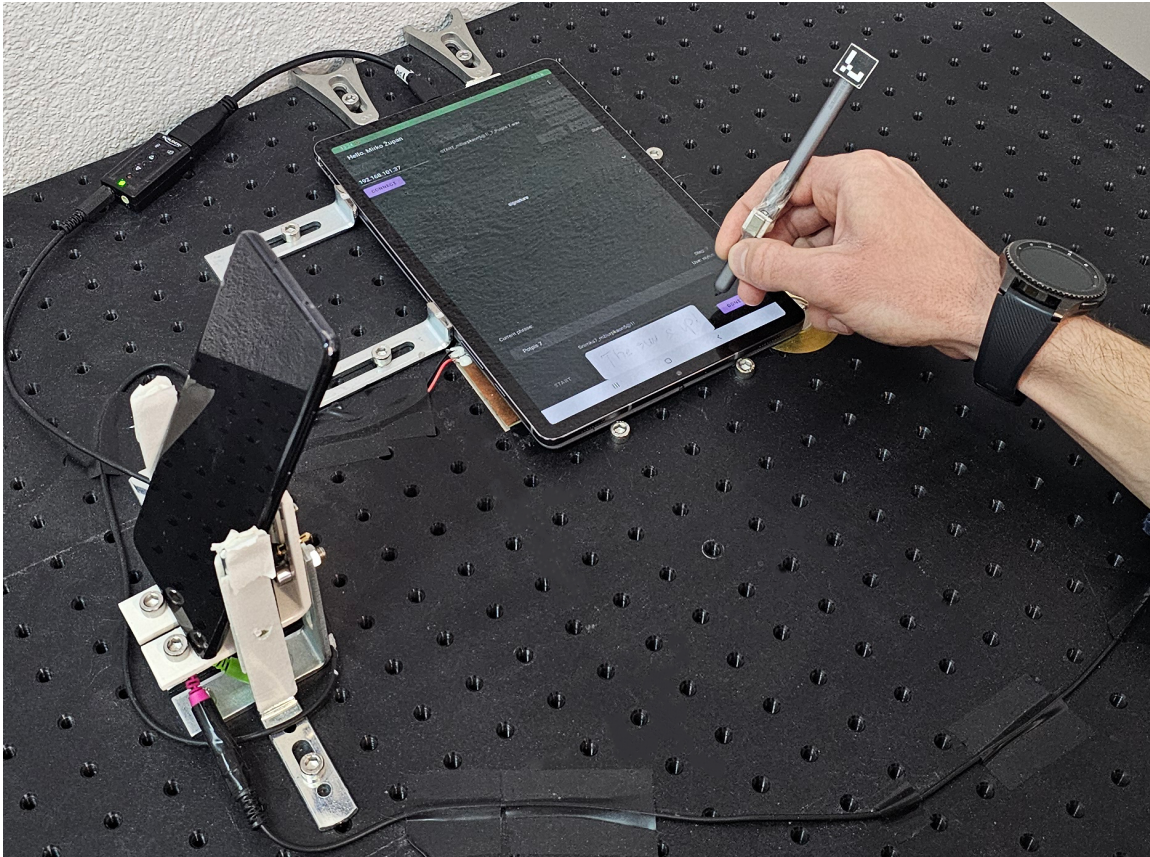
Figure 3.9: User conducting an experiment utilizing the proposed apparatus.

pressing the "Stop" button terminated sensor readings and saved the acquired measurements. This approach ensured data gathering occurred only within the time frame defined by the two defined events: pressing the "Start" and the "Stop" button. Moreover, it provided participants with the option to take breaks as required. The average duration of testing sessions per person ranged from 70 to 90 minutes, considering pauses.

At the beginning of the experiment, the participants were directed to reproduce their signature 25 times. They were instructed to replicate signatures in accordance with their typical signing practices on official documents, such as personal identification cards, driver's licenses, or other legal forms. Fig. 3.9 shows a user conducting an experiment using the proposed apparatus, writing on the tablet device with a stylus.

Regarding the composition of sentences, there were 10 phrases in total, evenly distributed between English and Croatian. They mainly comprised a selection of few brief terms, intentionally chosen to include most of the letters of both alphabets. More precisely, the sentences were composed of either two or three words. The participants were prompted to write each of the sentences a total of 5 times. Moreover, no limitations were

imposed regarding the utilization of printed or cursive writing.

Furthermore, 5 English and 5 Croatian words were selected employing the same methodology to incorporate the majority of the alphabet's letters. In this context, the selection method prioritized longer words. Similar to the procedure applied to sentences, participants were required to write each word a total of 5 times. Regarding both the sentences and words in Croatian, certain terms featured diacritical marks specific to the language (letters č, ć, đ, š, and ž). Three digraphs, representing combinations of two letters forming a single sound, were included as well (dž, lj, and nj).

Concerning the letters, all characters from the English alphabet were incorporated. However, digraphs from Croatian alphabet were excluded in this instance. The objective was to focus only on individual letters rather than combinations of characters. In addition to the capital letters, the collection of handwritten characters also contained lowercase letters. In comparison to phrases and words, each letter was written three times.

All handwritten forms, namely signatures, sentences, words, and letters, was systematically collected in duplicate. To be precise, after concluding the writing process with a stylus, the whole activity was repeated using a finger. The entire procedure followed a predefined schedule, according to the specified sequence for writing different handwriting forms. This configuration could be adjusted by modifying settings using the implemented Android application before the experiment commenced. Furthermore, a dedicated area on the screen was assigned for writing, with dimensions that adjusted based on the handwriting from currently being obtained. During the course of writing, participants were provided with prompts of the current sentence, word, or letter being collected to ensure correct input.

3.5. Sensor data overview

As a part of this research, a novel dataset is introduced encompassing readings derived from a multitude of sensors. This dataset was collected using the proposed apparatus described in Section 3.4. The resulting dataset comprises sensor measurements gathered from different devices during the handwriting activities, using both the stylus and the finger. Since readings are collected from a variety of sensor types, multiple data formats are employed for storage. The records for each subject are separately archived and cat-
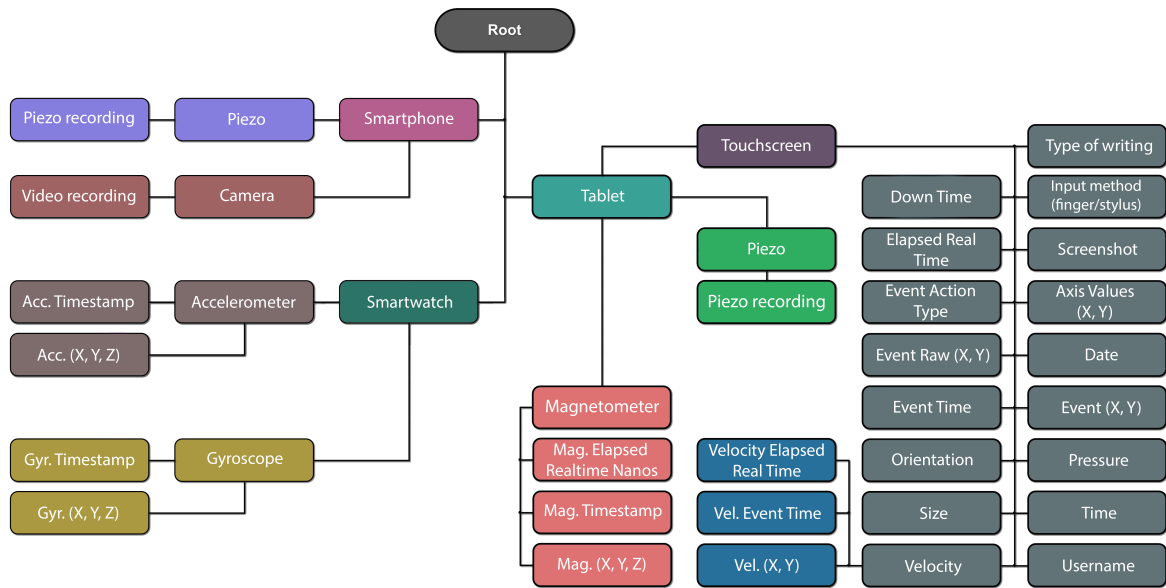
Figure 3.10: The data structure comprises the measurements acquired from apparatus employing three distinct devices, incorporating a diverse array of sensors. These readings are collected during the process of writing signatures, sentences, words, and individual letters on the tablet screen.

egorized by input modality (finger or stylus) and the type of device used to capture the measurements. An external hard drive was utilized for the storage of collected data. To ensure data integrity and prevent potential loss, regular backups were performed using cloud-based file hosting services.

Fig. 3.10 illustrates the dataset structure, providing details on the devices of the proposed apparatus, the sensors utilized, along with the data acquired by these sensors. The selection of particular devices was motivated by the inclusion of a broader array of sensors, encompassing both integrated and external types. As required, supplementary sensors can be incorporated by seamlessly integrating them with existing devices.

When utilizing a traditional pen and paper for writing, certain biometric information may inadvertently be omitted or remain unrecorded [104, 105]. The extensive dataset collected through proposed apparatus within this study is specifically intended to address this issue. The Fig. 3.11 presents an illustrative example of data obtained from a handwriting sample utilizing the proposed setup. As illustrated, each data point collected within the handwritten text encapsulates a multitude of information, distinct from that acquired through conventional pen and paper writing methods. For instance, each point can be associated with its respective x and y coordinates, timestamp, pressure, velocity,

size, and orientation. Additionally, unlike paper-based writing, it is feasible to acquire sensor readings from the magnetometer, gyroscope, and accelerometer. Consequently, examining the data from each corresponding point enables identification of the specific user's handwriting. The curvature, thickness, and width of the strokes are determined by the unique characteristics of each individual user. These attributes are derived from the analysis of various factors such as pressure level, velocity, touch size, stylus tilt, and other relevant parameters. Some characteristics can only be obtained during stylus use, while others are specific to finger input, as detailed in the following sections of the paper.

Through further analysis of the obtained data, it is possible to derive additional information about the characteristics of individuals' handwriting. For instance, by utilizing event information extracted from the MotionEvent object (such as stylus down, move, up), it is feasible to determine the number of gestures performed. Specifically, as indicated by the number inside the circle in Fig. 3.11, the count of individual strokes can be derived from the pen-down and pen-up events. In particular, the number 1 represents the initial gesture, corresponding to writing the letter "i" in a single motion. Following this, gesture 2 signifies the action where the user lifts the stylus and subsequently returns it to the screen to append the dot to the letter. Finally, gesture 3 denotes writing the letters "c" and "e" in one continuous movement without lifting the stylus. When another user writes the word "ice", it is possible that they will not write the letters "c" and "e" in a continuous stroke. Instead, they may interrupt the writing process after "c" by lifting the stylus and subsequently resume it to complete the letter "e". This indicates that
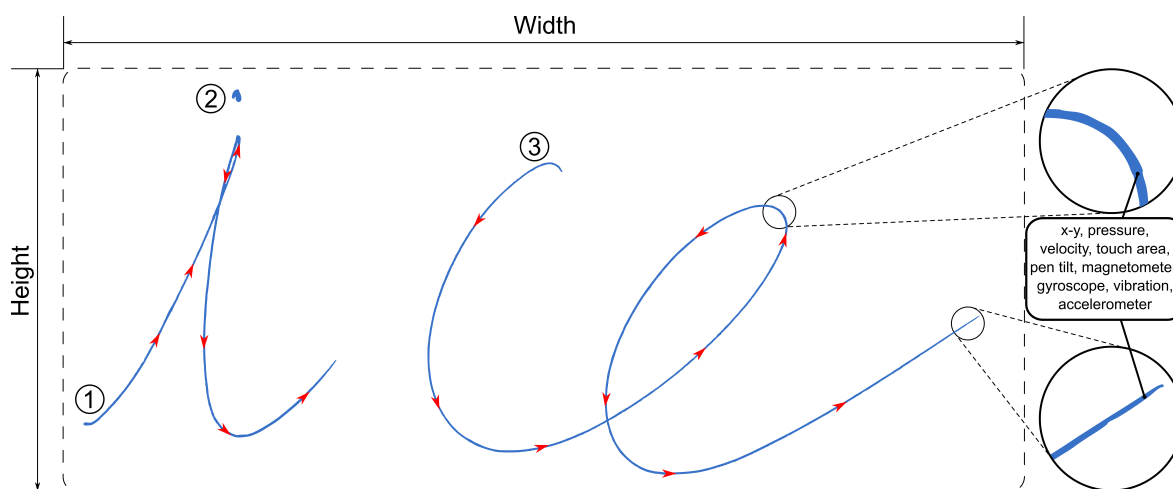
Figure 3.11: A few instances of the characteristics that may be acquired by employing the suggested apparatus while writing with a stylus on tablet device.

the different user could employ 4 distinct gestures to write the same word. This also demonstrates the variability among individuals, resulting in unique gesture patterns that differentiate their individual writing styles.

3.5.1. Touchscreen data

With regard to the tablet device, the touchscreen sensor enables the acquisition of a diverse array of data. As stated, the MotionEvent object in Android is employed to collect information about actions associated with the motion of a finger or stylus. Depending on the device specifications, it may contain absolute or relative motions, along with additional information. These movements are characterized through an action code and a collection of axis values. The action code indicates the state transition that took place, such as a pointer being lifted up or down. The axis values specify the location and additional features of the motion, such as size and orientation of the contact area, pressure level, and type of interaction. Therefore, a MotionEvent includes details about the type of action that occurred, as well as the coordinates and various other attributes of the action. For instance, as soon as the individual initially interacts with the screen, the device sends a touch signal with the action code ACTION_DOWN. Moreover, a collection of axis values encompasses the `x` and `y` positions of the input, along with details regarding the pressure, size, and orientation of the point of contact.

The MotionEvent object signifies user input interactions, encompassing aspects such as the location and motion of touch pointers on the screen. It encapsulates data pertaining to various aspects of a user interface event, involving:

- **Actions:** physical interactions with the device, including touching the screen, moving a pointer across the display, or hovering a pointer over the display area.

- **Pointers:** identifiers for objects that interact with the display, such as a stylus, finger, or mouse. Many screens support multi-touch functionality, wherein the system allocates a pointer for each finger, stylus, or other pointing object interacting with the display. A pointer index allows obtaining axis information for a specific pointer.

- **Axis:** the type of data, for example `x` and `y` coordinates, pressure, orientation, hover data (distance), and more.

Alongside the used devices and gathered sensor data, the Fig. 3.10 further contribute to a comprehensive understanding of the acquired MotionEvent parameters. The following sections present an extensive outline of the various touch event properties employed in this research scope. Certain attributes are only available when using a stylus input, while others can be obtained during finger-based writing activities. To store the aforementioned data, particularly for structuring information relevant to touch events, the JavaScript Object Notation (JSON) format was used.

When considering pointing devices such as touchscreens, the attributes $AxisValueX$ and $AxisValueY$ denote the value of the specific axis, or register a value of 0 in instances where the axis is unavailable. The $axis$ parameter of a method indicates the particular axis for which the value is to be retrieved. Axis values of the pointers can be retrieved utilizing the parameter $pointerIndex$. If the pointer index is not specified, the system will return the value associated with the first pointer, which is pointer zero (0). The units of measurement employed in this context are display pixels.

The properties $RawX$ and $RawY$ specify the raw coordinates of a touch event within the coordinate space of the device's display. This applies regardless of system decorations or multi-window mode. The obtained values are not modified by any adjustments or transformations performed on the display, such as screen orientation and scaling. The coordinates are measured in pixels, starting from the top-left corner of the display. For devices with subpixel precision, the value may include fractional components.

Furthermore, the view represents fundamental element for developing user interface components within the Android operating system. Each view performs the tasks of handling events related to user interactions, including the processing of touch events such as tapping and dragging. In contrast to the attributes $RawX$ and $RawY$, the attributes X and Y provide the coordinates of a touch input relative to the view that received it. In this instance, the X and Y coordinates are within the coordinate space of the view handling the motion event. These attributes are measured in pixels, originating from the top-left corner of the view. For devices that achieve subpixel precision, the value may include a fractional portion, indicating decimal values representing fractions of pixels. The differences between the X and Y coordinates and the $RawX$ and $RawY$ coordinates are illustrated in Fig. 3.12, highlighting the distinction between the two properties.

The parameter $Orientation$ represents the angular measurement of the touch region
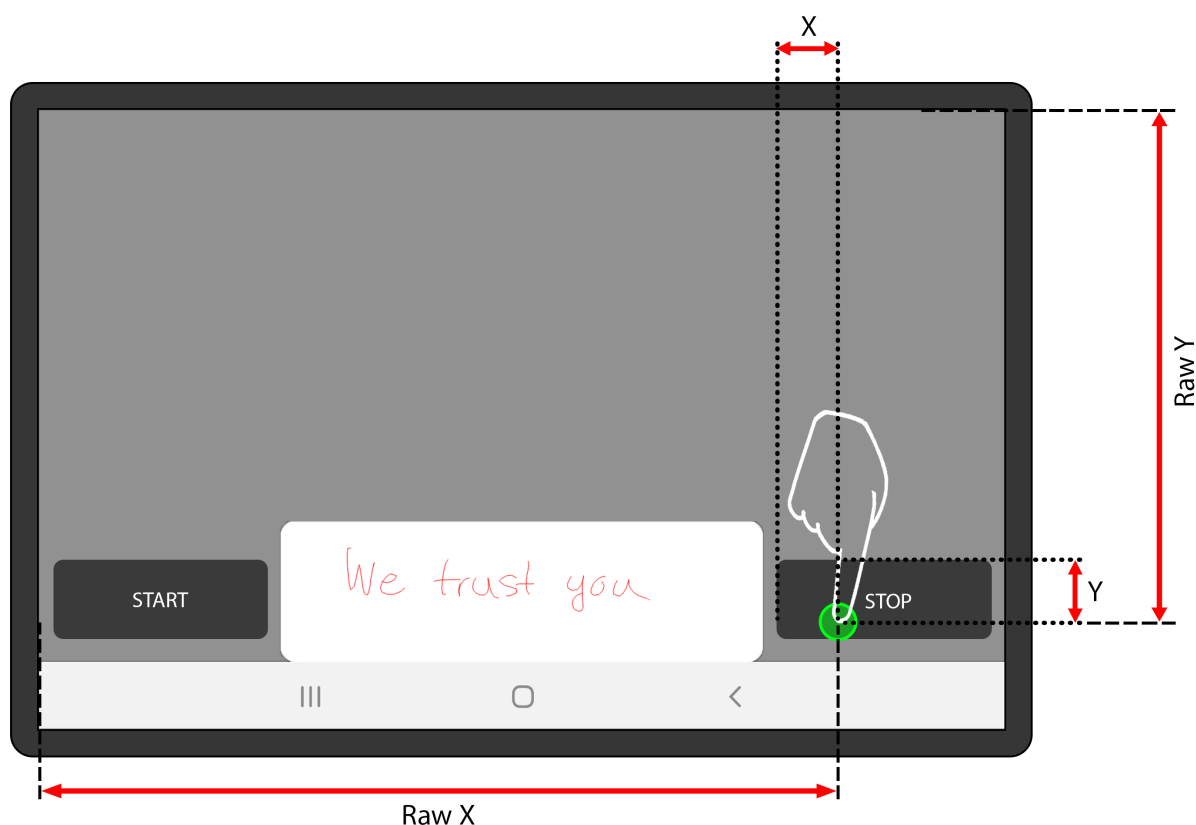
Figure 3.12: The X and Y values denote coordinates relative to the view that initiated the touch event. Here, the green circle indicates touch position on the button labeled *STOP*. Attributes *RawX* and *RawY* represent the absolute coordinates of a touch event, with respect to the screen of a device.

and tool area. The tool area refers to the specific region of the screen interacted with using a digital tool, such as a stylus. It is represented in radians, denoting the angle measured clockwise or counterclockwise from the reference axis, namely the vertical axis. Hence, this attribute specifies the direction or alignment of the touch event or gesture occurring with respect to this axis. For instance, an angle of 0 radians signifies that the touch or gesture is vertically aligned, ideally circular, or indicates an indeterminate orientation. Clockwise rotation is indicated by angles greater than 0 radians, while counterclockwise rotation is represented by values less than 0 radians. The complete range extends from $-\pi/2$ radians, indicating a stylus directed fully to the left, to $\pi/2$ radians, signifying an orientation entirely to the right. Fig. 3.13 illustrates the orientation of a stylus on a tablet device, wherein it is directed towards the left.

In Android, a sequence of motion events containing actions is used to represent an entire gesture. These actions determine the transitions and motions of the pointer state,
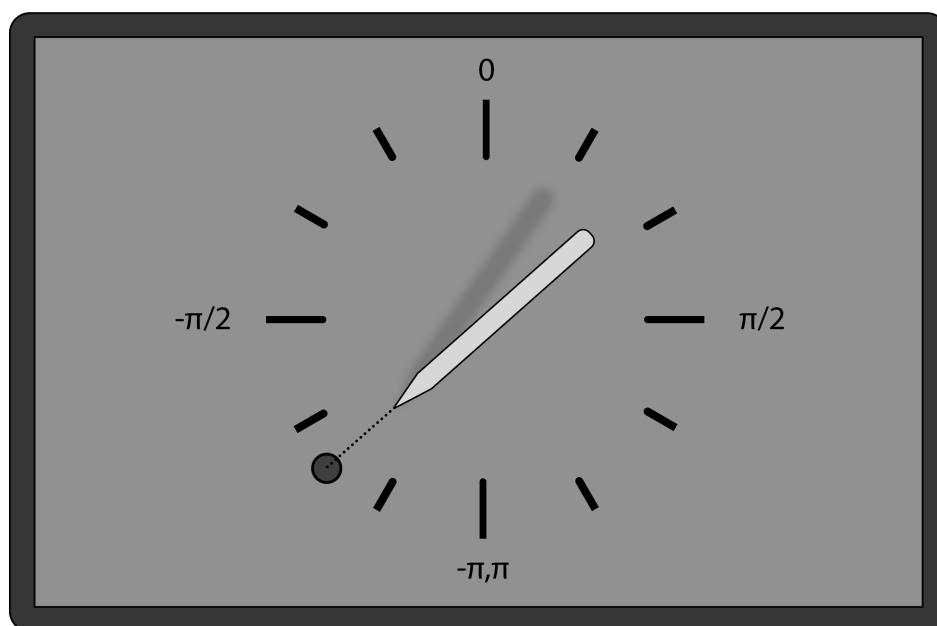
Figure 3.13: An example of the stylus tip orientation, indicating a direction to the left with an angle of approximately -2.27 radians.

specifically the stylus or finger. Thus, the *EventActionType* parameter specifies the particular category of action that has been performed. Different constants are employed to denote various types of touch events or actions executed on the screen, including pressing down, moving, or releasing a finger. For example, a gesture begins with a motion event comprising the *ACTION_DOWN* attribute, which indicates the position of the initial pointer down. This implies the primary pointer has made first contact with the screen or a particular area on the screen. The pointer movements following the initial touch are determined by motion events utilizing the *ACTION_MOVE* action constant. The action usually manifests when the user glides their finger or stylus over the screen without raising it. A gesture concludes when the last pointer is lifted off the screen with *ACTION_UP* or when the touch event is terminated with *ACTION_CANCEL*.

Furthermore, the parameter *Pressure* indicates the pressure level at a specific point applied to the screen. It is associated with a touch events, particularly when employing stylus input on devices with pressure-sensitive displays. The pressure levels usually range from 0 (which denotes the absence of pressure) to 1. However, values exceeding 1 can be attained in specific scenarios. This depends on the initial calibration or hardware characteristics of the input device. During the research, the acquisition of pressure data was unattainable when finger input was employed for writing. The touchscreen of the

Figure 3.14: Representation of the normalized stylus pressure: minimal pressure on the left (value 0), maximal pressure on the right (value 1).

tablet lacked pressure sensitivity for finger touch, consistently resulting in a measurement of 0. Fig. 3.14 illustrates how line thickness varies, reflecting different degrees of pressure applied by a stylus on the screen while writing. On the left side of the figure, labeled with number 0, the thinnest line indicates the minimum pressure level. It thickens progressively to the right as the pressure exerted by the stylus increases. The line is thickest on the rightmost side, marked with the number 1, signifying the maximum pressure applied.

The attribute $Size$ concerns the dimensions of the region on the display touched during a specific motion event. It indicates the estimated size of the touch interaction area between the user's finger or stylus and the touchscreen, typically obtained in pixels. The tapped surface area is approximated by normalizing the actual touch value in pixels with the range of values specific to the device. Afterward, these measurements are adjusted to span the interval from 0 to 1. This component includes more details regarding the touch interaction, distinguishing between accurate touches and general gestures. Moreover, it could be utilized to detect the presence of fat touch events.

The MotionEvent object provides two separate timestamps serving different purposes, specifically the $DownTime$ and the $EventTime$. The first attribute denotes the start of the touch interaction, indicating the time of initial contact. It serves as a reference for monitoring the duration of the touch event, remaining unchanged during a gesture. The $EventTime$ denotes timestamps for specific events within the touch gesture, for instance finger or stylus movements or releases. These values vary for each discrete event within the touch sequence, indicating the time of a specific action in relation to the start of the touch gesture. Both attributes are commonly expressed in milliseconds, either since the
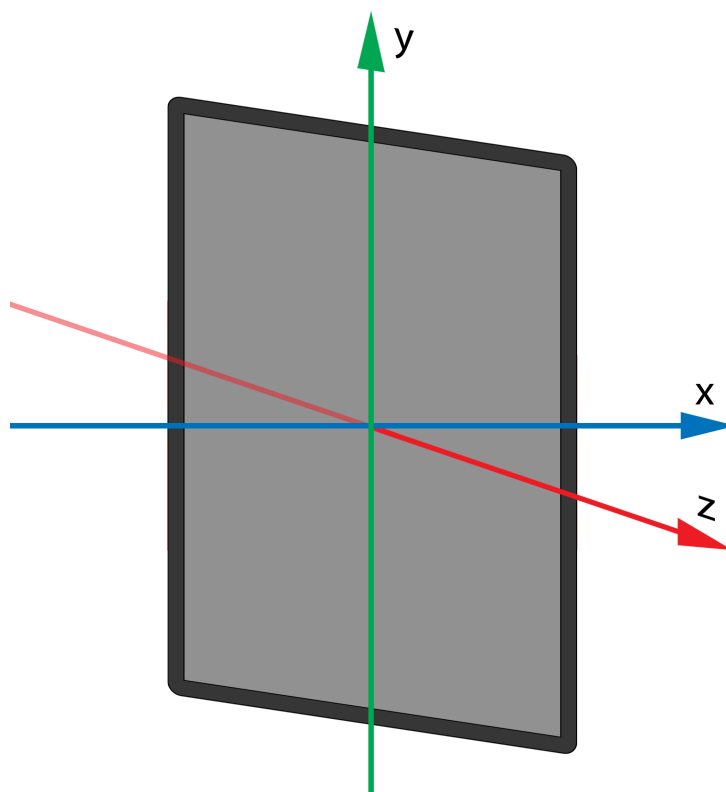
Figure 3.15: The coordinate system utilized by the Sensor API in relation to a device.

device was powered on or since the epoch [1].

The VelocityTracker API within the Android SDK was employed to determine the velocity of touch events. It facilitates the measurement of the velocity at which a pointer, such as a finger or stylus, traverses the screen. When determining the direction and speed of gestures, it is feasible to obtain the velocity of movement independently along the horizontal (x) and vertical (y) axes. Moreover, the parameter *units* can be specified for velocity measurement. For instance, a value of 1 represents units per millisecond, and 1000 denotes units per second, and so forth. The defined units are in accordance with those employed for reporting motion. More precisely, the measurement of motion along the x and y axes is expressed in terms of pixels.

