Performance and Security Evaluation of Behavioral Biometric Systems

Yris Brice Wandji Piugie

Normandie Université

UNIVERSITÉ CAEN NORMANDIE

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité INFORMATIQUE

Préparée au sein de l'Université de Caen Normandie

Performance and Security Evaluation of Behavioral Biometric Systems

Présentée et soutenue par
YRIS BRICE WANDJI PIUGIE

Thèse soutenue le 28/11/2023 devant le jury composé de		
MME HÉLÈNE LAURENT	Professeur des universités, INSA Centre Val de loire	Rapporteur du jury
M. WILLIAM PUECH	Professeur des universités, Université de Montpellier	Rapporteur du jury
M. PATRICK BOURS	Professeur des universités, Norwegian University of Science and Tech	Membre du jury
M. JOEL DI MANNO	Ingénieur de recherche, FIME EMEA	Membre du jury
MME STÉPHANIE SCHUCKERS	Professeur, Clarkson University	Membre du jury
M. CHRISTOPHE ROSENBERGER	Professeur des universités, ENSICAEN	Directeur de thèse
M. CHRISTOPHE CHARRIER	Maître de conférences, Université de Caen Normandie	Co-directeur de thèse

Thèse dirigée par CHRISTOPHE ROSENBERGER (Groupe de recherche en informatique, image, automatique et instrumentation) et CHRISTOPHE CHARRIER (Groupe de recherche en informatique, image, automatique et instrumentation)

ENSI CAEN
ÉCOLE PUBLIQUE D'INGÉNIEURS
CENTRE DE RECHERCHE

MIIS

GREYC
Laboratoire de recherche en sciences du numérique

Acknowledgments

I would like to take this opportunity to express my deep gratitude to all those who have contributed and supported me throughout this work.

I would like to express my sincere gratitude to the GREYC laboratory, the Normandy region, and Fime SAS for giving me the opportunity to carry out this Ph.D. thesis.

A special thanks to my Ph.D. directors, Christophe Rosenberger and Christophe Charrier, whose dedication and support made this thesis possible. I can never thank you enough for your unconditional availability, especially when deadlines approached. You have a penchant for challenges, and I sincerely hope I haven't overstretched you.

I would also like to express my gratitude to Joël Di Manno, whose role in supervising this Ph.D. thesis was essential. Your kindness was greatly appreciated.

My thanks also go to Hélène Laurent and William Puech, who accepted to evaluate this Ph.D. manuscript, as well as Stephanie Schuckers and Patrick Bours for joining them in my Ph.D. jury.

I would like to express my sincere appreciation to all my colleagues at Fime SAS and GREYC. In particular, I would like to thank the SAFE team at GREYC, in which I had the privilege of working in a pleasant atmosphere. My warmest thanks go to all the members of the Biometrics team at Fime SAS for their support throughout this thesis, and in particular to my colleague Dr. Abdarahmane Wone, who was a faithful and cheerful companion. I truly appreciate working with all my colleagues from both Fime and GREYC.

I would also like to extend my warmest thanks to the secretariats and the MIIS doctoral school, who provided me with invaluable support throughout this thesis.

I would also like to express my sincere gratitude to my father, Thomas Piugie, and my mother, Lydie Mireille Tsamo, for the love, education, guidance and advice they have given me since I was born.

Finally, I am grateful to all my fellow friends and family, both nuclear and spiritual.Their support and help throughout this thesis have been invaluable, even if circumstances have made our meetings less frequent. I would like to take this opportunity to express my sincere gratitude.

Abstract

Behavioral biometrics offers new prospects for strengthening security and enhancing the user experience by analyzing users' interactions with IT systems. So it is an approach for identification and authentication based on the analysis of users' interactions with computer systems. While it can enhance security and improve the user experience, it raises privacy concerns. This Ph.D. thesis proposes a generic method for analyzing behavioral biometrics, with applications such as keystroke dynamics and human activities. Additionally, it also explores the effectiveness of Classical Machine Learning techniques for identification, as well as Deep Learning methods for user authentication based on their behaviors, with a focus on human activity on smartphones and keystroke dynamics on laptops. This Ph.D. thesis also proposes an innovative method for processing raw biometric data considered as time series. This provides far results to those already available. The time series processing consists of transforming the behavioral biometric raw data into a 2D image color. This transformation process keeps all the characteristics of the behavioral signal. Time series does not receive any filtering operation with this transformation and the method is reversible. This signal-to-image transformation allows us to use the 2D convolutional networks to build efficient deep feature vectors. This allows us to compare these feature vectors to the reference template vectors to compute the performance metric. We evaluate the performance of the authentication system in terms of Equal Error Rate (EER) on benchmark datasets and we show the efficiency of the approach. The results demonstrate that these approaches can achieve good performance, but also highlight potential privacy issues. It shows the effectiveness of this innovative approach in enhancing security without disrupting the user experience. Data security is crucially important in ensuring the safety of users and the confidentiality of their information in the field of cybersecurity.

This is why many companies have begun to implement authentication systems to control and restrict access to their data. However, some traditional authentication methods have proved insufficient to ensure adequate data protection, which is why behavioral biometrics has gained importance. Despite promising results and a wide range of applications, biometric systems remain vulnerable to malicious attacks, particularly presentation attacks. That is why, in this Ph.D. thesis, we set out to deploy a presentation attack against an authentication system based on behavioral biometrics. Our approach is to use the most popular temporal adversarial generators (TimeGAN) to create synthetic behavioral biometric data, which could be used to impersonate an authorized user. These synthetic data are generated while preserving temporal dynamics, meaning that new sequences respect the original relationships between variables over time. Finally, we validated both the original data and the synthetic behavioral biometrics generated. This validation was carried out using qualitative and quantitative similarity measures, as well as by assessing predictive ability. In addition, an authentication system was set up to assess the effectiveness of the data generated. The results obtained, together with a visual inspection, indicate that TimeGAN can indeed generate behavioral patterns that can be used to fool and consequently test behavioral authentication systems.

Keywords: Behavioral Biometrics; Cybersecurity; Identification; Authentication; Time series to Image; Synthetic Behavioral Biometrics; Presentation Attack Instrument.

Résumé

La biométrie comportementale offre de nouvelles perspectives pour renforcer la sécurité et améliorer l'expérience utilisateur en analysant les interactions des utilisateurs avec les systèmes informatiques. Il s'agit donc d'une approche pour l'identification et basée sur l'analyse de ces interactions. Bien qu'elle puisse renforcer la sécurité et améliorer l'expérience utilisateur, elle soulève des préoccupations en matière de vie privée. Cette thèse de doctorat propose une méthode générique pour analyser des séries temporelles en biométrie comportementale, avec des applications telles que la dynamique de frappe au clavier et les activités humaines. De plus, elle explore également l'efficacité des techniques d'apprentissage machine classiques pour l'identification, ainsi que des méthodes d'apprentissage profond pour l'authentification des utilisateurs basée sur leurs comportements, en mettant l'accent sur l'activité humaine sur les smartphones et la dynamique de frappe sur les ordinateurs portables. Les données biométriques comportementales étant représentées par des series temporelles, ces signaux peuvent donc subir des opérations de traitement de signal. Le traitement des séries temporelles consiste à transformer les données biométriques comportementales (considérées comme une série temporelle) en une image couleur 2D. Ce processus de transformation conserve toutes les caractéristiques du signal comportemental, sans aucune opération de filtrage. Cette transformation signal-en-image nous permet d'utiliser des réseaux de convolution 2D pour créer des vecteurs de caractéristiques profonds efficaces. Cela nous permet de comparer ces vecteurs de caractéristiques avec les vecteurs de modèles de référence pour calculer la performance. Nous évaluons la performance du système d'authentification en termes de Taux d'Égal Erreur (TÉE) sur des ensembles de données de référence, et nous montrons l'efficacité

de l'approche. Les résultats montrent que ces approches peuvent obtenir de bonnes performances, mais soulignent également des problèmes potentiels de confidentialité. Cela démontre l'efficacité de cette approche innovante pour renforcer la sécurité sans perturber l'expérience utilisateur. La sécurité des données est cruciale pour garantir la sécurité des utilisateurs et la confidentialité de leurs informations dans le domaine de la cybersécurité. C'est pourquoi de nombreuses entreprises ont commencé à mettre en place des systèmes d'authentification pour contrôler et restreindre l'accès à leurs données. Cependant, les méthodes d'authentification traditionnelles se sont révélées insuffisantes pour assurer une protection adéquate des données, c'est pourquoi la biométrie comportementale a gagné en importance. Malgré des résultats prometteurs et une large gamme d'applications, les systèmes biométriques restent vulnérables aux attaques malveillantes, en particulier les attaques par présentation. C'est pourquoi, dans cette thèse de doctorat, nous nous sommes fixé pour objectif de déployer une attaque par présentation contre un système d'authentification basé sur la biométrie comportementale. Notre approche consiste à utiliser les générateurs adverses temporels les plus populaires (TimeGAN) pour créer des données biométriques comportementales synthétiques, qui pourraient être utilisées pour se faire passer pour un utilisateur autorisé. Ces données synthétiques sont générées tout en préservant les dynamiques temporelles, ce qui signifie que les nouvelles séquences respectent les relations originales entre les variables au fil du temps. Enfin, nous avons validé à la fois les données originales et les données biométriques comportementales synthétiques générées. Cette validation a été réalisée en utilisant des mesures de similarité qualitatives et quantitatives, ainsi qu'en évaluant la capacité prédictive. De plus, un système d'authentification a été mis en place pour évaluer l'efficacité des données générées. Les résultats obtenus, associés à l'inspection visuelle, indiquent que TimeGAN peut en effet générer des modèles comportementaux pouvant être utilisés pour tromper et tester les systèmes d'authentification comportementale.

Mots clés: Biométrie comportementale; Cybersécurité; Identification; Authentification; Série temporelle; Signaux Comportementaux Synthétiques; Instrument d'Attaque par Présentation.

Contents

List of Figures

List of Tables

Abbreviations

2FA	Two-Factor authentication
AI	Artificial Intelligence
AUC	Area Under the Curve
APCER	Attack Presentation Classification Error Rate
BPCER	Bonafide Presentation Classification Error Rate
CCPA	California Consumer Privacy Act
CMC Curve	Cumulative Match Characteristic Curve
CNIL	known as the French National Commission for Information Technology and Freedoms
DCNNs	Deep Convolutional Neural Networks
DET	Detector Trade-Off
EER	Equal Error Rate
FAR	False Acceptance Rate
FIDO Alliance	Fast IDentity Online Alliance

FMR	False Match Rate
FNMR	False Non Match Rate
FRR	False Reject Rate
FTA	Failure-To-Acquire
FTE	Failure-to-Enrol
GDPR	General Data Protection Regulation
GRU	Gated Recurrent Unit
HAR	Human Activity Recognition
IAPAR	Impostor Attack Presentation Accept Rate
IAPIR	Impostor Attack Presentation Identification Rate
ICT	Information and Communication Technologies
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
LR	Logistic Regression
LSTM	Long Short Term Memory
LSTM LN	LSTM Layer Normalization
MFA	Multi-Factor Authentication
MLP	MultiLayer Perceptron
NN	Neural Networks
PA	Presentation Attack
PAD	Presentation Attack Detection
PAI	Presentation Attack Instrument

PIPA	Personal Information Protection Act
PIPEDA	Personal Information Protection and Electronic Documents Act
RGB	Red Green Blue
ROC	Receiver Operating Characteristic
SAR (IAPMR)	Spoof Acceptance Rate
SVM	Support Vector Machine
t-SNE	t-Distributed Stochastic Neighbor Embedding
TimeGAN	Time Series Generative Adversarial Networks
TOE	Target Of Evaluation
TPIPAS	Taiwan's Personal Information Protection and Administration System

Introduction

The proliferation of biometric technology has opened the way for innovative and secure identification and authentication methods. Biometric technology includes a wide range of techniques and methodologies for recognizing and verifying individuals on the basis of their unique physiological or behavioral characteristics [Mekruksavanich and Jitpattanakul, 2021, Piugie et al., 2022]. In our exploration, we examine the enrolment, verification, and identification processes, as well as the architectural aspects of biometric systems. A deeper understanding of these elements is essential to grasp the nuances of behavioral biometric certification. Behavioral biometrics, a subset of biometrics, focuses on the unique behavioral patterns of individuals, such as keystroke dynamics, gait recognition, and so on. The motivation behind the use of behavioral biometrics lies in its potential to offer increased security, convenience, and versatility in a variety of fields. As technology advances, it becomes essential to understand the certification of these systems.

Certification of behavioral biometrics systems is a crucial aspect in the field of biometric technology, as it guarantees the accuracy and reliability of these systems for various applications. In this comprehensive introduction, we look at the background, motivation, assessment methods, and standards associated with the certification of behavioral biometrics, in order to highlight the importance of this field.

Assessing the performance of biometric systems is a fundamental step in certification. We look at the general methods used to assess the effectiveness of behavioral biometric systems. This includes assessing factors affecting system performance, measuring the quality of behavioral biometric models, and exploring techniques for generating logical attack bases.

The FIDO alliance, a leading organization in the field of authentication, has defined certification criteria for biometric systems. We also examine FIDO's biometric certification standards, giving an overview of their performance measurement and presentation of attack detection criteria. Understanding these standards is essential, as they influence the certification of behavioral biometric systems.

Another important standard in the field of behavioral biometrics is ISO/IEC 39794-17. We do examine the foreword, description, gait recognition data models, and gait recognition data flow as described in this standard. Understanding this standard is essential for aligning behavioral biometrics systems with global benchmarks.

The aim of this global study is threefold. Firstly, we seek to assess the various factors affecting the performance of behavioral biometrics systems. Secondly, we aim to establish methods for measuring the quality of behavioral biometric models. Finally, we aim to explore the generation of logical attack bases to improve the security of behavioral biometric systems. These objectives form the core of our research and guide our exploration throughout this study.

The objective of this thesis is to evaluate the factors influencing the performance of behavioral systems, define the quality of behavioral biometric models, and generate logical attack bases for these systems. These elements contribute to the overall understanding of the evaluation of behavioral biometric systems and their secure use in various domains.

In the field of biometrics, multimodal systems that combine several biometric features for authentication are gaining importance. We discuss decision-level fusion, score-level fusion, feature-level fusion, and sample-level fusion, all of which are crucial aspects of testing multimodal biometric implementations. Understanding these fusion techniques is essential for achieving robust and reliable biometric systems.

