Deep Learning–Based Behavioral Biometric Authentication Using Keystroke and Touch Patterns

by

Faisal Ahmed
21301186
MD. Shafiur Rahman
21201613
Mahmuda Akter Alif
21201149
Nazia Mumtahina
21201303
Murtasin Islam Mormo
21201466

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

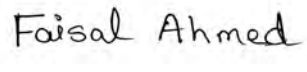
Department of Computer Science and Engineering
BRAC University
January 2026.

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

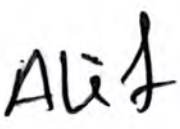
Student's Full Name & Signature:



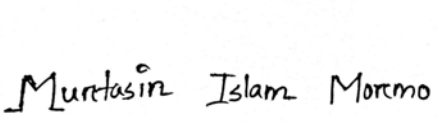
Faisal Ahmed
21301186

MD. Shafiur Rahman
21201613

Nazia Mumtahina
21201303



Mahmuda Akter Alif
21201149



Murtasin Islam Mormo
21201466

Approval

The thesis titled "Deep Learning–Based Behavioral Biometric Authentication Using Keystroke and Touch Patterns" submitted by

1. Faisal Ahmed (21301186)
2. MD. Shafiur Rahman (21201613)
3. Mahmuda Akter Alif (21201149)
4. Nazia Mumtahina (21201303)
5. Murtasin Islam Mormo (21201466)

of Fall, 2025 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science in Fall 2025.

Examining Committee:

Supervisor:
(Member)

Dr. Muhammad Iqbal Hossain

Associate Professor
Department of Computer Science and Engineering
Brac University

Thesis Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Dr. Sadia Hamid Kazi

Associate Professor & Chairperson
Department of Computer Science and Engineering
Brac University

Abstract

With the development of digitization, sensitive data protection and online transactions have become crucial. Traditional authentication methods, such as using a PIN and password on a regular basis, have to deal with credential theft, phishing, and brute force attacks. Unlike static authentication systems like knowledge-based authentication or physical biometric authentication, behavioral biometric authentication focuses on one's unique and habituated interactions with one's own device. This approach does not use anything that a user knows or has as a physical presence. Rather, it uses something that a user may as well be unaware of but follows on a daily basis, which is keystroke pattern and touch pattern. The technique fits well in the task of enhancing user verification without sacrificing usability. This research extends the domain of access control through the design of a flexible and scalable authentication system. For the assessment of the effectiveness of this technique, a cross-device web platform was built. By which keystroke and touch input of 30 users in four different sessions have been collected. Machine learning analysis of keystroke and touch behavior confirms the effectiveness of behavioral biometrics for identity verification. The initial findings from the model, like those of Siamese LSTM, demonstrate high accuracy (93% training accuracy and 82% test accuracy) are encouraging for keystroke. As a result, our solution makes sense as a substitute for conventional authentication methods.

Keywords: Traditional authentication, Credential theft, Phishing, Brute force, Multifactor authentication, Behavioral biometrics, Keystroke dynamics, Touch gestures, Continuous authentication, machine learning, Accuracy, Efficiency.

Acknowledgement

First and foremost, we thank and praise Allah for successful completion of our thesis without any major interruptions. Next, we express our sincere gratitude to our thesis supervisor, Professor Muhammad Iqbal Hossain, for his wholehearted support and guidance throughout the course of this work; he always showed up whenever we called for any help. Last but not least, to our parents, who are always a source of encouragement and support for this endeavor. Their encouragement and prayers have brought us this far.

˜*The Authors*

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

AUC Area Under the Curve

BiLSTM Bidirectional Long Short-Term Memory

CNN Convolutional Neural Network

DTW Dynamic Time Warping

ERR Equal Error Rate

FAR False Acceptance Rate

FN False Negative

FP False Positive

FRR False Rejection Rate

GRU Gated Recurrent Unit

KBA Knowledge-Based Authentication

KD Keystroke Dynamics

KVC Keystroke Verification Challenge

LOF Local Outlier Factor

LSTM Long Short-Term Memory

MCCV Monte Carlo Cross Validation

MFA Multi-Factor Authentication

NN Neural Network

PBRE Perturbation-Based Robustness Evaluation

PIN Personal Identification Number

RF Random Forest

ROC Receiver-Operating Characteristic Curve

SVM Support Vector Machine

TDD Temporal Dynamics Descriptors

TN True Negative

TP True Positive

Chapter 1

Introduction

1.1 Background

From mere information, personal data in modern times has turned out to be one of the most valued assets. Hence, its protection is just as important as protecting any assets. While the rapid emergence of digital platforms and services has brought convenience to life unimaginable in the past, it has also opened new avenues for seasoned cybercriminals to exploit the vulnerabilities of more traditional authentication methods using PINs and passwords. Whereas identity theft and data breaches have almost become commonplace, these static, easily compromised methods no longer suffice. With layered security at its core in confirming user identity, multifactor authentication has increasingly become a great solution to the problems. Among its numerous advances, behavioral biometrics is a groundbreaking method that uses a person's distinct interaction patterns—such as typing patterns, mouse movements, and use habits—to offer a safe and easy identification procedure. Behavioral biometrics are incredibly resistant to attacks because they are constantly evolving and adapting, unlike physical biometrics or static credentials. This study explores how behavioral biometrics may revolutionize authentication standards at a time when achieving a balance between user ease, security, and privacy is crucial. Understanding the urgent need for these developments serves as the starting point for the study and serves as a framework for the discussion that follows.

1.2 Motivation

The reliability and usability of authentication has been one of the most important issues because the phones have turned into a prime portal to personal and institutional secrets. Due to the ever demands of simultaneous authentication in the popular smartphones, the most common kind of authentication to be used is the traditional methods of authentication such as passwords and PIN which are memorized credentials. Nevertheless, there are certain limitations inherent in them, particularly in the environment with a high frequency of mobile devices use. The implementation of such a mechanism in mobile devices puts the burden primarily on users as they need to remember the credentials and, therefore, frequently, they do not comply with the rules, such as reusing or sharing them which makes the risk even greater [14]. These facts highlight the importance of authentication approaches

which on the one hand are simpler and user friendly yet on the other hand increase security.

The implementation of biometric authentication has been initiated as a measure to some of the mentioned issues; however, the most frequent physical biometrics like fingerprints and facial recognition present significant limitations in mobile contexts. Physical biometric platforms are usually dependent on direct user interaction, for example, the fingerprint sensor. They also might experience poor performance due to the environment, the quality of the sensor, or the physiological change of the user. Besides, physical biometric features are static and can be revoked only if the user deletes the related data, which can lead to long-term privacy issues if such data are leaked [24]. These drawbacks have compelled a growing interest not in biometric analysis but in behavioral biometric authentication as an alternative which is better and less visible.

Behavioral biometrics utilize the regularities that are reflected in user interaction patterns in the course of their natural use of a device and hence they are implemented in smartphones with a passive and continuous authentication approach. In contrast to physical characteristics, behavioral traits such as typing rhythm and touchscreen interaction patterns change over time, which allows the entity to adapt even to inter-user variability. The fact that these behaviors can be accumulated during regular interactions, stress the point of view that the behavioral biometric mechanisms are user-friendly and secure authentication processes can be achieved [23].

In this field, keystroke dynamics and touchscreen interaction patterns are at the forefront as the two most informative behavioral modalities common in smartphones. Keystroke dynamics observe the temporal aspects of typing behavior like typing speed, rhythm, and correction patterns, while no additional hardware is needed, making it cost-effective and maneuverable for installation on mobile devices [6]. Tapping, swiping and gesture based manipulation are examples of touchscreen interaction traits that take advantage of the rich interaction capabilities of contemporary cellphones as well as impacted by posture, motor control and usage context.

Though they come with their advantages, the existing behavioral biometric studies frequently assess physical keystroke dynamics and touchscreen interactions separately, or under limited laboratory tests that do not mimic real smartphone usage. Daily mobile interaction is characterized by such diverse activities as text entry, gaming, and visual content manipulation, all of which create different behavioral patterns. This leads to a situation where a system trained to operate on a limited number of interaction types could possibly fail to generalize across different contexts in the real world [23].

On account of lacking in these areas, this study presents a deep learning-based behavioral biometric authentication framework which apart from integrating keystroke also touch interaction patterns collected from smartphones includes realistic user activities such as playing games, viewing pictures using zooming and scaling functions.

The main aim of this research is to provide a clearer insight into the strengths and limitations of the diverse learning paradigms for continuous authentication on smartphones by evaluating deep learning models and classical machine learning classifiers on the same multimodal dataset. In addition to giving an unobtrusive and accurate

authentication system, the behavior variability modeling of the envisioned solution will also be its added high-performance stability.

1.3 Research Problem

Traditional authentication methods, such as passwords and PINs, are not as efficient as they once were, particularly as cyber security develops and becomes more complex. These traditional techniques are increasingly being challenged by more modern security challenges like phishing, brute force and credential theft. Nevertheless, the establishment of MFA is not quite simple [24].

Although there are advantages of MFA, it faces a great deal of issues particularly in practice. MFA may be compromised by advanced attacks that may breach several layers, consider it is credential theft through social engineering, eavesdropping, or other intelligent tricks. The other obstacle is the inconvenience of the very procedure: the time wastage and constant re-authentications can be a burden to a user. As an example, pattern recognition may be a challenge when there is noise, or the difference between the data samples is vast [11]

We would like to make MFA more functional, and behavioral biometrics, such as keystroke dynamics, mouse movement, touchscreen gesture, etc., are continuous and smooth. They bring in a new dimension of protection that becomes flexible and accommodating to the individual characteristics of a person through the analysis of behavior in real time. Nevertheless, accuracy has limits, spoofing is still possible, and ethical privacy issues.

MFA is having a hard time achieving its primary objective of ensuring that critical systems are not accessed by unauthorized personnel because of a multitude of reasons:

- Authentication of Fatigue: As users, we are sick of interruptions and therefore MFA requires non-reactive, continuous, behavioral checks, which will reduce nuisance.
- Modifying Threats: Attackers continue to refine their approach, specifically token compromise, biometric spoofing, and so on, and their practices are changing faster than the present MFA systems can handle.
- Integration and Scalability: Implementing MFA into different platforms (IoT, cloud, etc.) will become a problem due to the delays in infrastructure and real-time processing requirements. This may be made better with the aid of behavioral biometrics.
- Privacy: The constant gathering of behavioral information brings up legal concerns and predisposes some users to being constantly observed.
- Environmental Factors: External conditions are important, poor lighting will distort face recognition and voice-based security will be compromised under background noises [23].
- Usability: The security systems should be easy to use, and many are inflexible.

- Sophisticated Attacks: Hackers are capable of stealing session cookies of websites without cracking the credentials.

To address the issues mentioned above we should enhance the usability and security of MFA. Enhancing machine-learning models like CNN, RNN and LSTM with behavioral biometrics may increase system security. It was found that models with MLP and LSTM-RNN could handle the lowest equal error rate (EER) of 50.0% [8]. The suggested solution has the potential to reduce intrusive interventions and, hence, increase the user experience. The increasing complexity of cyber threats continues to strain the traditional authentication systems, particularly in relation to safeguarding contentious and confidential information, a factor that continues to increase.

Continued research in the field of advanced authentication methods is necessary because there is an urgent need for efficient multifactor authentication solutions that balance security, usability, and privacy. Currently, researchers and early technology adopters are attempting to integrate new sensors to be utilized in MFA systems [4]. Combining behavioral biometrics with continuous authentication methods is one of the most promising techniques to protect digital assets in a continuously changing threat landscape [23].

1.4 Research Objective

This paper has two main objectives, which include the design and test of a deep learning-based behavioral biometric authentication framework that rides on keystroke dynamics and touchscreen interaction patterns sourced from smartphones to effect continuous and secure user authentication in a naturalistic device usage setting.

The specific objectives are as follows:

1. Assess the efficacy of multivariate fusion keystroke and touchscreen interaction patterns being used as complementary behavioral biometric modalities for smartphone authentication based on recent findings suggesting that the typing and swipe fusion behaviors can enhance authentication reliability compared to single-modal methods [28].

2. Investigate the potential of continuous smartphone authentication under real-world user conditions by collecting behavioral data from daily tasks that incorporate soft-keyboard typing and touch-based interactions instead of controlled or one-time authentication scenarios [26].

3. Explore the expression of keystroke dynamics captured by smartphone soft keyboards in terms of temporal typing behavior formed during natural text entry activities, and their determination of specific discriminators.

4. Across diverse mobile interaction tasks, tapping, swiping, and gesture-based manipulation pattern interaction with the user across diverse mobile interaction tasks, evaluate the contribution of touchscreen interaction patterns to user authentication performance.

5. Explore whether learning-based authentication approaches can take into account behavioral variability arising from differences such as user interaction styles, task types, and session-to-session changes.

6. Evaluate the authentication performance of heterogeneous smartphone activities, e.g., text typing, gaming, and touch-based image manipulation, to find the impact of task diversity on the reliability of behavioral biometrics.

7. By underlining the non-intrusive features, adaptability to natural behavior of the user, and potential for deployment within the wild without creating extra hardware requirements, behavioral biometric security is highlighted.

By achieving these objectives, the study primarily aspires to contribute to the research of behavioral biometric by centering on integrated keystroke and touch-based behavioral patterns and their deployment in real-world smartphone systems.

1.5 Methodology in Brief

Recently, given all these computer threats, old authentication systems such as PINs and fingerprint are quite a puny item. This study attempts to enhance security by introducing behavioral biometric authentication parameters into a more comprehensive multi-factor machine-learned-based system.

The first step is to obtain required data that consists of such important bits as touch pressure, inter-key latency, and typing speed on the mobile and desktop. Once that information has been preprocessed we proceed to feature engineering, which involves principal component analysis and recursive feature elimination to identify the prominent patterns. Training and validating the ML models, which are the Random Forest and the Neural Networks, is the next step in which a set-splitting method is utilized. The backend is designed using Django and TensorFlow Lite, and the system can match patterns comfortably, whereas the frontend- HTML, CSS, JavaScript, or even Flutter maintains the best UX. We check the safety of all with OAuth2 APIs, session tokens, and SHA 256 hashing. AWS, Google cloud, Azure and Docker containers can also be scaled in the system.