3.5.2. Magnetometer data

Apart from the described touchscreen data, the proposed dataset comprises readings from various other sensors. An additional sensor employed from the tablet is the integrated magnetometer, accessed through the Android sensor framework. The obtained

---

[1]the count of non-leap seconds that have elapsed since $00:00:00$ UTC on Thursday, 1 January 1970

data pertains to measurements of the environmental geomagnetic field across three axes (x, y, and z). The values are reported individually for each coordinate axis within a single event. These readings are expressed in units of μT (microteslas).

To represent data values, the sensor API typically uses a standard 3-axis coordinate system. The coordinate system is specified relative to the display when the device is in its default position, as illustrated in Fig. 3.15. The orientation of the axes remains consistent despite any adjustments in the device's screen orientation. In the default orientation, the x axis lies horizontally, extending to the right, the y axis stands vertically and points upward, while the z axis projects outward from the front of the display. Within this system, coordinates positioned behind the display are represented with negative z values.

The sensor category utilized corresponds to an uncalibrated magnetic field type. It represents raw magnetic field measurements without any corrections for device-specific errors or environmental interferences. Therefore, the provided data does not include hard iron calibration, which addresses distortions caused by magnetized iron, steel, or permanent magnets on the device. However, these hard iron bias values are returned separately to facilitate custom calibrations. In the research, calibration was performed by assessing the environmental magnetic field across all three axes. Measurements were conducted over a total of 1000 steps, after which the average values were calculated. The corrected readings were obtained by subtracting these average values from the current readings. Additionally, the sensor operates without periodic calibration, ensuring continuity in the data stream during use.

Within the context of Android's sensor framework, there are various constants designed to regulate how often the measured values are delivered to the application. In this study, constant $SENSOR\_DELAY\_FASTEST$ was utilized to maximize the frequency of updates for magnetometer readings. This setting typically operates at the maximum sampling rate supported by the hardware and the operating system.

Among the reasons the tablet utilized in the apparatus was not updated to the latest available OS version was the sensor refresh rate constraint in Android 12. In this instance, the magnetometer sampling rate is limited to 200 Hz depending on the access to sensor data. To gather measurements at a higher rate, specific permission must be approved by the user, or a security violation will be indicated by an exception.

3.5.3. Accelerometer and gyroscope data

In the context of accelerometer and gyroscope sensor data, readings were acquired using a smartwatch during stylus or finger input. Both the accelerometer and gyroscope reference axes are depicted in Fig. 3.16. The Sensor API of the Tizen operating system was employed to access measurements from the sensors. For the accelerometer, the developed application captures changes in the velocity of a device while writing. It measures the acceleration vector along 3 axes (x, y, and z) relative to the body frame of a device. The acquired data represents a composite of linear acceleration and gravity components. Linear acceleration, which denotes the device's linear motion, can be derived by subtracting the gravitational components from accelerometer data. The accelerometer reference axes and the vector denoting user-driven acceleration are illustrated in Fig. 3.16a. The measured 3 Cartesian axis values are expressed in m/s^2, while the provided timestamp is given in microseconds. An acceleration of $1g$ is consistently observed on the axis aligned with Earth's gravitational field. While the device is idle, the readings indicate a value of $1g$ (the gravity offset) on one of its axes. This signifies the specific device axis that aligns with the direction of gravity.

The second sensor on the smartwatch pertains to the gyroscope, which detects angular
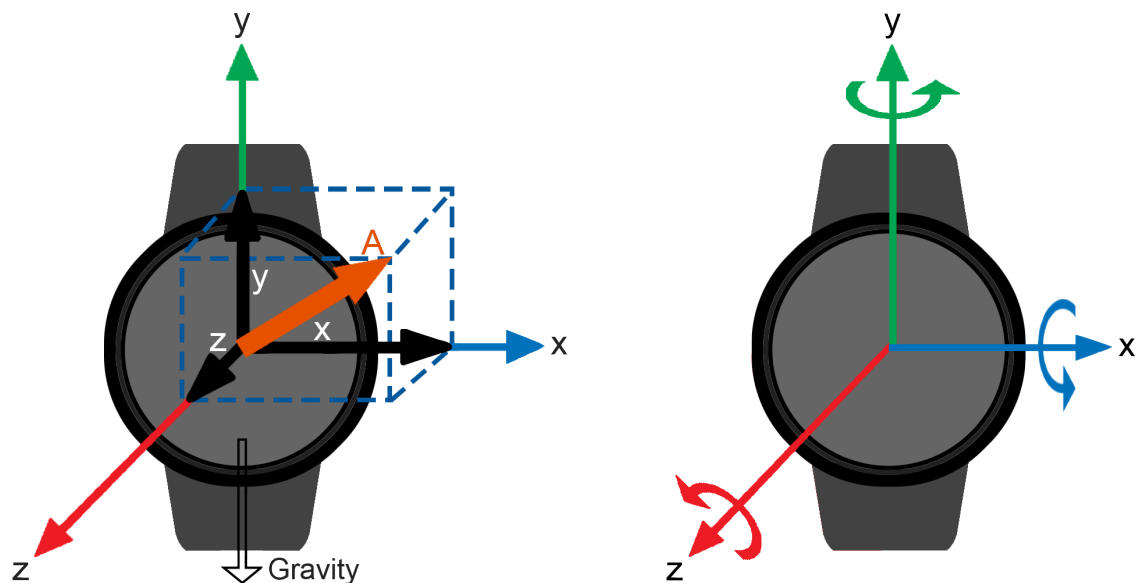
(a) The accelerometer measures velocity changes in 3 axes relative to the device's body frame. "A" denotes user-driven acceleration.

(b) The gyroscope detects and measures angular velocity or angular rates of a device along all 3 axes.

Figure 3.16: Accelerometer and gyroscope reference axes.

velocity or rates of rotation of a device. The 3D gyroscope data demonstrates high sensitivity when measuring incremental rotation angles. The sensor registers measurement data for all 3 axes in degrees per second ($°/s$), with timestamps captured in microseconds. The gyroscope reference axes are illustrated in Fig. 3.16b.

3.5.4. Camera tracking data

Due to the limitations of the S Pen Remote SDK discussed previously, an ArUco marker was positioned at the top of the stylus. Its function involves acquiring data related to the tilt of the stylus during the user's writing activity. The camera sensor from a smartphone was utilized to capture the entire process of handwriting with the stylus during the experiment. The recorded video files are intended to determine the tilt of the stylus by tracking ArUCo markers during writing. The video assets were encoded in MPEG-4 Part 14 (MP4) format, with a resolution of $720p$ and a frame rate of 30 frames per second (FPS). This format and quality ensure clear and detailed footage, while also maintaining a relatively small file size. Such attributes make the video suitable for analyzing the writing process and performing accurate ArUco marker detection.

Since the entire experiment process was recorded for each user, the clips were trimmed based on exact timestamps saved in a separate textual file while writing. These timestamps indicate the precise moments at which the writing process began and ended for each specific writing task. Each trimmed video recording represents a distinct instance of writing, whether it's an individual signature, sentence, word, or letter. This ensures that every clip corresponds to a specific written element, allowing for precise analysis of each separate writing instance. After trimming, the recordings were further processed for ArUco marker detection using a custom Python script. The 3D translation (position) and rotation vectors of the detected markers were then extracted and stored in separate textual files to determine the inclination of the stylus while writing.

As for the ArUco marker, it was generated using predefined dictionaries within the ArUco module. These dictionaries define a specific collection of distinct binary patterns that are employed to create individual markers. In the study, the standard ArUco dictionary was utilized, where the marker ID was 9 and its size measured 150 mm. Following the successful calibration process, the marker was tracked using trimmed video recordings

of signatures, sentences, words, and letters. The pose estimation process of the detected markers resulted in acquiring its orientation and position relative to the camera. The values returned correspond to rotation vectors, translation vectors, and object points.

Rotation vectors provide information about the alignment of detected markers in relation to the camera frame. The vectors do not explicitly represent the direction the marker is facing. When converted into rotation matrices, they indicate the 3D rotation (axis and angle of rotation) required to align the marker's local coordinate system with the coordinate system of a camera. By default, the coordinate system of the marker is centered at its midpoint, with the z axis oriented perpendicular to the surface of the marker. The coordinated of the four corners of the marker within its local coordinate system are presented as follows:

$$
\begin{aligned}
&\left(\frac{-\text{markerLength}}{2}, \frac{\text{markerLength}}{2}, \quad 0 \right), \\
&\left(\frac{\text{markerLength}}{2}, \frac{\text{markerLength}}{2}, \quad 0 \right), \\
&\left(\frac{\text{markerLength}}{2}, \frac{-\text{markerLength}}{2}, \quad 0 \right), \\
&\left(\frac{-\text{markerLength}}{2}, \frac{-\text{markerLength}}{2}, 0 \right).
\end{aligned}
\tag{3.1}
$$

Here, *markerLength* refers to the physical size of the marker, specifically the length of one of its square sides. Additionally, it is feasible for the coordinate system to be positioned in the upper left corner of the marker. The dataset proposed in the study contains information about the orientation of the marker in 3D space represented as a three-element vector. The first two elements specify the axis of rotation defined by the direction vector, while the third element indicates the angle of rotation around this axis.

Translation vectors describe the position of the detected markers in 3D space relative to the camera. They provide information about the displacement along all three axes from the camera's origin (usually center of the camera) to the center of the marker. In this research, these values were obtained and stored in the dataset as 3D vectors comprising parameters that specify the translation (movement) along the x, y, and z axes. This information facilitates the accurate identification of the spatial position of the ArUco markers in relation to the smartphone's camera, serving as the reference point.

Object points describe an optional parameter representing the known 3D coordinates
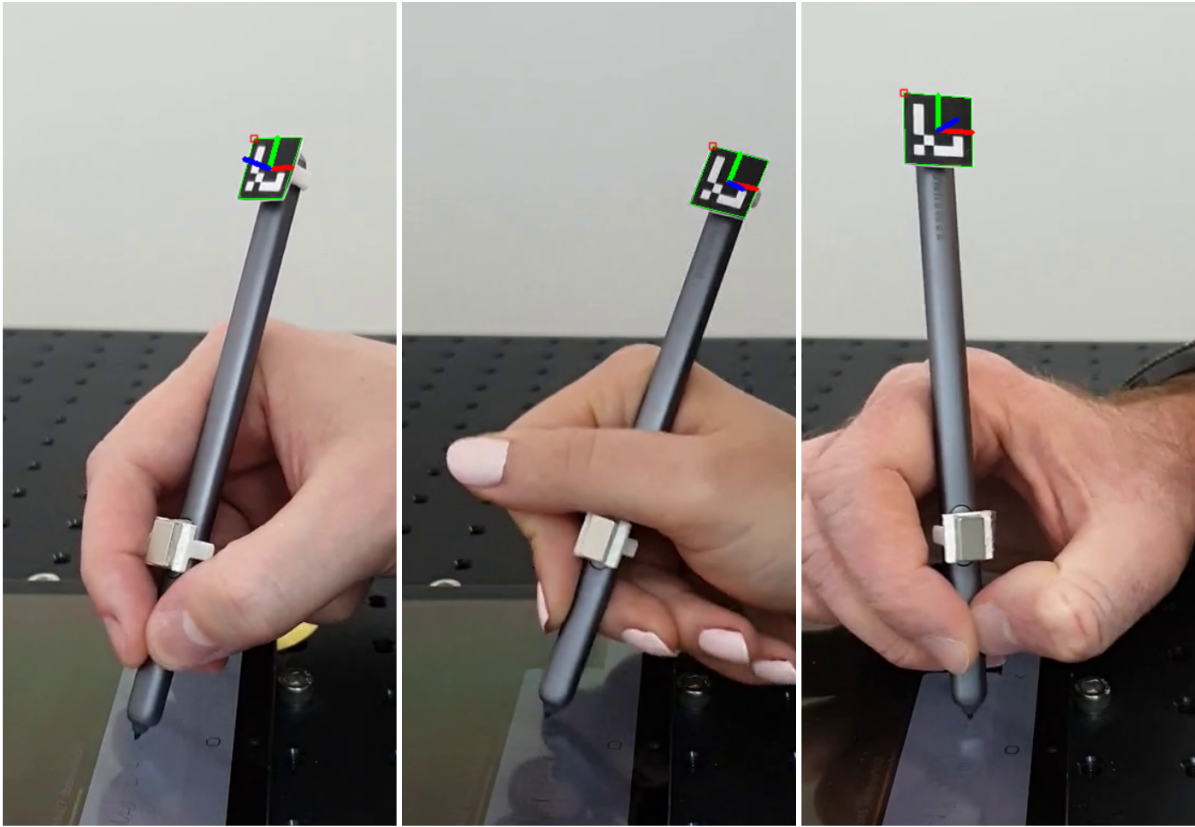
Figure 3.17: An example of ArUco marker detection during stylus writing, exhibiting variations in axis directions among three different users. The correspondence of axes to colors is as follows: `x` is represented by red, `y` by green, and `z` by blue.

of the four marker corners within its local coordinate system. This vector, combined with rotation and translation vectors, facilitates the complete reconstruction of the 3D pose of the marker. Alternatively, specifying the real physical dimensions of the marker enables automatic calculation of the coordinates. Fig. 3.17 displays examples of detected ArUco markers from a video recording during an experiment involving three different participants. While using a stylus for writing, participants demonstrate diverse approaches to gripping the device. Consequently, the acquired rotation and translation vectors exhibit values that vary among users, reflecting distinct writing styles.

Upon completion of the experiment, stylus tilt data were extracted from each video file by tracking the ArUco marker. The custom script was implemented for this purpose, performing camera calibration and marker detection from the trimmed videos. To establish an accurate mapping between a 3D point in the physical environment and its corresponding 2D projection in the image plane, it is essential to perform camera calibration. To calibrate the sensor, the process involves obtaining the distortion coefficients and intrinsic

parameters of the camera. This information allows for an accurate geometric interpretation of the captured images in relation to real-world coordinates. Once determined, the calibration parameters remain constant unless the camera's optics are modified. Details of the camera calibration procedure are provided in Appendix C.

3.5.5. Piezoelectric sensors data

Regarding the piezoelectric sensors, two were connected to both the smartphone and tablet to gather data on the vibrations and changes in pressure applied to the screen when writing. Considering their function as audio recording devices similar to conventional microphones, the data were stored in individual audio files. The recorded signal includes information about gliding, tapping, and applying pressure to either the touchscreen or the device itself. Readings from external piezoelectric sensors were saved using the Waveform Audio File Format (WAVE or WAV). This file format maintains the integrity of the original audio data, ensuring no loss in quality or information. The file was encoded using Pulse Code Modulation (PCM), a widely accepted method for converting analog audio signals into digital format [106]. All files within the proposed dataset consist of a single audio channel, maintaining a constant bit rate mode of 706 kbps. Furthermore, each sample in the audio file is represented with a bit depth of 16 bits, while the sampling rate was set to 44.1 kHz. Fig. 3.18 displays a waveform and spectrogram representation of
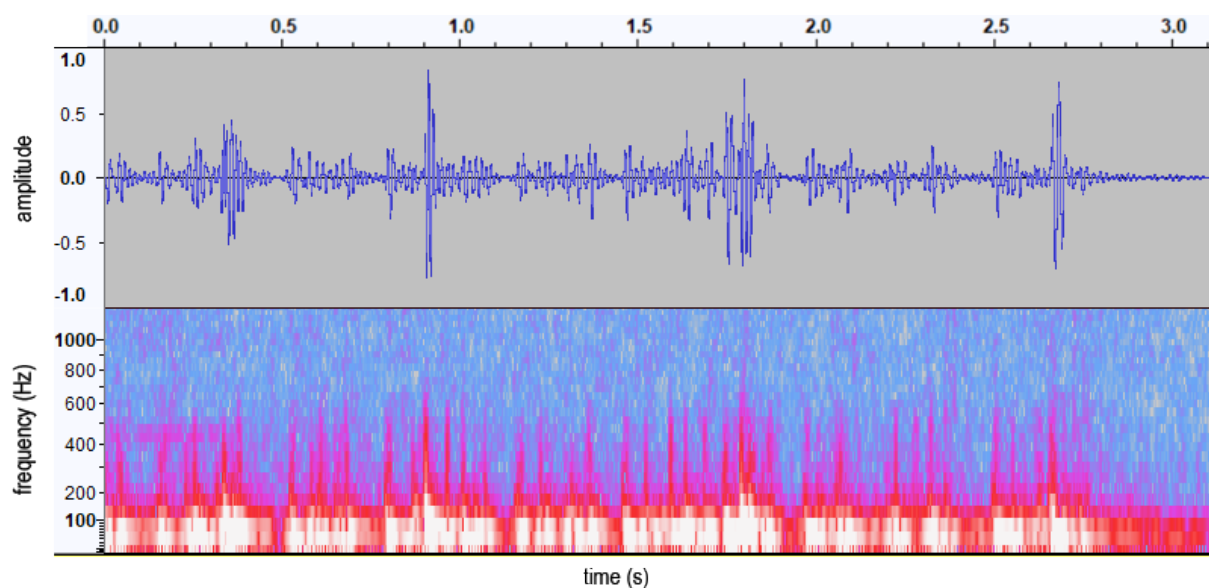
Figure 3.18: Time-domain and frequency-domain representations of a signal collected by a piezoelectric sensor while a user was writing on a tablet screen using a stylus.

64

a signal obtained from a piezoelectric sensor when writing with a stylus. The recordings provide detailed insights into the interaction between both the stylus and the finger with the surface. This allows for a deeper analysis of pressure dynamics and motion or vibration patterns during the writing process and user interaction with the touchscreen.

3.5.6.   Screenshots

Furthermore, a screenshot is taken at the conclusion of each writing session, for each trial of writing by each user, whether it involves signatures, sentences, words, or letters. This visual record helps in assessing characteristics like stroke patterns, variations in writing style (such as slant, curvature, and spacing), and the relative size and proportions of letters and words. The screenshot is saved in Portable Network Graphics (PNG) format, which preserves clarity and quality, providing a high-resolution, unaltered visual record for future reference and analysis. Although these screenshots were saved, they were not utilized in the construction of the deep learning model. However, they could serve as valuable resources for future research.

Fig. 3.19 displays handwriting samples of the sentence *"We trust you"* across three columns, where each column is associated with a distinct user. Within each column, the

Figure 3.19: Handwriting samples from multiple users, where each column represents successive attempts by the same individual to write the sentence *'We trust you'* using stylus. Similarities are observed within each user's attempts, while noticeable differences in letter formation and spacing are evident between different users.

user has made multiple attempts to write the same sentence. The figure shows that the attempts by the same user within each column are generally similar, displaying consistent characteristics such as the overall style of writing, the size of letters, and the spacing between words. For example, in each column, the shape and formation of letters like 'W' and 'T' remain relatively consistent across the attempts, indicating the unique writing pattern of each individual. However, noticeable differences between users across the columns can be observed. These variations include the way certain letters are written, such as the slant of the 'W', the crossing of the 'T', and the curvature of the 'Y'. Additionally, users exhibit different spacing between words and distinct handwriting flow, which can be seen in the overall structure and arrangement of the text. These differences present valuable perspectives into the unique handwriting characteristics of each user.

Chapter 4

MODEL DEVELOPMENT

Before delving into model development, this chapter describes the initial steps taken to explore the feasibility of handwriting-based person recognition. This included the visualization of handwriting data to uncover similarities in handwriting dynamics within the same individual and differences between various users. These observations are supported not only by graphical representations but also by numerical values described in Section 4.1. Specifically, Dynamic Time Warping (DTW) was applied to a preliminary dataset, further validating the consistency of intra-person similarities and inter-person differences. Additionally, it provided valuable information on the classification potential for user identification, as discussed in Section 4.2.

Once the classification potential was demonstrated, these initial steps guided the subsequent stages of model development using CNN. The following sections of this chapter outline the dataset utilized in the study, followed by a description of the data preprocessing steps required to prepare the data for analysis. Furthermore, it details the methods employed for feature extraction and the classification techniques utilized to identify individuals based on their handwriting.

4.1. Dynamic Time Warping (DTW)

Dynamic Time Warping (DTW) s a technique used to measure the resemblance between two time series, such as speech, gestures, handwriting, or any sequence that may differ in length or timing [107]. It identifies the best correspondence between sequences by

adjusting them non-linearly through expansion or contraction along the time dimension. DTW has been widely applied in speech recognition, gesture recognition, or handwriting analysis, where it remains relevant in contemporary research [108]. It determines the best non-linear route across two vectors, minimizing their total cumulative distance. For instance, considering two time series $Q = [q_1, q_2, \ldots, q_n]$ and $C = [c_1, c_2, \ldots, c_m]$, DTW aligns these sequences by constructing an $n \times m$ warping matrix in which each element $D(i, j)$ denotes the absolute difference between the corresponding elements q_i and c_j: $D(i, j) = |q_i - c_j|$ [109]. The objective is to identify a path through this matrix that reduces the overall cumulative distance to a minimum. The recurrence relation for computing the cumulative distance $\gamma(i, j)$ is:

$$\gamma(i, j) = D(i, j) + \min \begin{cases} \gamma(i, j-1) \\ \gamma(i-1, j-1) \\ \gamma(i-1, j) \end{cases} . \tag{4.1}$$

Here, $D(i, j)$ is the distance in the current cell, while $\gamma(i, j)$ represents the cumulative distance at that position, incorporating the smallest cumulative distances of the three neighboring cells. The path is traced from the first cell to the last, ensuring minimal distortion in the alignment of the sequences [110].

At the beginning of this research, prior to collecting the main dataset, an initial validation of the apparatus was conducted to provide proof-of-concept for its functionality. The primary objective was to demonstrate that the apparatus could effectively capture diverse biometric data, ensuring it was suitable for subsequent data collection and analysis. For this validation, an initial dataset was gathered from five participants, with each participant required to repeat the writing of signatures, sentences, words, and individual letters multiple times across different trials. This preliminary dataset is the same as the one described in Section 3.5., which was collected in an experiment with 60 participants, except it contains data from only five individuals. It included input from two modalities, with participants using both a stylus and their finger for writing.

To demonstrate the feasibility of signature-based (and more generally, handwriting-based) person recognition, DTW was applied to analyze handwriting patterns using a preliminary dataset. It measured both intra-person and inter-person differences in the

handwriting samples, with the objective of comparing writing patterns and identifying differences in handwriting dynamics. Intra-person refers to the variation within an individual's writing samples, examining how consistent a person's handwriting is across multiple trials. A lower intra-person distance indicates consistent handwriting, which is desirable in authentication contexts, as it reflects the stability of a user's handwriting. Inter-person distance refers to the differences between participants, highlighting how distinct their handwriting is when compared to one another. Higher inter-person value is preferable because it indicates that the system can easily distinguish between users' handwriting.

A detailed analysis of time-series data was performed using DTW metrics. The preliminary dataset included measurements from various sensors during the writing of the sentence *"We trust you"* by five different individuals. For each user, intra-person distance was computed by measuring pairwise distances between all trials, populating a distance matrix with the resulting values. Comparisons of identical trials were excluded, resulting in ten pairwise comparisons per user for five trials. Inter-person variability was calculated by comparing each user's trials against all other users' trials. This approach resulted in 25 inter-person distances for each pair of users. The average values of both intra-person and inter-person distances were then computed to evaluate the extent of similarity within individual users and the degree of differentiation between different users.

For sensors that capture multi-axis data, such as magnetometers, accelerometers, and gyroscopes, the magnitude was computed to simplify the complex multidimensional data for further analysis. In this context, magnitude represents a scalar quantity derived from measurements along the x, y, and sometimes z axes, depending on the sensor type. Instead of treating each dimension separately, the magnitude of the vector at each point can be calculated to reduce the three-dimensional data (x, y, z) into a single time series for each trial. It is obtained by calculating the Euclidean norm of the vector formed by these axes, providing a single magnitude value for each trial. The magnitude is calculated as:

$$\text{Magnitude} = \sqrt{x^2 + y^2 + z^2}, \qquad (4.2)$$

where $x, y,$ and z are the sensor measurements along the x, y, and z axes, respectively. In addition to DTW distances calculated for each individual axis, distances are also com-

puted between magnitude time series for each user. This comparison helps evaluate how consistent or varied a user's behavior is across different writing trials. Using magnitude as a unified metric simplifies the comparison of trials across different users by consolidating the data into a single dimension.

Section 5.1. presents the DTW analysis results, with numerical data highlighting consistent handwriting within individuals and notable differences between users. These findings are additionally supported by graphical representations that further illustrate and reinforce the observations.

4.2. Classification potential

The Classification Potential (CP) ratio was introduced to provide quantitative measure of distinguishing handwritings and maintaining consistency in person identification tasks. This ratio is computed by dividing the inter-person metric by the intra-person metric:

$$\text{CP} = \frac{\text{inter-person DTW}}{\text{intra-person DTW}}. \tag{4.3}$$

CP metric is considered favorable when the numerator (inter-person variability) is as high as possible, indicating that the handwriting of different individuals is easily distinguishable. On the other hand, the denominator (intra-person variability) should be as low as possible, showing that each individual's handwriting is highly consistent across trials. A higher CP value indicates better classification potential by showing that differences between individuals are much more pronounced than the variations within a single individual's attempts. In this scenario, handwriting is more distinct between people than it is within a person's own handwriting, resulting in more effective differentiation between individuals. In contrast, a lower CP suggests poor classification potential, either because handwriting is not distinct enough between individuals or there is significant inconsistency within the same individual's trials. The obtained values of the CP metric can be useful in the context of sensor selection and fusion optimization. Specifically, a combination of higher sensor cost (whether from purchase price or the effort needed for experimental integration) and a low CP value suggests that the sensor could be a candidate for exclusion from the fusion. In other words, if a researcher intends to remove certain sensors or groups

of sensors from the fusion for any reason, CP, along with the cost, serves as a relevant factor in making this decision. Within the scope of this research, it was crucial to take this into account before initiating the experiment, which required a considerable amount of time to conduct. This facilitated a preliminary analysis on a small dataset before proceeding with finding relevant participants for the experiment, collecting a larger amount of data, and applying neural networks for the classification task. The comparison of the CP metric values for various sensors and measurements, along with the intra-person and inter-person results, is provided in Chapter 5.2.

4.3.   Dataset

Handwriting samples were collected while users wrote on the tablet screen using a finger or stylus, and these samples were represented as a series of data points. Since multiple sensors are utilized, each data point corresponds to measurements from different internal and external sensors collected during the writing process. Consequently, every point within the handwriting sample has associated sensor readings that capture a range of writing characteristics, including pressure, velocity, magnetometer readings, gyroscope data, and accelerometer measurements, among others. All data from the dataset outlined in Section 3.5. were organized individually for each participant, categorized by input modality (finger or stylus), handwriting form (signature, sentence, word, letter), and the type of device from which the measurements were obtained.

The readings obtained from the sensors have been organized into 24 distinct vectors. Each vector contained particular measurements obtained from a specific sensor during handwriting sessions, resulting in one vector for each sensor feature. As an illustration, readings collected from the magnetometer along the x, y, and z axes formed three discrete vectors. Measurements from the gyroscope also produce three distinct vectors along these axes, as do the readings from the accelerometer. Additionally, six separate vectors were formed by the translation and rotation values obtained from the positional data of ArUco markers, both corresponding to the x, y, and z axes. The readings from the piezo sensors will result in a single vector for each piezo sensor, as they do not measure along any specific axis (e.g., x, y, or z), but instead detect changes in characteristics such as pressure or vibrations. Therefore, each sensor captures unique features, with each feature represented

as a vector of values. Below is a description of the 24 feature vectors, highlighting the specific characteristics obtained from the various sensors:

- **Touchscreen features:**

  - **relposx:** The x-coordinate of the touch event's position on the touchscreen, measured relative to the device's view area. It specifies the horizontal location where the stylus or finger touches the screen.

  - **relposy:** The y-coordinate indicates the vertical position of the touch event on the screen. Similar to the x-coordinate, it is measured relative to the active view space where the interaction occurs.

  - **velx:** The velocity of the touch event along the x-axis.