This introductory section provides an overview of the essential topics that are covered in our comprehensive study on the certification of behavioral biometric systems. We take a closer look at each of these areas, shedding light on the intricacies and significance of behavioral biometrics in the modern world of authentication and identification.

We take a closer look at each of these areas, highlighting the subtleties and importance of behavioral biometrics in the modern world of authentication and identification.

This work is a CIFRE [1] Ph.D. thesis (joint collaboration with a company) between the Biometric team at Fime SAS and the SAFE [2] team of the GREYC laboratory at ENSICAEN in Normandy University. Fime SAS is a company of Card payment and Biometrics (Facial, Fingerprint, Iris, Voice, Palm Vein systems) tests for certifications in France. It is one of the world's leading testing products for a wide range of customers and technologies with a gained know-how from over 20 years of testing products. The GREYC is a computer science research laboratory dedicated to modeling, methodological research, and practical application in the digital sciences. GREYC is renowned for its original contributions, hardware and software developments, validated experiments, as well as its interdisciplinary collaborations in the fields of human and social sciences, and the interaction between computer science, mathematics, and engineering sciences. It is also recognized for its concrete achievements thanks to solid academic and industrial partnerships.

Thesis outline

In Chapter 1, we lay the foundation by looking at the essential background of behavioral biometrics in the context of certification. We begin with an introduction to set the stage for our exploration, followed by an examination of the motivation driving research and development in this field. We look at the intricacies of biometric technology, including enrolment, verification, and identification, as well as the architectural aspects of biometric

[1] CIFRE stands for Convention Industrielles de Formation par la REcherche, i.e. Industrial Agreement of Training through Research. The research undertaken by a CIFRE fellow is within the framework of a public and private partnership between a French company and a laboratory and is formulated by both parties.

[2] Normandie Université, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

systems. We also discuss the evaluation of biometric systems, outlining the general methods used to assess their effectiveness. Furthermore, we explore the standards and criteria set by organizations such as FIDO Alliance and ISO/IEC $39794 - 17$, shedding light on their roles in the certification of Behavioral Biometrics.

Chapter 2 embarks on a comprehensive description of behavioral biometrics. First, we trace the evolution of biometric solutions, highlighting the transition to more subtle, behavior-based methods. We look at the objectives and issues involved in collecting behavioral biometric data, highlighting the challenges in terms of privacy and regulatory compliance. This analysis is complemented by an in-depth exploration of the main modalities of behavioral biometrics, its benefits, and a literature review. We also discuss the collection and analysis of behavioral biometric data, looking at collection methods, the types of sensors used and the precautions needed to guarantee data quality and security. Finally, we explore current applications and emerging trends in behavioral biometrics, examining its role in security and health, while identifying the potential limitations and risks associated with its use. We conclude this chapter by highlighting recent developments that are shaping this constantly evolving field.

Chapter 3 explores the transactional applications of behavioral biometrics with a focus on identification and authentication. We provide an overview of this segment, setting the stage for our in-depth exploration. It delivers the intricacies of related works in areas like time series analysis, human operations, and keystroke dynamics. A proposed architecture is presented, including the features generation, identification, authentication, and matching algorithms. We elucidate the experimental protocol, detailing the datasets used, performance metrics, pre-trained models, and classifier parameters. The chapter then looks at user identification, touching on classical machine learning and deep learning techniques, and engaging in discussions around these methodologies. User activity authentication is studied, with a review of performance expectations, single-activity performance, and multi-activity scenarios. Finally, user key authentication is discussed, with an analysis of performance on different datasets and scenarios.

Chapter 4 addresses the innovative field of synthetic behavioral biometric data generation using generative adverbial networks (GANs). We provide an overview of this cutting-edge

approach, followed by an introduction that contextualizes the importance of synthetic data generation. We review related work in the field and delve deeper into the principles and architecture of generative adversarial networks (GANs). We also explore TimeGAN, a specialized methodology for synthetic signal generation, including its learning process and techniques for improving signal generation. An important section is devoted to metrics for evaluating the models used to generate synthetic signals, including the performance evaluation of TimeGAN. Applications of synthetic signal generation are discussed, and the chapter ends with a stimulating discussion and a general conclusion.

Finally, the conclusion and ways of improving this exploratory work are discussed.

CHAPTER 1

Background: Certification of Behavioral Biometrics Systems

Summary

The introductory chapter sets the context for the certification of behavioral biometrics systems. It begins with an introduction to our exploration, followed by a review of the underlying motivations for research in this area. We then explore key aspects of biometric technology, including enrolment, verification, identification, and the architecture of biometric systems. The evaluation of biometric systems is also covered, as well as the standards set by organizations such as FIDO Alliance and ISO/IEC 39794-17 for certification in behavioral biometrics.

Keywords: Biometric technology; Enrolment; Verification; Identification; FIDO Alliance; Certification.

1.1 Introduction

The development of behavioral biometrics, which relies on user behavioral characteristics for authentication and identification, is gaining interest in the cybersecurity field. However, the evaluation and certification of these systems remain a major challenge for their widespread adoption. In this chapter, we explore the context for the certification of behavioral biometric systems. We examine the motivations and criteria for evaluating biometrics systems, as well as the standards established by organizations such as FIDO and ISO/IEC. We also discuss the objectives of certification and the methods used to evaluate the performance of behavioral biometrics models. Finally, we discuss the basics of logical attacks for behavioral biometrics systems.

1.2 Motivation and biometric technology

1.2.1 Motivation

Biometrics is today an indispensable tool for identification and authentication. It can be found in smartphones, in access to secure sites, in public identification systems in certain countries, etc.

Etymologically, the word biometrics means the study of living things. A biometric character is a unique and universal character. The CNIL [1], known as the French National Commission for Information Technology and Freedoms (primarily focuses on fostering the ongoing growth of new technologies and actively contributes to the establishment of digital ethics) defines biometrics as the set of computer techniques that allow the automatic recognition of an individual based on his or her characteristics from different modalities: physical, biological, or behavioral. Biometric characteristics are therefore traits that can be found in any individual but which are different from one person to another. The study of human behavior has shown that certain characteristics are specific to each individual, and can be used for authentication [Li et al., 2021]. Today, there are many solutions

[1] https://www.cnil.fr/en

such as keyboard typing dynamics [Migdal, 2019a, Ayotte et al., 2021b], mouse movement dynamics [Moskovitch et al., 2009, Monaro et al., 2020], accelerometer and position sensors on smartphones, and movement sensors on the screen of smartphones [Rayani and Changder, 2023], etc.

Extensive scientific research has been undertaken to establish methodologies for evaluating biometric systems and attack methods for morphological modalities such as fingerprint, face, or iris [Marcel et al., 2019, Tolosana et al., 2019]. Organizations such as the FIDO Alliance have been working also to set requirements for the evaluation of performance and the presentation of attack detection. Few works have been done on generating repeatable attack systems for behavioral biometrics as well as evaluating its performance and factors influencing the quality of behavioral signals. However, these biometric modalities achieve worse performance than morphological ones and are often easy to attack. As an example, an attack on a system based on the recognition of keystroke dynamics can be summed up by the impostor entering text on a keyboard, as well as capturing the typing dynamics.

This Ph.D. thesis proposes to make three scientific contributions for the evaluation of behavioral biometrics systems. The biometrics modalities concerned are the keystroke dynamics typing [Giot et al., 2009b], and gait from mobile signals [Gafurov, 2007] on human activity.

1.2.2 Biometric technology

1.2.2.1 Enrollment, verification, and identification

Biometric systems operate in three modes: enrollment, verification/authentication, and identification:

- **Enrollment**

  Enrollment is the first phase of any biometric system. It is the stage at which a user is registered in the system for the first time. It is common for both verification and identification. During enrollment, the biometric characteristic is measured using a biometric sensor to extract a digital representation. This representation is then

reduced, using a well-defined extraction algorithm, to reduce the amount of data to be stored, thus facilitating verification and identification. Depending on the application and the level of security required, the biometric template selected is protected (such as encrypted) and stored either in a central database or on a personal element specific to each individual;

- **Verification**

  Verification of identity involves checking whether the individual using the system is who he or she claims to be. The system compares the biometric information acquired with the corresponding biometric reference template stored in the database, known as one versus one (1 vs. 1). In this case, the system returns only a binary decision (yes or no), which can be weighted. The verification process can be formalized as follows: given the input vector C_U defining the biometric characteristics of user U extracted by the system, and M_U its biometric model stored in the database, the system returns a Boolean value following the computation of the function f defined by:

$$f(C_U, M_U) = \begin{cases} 1 & \text{if} \quad S(C_U, M_U) \geq \theta \\ 0 & \text{elseif} \end{cases} \tag{1.1}$$

  where S is the similarity function defining the correspondence between the two biometric vectors, and θ is the decision threshold at which the two vectors are considered identical;

- **Identification**

  In identification mode, the biometric system determines the user of an unknown individual from a database of identities, known as a one-to-all (1 vs N) test. In this case, the system can either assign the unknown individual the user corresponding to the closest profile found in the database (or a list of close profiles) or reject the individual. The identification process can be formalized as follows: Given an input vector C_U defining the biometric characteristics extracted by the system when a user U presents himself to it, identification amounts to determining the identity of I_t, $t \in \{0, 1, \ldots, N\}$ where $I_1, \ldots, I_N$ are the identities of users previously enrolled

in the system, and I_0 indicates an unknown user. The identification function f can thus be defined by:

$$f(C_U) = \begin{cases} I_k & \text{if} \quad \max_{1 \leq k \leq N} S(C_U, M_k) \geq \theta \\ I_0 & \text{elseif} \end{cases} \tag{1.2}$$

where M_k is the biometric model corresponding to identity I_k, S is the similarity function, and θ is the decision threshold.

1.2.2.2 Architecture of a biometric system

The architecture of a biometric system contains five modules, as shown in Figure 1.1:

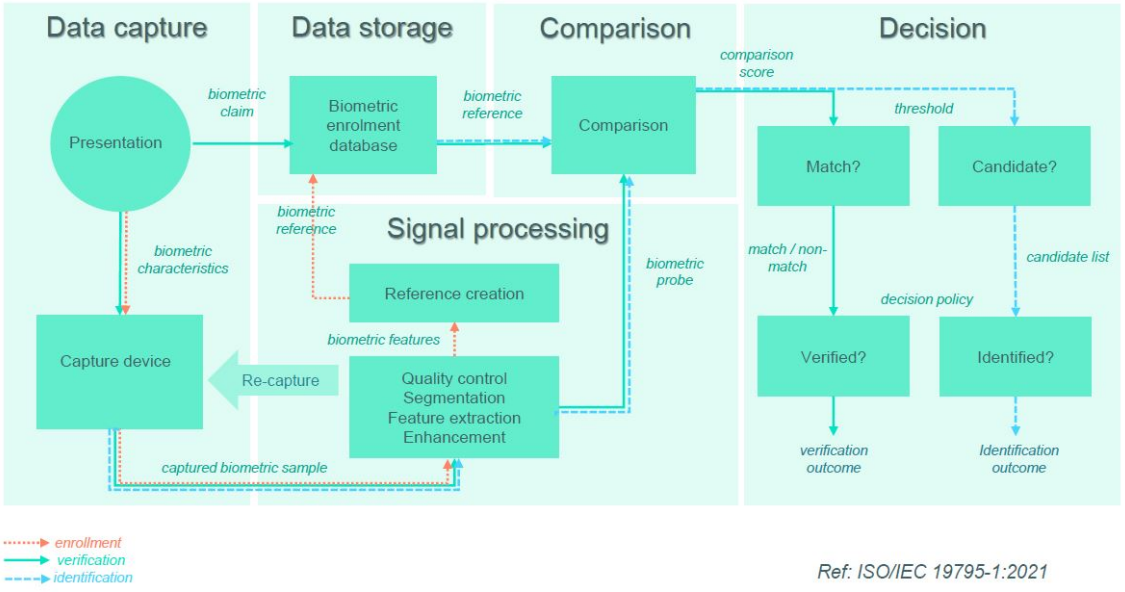


FIGURE 1.1: Generic architecture of a biometric system (extract from International Organization for Standardization ISO/IEC 19795-1) [ISO 19795-1, 2021].

1. The capture module acquires biometric data and extracts a digital representation. This representation is then used for enrolment, verification, or identification. The biometric sensor can be contact or contactless;

2. The signal processing module, reduces the extracted digital representation in order to optimize the amount of data to be stored during the enrolment phase or to facilitate

processing time during the verification and identification phase. This module can have a quality test to check the biometric data acquired;

3. The storage module containing the biometric templates of enrolled system users;
4. The similarity module, which compares the biometric data extracted by the feature extraction module with one or more previously stored templates. This module determines the degree of similarity (or divergence) between two biometric vectors;
5. The decision module determines whether the returned similarity index is sufficient to determine the identity of an individual.

1.3 Evaluation of biometric systems: general method

The aim of biometric system evaluation is to reduce limitations (performance, cultural limitations, vulnerability to specific attacks). The evaluation of these systems is generally carried out according to three evaluation aspects, as shown in Figure 4.9.

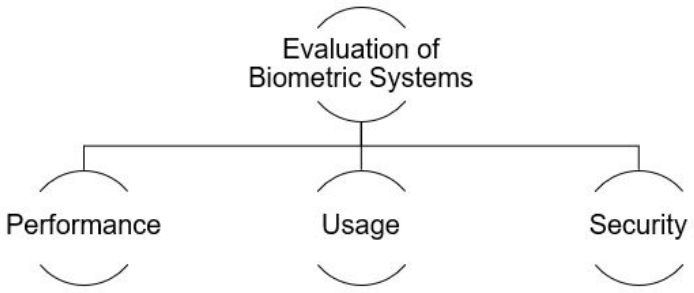


FIGURE 1.2: Evaluation aspects of biometric systems.

In order to compare various biometric systems, whether they belong to the same or different modalities, it is essential to evaluate them. Many studies have already addressed the evaluation of biometric systems [Conklin et al., 2004, Theofanos et al., 2008, Busch, 2023], and the purpose of this section is to present the commonly used methodologies. When evaluating a biometric system, there are three main aspects to consider:

1. **Performance** [ISO 19795-1, 2021]. It aims to quantify various statistical measures regarding system performance, such as Equal Error Rate (EER), False Acceptance

Rate (FAR), False Reject Rate (FRR), Failure To Enroll (FTE), Failure to Acquire (FTA), and ROC curves [ISO 19795-1, 2021, Busch, 2023]).