Lastly, learning under the supervision and without supervision enable the model to adapt to change of user-behaviors, and greatly enhance the accuracy of detection of legitimate users and reduction of cyber risks.

1.6 Scopes and Challenges

1.6.1 Scopes

The research paper attended to the design and evaluation of a behavioral biometric authentication framework hinged on keystroke dynamics and touchscreen interaction patterns collected from smartphones. This research was applied to the following scopes:

- **Behavioral Features Collected Through Cellphone Engagements:** This research analyzes behavioral characteristics that are collected through touch interaction, for example: taps, gestures, and swiped based action which also includes soft keyboard typing that did not require any extra hardware. These features are then used to simulate based on a specific user for authentication.
- **Learning-Based Methodologies:** The main goal of this study is obtaining user dependent engagement traits and time dependences from the collected data which was gathered through soft keyboard and touch as well as learning based behavioral analysis which explores the modelling of consecutive and engagement level parameters.
- **Continuous and Explicit Identification:** Rather than awaiting for one user's authorization, the suggested architecture should utilize touchscreen interfaces and soft keyboard tapping to verify the individuals in actual time as well as in the background during interaction with the device
- **Realistic Data Acquisition:** With the purpose of simulating realistic cellphone usage patterns, behavioral traits which are collected through a privately developed web based program throughout actual user activities, for example: typing texts, playing games based on interaction, and touch based manipulation of images.
- **Comparative Evaluation of Learning Paradigms:** The investigation is going to use the same multimodal behavioral dataset to evaluate and compare the different learning paradigms' performance, reliability, and applicability in the authorization of smartphone devices.

1.6.2 Challenges

The research study was mainly challenged by the pragmatic and methodological issues, resulting in problems with data collection, model training, and overall system performance:

- **Limited Dataset Size and Model Overfitting:** The small size of the group of data collectors and the 30 users that provided the behavior data are the main issue. According to the sequence learning model rules, which are based on the temporal sequence of linked inputs, this issue results in the effective number of training samples being less than it should. The model's overfitting condition was showing that while the model learns the data patterns in the training data, it fails to generalize the patterns to unseen data.

- **Unusual User Interaction While Collecting Data:** Users had no previous experience with the data collection interface, which was a web based application. Although the users had been instructed to complete the tasks as accurately as possible, it was evident that some users had very relaxed and inconsistent interactions with the application.Their relaxed approach to the application introduced noise to the data set, making it difficult for the models to identify stable and meaningful differentiating patterns in user behavior.

- **Technical Issues With Web Based Data Collecting Techniques :** Behavioral data collection applications that are functional and easy to use comparatively are generally difficult to create. To make sure the application is user friendly as well as the data is being captured accurately , a lot of time, effort, and use of different environments is placed. This includes with regard to the capture of keyboard inputs, specifically the special keys such as shift key.

- **Complex and Difficult Data Collecting Procedure:** For most users lacking technical skills, the process of data collection which demanded user participation and the completion of multiple activities presented a challenge.It was difficult to explain this to users not in the field of computer science and to maintain their interest during the lengthy data collection sessions.Thus, in part, due to the aforementioned reasons, several users did not fully cooperate in the data collection process, and this resulted in the volume of data collected being lower than anticipated.

These identify the real world difficulties related to collecting high quality behavioral biometric traits and training learning based authentication models in ordinary circumstances. Scaling, strengthening, and real world system implementation all rely upon resolving the situation.

1.7 Thesis structure

The behavioral biometrics provide a safe and versatile system of authentication as they are based on the analysis of the unique interaction patterns overcoming the constraints of the traditional system of PIN and password that can be stolen and spoofed. Keystroke dynamics and touchscreen interactions offer non-invasive and continuous authentication with high rates of acceptance and defeat the physical restrictions and usability barriers of traditional physical biometric systems. Multiple factors have presented a major challenge to multifactor authentication systems, which include: authentication fatigue, emerging threats, complexity of integration, privacy, and environmental threats, and thus machine learning advanced techniques should have been integrated with behavioral biometrics. The research problems are aimed at developing and testing a new multifactor authentication system based on combining the keystroke dynamics and touchscreen interaction patterns using deep learning models and fulfilling the requirements of security, usability, privacy, and scalability. The approach is to collect data on behavioral biometric, pre-process on feature engineering by training models with Random Forest, XgBoost and Neural Networks. The project scope is defined to include feature extraction of natural smartphone interaction, contextual authentication, cross-platform integration, and real-time continuous verification and admits issues in within-user and between-user similarity, spoofing attacks, dataset, and computational resource constraints.

Chapter 2

Literature Review

2.1 Preliminaries

Behavioral biometric authentication has been extensively evaluated as a substitute for conventional authentication techniques like passwords and PINs, which are prone to observation and replication attacks. Keystroke dynamics and touchscreen interaction patterns are ideal for smartphone-based authentication, thanks to the ease of their unobtrusive data collection and continuous monitoring capabilities. Earlier research has primarily dealt with the time-based features of keystroke dynamics and different machine learning or deep learning models utilized to identify users. However, in more recent advances, touch-based behavioral features have been seen as an extra layer of security in mobile environments, yet these modalities were often analyzed in isolation and in laboratory conditions. This literature review is an examination of available information on keystroke and touch based behavioral biometric authentication, emphasizing the methods used and the results reported in many studies. Primarily, the paper deals with the contrasting traditional machine learning classifiers to deep learning models. Concerns in the integration of different behavioral modalities and the consideration of diverse learning paths to use the authentic smartphone interaction data are the two issues that still stand even if the reported progress in authentication accuracy is there. Resolving the mentioned problems originated and pushed forward by the very research, which inquires into the deep learning approach on the basis of behavioral biometric authentication using keystroke and touchscreen patterns.

2.2 Review of Existing Research

Contemporary research on behavioral biometric authentication has continuously reported that keystroke dynamics are an unobtrusive and effective mode of user identification, particularly in smartphones and keyboards. Classical machine learning methods like the one implemented by Iyapparaja et al. (2024) [17], utilize features manually crafted from the timing of reciprocation in combination with classifiers such as Support Vector Machines, Random Forests, and Gradient Boosting. Their research findings establish that it is possible to determine distinct users' identification by means of keystroke timing information. Nevertheless, similar to other classical machine learning systems, these techniques also fall victim to a variation in behavior between users between sessions, which is undesirable as set features and decisions cannot readily accommodate the natural and normal variability of the typing behavior.

Based on the pioneer works on keystroke authentication, the current research refined the temporal features extraction process and, at the same time, resolved the issue of algorithmic robustness Risto and Graven (2024) [25]. discuss the problem of fixed text password authentication using K-Nearest Neighbor classification and Manhattan Distance classification where they achieve the True Accept Rates of 99% with False Accept Rates lower than 5%. The way they include the variables of length of passwords, complexity of characters, and frequency of repetition in their systemic analysis provide an excellent approach to the subject, and they expressly state that the literature had never considered those variables as dependent variables. Risto and Graven (2024) [25] validate that keystroke patterns are not hardware-bound security mechanisms used on both desktop and mobile environments; nevertheless, they mention limitations such as small sample sizes and time-based differences caused by user fatigue and progressive password knowledge in authenticating with a password repeatedly throughout the trial period.

Countering these shortcomings, studies issued from 2023 to 2024 have meant an upsurge in the area of learning approaches, lately, it is referred to as the deep learning paradigm. AbdelRaouf et al. (2024) [13] and Medvedev et al. (2023) [10] both state that deep learning architectures show better outcomes in sequential typing behavior than traditional machine learning acquisition. AbdelRaouf et al. (2024) [13] boost performance standards via transfer learning and GAN-based data augmentation, and Medvedev et al. (2023) [10] make use of a Siamese neural network to find out the user's similarities with respect to their keystroke patterns. What is common in these pieces of research is the fact that they show an apparent trend; deep learning models are the ones that no longer depend on manual features and thus are the ones that improve authentication accuracy. This performance improvement is, however, at the cost of certain tradeoffs such as more complex computing, reliance on the curated datasets, and unrealistic conditions of mobile interaction that have been scarcely assessed.

The simultaneous advances in transformer architectures emerge as alternatives to the conventional handcrafted feature engineering in keystroke-based verification systems.Simão et al. (2024) [29] suggest a novel model called TempCharBERT that combines dwell time and flight time with CharBERT's character-level embeddings to reach 90.14% identification accuracy under federated learning protocols that main-

tain privacy through decentralized training. This approach is a departure from explicit feature engineering; it removes the engineer from the project and rather relies on learned representations to encode behavioral signatures. Despite high performance, Simão et al. (2024) [29] point to the implementation difficulties arising from federated learning trade-offs, data quality dependencies, and generalization across diverse user populations which all curtail the deployment in uncontrolled mobile and desktop environments.

Lately, the spatio-temporal attention-based learning mechanism has been adopted in order to enhance behavioral biometric modeling considerably.Nguyen et al. (2023) [22] offer a spatio-temporal dual-attention transformer specially made for time-series behavioral biometrics, thus the model can realize both temporal dependencies and feature-level correlations at the same time. Their findings indicate that authentication is better than recurrent and convolutional models, thus showing the importance of advanced temporal representation learning in behavioral biometric authentication.

Moreover, the architectural design affects the results of keystroke-based authentication since various studies are conducted to check how architectural decisions swing the performance of deep learning.Kek et al. (2024) [18] analyze systematically different neural network models, such as CNN, GRU, and LSTM architectures, and show that authentication accuracy correlates highly with the network design and hyperparameter configuration. Whereas their outcomes validate deep learning models as powerful, there is still a lack of consistency across different architectures, thus bringing that question to the fore in the literature relating to the choice of model complexity against the feasibility of real-world deployment. Reminiscent of earlier deep learning investigations, the analysis remains confined to keyboard keystroke input, hence reiterating the idea that keystroke-only systems may not completely represent user behavior under typical conditions.

Combining the power of deep learning and hybrid architectures, further advantages have been showcased in temporal literacy for keystroke based authentication. Arsh et al. (2024) [16] investigate a CNN-GRU fusion architecture achieving 99.9% accuracy for static verification and 99.31% for continuous authentication using convolutional layers for spatial feature extraction and recurrent components for temporal sequence modeling of typing rhythms and inter-key latencies. Thus the hybrid structure is fully capable of escalating the pressure with temporal complexity on keystroke dynamics to help build robust multifactor systems. Nonetheless, Arsh et al. (2024) [16] note that there is a large intra-user variability due to fatigue, emotional state, and hardware heterogeneity which makes the approach unreliable in real-world scenarios despite the outstanding performance demonstrated under controlled conditions.

With the deep learning breakthroughs, alternative enhancement strategies have also been put forth in the keystroke-based authentication space. Shanmugavalli et al. (2023) [12] propose a hybrid machine learning structure that associates various soft biometric features with the user keystroke dynamics. Their findings indicate that incorporating behavioral features with additional user attributes can lead to higher recognition accuracy. However, hybrid systems that are compared with deep learning methods remain limited to selective feature engineering and do not use any features other than keystroke input such as. The juxtaposition demonstrates the approach

to divergence in the field, where progress is being made either with additional richer features or more expressive learning models, but seldom with both at the same time.

Recent developments of the scholarship field have taken tractable behavioral biometrics scanning to an additional level beyond singular keystroke features by applying an ovo-multimodal keystroke framework. Pryor et al. (2022) [8] used a systematic assessment of the tool to incorporate touch dynamics and phone movement measurements with the analysis of seven different classifiers, such as Random Forest, SVM, K-Nearest Neighbor, Naive Bayes, Logistic Regression, Multilayer Perceptron, and LSTM RNN, on the BioIdent and H MOG datasets to perform user authentication. The experiment they conducted showed insignificant differences in their results, as the mean Equal Error Rates of the two models MLP and LSTM-RNN were 50.0% in the groups of users, indicating a high level of difficulty in recognizing patterns with the help of such developed neural models. Their study researched the possibilities of single and multi modal behavioral biometrics that expose final trade-offs in regards to model complexity and authentication accuracy. Although the level of success is not so high, [8]gave information on the choice of neural network architecture in the context of a behavioral authentication system, detailing the fact that the application of standard deep learning solutions without the appropriate domain-specific architecture forms constitutes a high level of constraints. The key limitations would consist of a small size of the dataset as it impacts generalizability and attitudinal variability between sessions of training samples, and therefore, the need to establish a balance between false acceptance and false rejection within the form of multi-factor authentication.

Introducing behavioral biometric beyond the keystroke dynamics, questions have been put towards the same, inquiring of the tactile-based and multi-model smartphone interaction. The features Tang et al. [30] demonstrate are that a more effective authentication system can be achieved when gesture touchscreen features are combined with data obtained through inertial sensors. Also, the argument that deep learning models are more effective than classical classifiers in continuous swipe authentication has been reinforced by Sejjari et al. (2024) [28]. However, the two studies do not combine any type of interaction: in one study, the gesture - sensor fusion, the other, the swipe gestures, and in neither of the studies are the learning paradigms combined within the framework of learning.

Since the last years, the research related to behavioral authentication has extended to touchscreen modalities and is devoting attention to spatial, temporal and physiological interaction characteristics. In Wu et al. (2024) [31] they proposed BIOHOLD, a system based on hand-holding postures and Temporal Dynamics Descriptors and Spectral Complexity Descriptors to reach an Equal Error Rate of only 1.25% with 40 training samples. This combination of physiological biometrics and behavioural patterns demonstrates that it is possible to achieve high security performance without being super data-efficient or fighting off zero-effort attacks, smudge attacks and visual eavesdropping. The disadvantage is that the system remains dependent on certain hardware and rather narrow dataset, which makes the generalization to a variety of mobile devices hard and restricts the real-world usage.