  - **vely:** The velocity of the touch event along the y-axis.

  - **pressure:** Pressure applied by the stylus on the screen. It is captured by the sensor embedded in the touchscreen, and it varies based on the force exerted during the writing. Not valid for finger input.

  - **size:** Finger touch size on the screen. Represents the area of contact when a finger touches the display. This value indicates the size of the pressed surface on the touchscreen, normalized between 0 and 1. Not valid for stylus input.

  - **orientation:** The orientation of the stylus as it interacts with the screen. It is measured in radians and represents the angle of the stylus in relation to the vertical axis of the display. Not valid for finger input.

- **Magnetic sensor readings:**

  - **magx, magy, magz**: The readings from the magnetometer sensor, measuring the magnetic field across the x, y, and z axes during the writing process. The magnetometer detects magnetic field disturbances from the stylus or finger, both equipped with magnets. These values help identify the unique handwriting pattern by tracking magnetic interference during writing. The magnetometer readings were adjusted by calibration measurements of the ambient magnetic field.

- **Piezoelectric sensor readings (attached to the tablet and smartphone):**

  - **tablet_piezo:** Readings from the piezoelectric sensor, connected to the tablet and positioned underneath, capture vibrations and mechanical strain applied on its surface while writing with a stylus or a finger.

  - **smartphone_piezo:** Similar to the tablet piezo, but located in a different position beneath the tablet. It is connected to the smartphone and provides complementary data on vibrations and pressure.

- **Smartwatch sensor readings:**

  - **rx, ry, rz:** Rotation rates around the x, y, and z axes, which are recorded by the gyroscope sensor from the smartwatch. It measures angular velocity, tracking how the user's hand rotates while writing using stylus or a finger.

  - **ax, ay, az:** Acceleration readings from the smartwatch in the x, y, and z directions, captured by the accelerometer sensor. It detects the linear acceleration of the user's hand, tracking how quickly the hand moves in any direction, whether the writing is performed using a stylus or a finger.

- **Camera-based tracking (ArUco markers):**

  - **aruco_tx, aruco_ty, aruco_tz:** 3D translation of the marker in space, providing the x, y, and z coordinates of the marker, contributing to the analysis of the pen's exact movement during the writing process.

  - **aruco_rx, aruco_ry, aruco_rz:** 3D rotation of the marker in space, determining the rotation angles around the x, y, and z axes. The rotation values help in determining how the stylus is oriented and rotated while writing.

Furthermore, each of the 24 vectors was organized into 6 distinct categories, referred to as sensor subsets in the text. This categorization was introduced to simplify the analysis and was based on the type of data and the sensors used to obtain them. Every category included specific measurement parameters:

- **touchscreen**: Includes touch positions (x, y) and touch velocities (in x and y directions) while writing.

– **magnetometer**: Measurements of magnetic fields along the x, y, and z axes.

– **input specific**: For the stylus, the values represent stylus tilt and pressure, while for the finger, they correspond to the touch size.

– **piezos**: Obtained piezoelectric data from two piezoelectric sensors connected to tablet and smartphone.

– **smartwatch**: Gathers rotational data and acceleration data across three axes from gyroscope and accelerometer.

– **visual tracking**: Monitors the translation and rotation of a stylus using ArUco markers tracked from the camera recording.

The dataset gathered consisted of data from various sensors outlined in previous chapters, captured during the handwriting with either a finger or a stylus. Before preprocessing (resampling, normalization, and padding), the dataset was split into train and test sets. Additionally, as illustrated in Fig. 4.1, the test set was further split into gallery and query subsets:

– **gallery subset:** Functions as the system's reference or "enrollment" set. It contains pre-enrolled handwriting samples from the participants, which act as templates for identification.

– **query subset:** Contains samples that are "queried" or compared against the gallery. The system receives these samples and attempts to identify the individual by matching the query to one of the enrolled samples in the gallery.
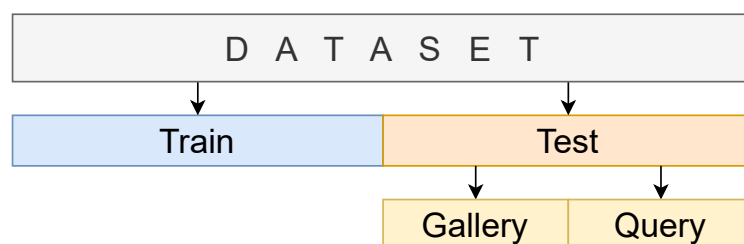
Figure 4.1: Visualization of the dataset split into training and test sets, with the test set further divided into gallery and query subsets.

The train set size represents all biometric samples (recordings) from 10, 20, or 30 randomly selected participants. The size of the test set, comprising the gallery and query subsets, depends on the train set, as it is derived from the remaining participants (50, 40, or 30, respectively). Depending on the input method (stylus or finger), handwriting form (signature, sentence, word, letter), and train size, $2 \times 4 \times 3$ datasets were generated and observed independently.

Furthermore, specific rules are applied for each handwriting form to allocate samples to the gallery and query subsets. For the sentence and word handwriting forms, the gallery subset comprised the second and fourth samples, while the query subset contained the first, third, and fifth samples. For the letter handwriting form, the second sample was included in the gallery subset, while the first and third samples were assigned to the query subset. Since there were 25 samples for signatures, the assignment to gallery and query subsets was determined using their ordinal positions after applying a modulo operation with the value 5. This facilitated the same selection pattern applied to the sentence and word samples, ensuring a consistent and systematic distribution across the gallery and query subsets.

4.4. Data preprocessing

Data preprocessing ensures that all handwriting data collected from various sensors (touchscreen, piezo sensors, gyroscope, accelerometer, magnetometer, etc.) is uniformly formatted and ready for use in training deep learning models. Resampling, normalization, and padding address the issues of varying sample sizes and scales.

As various sensors typically record at different frequencies, the vectors contain raw sensor measurements that vary in size. Consequently, the length of all vectors (number of data points) is determined by the number of feature points collected by specific sensors during the writing process. In other words, it depends on factors such as the duration of the recording while the user is writing and the sampling rate of each sensor. For instance, the touchscreen might record touch events every ten milliseconds, while the piezo sensor may sample every five milliseconds, leading to a mismatch in the number of data points. Additionally, factors like writing speed and the complexity of the stroke affect the number of samples collected. For example, a person who writes slowly generates more data points

than someone who writes quickly. Other considerations include differences in writing styles, the pressure applied, and pauses during writing, all of which can influence the amount of data captured by the sensors.

Resampling standardizes these differences by adjusting the number of samples in each vector to achieve uniformity in size. This crucial preprocessing step enables the comparison or fusion of data originating from different sensors. If vectors are not resampled to the same length, they cannot be used together in a consistent way for deep learning models. To maintain consistency, each vector was adjusted to contain the same number of samples. This number, denoted as N_s, matches the minimum sample count obtained from all sensors. In this specific context, N_s represents the length of the vector derived from the tablet's touchscreen sensor, which recorded the fewest samples. Consequently, the lengths of each feature vector were either reduced or increased to align with N_s. This adjustment guarantees that all 24 feature vectors are aligned in terms of sample size, allowing for direct comparison and utilization in deep learning models.

The maximum lengths of vectors per collected handwriting form were as follows: signature 613, sentence 777, word 607, and letter 130. Based on the number of samples N_s in the dataset, a predefined maximum size (MAX_SIZE) of 1024 samples was selected.

Once the data is resampled to a consistent size, the next step was normalization. This process ensures that all data from different sensors is on the same scale, as different sensors can produce vastly different ranges of values. The influence of sensor-specific value ranges is mitigated, preventing any single sensor from disproportionately affecting the training process. For instance, stylus pressure levels range from 0 to 1, while magnetometer readings can vary considerably based on the magnetic field. Min-max normalization is applied to each of the 24 vectors representing sensor readings, scaling all values to a range between 0 and 1. To achieve a uniform data representation across all vectors, min-max formula is applied:

$$\text{vector}[i] = \frac{\text{vector}[i] - \min(\text{vector})}{\max(\text{vector}) - \min(\text{vector})}, \forall i \in [0, N_s]. \tag{4.4}$$

Here, vector$[i]$ represents the original value at index i, $\min(\text{vector})$ is the minimum value in the vector, and $\max(\text{vector})$ is the maximum value.

To ensure every vector has the same size, with a target length of 1024 (MAX_SIZE) in

76

this case, padding is applied. The process involved extending the vectors by appending zeros to the right side of the samples until the maximum size is achieved (right-padding).

4.5. Convolutional neural network

In online signature verification or identification, various types of neural networks are frequently employed to capture and analyze the distinctive features and dynamic characteristics of handwritten signatures. Based on the complexity of the task and the nature of the dataset, some common architectures of neural networks used include: convolutional neural networks (CNNs) [111], recurrent neural networks (RNNs) [112], Siamese networks [113], and autoencoders [114]. Each of the network types has specific strengths in processing different aspects of signature data, from spatial features to temporal sequences.

The CNNs represents one of the most widely utilized and prominent architectures within the domain of deep learning (DL) networks [115]. Due to their effectiveness in feature learning, they are commonly employed across various domains [116]. These include applications in autonomous vehicles, object detection, character recognition, computer vision, medical image processing, natural language processing, and handwritten signature verification or identification [117]. CNNs, initially introduced by LeCun et al. [118], were primarily developed for image processing tasks.

Unlike their 2D counterparts tailored for images, 1D CNNs are well-suited for capturing local patterns in one-dimensional signals. They are designed to extract features from sequential or temporal data, such as financial time series, audio, and text. The architecture is built upon the same principles as CNNs for images, but the convolution operations are applied along one dimension [119]. In a 1D CNN, the core operation is the convolution, which involves sliding a one-dimensional filter (or kernel) over the input sequence to compute local weighted sums. The general mathematical formulation for the convolution operation is given by:

$$y(t) = \sum_{i=0}^{k-1} x(t+i) \cdot w(i) + b, \tag{4.5}$$

where x is the input sequence, w is the filter of length k, b is a bias parameter, and t indexes the output positions. The equation computes the output $y(t)$ at each time step

by summing the element-wise products of the input sequence and the filter weights [120]. This operation allows the network to capture local patterns and dependencies within the data. The process is repeated as the filter slides along the input, producing a feature map that highlights the presence of specific patterns. The formula is often adapted based on several important hyperparameters, including kernel size, stride, padding, and the number of filters. Specifically, the kernel size determines the window over which the convolution is computed, the stride specifies the step size of the sliding filter, and padding can be used to control the dimensionality of the output relative to the input [121].

The common architecture of a 1D CNN starts with an input layer that feeds raw sequential data into one or more convolutional layers. Each convolutional layer applies several one-dimensional filters to extract different features from local regions of the signal, producing its own feature map. Typically, the output of a convolution is passed through a non-linear activation function (such as ReLU, sigmoid, or tanh), which allows the network to model complex relationships within the data [122, 123]. Convolution layers are often combined with pooling operations, such as max pooling or average pooling, which downsample the feature maps to reduce data dimensionality and make the network less sensitive to small shifts in the input [124]. To further improve generalization and prevent overfitting, regularization techniques such as dropout and batch normalization are often applied between layers. Finally, the extracted features are typically passed through one or more fully connected layers that combine the features extracted by the convolution and pooling layers. This step allows the network to perform tasks such as regression or classification based on the combined features. For tasks involving classification, the softmax function is commonly applied [125].

In this research, alternative architectures such as ResNet50, VGGNet, EfficientNet, and DenseNet were considered by transforming all vectors into 2D images, with each row representing a vector. The chosen 1D-CNN architecture is justified by several key considerations, offering advantages compared to the other models:

- reduced parameter count and model size:

    - considering all 24 vectors, the model trains around 5 million parameters compared to approximately 25.6 million in models like ResNet50

    - the memory footprint of a model is only 39 MB, providing the best trade-off

between size and accuracy

- comparable accuracy:

  – despite its reduced complexity, it achieves comparable accuracy to the larger and more advanced architectures

Regarding the network depth, experimental results from preliminary testing demonstrated that increasing the number of convolutional layers from two to three resulted in a notable improvement in accuracy. Adding a third convolutional layer enhanced feature extraction capability, allowing it to capture more complex patterns in the data.

4.6. Feature extraction

After data preprocessing, the next step involves feature extraction (FE). In this phase, the normalized, padded, and resampled raw sensor data is transformed into a lower-dimensional form that is appropriate for model predictions. The main method for feature extraction is the 1D CNN (Conv1D), which is appropriate for processing sequential or time-series data, such as sensor measurements in this case. Convolutional layers, key to this process, apply filters (kernels) that slide across short, consecutive segments of the input data. These filters act as pattern detectors, capturing important local features such as transitions in pressure, velocity, and changes in writing direction.

As illustrated in Fig. 4.2, the structure consists of three convolutional layers, which identify local patterns within the input data. The layers are followed by the rectified linear unit (ReLU) activation layer, the batch normalization layer and the dropout layer to prevent overfitting. Lastly, a fully connected linear layer reduces the number of features, after which a Softmax function is applied for classification. Each of these layers and their functions are discussed further in the subsequent paragraphs.

The training process for feature extraction focuses on identifying relevant features from the train subset. Each of the 24 input vectors, representing data from different sensors (e.g., accelerometer, gyroscope, magnetometer), is processed by its dedicated Conv1D-based feature extractor network. As a result, a separate feature extractor is trained for each vector, consisting of convolutional layers designed to identify local patterns in the
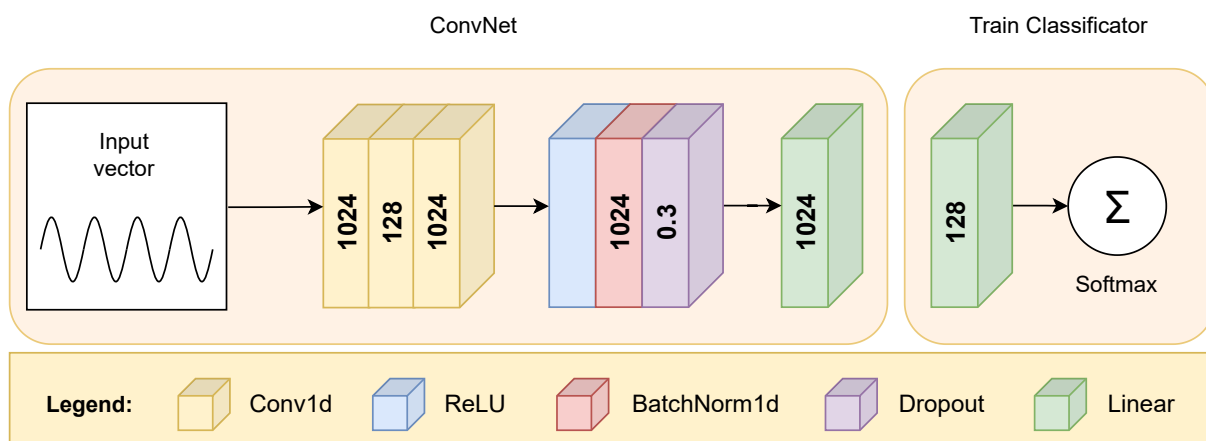
Figure 4.2: The configuration of a convolutional neural network specifically designed for the objective of feature extraction, accompanied by a fundamental training classifier.

input data. Stacking multiple convolutional layers enables hierarchical feature extraction, allowing deeper layers to capture more complex patterns by integrating simpler features from earlier layers.

The first convolutional layer operates on the full length of the input vector, which contains 1024 data points. This layer applies 128 convolutional filters (feature maps), each with a kernel size of 3, meaning that each filter processes three consecutive data points at a time. A kernel of this size is suitable for detecting short-range relationships between data points, such as slight changes in pressure or velocity in handwriting. The stride of 1 ensures that the convolution moves by one data point at a time, thereby evaluating every possible sequence of three points across the input vector. This helps the network capture subtle details in the input. To preserve the original input size of 1024 points, a zero was added to the start and end of the vector. It allows the model to recognize local patterns in the handwriting and ensures that the convolution operation does not reduce the dimensionality of the output. As a result, the first layer generates 128 distinct feature maps, each representing a specific patterns present in the handwriting data. In the second convolutional layer, the network applied another set of 128 filters to the feature maps generated by the first layer. Instead of operating directly on the raw input data, this layer processes the 128 feature maps that were learned previously. The objective was to refine the learned feature representations, combining them into more complex representations that capture relationships between different aspects of the handwriting.

Once the convolutional layers have extracted the relevant features, ReLU is applied to

incorporate non-linearity into the model. Convolutional layers by themselves only perform linear operations, which would limit the network's ability to understand detailed patterns [122]. Without ReLU, the network would only learn linear transformations, making it unable to recognize extensive characteristics in handwriting data. Attributes related to writing, captured through sensor fusion, involve highly non-linear relationships among input signals. For instance, subtle variations in stroke pressure, speed or stylus angle reflect distinctive, user-specific writing patterns. This activation function ensures that these variations are recognized by emphasizing only significant activations. By applying ReLU, the model is able to learn more intricate, non-linear connections in the data, important for recognizing subtle differences in handwriting styles and characteristics [126]. Unlike other activation functions that require advanced calculations (like sigmoid or tanh), ReLU is fast and straightforward, which is important for processing large amounts of sensor data. This simplicity contributes to faster convergence during training, enabling the network to learn more quickly from large volumes of multi-sensor data. Furthermore, this activation function is typically the primary choice suggested for deployment in the majority of feedforward neural networks [127].

After generating the ReLU output, the resulting values can vary considerably, with large fluctuations influenced by different users or writing styles. This variability can be difficult for the network to learn efficiently from the activations in the following layers. As a result, it may focus on overly large activations while overlooking smaller, important activations that represent detailed handwriting characteristics. By applying batch normalization after ReLU, the method prevents any individual feature, such as high pressure or fast movements, from dominating the learning process. BatchNorm1D normalizes the inputs to the next layer by modifying and rescaling the activations, ensuring the training process remains stable. This enables the network to concentrate on identifying the key patterns instead of being affected by variations in the sensor data. This is particularly important when data from multiple sensors includes its own noise or range of values. Additionally, BatchNorm after ReLU helps the network train faster, minimizes the risk of overfitting, and enhances performance by allowing for higher learning rates [128].

Dropout was implemented following the convolutional layers, ReLU activations, and batch normalization, just prior to the fully connected layers. The convolutional layers capture fine-grained patterns in the sensor data, such as pressure variations and sudden

velocity changes. By introducing dropout, the model learns to generalize these features across different users' handwriting styles instead of becoming overly specific to the training data. Since handwriting data is subject to variability between different users, dropout enables the system to focus on consistent patterns rather than individual writing details. For example, instead of relying on a specific sensor input (such as a pressure peak from a particular writing instance), the network learns to interpret the overall variation in pressure across different users. This approach makes the model more robust when identifying individuals based on handwriting, as it does not depend on specific sensor measurements from the train set. Dropout contributes to avoiding overfitting to the limited train samples. This technique ensures that the model generalizes better when evaluating new, unseen handwriting samples [129].

In the training phase of the feature extractor, a simple classifier was used with a single, fully connected linear layer, whose output size corresponds to the train set size. This layer functions to map the high-level features learned by the convolutional layers to the final prediction, transforming these features into a format suitable for classification. A softmax layer was applied after the linear layer, which converts a vector of arbitrary real values into a probability distribution, enabling the classification of handwriting samples. Additionally, during this phase, a cross-entropy loss function was used with the adaptive moment estimation (Adam) optimizer to effectively classify handwriting, minimize classification errors, and optimize model parameters [130]. The values of the hyperparameters were determined by empirical fine-tuning, with the *learning rate* set to 10^{-4} and the *weight decay* set to 10^{-5}. The training was performed for 1000 epochs with a *batch size* of 32. To facilitate the training process, the Nvidia RTX 4090 GPU with 24 GB VRAM was utilized.

4.7. Classification

Following the training of the feature extractors, all layers of each respective feature extractor were frozen, meaning their parameters were no longer updated during classifier training process. By employing this approach, the weights within the feature extractor network remain unchanged. This method ensured that the classifier's training process only modifies its own weights without affecting the previously learned representations of the

FEs. The feature extractor layers function as fixed tools for converting raw input data to higher-level feature representations. In the context of classifier training, the gallery subset was used as detailed in Section 4.4. The classifier learns to differentiate users using the known handwriting data of various individuals (collected through multiple sensors) from this gallery set. Subsequently, its performance was assessed using the query subset for evaluation purposes, with testing conducted on new, unseen samples from this set.

Regarding classifier models, the feature vectors derived from each of the N feature extractors are concatenated vertically, resulting in a a matrix of dimensions $N \times 128$. The matrix in question consists of N rows, each pertaining to an individual feature extractor, and 128 columns designated for every feature vector sample. In this context, the number N depends on the number of sensors employed; when all sensors are taken into account, N equals 24. As each feature vector contains 128 elements, and with 24 such vectors from each sensor, the resulting matrix had a size of 24×128. Thus, each row corresponded to a feature vector from a different sensor, and each column represented individual features of these vectors. This matrix was the input to the classifier, which tried to identify the individual (or handwriting owner) based on these stacked feature representations. During ablation study, different configurations of feature vectors were used (i.e., using fewer sensor measurements), leading to variations in the matrix size.

The classifier architecture illustrated in Fig. 4.3 employed a series of layers designed to process the stacked feature vectors extracted from multiple sensors. The architecture begins with a 1D convolutional layer performing a convolution operation over the input matrix of stacked feature vectors with dimensions N $\times 128$, transforming it into an output vector of size 1×128. By utilizing a kernel size of 3, this layer captures local dependencies
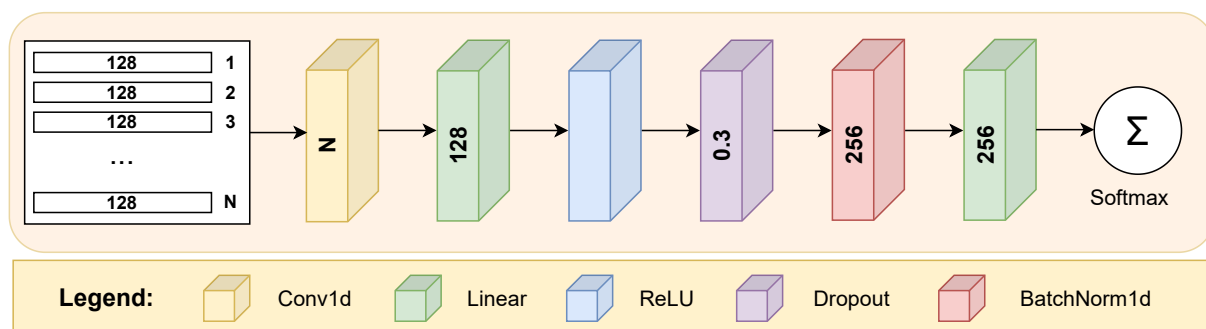
Figure 4.3: The structural arrangement of the classifier model undergoing training on the gallery dataset and subsequently assessed on a designated query subset.

and patterns within the feature data. Following the convolutional layer, a fully connected layer links each neuron in the convolutional layer to the next layer, which helps to expand learned features into a higher-dimensional space. Specifically, the first linear layer mapped the output from the convolutional layer (of size 128) to a new representation with 256 dimensions. To introduce non-linearity into the model, a ReLU activation function was used following the fully connected layer. It helps the model learn complex patterns by allowing it to capture non-linear relationships in the data. Additionally, the activation pattern generated by the ReLU function enhances generalization by making the network less likely to overfit the training data.

To further address the challenge of overfitting, which occurs when the model excessively learns from the training data and leads to weak performance on unseen data, a dropout was added after the ReLU activation. By applying a dropout rate of 0.3, a subset of neurons is randomly deactivated in every training iteration. Following the dropout layer, a batch normalization was utilized to stabilize the training process. This layer modifies and scales the activations to normalize the output from the preceding layer. By reducing the effects of internal covariate shift, in which the input distribution varies throughout training, batch normalization accelerates model convergence.

The architecture proceeds with a final linear layer that reduces the output from 256 dimensions to a value that corresponds to the unique classes (number of users in the gallery and query subset) for each user's handwriting. This layer acts as a mapping function, converting the learned feature representation into class scores. Finally, the softmax activation function transforms the raw outputs from the last layer into a probability distribution for the class labels.

Chapter 5

RESULTS AND DISCUSSION

This chapter outlines, analyzes, and addresses the findings from the research carried out within the scope of this dissertation. The results of the DTW analysis are presented, along with graphs that visually support the findings, demonstrating similarities in handwriting dynamics within the same individual and noticeable differences between different users. Additionally, it provided valuable insights into the classification potential of each sensor for identifying individuals.

The obtained results showed that the proposed handwriting recognition model achieves high accuracies, with CNN used for feature extraction and classification tasks. Performance varied across different handwriting forms, with the highest accuracy for signatures and the lowest for individual letters. Furthermore, the outcomes of the three-way repeated measures ANOVA are presented, investigating the main effects with regard to the influence of three different factors (handwriting form, input modality, and train set size) on the system's accuracy. Finally, the findings of an ablation study are provided, analyzing the impact of individual sensors within the fusion-based setup.

5.1. DTW results

In the following paragraphs, the outcomes of the DTW analysis are outlined, along with graphs that visually support the findings. The results show clear similarities in handwriting dynamics within the same individual and noticeable differences between different users. These observations are supported not only by numerical data but also by graphi-

cal representations, which further confirm the consistency of intra-person similarities and inter-person differences. The analysis focuses on sensor measurements collected during the task of writing the randomly chosen phrase "*We trust you*" using both a stylus and a finger. Different individuals completed the task over multiple attempts, consistently writing the same phrase each time. All the graphs and tables presented on the following pages are based on data collected while writing this specific phrase.

As described, a smartwatch is utilized to collect data through its built-in accelerometer and gyroscope sensors. Fig. 5.1 provides a visual representation of the acceleration in three directions for three distinct users, each making multiple attempts to write the same phrase. Upon closer inspection of the acceleration values across the x, y, and z axes, it is evident that the data from the same individual exhibits a high degree of consistency across the different writing trials. This consistency suggests that each individual's writing motion is relatively stable within their own attempts. However, when comparing the data across distinct users, there are noticeable differences between their attempts. These variations can likely be attributed to differences such as wrist movement and posture, writing speed, or hand positioning, which are reflected in the sensor readings. The accelerometer sensor thus captures these subtle variations, demonstrating how user-specific factors, such as writing pace, can be used to distinguish between users. Table 5.1 presents a comparison of intra-person and inter-person variability for accelerometer readings, as well as the magnitude, when using a stylus and a finger for writing. It can be noticed that intra-person variability is consistently lower than inter-person variability in all accelerometer sensor measurements.