2. **Acceptability**, which measures the acceptability and satisfaction of users when using biometric systems [Theofanos et al., 2008]. It is more about the perception and adherence of individuals, providing information on these aspects rather than on the error performance of the system.

3. **Security**, which measures the robustness of a biometric system (sensor and algorithms) against fraud ISO [19792, 2008]. It assesses the safety level of the system by measuring the potential number of frauds that an impostor or concealer can commit.

Where:

- FAR (False Acceptance Rate) represents the percentage of impostors wrongly accepted by the system.
- FRR (False Rejection Rate) represents the percentage of users wrongly rejected.
- EER (Equal Error Rate) represents the error rate corresponding to a setting of the biometric system's decision threshold so that the FAR value is equal to FRR.
- The ROC curve is used to represent the efficiency of a biometric system. It represents the evolution of the FAR as a function of the FRR.

In an ideal world, a perfect system has an $EER = 0$. In practice, this is nearly impossible since it is complicated to get FAR and FRR close to zero given the intrinsic variability of biometric data capture. So low FRR and low FAR equal to a better security.

The evaluation of biometric systems is a major issue in biometrics for several reasons. Firstly, it provides researchers and developers with a tool to better test and evaluate their systems with those that exist in the state-of-the-art. Secondly, it allows users' behavior to be taken into account during the evaluation process, enabling us to better understand their needs and better deploy this technology in our daily lives. Finally, it allows us to identify industrial applications for each system, based on various criteria such as performance, usage, security, and the cost of deploying the technology [El-Abed, 2011].

The ISO/IEC 19795-1 standard [ISO 19795-1, 2021] covers the evaluation of biometric systems in terms of performance. The standard proposes statistical measures designed to quantify and evaluate biometric systems.

All these evaluation approaches have to be taken into account when comparing various biometrics systems. It seems strange to consider a system to be good if it has very low error rates (i.e. very good performance) while having a very low user acceptance (i.e. a high probability of being unused). Comparing biometric systems can be realized within three types of performance testing [ISO 19795-1, 2021]:

1. **Technology evaluation**: involves testing only one component of the biometric system, such as a matching or feature extraction algorithm.
2. **Scenario-based evaluation**: involves testing the whole system.
3. **Operational evaluation**. Similar to scenario-based evaluation, but the system is integrated into a real application, with real end-users.

In preparation for a capture session, it is imperative to consider and integrate several essential steps. These include the administrative management of subjects, including verification/authentication, processing of agreement forms, and compensation distribution. In addition, careful management of biometric data is essential, including tasks such as secure copying and storage of data, as well as the implementation of rigorous quality controls. In addition, the optimal configuration of capture conditions, including elements such as adequate lighting arrangements and the installation of appropriate supports, is crucial to the success of the session. Typically, this total duration does not exceed 30 minutes for one single test subject for each standard. Figure 1.3 draws industrial standards for biometric solutions according to their factors.

In the next formal section, we are more focused on FIDO Alliance standards as an example of biometric certification content.

1.4 What FIDO says about biometric certification?

The FIDO (Fast IDentity Online) alliance is an organization dedicated to creating authentication standards that reduce reliance on passwords. The FIDO alliance is a benchmark for the development, use and enforcement of authentication standards in a security-focused industry.

1.4.1 FIDO Biometric Certification

FIDO was founded in July 2012 and has been publicly active since 2013. Currently, it brings together participants from several well-known Internet companies. Its main objective is to establish industry standards for the use of authenticators in the web application authentication process, either as a single factor (no password) or as an additional factor such as Two-Factor Authentication / Multi-Factor Authentication (2FA/MFA) [2].

FIDO biometric evaluation is applicable to a variety of biometric technologies, such as facial recognition, fingerprints, iris, and voice. During FIDO biometric certification, two fundamental aspects are evaluated: performance measurement and detection of presentation attacks. Both are analyzed during a test session that encompasses several processes.

[2] `https://fidoalliance.org/certification/biometric-component-certification/`

Standards	Modality	Form factor
ISO 19795-1 to -10	Depending on the part of the specification	Nonspecific
ISO 30107-3	Nonspecific	Nonspecific
ISO 19989-1 to -3	Nonspecific	Nonspecific
NIST	Fingerprint, face, iris, voice, DNA, and multimodal	Component, system, device
FIDO	Fingerprint, face, iris, voice and others	Component, system, device
Mastercard	Fingerprint, face	Biometric card, sensor, device
Visa	Fingerprint	Biometric card
Android™	Fingerprint, face, voice, iris, and others	Component, sensor, system, device
Windows Hello	Fingerprint, face, iris (mobile)	Component, sensor, device

FIGURE 1.3: Industrial standards for biometric solutions.
NIST National Institute of Standards and Technology.

Full completion of these processes is essential to accurately assign FIDO biometric evaluation status to the product.

The FIDO Biometric Certification Program uses False Reject Rate (FRR), False Accept Rate (FAR), Failure-to-Enrol (FTE), and Failure-to-Acquire (FTA) to measure Biometric Performance.

Following the definitions from ISO/IEC 2382-37 [ISO 2382-37, 2022] and provided in *Biometric Data* and *Evaluation Terms*, a *verification attempt* results in a biometric comparison while a *verification transaction* results in a resolution of the biometric claim (accept or reject). The ISO definition for real-time acquisition steps for **enrolment**, **verification** or **identification** transactions involve a series of capture trials, in accordance with the corresponding decision policy. Each capture trial may include one or more presentations, depending on the sensor operating mode, sample quality criteria and any constraints governing the number of presentations or duration allowed per trial [ISO 19795-1, 2021].

The FAR, FRR, FTA, and FTE according to the FIDO alliance are defined in terms of verification operations as follows:

1. **FAR** is computed through offline testing based on biometric references and stored verification transactions collected during online testing.
   FAR(%) = (Number of zero-effort imposter transactions for which decision is Accepted) / (Number of zero-effort imposter transactions conducted) * 100

2. **FRR** shall be estimated by the equation given in standard [ISO 19795-1, 2021]. The computation of FRR shall be based on:
   FRR(%) = (Number of mated transactions for which decision is rejected or FTA happens for all attempts) / (Number of mated transactions conducted) * 100

3. **FTA**, Failure-to-Aquire rate is the proportion of verification or identification attempts for which the system fails to capture or locate an image or signal of sufficient quality [ISO 19795-1, 2021].

4. **FTE**, Failure-to-Enrol rate is the proportion of the population for whom the system fails to complete the enrolment process [ISO 19795-1, 2021].

1.4.2 General overview of FIDO certification criteria

Standard $30107-3:2017$ defines a metric for the Presentation Attack Detection (PAD) called Impostor Attack Presentation Match Rate (**IAPMR**). A correction has been made to this term in [ISO 30107-3, 2023], which changes the name to "Impostor Attack Presentation Accept Rate" (**IAPAR**) such that it is consistent with biometric performance metrics. The IAPAR is defined as:

IAPAR(%) = (Number of Impostor Presentation Attack Transactions for which Decision is Accept) / (Total Number of Impostor Presentation Attack Transactions Conducted) * 100

The IAPAR shall be computed for each instrument's attacks (PAI Species).

TABLE 1.1: Biometric requirements by levels [Schuckers et al., 2023]. **M**: Mandatory at; **O**: Optional at.

	BioLevel 1	**BioLevel 1+**	**BioLevel 2**	**BioLevel 2+**
#Subjects for FAR/FRR	25	245	25	245
#Subjects for PAD	15	15	15	15
Lab Tested FAR	1%	.01%	1%	.01%
Lab Tested FRR	7%	5%	7%	5%
Lab Tested IAPAR	15%	15%	7%	7%
#Species A/B	6/8	6/8	6/8	6/8
#IAPAR Subjects	15	15	15	15
Self Attestation FAR	**M** $<= 1/10k$	**O** $<= 1/10k$	**M** $<= 1/10k$	**O** $<= 1/10k$
Self Attestation FRR	**M** $<= 5\%$	**O** $<= 5\%$	**M** $<= 5\%$	**O** $<= 5\%$

The requirements for certification of FIDO biometric components are listed in Table 4.4. Unless noted as **Optional** (O), all these requirements are necessary for certification. There are two levels of certification which have different thresholds for **IAPAR** metric for PAD assessment. Otherwise, the testing procedure is the same for both levels.

1.4.2.1 Performance measurement

The aim of the performance measurement test is to evaluate the biometric performance score of a **Target of Evaluation** (TOE), i.e. the product under test, by comparing its results with a database containing biometric samples from real individuals.

This test leads to an evaluation based on the FRR, i.e. the percentage of verification transactions where genuine identity declarations are unfairly rejected, and the FAR, which represents the expected proportion of inauthentic (zero-effort non-genuine transactions) transactions incorrectly accepted. These rates are computed using a bootstrap method.

1.4.2.2 Presentation Attack Detection

The aim of the Presentation Attack Detection (PAD) test is to evaluate the reaction of a biometric security product to various presentation attack instruments (PAI), also known as spoofs. This test results in a verdict based on the IAPAR, i.e. the proportion of impostor attack presentations using the same types of PAI that are accepted.

The criteria to be met are as follows:

- $\text{IAPAR}_{TOE,PAIx} < 7\%$ (maximum of **10 errors out of 150** attempts per PAI species)
  In this case, TOE is considered as meeting **BioLevel** 2 or **BioLevel** 2+ PAD requirements, or
- $\text{IAPAR}_{TOE,PAIx} < 15\%$ (maximum of **22 errors out of 150** attempts per PAI species)
  In this case, TOE is considered as meeting **BioLevel** 1 or **BioLevel** 1+ PAD requirements

1.5 What does standard ISO/IEC 39794-17 say about behavioral biometrics ?

1.5.1 Foreword

The ISO (International Organization for Standardization) and the IEC (International Electrotechnical Commission) form the specialized system of worldwide standardization. They define criteria for both physiological and behavioral biometrics. For example, the ISO/IEC

39794-17 [ISO 39794-17, 2021] specification addresses the topic of behavioral biometrics, specifically gait. Figure 1.4 shows the components of a human activity image sequence biometric system as defined in the standard.

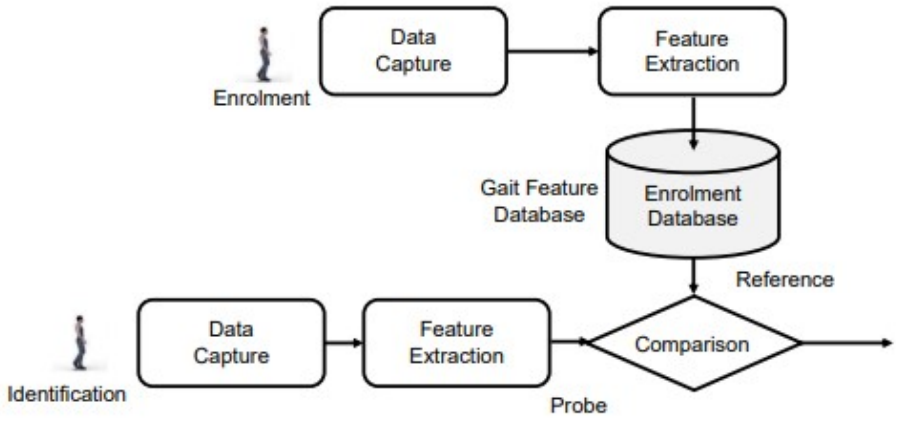


FIGURE 1.4: standards ISO/IEC 39794 – 17 components of a gait image sequence biometric system [ISO 39794-17, 2021].

1.5.2 Description

Many countries worldwide use biometric recognition systems for law enforcement and border control. Many of these systems are not limited to face recognition purposes. Different groups are working on technical documents, guidelines, and best practice recommendations to ensure consistency in these deployments and processes. However, these documents focus primarily on travel documents and border control systems, as well as technical and operational issues related to the planning and deployment of these systems. Gait recognition is used as a secondary biometric mode, in addition to whole-body biometric recognition or for forensic purposes.

Standards ISO/IEC 39794-17 reported little guidance regarding gait imagery for cross-border interoperability or law enforcement services. Thus, there is a need for guidance on the use of high-quality digital cameras and video surveillance devices to record gait image sequence data. This standard is not limited to whole-body gait image sequence data. For example, it may be possible to extract only head movement data for recognition.

Currently, border guards use local practices for biometric enrolment, verification, and identification of gait in videos.

This part of ISO/IEC 39794 is intended to provide advice on the use of body image data for gait and upper body movement recognition applications requiring an exchange of gait image sequence data and upper body movement data. Typical applications are:

- automated body biometric verification and identification (one-to-one as well as one-to-many comparison),
- support for human biometric verification by comparison of persons based on video and still gait images,
- Support for human examination of video and still gait images with sufficient resolution to allow a human examiner to perform biometric verification.

1.5.2.1 Data models for gait recognition

Gait recognition systems can be grouped into three categories based on the sensors used namely: 1) motion imaging (vision) based, 2) wearable sensor-based, and 3) spatial (floor) sensor-based. Motion imaging can itself be subdivided into two groups: a) appearance-based methods and b) model-based methods. Appearance-based methods can be classified into two types: state-space-based methods and spatiotemporal methods [Ali et al., 2011]. Figure 1.5 illustrates the scope of standards ISO/IEC 39794-17 marked with green background shading on the figure.

1.5.2.2 Data flow of gait recognition

For comparison methods, conventional feature-based model sets can be used, or feature vectors from deep convolutional neural networks (DCNNs). Once the feature vectors are generated through gait signatures and DCNN processing, the comparison is based on one of several classification algorithms commonly used in machine learning, such as the Bayesian classifier or the Euclidean classifier. In the following section, we present the different

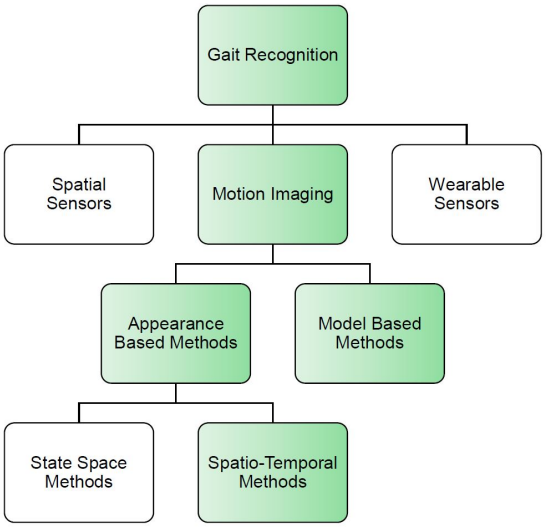


FIGURE 1.5: Classification of gait recognition systems. The scope of standards ISO/IEC 39794-17 is marked with green background shading on the figure above.

types of fusion defined by ISO, providing details on technology assessment and scenario evaluation.