The trend towards naturalistic interaction in touch based authentication has seen researchers gather data in real life situations, such as gaming, rather than laboratory

controllable conditions. Nasciment et al.(2024) [21] extracted Angular Velocity, Path Tangent, and Acceleration angle measurements of real Minecraft on Samsung tablet sessions, which reflect behavioral signatures based on gestures. Their Support Vector Classifier achieved a 90 percent accuracy, and 0.97 AUC. The best bit here is that the information is obtained in a real play and not in a staged activity. Nevertheless, they also note that their sample heterogeneity is small, they have used single device and their models are prone to overfitting which restricts their applicability to other mobile verification scenarios.

Capacitive sensing is now an attractive approach to enhance touchscreen-based authentication. Lertvittayakumjorn et al.(2024) [19] investigated raw capacitive sensor images of keyboards and reported that detailed spatial information on the area of finger contact, finger shape and pressure distributions are much more effective than the traditional coordinate-based representations. They demonstrated by experiment that the addition of capacitive touch images significantly enhances keyboard decoding, particularly in case of the ambiguous keypresses that border the labels, and in hinting at the possibility of user-specific capacitive signatures as an authentication consideration.

Massive researches still support the advantages of multimodal behavioral validation and indicate existing gaps. A thorough study organized by Stragapede et al. (2022) [32] on touchscreen interactions and background sensor data of 600 users revealed that the fusion of multiple modalities of the scores could lead to a significant error rate in authentication. Nevertheless, the research was constrained by set tasks and did not investigate the performance of various deep learning and traditional classifiers in different and task-oriented smartphone activities such as typing, gaming, or zoom-sourcing touches in actual life.

Behavioral biometrics are proving to have wider potential as keystroke analytics is now extending into mental health. Lim et al(2024),[20] investigated patterns of typing on smartphones as indicators of social isolation and loneliness, and discovered that keyboards behavior evolves over time, which is associated with any psychological state. Their machine learning models were able to effectively identify the users based on their levels of loneliness based on typing rhythms, dwell times and interkeys time with success. They used to admit the ethical issues of continuous attention to behavior and the collection of longitudinal data, which are serious obstacles to health-oriented biometric applications.

2.3 Summary of Key Findings

The literature reviewed indicates a shift from traditional machine learning classifiers towards the application of deep learning models in the area of behavioral biometric authentication. Current studies conducted using Long Short-Term Memory, Convolutional Neural Networks, and a combination of CNN and GRU report improved performance with accuracy rates higher than 99% and Equal Error Rates lower than 3%. The combination of multimodal fusion with keystroke dynamics and touchscreens has shown improved robustness over the use of a single modality, with higher discrimination power and improved resilience against spoofing attacks. Despite the promising results obtained from the application of the above-mentioned approaches, there are a number of critical issues that need to be addressed. The studies conducted have a small number of participants, ranging from 15 to 100 users. Contextual variability from user fatigue, emotional states, and environmental factors significantly impacts behavioral patterns, yet adaptation mechanisms remain underexplored. Additionally, evaluation methodology inconsistencies, computational complexity concerns for mobile deployment, and insufficient privacy considerations represent substantial gaps.

These limitations motivate the present research, which addresses dataset diversity through multi-session collection from 30 users, employs Siamese LSTM architectures for similarity-based verification, and integrates keystroke dynamics with touchscreen gestures to develop scalable, accurate behavioral biometric authentication systems suitable for real-world mobile environments.

Chapter 3

Requirements, Impacts and Constraints

3.1 Final Specifications and Requirements

We seek to develop a deep learning-based system of behavioral biometric authentication by touch gesture pattern and key stroke dynamics. The last system must be able to work as a perpetual and easy to use authentication system without disturbing the normal user interaction.

The proposed authentication system will be developed on the basis of deep learning based Siamese neural network architectures, where Siamese LSTM is the major model that will be used to extract temporal correlations in behavioral biometric data. The methodology applies the use of sequence based data preparation techniques such as normalization, temporal padding and structured pair generation to facilitate similarity learning. As opposed to the traditional multi class classification methods, the system is based on the framework of verification whereby behavioral embeddings are compared in order to identify the authenticity of the user. To balance between security and usability, similarity threshold of 65 percent is optimized empirically to carry out authentication decisions.

The system is deployed in Python based development environment with modern deep learning solutions like TensorFlow or PyTorch. Scientific computing libraries are used to perform data processing and feature engineering and analytical visualization tools are used to provide insights into performance. The datasets needed are time series keystroke and touch gesture interactions in the form of genuine and imposter pairs to train Siamese. The solution also works in a fully operating systems environment and can use standard input interfaces, like touchscreens and virtual keyboards, and does not require specialized hardware, but can be scaled to web-based and mobile deployment.

3.2 Societal Impact

The proposed framework of behavioral biometric authentication would improve the online level of security as it would decrease the reliance on the systems of credentialing that are not based on movement. With its ability to exploit the natural user interaction patterns, the system ensures higher levels of protection to sensitive information and still has high level of usability and low level of user burden. This strategy facilitates the safe use of important digital applications such as financial applications, healthcare systems, and online trade. The decrease in the cases of credential compromise would help to increase the levels of trust the population has in digital technologies and support the cybersecurity ecosystem as a whole.

3.3 Environmental Impact

The system is largely a software-based one and it is based on digital data acquisition, therefore the use of minimal physical resources. No special biometric equipment minimizes electronic waste and production impact. Even though model training is energy-consuming, model optimized network structures and efficient inference mechanisms do not restrain energy usage over time. The solution suggested, therefore, supports sustainable computing methods due to its focus on efficient digital security.

3.4 Ethical Issues

There are serious ethical issues regarding privacy, data security and fairness with behavioral biometric systems. User interaction data is sensitive personal information, and it should be secured with the help of the efficient storage and encryption systems. Data collection procedures ought to be performed in an informed user consent and honesty in communication on data use. In order to achieve fair system operation, the models should be tested by different groups of users to avoid bias due to behavioral or device variation. Adherence to data protection laws and biometric information responsible use are the keys to the continued trust of the user and system security.

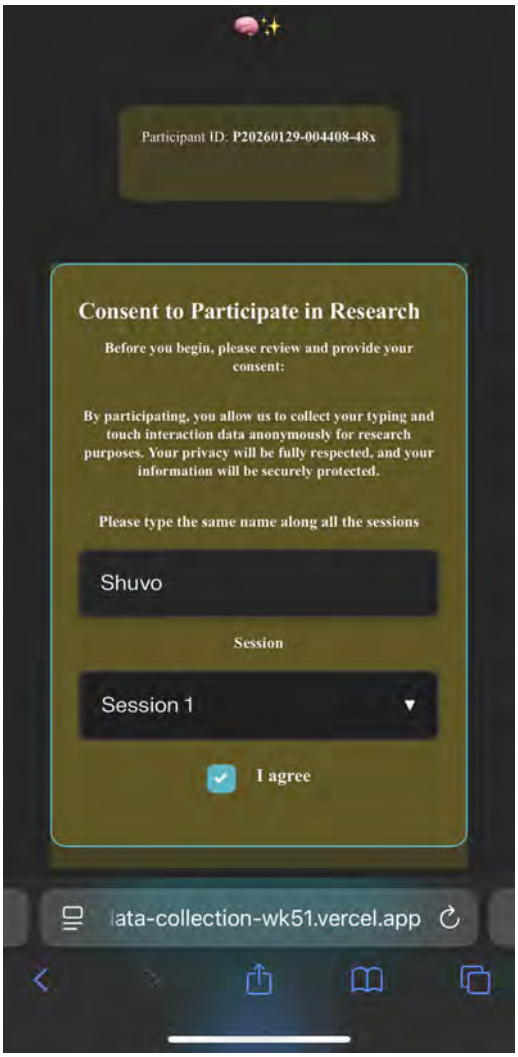


Figure 3.1: Consent Page of Data Collection Platform

3.5 Project Management Plan

The project was developed in a systematic development lifecycle, starting with the requirement analysis and literature review as the foundation of determining the system objectives and technical feasibility. Cross-device web platform was then used to collect and preprocess data to obtain behavioral biometric interactions. Model development The development of sequential data representation optimized Siamese neural network architectures. The optimization phase and evaluation phase entailed the performance optimization and robust testing to various experimental conditions. The last stage was system documentation and research reporting. The main sources of development included open-source software tools, publicly available datasets, and standard computing infrastructure, which allows implementing projects efficiently with a small amount of financial resource use.

3.6 Risk Management

The major risks in the project were data noise, behavioral variability, model overfitting, computational constraints and privacy. These issues were addressed by high levels of preprocessing, normalization and feature engineering. Validation methods and regularization measures were used to overfit. Neural alternatives were also considered to be lightweight in order to overcome computational limitations. The variability of behavior was dealt with by modeling sequence and optimization of similarity threshold. The threat of privacy was reduced by practices of anonymization and policies of secure data handling. Such mitigation measures provided robustness and reliability of the system.

3.7 Economic Analysis

The main expenses of the system are the time and resources spent on the development of the system and the use of the model to train and evaluate it. Nevertheless, an all-software design removes costs associated with the deployment of specialized hardware. The system is influential in terms of its long term benefits as it saves on the maintenance of authentication, causes less fraud related losses and it also enhances user experience. It is also cost effective due to its scalability through digital channels. The behavioral biometric system offers high-performance security and usability unlike other traditional credential-based system, and the high operational overhead is also lower, which is economically viable when it comes to a large-scale deployment.

Chapter 4

Proposed Methodology

4.1 Methodology Overview

To do user authentication perpetually, an authentication approach is presented in this study with behavioral biometrics touch gestures and keystrokes. The time-series sequence and the applications of deep learning architectures are used to facilitate real versus fake touch behavior discrimination. The degree of similarity between gesture patterns is evaluated in a preset determination threshold of 75%. The methodology consisted of two interrelated components, namely the design process and the system workflow.

4.1.1 Design Process

The system workflow begins with the collection of raw behavioral data through the interaction of the user with a particular web platform using their own devices. This aids in gathering real world touch and keystroke behavior and therefore continuous authentication can be achieved. Once the data has been acquired, the data is then cleaned, which addresses issues of missing values, and noisy data.

In the extraction of the features, behavioral features are developed in both kinds of data. In the case of keystroke, such features as the target character code, X, and Y coordinates, deletion, and flight time are processed. These features are also integrated with other features such as touch accuracy, error and log flight time. In case of touch data, such features as the location of the spatial coordinates, states, and path lengths are obtained, and other features such as movement velocity, directional angle, and pressure change are also computed. Standardizing all the numerical features is done using StandardScaler.

A two-pathway assessment is applied during the design process. The first to be used are the standard machine learning classifiers such as RF, and XGBoost. Such models are trained based on multi-class classification strategy to identify the users using their distinct behavioral feature vectors.

Secondly, the structured data in the context of the continuous authentication model is modeled as fixed-length sequences, i.e. 40 touch pattern timesteps and 35 keystroke pattern timesteps with a 50% overlap, by use of a sliding window method. The Siamese LSTM and Hybrid CNN-BiLSTM are fed with these sequences as inputs. The system forms pairs of sequential sequences of the same users as positive pairs

and random pairs with different users as negative pairs; in a 1:1 relationship. The structured information would be partitioned at the user level in order to prevent leakage. The Siamese network is trained to learn a similarity mode of two pairs in terms of distance measures such as Euclidean (L2) and Manhattan (L1) distance. The evaluation of the performance is done on the basis of accuracy, precision, recall and F1-score. The Siamese network operates on a similarity threshold which is normally 65% in order to decide whether to accept or reject users.

4.1.2 Workflow Diagram

Our workflow sequence throughout the process is given below

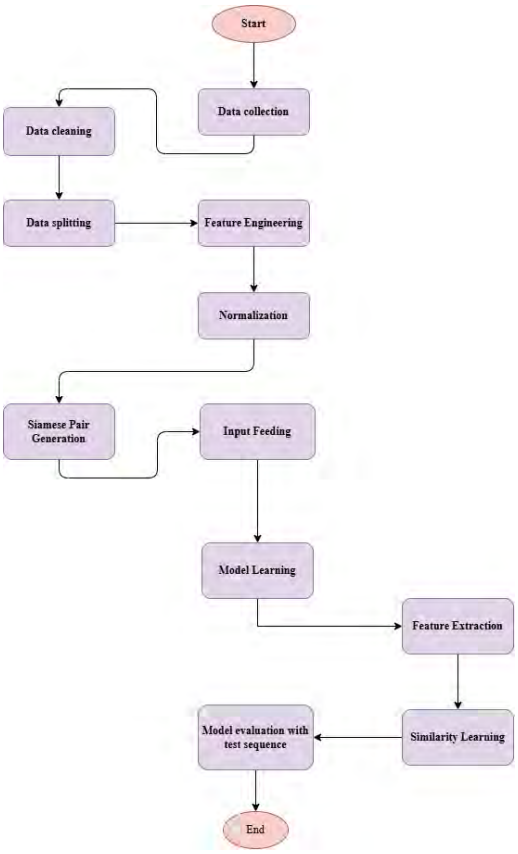


Figure 4.1: Work-Flow Diagram

4.2 Preliminary Design or Design (Model) Specification

A variety of ML and DL algorithms were studied in behavioral biometric authentication depending on the pattern of inputs modeling. Random Forests (RF) and XGBoost were used as traditional methods of classification. Although these models would have a high weighted average precision (just above 90%), they were found to have serious limitations to deal with unseen users, and recall on new samples even falls to zero in some cases. These models are thought to be insufficient in the adaptive systems since they need retraining of each new user. Moreover, in general, they cannot generalize among users without per-user information. Deep learning methods and especially Convolutional Neural Networks (CNNs) and BiLSTM were considered due to their capacity to acquire complex temporal relationships with sequential information. While CNNs are useful in spatial patterns, LSTM models are more useful in behavioral modalities such as keystroke dynamics and touch gestures. In the end, the final design adopted the use of the Siamese Neural Networks since it is concerned with the similarity metrics instead of direct classification. A Siamese model is composed of two equal subnetworks which process two inputs independent of each other with a shared weight. This enables the system to check whether a new behavior sequence is the same as a stored profile without having to use a multi-class classifier on each user. The behavioral sequences are analyzed in the system through an improved architecture of Siamese LSTM which is Architecture Components. Each branch is regularized with Layer Normalization, Batch Normalization, and Dropout (0.2-0.3) and encodes with a multi-layer LSTM encoder (between 128 and 512 hidden units).