Consistent measurements can be observed in Fig. 5.2 pertaining to the gyroscope sensor integrated into the smartwatch. The provided graph displays distinct peaks in the gyroscope readings, indicating repetitive patterns occurring at certain time intervals throughout identical user attempts. This is particularly noticeable when examining the rotation rate of the gyroscope along the x-axis, as shown by the blue lines on the graph. Similar trends are visible on the y and z axes, indicating consistent movement patterns across different directions. However, variations between users are evident, with some exhibiting sharper and more pronounced peaks, particularly on the x-axis. These differences likely reflect variations in wrist movement dynamics and hand motion patters during writing. Furthermore, intra-person consistency remains high, as trials for the same individual
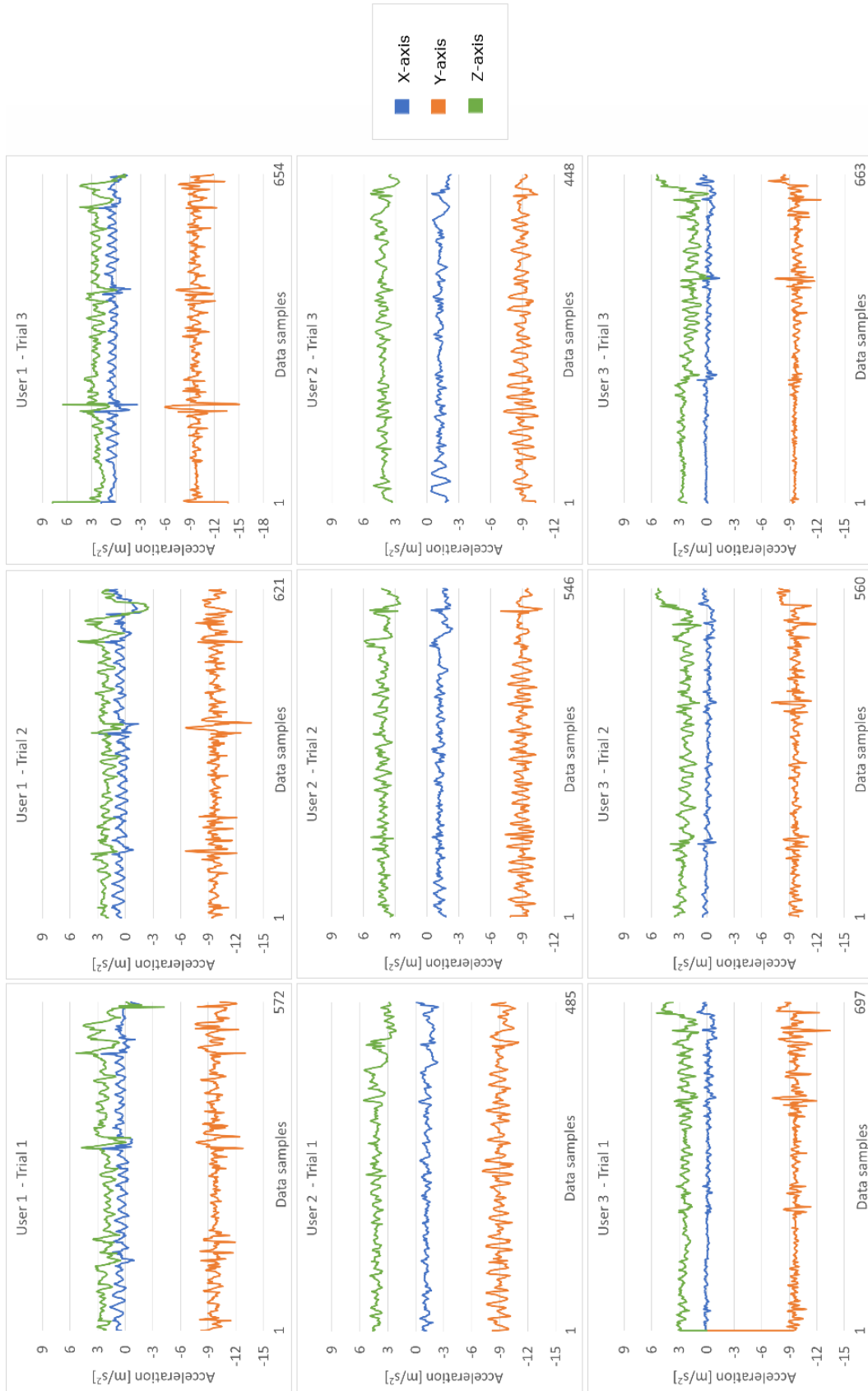
Figure 5.1: Smartwatch accelerometer data shows variations in acceleration between different users along all three axes, with the stylus serving as the input modality.
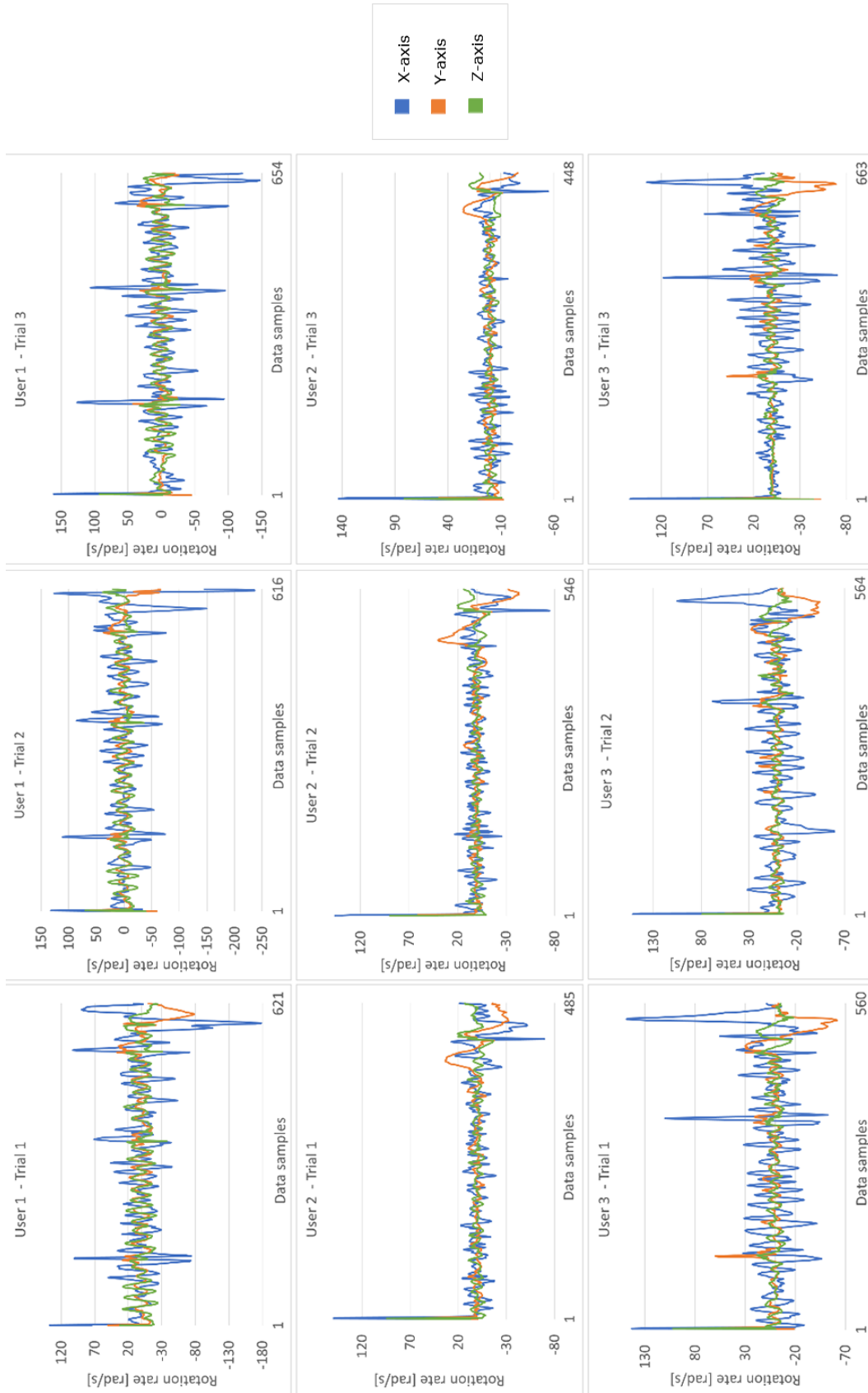
Figure 5.2: The gyroscope time-series sensor data was acquired from a smartwatch during the period in which three distinct individuals wrote on the tablet device using a stylus.

Table 5.1: Comparison of intra-person and inter-person measurements for stylus and finger using data obtained from an accelerometer. Five participants in the experiment wrote the same phrase in five different attempts.

Measurement	Axis	Stylus		Finger	
		Intra-Person	Inter-Person	Intra-Person	Inter-Person
	X	4.19	20.91	4.56	17.23
	Y	8.58	110.47	8.33	78.13
Accelerometer [m/s^2]	Z	8.01	58.35	6.05	94.59
	Magnitude	8.50	117.23	4.70	124.68

exhibit similar rotational profiles. This suggests stable movement behavior during writing, while inter-person variations highlight distinct movement styles that can be attributed to the specific techniques of an individual's handwriting.

The gyroscope-based values in Table 5.2 demonstrate similar patterns to the accelerometer readings, with intra-person variability being notably lower than inter-person variability. For both stylus and finger input modalities, intra-person variability is consistently lower than inter-person variability across all axes (x, y, z) and the magnitude. In conclusion, the gyroscope data shows consistency with similar results observed from the accelerometer data. This further supports the idea that measurements from both the accelerometer and gyroscope sensors integrated into the smartwatch show greater variability between different individuals, while remaining more consistent within the same person across multiple writing trials.

Table 5.2: Analysis of intra-person and inter-person differences in writing with a stylus versus a finger, based on gyroscope data. Each of the five participants completed five distinct trials, writing the same phrase each time.

Measurement	Axis	Stylus		Finger	
		Intra-Person	Inter-Person	Intra-Person	Inter-Person
	X	244.18	541.28	210.34	408.35
	Y	94.49	215.95	109.69	246.60
Gyroscope [rad/s]	Z	91.60	196.48	81.02	174.83
	Magnitude	198.83	495.22	162.89	405.33

Fig. 5.3 presents magnetometer measurements along the x, y, and z axes captured from a tablet device during five trials conducted by three different participants. As explored in the author's previous research [34], the magnetometer detects subtle changes in the magnetic field caused by strong neodymium magnets, identical in size and characteristics, moving in front of a tablet. These magnets, affixed to the stylus and a ring worn on the user's finger, cause interference of the magnetic field during the writing process. Such fluctuations can be analyzed and stored as a unique "magnetic signature". It is important to highlight that the magnet remained in a relatively constant starting position, and the magnetometer was consistently calibrated at the beginning of each experiment session.

Notable variations in the magnetometer sensor readings were observed across different users. As shown in Fig. 5.3, these differences were most pronounced in the magnetic field readings along the x and y axes, while the z-axis measurements exhibited less variation. Furthermore, the minimum and maximum values for the same users show minimal deviation from the mean across all five writing trials, suggesting consistency in their writing. The graph illustrates that each user may have a unique activity surface, reflecting a distinct set of writing attempts. For person recognition, the current writing attempt must fall within the existing activity surface for a specific user, as determined by the magnetometer readings.

Similar to an accelerometer and gyroscope in a smartwatch, the magnetometer data reveals notable differences between intra-person and inter-person variability, as shown in Table 5.3. For magnitude readings obtained using the tablet device with an integrated

Table 5.3: Comparison of intra-person and inter-person variability in magnetometer readings across the x, y, and z axes, as well as their magnitudes, for stylus and finger inputs. These readings were collected from five different participants, each of whom completed five distinct writing trials.

Measurement	Axis	Stylus		Finger	
		Intra-Person	Inter-Person	Intra-Person	Inter-Person
	X	70.98	143.40	63.84	290.57
Magnetometer [μT]	Y	83.23	540.27	104.12	625.94
	Z	78.78	441.09	96.61	604.51
	Magnitude	63.95	493.07	80.61	542.80
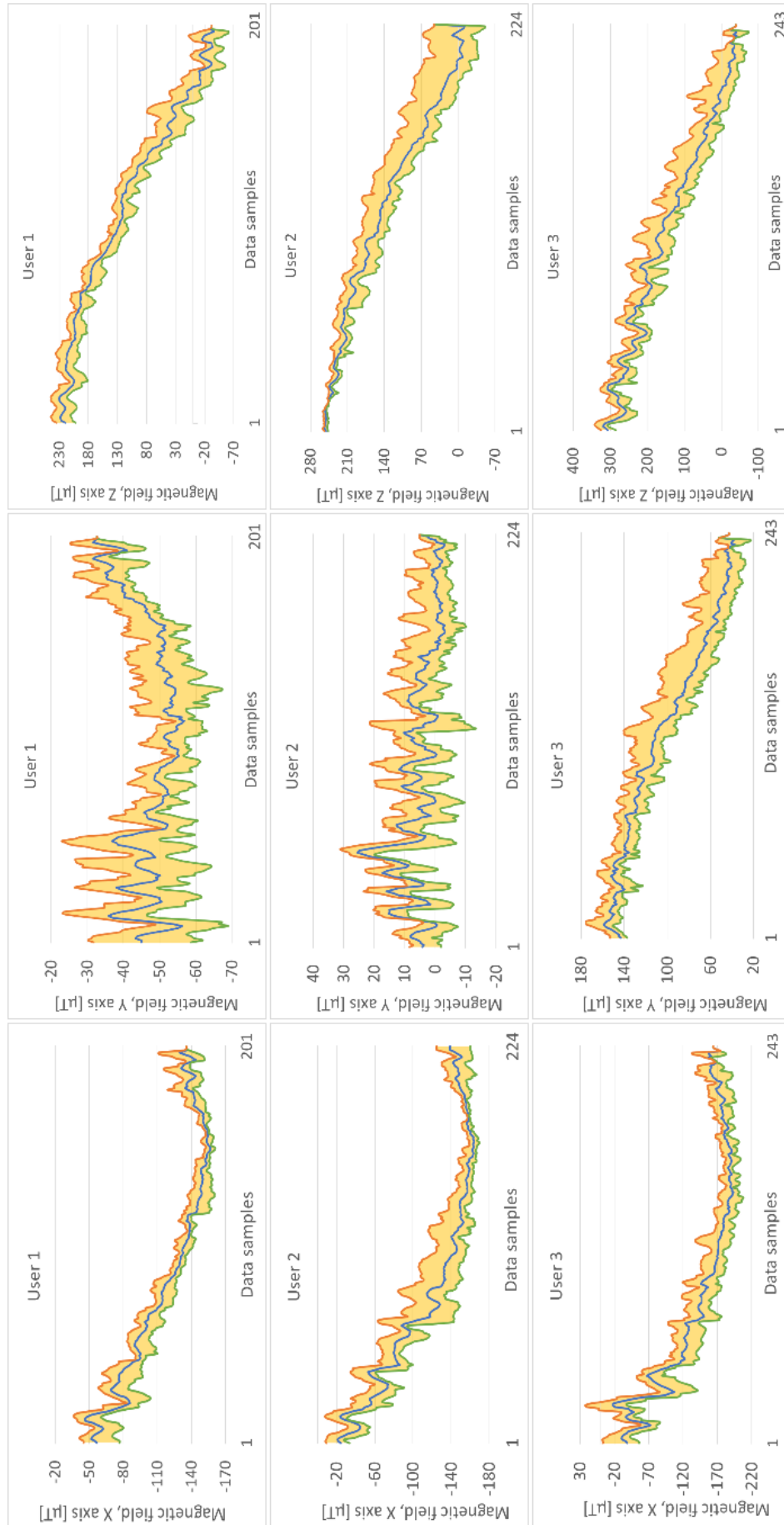
Figure 5.3: Magnetometer data collected from three users throughout five trials of writing with a stylus. The minimum, average, and maximum magnetometer reading values of the sensor are represented by the green, blue, and orange lines, respectively. The values shown are the average of all five trials per user.

magnetometer, the intra-person distance with the stylus is notably lower than the inter-person distance. Similarly, for the finger writing, the inter-person distance is considerably greater when compared to the intra-person distance. This suggests that an individual likely has a unique "magnetic signature" that distinguishes them from other users.

The velocity at which a pointer, namely a finger or stylus, moves across the screen is monitored using the VelocityTracker Android API. When determining the direction and velocity of gestures, it can be measured independently along the horizontal (x) and vertical (y) axes, as seen in Fig. 5.4a and 5.4b. Compared to the readings from other sensors, the averaged minimum and maximum values for a particular user exhibit a marginally greater deviation from the mean reading values. However, the average velocity lines (in red) remain relatively stable across trials for each user, indicating that each individual maintains a consistent writing velocity. Additionally, considerable differences in velocity are observed along both the x and y axes across distinct users. Some individuals demonstrate a wider range of velocities, suggesting more noticeable fluctuations in their writing velocity. In contrast, others exhibit a more limited range, indicating a more consistent and stable writing velocity throughout the trials. This pattern can be anticipated, as the writing velocity tends to remain consistent for each individual, while different individuals could write the same sentence at varying velocities.

Table 5.4 underscores that writing velocity measurements display lower variability within individuals compared to variability between individuals for both stylus and finger inputs. The magnitude data indicates comparable differences in intra-person and inter-person variability, consistent with the x and y-axis measurements. These results suggest that while velocity measurements for an individual remain relatively stable, there

Table 5.4: Comparison of intra-person and inter-person variability in velocity readings for stylus and finger inputs. The data was collected from five participants, each completing five separate trials.

Measurement	Axis	Stylus		Finger	
		Intra-Person	Inter-Person	Intra-Person	Inter-Person
	X	19,465.04	19,883.20	1112.40	1419.23
Velocity [pixels/s]	Y	328,021.36	325,161.88	1958.50	2786.84
	Magnitude	329,110.20	319,536.78	1561.99	2093.88
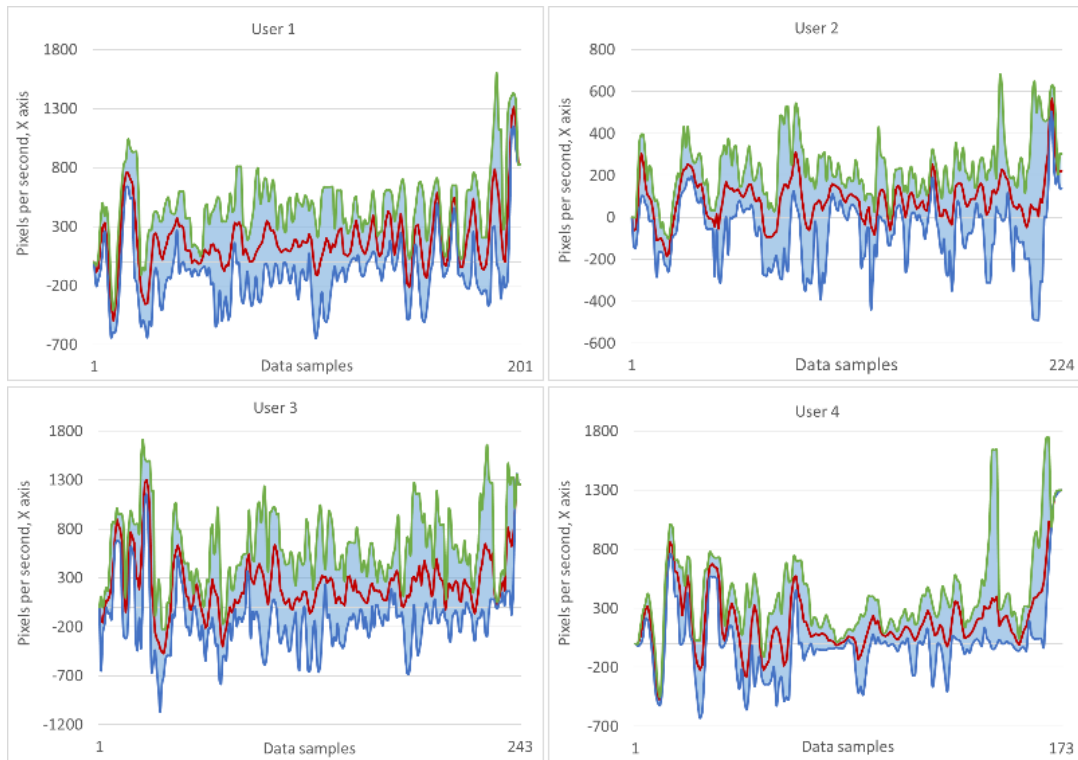
(a)
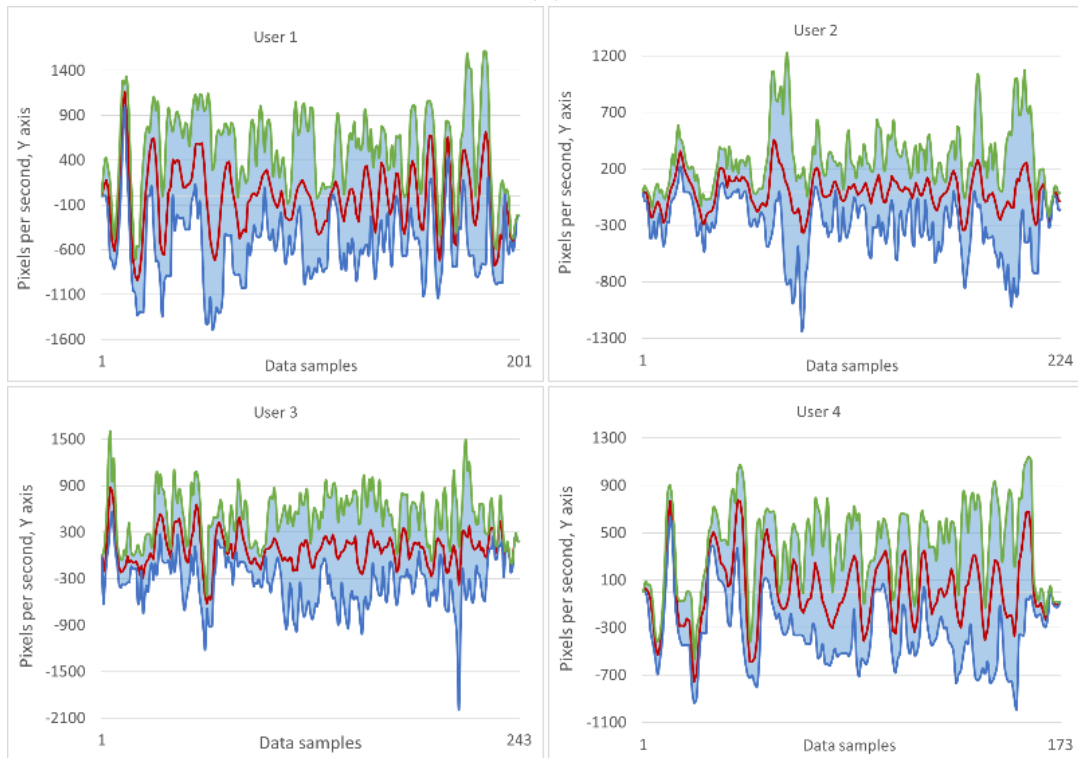
(b)

Figure 5.4: The minimum (blue), average (red), and maximum (green) measurements of the velocity in the x-direction (a) and y-direction (b) for four users during five trials of writing using a stylus.

is noticeable variability between different people. This could be attributed to individual differences in motor control, writing habits or techniques. Additionally, the stylus demonstrates notably higher velocity values than the finger, possibly because it facilitates smoother and faster motion compared to finger input. In conclusion, both input modalities show greater intra-person consistency than inter-person variation.

An ArUco marker was used to overcome S Pen API limitations, capturing inclination data and extracting x, y, and z coordinates from video recordings to track stylus rotation and translation. Fig. 5.5 illustrates variations in stylus tilt measurements, with only x and y coordinates shown for simplicity, based on multiple attempts by different users. Data points with similar hues (same user across multiple trials) show a tendency to form clusters, suggesting that individuals may exhibit relatively consistent stylus tilt patterns across attempts, though slight variations are present within the clusters. Clusters represented by different colors (different users) appear to have some level of separation, indicating potential differences in stylus tilt patterns among users. However, overlapping regions suggest that these differences may not always be clearly distinguishable.

The numbers in Tables 5.5 and 5.6 show a notable difference between intra-person and
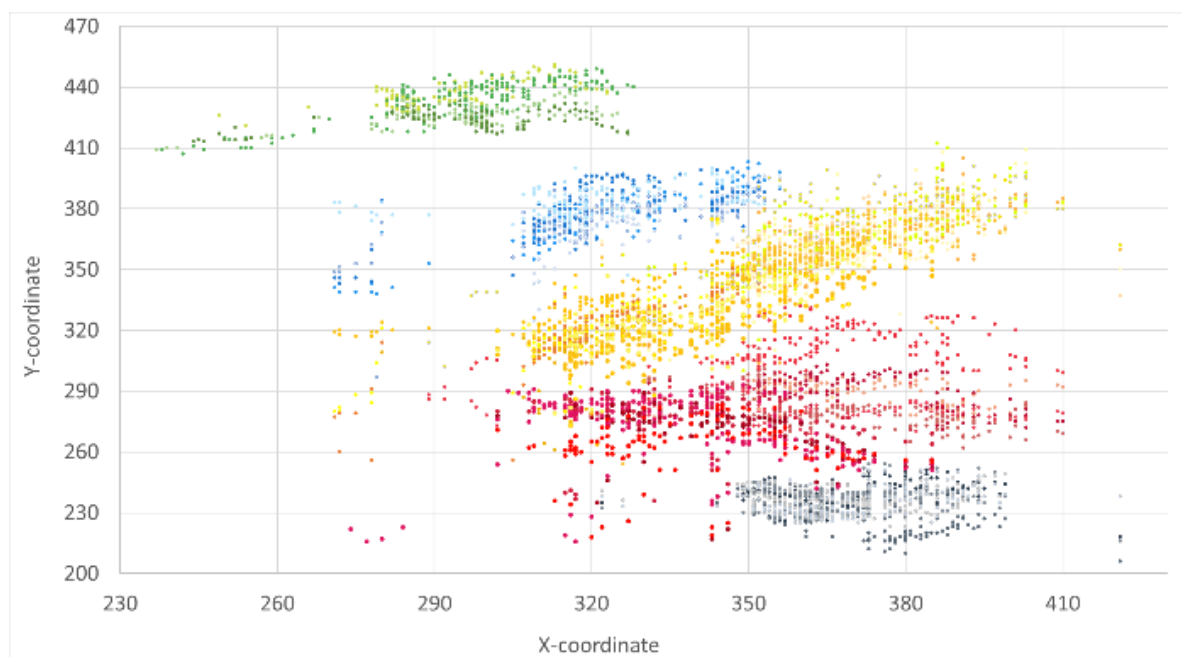
Figure 5.5: The x and y coordinates of the stylus tilt are determined by using the ArUco marker positioned on the tip of the stylus. Distinct hues of a single color signify multiple trials of the same person, whereas varying colors denote the inclinations of the stylus for different individuals.

Table 5.5: The rotation vector variability for stylus tilt readings examined across the x, y, and z axes, as well as its magnitude, within individuals and between users. Five participants completed five separate trials, with no data collected for finger writing.

Measurement	Axis	Intra-Person	Inter-Person
Stylus rotation [rad]	X	9.11	18.83
	Y	7.33	14.68
	Z	2.18	4.37
	Magnitude	5.42	18.52

Table 5.6: Intra-person and inter-person variability of the translation vector for stylus tilt across the x, y, and z axes and magnitude. Data for finger writing was not available. Each of the five participants performed five separate trials using a stylus.

Measurement	Axis	Intra-Person	Inter-Person
Stylus translation [mm]	X	0.65	2.62
	Y	0.85	3.52
	Z	1.62	6.78
	Magnitude	1.94	8.06

inter-person variability for stylus tilt readings, with intra-person variability being lower across all three axes for both rotation and translation vectors. For the magnitude of the rotation vector, intra-person variability is lower than inter-person variability, which exceeds three times the former. Similarly, for the magnitude of the translation vector, intra-person variability is lower, with inter-person variability being more than four times higher. These comparisons highlight the greater variability in stylus tilt during writing across different individuals, as opposed to within a single person. The reason could be that each individual develops a unique way of holding and using the stylus, similar to how one holds a pencil. This may be influenced by factors such as motor coordination, ergonomic preferences, and writing experience. As shown in Fig. 3.17 from Chapter 3, the way the stylus is held further supports this claim, reflecting variations in individual handling and grip preferences. In contrast, within an individual, these factors tend to relatively consistent over time, resulting in less variability in stylus tilt.