1.6 Testing of multimodal biometric implementations

This section sets out procedures for evaluating and reporting the performance of multimodal biometric algorithms and systems. In accordance with standard ISO/IEC 19795-2:2007 [ISO 19795-2, 2015], multimodal biometric implementations can be used to accomplish the following purposes:

- to support users who cannot present one or more requested modalities to the system, in other words, to improve the failure-to-enroll rate;
- to improve biometric system flow rate;
- to improve recognition performance (e.g. through reduction of false negative identification rates);
- to improve usability and;
- to increase robustness against presentation attacks.

The standard ISO/IEC TR 24722 [ISO 24722, 2015] defines the following multimodal fusion levels:

- decision-level;
- score-level;
- feature-level;
- sample-level.

Multimodal fusion approaches vary at each level. Even if multimodal data are obtained using identical sensors, the results may vary according to the degree of fusion adopted. Hence the need for the experimenter to define the system or application to be evaluated. The evaluation must precisely specify the level of fusion involved, the constituent elements of multimodal fusion, and the criteria required for evaluations at each stage of fusion.

1.6.1 Decision-level fusion

An example of decision-level fusion is shown in Figure 1.6. Decision-level fusion systems combine decision results from separate biometric sub-systems [ISO 24722, 2015].

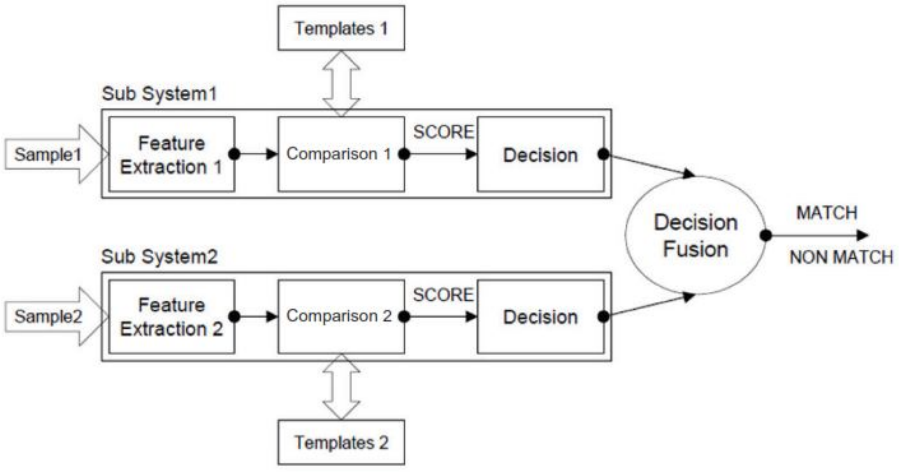


FIGURE 1.6: Fusion on the decision level [ISO 19795-2, 2015].

Requirements for repeatability of decision-level fusion technology evaluation results are as follows:

- The decision fusion logic shall be identical;
- The function configurations (i.e. feature extraction, comparison, and decision) of Sub-System 1 and Sub-System 2, respectively, shall remain consistent across all tests;

Sub-System 1 and Sub-System 2 can have different function configurations and user-specific thresholds can differ for different users such as:

- The combination of Sample 1 and Sample 2 fed into each feature extraction function shall be identical;
- The combination of Template 1 and Template 2 shall be identical.

Consistent data selection methods for samples and templates are also required for evaluation repeatability. If Sub-System 1 and Sub-System 2 are independent and separate, the evaluation report should include the following:

- Identifying information for Sub-System 1 and Sub-System 2;
- Identifying information for decision fusion logic;
- Fusion level.

Requirements for repeatability of decision-level fusion evaluation results are as follows the function configurations (i.e. capture, feature extraction, comparison, and decision) of Sub-System 1 and Sub-System 2, respectively, shall remain consistent across all tests;

Sub-System 1 and Sub-System 2 can have different function configurations and user-specific thresholds may differ for different users such as:

- The decision fusion logic shall be identical;
- The combination of Sample 1 and Sample 2 fed into each feature extraction function shall be based on the same subject and position (e.g. right iris);

- The combination of Template 1 and Template 2 shall be based on the same subject and position.

Consistent data selection methods for samples and templates are also required for evaluation repeatability. If Sub-System 1 and Sub-System 2 are independent and separate, the evaluation report should include the following:

- Identifying information for Sub-System 1 and Sub-System 2;
- Identifying information for decision fusion function;
- Fusion level.

1.6.2 Score-level fusion

Fusion on the score level is illustrated in Figure 1.7. Score-level fusion systems utilize score results from separate biometric subsystems [ISO 24722, 2015].

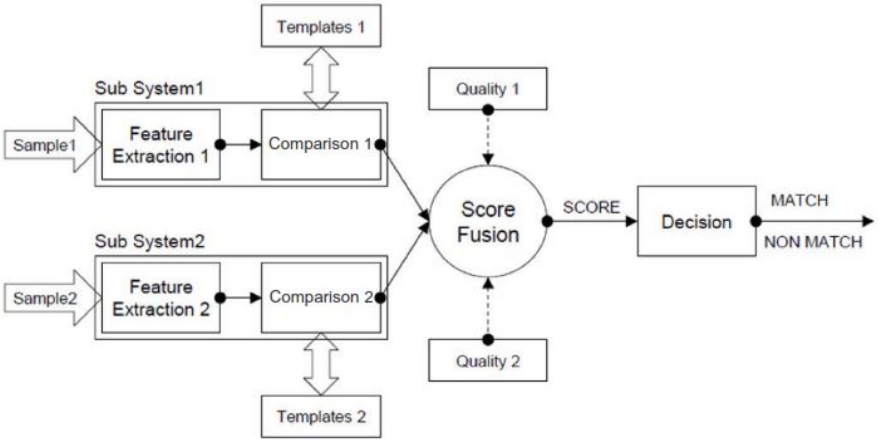


FIGURE 1.7: Fusion on the score level [ISO 19795-2, 2015].

Score-level fusion may use sample quality in scenario or technology evaluations. Score-level fusion systems might be used to improve the false match rate (FMR) and false non-match rate (FNMR).

Requirements for repeatability of score-level fusion technology evaluation results are as follows:

- the score fusion function and decision function shall be identical;
- the function configurations (i.e. feature extraction and comparison) of Sub-System 1 and Sub-System 2, respectively, shall remain consistent across all tests;

Sub-System 1 and Sub-System 2 can have different function configurations, and user-specific thresholds might differ for different users.

- the combination of Template 1 and Template 2 shall be identical;
- the combination of Sample 1 and Sample 2 fed into each feature extraction function shall be identical.

Requirements will be necessary for the data selection method for samples and templates, in order to keep repeatability. If Sub-System 1 and Sub-System 2 are independent and separate, the evaluation report should include the following:

- identifying information for Sub-System 1 and Sub-System 2;
- identifying information for score fusion function and decision function;
- fusion level.

Requirements for repeatability of score-level fusion scenario evaluation results can be stated as follows the function configurations (i.e. capture, feature extraction, and comparison) of Sub-System 1 and Sub-System 2, respectively, shall remain consistent across all tests. Sub-System 1 and Sub-System 2 can have different function configurations and user-specific thresholds can differ for different users such as:

- The score fusion function and decision function shall be identical;
- The combination of Template 1 and Template 2 shall be based on the same subject and position;
- The combination of Sample 1 and Sample 2 fed into each feature extraction function shall be based on the same subject and position.

Consistent data selection methods for samples and templates are also required for evaluation repeatability. If Sub-System 1 and Sub-System 2 are independent and separate, the evaluation report should include the following:

- Identifying information for Sub-System 1 and Sub-System 2;
- Identifying information for score fusion function and decision function;
- Fusion level.

1.6.3 Feature-level fusion

Fusion on the feature level is illustrated in Figure 1.8. Feature-level fusion systems utilize results from separate feature extraction components [ISO 24722, 2015].

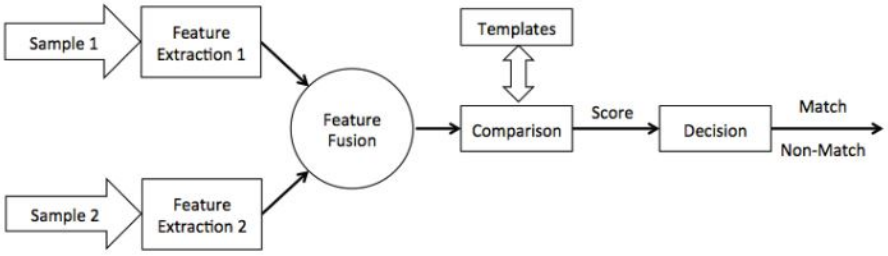


FIGURE 1.8: Feature-level fusion [ISO 19795-2, 2015].

Requirements for repeatability of feature-level fusion technology evaluation results are as follows:

- Feature fusion function, comparison function, and decision function shall be identical;
- The function configurations (i.e. feature extraction) of Feature Extraction 1 and Feature Extraction 2, respectively, shall remain consistent across all tests;

Feature Extraction 1 and Feature Extraction 2 can have different function configurations and user-specific thresholds can differ for different users as follows:

- The combination of Sample 1 and Sample 2 fed into each feature extraction function shall be identical;

- The combination of Sample 1 and Sample 2 at the time of template creation shall be identical.

Consistent data selection methods for samples and templates are also required for evaluation repeatability. If Feature Extraction 1 and Feature Extraction 2 are independent and separate, the evaluation report should include the following:

- Identifying information for Feature Extraction 1 and Feature Extraction 2;
- Identifying information for the feature fusion function, comparison function, and decision function;
- Fusion level.

Requirements for repeatability of feature-level fusion scenario evaluation results are as follows:

- Feature fusion function, comparison function, and decision function shall be identical;
- The function configurations (i.e. capture and feature extraction) of Feature Extraction 1 and Feature Extraction 2, respectively, shall remain consistent across all tests;

Feature extraction 1 and Feature extraction 2 can have different function configurations and user-specific thresholds can differ for different users.

- The combination of Sample 1 and Sample 2 fed into each feature extraction function shall be based on the same subject and position;
- The fused sample template production process shall be identical. Consistent data selection methods for samples and templates are also required for evaluation repeatability.

If Feature Extraction 1 and Feature Extraction 2 are independent and separate, the evaluation report should include the following:

- Identifying information for Feature Extraction 1 and Feature Extraction 2;
- Identifying information for the feature fusion function, comparison function, and decision function;
- Fusion level.

1.6.4 Sample-level fusion

Fusion on the sample level is illustrated in Figure 1.9. Sample-level fusion systems utilize samples from separate capture systems [ISO 24722, 2015].

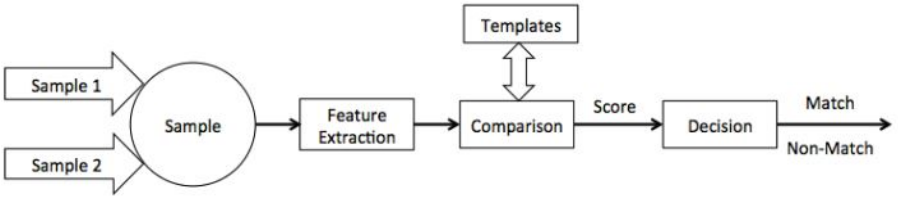


FIGURE 1.9: Sample-level fusion [ISO 19795-2, 2015].

To ensure reproducible results when evaluating fusion technologies at the sample level, the following criteria must be met:

- Sample fusion, feature extraction, comparison, and decision functions must remain identical,
- The combinations of samples 1 and 2, applied to each sample fusion function, must remain constant,
- The combinations of samples 1 and 2 used during model creation must be the same.

It is also necessary to adopt consistent methods for the selection of data used for samples and models, in order to guarantee reproducibility in technology assessment. For the evaluation of fusion scenarios at the sample level, the following conditions must be met to ensure the reproducibility of results:

- The fusion sample collection process must remain identical,

- Sample fusion, feature extraction, comparison, and decision functions must remain unchanged,
- Function configurations (including capture) must remain consistent across all tests.

As with technology evaluation, the use of consistent methods for sample and model data selection is essential to ensure reproducible results when evaluating scenarios.

In resume, maintaining consistency of function and process while adopting consistent data selection methods is imperative to ensure the reproducibility of assessments, whether for fusion technologies or fusion scenarios.

1.7 Ph.D. thesis Objectives

This Ph.D. thesis proposes to make several scientific contributions to the evaluation of behavioral biometric systems. The biometric modalities concerned are the keystroke dynamics [Giot et al., 2009b] and human activities (e.g. *sitting*, *sleeping*, *walking*, *gait*...) from mobile signals [Gafurov, 2007]. The aim is to propose a generic method for analyzing behavioral biometrics. Furthermore, considering the most popular temporal adversary generators, the other objective is to create synthetic behavioral biometrics, which could be used to impersonate an authorized user in a certification scenario.

The central aim of this Ph.D. thesis is to examine the parameters impacting the performance of behavioral systems, to define the quality of behavioral biometric models, and to generate databases for logical attacks against these systems. These aspects are crucial for a complete understanding of the evaluation of behavioral biometric systems, and for ensuring their secure use in various application domains.

1.7.1 Factors affecting the performance of behavioral systems

These factors include the method of assessing the quality of the bio-signals. Few research has been conducted in this area. Another factor is the variation in human behavior and stability over time for authentication. How does the evolution of behavioral signals over

time impact the performance of systems? To certify a behavioral biometrics solution to secure a payment, for example, it is necessary to know these factors to be able to test it.

1.7.2 Quality measurement of behavioral biometric templates

The nature of behavioral biometric data is by definition less stable over time than morphological data. In order to guarantee a good performance of biometric systems, it is necessary to define quality metrics for biometric data to optimize user enrollment. For morphological data, there are standardized metrics, notably for fingerprints [Bausinger and Tabassi, 2011], but few works have concerned behavioral data. The definition of such a metric for behavioral data (probably based on several samples) is essential to improve the performance of such systems.