Encoding Behavioral Sequences: Encoding behavioral sequences into embedding vectors $f(x)$. In the case of touch gestures, the three-layer LSTM encoder (512–256–128 units) generates a 64-dimensional embedding.

Similarity Computation: The system computes similarity using Euclidean distance, Manhattan distance, and cosine similarity, defined as

$$s = |f(x_1) - f(x_2)| \tag{4.1}$$

Prediction: The combined distance features are input to a dense unit with a sigmoid activation function, which outputs a likelihood score indicating whether two samples belong to the same user. The prediction is represented as a single value in the range $[0, 1]$.

Loss: The model is optimized using Binary Cross-Entropy loss, defined as

$$L(y, \hat{y}) = -\big[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})\big] \tag{4.2}$$

The iterative design procedure made use of intense preprocessing of data of 30 users: Feature Engineering: 12-13 features were processed per modality. Keystroke features were flight time log, error angle and touch precision. Touch gestures were implemented with features of the touch pressure variation, velocity, and movement angle. Temporal Sequencing: The rhythms of behavior were recorded by setting temporal limits (35 keystrokes and 40 touch). These limits were fixed length sequences.

Optimization: Architectures were designed in a way of establishing a trade- off between security and usability. Internal tests and holdout data of three unknown users were used to test impersonator rejection using the final trained models.

Empirical Threshold: It was set to produce more improved results, in comparison to the traditional systems setting the threshold to 75 percent. The internal test error was 86.96 percent and external test error was 89.17 percent respectively of the keystroke with the Siamese LSTM with a remarkable False Acceptance (FAR. The AUC of validation of the touch authentication model was 0.96+.

Important Observation: Experimental results indicate that generalizing traditional class models to holdout users is not possible. Therefore, they do not support dynamic authentication. Instead, the Siamese formulation was scalable, allowing personalized verification of the boundaries of classes without necessarily learning them. In long sequences, the LSTM layers had the ability to capture complex behavioural patterns such as swipes and typing rhythms. The system provides an efficient way of verifying behavioral biometrics through multi-metric similarity computation and improved embedding layers.

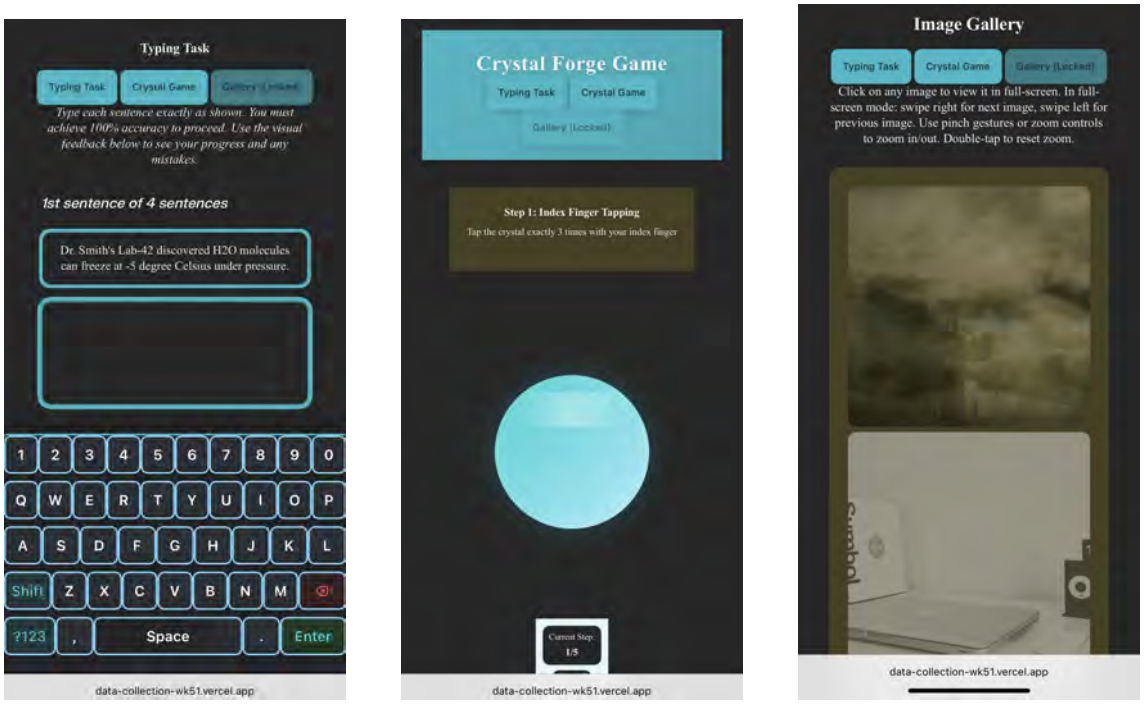
4.3 Data Collection

In this research, the goal was to develop a continuous user authentication system based on behavioral biometrics, with special emphasis on touch gestures and swipe dynamics. For collecting raw data from users, a specialized web application was built using vanilla javascript. Further, it was hosted using Vercel to run on different devices. A total of 30 user's data was collected in 4 different sessions. The web application contains 4 different sentences to type for storing keystroke data and 5 different touch based tasks, along with gallery interactions to store touch gesture data.

Dataset:

Keystroke Behavioral Data Collection

Keystroke Behavioral Data Collection In order to come up with the continuous behavioral biometric authentication system, a web-based data collection platform was constructed and implemented using the conventional web technologies and has been hosted on a cloud environment to make it accessible to many devices. To obtain variability in interaction, the data were gathered when the participants were exposed to natural interaction in 30 participants in four independent sessions. In case of keystroke biometrics the user was asked to type four key sentencing words, so that there would be uniformity and yet to maintain unique typing style. Every participant had a unique anonymized identifier that was created by combining elements based on the timestamp and randomness so that they could be tracked but not broken. To differentiate typing activity among other interaction tasks, all the entries on keystroke were identified by a fixed task identifier. Timing stamps were kept on each typing event to get a fine temporal typing dynamics. The typed characters were regularized so that they could face the differences like punctuation styles, special



(a) Typing data collection page

(b) Touch gesture data collection page (Task 1)

(c) Touch gesture data collection page (Task 2)

Figure 4.2: Mobile application screenshots for behavioral biometric authentication

symbols and so on in such a way that it would represent the input in a uniform manner irrespective of the users. To measure the behavior of finger location on the screen, the spatial touch coordinates on the screen were measured, and the center position of the virtual key was recorded when a custom keyboard interface was implemented to measure the spatial accuracy and motor control patterns. Deletion acts were particularly identified to indicate typing errors and correction behavior which are highly individualistic. Moreover, the inter- keystroke flight time was determined according to the time interval between the subsequent key presses, which represented the rhythm of typing of each user. All these temporal, spatial, and behavioral characteristics created a whole picture of a keystroke profile that was able to reflect some individual patterns of interaction of the user.

Touch Gesture Behavioral Data Collection

To record touch gesture biometrics, interactions were recorded in a sequence of interactive activities within the same web interface such as guided touch activities and naturalistic interaction in a gallery to record the controlled and uncontrolled user activities. Just like the keystroke information, the individual gesture sequences were associated with a distinct participant identifier and identified with the task-specific identifiers to distinguish interaction situations. In the case of task based activities, the number of trials was noted to track the development of gestures and behavioral consistency on repeat trials. The timestamps and timestamps of each touch were recorded so as to maintain motion over time. The finger contact on the screen was recorded using the spatial coordinates to model the gesture trajectories and movement accuracy. The system was able to capture the state of each touch interaction, separating initiation, movement and termination of contact to capture

the complete lifecycle of a gesture. The concurrent touch points were recorded to trace the behavior patterns of single-finger and multi-touch. Besides, cumulative gesture path length was calculated by adding the distance between two consecutive touch points and it is a quantitative measure of smoothness of gesture motion, speed, and spatial demarcation. All these features defined the dynamic motor behavior of users and helped to extract the signature of unique touch gestures which could be used in continuous authentication process.

4.3.1 Data Cleaning

The raw data, which consists of more than 49,015 keystrokes and 178,071 touch gestures, with 30 users, was first cleaned.

- Handling of Missing Values: Null values were removed.
- Encoding of Categorical Values: Labels like user identity, or Class, and reference characters, or refchar, were encoded into numerical values using LabelEncoder. Touch state, for example, in the touch gestures data, was also encoded.
- Removing Duplicates: Duplicates were removed for traditional machine learning models like XGBoost and RF.

4.3.2 Data Transformation

The phase of data transformation is present in the preparation of the features to the neural network model. Several data transformation methods were used to preprocess the data that was used in the neural network model. In the case of the numeric features, the StandardScaler was used to standardize all the datasets. It is significant to the neural network model, where the features must have a mean equal to 0 and a standard deviation equal to 1. In the case of categorical features, the user labels, touch states, including touchstart and touchmove and the reference characters of the keystroke data were applied in the LabelEncoder.

In both datasets, raw behavioral data were converted to a sequence of features to obtain unique patterns of each user. With the touch gesture data, the velocity of the movement direction, and variation in the touch pressure were engineered. Likewise, in case of the keystroke data, the touch accuracy, error range, and flight time were extracted. After the engineering of the features on the data, the data was converted into a fixed sequence of 40 items as in the touch gestures and 35 items as in the keystroke patterns. This was to suit the input requirement of the Siamese LSTM and Hybrid CNN-BiLSTM models. A predetermined order was created with a 50% stride of controlled overlap strategy. This has been done to enhance the quantity of data and the integrity of the data to the specific task or gesture. Each user was normalized on the global scaling of the engineered behavioral features to obtain the data.

4.3.3 Data Integration

Each of the datasets, keystroke and touch gestures, was processed separately using independent processing chains to retain their respective modality-specific information. The normalized data were then processed into pairwise inputs to train the Siamese model.

Positive and negative pairs were constructed using the following:

→ Positive pairs: Interactions of the same user, e.g., two different typing sequences or two different touch gestures by the same user.

→ Negative pairs: Interactions by different users.

These pairwise inputs enabled the model to identify a similarity function between user interactions without requiring any fusion of the data at the feature level. This strategy retained the credibility of each modality while allowing the model to generalize across both modalities by learning a shared similarity metric in the embedding space.

4.3.4 Data Reduction

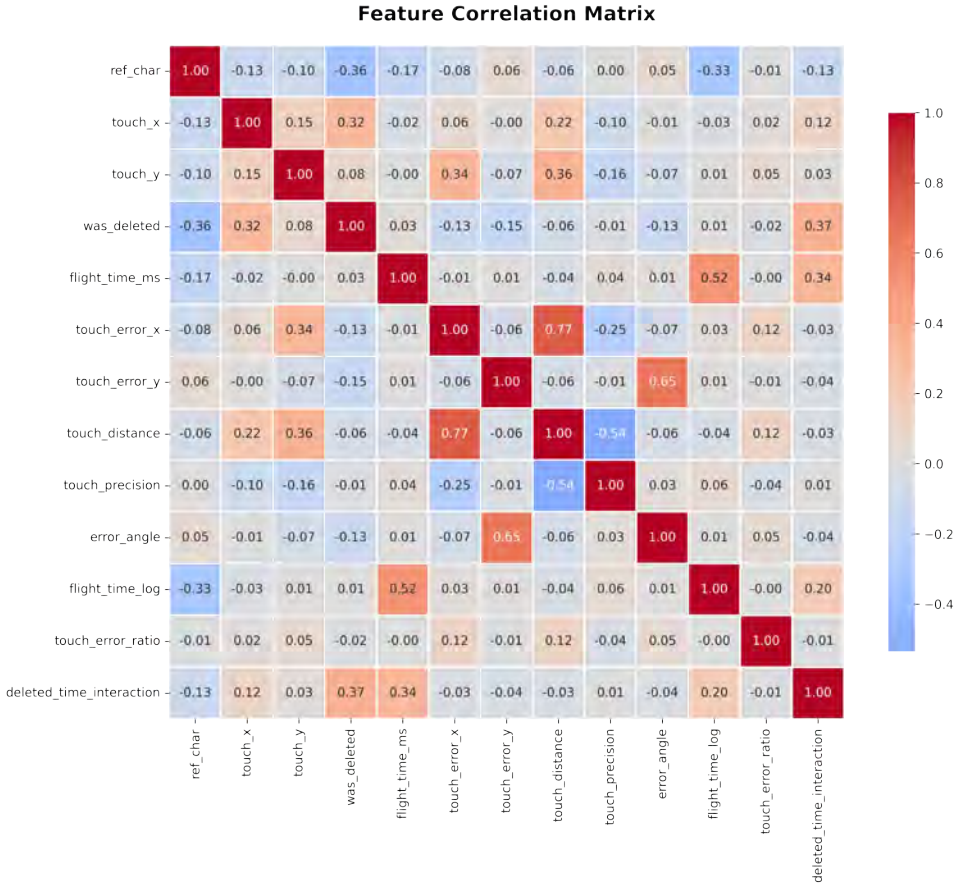
In order to simplify the data and to enhance the performance of the model, the different data reduction strategies were used. The datasets were not included in irrelevant and redundant features like eventid, timestamp, and taskid. In the keystroke dataset, raw spatial variables, that is, keyx and keyy were removed so that the model can focus on the temporal patterns as opposed to the spatial arrangement of the keystroke. Dense embedding layers in the neural network architecture also were used to reduce dimensionality of categorical features, e.g., reference characters (ref_char).