Since the stylus S Pen Remote SDK could not provide pressure level data from the stylus, the MotionEvent object was used to capture pressure readings instead. Pressure typically varies from 0 (no pressure) to 1 (normal pressure), based on how the input device
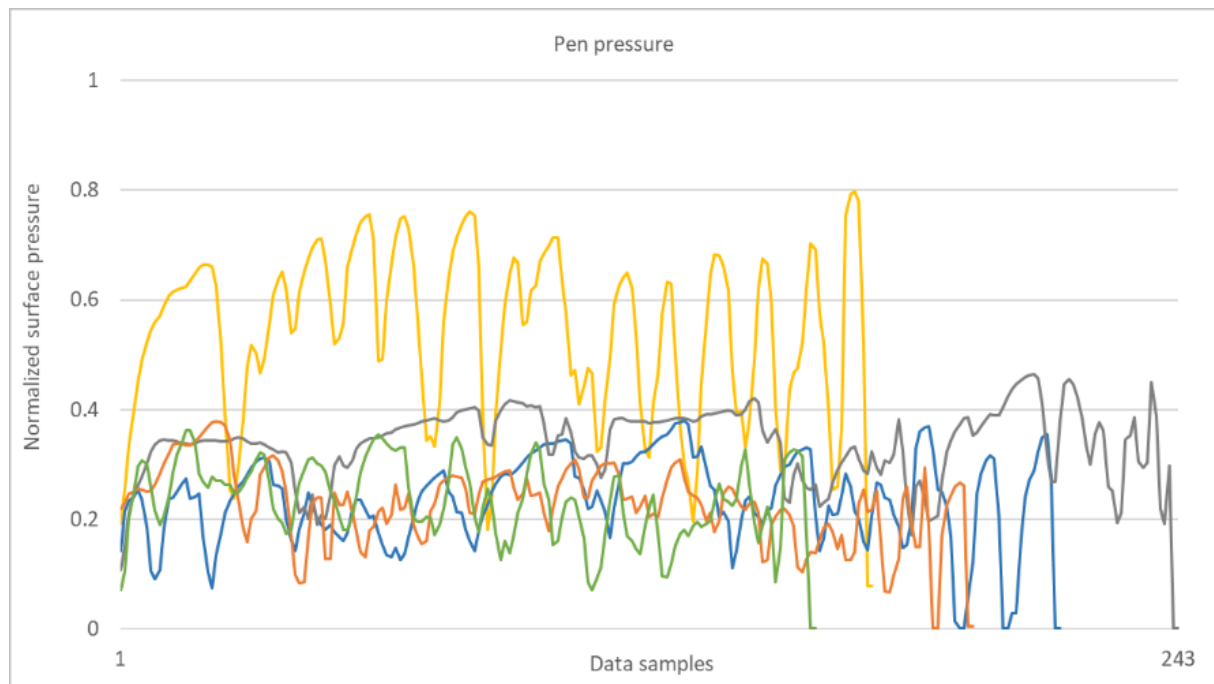
Figure 5.6: Average pressure levels recorded during five writing trials using a stylus, with each color representing one of the five users.

is calibrated. Figure 5.6 illustrates variations in the normalized mean pressure levels across five users, each writing the same phrase five times. It can be seen that some variations in pressure patterns exist between users, with one individual (yellow line) exhibiting a much higher and fluctuating pressure range compared to others. The remaining four users display relatively lower and steadier pressure trends, though with some variability. This indicates that differences in pressure during stylus-based writing can be captured, highlighting variations in writing dynamics and pressure application across individuals. These variations could be associated with personal variations in writing technique, grip strength, and physical interaction with the stylus, which vary more between people than within a single individual. In contrast, variability between users is lower as each person's writing style remains more consistent across trials. However, it should be noted the MotionEvent object does not support capturing pressure data for finger-based input.

Fig. 5.7 shows variations in pressure application across users, while trials from the same user tend to be more consistent. Some individuals distribute pressure more evenly, creating smoother and more uniform color transitions across trials, whereas others exhibit greater variation, resulting in noticeable shifts in color intensity within their strokes. The patterns of pressure distribution suggest distinct writing styles, with some marked by
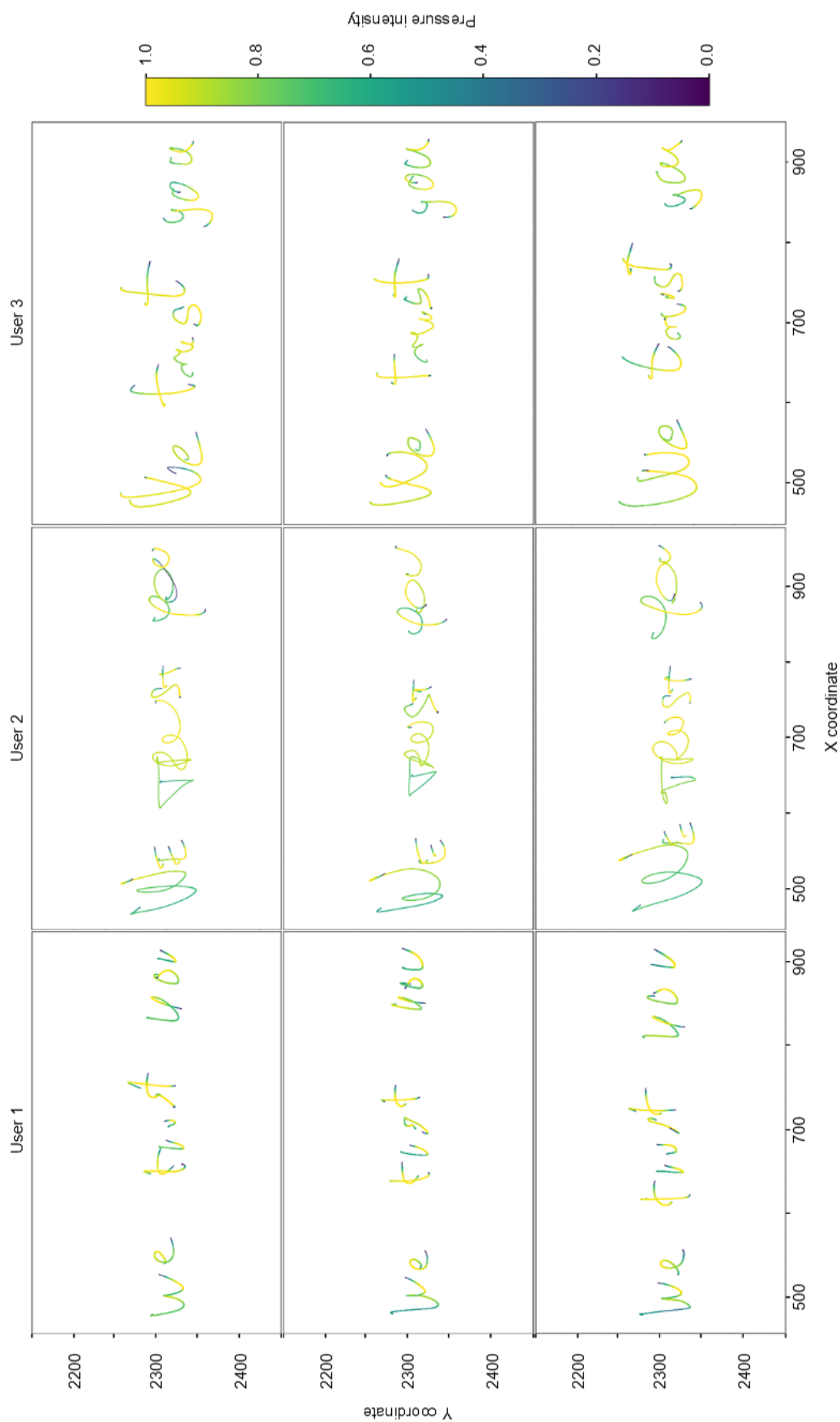
Figure 5.7: Reconstructed handwriting strokes based on x, y coordinates, and pressure values from three users, with each column representing three different trials from the same user. Color variations indicate pressure intensity, with darker tones representing higher pressure. Cubic spline interpolation is applied for smoother trajectories.

Table 5.7: Pressure data was collected for stylus writing only, with no measurements available for finger writing. Five participants completed five trials each, with intra-person and inter-person measurements compared.

	Stylus		Finger	
Measurement	Intra-Person	Inter-Person	Intra-Person	Inter-Person
Pressure	0.34	0.97	N/A	N/A

consistently higher pressure, leading to more saturated colors, while others apply a lighter touch with lower overall pressure. Despite these differences, repeated trials from the same user generally maintain a similar pressure pattern, reflecting a stable writing behavior.

For pressure metrics, the stylus data from Table 5.7 suggests a distinction between intra-person and inter-person distances, indicating that pressure may help differentiate between users, although the values are smaller compared to other metrics. This implies that pressure could be a more subtle, yet potentially useful, metric for user identification.

While exact pressure data for finger-writing input cannot be obtained, the Motion-Event object estimates the touched area on the screen. When a touch occurs, the normalized pixel value of the touch area is given, which is adjusted based on the device's pixel range. Fig. 5.8 illustrates the average results of five writing attempts by four distinct users, showing the estimated finger-pressed area on the screen. The mean touched area varies across individuals, irrespective of them writing the identical phrase. Each participant appears to exhibit distinct touch size patterns, where variations likely reflect individual differences in touch dynamics, finger contact area, applied finger pressure while writing, or interaction with the touchscreen surface. Certain individuals displayed a wider range of touch sizes, suggesting larger fluctuations in pressure or contact area. While some participants consistently had a larger touch area, potentially due to a flatter finger position or greater pressure, others exhibited a smaller touch area, possibly reflecting a more precise or lighter touch.

For touch size metrics, as shown in Table 5.8, finger input reveals a difference between intra-person and inter-person distances. This suggests that touch size may help differentiate users, but it is less distinct compared to other metrics. Data for the stylus was not available due to the inability to estimate the surface area touched while writing with the stylus, so no comparison can be made for that input method.
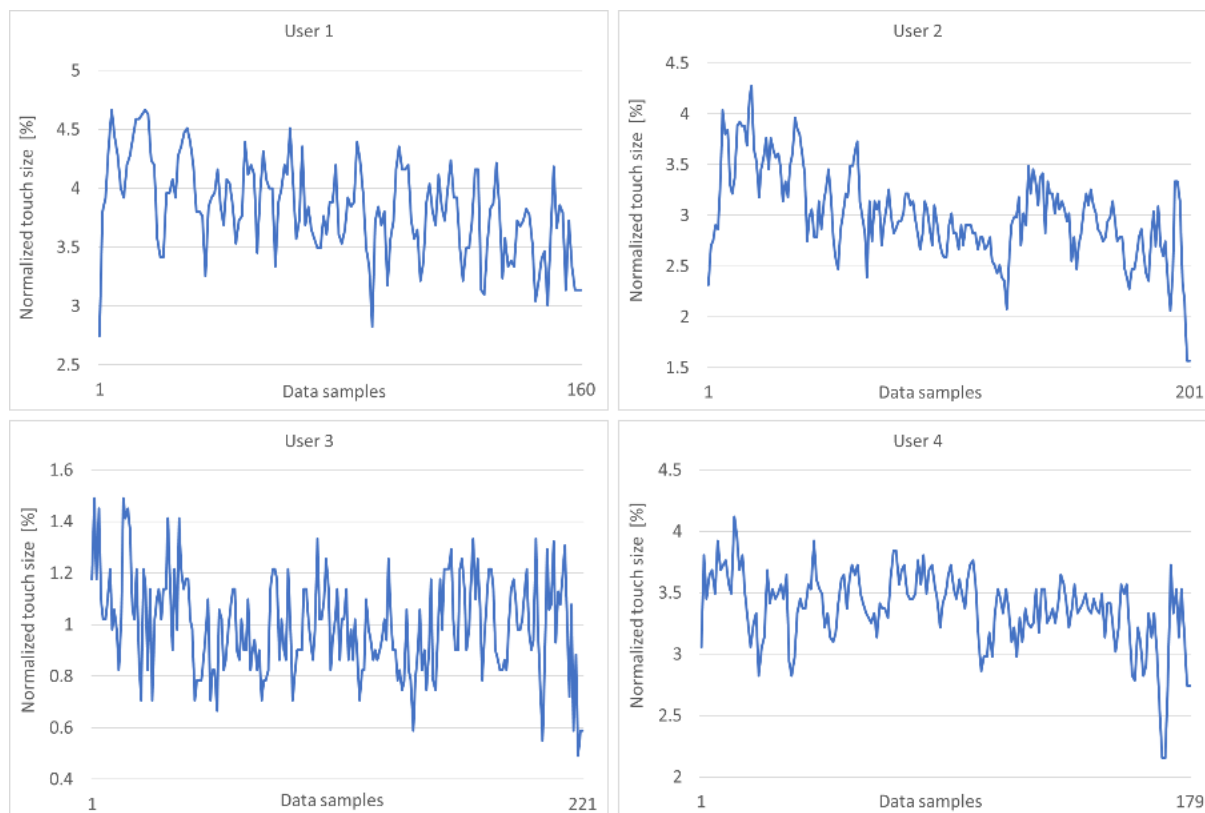
Figure 5.8: Estimation of the screen surface area touched while writing with a finger. The values represent the averages from four users, each completing five trials, with data for stylus input unavailable.

Table 5.8: Variability in touch size area measurements for finger input during writing tasks. Data was collected from four individuals, each completing five trials, with no data available for stylus input.

	Stylus		Finger	
Measurement	Intra-Person	Inter-Person	Intra-Person	Inter-Person
Touch size area	N/A	N/A	0.02	0.09

Fig. 5.9 shows five attempts by four different users to write the same phrase "*We trust you*". Since each user tends to form letters and words uniquely, the figure highlights variations in the writing process among users, such as differences in gesture durations and pause intervals. A "*gesture*" refers to the time between placing the stylus on the screen and lifting it off, encompassing the writing of a single character (e.g., a letter, period, or comma) or multiple characters when the user combines them while writing. Although there are variations in the writing patterns across trials for each individual, certain trends suggest that users tend to follow relatively consistent patterns within their own writing
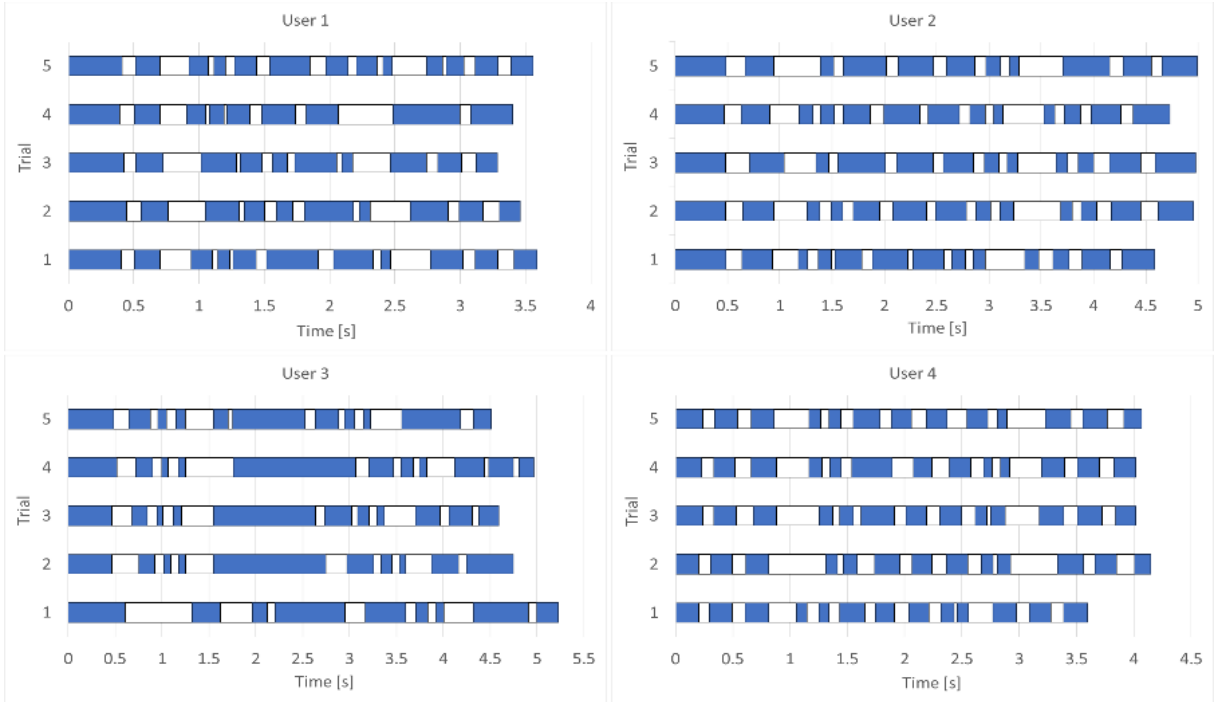
Figure 5.9: The number of stylus gestures performed by four users across five distinct writing trials. The writing periods are highlighted in blue, indicating active writing, while the white spaces represent pauses in writing.

process. Overall, there is noticeable variability in the duration and frequency of pauses across users. Some writing patterns appear more steady, with fewer interruptions, while others show frequent pauses, indicating differences in writing fluency and pacing. The visualization also illustrates the varying numbers of gestures performed, which differ from user to user, reflecting individual writing styles during the task.

The variability of x and y coordinates and their magnitudes, during five trials conducted using a stylus or finger as input methods, is presented in Table 5.9. Relative

Table 5.9: Comparison of intra-person and inter-person variability in relative position measurements (x, y coordinates) and magnitudes for stylus and finger writing, based on five trials from each of the five participants.

Measurement	Axis	Stylus		Finger	
		Intra-Person	Inter-Person	Intra-Person	Inter-Person
	X	101.73	306.44	93.27	219.74
Relative Position (X,Y) [pixel]	Y	79.34	129.70	99.56	202.03
	Magnitude	103.37	220.40	100.89	226.56
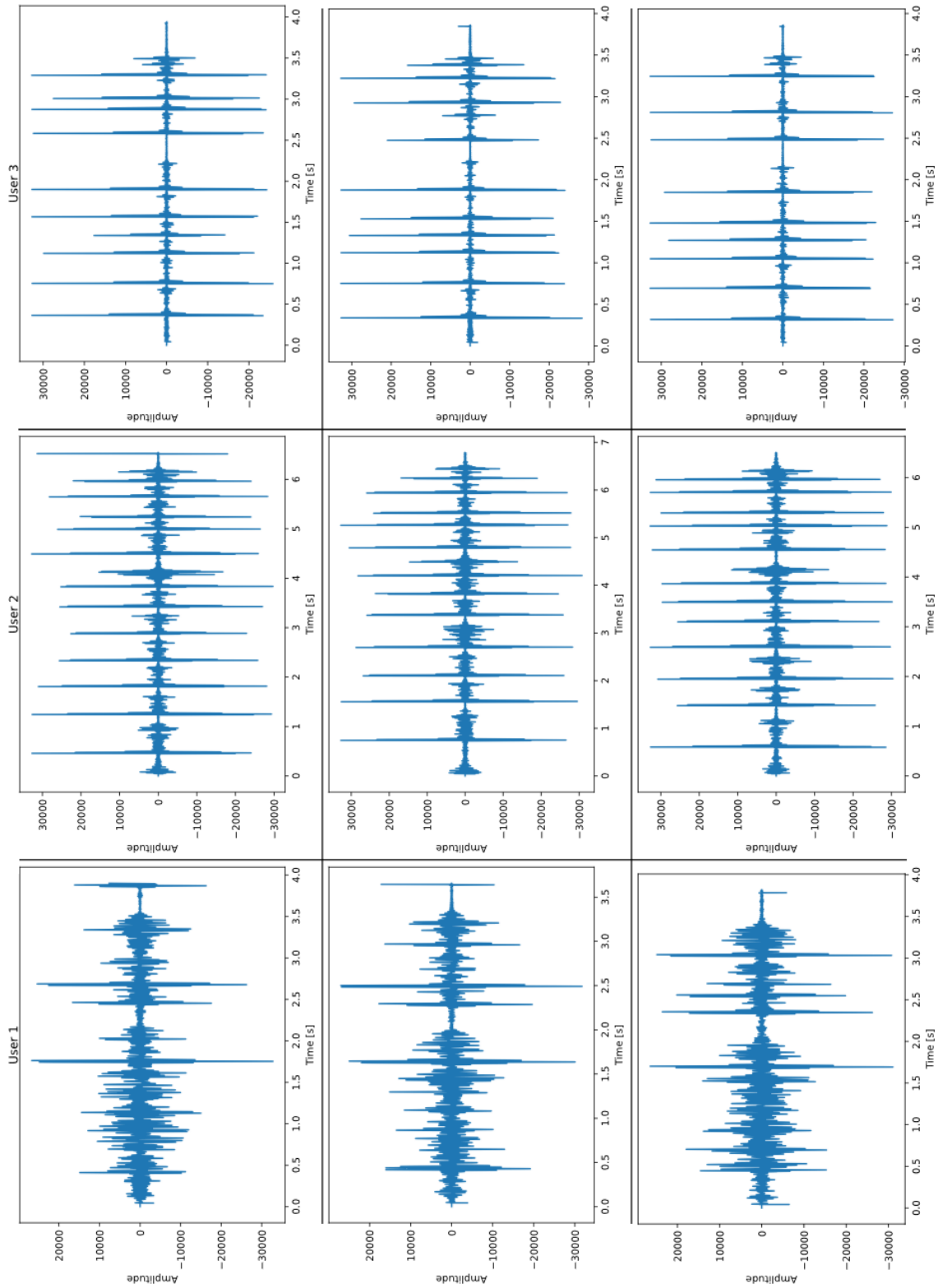
Figure 5.10: A visual representation of vibration and pressure changes captured by a piezoelectric sensor beneath the tablet.

Table 5.10: Piezoelectric sensor data from five participants, each completing five trials with a stylus, comparing intra-person and inter-person variations. *Piezoelectric sensor 1* corresponds to the rectangular model, while *piezoelectric sensor 2* corresponds to the circular model.

Measurement	Stylus		Finger	
	Intra-Person	Inter-Person	Intra-Person	Inter-Person
Piezoelectric sensor 1	73,069.27	103,811.43	81,153.56	110,985.47
Piezoelectric sensor 2	16,213.3	26,570.19	13,116.84	24,172.57

coordinates represent the position of an event within a specific view, with the view space, in this case, being the area on the tablet screen where users can write. The findings indicate that intra-person variability is consistently lower than inter-person variability, suggesting that users maintain stable and repeatable input behaviors across trials, resulting in distinct movement trajectories on the tablet surface.

Fig. 5.10 displays vibration readings and pressure variations from a piezoelectric sensor during tablet writing by three users, each writing the same sentence three times. Although two distinct piezo sensor models with different properties were positioned underneath the tablet, the data relates to the larger sensor, which captured a higher amplitude signal. Notable disparities are seen in the quantity and intensity of signal spikes, signal duration, as well as in the variation of amplitude strength throughout the writing. The observed variations in the piezo sensor readings may reflect individual differences in writing habits, speed, the intensity and direction of the exerted force, and stylus handling.

The data in Table 5.10 indicates that both piezoelectric sensors exhibit higher variation between different people compared to repeated actions by the same person. When comparing the two sensors, the first sensor generally exhibits higher variability than the second, both in intra-person and inter-person measurements. This suggests that the first sensor may be more responsive to variations in pressure or vibration, whether observed between different individuals or in repeated attempts by the same person.

5.2. Classification potential results

As discussed in Section 4.1., the proposed CP metric that evaluates the trade-off between inter-person and intra-person variability helps guide decisions about sensor inclusion or exclusion from the fusion. Tables 5.11 and 5.12 compare intra-person and inter-person variability, along with classification potential, for various sensors and measurements. The data is sorted in descending order, from the highest to the lowest CP, to highlight the sensors or measurements with the highest classification potential. In both tables, "Piezoelectric sensor 1" designates the larger rectangular model, while "Piezoelectric sensor 2" represents the smaller circular model.

The accelerometer appears to outperform other sensors or measurements in terms of CP, especially in the finger-writing task. This reveals that acceleration patterns, influenced by unique hand movements and writing styles during handwriting, may be highly individualized and remain consistent within each person.

The magnetometer seems to be a valuable component for both input modalities,

Table 5.11: A comparison of the classification potential (CP) for different sensors and measurements from the proposed apparatus. The data was collected during stylus-writing tasks where five participants completed five trials. The units are consistent with those from tables in Chapter 5.1.

Sensor/Measurement	Intra-Person Variability	Inter-Person Variability	Classification Potential (CP)
Accelerometer	8.5	117.23	13.79
Stylus orientation	0.28	2.76	9.86
Magnetometer	63.95	493.07	7.71
Stylus translation	1.94	8.06	4.15
Stylus rotation	5.42	18.52	3.42
Pressure	0.34	0.97	2.85
Gyroscope	198.83	495.22	2.49
Touch position	103.37	220.4	2.13
Piezoelectric sensor 2	16,213.3	26,570.19	1.64
Piezoelectric sensor 1	73,069.27	103,811.43	1.42
Velocity	329,110.2	319,536.78	0.97

Table 5.12: A comparison of classification potential for various sensors, based on finger-writing tasks where five participants completed five trials. The units match previous tables in Chapter 5.1.

Sensor/Measurement	Intra-Person Variability	Inter-Person Variability	Classification Potential (CP)
Accelerometer	4.7	124.68	26.53
Magnetometer	80.61	542.8	6.73
Touch size area	0.02	0.09	4.5
Gyroscope	162.89	405.33	2.49
Touch position	100.89	226.56	2.25
Piezoelectric sensor 2	13,116.84	24,172.57	1.84
Piezoelectric sensor 1	81,153.56	110,985.47	1.37
Velocity	1561.99	2093.88	1.34

demonstrating notable CP values across both stylus and finger writing tasks. The sensor captures subtle variations in the magnetic field during the writing process, suggesting that the manner in which users move the stylus or finger through space while writing creates distinct and unique "magnetic signatures". These signatures can help distinguish between different individuals, whether they are writing with a stylus or a finger.

Stylus orientation values, which are only available for stylus writing, show notably lower intra-person and inter-person variability. However, these values also stand out with a higher CP compared to other sensor readings. This implies that the way individuals hold and angle a stylus, as indicated by the touch area orientation, may contain personalized characteristics that could be useful for identification purposes. For finger writing, the touch size area might be an important factor in terms of CP, although it shows notably lower intra-person and inter-person variability comparing to other measurements.

Regarding the remaining sensors and measurements from Tables 5.11 and 5.12, they provide varying levels of classification potential. For instance, stylus translation and rotation demonstrate moderate classification potential compared to other measurements, while both piezoelectric sensors exhibit relatively low CP values across both writing tasks. Similarly, velocity readings for both input modalities show the lowest CP, indicating poor classification potential and minimal expected contribution to the sensor fusion setup.

5.3.   Classification results

To account for the inherently non-deterministic nature of neural networks and their training process, the feature extractor and classifier were each trained and evaluated over 30 repeated runs. The accuracies obtained from these runs were then averaged, and the standard deviation was calculated.

The mean values and standard deviations of model accuracies are presented in Fig. 5.11. As anticipated, the highest scores were predictably obtained for signatures, regardless of input modality (stylus or finger). This trend was observed even with the smallest train set size (10). It can be likely attributed to the unique and distinctive nature of signatures as a form of handwriting [22]. In biometrics, handwritten signatures are classified as behavioral traits, representing the learned patterns and habits inherent in an individual's handwriting. The repetitive strokes and patterns in signature writing may help the model to more effectively learn and recognize these unique features. In contrast, all participants wrote the same sentences, words, and letters, which may have contributed to less variability. Since each person's signature was unique, introducing noticeable inter-person differences, it can be argued that this may explain why accuracies were highest specifically for signature writing.

In all instances, higher accuracies were observed when participants wrote individual words as opposed to sentences. While sentence writing yields more dynamic information than writing specific words, the intra-person variability in sentence composition introduces noise, making it more challenging to distinguish between individuals. This could be attributed to potential variations in spacing between words in a sentence, a variability absent when only single words are written exclusively. Since writing sentences takes more time and requires more space, some degree of variation across different attempts by the same participant is to be expected. When focusing only on words, people could experience reduced cognitive load [131], enabling them to concentrate better on accurately forming each word. Additionally, repeatedly writing shorter words may promote greater consistency in execution, compared to writing longer, more complex sentences that could require greater focus and increased mental processing and planning [132].