1.7.3 Synthetic generation for behavioral biometric systems

Evaluating the resistance of behavioral systems to attack is a major concern. The issue of cost and evaluation time is paramount. In the last 5 years, a lot of work has been done on Presentation Attack Detection (PAD) on modalities like fingerprint, face, and iris, including the use of convolutional neural networks [Pérez-Cabo et al., 2019, Engelsma and Jain, 2019]. Few works [Khan et al., 2020] have considered the problem of behavioral bio-signals in attack instrument presentations. Work done at GREYC [Migdal and Rosenberger, 2019] on generating (statistical modeling) synthetic data for keystroke dynamics can be generalized to other behavioral modalities and contexts (including presentation attacks).

Behavioral biometrics can be extremely useful when merging biometrics.

1.8 Conclusion

Certification of behavioral biometric systems is a crucial aspect of ensuring the reliability and security of these systems. In this chapter, we examined the motivations and criteria

for evaluating biometric systems, highlighting the importance of thorough evaluation to ensure their effectiveness.

We also examined FIDO's recommendations for biometric certification, highlighting the organization commitment to establishing rigorous certification standards for behavioral biometric systems.

In addition, we explored ISO guidelines for behavioral biometrics, specifically ISO/IEC DIS 39794-17, which provides data models and data streams for gait recognition. These standards play a key role in establishing best practices and certification criteria for behavioral biometric systems.

This Ph.D. thesis aims to analyze the parameters influencing the performance of behavioral systems and define the quality of behavioral biometric models. In addition, it aims to create databases for assessing system vulnerabilities to logical attacks, thus contributing to the secure use of these systems in various application domains.

In summary, the evaluation of behavioral biometric systems is a complex and necessary process to ensure their reliability. Evaluation criteria, FIDO recommendations, and ISO standards are all valuable tools for establishing sound evaluation practices. By continuing our research in this area, we can continue to improve the security and effectiveness of behavioral biometric systems, opening up exciting new opportunities in the field of individual identification and verification.

CHAPTER 2

Description of Behavioral Biometrics

Summary

The rise of behavioral biometrics has brought significant advances in the field of cybersecurity and authentication. This chapter explores this field in depth, examining the evolution of biometric solutions over time. We also look at the objectives and issues involved in data collection, including privacy challenges and regulations in different regions. Behavioral biometrics, with its key modalities, is at the core of our discussion, highlighting the advantages it offers over other biometric methods. We'll dive into the details of collecting and analyzing behavioral biometric data, addressing the different types of sensors used and the security measures needed to guarantee data quality. In addition, we explore current applications for behavioral biometrics, from enhancing security to contributing to advances in healthcare. Finally, we discuss the limitations and potential risks of using this data, as well as recent developments that are shaping the future of behavioral biometrics. This chapter thus provides a foundation for understanding this constantly evolving field.

Keywords: Behavioral biometrics; Cybersecurity; Authentication.

2.1 Introduction

The term "*biometrics*" originates from the Greek words "*bios*", meaning *life*, and "*metricos*", meaning *measurement*, and refers to the examination of biological features. Biometrics involves utilizing physiological or behavioral characteristics in an automated manner to establish identity, with verification achieved through the measurement of an individual's physiological or behavioral traits. Various biometric techniques, such as fingerprint, palm print, hand geometry, face, ear, iris, voice, signature, body odor, and so on, have been proposed by researchers for human identification and authentication purposes [Saeed, 2016]. However, it is important to note that the use of this behavioral data must be done in an ethical manner that respects the privacy of individuals. Healthcare professionals must ensure that the data collected is securely stored and used in accordance with applicable regulations and standards.

The statistical nature of biometric traits makes the system highly reliable and unique, particularly when a sufficient amount of data is available for analysis. Biometrics operates across various modalities, utilizing measurements of an individual's physical features and body, as well as their behavioral patterns. These modalities are classified based on the individual's biological traits and typically fall into two main types: 1) physiological and 2) behavioral [Saeed, 2016]. Physiological biometrics, like fingerprint recognition, were characterized as stable and unchanging and, therefore, reliable; behavioral biometrics, like speech recognition, were untrustworthy because they depended on variable and unpredictable human behavior. Table 2.1 below illustrates some points of different biometrics technologies traits.

TABLE 2.1: Different types of biometrics with their features [Alsaadi, 2021].

Biometrics Traits	**Type**	**User Acceptance**	**Reliability**	**Universality**
Face Recognition	Physical	Medium	High	High
Voice Recognition	Behavioral	High	Medium	Medium
Finger Recognition	Physical	Medium	High	Medium
Signature Recognition	Behavioral	High	Medium	Lower
Iris Scanning	Physical	Medium	High	High
Gait Recognition	Behavioral	High	High	Medium
Keystroke Dynamics	Behavioral	Medium	Medium	Lower
Hand Geometry	Physical	Medium	Medium	Medium

Figure 2.1 presents a taxonomy of biometrics. Behavioral biometrics involves measuring a user's behavioral tendencies, which can include gait, voice recognition, signature verification, keystroke dynamics, mouse dynamics, and Graphical User Interface (GUI) usage analysis [Bailey et al., 2014]. According to Bailey *et al.*, behavioral biometrics has not been as widely adopted as physiological biometrics due to the variability of the human body and mind [Bailey et al., 2014]. It is worth noting that analyzing user activities does not require additional hardware.

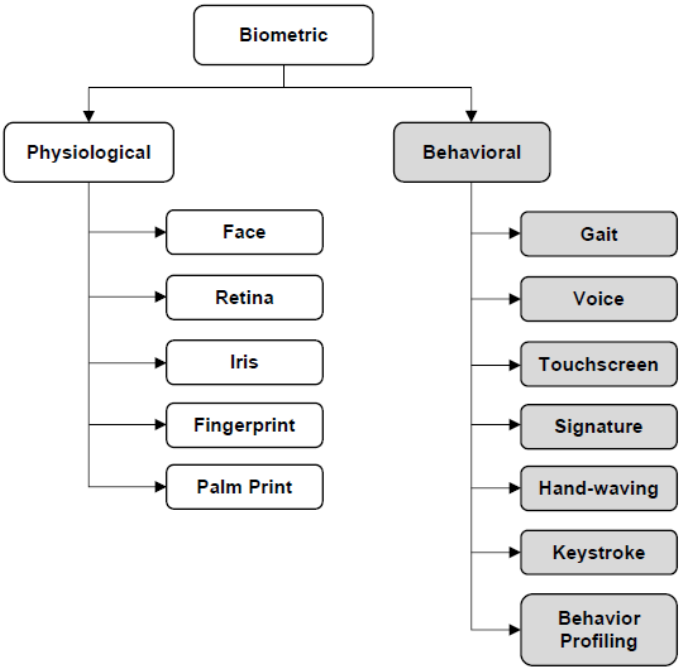


FIGURE 2.1: Taxonomy of biometrics modalities.

It is noticed that behavioral biometrics is a measurement that characterizes how an individual interacts with his or her environment, and this interaction is captured by a device, such as keystroke dynamics, eye movements, voice, signature, or mouse handling as illustrated in Figure 4.8. These measures can be used to improve computer security by enabling more reliable authentication, but also to optimize the user experience by personalizing interfaces and improving performance.

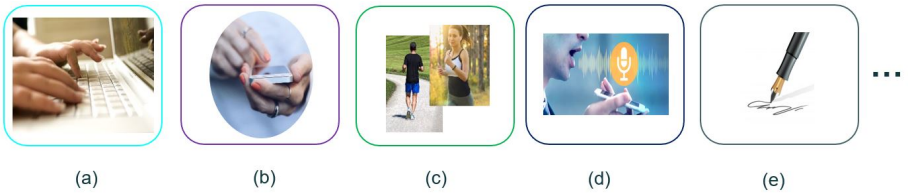


FIGURE 2.2: Behavioral traits: (a) keystroke dynamics, (b) touchscreen, (c) gait or human activity (d) voice, and (e) signature.

Behavioral biometric data are also useful in scientific research to study cognitive disorders and analyze human behavior. However, the use of such data raises ethical and data protection issues [Bailey et al., 2014], including abusive collection, excessive surveillance, and discrimination. It is therefore important to comply with applicable regulations and adopt good practices to ensure the protection of individuals' personal data.

The behavioral category of biometric modalities includes features that we acquire through our interactions with the environment and nature throughout our lives. This modality involves changes in human behavior over time.

2.1.1 Evolution of biometric solutions

Biometrics were introduced as early as the ancient Babylonian empire. However, the modern biometrics industry did not begin until the 1800s. Alphonse Bertillon is considered to be the pioneer of modern biometrics, as he developed the Bertillon body measurement system. Biometric solutions have evolved over time, to become more reliable and efficient. For example, fingerprint recognition has become more accurate, and facial recognition technology has improved. Additionally, newer biometric solutions, such as iris scanning or behavioral biometrics are becoming increasingly popular. These advancements allow for more accurate and secure biometric solutions. Figure 4.10 shows the timeline of the evolution of biometric solutions.

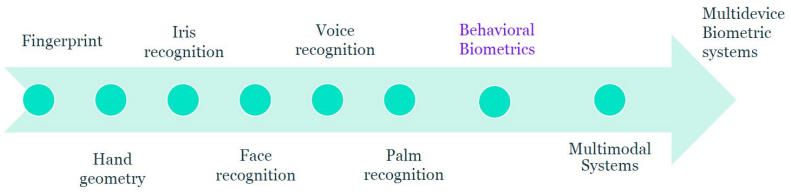


FIGURE 2.3: Evolution of biometric solutions.

The hybrid modality involves a combination of physical and behavioral traits, such as voice recognition, which relies on both the anatomical features of the vocal cords, nasal and mouth cavities, and lips (physical traits) as well as the emotional status, age, and illness of an individual (behavioral traits). The hybrid modality is also classified as a type of multimodality, which involves the use of more than one mode to authenticate an individual's identity [Saeed, 2016]. This means that to have access to this behavioral data, we need to collect them.

2.1.2 Objectives and issues of the data collection

Each person's behavioral pattern is made up of a diverse range of unique behaviors, which are all combined to form a larger, distinctive profile. This behavior pattern is not only shaped by biometric features but also influenced by social and psychological factors, making it impossible to replicate someone else's behavior [Saeed, 2016]. The behavioral pattern of the person is compared with the stored pattern. Matching scores from similarities scores are computed to recognize or authenticate users. This helps us to define in the next section the objectives of data collection.

The purposes for collecting behavioral biometric data are diverse. They can be related to improve IT security, optimize the user experience, or scientific research. Here are some examples:

- **Fraud detection**:
  Behavioral biometrics data can be used to detect fraudulent activities such as the use of stolen login credentials or the creation of fraudulent accounts [Migdal, 2019a].
- **User authentication**:
  Behavioral biometrics can be used to verify a user's identity by comparing measurements collected during a session with previously recorded measurements [Parkinson et al., 2021].
- **Interface personalization**:
  Behavioral biometrics data can be used to personalize interfaces based on a user's

preferences and habits, for example by providing product recommendations or relevant content [Jaouedi et al., 2020].

- **Study of cognitive disorders**:
  Behavioral biometrics can be used to study cognitive disorders such as dyslexia or Alzheimer's disease by analyzing individuals' typing patterns or eye movements [Piugie et al., 2019].

The challenges of collecting behavioral biometric data are related to the privacy and security of personal data. This data can be sensitive and reveal personal information about individuals. It is therefore essential to guarantee the confidentiality and security of this data by adopting appropriate protection measures. Current regulations, such as the General Data Protection Regulation (GDPR) in Europe, strictly control the collection and use of this data. It is therefore important to comply with these regulations and to adopt good practices to ensure the security and confidentiality of behavioral biometric data.

In France, any campaign to collect biometric data must be declared in advance to the "Commission Nationale de l'Informatique et des Libertés" (CNIL). This is an independent administrative authority in France responsible for ensuring the protection of personal data and the preservation of privacy in the context of the use of information technologies. Its main role is to ensure that the data protection rights of individuals are respected and that the processing of personal data complies with the law. The CNIL is also responsible for regulating and supervising activities relating not only to the collection but also the storage, use, and dissemination of personal data.

2.1.3 Ethical and legal issues

2.1.3.1 Challenges of privacy and personal data protection

The collection and use of behavioral biometrics data raises major privacy and data protection concerns. Individuals may be reluctant to share this data because it is personal and often reveals intimate information about their habits, behaviors, and emotions.

Organizations and companies that collect and use this data must comply with privacy and data protection laws and regulations, such as the European Union's **GDPR**. However, it should be noted that the GDPR only applies to companies that operate and store data on European soil, which means that not all companies are affected. These regulations impose strict requirements on the collection, storage, processing, and transfer of personal data, as well as severe penalties for violations.

It is also important that companies and organizations clearly communicate to individuals how their data is used and with whom it is shared. Individuals must have the right to give informed consent to the collection and use of their behavioral biometric data and to withdraw it at any time.

Finally, companies and organizations must ensure the security of behavioral biometric data by storing it securely and taking steps to prevent unauthorized access or misuse.

2.1.3.2 Laws and regulations by region

Laws and regulations applicable to the collection and use of behavioral biometric data vary by country and region. Examples of important laws and regulations in this area include:

1. **European Union's General Data Protection Regulation (GDPR [1])**
   The **GDPR** establishes strict standards for the collection, processing, and transfer of personal data, including behavioral biometric data.

2. **Taiwan's Personal Information Protection and Administration System (TPIPAS [2])**
   TPIPAS imposes strict rules for the collection, use, and disclosure of personal data, including behavioral biometric data.

[1] `https://gdpr.eu/what-is-gdpr/`
[2] `https://www.tpipas.org.tw/`

3. **Singapore's Personal Information Protection Act (PIPA [3])**

   PIPA imposes strict rules for the collection, use, and disclosure of personal data, including behavioral biometric data.

4. **Canada's Personal Information Protection and Electronic Documents Act (PIPEDA [4])**

   PIPEDA establishes rules for the collection, use, and disclosure of personal information, including behavioral biometric data.

5. **California's Personal Information Protection Act (PIPA [5])**

   PIPA imposes transparency and consent obligations for the collection and use of personal data, including behavioral biometrics.

   As of January 1, 2004, PIPA, the Personal Information Protection and Electronic Documents Act are in effect. It is accompanied by the PIPA regulations. PIPA provides individuals with the right to access their personal data, while it provides private organizations with a framework for the collection, use, and disclosure of that data. Private entities subject to PIPA include corporations, unincorporated associations, professional regulatory associations, unions, partnerships, private schools or colleges, and any person engaged in commercial activity. Non-profit organizations engaged in commercial activities are subject to PIPA on a limited basis.