Normalization and standardized temporal windows were further used to reduce the high dimensional behavioral vectors. The length of all interaction sequences was limited to an equal number to make all of the input samples of the same shape. On the codes, the size of the temporal window was determined as constant (i.e. 35 steps in keystroke and 40 steps in touch gestures). This simplified the input side and enabled the model to generate appropriate behavioral pattern.

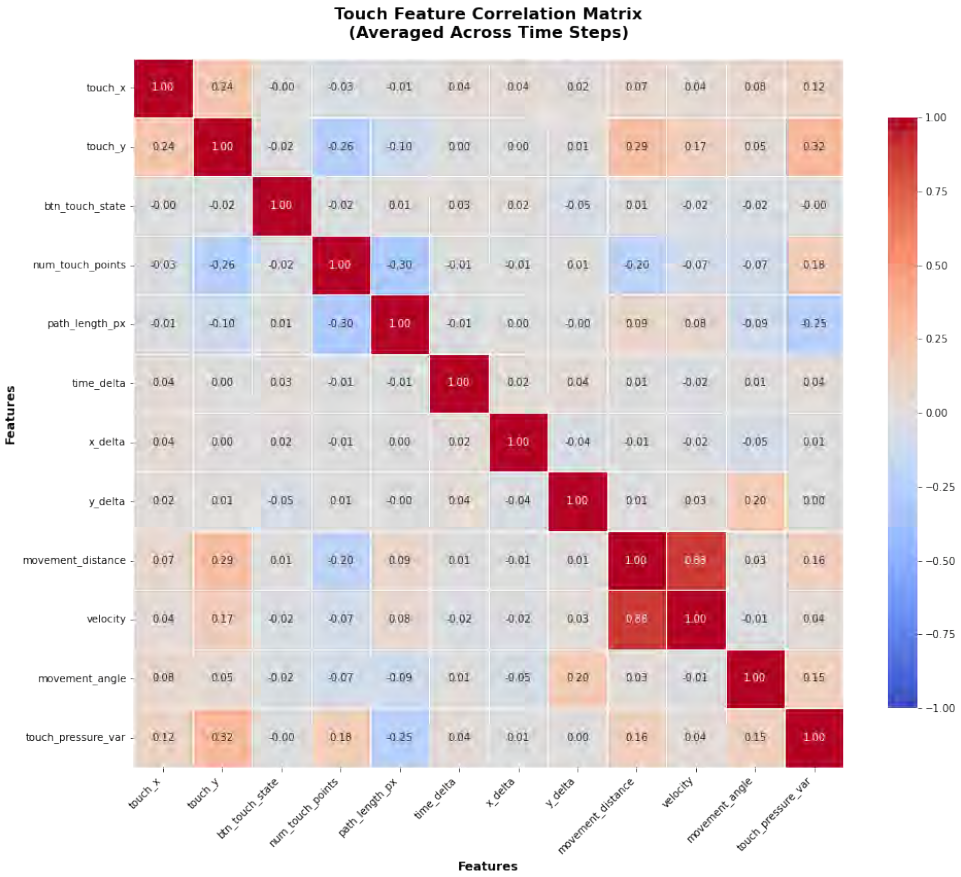
4.3.5 Summary of Preprocessed Data

After that, the datasets were transformed to appropriate formats of classic machine learning and deep learning structures. Despite the fact that the classic machine learning models, such as the Random Forest, and the XGBoost, were based on two-dimensional feature vectors, consisting of weighted averages, the format of the Siamese models was more complex. The keystroke data were transformed into a three-dimensional array, which contains time steps, precision-based characteristics, which are flight time log, touch precision and error angle. The touch gesture data were also reformulated to have numeric aspects like velocity, movement angle, and change in touch pressure. These pre-processed datasets, standardized to 35 steps for keystrokes and 40 steps for touch, formed the basis of the traditional machine learning phase. These datasets were then used to create thousands of balanced pairwise inputs to train the Siamese LSTM and Hybrid CNN-BiLSTM models. The datasets were then mapped to a common space, and their absolute difference was computed to learn to calculate behavioral similarity scores. This was found to be more efficient than traditional models, where the traditional ML models were unable to generalize to unseen users, whereas the Siamese model was able to achieve an internal accuracy of 86.96% and successfully verified identities based on empirical similarity threshold values from real-time biometric interactions.

Fig 3.2 shows correlation matrix of the preprocessed data of keystroke features and fig 3.3 shows correlation matrix of the preprocessed data of touch gesture features.



(a) Correlation matrix of preprocessed data of keystroke dataset



(b) Correlation matrix of preprocessed data of touch gesture dataset

Figure 4.3: Correlation analysis of extracted behavioral biometric features

4.4 Implementation of Selected Design

The implementation phase brought the aforementioned behavioral biometric authentication system to life, converting conceptual models into a functional and testable prototype. The principal objective was to develop a resilient, device-agnostic authentication system to verify user identity by observing touch and typing behavior in real-time. This was achieved through the deployment of multiple deep learning architectures. Specifically, Siamese networks and exhaustive testing with various statistical and perturbation-based methods.

4.4.1 Model Architectures

To compare the performance of different neural architectures on time-series behavioral data, two baseline machine learning model and two siamese deep learning model were trained and tested:

a) Random Forest (RF)
Random Forest was adopted to serve as a classical bagging-based ensemble method to give a reference to multi-class classification. The model tried to classify interactions into certain user buckets by creating a large number of decision trees in the course of the training method.
Just like other classical ML methods used in the study, the Random Forest model performed well in terms of high weighted average but had poor generalization with a macro average recall. It generated a poor accuracy inexternal test case. Its accuracy was grounded, by depicting that the traditional decision-tree structures could not easily be used to define the fluidity of behavioral biometric profiles without retraining every new user.

b) XGBoost (Extreme Gradient Boosting)
As a high-performance ensemble baseline, XGBoost was applied to assess the effectiveness of gradient-boosted decision trees on aggregated behavioral features. While it achieved a high weighted average precision on training data, its real-world utility was limited by its inability to handle unseen users.
Performance metrics demonstrated a significant drop in recall when faced with new subjects, resulting in an average overall accuracy. This confirmed that while XGBoost is computationally efficient for static classification, it lacks the adaptability needed for the dynamic similarity-matching tasks important to continuous authentication.

c) Siamese LSTM (Proposed Architecture For Touch Gesture)
Siamese LSTM architecture was employed as the primary design since it was most appropriate for extracting temporal relationships in user behavior sequences. The Siamese network consisted of each branch composed of LSTM layers that projected behavioral sequences (such as swipe actions or keystroke durations) into vectors of embeddings. The absolute difference between the embeddings was taken and input into a sigmoid-activated dense layer to predict similarity. This model performed better at modeling long-term temporal dependencies in keystroke and gesture sequences. It was highly resilient to input variation and noise, with stable accuracy and generalization across subjects. Its primary drawback, however, is that it is computationally costly; the model requires more memory and training time and is

therefore less amenable to deployment on low-power or mobile devices.

Siamese LSTM model shows high results in touch gesture authentication by scoring 92.41 percent in self-testing enrolled users. The model has strong external session (new authentication attempt) generalization (81.67% accuracy) with 18.33% EER and a False Accept Rate (FAR) of just 9.26 on holdout testing. The findings confirm the ability of the model to effectively differentiate between genuine and malicious users of the site according to their distinctive pattern of touch gestures.

d) Hybrid CNN-BiLSTM (Proposed Architecture For Keystroke)
The proposed hybrid architecture uses Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) layers. CNN layers are utilized as feature extractors that automatically extract small patterns in typing rhythms or gesture pressures. On the contrary, BiLSTM layers process the input sequences in forward and backward directions to comprehend the context of the behavior.

The hybrid CNN BiLSTM model is a decent performer when it comes to touch gesture authentication because this model can effectively learn both spatial and temporal features in the behavior of its users. The CNN layers are employed in the extraction of local spatial features on touch coordinates and BiLSTM element is employed in the process of estimating sequental dependencies and Dynamics of sequence patterns. This is a very discriminating architecture with an accuracy of 93.24% and Equal Error Rate (EER) of 8.33 which is quite discriminative between a real user and an impostor user. Thus, it is befitting in the authenticating systems in the real world.

4.4.2 Evaluation Strategies

For properly evaluating the performance, reliability, and practical usability of the used behavioral biometric authentication models, three evaluation strategies were used. These strategies ensured that the robustness, generalizability, and predictive criteria of the system were properly tested through various scenarios.

1. External Test Set Evaluation
This approach consisted of testing the model on the totally new session of data of users who already have data in the training set. The system was tested to investigate intra-user variability using a separate dataset that was gathered at a different time as compared to the first training data. This is a simulation of real world scenarios in which typing or gesture rhythm may change marginally depending on mood, stress or environment by the user.
Result: The Hybrid CNN-BiLSTM model showed a high degree of strength to the point that its accuracy on external keystroke data was actually better at 92.50%. On the same note, the Siamese LSTM touch gesture model had a high AUC, which demonstrated that the models acquired consistent behavioral signatures instead of memorizing a single session of data.

2. Holdout Test Set Evaluation (Impostor Rejection)
In this method a small number of users (3 users, which is 10% of the dataset) were not included in the training stage at all. These are the holdout users who were considered as unknown impostors in the testing stage. The model was tested with regards to the capability to accurately reject such non-observable users when compared to the enrolled users (Zero-shot testing). It is a very important security test in order to gauge False Acceptance rate (FAR) of random attackers.
Result: The model was able to separate enrolled users and the holdout impostors. The system was able to have high security capabilities and effective open-set generalization by systematically giving low similarity scores to pairs of pairs with holdout users, which showed that it could protect the system against unauthorized users that the system has never previously met.

3. Baseline Train-Test Split Evaluation
The 27 enrolled users were split (80/20) as a baseline evaluation. The approach was easy to reproduce and be compared to the past. It also enabled the end-to-end optimization of hyperparameters and performance based on conventional metrics of validation such as AUC and Accuracy.
Result: The 80/20 ratio showed a steady performance having a validation accuracy of about 89-94% and an AUC of 0.96. Similarity threshold was set empirically in order to trade off false acceptance and false rejection rates to obtain a stable point of decision of user authentication.

A combination of these test methods enabled an in-depth study of the model under varied operating conditions. They allowed the robustness of the model to be fully tested.

4.4.3 Target metrics:

The main aim of the model is to be highly accurate and precise to authenticate the legitimate users in a proper manner and reduce false acceptances. The model is also aimed to moderately reciprocal and F1-score, and, thus, a trade-off between sensitivity and overall strength. This merge of measures is a representation of a real trade off, where there is a lopsidedness in the correct recognition of real users and a logical level of tolerance in the false rejection. Such performance metrics are especially critical in systems of behavioral biometrics, where the natural inconsistency can be generated by the variability of the patterns of user input.

4.4.4 Implementation Summary

The implementation was able to change the desired behavioral authentication system into reality in that the proposed system was initially a theoretical design that was converted into a running deep learning framework that can perform real-time user authentication. The system was designed so that it encodes user behavior in high dimensional embeddings and enables it to be authenticated based on similarity metrics instead of strict classification.

Key milestones are

- **Deep Learning to Evolution:** Early development stages were based on classical machine learning classifiers, i.e. Random Forest (RF) and XGBoost. Empirical testing however showed that these models were not sufficient to capture the non-linear and complex temporal dependencies of behavioral data. As a result, the application shifted to the use of the Neural Networks based on the Siamese neuron (with LSTM and Hybrid CNN-BiLSTM architectures) that showed a better ability to capture sequential patterns than the initial baselines.

- **Strong Data Processing Pipeline:** A behaviorally heterogeneous preprocessing pipeline was created that is scaleable. This included the normalization of the variable-length sequence of keystroke and touch gestures, the use of the sliding window segmentation, and outlier elimination to guarantee the high quality of the input to the neural networks.

- **Broad-based Empirical Validation:** The models have been empirically subjected to extensive empirical. Comparisons of three evaluation strategies Internal Split, External Session and Holdout Impostor Rejection. This made the model more examined since it was confirmed to be accurate as well as its lastingness through time and resistance to unknown enemies.

- **Architectural Optimization:** The validation findings proved that the Siamese LSTM/Hybrid architecture was the best design solution. The Siamese network was able to effectively learn the space of similarity metrics unlike static decision boundaries of XGBoost, and thus greatly resistant to noise in inputs and able to learn the dynamism of human-computer interaction.

This leads to a strong behavioral biometric authentication system which works far better than the traditional ML methods. It provides functionable, high potential solution to prevalent mobile environment, which needs daily, low-friction and high security authentication.

Chapter 5

Result Analysis

This part presents a detailed analysis of the experimental results and performance obtained from the proposed behavioral biometric authentication system. Two independent datasets were created for this study, one for keystroke dynamics and one for touch gesture patterns.

To evaluate the robustness and generalization capability of the system, two machine learning model and two deep learning model were trained and tested on the datasets:

- Random Forest (RF)
- XGBoost
- Siamese LSTM
- Hybrid CNN-BiLSTM

The performance of each model were evaluated using multiple evaluation metrics that are commonly used in behavioral biometric authentication systems.

5.1 Performance Evaluation

Model performances were assessed using a combination of classification metrics and biometric-specific error rates, such as Accuracy, F1-score, Recall, Precision, and ROC curve. Each metric denotes how the model is performing and what room there is for improvement. These were used to authenticate users based on keystroke dynamics and touch gesture patterns. To calculate model accuracy and classification ability, the following universally accepted metrics were used, translated in the case of user authentication:

1. **Accuracy:** Illustrates the overall proportion of correctly classified samples, genuine users, and impostors.
2. **Precision:** Refers to the proportion of users identified as authentic who are genuine.
3. **Recall:** Approximates the system's ability to correctly classify all authentic user attempts.