The lowest accuracies were found for letters across both input modalities, with participants using a stylus and a finger. This may be due to the shorter time required to write
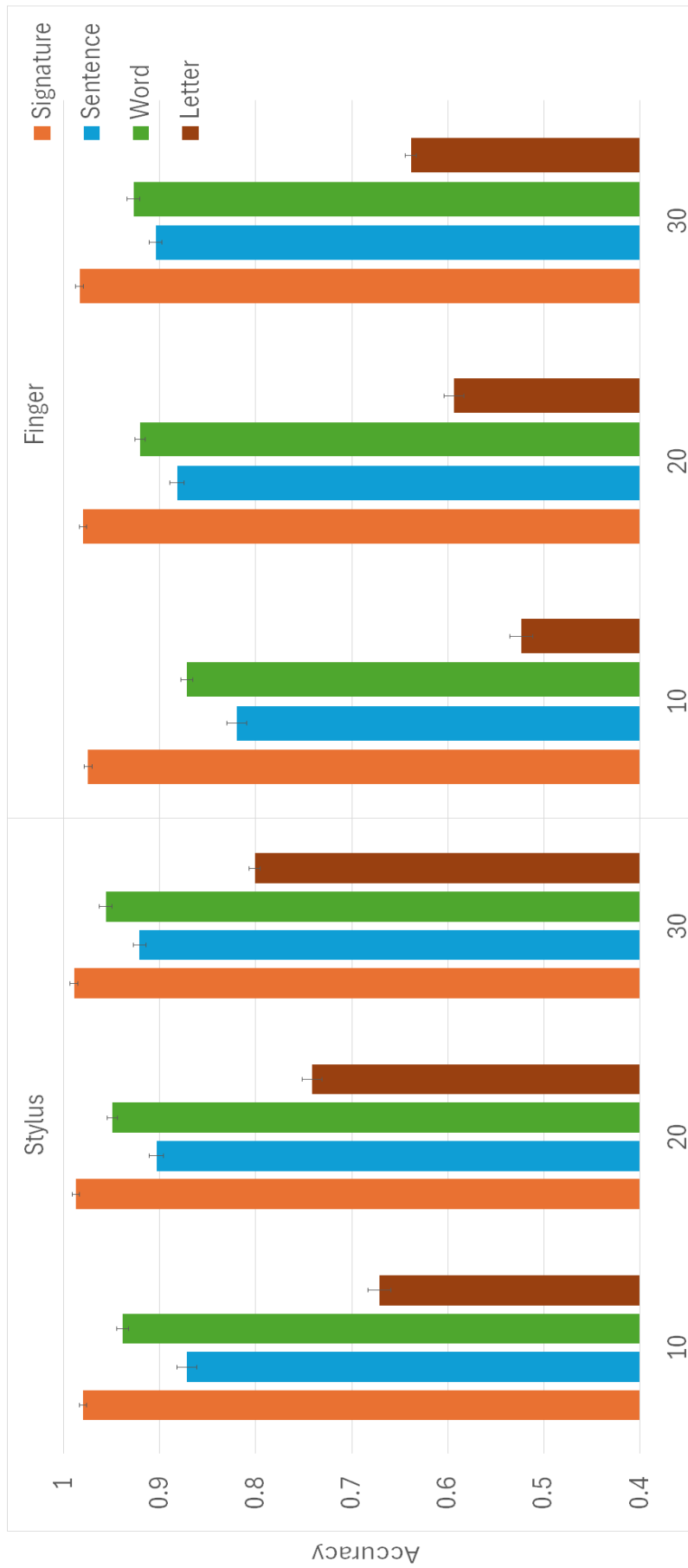
Figure 5.11: The mean values and standard deviations of classification accuracies for two input modalities (stylus and finger) are presented for three train set sizes (10, 20, and 30) and four collected handwriting forms (signature, sentence, word, and letter).

individual letters compared to sentences and words, which results in a smaller amount of data available for analysis. The limited information gathered from writing individual letters could make distinguishing between individuals more challenging. With reduced data, identifying enough unique features for the model might prove difficult, especially given the less precise nature of finger input [133].

The higher accuracy achieved with a stylus compared to a finger is likely due to the greater precision and control it offers. Using a stylus provides finer motor control, resulting in greater accuracy and consistency in writing [134]. Many individuals find that writing with a pen-like tool feels more intuitive, as it closely mimics the experience of writing on paper. In addition, using a stylus can minimize screen smudging and improve visibility, a frequent issue when writing with fingers. The lower accuracy observed with finger input can be attributed to the limited precision of using a finger on a touchscreen, commonly known as the fat-finger effect [135]. Unlike a stylus, fingers are broader and less stable, leading to more variable and inconsistent writing [136].

5.4. Main effects

A three-way $(2 \times 3 \times 4)$ repeated measures ANOVA was conducted to assess the influence of three within-subject factors on the model's accuracy: input modality (stylus, finger), train set size $(10, 20, 30)$, and collected handwriting forms (signatures, short sentences, words, and individual letters). To address violations of sphericity, the Greenhouse-Geisser correction (ϵ) was applied in instances where violations of sphericity had been detected. When statistically significant effects were identified, post hoc pairwise comparisons were performed using the Bonferroni correction.

Mauchly's test indicated a violation in the assumption of sphericity only for the variable related to the collected handwriting form: $\chi^2(5) = 23.866, p < 0.001$.

The results of the repeated measures RM ANOVA indicated significant main effects for the examined factors as follows:

1. **Input modality** – Model's mean accuracy differs statistically significantly between two input modalities, $F(1, 29) = 10915.991, p < 0.001, \eta^2 = 0.997$.

2. **Train set size** – A size of the train set significantly affects the accuracy of the

model, $F(2, 58) = 3615.583$, $p < 0.001$, $\eta^2 = 0.992$.

3. **Handwriting form** – Significant main effect was observed for collected handwriting form as well, $F(2.193, 63.609) = 38213.282$, $p < 0.001$, $\epsilon = 0.731$, $\eta^2 = 0.999$.

Consequently, in terms of input modality, stylus-based handwriting yields higher model performance (0.892 ± 0.103) compared to finger-based writing (0.835 ± 0.16), with a statistically significant difference ($p < 0.001$). Utilizing a stylus appears to produce more consistent handwriting strokes and patterns across multiple writing attempts. On the other hand, writing with a finger on a touchscreen tends to introduce greater variability, likely due to reduced precision and control over writing movements.

Regarding the train set size, post hoc analysis showed how the model's accuracy significantly improves as the train set size increases. In particular, the model showed a significant improvement in accuracy with a train set size of 30 (0.89 ± 0.118), as opposed to train set sizes of 20 (0.87 ± 0.136) and 10 (0.831 ± 0.16), with $p < 0.001$ for both pairwise comparisons. The accuracy difference between the train set sizes of 10 and 20 was also found to be statistically significant ($p < 0.001$). This supports the initial expectations, as a larger dataset helps to minimize overfitting by reducing the likelihood of the model memorizing specific examples from the train set.

In terms of the collected handwriting form, the model achieves the highest accuracy when using signatures (0.982 ± 0.005), followed by words (0.927 ± 0.03), sentences (0.884 ± 0.036), and individual letters (0.661 ± 0.1). Post hoc analysis indicated a significant difference in all pairwise comparisons observed herein, with $p < 0.001$ for each case.

5.5. Comparison to existing approaches

When comparing with other studies, Table 5.13 provides a summary of relevant research in this field. Unlike most studies in the related work that focus on verification systems, the table highlights research that, like this study, identifies individuals through handwriting. It is important to emphasize that the number of such studies is notably lower compared to those focusing on verification. Therefore, among many studies, only those related to the recognition of individuals through handwriting were considered.

In addition to the list of sensors employed, the table presents the different handwriting forms used for user identification, along with the reported accuracy from the corresponding studies. It shows that the handwriting forms used mainly refer to signatures, in contrast to this research concerning broader handwriting-based identification. The comparison becomes more relevant when focused only on studies that involve online (dynamic) handwriting, particularly those conducted on smartphones or tablets. Compared to similar research, the proposed approach supports the inclusion of a larger number of sensors for person recognition through handwriting, instead of using only one or a rather small number of specific sensors. As shown in the table, many existing studies rely on image-capturing devices or digital tablets, rather than utilizing sensor fusion.

The proposed solution has achieved high accuracy in signature recognition (0.982), which is adequately comparable to the results of other studies reviewed. Additionally, it has shown high accuracy in recognizing words (0.927) and short sentences (0.884), while exhibiting lower accuracy for individual letters (0.661). The results obtained reinforce the concept of using different forms of handwriting, extending beyond signatures, as potential alternative methods for recognizing individuals through handwriting. It is important to highlight that the accuracies provided here represent the mean classification accuracies for two input modalities (stylus and finger) and three different train test sizes (10, 20, 30). Therefore, the results of the ablation study presented next may be more relevant for comparison with related studies. Despite achieving competitive results, certain challenges remain, such as the lower accuracy for letter-based recognition, highlighting potential areas for future improvements.

Study	Device/sensor	Handwriting form	Accuracy (%)
Li et al. (2024) [93]	Image-capturing device	Handwritten signatures	98.50
Rahim et al. (2024) [95]	Tablet and a digital pen	Bengali handwriting samples, focusing on 10 distinct keywords	94.62
Leghari et al. (2024) [79]	Smartphone with a stylus	Handwritten signatures	96.00
Hasan et al. (2024) [85]	Tablet and a digital pen	10 specific keywords	98.31
Çiftçi and Tekin (2024) [86]	Image-capturing device	Handwritten signatures	98.77
Chuen et al. (2023) [91]	Microsoft Kinect camera	In-air signatures	93.00
Khoh et al. (2023) [73]	Microsoft Kinect camera	In-air signatures	97.43
Culqui-Culqui et al (2022) [89]	Image-capturing device	Handwritten signatures	98.03
Rexit et al. (2022) [82]	Image-capturing device	Signatures in Chinese, Uyghur	92.95
Rexit et al. (2022) [83]	Image-capturing device	Handwritten signatures in Uyghur, Kazakh, and Han languages	98.40
Begum et al. (2021) [77]	Tablet and a digital pen	Defined keywords and phrases	98.00
Kette et al. (2021) [74]	Image-capturing device	Handwritten signatures	90.00
Ghosh et al. (2021) [96]	Leap Motion sensor	In-air signatures	94.63
Sriwathsan et al. (2021) [76]	Image-capturing device	Handwritten signatures	96.87
Seki (2021) [92]	Anoto digital pen	Four Kanji handwriting samples combined into a "pseudo-signature"	95.62
Akash et al. (2020) [78]	Tablet and a digital pen	Defined keywords and phrases	87.00
Pokharel et al. (2020) [87]	Tablet and a digital pen	Handwritten signatures	95.20
Gumusbas and Yildirim (2019) [94]	Image-capturing device	Handwritten signatures	98.80
Al-Shamaileh et al. (2019) [84]	Digital tablet device	Arabic handwriting, specific text-dependent words	81.35
Çalik et al. (2019) [88]	Image-capturing device	Handwritten signatures	98.30
Dargan et al. (2019) [80]	Image-capturing device	Handwritten Devanagari characters	91.53
Mo et al. (2019) [81]	Image-capturing device	Signatures in Kirgiz and Uyghur	97.95
Hezil et al. (2018) [75]	Image-capturing device	Handwritten signatures	97.30
Behera et al. (2018) [90]	Leap Motion sensor	In-air signatures	95.00

Table 5.13: Comparison of recent studies on handwriting-based person recognition, detailing the employed devices or sensors, handwriting forms, and the reported performance in terms of accuracy metrics.

5.6. Ablation study

The purpose of the ablation study is to assess how each of the six sensor subsets influences the accuracy of handwriting-based person recognition. This procedure entails removing one sensor subset at a time to evaluate its effect on the model's accuracy with the remaining subsets. For example, if excluding a specific sensor subset leads to an improvement in accuracy, it indicates that this subset negatively affects user identification in the given context. Therefore, the model training is additionally conducted for six setups independently, each of which had five active sensor subsets.

The results regarding the obtained accuracies are presented in Table 5.14. In the respective procedure, the train set size was fixed to 30, as this particular size demonstrated the highest accuracy results. The achieved accuracies are presented for two input modalities (stylus or finger) and different handwriting forms (signatures, sentences, words, letters). Each row represents cases where a particular sensor subset was excluded, with the exception of the final row, which presents results when all sensor subsets in the fusion were enabled. The table data indicates that excluding certain sensor subsets influences accuracy differently across the observed categories.

For recognizing a signature using a stylus, the model achieves its highest accuracy (0.9911) when either the touchscreen or piezo sensors are excluded. Including all sensors in the fusion results in a slight decrease in the model's accuracy for both stylus (0.9866) and finger (0.9844) inputs. For input with a stylus, the outcome suggests that these sensors do not contribute to improving signature recognition. A possible reason for this is the nature of signature dynamics, which involve a pre-learned action. A signature is generally a short, established, and specific movement pattern that often involves predetermined motions, leading to more stable and repeatable dynamics. Therefore, their shape and execution remain relatively stable across instances, especially when compared to writing sentences or words, which involve more variability in motion and structure. As a result, signatures may not depend on detailed touch position tracking (from the touchscreen) or precise force and pressure tracking (offered by piezo sensors) to attain high recognition accuracy. Additionally, for all other handwriting forms (sentences, words, and letters), recognition accuracies are consistently higher when all sensors are included. Readings from multiple sensors appears to be beneficial for these forms, as they likely depend on

Table 5.14: Impact of different sensor subsets on handwriting-based person recognition accuracy with a train set size of 30. The table displays accuracy scores for various handwriting forms (signatures, sentences, words, letters) using stylus and finger inputs, with the highest scores highlighted in bold.

(a) Recognition accuracy using a stylus

Excluded sensor subset	Signatures	Sentences	Words	Letters
Touchscreen	**0.9911**	0.9141	0.9554	0.7945
Magnetometer	0.9866	0.8862	0.9018	0.7570
Input specific	0.9821	0.8806	0.9275	0.7683
Piezos	**0.9911**	0.9241	0.9531	0.8060
Smartwatch	0.9799	0.8973	0.9241	0.7144
Visual tracking	0.9799	0.9018	0.9230	0.6719
All included	0.9866	**0.9263**	**0.9576**	**0.8130**

(b) Recognition accuracy using a finger

Excluded sensor subset	Signatures	Sentences	Words	Letters
Touchscreen	0.9710	0.8884	0.9163	0.6347
Magnetometer	0.9665	0.6953	0.7020	0.5019
Input specific	0.9821	**0.9163**	0.9163	0.6344
Piezos	**0.9911**	0.9141	**0.9286**	0.6293
Smartwatch	0.9821	0.8683	0.8683	0.4973
Visual tracking	0.9821	0.9107	0.9241	**0.6447**
All included	0.9844	0.9051	0.9185	0.6369

the fusion of diverse signals to capture a broader range of variations that occur when writing sentences and words. In different writing attempts, each word or sentence is more prone to vary as a person may change writing speed, punctuation, spacing, letter height, and other aspects. Such changes lead to greater variations in the structure and dynamics of the movement, especially when transitioning from one word to another.

In contrast, when finger input is used, the model achieves its highest accuracy (0.9911) when the piezo sensor is excluded, suggesting it negatively affects accuracy in this case. Compared to a stylus, which allows for more precise control and input, a finger is less accurate and produces broader, less defined touch points. Furthermore, finger input causes more variability in the interaction with the touchscreen due to the larger touch area and inconsistent pressure. The piezo sensor, which measures pressure variations, may not contribute notably to finger input, as the touch pressure and force applied with a finger are less stable and harder to capture accurately. This indicates the variability in a finger's touch dynamics makes piezo sensors less beneficial for recognizing persons. The ablation

study confirms the relevance and validity of the introduced CP metric, as seen in both Tables 5.11 and 5.12 from Section 5.2. The tables show that the CP values for both piezoelectric sensors are notably lower compared to the measurements from other sensors.

Regarding the impact of magnetic field measurements, excluding the magnetometer leads to a slight decrease in accuracy for both stylus and finger inputs. The magnetometer appears beneficial in capturing distinct patterns and movements during the writing process. The sensor helps differentiate the dynamics and motion involved in writing, regardless of whether a permanent magnet is mounted on the stylus or the finger. This is also consistent with the previously obtained data related to the CP metric, as shown in both Tables 5.11 and 5.12. It is evident that for both input modalities, the CP values for the magnetometer are higher compared to the measurements from other sensors.

In tasks involving sentence writing with a stylus, the model achieves its best accuracy (0.9263) when the full fusion setup is enabled. This indicates that a multi-sensor approach provides valuable information for sentence recognition with stylus input, which involves more complex and extended strokes compared to signatures. On the other hand, when only finger input data is used, the model achieves its peak accuracy (0.9163) when input-specific sensors are omitted. The decrease in accuracy can be attributed to the lack of input specific sensor data when using finger input, which introduces noise into the analysis. While stylus input provides valuable data on tilt and pressure levels, these measurements are not available with finger input, where they are replaced with zeros. This substitution can introduce noise and inconsistencies, leading to variations that negatively affect the model's accuracy. Interestingly, in sentence-based tasks with finger input, excluding the magnetometer measurements notably reduces the model's accuracy, dropping to 0.6953. As with stylus-writing, this highlights the importance of magnetometer data, showing its ability to detect subtle changes in the magnetic field during writing with a finger. Even slight differences in how a person moves their finger, like variations in angle, tilt, and rotation, can form unique patterns that the magnetometer can detect. All of this further confirms the importance of the introduced CP metric, with magnetometer and stylus orientation readings showing higher CP values than other sensors.

The recognition of written words follows a similar pattern, with the model achieving its highest accuracy of 0.9576 for stylus input when all sensors are active. This outcome reaffirms the importance of sensor fusion utilized in this research, this time for word

recognition tasks. In comparison, the model based on finger input achieves its best performance (0.9286) when the readings from the piezo sensor are excluded. Since the same observation applies to finger-based signatures, it suggests that the piezo sensor may introduce noise. This could potentially reduce its effectiveness in capturing the tactile details of finger movements. As mentioned, the reason could be that the less precise nature of finger input, along with variable pressure and a broader contact area, might contribute to greater inconsistency in touchscreen interaction when writing words.

Similar to sentence writing, removing the magnetometer sensor data when writing words notably decreases the model's accuracy, especially for finger input. This highlights the important contribution of the magnetometer to word classification tasks. Excluding the smartwatch sensors (accelerometer and gyroscope) also leads to a decrease in model accuracy during word writing, especially when considering finger-writing (0.8683). This could be attributed to the larger and more varied wrist movements involved in finger-based writing, which the smartwatch is capable of capturing. In contrast, when using a stylus, the wrist remains relatively stable, and finer control is managed by the fingers, reducing the smartwatch's impact on recognition accuracy. This once again justifies the introduction of the CP metric, where Tables 5.11 and 5.12 demonstrate a notably higher classification potential for the accelerometer and gyroscope readings when compared to the measurements from other sensors.

Letter recognition follows a a slightly different trend, with the model achieving its highest accuracy for stylus input (0.813) when all sensors are included. Here, the fusion of sensors also demonstrates its benefits, indicating that detailed multi-sensor data is essential for interpreting finer handwriting details. In contrast, the model based on finger input reaches its peak accuracy (0.6447) when the visual tracking sensor subset is excluded. These results could be anticipated, as the ArUco marker was only used during stylus writing to detect tilt, which was not applicable to finger input. Consequently, the absence of samples in this scenario could potentially introduce noise and inconsistencies in the analysis. Similar to word writing, excluding the smartwatch sensors during individual letter writing leads to a notable decrease in model accuracy (0.4973). This further underscores their importance in capturing the fine-grained movements and dynamics of finger writing, particularly for handwriting forms like individual letters.

Notably, accuracies for letter classification tend to be lower across most sensor subsets,

suggesting that letters are inherently more challenging to classify accurately. This could be attributed to the subtle differences in writing certain letters, the short writing duration, and the similarities in the shapes used to form specific letters.

Overall, the results indicate that the best sensor fusion subset is greatly influenced by both the handwriting form and the input method. In cases when all sensors are included, the accuracy of the model improves for recognizing sentences, words, and individual letters written with a stylus. This underscores the benefit of employing a comprehensive multi-sensor setup for capturing handwriting dynamics. Hence, it can be concluded that incorporating additional sensors can certainly enhance the person recognition accuracy, thereby justifying their inclusion into the experiment apparatus.

Chapter 6

CONCLUSION

The presented research focuses on biometric identification systems by analyzing the unique patterns and characteristics of an individual's handwriting, with a particular emphasis on the dynamic writing process on touchscreen devices. Existing studies in this field highlight the dominance of specific research approaches. First, most of the reviewed studies focus on verification systems, which intends to differentiate authentic from forged handwriting samples. Second, the emphasis is primarily on analyzing signatures, while the broader context of user recognition across different forms of handwriting is often overlooked. Third, in most cases, the proposed solutions rely on custom-designed devices or specialized sensors rather than off-the-shelf components to gather biometric data.

To address the identified research gap, a novel handwriting-based user recognition technique is introduced. It broadens the focus from verification to a wider range of handwriting identification tasks, involving a general process of detecting the handwriting owner through multi-class classification. This approach extends beyond signatures for handwriting-based person recognition by incorporating various forms of handwriting, namely short sentences, words, and individual letters. The objective was to explore alternative approaches for person recognition, moving beyond the traditional dependence on signatures, which are generally regarded as pre-learned actions. To collect data on the dynamics of handwriting and signing, commonly available smart devices were utilized, leveraging their built-in sensors along with additional external sensors and accessories. Unlike most existing systems, which primarily rely on a single sensor (or a rather small number of sensors), the novel experimental apparatus incorporates a sensor fusion concept,

aiming to collect and leverage a much larger volume of handwriting dynamics information. Additionally, two different methods of handwriting input are analyzed, based on a stylus and a finger.

In a controlled experiment focused on acquiring biometric characteristics of touchscreen handwriting, a total of sixty participants were involved. The outcome of the conducted experiment is an original dataset that includes sensor readings from various sensors, specifically utilizing a camera, accelerometer, gyroscope, magnetometer, touchscreen, and two piezo sensors. These readings were collected during both stylus and finger input across a variety of writing tasks. In addition to providing their own signatures, users also wrote short sentences, words, and individual letters. This diverse range of writing tasks enabled the collection of detailed biometric handwriting data, resulting in a comprehensive dataset of handwriting dynamics across different input modalities.

The hypothesis that dynamic features from touchscreen handwriting show consistency within individuals and variability across individuals is supported by sensor measurements, which reveal notable differences between intra-person (within the same user) and inter-person (between different users) variability. This can be clearly observed in the tables displaying the DTW metric, as well as in the corresponding graphs, which further highlight the potential of these measurements in differentiating between users. These distinctions are consistently observed across both stylus and finger input modalities.

To facilitate handwriting recognition, a CNN-based model was developed for feature extraction and classification tasks. The accuracy of the model was analyzed with regard to the influence of three different factors: train set size, input modality, and handwriting form. During the model training process, train set sizes of 10, 20, and 30 were used. The results demonstrate a statistically significant effect of set size on the model's accuracy, with the highest accuracies achieved using the largest train set size.

The statistical analysis also revealed a significant difference in the model's accuracy when writing with a stylus compared to a finger. In line with findings from other studies, finger-writing is characterized by larger, less precise motions, whereas stylus usage improves precision and ensures greater consistency. Experiment participants also reported that the stylus provided greater precision than finger, closely resembling the experience of writing with a traditional pen and paper.

Among the different handwriting forms examined, the highest accuracy was achieved

for signatures (0.982), which is anticipated, as each signature reflects a pre-learned, unique action specific to the individual participant. However, recognition accuracy was also found to be quite high for other handwriting forms, particularly for words (0.927). The model showed a statistically significant improvement in accuracy when recognizing words, compared to both sentences (0.884) and letters (0.661). The lower accuracy for sentences may be attributed to their typically inconsistent spacing and the greater time and cognitive effort required for composition. In contrast, writing shorter words repetitively may reinforce muscle memory, potentially reducing the mental effort involved and allowing for more consistent word formation. The lowest accuracy of the model in recognizing individual letters may stem from the small number of strokes involved, which consequently leads to a considerably reduced amount of sensor data being collected.

Although the model's accuracies are slightly lower for words and sentences, the outcomes convincingly support the hypothesis that, besides signatures, other forms of handwriting can be used for person identification. Furthermore, these findings also confirm that deep learning models based on CNNs can be developed for successful person recognition based on touchscreen handwriting, for both stylus and finger input.

An ablation study was carried out to analyze the impact of individual sensors within the fusion-based setup. The outcomes from this analysis provide strong evidence supporting the hypothesis that the implementation of a sensor fusion approach could improve the accuracy of person recognition based on touchscreen handwriting. Overall, the results indicate that the best subset of sensors for fusion is strongly influenced by both the type of handwriting and the input modality used. When all sensors are enabled, the model's accuracy improves for recognizing sentences, words, and letters written with a stylus. This highlights the benefit of utilizing a comprehensive multi-sensor setup. Hence, it can be concluded that incorporating additional sensors can certainly enhance the accuracy of the model, thereby justifying their integration into the experiment apparatus.

Concerning limitations, the experimental setup faced several notable challenges, primarily associated with device connectivity and battery life. These issues impacted both the overall performance and the data collection process. One of the foremost challenges was the connectivity issue between the tablet and the smartwatch devices. While a connection was successfully established between the smartwatch and the smartphone, it failed to pair with the tablet. This limitation required routing commands for initiating and ter-

minating data collection from the tablet to the smartphone, which then communicated with the smartwatch. The indirect communication setup introduced additional complexity and dependency, which could be easily avoided if the manufacturer allowed direct pairing between the smartwatch and the tablet devices.

The battery life of the devices also proved to be a considerable constraint throughout the experiment. Continuous operation of the sensors, such as the accelerometer and gyroscope, was not feasible due to the high energy demands, which limited the duration of the experiment and the number of participants who could be included. To address this, the system could be optimized to minimize energy consumption while ensuring efficient data collection for a larger sample size. This could possibly be achieved by adjusting sensor parameters, such as reducing the sampling rate.

Another important limitation was due to the functionality of the S Pen Remote SDK, which does not support the collection of pressure levels or tilt data from the stylus. This constraint was overcome by developing a workaround that utilizes external sensors, including piezoelectric sensors for pressure measurement and ArUco marker-based video tracking for tilt estimation. While the workaround solution allowed for the extraction of stylus tilt data through ArUco marker tracking, it had a major drawback with the high energy consumption required for video recording. Enhancing the SDK to support these functionalities natively would reduce energy consumption and eliminate the need for the current solution. Another approach would involve using styluses with open-source SDKs or alternative digital pens that inherently support pressure and tilt data acquisition.

Regarding future work, research could focus on the consistency of biometric features through a longitudinal study to examine how an individual's handwriting characteristics evolve over time. Such a research would explore the impact of factors such as aging, changes in writing habits, and physical conditions on handwriting dynamics.

Additionally, future studies could investigate handwriting biometrics across different languages and scripts. Collecting samples from diverse character sets would provide insight into the universality of handwriting recognition techniques and help determine whether adaptations are necessary for non-Latin scripts.

The effectiveness of the proposed solution could be explored by integrating new sensors or devices into the existing fusion-based apparatus. For example, incorporating motion capture devices or electromyography (EMG) sensors could enable a more detailed analysis

of muscle activity during handwriting. Optical or infrared sensors could also be utilized to track hand movements, potentially capturing subtle biometric markers that may enhance the model's ability to distinguish individuals. Different placements of piezoelectric sensors could also be examined to determine the best positions for capturing pressure and vibrations on the tablet screen. By experimenting with various models of piezoelectric sensors, researchers could gain valuable insights into their performance, sensitivity, and overall suitability for analyzing handwriting dynamics.