6. **California Consumer Privacy Act (CCPA [6])**

   The **CCPA** is a 2018 California law to protect consumer privacy, giving consumers more control over the personal information collected about them by businesses. CCPA regulations provide guidance for implementing this landmark law, which grants new privacy rights for Californians, such as the right to know what personal information is being collected, the right to delete that information (with some

[3] https://sso.agc.gov.sg/Act/PDPA2012/

[4] https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda

[5] https://oipc.ab.ca/legislation/pipa/

[6] https://theccpa.org/

exceptions), the right to refuse the sale or sharing of that information and the right to non-discrimination in the exercise of those data protection rights.

In addition to these laws, there are also standards and guidelines, such as ISO 29100 and ISO 27001, that provide guidance for managing the privacy and security of personal data, including behavioral biometric data.

It is important that companies and organizations comply with these laws and regulations to ensure that behavioral biometric data is collected, stored, and used in an ethical, transparent, and privacy-friendly manner.

The chapter is organized as follows. Section 2.2 presents the state-of-the-art in behavioral biometrics. The section compiles and analyzes the methods of behavioral biometrics collection. Section 2.3 draws applications and trends in behavioral biometrics and additionally, details recent developments in the field of behavioral biometrics. Section 2.4 concludes this work and provides some perspectives.

2.2 Behavioral biometrics

2.2.1 Main modalities

Behavioral biometrics can be collected using a variety of physical measurements and characteristics, such as heart rate, body movements, eye movements, voice, typing speed, and Internet browsing habits. There are several types of behavioral biometrics, each of which can characterize an individual's behavior when interacting with a computer system:

- **Keystroke dynamics**

  Keystroke dynamics is a behavioral biometric modality consisting in analyzing the way someone types on a keyboard [Migdal, 2019a, Ayotte et al., 2021b]. It is based on an individual's typing habits. Characteristics such as speed, rhythm, keystroke pressure, duration of pauses, and typing errors can be measured to identify and authenticate an individual.

- **Mouse dynamics**

  Behavioral biometrics offers another example of mouse dynamics recognition. This method evaluates how the user interacts with his computer using his mouse [Ahmed and Traore, 2010, Antal et al., 2021, Monaro et al., 2020]. A behavioral profile is created by analyzing specific mouse movements. Although mouse dynamics and typing are complementary, the mouse is preferred for graphical interfaces, while the keyboard is more commonly used for text entry and commands [Bhatnagar et al., 2013]. The combination of these two methods enhances computer security [Sharma and Elmiligi, 2022]. Mouse dynamics data is based on an individual's mouse movements and clicks. Characteristics such as speed, rhythm, trajectory, and click count can be measured to identify and authenticate an individual.

- **Eye movements**

  Eye movement data is based on an individual's eye movements when interacting with a computer system. Characteristics such as fixation time, travel distance, and blinking frequency can be measured to identify and authenticate an individual.

- **Voice**

  Voice data is based on an individual's vocal characteristics such as frequency, pitch, rhythm, rate, and accent. These characteristics can be measured to identify and authenticate an individual.

- **Signature**

  Signature data is based on the characteristics of an individual's handwritten signature. Characteristics such as shape, size, slant, pressure, and speed can be measured to identify and authenticate an individual.

Each of these types of behavioral biometrics has advantages and limitations. For example, keystrokes can be used to detect fraudulent activity but can be affected by variations in language and keyboard use. Voice can be used for voice authentication, but can be affected by variations in the sound environment. Therefore, it is important to choose the most appropriate measure based on the purpose and characteristics of the user and the computer system.

2.2.2 Advantages of behavioral biometrics

There are many advantages of behavioral biometrics over physical biometrics [Sharma and Elmiligi, 2022]. The following points illustrate these advantages which are non-exhaustive [Liu and Silverman, 2001, Yampolskiy and Govindaraju, 2010b, Alsaadi, 2021]:

- **Continuous collection and authorization**

  Behavioral biometrics provides continuous user monitoring, ensuring that only authorized individuals access the system, even after initial identity verification.

- **Non-intrusive collection**

  Behavioral data can be collected seamlessly, without disrupting normal service usage.

- **No special equipment required**

  Behavioral data can be collected using a standard camera or voice recorder, without requiring special equipment. Video or audio recordings are processed to extract data for later authorization.

- **Useful for authorization**

  Behavioral biometrics provides continuous user authentication and is a powerful defense. However, it should only be used as a supplement to one-time authentication techniques such as passwords, PINs, and other physiological biometrics.

- **Universality**

  With respect to universality, when behavioral biometrics is applied to a large population, it has low universality because of the small degree of difference in behavior [Alsaadi, 2021]. This criteria is not only specific to behavioral biometrics.

- **Circumvention**

  With respect to circumvention, the characteristics of behavioral biometrics are very difficult to imitate or copy [Sharma and Elmiligi, 2022]. This criteria is not only

specific to behavioral biometrics.

- **Unique combination**

  Regarding unique combination, behavioral biometrics is mainly a unique combination of behavioral features analyzed for each real person. This criteria is not only specific to behavioral biometrics.

- **Smooth integration**

  Regarding smooth integration, once the behavioral biometrics model is defined, it can be easily integrated into existing security systems. For example, the usual video surveillance system can be used to implement a behavioral biometric system. This criteria is not only specific to behavioral biometrics.

- **Good verification accuracy**

  Regarding verification accuracy, behavioral biometrics has shown good verification accuracy in multimodal identification systems [Sharma and Elmiligi, 2022]. This criteria is not only specific to behavioral biometrics.

- **Acceptability**

  Regarding acceptability, behavioral biometrics are often collected without user participation, making them very acceptable. However, they can face several privacy and ethical objections.

Table 2.2 shows the different application areas for behavioral biometrics, such as security, health, and behavior-based authentication systems.

TABLE 2.2: Behavioral biometrics commercial organizations [Sharma and Elmiligi, 2022].

Company Name	Year	Types	Used by
BioCatch [7]	**2011**	Typing speed, Swipe pattern, mouse clicks	HSBC, Itau, BARCLAYS, nab, American Express, citi VENTURES, 86400 banks, NatWest
Simprints [8]	**2012**	Wireless Fingerprint scanners	BRAC, Cohesu
Plurilock [9]	**2016**	Keystroke dynamics, Pointer dynamics	US federal agencies
TypingDNA [10]	**2016**	Keystroke dynamics	Microsoft Azure, ForgeRock, Optimal IdM, BBVA, Proctoru, Capgemini
ThreatMark [11]	**2015**	Mouse events, keystroke dynamics, site navigation patterns, interaction with website elements	SLOVENSKA SPORITEIŇA(Bank), SBERBANK
DiVi [12]	**2011**	Facial Recognition, Skeleton tracking	Intel, Adidas, LG, Orbbec
Zighra [13]	**2010**	Task-based authentication using behaviors such as holding the phone and swiping across the screen	Government of Canada innovation Fund
VoiSentry [14]	**2018**	Speaker identification and verification system	ForgeRock, University of York, MyForce

Cynet [15]	**2018**	Behavior analytic System to continuous monitoring	Darktrace, Microsoft Azure, Vectra Networks
BehaioSec Inc. [Stiller et al.,]	**2010**	The API can turn behavior into actionable intelligence with just a few lines of code	IDG, Gartner, Goode Intelligence
SecureAuth Inc. [16]	**2015**	Identity Security Without Compromise	Xerox, Michaels, Unisys
Unify Id	**2015**	Passive behavioral authentication platform designed to identify users without any conscious user action	US banks
SecureTouch Inc.	**2014**	Deliver continuous authentication technologies to strengthen security and reduce fraud	Zaraz, Neon Media, TimeRack

[7] https://www.biocatch.com/
[8] https://www.simprints.com/
[9] https://plurilock.com/
[10] https://www.typingdna.com/
[11] https://www.threatmark.com/whythreatmark/
[12] https://3divi.ai/
[13] https://zighra.com/
[14] https://www.aculab.com/
[15] https://www.cynet.com/platform/threatprotection/uba-user-behavioranalytics/
[16] https://www.secureauth.com/

Having examined Table 2.2, we can see that the deployment of kits using behavioral biometrics applications within the population is very varied, with the active participation of several leaders in the digital and banking sectors.

2.2.3 Literature review

Behavioral biometrics is attracting strong interest among researchers and industry experts, with significant influence in various fields such as user profiling, user modeling, adversary modeling, criminal profiling, jury profiling, etc. [Yampolskiy and Govindaraju, 2010b]. The data used for behavioral analysis comes from various sources such as sensors, cameras, keyboard and mouse usage, devices, audit logs, signatures or handwriting, programming style, language, smell, etc. [Yampolskiy and Govindaraju, 2010b]. In addition, physical characteristics such as smell, heart rate, and even DNA are also exploited in certain applications. Researchers are also exploring ECG, brainwaves, and movement to analyze behavioral traits [Yampolskiy and Govindaraju, 2010b].

One of the most widely used methods in behavioral biometrics is keystroke dynamics. This method has been used for years to authenticate users based on keystroke patterns extracted from raw data, whether standard or non-standard passwords. These characteristics can be used to create a unique profile for each user, authorizing subsequent access to resources [Choi et al., 2021, El Zein and Kalakech, 2018, Halakou, 2013]. In addition, they can be used to recognize a person's emotions [Qi et al., 2021]. For example, emotion recognition based on typing patterns is achieved by asking users to type a specific sentence. Using feature extraction techniques, predictive models can be developed to classify different emotions. One study defined and created an emotion detection model based on users' typing and swiping habits, with an accuracy rate of 73% [Ghosh et al., 2019]. Typing and swiping patterns are used in many applications to detect the emotions of smartphone users [Ghosh et al., 2019].

Another example of behavioral biometrics is mouse dynamics, where recognition of a user's profile is based on the way they use their mouse on the computer [Wang and Geng, 2009, Antal et al., 2021, Monaro et al., 2020]. The behavioral profile is created by extracting specific features related to a user's mouse movements. Mouse and typing dynamics are

complementary and closely linked. Mouse use is paramount in graphical user interface applications, while the keyboard is commonly used in word processing and command line applications [Bhatnagar et al., 2013]. Analysis of mouse and keystroke dynamics plays an essential role in improving IT security.

Gait analysis (GAIT) is one of the most exciting areas of research in the field of behavioral biometrics. It is used to authenticate users based on their walking style or manner [Katiyar et al., 2013, Chai et al., 2022]. GAIT analysis systems are mainly based on the use of a video camera that captures images of people in motion. These images are processed to extract appropriate user characteristics, such as joint angles or silhouettes, and the resulting values are then compared with the recorded signatures and walking profiles of authorized persons. One of the main advantages of GAIT analysis is its non-intrusive nature, as it does not require the individual's cooperation and can operate at moderate distances from the person being observed.

Biotouch is another framework based on behavioral biometrics and localization, used for continuous authentication in mobile banking applications [Estrela et al., 2020]. Biotouch uses touch patterns to profile users as they tap or hold the device. This data is then used to build predictive models and for authorization.

A new approach to user behavioral profiling involves creating profiles based on their playing style. This technique analyzes the strategies used during a game and creates a behavioral profile based on these strategies, similar to behavioral biometrics. These profiles are then used to continuously observe the player and authorize his or her access to servers [Yampolskiy and Govindaraju, 2010a]. For example, this approach is used to explore the strategies employed during a poker game in order to create behavioral biometric profiles [Yampolskiy and Govindaraju, 2010a]. Once the profile has been created, it can be used to authorize the player on the move.

Another interesting approach is to use smell as a biometric element to identify individuals [Gibbs, 2010]. In this method, small quantities of permanently evaporating odor-generating molecules, called odorants, are detected by a special sensor called an e-nose. The e-nose is a chemical sensor capable of collecting data unique to each participant. This data can be used to form classification models and for user authentication [Borowik et al.,

2020]. The e-nose is a fast, non-invasive, and intelligent online instrument with feasible and efficient odor recognition. It consists of a set of sensors and constitutes a suitable pattern recognition system, capable of identifying specific odors.

User classification has made extensive use of facial recognition and emotion detection in many applications. Gabor wavelets represent a method for extracting features from images for recognition. For example, in the context of facial recognition, facial images are analyzed by pre-processing or normalizing the facial image [Amin and Yan, 2010]. In general, eyes and mouth are aligned at approximately the same location in images of similar size for face processing. Gabor filters at different scales and orientations are applied to each face image to generate feature vectors used to train machine learning models.

Many researchers consider handwriting biometrics to be behavioral biometrics, as they are based on the actions performed by a specific subject. Handwriting recognition involves transforming a language represented as spatial graphical marks into a symbolic representation [Plamondon and Srihari, 2000].

Speech recognition is another behavioral biometric that can be used to identify a speech pattern based on the most common sound variations in a person's speech. Speaker identification and verification can be achieved by capturing key speaker features in a narrow band, such as pitch and formants [Ravanelli and Bengio, 2018]. This technique is used for biometric authentication, forensics, security, speech recognition, and speaker diarization.

Table 2.3 summarizes a brief overview of previous studies.

2.2.4 Collection and analysis of behavioral biometric data

The methods and techniques employed for gathering behavioral biometric data include wearable sensors, mobile applications, and online platforms.

2.2.4.1 Behavioral biometric data collection

The collection and analysis of behavioral biometric data involve several key steps such as:

TABLE 2.3: Behavioral biometrics research work.

Behavioral Biometrics	**Purpose**
Keystroke Dynamics	To recognize a person using keystroke dynamics [Shadman et al., 2023]
Keystroke and Mouse Dynamics	Identity theft issues by verifying users based on their keystroke dynamics and mouse activities [Shi et al., 2023]
Touch and hold a device	Emotion detection from touch interactions during text entry on smartphones [Li et al., 2023b]
Touch Patterns	continuous authentication on mobile banking applications [Stragapede et al., 2023]
Mouse Dynamics	Computer user recognition based on the way a user uses his/her mouse [Monaro et al., 2020]
GAIT	Authorization process based on style or manner of walking [Akber et al., 2023]
Strategy	Player profile is used to authorize the player on the go [Giles et al., 2023]
Odor	Human recognition through the odor authentication [Manikantaa and Saranya, 2023]
Gabor wavelets	To extract features from an image for recognition [Li et al., 2023a]
Handwriting Biometric	A process of transforming a language represented in its spatial form of graphical marks into its symbolic representation [Wang et al., 2023]
Speech	Useful for biometric authentication, forensics, security, speech recognition, and speaker diarization [Mishra et al., 2023]

- **Data collection**

  Behavioral biometric data can be collected using different types of sensors, such as motion sensors, cameras, heart rate sensors, microphones, pressure sensors, etc. Data can be collected continuously or on an ad hoc basis depending on the needs of the application.