4. **F1-Score:** The harmonic mean of precision and recall, used to provide an unbiased assessment of the model's performance.

5. **Equal Error Rate (EER):** The position where the False Acceptance Rate and False Rejection Rate are equal. It is commonly user in biometric systems.

6. **False Acceptance Rate (FAR):** The rate at which impostors are incorrectly authenticated.

7. **False Rejection Rate (FRR):** The rate at which true users are incorrectly declined.

These measurements were complemented by graphical tools to facilitate interpretability, particularly under noisy or class-imbalanced conditions like:

- ROC (Receiver Operating Characteristic) Curves
- Precision-Recall Bar chart.
- FAR-FRR crossover graphs

All these visualizations made it easier to understand the learning behavior of participants and the decision boundaries of the models.

5.2 Analysis of Design Solutions

The study proposes multiple design solutions for behavioral biometric authentication using keystroke anf touch gesture data. Both classical machine learning models like Random Forst and XGBoost, and deep learning architectures like Siamese LSTM and Hybrid CNN-BiLSTM are used.

Traditional machine learning models shows high efficiency and low computational cost and it makes them suitable for systems with limited resources. However, their performance degrades under external testing conditions that indicated limited generalization capability.

Deep learning based architectures achieved better generalization and balanced error rates, especially under cross session and external testing. However, those model require higher computational resources and longer training time.

Overall, the design solutions shows a trade off between robustness and computational efficiency. Here, deep learning models offer a enhanced security performance at a higher computational cost.

5.3 Final Design Adjustments

To optimize the systems performance and address generalization challenges, some critical adjustments were implemented during the final design phase. This improvements targeted three key areas like regularization, personalized normalization and training data diversity.

Dropout Regularization: Dropout layers were implemented after convolutional and dense layers. This regularization prevents overfitting by randomly removing neurons during training. It forces the network to learn robust, generalizable features rather than memorizing patterns.

Per User Normalization: For 27 enrolled users, individual feature scalers were fitted. It normalizes touch gesture features(pressure, velocity, movement distance) relative to each user's personal behavioral nuances rather than global population average. This personalized approach enhances the model's ability to differentiate genuine user variations from impersonation attempts and preserves individual behavioral signatures.

Aggressive Data Adjustmentation: The training data was expanded from 6555 to 19663 sequences (200% increase) through 5 complementary transformations such as spatial shifts, time resampling, rotation, noice injection and scaling. This augmentation strategy simulates realistic gesture execution and maintains bias power for user authentication.

5.4 Statistical Analysis

This section presents a statistical analysis of all four implemented models based on accuracy, computational efficiency, deployment feasibility, and overall authentication reliability.

- The classical machine learning models- Random Forest and XGBoost offered fast training and low computational cost but showed limited generalization and higher behavioral biometric error rates.
- Hybrid CNN-BiLSTM achieved best performance for keystroke dataset by combining spatial and temporal learning but required higher computational resources.
- Siamese LSTM delivered the best result in touch gesture dataset, particularly under external and holdout testing conditions, due to its ability to model sequential behavioral pattern.

5.4.1 Keystroke Dataset Result Analysis

The accuracy metrics all four models on keystroke dataset is simmarized in Table 5.1. It compares training accuracy, test accuracies and holdout accuracy. The other metrics (precision, recall, F1 score and EER) are shown in Table 5.2.

Table 5.1: Model Performance (Accuracy Metrics)

Model	**Train Acc (%)**	**Internal Test Acc (%)**	**External Test Acc (%)**	**Holdout Acc (%)**
Random Forest	N/A	88.99	74.21	N/A
XGBoost	N/A	90.03	79.49	N/A
Siamese LSTM	95.56	88.98	89.17	83.89
Hybrid (CNN–BiLSTM)	92.53	93.24	92.50	88.15

Table 5.2: Model Performance (Precision, Recall, F1, and EER)

Model	**EER (%)**	**Precision (%)**	**Recall (%)**	**F1 (%)**
Random Forest	12.84	99.70	74.21	83.48
XGBoost	11.16	100	79.49	87.58
Siamese LSTM	11.67	85.74	93.52	89.46
Hybrid (CNN–BiLSTM)	8.33	88.60	99.26	93.62

From the above tables, we can clearly observe that the Hybrid CNN-BiLSTM model outperforms all other models based on overall performance mertices. It showed the highest external accuracy (92.50%) and the lowest EER (8.33%). Compared to the traditional ML models, the deep learning approaches—especially Siamese LSTM and the Hybrid model—demonstrated significantly better generalization and reliability. These results confirm that deep learning architectures are more effective for keystroke-based behavioral biometric authentication.

The comparative ROC curve of the deep learning models is presented in the figure 5.1.

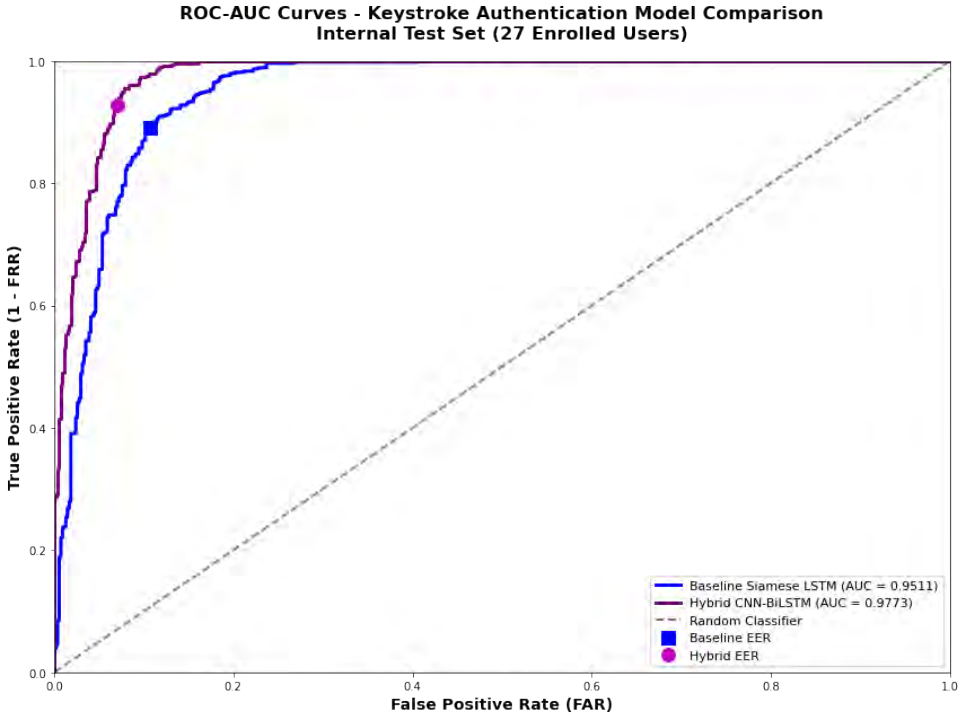


Figure 5.1: AUC Comparison for Keystroke-Based Models

5.4.2 Touch Gesture Dataset Result Analysis

A similar evaluation was performed for the touch gesture dataset. The comparative results are presented in Table 5.3 and 5.4.

Table 5.3: Model Performance (Accuracy Metrics)

Model	**Train Acc (%)**	**Internal Test Acc (%)**	**External Test Acc (%)**	**Holdout Acc (%)**
Random Forest	N/A	99.74	69.33	N/A
XGBoost	N/A	99.90	67.57	N/A
Siamese LSTM	94.52	92.41	81.67	94.63
Hybrid (CNN–BiLSTM)	91.27	85.56	75.00	49.26

Table 5.4: Model Performance (Precision, Recall, F1, and EER)

Model	**EER (%)**	**Precision (%)**	**Recall (%)**	**F1 (%)**
Random Forest	5.18	97.91	69.33	73.32
XGBoost	16.53	83.43	66.57	68.35
Siamese LSTM	18.33	78.79	86.67	82.54
Hybrid (CNN–BiLSTM)	26.67	74.19	76.67	75.41

For touch-based authentication, the Siamese LSTM model showed the strongest overall performance with an external accuracy of 81.67%. It outperformed both

Random Forest and XGBoost. Although the Hybrid CNN-BiLSTM model did not generalize well to external data, but still the deep learning approaches still demonstrated better consistency than the traditional ML models. These results indicate that temporal deep learning architectures remain more robust for modeling touch-gesture behavioral patterns.

The comparative ROC curve of the deep learning models is presented in the figure 5.2.

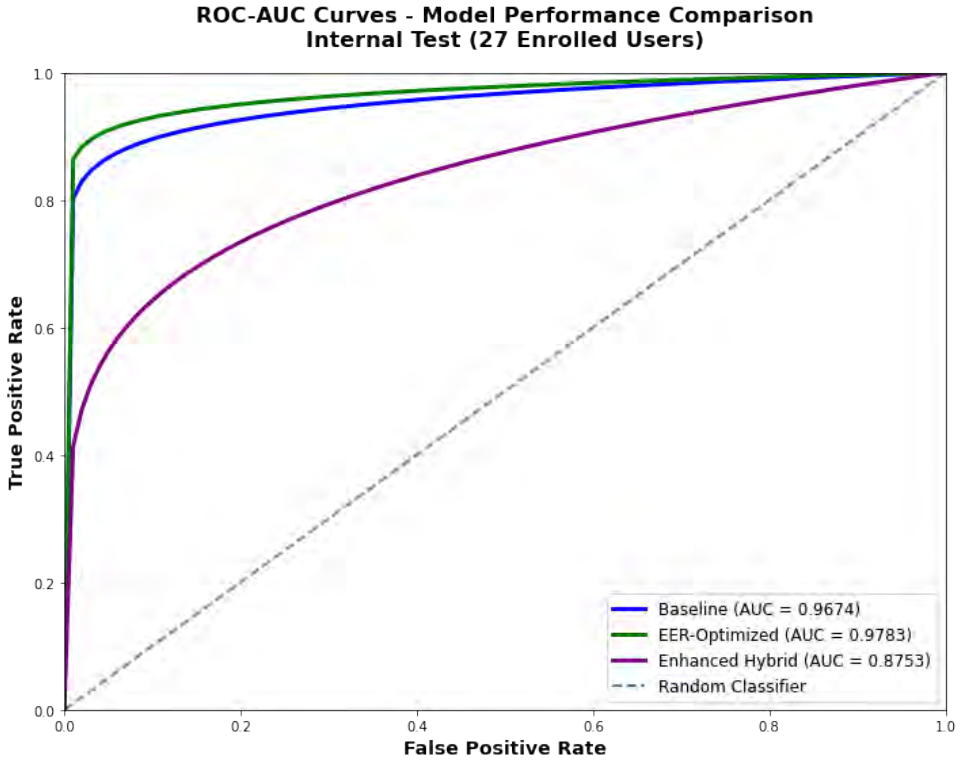
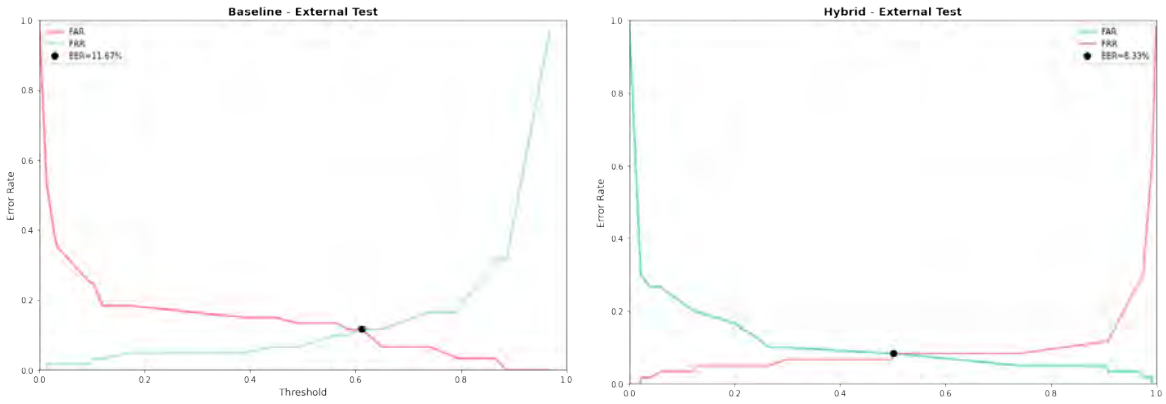


Figure 5.2: AUC Comparison for Touch-Based Models

5.4.3 Biometric Error Rate Analysis

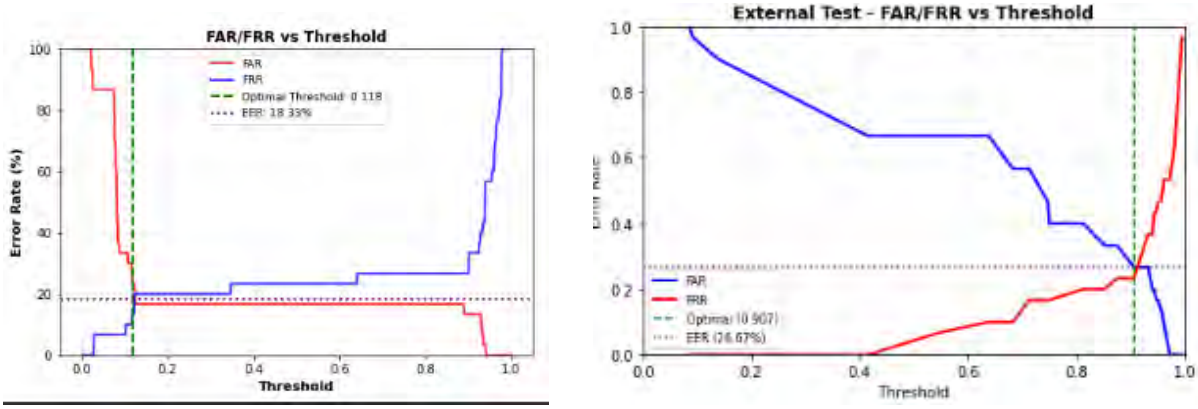
The hybrid CNN-BiLSTM model for keystroke dynamics data consistently achieved the lowest EER. It indicates better separation between genuine users and impostors. So here is the FAR vs FRR graph of the hybrid model and Siamese LSTM model for keystroke in 5.3.