The presented apparatus can also extend beyond identification, particularly in the field of human-computer interaction. Since it can monitor all types of touchscreen interactions, it offers the potential to study users' interaction patterns during activities such as touch-drawing or performing basic operations like swipe, pinch, drag, and point-and-select. Finally, another important consideration is the practical application of the proposed solution in various real-world scenarios that, from a security perspective, either support or require handwriting-based identification.

LITERATURE

[1] Waleed A. Hammood, Ruzaini Abdullah, Omar A. Hammood, Salwana Mohamad Asmara, Mohammed A. Al-Sharafi, and Ali Muttaleb Hasan. A review of user authentication model for online banking system based on mobile imei number. *IOP Conference Series: Materials Science and Engineering*, 769(1):012061, February 2020. ISSN 1757-899X. doi: 10.1088/1757-899x/769/1/012061.

[2] Hongfang Shi, Wenying Zhang, Zimin Zhang, and Dewen Ding. Vulnerability analysis of chinese digital passwords related to atm pin using deep learning. *IEEE Transactions on Dependable and Secure Computing*, 20(4):2825–2835, July 2023. ISSN 2160-9209. doi: 10.1109/tdsc.2022.3188505.

[3] Joaquin Zermeno-Saldana and Jesus Arturo Perez-Diaz. A novel non-biometric multi-factor authentication system using audios and relationships. In *2023 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE, December 2023. doi: 10.1109/ieem58616.2023.10406725.

[4] Shoroog Albalawi, Lama Alshahrani, Nouf Albalawi, Reem Kilabi, and Aaeshah Alhakamy. A comprehensive overview on biometric authentication systems using artificial intelligence techniques. *International Journal of Advanced Computer Science and Applications*, 13(4), 2022. ISSN 2158-107X. doi: 10.14569/ijacsa.2022.0130491.

[5] Anfal Ahmed Aleidan, Qaisar Abbas, Yassine Daadaa, Imran Qureshi, Ganeshkumar Perumal, Mostafa E. A. Ibrahim, and Alaa E. S. Ahmed. Biometric-based human identification using ensemble-based technique and ecg signals. *Applied Sciences*, 13(16):9454, August 2023. ISSN 2076-3417. doi: 10.3390/app13169454.

[6] Shaymaa Adnan Abdulrahman and Bilal Alhayani. A comprehensive survey on the

biometric systems based on physiological and behavioural characteristics. *Materials Today: Proceedings*, 80:2642–2646, 2023. ISSN 2214-7853. doi: 10.1016/j.matpr. 2021.07.005.

[7] Abdeljebar Mansour, Nahid Eddermoug, Mohamed Sadik, Essaid Sabir, Mohamed Azmi, and Mostafa Jebbar. A lightweight seamless unimodal biometric authentication system. *Procedia Computer Science*, 231:190–197, 2024. ISSN 1877-0509. doi: 10.1016/j.procs.2023.12.192.

[8] O. N. KADHIM and M. H. ABDULAMEER. Biometric identification advances: Unimodal to multimodal fusion of face, palm, and iris features. *Advances in Electrical and Computer Engineering*, 24(1):91–98, 2024. ISSN 1844-7600. doi: 10.4316/aece.2024.01010.

[9] Harbi Al-Mahafzah, Tamer AbuKhalil, Malek Alksasbeh, and Bassam Alqaralleh. Multi-modal palm-print and hand-vein biometric recognition at sensor level fusion. *International Journal of Electrical and Computer Engineering (IJECE)*, 13(2):1954, April 2023. ISSN 2088-8708. doi: 10.11591/ijece.v13i2.pp1954-1963.

[10] David Palma and Pier Luca Montessoro. *Biometric-Based Human Recognition Systems: An Overview*, chapter 2. IntechOpen, July 2022. ISBN 9781803554570. doi: 10.5772/intechopen.101686.

[11] Saleh Alwahaishi and Jaroslav Zdralek. Biometric authentication security: An overview. In *2020 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*. IEEE, November 2020. doi: 10.1109/ccem50674.2020.00027.

[12] Pietro Melzi, Christian Rathgeb, Ruben Tolosana, Ruben Vera-Rodriguez, and Christoph Busch. An overview of privacy-enhancing technologies in biometric recognition, 2022.

[13] Mitesh Parmar, Nupur Puranik, Dhruva Joshi, Sonal Malpani, and Bhushan S. Thakare. State of Art Survey Signature Verification Techniques 2019. In *Asian Journal For Convergence In Technology (AJCT)*, 2020. URL https://api. semanticscholar.org/CorpusID:231661238.

[14] Moises Diaz, Miguel A. Ferrer, Donato Impedovo, Muhammad Imran Malik, Giuseppe Pirlo, and Réjean Plamondon. A perspective analysis of handwritten signature technology. *ACM Computing Surveys*, 51(6):1–39, January 2019. ISSN 1557-7341. doi: 10.1145/3274658.

[15] S. Pal, M. Blumenstein, and U. Pal. Off-line signature verification systems: a survey. In *Proceedings of the International Conference & Workshop on Emerging Trends in Technology - ICWET '11*, ICWET '11. ACM Press, 2011. doi: 10.1145/1980022. 1980163.

[16] Saba Mushtaq and A. H. Mir. Signature verification: A study. In *2013 4th International Conference on Computer and Communication Technology (ICCCT)*. IEEE, September 2013. doi: 10.1109/iccct.2013.6749637.

[17] M. Muzaffar Hameed, Rodina Ahmad, Miss Laiha Mat Kiah, and Ghulam Murtaza. Machine learning-based offline signature verification systems: A systematic review. *Signal Processing: Image Communication*, 93:116139, April 2021. ISSN 0923-5965. doi: 10.1016/j.image.2021.116139.

[18] Ashok Kumar and Karamjit Bhatia. A survey on offline handwritten signature verification system using writer dependent and independent approaches. In *2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Fall)*. IEEE, September 2016. doi: 10.1109/icaccaf.2016.7748998.

[19] L. B.Mahanta and Alpana Deka. A study on handwritten signature. *International Journal of Computer Applications*, 79(2):48–52, October 2013. ISSN 0975-8887. doi: 10.5120/13717-1489.

[20] Anjali Rohilla and Rajesh Kumar Bawa. *A Survey of Digitized Handwritten Signature Verification System*, pages 133–151. Springer Singapore, 2020. ISBN 9789811569074. doi: 10.1007/978-981-15-6907-4_8.

[21] Sudeep Tanwar, Mohammad S. Obaidat, Sudhanshu Tyagi, and Neeraj Kumar. *Online Signature-Based Biometric Recognition*, pages 255–285. Springer International Publishing, October 2018. ISBN 9783319987347. doi: 10.1007/978-3-319-98734-7_10.

[22] Zainab Hashim, Hanaa M. Ahmed, and Ahmed Hussein Alkhayyat. A comparative study among handwritten signature verification methods using machine learning techniques. *Scientific Programming*, 2022:1–17, October 2022. ISSN 1058-9244. doi: 10.1155/2022/8170424.

[23] Elias N. Zois, Dimitrios Tsourounis, Ilias Theodorakopoulos, Anastasios L. Kesidis, and George Economou. A comprehensive study of sparse representation techniques for offline signature verification. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 1(1):68–81, January 2019. ISSN 2637-6407. doi: 10.1109/tbiom. 2019.2897802.

[24] Amruta B. Jagtap, Dattatray D. Sawat, and Ravindra S. Hegadi. *Review on Offline Signature Verification: Datasets, Methods and Challenges*, pages 458–468. Springer Singapore, 2021. ISBN 9789811605079. doi: 10.1007/978-981-16-0507-9_38.

[25] J.F. Vargas, M.A. Ferrer, C.M. Travieso, and J.B. Alonso. Off-line signature verification based on grey level information using texture features. *Pattern Recognition*, 44(2):375–385, February 2011. ISSN 0031-3203. doi: 10.1016/j.patcog.2010.07.028.

[26] Mehr Yahya Durrani, Salabat Khan, and Shehzad Khalid. Versig: a new approach for online signature verification. *Cluster Computing*, 22(S3):7229–7239, August 2017. ISSN 1573-7543. doi: 10.1007/s10586-017-1129-4.

[27] Zhaoxiang Zhang, Kaiyue Wang, and Yunhong Wang. *A Survey of On-line Signature Verification*, pages 141–149. Springer Berlin Heidelberg, 2011. ISBN 9783642254499. doi: 10.1007/978-3-642-25449-9_18.

[28] Harmandeep Kaur and Munish Kumar. Signature identification and verification techniques: state-of-the-art work. *Journal of Ambient Intelligence and Humanized Computing*, 14(2):1027–1045, June 2021. ISSN 1868-5145. doi: 10.1007/ s12652-021-03356-w.

[29] Sonia Garcia-Salicetti, Nesma Houmani, Bao Ly-Van, Bernadette Dorizzi, Fernando Alonso-Fernandez, Julian Fierrez, Javier Ortega-Garcia, Claus Vielhauer, and Tobias Scheidat. *Online Handwritten Signature Verification*, pages 125–165. Springer London, 2009. ISBN 9781848002920. doi: 10.1007/978-1-84800-292-0_6.

[30] Ibraheem M Alharbi. Efficient handwritten signatures identification using machine learning. *Int. J. Adv. Comput. Sci. Appl.*, 14(3), 2023.

[31] Zainab Hashim, Hanaa Mohsin, and Ahmed Alkhayyat. Offline handwritten signature identification based on hybrid features and proposed deep model. *Iraqi Journal For Computer Science and Mathematics*, 5(1):220–236, February 2024.

[32] Alen Salkanovic, David Bačnar, Diego Sušanj, and Sandi Ljubic. A sensor-fusion-based experimental apparatus for collecting touchscreen handwriting biometric features. *Applied Sciences*, 14(23):11234, December 2024. ISSN 2076-3417. doi: 10.3390/app142311234.

[33] Alen Salkanovic, Diego Sušanj, Luka Batistić, and Sandi Ljubic. Beyond signatures: Leveraging sensor fusion for contextual handwriting recognition. *Preprints*, March 2025. doi: 10.20944/preprints202503.1018.v1. URL `https://doi.org/10.20944/preprints202503.1018.v1`.

[34] Sandi Ljubic, Franko Hržić, Alen Salkanovic, and Ivan Štajduhar. Augmenting around-device interaction by geomagnetic field built-in sensor utilization. *Sensors*, 21(9):3087, April 2021. ISSN 1424-8220. doi: 10.3390/s21093087.

[35] Mingda Han, Huanqi Yang, Tao Ni, Di Duan, Mengzhe Ruan, Yongliang Chen, Jia Zhang, and Weitao Xu. mmSign: mmWave-based Few-Shot Online Handwritten Signature Verification. *ACM Transactions on Sensor Networks*, 20(4):1–31, May 2024. ISSN 1550-4867. doi: 10.1145/3605945.

[36] Wanqing Li, Tongtong He, Nan Jing, and Lin Wang. mmHSV: In-Air Handwritten Signature Verification via Millimeter-Wave Radar. *ACM Transactions on Internet of Things*, 4(4):1–22, November 2023. ISSN 2577-6207. doi: 10.1145/3614443.

[37] Feng Ding, Dong Wang, Qian Zhang, and Run Zhao. ASSV: Handwritten Signature Verification Using Acoustic Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(3):1–22, September 2019. ISSN 2474-9567. doi: 10.1145/3351238.

[38] Mengqi Chen, Jiawei Lin, Yongpan Zou, and Kaishun Wu. Acoustic Sensing Based on Online Handwritten Signature Verification. *Sensors*, 22(23):9343, November 2022. ISSN 1424-8220. doi: 10.3390/s22239343.

[39] Run Zhao, Dong Wang, Qian Zhang, Xueyi Jin, and Ke Liu. Smartphone-based Handwritten Signature Verification using Acoustic Signals. *Proceedings of the ACM on Human-Computer Interaction*, 5(ISS):1–26, November 2021. ISSN 2573-0142. doi: 10.1145/3488544.

[40] Mustafa Semih Sadak, Nihan Kahraman, and Umut Uludag. Handwritten Signature Verification System Using Sound as a Feature. In *2020 43rd International Conference on Telecommunications and Signal Processing (TSP)*, pages 365–368. IEEE, July 2020. doi: 10.1109/tsp49548.2020.9163442.

[41] Gen Li and Hiroyuki Sato. Handwritten Signature Authentication Using Smartwatch Motion Sensors. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, volume 2, pages 1589–1596. IEEE, July 2020. doi: 10.1109/compsac48688.2020.00-28.

[42] Raghavendra Ramachandra, Sushma Venkatesh, Kiran Raja, and Christoph Busch. Handwritten Signature and Text based User Verification using Smartwatch. In *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, January 2021. doi: 10.1109/icpr48806.2021.9412048.

[43] Alona Levy, Ben Nassi, Yuval Elovici, and Erez Shmueli. Handwritten Signature Verification Using Wrist-Worn Devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):1–26, September 2018. ISSN 2474-9567. doi: 10.1145/3264929.

[44] Gen Li, Lingfeng Zhang, and Hiroyuki Sato. In-air Signature Authentication Using Smartwatch Motion Sensors. In *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE, July 2021. doi: 10.1109/compsac51774.2021.00061.

[45] Gen Li and Hiroyuki Sato. Sensing In-Air Signature Motions Using Smartwatch:

A High-Precision Approach of Behavioral Authentication. *IEEE Access*, 10:57865–57879, 2022. ISSN 2169-3536. doi: 10.1109/access.2022.3177905.

[46] Yuheng Guo and Hiroyuki Sato. Smartwatch In-Air Signature Time Sequence Three-Dimensional Static Restoration Classification Based on Multiple Convolutional Neural Networks. *Applied Sciences*, 13(6):3958, March 2023. ISSN 2076-3417. doi: 10.3390/app13063958.

[47] Yubo Shao, Tinghan Yang, He Wang, and Jianzhu Ma. AirSign: Smartphone Authentication by Signing in the Air. *Sensors*, 21(1):104, December 2020. ISSN 1424-8220. doi: 10.3390/s21010104.

[48] Attaullah Buriro, Rutger Van Acker, Bruno Crispo, and Athar Mahboob. AirSign: A Gesture-Based Smartwatch User Authentication. In *2018 International Carnahan Conference on Security Technology (ICCST)*, pages 1–5. IEEE, October 2018. doi: 10.1109/ccst.2018.8585571.

[49] Muhammad Taimoor, Huma Butt, Tanveer Khadim, M. Ehatisham-ul Haq, Aasim Raheel, and Aamir Arsalan. REALME: An Approach for Handwritten Signature Verification based on Smart Wrist Sensor. In *2020 IEEE 23rd International Multitopic Conference (INMIC)*, volume 2004, pages 1–6. IEEE, November 2020. doi: 10.1109/inmic50486.2020.9318184.

[50] Elyoenai Guerra-Segura, Aysse Ortega-Pérez, and Carlos M. Travieso. In-air signature verification system using Leap Motion. *Expert Systems with Applications*, 165: 113797, March 2021. ISSN 0957-4174. doi: 10.1016/j.eswa.2020.113797.

[51] Satoshi Kamaishi and Ryuya Uda. Biometric Authentication by Handwriting Using Leap Motion. In *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication*, volume 4677 of *IMCOM '16*, pages 1–5. ACM, January 2016. doi: 10.1145/2857546.2857583.

[52] Santosh Kumar Behera, Debi Prosad Dogra, and Partha Pratim Roy. Analysis of 3D signatures recorded using leap motion sensor. *Multimedia Tools and Applications*, 77(11):14029–14054, July 2017. ISSN 1573-7721. doi: 10.1007/s11042-017-5011-4.

[53] Kohei Nakanishi, Shin-ichi Ito, Momoyo Ito, and Minoru Fukumi. Biometrics authentication of aerial handwritten signature using a convolutional neural network. In *2017 Proceedings of The 5th IIAE International Conference on Industrial Application Engineering*, 2017.

[54] Michał Lech and Andrzej Czyżewski. *Handwritten Signature Verification System Employing Wireless Biometric Pen*, pages 307–319. Springer International Publishing, Cham, 2019. ISBN 978-3-319-77604-0. doi: 10.1007/978-3-319-77604-0_22. URL https://doi.org/10.1007/978-3-319-77604-0_22.

[55] Divas Subedi, Isabella Yung, Digesh Chitrakar, and Kevin Huang. Inertial-Measurement- Based Biometric Authentication of Handwritten Signature. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, volume 11, pages 4320–4324. IEEE, July 2022. doi: 10.1109/embc48229.2022.9871781.

[56] Divas Subedi, Digesh Chitrakar, Isabella Yung, Yicheng Zhu, Yun-Hsuan Su, and Kevin Huang. Biometric Signature Authentication with Low Cost Embedded Stylus. In *2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, volume 1, pages 834–839. IEEE, June 2023. doi: 10.1109/aim46323.2023.10196285.

[57] Mariusz Kurowski, Andrzej Sroczyński, Georgis Bogdanis, and Andrzej Czyżewski. An automated method for biometric handwritten signature authentication employing neural networks. *Electronics (Basel)*, 10(4):456, February 2021.

[58] Yiwen Zhou, Jianbin Zheng, Huacheng Hu, and Yizhen Wang. Handwritten Signature Verification Method Based on Improved Combined Features. *Applied Sciences*, 11(13), 2021. ISSN 2076-3417. doi: 10.3390/app11135867. URL https://www.mdpi.com/2076-3417/11/13/5867.

[59] Muhammad Imran Malik. ICDAR 2009 Signature Verification Competition (SigComp2009). https://tc11.cvc.uab.es/datasets/SigComp2009_1, 2009. [Accessed 06-03-2024].

[60] Dit-Yan Yeung, Hong Chang, Yimin Xiong, Susan George, Ramanujan Kashi, Takashi Matsumoto, and Gerhard Rigoll. SVC2004: First International Signature Verification Competition. In *Biometric Authentication*, volume 5, pages 16–22, 01 2004. ISBN 978-3-540-22146-3. doi: 10.1007/978-3-540-25948-0_3.

[61] J. Ortega-Garcia, Julian Fierrez, Danilo Simon, J. Gonzalez, Marcos Faundez-Zanuy, Vito Espinosa, A. Satue, Inmaculada Hernáez, Juan Igarza, Carlos Vivaracho-Pascual, David Escudero-Mancebo, and Q.Isaac Moro-Sancho. MCYT baseline corpus: a bimodal biometric database. IEE Proc Vis Image Signal Process Spec Issue Biom Internet. *IEEE Proceedings - Vision Image and Signal Processing*, pages 395 – 401, 12 2003. doi: 10.1049/ip-vis:20031078.

[62] A. Kholmatov and B. Yanikoglu. "SUSIG: an on-line signature database, associated protocols and benchmark results". *Pattern Analysis and Applications*, 12(3):227–236, 2009.

[63] J. Fierrez, J. Galbally, J. Ortega-Garcia, M. R. Freire, F. Alonso-Fernandez, D. Ramos, D. T. Toledano, J. Gonzalez-Rodriguez, J. A. Siguenza, J. Garrido-Salas, E. Anguiano, G. Gonzalez-de Rivera, R. Ribalda, M. Faundez-Zanuy, J. A. Ortega, V. Cardeñoso-Payo, A. Viloria, C. E. Vivaracho, Q. I. Moro, J. J. Igarza, J. Sanchez, I. Hernaez, C. Orrite-Uruñuela, F. Martinez-Contreras, and J. J. Gracia-Roche. Biosecurid: a multimodal biometric database. *Pattern Analysis and Applications*, 13(2):235–246, 2 2009. ISSN 1433-755X. doi: 10.1007/s10044-009-0151-4. URL http://dx.doi.org/10.1007/s10044-009-0151-4.

[64] Lingfeng Zhang, Yuheng Guo, Yepeng Ding, and Hiroyuki Sato. 1-D CNN-Based Online Signature Verification with Federated Learning. In *2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pages 2698–2705, 2023. doi: 10.1109/TrustCom60117.2023.00376.

[65] Ashwin Shenoy M, Saritha Suvarna, and Rajgopal K T. Online Digital Cheque Signature Verification using Deep Learning Approach. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, pages 866–871, 2023. doi: 10.1109/ICECAA58104.2023.10212410.

[66] Nan Ji, Bin Liu, Zhiwei Zhao, Yan Lu, Qi Chu, Zhenchao Jin, and Nenghai Yu. Content-Independent Online Handwriting Verification Based on Multi-Modal Fusion. In *2021 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, 2021. doi: 10.1109/ICME51207.2021.9428239.

[67] Jianbin Zheng, Ziyao Chen, Liping Huang, Yifan Gao, Xiangxiang Yu, Hui Wang, Jiamei Yang, and Yu Wang. DWSCNN Online Signature Verification Algorithm Based on CAE-MV Feature Dimensionality Reduction. *IEEE Access*, 12:22144–22157, 2024. doi: 10.1109/ACCESS.2024.3355449.

[68] Mehwish Leghari, Shahzad Memon, Lachhman Das Dhomeja, Akhtar Hussain Jalbani, and Asghar Ali Chandio. Online signature verification using deep learning based aggregated convolutional feature representation. *Journal of Intelligent & Fuzzy Systems*, 43(2):2005–2013, June 2022. ISSN 1875-8967. doi: 10.3233/jifs-219300.

[69] Katarzyna Roszczewska and Ewa Niewiadomska-Szynkiewicz. Online signature biometrics for mobile devices. *Sensors*, 24(11), 2024. ISSN 1424-8220. doi: 10.3390/s24113524. URL https://www.mdpi.com/1424-8220/24/11/3524.

[70] Mei Wang, Ke Zhai, Chi Harold Liu, and Yujie Li. A mobile computing method using CNN and SR for signature authentication with contour damage and light distortion. *Wirel. Commun. Mob. Comput.*, 2018(1):1–10, January 2018.

[71] Abigail Singh and Serestina Viriri. Online Signature Verification using Deep Descriptors. In *2020 Conference on Information Communications Technology and Society (ICTAS)*, pages 1–6, 2020. doi: 10.1109/ICTAS47918.2020.233999.

[72] Chandra Sekhar Vorugunti, Avinash Gautam, and Viswanath Pulabaigari. OSV-ConTramer: A Hybrid CNN and Transformer based Online Signature Verification. In *2023 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10, 2023. doi: 10.1109/IJCB57857.2023.10449120.

[73] Wee How Khoh, Ying Han Pang, and Hui Yen Yap. In-air hand gesture signature recognition: An ihgs database acquisition protocol. *F1000Research*, 11:283, 5 2023. ISSN 2046-1402. doi: 10.12688/f1000research.74134.2.

[74] E K D Kette, D R Sina, and B S Djahi. Digital image processing: Offline handwritten signature identification using local binary pattern and rotational invariance local binary pattern with learning vector quantization. *Journal of Physics: Conference Series*, 2017(1):012011, 9 2021. ISSN 1742-6596. doi: 10.1088/1742-6596/2017/1/012011.

[75] Hedjaz Hezil, Rafik Djemili, and Houcine Bourouba. Signature recognition using binary features and knn. *International Journal of Biometrics*, 10(1):1, 2018. ISSN 1755-831X. doi: 10.1504/ijbm.2018.090121.

[76] W. Sriwathsan, M. Ramanan, and A. R. Weerasinghe. Offline handwritten signature recognition based on sift and surf features using svms. *Asian Research Journal of Mathematics*, pages 84–91, 1 2020. ISSN 2456-477X. doi: 10.9734/arjom/2020/v16i130170.

[77] Nasima Begum, Md Azim Hossain Akash, Sayma Rahman, Jungpil Shin, Md Rashedul Islam, and Md Ezharul Islam. User Authentication Based on Handwriting Analysis of Pen-Tablet Sensor Data Using Optimal Feature Selection Model. *Future Internet*, 13(9), 2021. ISSN 1999-5903. doi: 10.3390/fi13090231. URL `https://www.mdpi.com/1999-5903/13/9/231`.

[78] Md. Azim Hossain Akash, Nasima Begum, Sayma Rahman, Jungpil Shin, Md Amiruzzaman, and Md Rashedul Islam. User authentication through pen tablet data using imputation and flatten function. In *202020 3rd IEEE International Conference on Knowledge Innovation and Invention (ICKII)*, pages 208–211. IEEE, 8 2020. doi: 10.1109/ickii50300.2020.9318975.

[79] Mehwish Leghari, Asghar ali Chandio, Muhammad Ali Soomro, Shah Zaman Nizamani, and Muhammad Hanif Soomro. A comparative analysis of machine learning algorithms for online signature recognition. *VFAST Transactions on Software Engineering*, 12(2):231–240, 6 2024. ISSN 2411-6246. doi: 10.21015/vtse.v12i2.1845.

[80] Shaveta Dargan, Munish Kumar, Anupam Garg, and Kutub Thakur. Writer identification system for pre-segmented offline handwritten devanagari characters using

k-nn and svm. *Soft Computing*, 24(13):10111–10122, 11 2019. ISSN 1433-7479. doi: 10.1007/s00500-019-04525-y.

[81] Long-Fei Mo, Mahpirat, Ya-Li Zhu, Hornisa Mamat, and Kurban Ubul. *Off-Line Handwritten Signature Recognition Based on Discrete Curvelet Transform*, pages 424–434. Springer International Publishing, 2019. ISBN 9783030314569. doi: 10.1007/978-3-030-31456-9_47.

[82] Aliya Rexit, Mahpirat Muhammat, Xuebin Xu, Wenxiong Kang, Alimjan Aysa, and Kurban Ubul. Multilingual handwritten signature recognition based on high-dimensional feature fusion. *Information*, 13(10):496, 10 2022. ISSN 2078-2489. doi: 10.3390/info13100496.

[83] Aliya Rexit, Mahpirat Muhammat, Ruxianguli Abudurexiti, Yusnur Muhtar, and Kurban Ubul. Multi-lingual offline signature recognition based on lomo feature. In *2022 International Conference on Virtual Reality, Human-Computer Interaction and Artificial Intelligence (VRHCIAI)*, pages 201–206. IEEE, 10 2022. doi: 10.1109/vrhciai57205.2022.00041.

[84] Mohammad Z. Al-Shamaileh, Ahmad B. Hassanat, Ahmad S. Tarawneh, M. Sohel Rahman, Ceyhun Celik, and Moohanad Jawthari. New online/offline text-dependent arabic handwriting dataset for writer authentication and identification. In *2019 10th International Conference on Information and Communication Systems (ICICS)*, pages 116–121. IEEE, 6 2019. doi: 10.1109/iacs.2019.8809080.

[85] Tahmid Hasan, Md. Abdur Rahim, Jungpil Shin, Satoshi Nishimura, and Md. Najmul Hossain. Dynamics of digital pen-tablet: Handwriting analysis for person identification using machine and deep learning techniques. *IEEE Access*, 12:8154–8177, 2024. ISSN 2169-3536. doi: 10.1109/access.2024.3352070.

[86] Bahar Çiftçi and Ramazan Tekin. Deep learning based offline handwritten signature recognition. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 13(3):871–884, 9 2024. ISSN 2147-3129. doi: 10.17798/bitlisfen.1527670.

[87] Suresh Pokharel, Santosh Giri, and Subarna Shakya. Deep learning based hand-

written signature recognition. *NCE Journal of Science and Engineering, Volume I, Issue I*, 1(1), 2 2020. ISSN 2717-4794.

[88] Nurullah Çalik, Onur Can Kurban, Ali Rıza Yilmaz, Tülay Yildirim, and Lütfiye Durak Ata. Large-scale offline signature recognition via deep neural networks and feature embedding. *Neurocomputing*, 359:1–14, 9 2019. ISSN 0925-2312. doi: 10.1016/j.neucom.2019.03.027.

[89] Germán Culqui-Culqui, Sandra Sanchez-Gordon, and Myriam Hernández-Álvarez. An algorithm for classifying handwritten signatures using convolutional networks. *IEEE Latin America Transactions*, 20(3):465–473, March 2022. ISSN 1548-0992. doi: 10.1109/tla.2022.9667145.