- **Data storage**

  Collected data must be stored in secure databases to ensure confidentiality and integrity. Data must also be stored in accordance with applicable laws and regulations.

- **Data preprocessing**

  Collected data must be preprocessed to make it usable. This step typically involves cleaning the data, eliminating missing data, normalizing the data, and transforming the data so that it can be used in analytical algorithms.

- **Data analysis**

  The pre-processed data is then analyzed using statistical analysis and modeling techniques. These techniques may include analysis of variance, regression analysis, factor analysis, classification, clustering, etc.

- **Interpretation of results**

  The results of the analysis must be interpreted to obtain useful information about users or computer systems. The results can be used to improve the user experience, identify system performance problems, detect fraud, etc.

- **Data security**

  Data security is an important aspect of behavioral biometric data collection and analysis. Data must be stored securely, access to the data must be limited, and the data must be handled in accordance with applicable laws and regulations.

In sum, the collection and analysis of behavioral biometric data can provide valuable information about users and computer systems. However, it is important to ensure data confidentiality and security to avoid any risk of abuse or privacy violation.

2.2.4.2 Different types of sensors used for data collection

There are different types of sensors integrated for example in a smartphone that can be used to collect behavioral biometric data such as presented in Figure 2.4 and Figure 2.5.

Motion sensors, *cameras*, *heart rate sensors*, *microphones*, *pressure sensors*, *keystroke speed sensors*, and *web browsing sensors* are commonly used measurement tools to assess various aspects of the human body and behavior, such as physical activity, posture, gestures, heart rate, voice and speech, surface pressure, keystroke speed, and web browsing habits.

FIGURE 2.4: Smartphone: integrated sensors and actuators, along with a variety of available operating systems and general apps [Rayani and Changder, 2023].

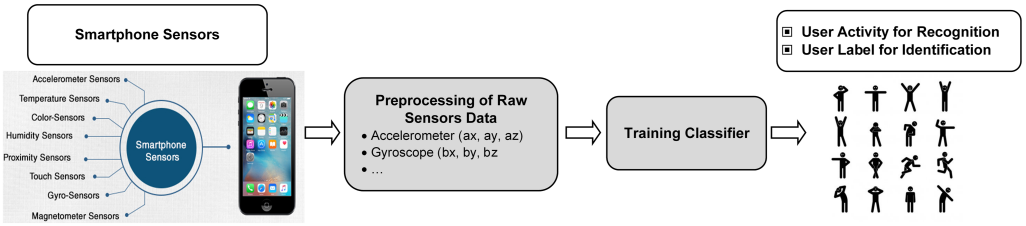


FIGURE 2.5: Smarthphone sensors usage.

These different types of sensors can be used alone or in combination to collect behavioral biometrics data. The choice of the sensor depends on the characteristics being measured, and the application being considered. Figure 2.6 describes how behavioral biometrics works from data collection to the matching process.

The different types of sensors can be used to collect behavioral biometrics data depending on the needs of the application.

2.2.4.3 Precautions to be taken to ensure data quality and security

When collecting and analyzing behavioral biometric data, it is important to take precautions to ensure data quality and security. Some of the most common precautions include [Sharma and Elmiligi, 2022, Stragapede et al., 2023] :

- **Respect privacy through a consent form**

  Before collecting behavioral biometrics, it is important to obtain informed consent

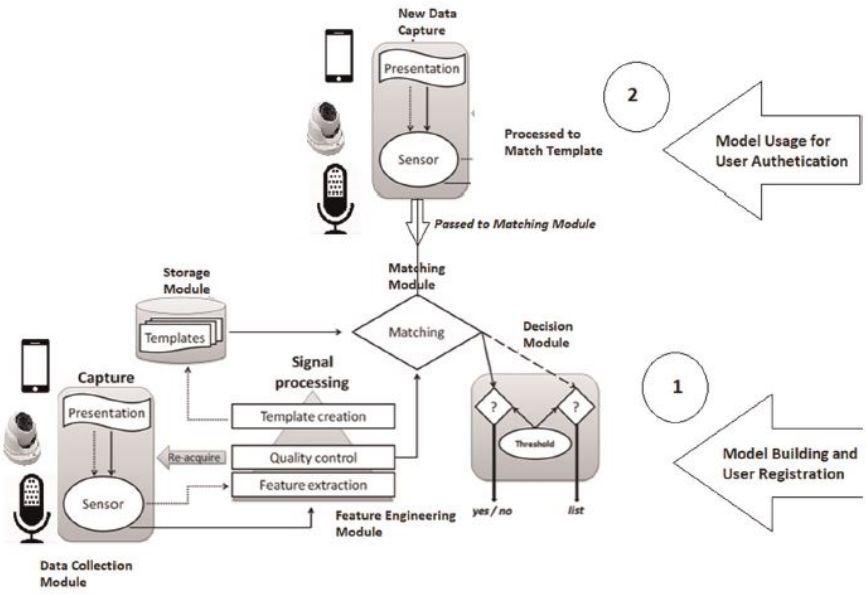


FIGURE 2.6: Behavioral biometric model [Sharma and Elmiligi, 2022].

from the individual. This means that the individual must clearly understand why the data is being collected, how it is used, and what the potential consequences are.

- **Use reliable data collection tools**

  It is important to use reliable data collection tools to ensure that the data collected is accurate and reliable. This may include the use of high-quality motion sensors or validated data collection software.

- **Protect the data collected**

  Behavioral biometric data can be considered as sensitive data since it can be used to identify individuals. It is therefore important to protect this data by using secure storage and transmission methods.

- **Respect privacy**

  It is important to respect the privacy of the individuals whose data is being collected. This may include using pseudonyms to protect the identity of individuals and limiting access to data to authorized individuals.

- **Follow regulations and standards**

  There are specific regulations and standards that govern the collection and use of

behavioral biometric data. It is important to follow these regulations and standards to ensure compliance and data security which depends from one country to another ones.

In order to use the system in a real-world context, it is imperative to accurately measure its quality. This involves defining the context of use, as well as the effectiveness and robustness of the logic, to determine whether it meets the specific requirements of an application and whether it is based on logical or physical access. It is crucial to compare different biometrics modalities to analyze their respective advantages and disadvantages. Performance evaluation is also an essential step to facilitate research in this area. Evaluation techniques are used to measure the performance of behavioral biometrics systems, and there is a need for a reliable method to analyze the benefits of the system.

Quality is an essential factor because it has a direct impact on the successful use of the system in a real-life context. Controlling the quality of acquired samples ensures that system performance measurements are reliable and accurate. Indeed, enrolling a poor-quality sample can distort results, compromise the system's ability to meet specific application requirements, and lead to logical or physical access problems. Consequently, the quality of the input data is a critical element in ensuring the effectiveness and robustness of the behavioral biometrics system, as well as facilitating research in this field by enabling a relevant assessment of the system's advantages over other biometric modalities.

In summary, it is essential to take precautions to ensure the quality and security of behavioral biometric data. This includes respecting privacy, protecting the data collected, using reliable data collection tools, and following applicable regulations and standards.

2.3 Application and trends in behavioral biometrics

Behavioral biometrics is a commonly used method in information security to identify individuals based on the unique characteristics of their activities, whether conscious or unconscious. Recently, behavioral biometrics have been used in several interesting applications. Researchers have proposed methodologies for speaker recognition by studying lip

movements, biometric verification by analyzing finger movements, as well as extraction of biometric features from the voice for user identification [Saeed, 2016].

TABLE 2.4: Behavioral biometrics timeline [Sharma and Elmiligi, 2022].

Year	**Biometric Used**	**Used for**
1960	First model speech production using X-rays of speaking subjects	Authentication
1960	Facial recognition	Identification
1965	Signature recognition system	Identification
1970	Dynamic signature and fingerprints recognition	Identification
1970	An early form of biometric modeling using full-motion x-rays and the previous work of Drs. Fant and Stevens, even used today	Authentication
1980	Speech Group to promote voice recognition tech	Recognition
1991	Real time face recognition	Recognition
1996	Hand geometry recognition gets deployed at Olympics	Identification
1999	ICAO initiates study on biometrics and MRTD	Issuance and acceptance
2001	Face recognition is deployed at the Super Bowl	Recognition
2001	attacks on the World Trade Center draw attention to the need for continuous authentication as a new security measure in global information systems	Security
2002	DARPA launches Total Information Awareness (TIA), the first large-scale use of technologies designed to mine data sets for identifying biometric information	Identification

2004	US-VISIT (United States Visitor and Immigrant Status Indication Technology) becomes operational	Authorization
2006	Keystroke Dynamics embedded in consumer products	Continuous authorization
2010	Osama bin Laden's body gets identified with biometrics	Identification
2011	Mobile biometrics	Authorization
2013	Continuous authentication for mobile application security	Authorization
Mid 2010s	Biometric systems to improve security as well as the system performance	Authorization
Late 2010s	Electric vehicles with face biometrics	Authorization
2018	World's first phone with under-display fingerprint sensor	Identification

Behavioral biometrics has a history that predates the rise of artificial intelligence and machine learning, such as convolutional neural networks (CNN) and other related techniques. However, these technological advances have considerably enriched and expanded the possibilities of behavioral biometrics, making it more precise and adaptable. In the 1960s, researchers Gunnar Fant and Kenneth Stevens developed the very first model of speech production using X-rays. Subsequently, in 1970, researcher Joseph Perkell exploited these advances to create a biometric model of speech recognition. Table 2.4 draws the usage timeline [Sharma and Elmiligi, 2022] of Behavioral Biometric. The table highlights the use of data in various fields such as security, health, financial services, and research.

- **Security**

  Behavioral biometrics can be used for the authentication and identification of individuals. For example, keystroke patterns or hand movements can be used to identify a person or to detect an identity theft attempt.

- **Health**

  Behavioral biometrics can be used to monitor the health and fitness of individuals. For example, walking patterns can be used to detect changes in mobility or to monitor symptoms of certain diseases.

- **Financial Services**

  Behavioral biometrics can be used in financial services to help prevent fraud and identity theft. For example, keystroke patterns can be used to identify online fraud or suspicious transactions.

- **Research**

  Behavioral biometrics can be used in research to understand human behavior and patterns. For example, analysis of walking patterns can be used to study the mobility patterns of the elderly or people with certain diseases.

It is important to note that the use of behavioral biometric data must be done in an ethical manner that respects the privacy of individuals. Companies and organizations that use this data must comply with applicable regulations and standards to ensure the security and protection of the data collected.

2.3.1 Security applications

Behavioral biometrics are used in many security applications, including authentication, fraud detection, and surveillance.

For the purpose of identification or verification, behavioral biometrics data is first collected and stored. The data is processed further to prepare a signature profile. Using algorithms, predictive models are trained, developed, and evaluated. Later, this model is used as a comparison tool, whenever the user runs the application. Using behavioral patterns, the model is used to continuously verify the user's profile throughout their working sessions. The generic architectures of biometric systems consist of five main modules namely data collection, feature engineering, storage, matching, and decision module.

The data collection module collects biometric data, the feature engineering module prepares it for analysis, the storage module keeps it secure, the matching module compares it, and the decision module makes a decision on user authentication. Together, these modules create a biometric system capable of reliably identifying and verifying individuals.

2.3.2 Health applications

Improvement in information and communication technologies (ICT) and ambient intelligent technologies, such as sensors and smartphones, have facilitated the swift progress of smart environments [Lytras et al., 2018, Visvizi et al., 2020]. A significant amount of resources can be conserved by using sensors to record and monitor patients or automatically detect any irregular behavior [Rasekh et al., 2014, Piugie et al., 2019, Piugie et al., 2022], as depicted in Figure 2.7.

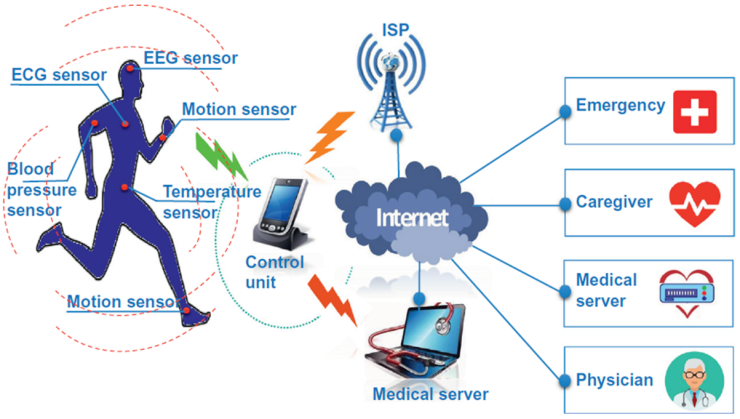


FIGURE 2.7: Framework for intelligent health care monitoring systems (HCMS) [Subasi et al., 2020].

Behavioral biometric data can be used in many health applications, including patient monitoring, disease prevention, and chronic disease management. Some examples of applications of this data for health include:

- **Patient Monitoring**

  Behavioral biometric data can be used to monitor patient health. For example, walking patterns can be used to detect changes in patient mobility, which can help identify health problems such as neurological disorders or heart disease.

- **Disease Prevention**

  Behavioral biometric data can be used to prevent disease. For example, monitoring dietary habits and physical activity levels can help identify risks for diseases such as obesity and diabetes, which can enable healthcare professionals to recommend appropriate preventive interventions.

- **Chronic Disease Management**

  Behavioral biometrics can be used to manage chronic diseases. For example, monitoring sleep patterns and heart rate can help monitor the condition of patients with heart disease, which can enable healthcare professionals to provide appropriate care.

- **Population Data Analysis**

  Behavioral biometric data can be used for population data analysis. For example, analyzing how people move around a city can help health authorities understand the lifestyle patterns and health risks of the population.

By using behavioral biometric data for health, healthcare professionals can better understand patient behaviors and patterns, which can lead to earlier intervention and more effective disease management.

2.3.3 Limits and risks of using behavioral data

The use of behavioral biometrics also has certain limitations and risks that must be considered. Here are some of them:

- **Data Reliability**

  Behavioral biometric data can be influenced by many factors, such as emotional state or fatigue. This can affect the reliability of the data and lead to inaccurate results.