(a) Siamese LSTM FAR/FRR Curve (b) Hybrid FAR/FRR Curve

Figure 5.3: Comparison of FAR/FRR curves for Hybrid and Siamese LSTM models

On the other hand, the Siamese LSTM model for touch dynamics data achieved the lowest EER. So here is the FAR vs FRR graph of the hybrid model and Siamese LSTM model for touch data in 5.4.



(a) Siamese LSTM FAR/FRR Curve (b) Hybrid FAR/FRR Curve

Figure 5.4: Comparison of FAR/FRR curves for Hybrid and Siamese LSTM models (Touch)

Although Siamese LSTM and Hybrid CNN-BiLSTM models incurred higher computational costs, it's superior accuracy and lower error rates justify its use in high-security authentication systems.

Here is the comparison table Table 5.5 of the computational cost for the applied models

Table 5.5: Computational complexity and deployment feasibility comparison

Model	**Training Cost**	**Inference Cost**	**Memory Usage**	**Deployment Suitability**
Random Forest	Low	Low	Medium	High
XGBoost	Medium	Medium	Medium	Medium
Siamese LSTM	High	Medium	High	Medium
Hybrid (CNN–BiLSTM)	High	High	High	Low–Medium

This table is the summary of computational complexity and deployment of every model. The training and inference costs of traditional ML models (Random Forest, XGBoost) are lower, whereas deep learning models particularly Siamese LSTM and the Hybrid CNN-BiLSTM models require more computation and memory, so they cannot be deployed in as light or real-time.

5.4.4 Comparison of Architectural Effectiveness Across Modalities

According to the comparative study of the experimental findings, the difference in performance between the hybrid (CNN-BiLSTM) and Siamese LSTM models in different modalities can be explained by the nature of the feature spaces. The hybrid architecture is better in keystroke dynamics since typing behavior has different local motifs including particular digraph flight time, which the 1D-CNN layers can well-extract as low-level rhythmic features, and the BiLSTM layers can capture the long-range time-dependent relationships between typing flow. In contrast, with touch gesture authentication, the Siamese LSTM is much better than the hybrid model due to the fact that touch data is continuous, high-dimensional, and displays a high intra-user variance. The hybrid model is prone to overfitting to the spatial coordinates or noisy local variations of touch path and the model cannot generalize across sessions. The Siamese LSTM however uses a distance-related metric learning model that seeks the comparative temporal stability in addition to relative shape-similarity of two gesture pairs, and thus much more immune to the natural variations of touch interactions compared to the fixed feature extraction maps of the hybrid network.

5.5 Comparisons with Existing Research

5.5.1 Overview of Comparative Study

This part presents a comparison of the proposed authentication system with the existing research of behavioral biometrics, in particular, the keystroke dynamics and the touch gesture based authentication. The comparison is concentrated on the representative studies, that are examples of classical machine learning models and deep learning architectures. The purpose is to find out how the proposed models perform relative to existing work in terms of accuracy, EER and robustness under real world conditions.

In contrast to many studies, which are based on interval or benchmark datasets, the current proposal uses newly collected datasets obtained through a custom built web application, which makes the evaluation more realistic compared to the practical deployment scenarios.

5.5.2 Comparison with Deep Learning–Based Authentication

Keystroke:

Contemporary studies have shown that deep learning architectures, perform much better at capturing temporal dependencies of keystroke data compared to classical approaches. Acien et al. [5] reported the effecting use of LSTM networks for improving the authentication rate is to have lower Equal Error Rates (EER) in comparison with standard techniques.
The proposed Hybrid CNN-BiLSTM model is inline with this trend and it goes a step further by introducing multi-level evaluation such as internal testing, external testing and holdout validation. The model had the following:

- **93.24% internal test accuracy (80-20 percent split testing)**
- **92.50% external test accuracy (testing with new session of the existing users)**
- **88.15% holdout test accuracy (testing with completely new users)**

These results indicate that the model has performed excellently in generalization and robustness. Thus, it has outperformed the classical models and remains competitive with existing deep learning based approaches reported in the literature.

Iyapparaja M et al. [17] have successfully made the case of unsupervised anomaly detection useful and the ensemble methods achieved 93.5% accuracy on fixed-text data. Nevertheless, they cannot extract complex temporal dependencies by their model, as their method of feature extraction and conventional machine learning is based on manual feature extraction. Our Hybrid CNN-BiLSTM keystroke model, in its turn, is superior, as it uses the deep feature learning process to automatically identify more complex typing patterns, which gives a stronger authentication mechanism, without the limitations of handcrafted features.

Sawant et al. [27] applies keystroke dynamics with the help of SVM, Random Forests, and Neural Networks to improve security by engaging in adaptive authentication and liveness detection. Their work is successful in tackling the conventional vulnerabilities surrounding passwords, but they use the standard classifiers, which usually have to be retrained with new users. On the contrary, we employed a Hybrid CNN-BiLSTM Siamese Network to perform few-shot learning, which can allow verification to be performed without retraining. Moreover, our multi-modal system provides better results with 89.17 accuracy and 95.19 recall rate by combining Touch Gestures with key strokes, which will provide a secure protection even when used in the mobile setting.

Kek et al. [18] compare CNN, GRU, and LSTM networks, with a CNN-GRU hybrid being the most stable with an 87 percent accuracy rate. Their emphasis on multi-class classification however limits the system to a predetermined number of users. Our Siamese Network architecture, conversely, supports few-shot verification, with better performance of 89.17 percent accuracy and 95.19 percent recall. This will facilitate scalable and accurate identification of outlawed users without having to retrain the model to accommodate the new persons.

Table 5.6: Comparison of Keystroke-Based Behavioral Biometric Authentication Approaches

Study / Method	**Modality**	**Model / Technique**	**Evaluation Metrics**	**Best Reported Performance**	**Key Limitations**
Iyapparaja M et al.'s [17] study on Enhancing User Authentication	Keystroke (fixed-text)	Isolation Forest, Random Forest, Gradient Boosting, Support Vector Machine	Accuracy, Recall, F1-score, False Positive Rate	93.5% accuracy (Random Forest and Gradient Boosting), 90% accuracy (Isolation Forest)	Fixed-text only; lacks Equal Error Rate and cross-environment evaluation
Kek at al's [18] study on user Authentication	Keystroke (free-text)	Convolutional Neural Network, Gated Recurrent Unit, Long Short-Term Memory, CNN–GRU Fusion	Accuracy, F1-score	87% accuracy (CNN–GRU fusion)	No FAR, FRR, or EER reported; sensitive to hyperparameter selection
Sawant et al.'s [27] study on Keystroke Dynamics	Keystroke	Traditional machine learning with adaptive classification	Accuracy (qualitative evaluation)	Improved authentication reliability (qualitative)	Lacks quantitative performance metrics and comparative error analysis
Proposed System (This Work)	Keystroke	Random Forest, XGBoost, Siamese LSTM, Hybrid CNN–BiLSTM	Accuracy, Precision, Recall, F1-score, EER	88.15% holdout accuracy, F1-score = 93.62%, EER = 8.33% (Hybrid)	Smaller dataset size and controlled experimental setup

Touch:

Touch gesture based authentication has been extensively studied in mobile security. [3] proposed a product called Touchalytics, which they used to demonstrate the functionality of swipe dynamics with behavioral features, whose EER values were promising under experimental conditions. [2] considered gesture-based continuous authentication by means of touchscreen, centering on the aspect of temporal

coherence.

The system analyzes touch gesture authentication through the Siamese LSTM and Hybrid CNN + BiLSTM model in deep learning. Among these two model, Siamese LSTM model showed good result with metrics value:

- **92.96% internal test accuracy (80-20 percent split testing)**
- **85.00% external test accuracy (testing with new session of the existing users)**
- **90.74% holdout test accuracy (testing with completely new users)**

In spite of the hybrid model displaying high internal accuracy (83.15%), its performance fell under external (63.33%) and holdout (11.11%) tests. The drop in performance reflects the increased variability inherent in real world gesture input and aligns with challenges noted in prior research.

Sejjari et al. [28] got impressive results with deep autoencoders and confidence mechanisms; they reached to the Equal Error Rate (EER) of 0.20 per cent on large-scale data. Although they have a powerful one-class classification framework, which is appropriate in detecting anomalies, they can be quite demanding in terms of training data per user to establish a boundary. The unique feature of our approach is a large direct similarity metric, which is learned by a Siamese CNN-BiLSTM-Attention architecture, which enables the highly accurate authentication even in few-shot situations when the data is scarce in comparison to the giant BrainRun data.

Fereidooni et al. [9] use motion sensors and few-shot learning to reach a 97% F1-score yet the system can only work in scenarios that involve device motion. Our approach fills this gap by concentrating on direct interaction modalities (Keystroke and Touch), ensuring 95.19% Recall and 89.81% Accuracy even in the situation when the device is not in motion and motion data is not available.

Bajaber et al. [7] defined the efficacy of hybrid deep learning architecture such as CNN-LSTM with a maximum accuracy of 96 on normal datasets. Compared to this standard implementation, our work has a higher level of performance because it has incorporated an Attention mechanism and a Siamese network-structure. This gives our model superiority in dealing with open-set situations (rejecting unseen impostors/holdout users) as opposed to the closed-set classification models used in their principal study.

Akiirne et al. [15] were capable of postulating the conceptual merit of applying Similarity Learning and Siamese networks to continuous authentication. This base is furthered by our approach through applying particular architectural optimizations such as aggressive data augmentation (200%+) and strict pair balancing. In places where their work presents the metric learning concept, our implementation gains superiority by such improvements to engineering stability and generalization to skilled impostors.

Table 5.7: Comparison of Touch Gesture–Based Authentication Approaches

Study / Method	**Modality**	**Model / Technique**	**Evaluation Metrics**	**Best Reported Performance**	**Key Limitations**
Sejjari et al's [28] study on Dynamic Authentication on Mobile Devices	Touch (swipe gestures)	One-Class Support Vector Machine, One-Class Gaussian Mixture Model, Isolation Forest, One-Class k-Nearest Neighbors, Autoencoder	Equal Error Rate	0.20% EER (Autoencoder, unbalanced data)	Can lead to higher False Acceptance Rates when an imposter's swipe is simple enough to be reconstructed by the model
Bajaber et al's [7] study on Evaluation of Deep Learning Models for Touch Gesture Authentication	Touch (gesture dynamics)	Long Short-Term Memory, Gated Recurrent Unit, CNN–LSTM, CNN–GRU	Accuracy, FAR, FRR	Up to 96% accuracy (CNN–GRU / CNN–LSTM)	Lack of scalability, as the system is rigid and bound to the specific set of users in the training dataset
Akiirne et al.'s [15] study on Dynamic Swipe Gestures based Continuous Authentication	Touch (swipe gestures)	Siamese Network using similarity learning	Accuracy, distance-based similarity	High authentication reliability (qualitative)	Loss of temporal resolution, as acceleration, pauses, and velocity changes are flattened into a static representation
Proposed System (This Work)	Touch + Gesture	Random Forest, XGBoost, Siamese LSTM, CNN–BiLSTM (Hybrid)	Accuracy, FAR, FRR, EER	90.74% external accuracy, EER = 15.00% (Siamese LSTM)	Limited dataset size; some metrics are still under update

5.5.3 Summary of Comparative Findings

Based on the comparative analysis, deep learning based authentication methods usually have a better ability to be robust and generalize compared to the traditional machine learning models. Although ensemble techniques are highly accurate in the controlled environment, they deteriorate in performance in the external testing. Models that rely on similarity-learning, specifically those that are designed as siamese, are less unbalanced and are better at impostor detection. The proposed system has competitive performance as compared to the existing ones, even though the dataset is smaller, and the system focuses on security-oriented measures like FAR, FRR, and EER. All in all, the results prove that multimodal behavioral biometrics offer an effective and efficient approach to constant user authentication.

5.6 Comparative Analysis of Methodology

Although the performance of the Deep Learning model can be measured by quantitative factors such as Accuracy and Equal Error Rate (EER), the viability of the system can be assessed based on usability and security factors. Using the framework developed by [1], the viability of the proposed behavioral biometric model can be assessed against traditional knowledge-based approaches such as passwords or PINs and physical biometric approaches such as fingerprint or face recognition.

Table 5.8: Authentication System Comparison Based on the UDS Framework

Category	**Criteria**	**Knowledge-based**	**Physical Biometric**	**Proposed System**	**Justification for Proposed System**
Usability	U1: Memorywise-Effortless	Fail (Must remember complex strings)	Pass (Nothing to remember)	Pass	Normal typing and swiping authenticate the user without memorization.
	U4: Physically-Effortless	Fail (Typing specific password/pin)	Quasi (Requires specific scan action)	Pass	Passive and continuous authentication without explicit scan actions.
Deployability	D1: Accessible	Pass	Fail (Injury, skin issue)	Pass	Adapts to user behavior regardless of physical conditions.
	D2: Negligible-Cost-per-User	Pass	Fail (Requires specific scanner/hardware)	Pass	Requires no additional hardware beyond standard input devices.
	D4: Browser-Compatible	Pass	Quasi (Requires hardware drivers)	Pass	Uses standard HTML5 and JavaScript events for data capture.
Security	S1: Resilient-to-Observation	Fail	Pass (Cannot be stolen)	Pass	Timing and pressure features cannot be visually replicated.
	S2: Resilient-to-Impersonation	Fail	Pass	Pass	Subconscious motor behavior resists targeted attacks.
	S8: Resilient-to-Theft	Fail (Can be stolen/phished)	Quasi (Standard can be spoofed/lifted)	Pass	Dynamic behavioral traits cannot be stolen or reused.
	S11: Unlinkable	Pass	Fail (Unique identifier)	Pass	Context-dependent patterns prevent cross-site linking.