[90] S. K. Behera, A. K. Dash, D. P. Dogra, and P. P. Roy. Air Signature Recognition Using Deep Convolutional Neural Network-Based Sequential Model. In *2018 24th International Conference on Pattern Recognition (ICPR)*, volume 26, pages 3525–3530. IEEE, August 2018. doi: 10.1109/icpr.2018.8546265.

[91] Alvin Lim Fang Chuen, Khoh Wee How, Pang Ying Han, and Yap Hui Yen. In-air hand gesture signature recognition using multi-scale convolutional neural networks. *JOIV: International Journal on Informatics Visualization*, 7(3–2):2025, 11 2023. ISSN 2549-9610. doi: 10.30630/joiv.7.3-2.2359.

[92] Yoko Seki. A preliminary study on handwriting classification using pen pressure patterns. In *2021 International Conference on Engineering and Emerging Technologies (ICEET)*, pages 1–6, 2021. doi: 10.1109/ICEET53442.2021.9659611.

[93] Wanying Li, Mahpirat Muhammat, Xuebin Xu, Alimjan Aysa, and Kurban Ubul. Sig-rfpnet: multi-branch network and multi-classifier method for offline signature recognition. *Signal, Image and Video Processing*, 18(11):7847–7856, 7 2024. ISSN 1863-1711. doi: 10.1007/s11760-024-03433-5.

[94] Dilara Gumusbas and Tulay Yildirim. Offline signature identification and verification using capsule network. In *2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)*, pages 1–5. IEEE, 7 2019. doi: 10.1109/inista.2019.8778228.

[95] Md. Abdur Rahim, Fahmid Al Farid, Abu Saleh Musa Miah, Arpa Kar Puza, Md. Nur Alam, Md. Najmul Hossain, Sarina Mansor, and Hezerul Abdul Karim. An enhanced hybrid model based on cnn and bilstm for identifying individuals via handwriting analysis. *Computer Modeling in Engineering & Sciences*, 140(2):1689–1710, 2024. ISSN 1526-1506. doi: 10.32604/cmes.2024.048714.

[96] Souvik Ghosh, Spandan Ghosh, Pradeep Kumar, Erik Scheme, and Partha Pratim Roy. A novel spatio-temporal siamese network for 3d signature recognition. *Pattern Recognition Letters*, 144:13–20, 4 2021. ISSN 0167-8655. doi: 10.1016/j.patrec.2021.01.012.

[97] Man Lok Fung, Michael Z. Q. Chen, and Yong Hua Chen. Sensor fusion: A review of methods and applications. In *2017 29th Chinese Control And Decision Conference (CCDC)*. IEEE, May 2017. doi: 10.1109/ccdc.2017.7979175.

[98] Ramon F. Brena, Antonio A. Aguileta, Luis A. Trejo, Erik Molino-Minero-Re, and Oscar Mayora. Choosing the best sensor fusion method: A machine-learning approach. *Sensors*, 20(8):2350, April 2020. ISSN 1424-8220. doi: 10.3390/s20082350.

[99] Asahi Kasei Microdevices/AKM. Ak09918 3-axis electronic compass, November 2016. URL `https://www.akm.com/content/dam/documents/products/electronic-compass/ak09918c/ak09918c-en-datasheet.pdf`. Accessed: 22.10.2022.

[100] How to get raw data from Galaxy S Pen Remote SDK? - Samsung Developers Forums. `https://forum.developer.samsung.com/t/how-to-get-raw-data-from-galaxy-s-pen-remote-sdk/8455`, 2023. Accessed: 2023-09-19.

[101] Android Developers. MotionEvent Class, 9 2023. URL `https://developer.android.com/reference/android/view/MotionEvent`. Accessed: 2023-09-28.

[102] S. Garrido-Jurado, R. Muñoz-Salinas, F.J. Madrid-Cuevas, and M.J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292, 6 2014. ISSN 0031-3203. doi: 10.1016/j.patcog.2014.01.005.

[103] STMicroelectronics. *LSM6DSL iNEMO 6DoF inertial measurement unit (IMU)*, September 2017. URL `https://www.st.com/resource/en/datasheet/lsm6dsl.pdf`.

[104] Li Yang, Xiaoyan Jin, and Qi Jiang. Online handwritten signature verification based on the most stable feature and partition. *Cluster Computing*, 22(S1):1691–1701, February 2018. ISSN 1573-7543. doi: 10.1007/s10586-018-1749-3.

[105] K.S. Manjunatha, S. Manjunath, D.S. Guru, and M.T. Somashekara. Online signature verification based on writer dependent features and classifiers. *Pattern Recognition Letters*, 80:129–136, September 2016. ISSN 0167-8655. doi: 10.1016/j.patrec.2016.06.016.

[106] Gerasimos Pagiatakis, Nikolaos Voudoukis, and Dimitrios Uzunidis. An integrated approach to teaching analog-to-digital conversion and its implementation in telecom networks. In *INTED2023 Proceedings*, volume 1 of *INTED2023*, pages 3778–3782. IATED, March 2023. doi: 10.21125/inted.2023.1009.

[107] Mandlenkosi Victor Gwetu. *A Dynamic Time Warping and Deep Neural Network Ensemble for Online Signature Verification*, pages 141–153. Springer International Publishing, 2021. ISBN 9783030708665. doi: 10.1007/978-3-030-70866-5_9.

[108] Shaohua Jiang and Zheng Chen. Application of dynamic time warping optimization algorithm in speech recognition of machine translation. *Heliyon*, 9(11):e21625, November 2023. ISSN 2405-8440. doi: 10.1016/j.heliyon.2023.e21625.

[109] Manabu Okawa. Template matching using time-series averaging and dtw with dependent warping for online signature verification. *IEEE Access*, 7:81010–81019, 2019. ISSN 2169-3536. doi: 10.1109/access.2019.2923093.

[110] Bhushan V. Patil and Punam R. Patil. An efficient dtw algorithm for online signature verification. In *2018 International Conference On Advances in Communication and Computing Technology (ICACCT)*, volume 2, pages 1–5. IEEE, February 2018. doi: 10.1109/icacct.2018.8529614.

[111] Zhang, Lingfeng and Guo, Yuheng and Ding, Yepeng and Sato, Hiroyuki. 1-D CNN-Based Online Signature Verification with Federated Learning, 2024.

[112] Chirag Nathwani. Online signature verification using bidirectional recurrent neural network. In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, May 2020. doi: 10.1109/iciccs48265.2020.9121023.

[113] Qi Shen, Fangjun Luan, and Shuai Yuan. Multi-scale residual based siamese neural network for writer-independent online signature verification. *Applied Intelligence*, 52 (12):14571–14589, March 2022. ISSN 1573-7497. doi: 10.1007/s10489-022-03318-5.

[114] Mohammad Hajizadeh Saffar, Mohsen Fayyaz, Mohammad Sabokrou, and Mahmood Fathy. Online signature verification using deep representation: A new descriptor, 2018.

[115] Anamika Dhillon and Gyanendra K. Verma. Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, 9(2):85–112, December 2019. ISSN 2192-6360. doi: 10.1007/s13748-019-00203-0.

[116] Zewen Li, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12):6999–7019, December 2022. ISSN 2162-2388. doi: 10.1109/tnnls.2021.3084827.

[117] T. Vasudeva Reddy, D. Harikrishna, V. Hindumathi, P. Asha Rani, and T. Keerthi. *Analysis on Identification and Detection of Forgery in Handwritten Signature Using CNN*, pages 127–143. Springer International Publishing, 2023. ISBN 9783031236020. doi: 10.1007/978-3-031-23602-0_8.

[118] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. ISSN 0018-9219. doi: 10.1109/5.726791.

[119] Mohammad Mustafa Taye. Understanding of machine learning with deep learning: Architectures, workflow, applications and future directions. *Computers*, 12(5):91, April 2023. ISSN 2073-431X. doi: 10.3390/computers12050091.

[120] Yun Liu, Ming-Ming Cheng, Xiaowei Hu, Jia-Wang Bian, Le Zhang, Xiang Bai, and Jinhui Tang. Richer convolutional features for edge detection. *IEEE Transactions*

on *Pattern Analysis and Machine Intelligence*, 41(8):1939–1946, August 2019. ISSN 1939-3539. doi: 10.1109/tpami.2018.2878849.

[121] Ilaria Cacciari and Anedio Ranfagni. Hands-on fundamentals of 1d convolutional neural networks—a tutorial for beginner users. *Applied Sciences*, 14(18):8500, September 2024. ISSN 2076-3417. doi: 10.3390/app14188500.

[122] Nair, Vinod and Hinton, Geoffrey E. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, page 807–814, Madison, WI, USA, 2010. Omnipress. doi: 10.5555/3104322.3104425.

[123] Yann A. LeCun, Léon Bottou, Genevieve B. Orr, and Klaus-Robert Müller. *Efficient BackProp*, pages 9–48. Springer Berlin Heidelberg, 2012. ISBN 9783642352898. doi: 10.1007/978-3-642-35289-8_3.

[124] Afia Zafar, Muhammad Aamir, Nazri Mohd Nawi, Ali Arshad, Saman Riaz, Abdulrahman Alruban, Ashit Kumar Dutta, and Sultan Almotairi. A comparison of pooling methods for convolutional neural networks. *Applied Sciences*, 12(17):8643, August 2022. ISSN 2076-3417. doi: 10.3390/app12178643.

[125] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge, 2014.

[126] Kevin Jarrett, Koray Kavukcuoglu, Marc'Aurelio Ranzato, and Yann LeCun. What is the best multi-stage architecture for object recognition? In *2009 IEEE 12th International Conference on Computer Vision*, pages 2146–2153, 2009. doi: 10.1109/ICCV.2009.5459469.

[127] Ian Goodfellow and Yoshua Bengio and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[128] Sergey Ioffe and Christian Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*, ICML'15, page 448–456. JMLR.org, 2015.

[129] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014. URL `http://jmlr.org/papers/v15/srivastava14a.html`.

[130] Roseline Oluwaseun Ogundokun, Rytis Maskeliunas, Sanjay Misra, and Robertas Damaševičius. Improved cnn based on batch normalization and adam optimizer. In Osvaldo Gervasi, Beniamino Murgante, Sanjay Misra, Ana Maria A. C. Rocha, and Chiara Garau, editors, *Computational Science and Its Applications – ICCSA 2022 Workshops*, pages 593–604, Cham, 2022. Springer International Publishing. ISBN 978-3-031-10548-7.

[131] Kun Yu, Julien Epps, and Fang Chen. Cognitive load evaluation of handwriting using stroke-level features. In *Proceedings of the 16th International Conference on Intelligent User Interfaces*, IUI '11, page 423–426, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450304191. doi: 10.1145/1943403. 1943481. URL `https://doi.org/10.1145/1943403.1943481`.

[132] Lorna Bourke and Anne-Marie Adams. The relationship between working memory and early writing assessed at the word, sentence and text level. *Educational and Child Psychology*, 20(3):19–36, 2003. ISSN 2396-8702. doi: 10.53841/bpsecp.2003. 20.3.19.

[133] An-Hsiang Wang, Hsin-Yu Lin, and Sheng-He Xiao. Effects of text input methods on performance in operating handheld devices between young and elderly users. *Displays*, 65:101974, December 2020. ISSN 0141-9382. doi: 10.1016/j.displa.2020. 101974.

[134] Huawei Tu, Xiangshi Ren, and Shumin Zhai. Differences and similarities between finger and pen stroke gestures on stationary and mobile devices. *ACM Trans. Comput.-Hum. Interact.*, 22(5), 8 2015. ISSN 1073-0516. doi: 10.1145/2797138. URL `https://doi.org/10.1145/2797138`.

[135] Ku Bong Min and Jinwook Seo. Efficient typing on ultrasmall touch screens with

in situ decoder and visual feedback. *IEEE Access*, 9:75187–75201, 2021. ISSN 2169-3536. doi: 10.1109/access.2021.3081173.

[136] Domenico Prattichizzo, Leonardo Meli, and Monica Malvezzi. Digital handwriting with a finger or a stylus: A biomechanical comparison. *IEEE Transactions on Haptics*, 8(4):356–370, October 2015. ISSN 1939-1412. doi: 10.1109/toh.2015.2434812.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

The following abbreviations are used in this dissertation:

2D	Two dimensional space
3D	Three dimensional space
AMOLED	Active-matrix organic light-emitting diode
API	Application programming interface
AUC	Area under the curve
CIR	Channel impulse response
CNN	Convolutional neural network
CWT	Continuous wavelet transform
dB	Decibel
DCT	Discrete cosine transform
DL	Deep learning
DT	Decision tree
DTW	Dynamic time warping
EER	Equal error rate
FAR	False acceptance rate
FFT	Fast Fourier transform
FPS	Frames per second
FRR	False rejection rate
HCI	Human-computer interaction
HDR	High dynamic range
IC	integrated circuit
IMU	Inertial measurement unit

IoT	Internet of things
kHz	Kilohertz
KNN	k-nearest neighbors
LR	Logistic regression
LSB	Least significant bit
LSTM	Long short-term memory
MB	Megabyte
MBS	Multimodal biometric system
MFA	Multi-factor authentication
MLP	Multilayer perceptron
OTP	One-time password
PCM	Pulse code modulation
PPI	Pixels per inch
RAM	Random-access memory
RBF	Radial basis function
ResNet	Residual network
RF	Random forest
RM	Repeat measures
RNN	Recurrent neural network
SDB	Smart development bridge
SDK	Software development kit
SiP	System-in-package
SNR	Signal-to-noise ratio
SVM	Support vector machine
T	Tesla
TAR	True acceptance rate
UBS	Unimodal biometric system
UI	User interface
USB	Universal serial bus
WAVE or WAV	Waveform audio file format
PIN	Personal identification number

APPENDIX

A. User Testing Informed Consent Form

User Testing Informed Consent Form

under Ethics committee approval no. _____

Study administrator: _____

Participant: _____

Participant ID (if applicable): _____

This is a study on **the execution of handwritten signatures and handwritten letters/words/phrases on a touchscreen tablet**. When interacting with a mobile device via finger or stylus, data is retrieved from different sensors, based on the concept of sensor fusion. The collected data can be used to extract biometric features of handwriting and signature. The final goal of the research is to build a system for advanced user authentication, and your participation in the experiment will help construct a data set that will be amenable to the application of machine learning methods.

In this session, you will work in an experimental environment that, in addition to a tablet, includes a smartphone, a smartwatch, and a specially enriched stylus. Information about interaction with the touchscreen will be monitored through several built-in and external sensors. You will be asked to perform tasks typical users might do, such as signing and writing letters/words/short phrases. The study administrator(s) will be in the same room, quietly observing and, if necessary, taking notes. If you have any problems or questions regarding the exam session, you can ask for help from study administrator(s).

All information collected in the test session(s) belongs to **Faculty of Engineering, University of Rijeka**, and will be used for internal purposes. Test session **may be** videotaped, audiotaped, and/or photographed. We may publish our results from this and other sessions in our reports/papers, but all such reports/papers will involve confidentiality and will not include your name. Also, in addition to anonymizing data, it will be ensured that sensitive data is not further shared.

This is a test of the **interaction system**. We are not testing you. You may take breaks as needed and stop your participation in the study at any time.

Statement of Informed Consent

I have read the description of the study and of my rights as a participant. I voluntarily agree to participate in the study.

Print Name: _____

Signature: _____

Date: _____

B. Ethics Committee Approval

Tehnički fakultet u Rijeci

Etičko povjerenstvo

KLASA: 640-08/23-01/3

URBROJ: 2170-1-43-29-23-2

Rijeka, 6. studenoga 2023.

Predmet: Mišljenje Etičkog povjerenstva o istraživanju asistenta Alena Salkanovića za potrebe izrade

znanstvenog rada

Etičko povjerenstvo je putem elektroničke pošte i urudžbenog zapisnika zaprimilo zahtjev asistenta Alena Salkanovića za davanje mišljenja o istraživanju za potrebe izrade znanstvenog rada. Asistent Alen Salkanović je u zahtjevu objasnio ciljeve istraživanja i proces prikupljanja podataka kako slijedi u nastavku:

Ciljevi istraživanja

- Prikupljati podatke ispitnih korisnika tijekom njihovog pisanja na zaslonu tablet računala
- Dobivanje biometrijskih odrednica ispitanika koristeći aparatus koji se sastoji od tablet računala, pametnog sata i mobilnog uređaja sa zaslonom osjetljivim na dodir
- Koristiti senzore tvornički ugrađene u navedene uređaje, zajedno s dodatnim piezo senzorom koji prikuplja dopunske podatke poput vibracija i razine pritiska na ekran
- Istraživati odnose između dobivenih biometrijskih podataka, s ciljem istraživanja te poboljšanja procesa dinamičke verifikacije potpisa
- Izrada skupa podataka pomoću opisanih metoda, koji bi naknadno bio publiciran, uz anonimiziranje podataka ispitnih korisnika

Proces prikupljanja podataka

Podaci će se prikupljati u kontroliranim uvjetima, na unaprijed određenoj lokaciji i naputcima koji će biti detaljno objašnjeni sudionicima eksperimenta. Aparatus će bilježiti podatke koristeći ugrađene i vanjske senzore ranije navedenih uređaja:

- akcelerometar
- kamera
- žiroskop
- dva vanjska piezoelektrična senzora
- zaslon osjetljiv na dodir

Mišljenje Etičkog povjerenstva

Nakon proučenog zahtjeva asistenta Alena Salkanovića, Etičko povjerenstvo smatra da nema razloga zbog kojeg se istraživanje ne bi izvršilo. Sudionici će biti informirani o istraživanju, te će se pribaviti njihov informirani pristanak u pisanom obliku. Etičko je povjerenstvo mišljenja da se navedenim ciljevima i procesima prikupljanja podataka ne krši Etički kodeks Sveučilišta u Rijeci. Dodatno želimo naglasiti kako ovaj upit možda nije ni trebao biti usmjeren prema Etičkom povjerenstvu jer više ima dodirnih točaka s Općom uredbom o zaštiti osobnih podataka (GDPR).
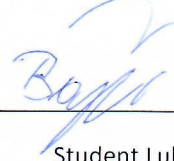
Članovi Etičkog povjerenstva:

Prof. dr. sc. Miroslav Vrankić

Prof. dr. sc. Marko Čanađija

Student Luka Bajić

Dostaviti:

1. Dekan

2. Asistent Alen Salkanović

3. Arhiva Fakulteta

C. Camera calibration procedure

The distortion coefficients form a vector containing 5 or more elements, which describes the distortion caused by a smartphone camera. The two primary types of distortion include radial distortion and tangential distortion. The first type causes straight lines to bend, and the effect becomes more noticeable as points move further from the center of the image. It can be expressed in the following manner:

$$
x_{distorted} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6),
$$
$$
y_{distorted} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6).
$$

(C.1)

In this context, x and y denote the position of a point in the undistorted image. The altered coordinates of a point that was originally positioned at (x, y) in the image are indicated as $x_{distorted}$ and $y_{distorted}$. These coordinates represent the distorted position of a point resulting from radial distortion correction. Furthermore, the coefficients k_1, k_2, and k_3 quantify the degree of radial distortion in the image. Finally, the variable r signifies the radial distance between the image center and the point (x, y).
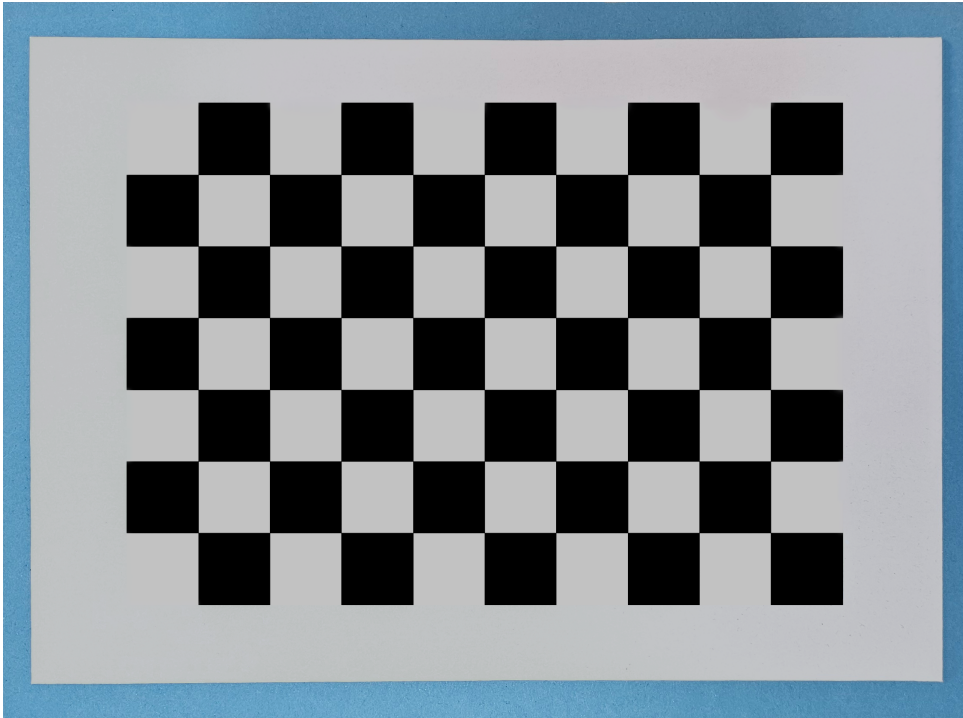

Figure C.1: Example of a checkerboard calibration pattern employed for camera calibration in the context of ArUco marker detection.

Tangential distortion arises due to the camera lens not being precisely oriented parallel to the imaging plane. Consequently, some regions of the image may seem nearer than anticipated. This type of distortion can be quantified as follows:

$$x_{distorted} = x + [2p_1 xy + p_2(r^2 + 2x^2)],$$
$$y_{distorted} = y + [p_1(r^2 + 2y^2) + 2p_2 xy].$$

(C.2)

Here, $x_{distorted}$ and $y_{distorted}$ indicate the distorted coordinates of a point in the image, after applying the distortion to the original coordinates (x, y). Thus, x and y denote the original, undistorted coordinates of the same point. The coefficients p_1 and p_2 represent the effects of tangential distortion. The radial distance r between the point and the center of the image, typically the principal point of the camera lens, is calculated as $r^2 = x^2 + y^2$. To summarize, camera calibration requires the determination of 5 distortion coefficients:

$$Distortion\ coefficients = \begin{pmatrix} k_1 & k_2 & p_1 & p_2 & k_3 \end{pmatrix}$$

(C.3)

Intrinsic parameters include information like focal length and optical centers of the particular sensor. The Python script is used to generate a camera matrix for correcting distortion caused by the lens. It is specific to a particular camera type and model, represented by a 3×3 matrix:

$$Camera\ matrix = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}.$$

(C.4)

The parameters f_x and f_y denote the focal length of a camera, while c_x and c_y represent the optical centers. To determine these values, sample images featuring a clearly defined pattern must be provided. These images are employed to establish correspondences between environmental points and their projections in the camera image from various viewpoints. In this study, these correspondences are derived from the corners of chessboard patterns, as shown in the Fig. C.1. As a result of successful calibration, camera calibration parameters were obtained: the camera matrix and distortion coefficients necessary for tracking ArUco markers. These parameters are crucial for correcting distortions

and ensuring accurate spatial measurements in the captured images. Once the camera matrix and distortion coefficients are known, they can be applied to accurately detect the tilt of the stylus in relation to the camera.

BIOGRAPHY

Alen Salkanović was born on September 2, 1994, in Rijeka, Croatia. He started his undergraduate studies in Computer Science at the Faculty of Engineering, University of Rijeka in 2013 and graduated with a bachelor's degree in 2016. He continued his education at the same faculty, enrolling in graduate studies in Computer Science with a focus on the Computer Systems, and graduated with a Master's degree in 2018.

In 2019, he began working at the Faculty of Engineering, University of Rijeka, as an assistant in the Department of Computer Engineering. That same year, he enrolled in postgraduate doctoral studies in Computer Science, where he focused on both teaching and research activities. He is currently affiliated with the AIRI: Center for Artificial Intelligence and Cybersecurity at the University of Rijeka. His research interests are primarily centered around mobile Human-Computer Interaction (HCI).

LIST OF PUBLICATIONS

[1] A. Salkanovic, I. Štajduhar and S. Ljubic, Floating Hierarchical Menus for Swipe-Based Navigation on Touchscreen Mobile Devices, in Human-Computer Interaction. Multimodal and Natural Interaction, Springer International Publishing, 2020, pp. 509-522, doi: 10.1007/978-3-030-49062-1_34.

[2] A. Salkanovic, M. Stojkovic and S. Ljubic, The Effect of Dwell Time on Swipe-based Pie-Menu Navigation Efficiency, International Journal of Online and Biomedical Engineering (iJOE), vol. 16, no. 15, pp. 4-22, Dec. 2020, doi: 10.3991/ijoe.v16i15.17763.

[3] A. Salkanovic and S. Ljubic, Touchless Interaction on Mobile Devices Using Embedded Ambient Light Sensor, in Distributed, Ambient and Pervasive Interactions, Springer International Publishing, 2021, pp. 153-163, doi: 10.1007/978-3-030-77015-0_11.

[4] S. Ljubic, F. Hržić, A. Salkanovic and I. Štajduhar, Augmenting Around-Device Interaction by Geomagnetic Field Built-in Sensor Utilization, Sensors, vol. 21, no. 9, p. 3087, Apr. 2021, doi: 10.3390/s21093087.

[5] A. Salkanovic, S. Ljubic, L. Stankovic and J. Lerga, Analysis of Cryptography Algorithms Implemented in Android Mobile Application, Information Technology and Control, vol. 50, no. 4, pp. 786-807, Dec. 2021, doi: 10.5755/j01.itc.50.4.29464.

[6] D. Selimovic, A. Salkanovic and M. Tomic, Application of MQTT Based Message Brokers for IoT Devices Within Smart City Solutions, in 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), IEEE, May 2022, pp. 428-433, doi: 10.23919/mipro55190.2022.9803388.

[7] A. Salkanovic and S. Ljubic, LuxBoard: Ambient Light Manipulation for Contactless Text Entry on Mobile Devices, in Human-Computer Interaction. Technological Innovation, Springer International Publishing, 2022, pp. 126-139, doi: 10.1007/978-3-031-05409-9_10.

[8] S. Ljubic and A. Salkanovic, Generating Representative Phrase Sets for Text Entry Experiments by GA-Based Text Corpora Sampling, Mathematics, vol. 11, no. 11, p. 2550, Jun. 2023, doi: 10.3390/math11112550.

[9] V. Ecimovic, L. Skitarelic, A. Salkanovic and S. Ljubic, A Comparison of Finger, Stylus and Mouse in Trajectory Tracing Task Execution on Touchscreen Mobile Devices, in Human-Computer Interaction, Springer Nature Switzerland, 2023, pp. 433-449, doi: 10.1007/978-3-031-35596-7_28.

[10] S. Ljubic, A. Skoki, F. Hržić and A. Salkanovic, Automated Online Assessment and Cloud-Based Programming: Advancing Computer Engineering Education, in Higher Education Learning Methodologies and Technologies Online, Springer Nature Switzerland, 2024, pp. 78-94, doi: 10.1007/978-3-031-67351-1_6.

[11] A. Salkanovic, D. Bačnar, D. Sušanj and S. Ljubic, A Sensor-Fusion-Based Experimental Apparatus for Collecting Touchscreen Handwriting Biometric Features, Applied Sciences, vol. 14, no. 23, p. 11234, Dec. 2024, doi: 10.3390/app142311234.

[12] A. Salkanovic, D. Sušanj, L. Batistić and S. Ljubic, Beyond Signatures: Leveraging Sensor Fusion for Contextual Handwriting Recognition, Preprints, Mar. 2025, doi: 10.20944/preprints202503.1018.v1.