5.7 Discussions

The Hybrid Siamese architecture (CNN-BiLSTM-Attention) was introduced to address the shortcomings of the tree-based classifiers, the closed-set nature of which, was overcome by transitioning to the more flexible verification paradigm. The attention mechanism was effective at isolating high-variance biometric markers by making use of CNNs to extract local spatial features, rather than absolute coordinates. It makes the model less dependent on the device. Nevertheless, the analysis showed that there were major data set shortcomings; the greater use of aggressive data augmentation (¿200%) shows that there was no natural entropy and longitudinal divergence in the unprocessed samples. Amazingly high levels of accuracy in turn are therefore probably due to strong resistance to synthetic mathematical noise, and not actual behavioral drift. In addition, the deployment is limited by the high computational latency of the model on edge devices and validation gap because the system was only tested on random imposters and not skilled forgeries.

Chapter 6

Conclusion

6.1 Summary of Findings

The results of the experiment show that various model architectures are more appropriate to various behavioral modalities. In keystroke-based authentication, the hybrid CNN-BiLSTM model displayed the most consistent and reliable performance and good generalization in internal, external, and holdout assessment. However, in contrast, with respect to touch gesture based authentication, the performance of the Siamese LSTM model was greater as compared to the other models, since it attained a more balanced error rates, which implies that it is more efficient in generating similarity patterns among touch dynamics. The results of these findings indicate the significance of mode-specific architectures in the efficient behavioral biometric authentication.

6.2 Contributions to the Field

This thesis involves in the body of behavioral biometric security to design and test a multimodal continuous authentication system, which combines the dynamics of keystroke with touchscreen gesture behavior through a single, learning-based framework. One of the contributions is the creation of a realistic cross-device web-based data collection system that can be used to record fine grained temporal and spatial behavioural information using specific tasks and thus can be used in real world practical application.

The paper contributes to the development of authentication studies by providing a comparative study of classical machine learning methods and current deep learning paradigms of the identical behavioral dataset. The research underlines the role of temporal representation learning in behavioral biometrics by showing that Siamese similarity learning is better at gesture verification and hybrid CNN-BiLSTM learning model works better at keystroke authentication. Moreover, the use of the continuous and implicit authentication surpasses the stationary systems of logins and responds to the increasing necessity of frictionless but safe check-in of digital identities.

The other contribution is the fact that a credible biometric authentication can be acquired by the use of common smartphone interaction sensors and no special equipment is required, which can dramatically enhance the level of scalability and the

ability to deploy it. The study thus fills the gap between theoretical and practical concepts of behavior modeling and mobile security applications.

6.3 Recommendations for Future Work

Although this paper has established that it is possible to have continuous behavioral biometric authentication by using keystroke and touch interaction pattern, more studies are needed to improve the scalability, robustness, and practicability in the real world. Further efforts in future should focus on gathering bigger and more heterogeneous data sets in various demographic groups, devices, and usage conditions to enhance the generalization of models, as well as to lower sensitivity to behavioral heterogeneity. Longer data gathering during more prolonged timeframes would allow studying the consistency of the behavior and the natural development of the user interaction patterns which would lead to more adaptive authentication systems.

Future progress in the design of models can also augment the performance of the systems through the introduction of attention-based models, transformer architectures, and better temporal feature fusion methods. Strategy of online learning might enable the authentication models to revise with time as user information is received and thus remain accurate despite changing behavioral trends. Also, a larger-scale security testing in adversarial conditions, such as imitation attempts and signal manipulation, would be necessary to confirm robustness before going in the field.

A significant practical application of the work is to integrate the authentication system directly as a part of mobile operating systems, and not through application-level implementations. Integration would be offered on an OS level to allow access to native input streams and system-wide event interaction on a more fidelitous level, allowing continuous authentication between applications to occur. Moreover, integration of behavioral biometrics with already existing physical biometric operations like fingerprint scanning or facial recognition, and knowledge-based authentication like PINs or passwords, has a great potential of greatly enhancing mobile security without compromising user convenience. The use of secure execution environments and guarded storage modules in the current mobile operating systems would further improve privacy protection and protection against unauthorized access to the system. All these developments would bring the behavioral biometric authentication closer to mass commercialization.

Bibliography

[1] J. Bonneau, C. Herley, P. C. van Oorschot, and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes," in *2012 IEEE Symposium on Security and Privacy*, IEEE, 2012, pp. 553–567. DOI: 10.1109/SP.2012.44.

[2] T. Feng et al., "Continuous mobile authentication using touchscreen gestures," in *2012 IEEE Conference on Technologies for Homeland Security (HST)*, IEEE, 2012, pp. 451–456. DOI: 10.1109/THS.2012.6459891.

[3] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, *Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication*, 2012. arXiv: 1207.6231 [cs.CR]. [Online]. Available: https://arxiv.org/abs/1207.6231.

[4] A. Ometov, S. Bezzateev, N. Mäkitalo, S. Andreev, T. Mikkonen, and Y. Koucheryavy, "Multi-factor authentication: A survey," *Cryptography*, vol. 2, no. 1, p. 1, 2018. DOI: 10.3390/cryptography2010001.

[5] A. Acien, A. Morales, J. V. Monaco, R. Vera-Rodriguez, and J. Fierrez, "Typenet: Deep learning keystroke biometrics," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 4, no. 1, pp. 57–70, 2021. DOI: 10.1109/TBIOM.2021.3112539.

[6] I. Alsaadi, "Study on most popular behavioral biometrics, advantages, disadvantages and recent applications : A review," *International Journal of Scientific Technology Research*, vol. 10, pp. 15–21, Jan. 2021.

[7] A. Bajaber, M. Fadel, and L. Elrefaei, "Evaluation of deep learning models for person authentication based on touch gesture," *Computer Systems Science and Engineering*, vol. 42, no. 2, pp. 465–481, 2022. DOI: 10.32604/csse.2022.022003. [Online]. Available: https://doi.org/10.32604/csse.2022.022003.

[8] L. Pryor, J. Mallet, R. Dave, N. Seliya, M. Vanamala, and E. S. Boone, *Evaluation of a user authentication schema using behavioral biometrics and machine learning*, 2022. arXiv: 2205.08371 [cs.CR]. [Online]. Available: https://arxiv.org/abs/2205.08371.

[9] H. Fereidooni et al., *Authentisense: A scalable behavioral biometrics authentication scheme using few-shot learning for mobile platforms*, 2023. DOI: 10.48550/arXiv.2302.02740. arXiv: 2302.02740 [cs.CR]. [Online]. Available: https://arxiv.org/abs/2302.02740.

[10] V. Medvedev, A. Budžys, and O. Kurasova, "Enhancing keystroke biometric authentication using deep learning techniques," in *2023 18th Iberian Conference on Information Systems and Technologies (CISTI)*, 2023, pp. 1–6. DOI: 10.23919/CISTI58278.2023.10211344.

[11] A. H. Y. Mohammed, R. A. Dziyauddin, and L. A. Latiff, "Current multifactor of authentication: Approaches, requirements, attacks and challenges," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 1, 2023. DOI: 10.14569/IJACSA.2023.01401xx.

[12] V. Shanmugavalli, S. Suresh Kumar, and S. Nithya Kalyani, "A hybrid machine learning technique for multiple soft biometric based dynamic keystroke pattern recognition system," *Neural Processing Letters*, vol. 55, no. 8, pp. 10 845–10 871, 2023. DOI: 10.1007/s11063-023-11354-6.

[13] H. AbdelRaouf, M. M. Fouda, and M. I. Ibrahem, "Secure and robust user authentication using transfer learning and ctgan-based keystroke dynamics," in *2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI)*, 2024, pp. 1–6. DOI: 10.1109/ICMI60790.2024.10585811.

[14] T. Adeniran, R. Jimoh, E. Abah, N. Faruk, E. Alozie, and A. Imoize, "Vulnerability assessment studies of existing knowledge-based authentication systems: A systematic review," *SLU Journal of Science and Technology*, vol. 8, pp. 34–61, Jun. 2024. DOI: 10.56471/slujst.v7i.485.

[15] Z. Akiirne, Z. Naji, A. Saoudi, and D. Bouzidi, "Dynamic swipe gestures based continuous authentication using similarity learning," in *NISS 2024: The 7th International Conference On Networking, Intelligent Systems and Security*, ACM, Aug. 2024, pp. 1–6. DOI: 10.1145/3659677.3659739. [Online]. Available: https://doi.org/10.1145/3659677.3659739.

[16] A. Arsh, N. Kar, S. Das, and S. Deb, "Multiple approaches towards authentication using keystroke dynamics," *Procedia Computer Science*, vol. 235, pp. 2609–2618, 2024, International Conference on Machine Learning and Data Engineering (ICMLDE 2023).

[17] M. Iyapparaja, I. Karunanithi, and S. U. Bhat, "Enhancing user authentication through keystroke dynamics analysis using isolation forest algorithm," in *Proceedings of the Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)*, 2024, pp. 1–5. DOI: 10.1109/ICETITE58242.2024.10493648.

[18] X.-J. Kek, Y.-B. Leau, and S. F. Tan, "User authentication with keystroke dynamics: Performance evaluation in neural network," in *2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, 2024, pp. 30–35. DOI: 10.1109/IICAIET62352.2024.10730363.

[19] P. Lertvittayakumjorn, S. Cai, B. Dou, C. Ho, and S. Zhai, "Can capacitive touch images enhance mobile keyboard decoding?" In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '24, Pittsburgh, PA, USA: Association for Computing Machinery, 2024, ISBN: 979-8-4007-0628-8. DOI: 10.1145/3654777.3676420. [Online]. Available: https://doi.org/10.1145/3654777.3676420.

[20] S. Lim, C. Kim, B. H. Cho, S.-H. Choi, H. Lee, and D. P. Jang, "Investigation of daily patterns for smartphone keystroke dynamics based on loneliness and social isolation," *Biomedical Engineering Letters*, vol. 14, no. 2, pp. 235–243, 2024, ISSN: 2093-985X. DOI: 10.1007/s13534-023-00337-0. [Online]. Available: https://doi.org/10.1007/s13534-023-00337-0.

[21] P. G. do Nascimento, P. Witiak, T. MacCallum, Z. Winterfeldt, and R. Dave, *Your device may know you better than you know yourself – continuous authentication on novel dataset using machine learning*, 2024. arXiv: 2403.03832 [cs.AI]. [Online]. Available: https://arxiv.org/abs/2403.03832.

[22] K.-N. Nguyen, S. Rasnayaka, S. Wickramanayake, D. Meedeniya, S. Saha, and T. Sim, "Spatio-temporal dual-attention transformer for time-series behavioral biometrics," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 6, no. 4, pp. 591–601, 2024. DOI: 10.1109/TBIOM.2024.3394875.

[23] S. Oduri, "Continuous authentication and behavioral biometrics: Enhancing cybersecurity in the digital era," *International Journal of Innovative Research in Science Engineering and Technology*, vol. 13, pp. 13 632–13 640, Jul. 2024.

[24] I. Riaz, A. N. Ali, and H. Ibrahim, "Loss of fingerprint features and recognition failure due to physiological factors: A literature survey," *Multimedia Tools and Applications*, vol. 83, no. 39, pp. 87 153–87 178, 2024. DOI: 10.1007/s11042-024-19848-8.

[25] H. N. Risto and O. H. Graven, "Fixed-text keystroke dynamics authentication data set—collection and analysis," *Annals of Telecommunications*, vol. 79, no. 11, pp. 731–743, 2024. DOI: 10.1007/s12243-024-01039-z.

[26] E. A. Sağbaş and S. Ballı, "Machine learning-based novel continuous authentication system using soft keyboard typing behavior and motion sensor data," *Neural Computing and Applications*, vol. 36, no. 10, pp. 5433–5445, 2024. DOI: 10.1007/s00521-023-09360-9. [Online]. Available: https://doi.org/10.1007/s00521-023-09360-9.

[27] S. A. Sawant, S. V. Kalambe, and R. Pashte, "Keystroke dynamics: A machine learning approach to behavioural biometric authentication," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 10, no. 2, 2024. [Online]. Available: https://www.ijariit.com/manuscripts/v10i2/V10I2-1164.pdf.

[28] A. Sejjari, C. Moujahdi, N. Assad, and H. Abdelfatteh, "Dynamic authentication on mobile devices: Evaluating continuous identity verification through swiping gestures," *Signal, Image and Video Processing*, vol. 18, no. 12, 2024, ISSN: 1863-1711. DOI: 10.1007/s11760-024-03532-3. [Online]. Available: https://doi.org/10.1007/s11760-024-03532-3.

[29] M. Simão, F. Prado, O. A. Wahab, and A. Avila, "Tempcharbert: Keystroke dynamics for continuous access control based on pre-trained language models," in *2024 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2024, pp. 1–6. DOI: 10.1109/WIFS61860.2024.10810697.

[30] G. Tang, J. Zhao, Z. Xu, Y. Wang, H. Yang, and Z. Zhang, "Continuous smartphone user authentication based on gesture-sensor fusion," in *2024 International Conference on Networking and Network Applications (NaNA)*, 2024, pp. 516–522. DOI: 10.1109/NaNA63151.2024.00091.

[31] C. Wu et al., "It's all in the touch: Authenticating users with host gestures on multi-touch screen devices," *IEEE Transactions on Mobile Computing*, vol. 23, no. 10, pp. 10 016–10 030, 2024. DOI: 10.1109/TMC.2024.3371014.

[32] G. Stragapede et al., "Kvc-ongoing: Keystroke verification challenge," *Pattern Recognition*, vol. 161, p. 111 287, 2025, ISSN: 0031-3203. DOI: 10.1016/j.patcog.2024.111287. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320324010380.