- **Privacy**

  The collection of behavioral biometric data may raise privacy concerns. Individuals may be reluctant to share this data with third parties, especially if they do not understand how it is used.

- **Data Security**

  Behavioral biometric data is sensitive data that must be stored securely to prevent unauthorized access.

- **Human error**

  Human error can occur in the collection and analysis of behavioral biometric data. This can affect the quality of the data and lead to inaccurate results.

- **Bias**

  Algorithms for analyzing behavioral biometric data can be biased due to the subjective nature of the data. This can lead to discriminatory decisions or unfair results.

- **Misuse**

  Behavioral biometric data can be misused or abused in ways such as identity theft, illegal surveillance, or ad targeting.

It is important that companies, organizations, and healthcare professionals use behavioral biometric data in a way that is ethical, transparent, and respectful of individual privacy. Security regulations and standards must be followed to ensure that data is collected, stored, and used appropriately. Individuals must also be informed and aware of the potential risks associated with the collection and use of their behavioral biometric data.

2.3.4 Recent developments in behavioral biometrics

Behavioral biometrics is a rapidly evolving field with many emerging trends. Here are some of the most recent trends:

- **Use of Artificial Intelligence (AI)**

  AI is increasingly being used to improve the accuracy of behavior-based identification and authentication. Machine learning algorithms can be used to analyze behavioral patterns and identify unique characteristics that distinguish individuals. This approach is often referred to as "AI-based behavioral biometrics" or "behavioral AI."

- **Combining multiple biometric factors**

  Security experts increasingly recognize that behavioral biometrics should not be used as a single authentication factor. Instead, they recommend combining behavioral biometrics with other biometric factors, such as facial recognition or fingerprinting, to improve accuracy and security.

- **Use for fraud detection**

  Behavioral biometrics are increasingly being used to detect fraud. Banks and credit card companies often use behavioral biometrics to monitor transactions and detect fraudulent behavior patterns.

- **Healthcare applications**

  Behavioral biometrics is increasingly being used to monitor patient health. Wearable devices and mobile apps can track eating habits, sleep, physical activity, and stress levels, allowing doctors and patients to track progress and identify health issues.

- **Integration into existing security technologies**

  Behavioral biometrics are increasingly being integrated into existing security technologies, such as access control systems and video surveillance systems. This allows for more accurate monitoring and faster identification of potential threats.

Artimetrics is a term that refers to the use of biometric characteristics to authenticate artificial entities such as industrial robots, intelligent software agents, and virtual world avatars [Saeed, 2016]. It involves applying the principles and techniques of biometrics to verify the identity of non-human entities. The concept is still relatively new and there is ongoing research into its development and applications.

Biometric data is also used to enhance the security of cryptographic systems, and new algorithms are being developed by researchers to filter this data [Crihan et al., 2023]. System performance evaluation is essential for the following reasons.

Overall, behavioral biometrics continues to grow and evolve, providing new opportunities for identification and authentication, fraud detection, and health monitoring. The use of AI is also expected to significantly improve the accuracy of these applications in the future.

2.4 Conclusion

In conclusion, behavioral biometrics is personal information that is collected from individual behaviors, such as how one enters a password or walks. This data is increasingly used for identification and authentication, fraud detection, and health monitoring.

The key characteristics of behavioral biometrics are its unique and individual nature, its ability to be collected non-intrusively and continuously, and its ability to be used in combination with other biometric factors to improve the accuracy of identification and authentication.

Behavioral biometrics can improve application security, including user authentication and intrusion detection, with minimal impact on users. However, its effectiveness depends on the method of implementation, such as keystroke dynamics that are influenced by keyboard type. Multimodal systems benefit more from behavioral biometrics, which involves the simultaneous use of multiple types of biometric systems, than unimodal systems that rely on a single type of biometric. According to [Saeed, 2016], multiple spoofing attacks can pose a threat to the security of behavioral biometrics.

The prospects for the future use of this data are very promising. As AI adoption increases, machine learning algorithms are able to analyze even larger amounts of data to improve the accuracy of identification and authentication. Applications of behavioral biometrics are also expected to grow in areas such as healthcare, security monitoring, and fraud prevention.

However, it is important to note that the use of behavioral biometrics raises privacy and data security concerns. Laws and regulations must be put in place to ensure that this data is collected, stored, and used ethically and securely.

CHAPTER 3

Transactional Applications of Behavioral Biometrics: Identification and Authentication

Summary

This part explores transactional applications of behavioral biometrics, focusing on identification and authentication. It provides an overview of this field, outlines related work in areas such as time series analysis, and presents an architecture including feature generation, identification, authentication, and matching algorithms. The experimental protocol is detailed, covering datasets, performance measures, pre-trained models, and classifier parameters. User identification is explored through machine learning and deep learning techniques, with an analysis of expected performance and multi-activity scenarios. Finally, user key authentication is discussed, with performance evaluation on different data and scenarios.

Keywords: Transactional application; Identification; Authentication; Machine Learning; Deep Learning.

3.1 Introduction

Biometrics enables a person to be identified and authenticated using recognizable, verifiable, unique, and specific data. The aim is to capture an item of biometric data from this person. It can be a photo of their face, a record of their voice, or an image of their digital fingerprint. This data is then compared to the biometric data of several other persons kept in a database. Figure 3.1 above resumes the identification and authentication configuration.

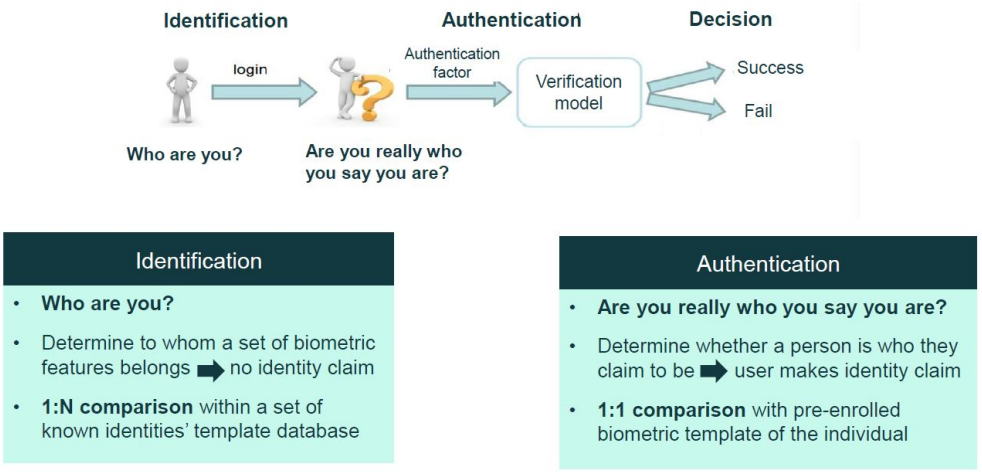


FIGURE 3.1: How do we use biometrics ?

Biometric identification is used to determine a person's identity. This involves capturing specific biometric data about the person. This data is then compared with the biometric data of several other people stored in a database. This time, the question is simple: "Who are you?" In contrast, biometric authentication is a process that compares a person's characteristics with a stored biometric "template" to determine whether they match. To do this, the reference template is first stored. Then, the biometric data of the person to be authenticated is compared with this stored data to answer the question: "Are you Mr. or Mrs. X?"

With the increasing use of smartphones to store personal and sensitive information such as bank account details, personal IDs, passwords, and credit card information, people remain constantly connected and their mobile devices are at risk of security and privacy breaches by malicious actors [Mekruksavanich and Jitpattanakul, 2021, Nugier et al., 2021, Piugie

et al., 2022]. Traditional forms of protection such as passcodes, PINs, patterns, facial recognition, and fingerprint scans are all vulnerable to various forms of attack, including smudge attacks, side-channel attacks, and shoulder-surfing attacks [Mekruksavanich and Jitpattanakul, 2021, Piugie et al., 2021, Piugie et al., 2022].

The development of Information and Communication Technologies (ICT), as well as improvements in ambient intelligent technologies, such as sensors and smartphones, have led to the growth of smart environments [Lytras et al., 2018, Visvizi et al., 2020, Piugie et al., 2021]. Using sensors, staff can save resources by recording and monitoring users or automatically reporting any unusual behavior [Rasekh et al., 2014, Piugie et al., 2019, Piugie et al., 2021]. For instance, in payment systems, to ensure strong customer authentication, it is necessary to implement adequate security features based on authentication factors [1] such as **knowledge**, **possession**, and **inherent** or **biometric factors** [Cherrier, 2021]. **Knowledge factors** are based on information that the user knows, such as a password, PIN, or shared secret. **Possession factors** rely on an object that the user possesses, like a smart card, USB key, smartphone, or security token. **Inherent** or **biometric factors** are directly related to the user and are useful in reducing the risk of unauthorized parties discovering, disclosing, and using elements such as algorithm specifications, key length, and information entropy [Migdal, 2019b]. When Multi-Factor Authentication (MFA) is requested, using Seamless biometrics, as behavioral, improves security without decreasing the User Experience (UX). Increasing performance of such biometrics is a high need of current industrials [Piugie et al., 2022].

User authentication for logical access control, such as browsing the Internet on a laptop, is now commonly done using biometrics [Yohan et al., 2018, Migdal, 2019b, Piugie et al., 2022]. Experts employ various biometric modalities, among fingerprint, retina, and voice recognition, to design recognition systems using artificial intelligence techniques like Machine Learning and Deep Learning. Each approach has its own pros and cons, with fingerprint recognition being well-established and available in commercial products. However, these systems require input readers, such as sensors, which can vary in cost on the

[1] `http://data.europa.eu/eli/reg_del/2018/389/oj`

market [Yohan et al., 2018, Kim et al., 2020]. Moreover, some of these biometric modalities usage is not frictionless for the subject as they have to do an additional action to authenticate themselves.

This led us to ask the question of how easily it is possible to identify a person based on his/her behavior such as with the keystroke dynamics even when users enter the same password. Can we recognize a person by the activities he or she has carried out? Identification is a multi-class classification problem. From the data, a chosen classifier distinguishes and identifies the user who has generated a given characteristics and feature sample, by returning the user ID (identification number or class) of the user to whom these characteristics belong [Bailey et al., 2014]. However, biometric identification has an advantage over passwords as it is based on features that are specific to an individual and are not easy to duplicate or steal [Rosenberger, 2020]. Another advantage for security (and maybe a drawback for privacy) is the possibility to use behavioral biometrics for transparent authentication solutions [Ashibani and Mahmoud, 2020] where user behaviors are constantly analyzed.

Behavioral biometrics involves measuring user's behavioral tendencies, which can include gait, human activity, keystroke dynamics, voice recognition, signature verification, mouse dynamics, and Graphical User Interface (GUI) usage analysis [Bailey et al., 2014]. According to Bailey *et al.*, behavioral biometrics has not been as widely adopted as physiological biometrics due to the variability of the human body and mind [Bailey et al., 2014]. It is worth noting that analyzing user activities does not require additional hardware.

Human activity can be one solution to enhance the security of password authentication without adding any disruptive handling for users. Industries are looking for more security without impacting too much user experience. Considered as a frictionless solution, human activity is a powerful solution to increase trust during user authentication without adding charge to the user like keystroke dynamic as a behavioral modality.

Behavioral biometrics identification/authentication methods have lower performance compared to morphological modalities [Bailey et al., 2014]. The proposed methods aim to introduce an alternative approach using Deep Learning for behavioral biometrics described as time series.

Traditional identification methods based on physiological features outperform behavioral biometric identification [Bailey et al., 2014]. We investigate the effectiveness of a fundamental Machine Learning approach for user identification based on behavioral biometric data. We compare various Machine Learning algorithms, including Deep Learning, for user identification based on user behavior described by time series. In addition, this work highlights Orange's data mining software for prototyping the processing workflow and proposes a highly efficient and simplified method for performing data analysis using Machine Learning classifiers. We examine two modalities of behavioral biometrics: physical activities (such as "laying", "sitting", "standing", "walking", "walking downstairs", and "walking upstairs") captured from a smartphone, as well as keystroke dynamics on a laptop.

Figure 3.2 gives the different use cases of keystroke dynamics. We consider passphrase authentication where all users type the same password, the authentication is realized by only analyzing the way of typing. This approach is convenient for users as no password has to be remembered but is more challenging for research in terms of performance.

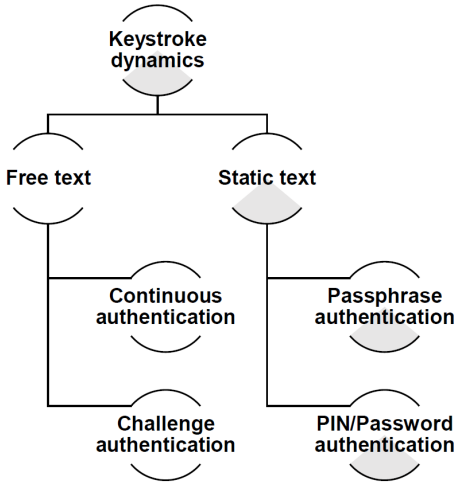


FIGURE 3.2: Overview of the different use cases of keystroke dynamics systems.

We then proposed a method to implement an authentication system using the behavioral biometrics of keystroke dynamics. We assume that by extracting only the keystroke characteristics of each user, it is possible to apply a promising and low-cost authentication system compared to many other biometrics systems, as it does not require any additional sensors and is easy for the user to perform. Keystroke dynamics are characterized by

the time taken to press the keys on the keyboard. The use of these typing characteristics demonstrates the possibility of authenticating a person based on their typing style. Research has also been carried out in recent years to find the best algorithm for this authentication task. Our aim is to evaluate the effectiveness of Deep Learning approaches for password-based user authentication using keystroke dynamics data. We present a proof of concept and examine the performance of different Deep Learning architectures in the section.

We have proposed a solution applied to two behavioral biometric modalities to demonstrate the genericity of the approach. The proposed research uses an image-based architecture (for a chosen behavioral biometric modality: gait analysis). A Deep Learning process for user authentication based on human activity is proposed. We consider only one behavioral biometrics modality which refers to the following physical activities including *laying*, *sitting*, *standing*, *walking*, *walking downstairs*, and *walking upstairs*, all of them being acquired by a smartphone. We tested many architectures for identification/authentication purposes. Generated deep features are fused through different strategies. The obtained performance on a dataset used by the research community outperforms results from the state-of-the-art.

The work is organized as follows. Section 3.2 contains related works on authentication systems from human activity. Section 3.3 draws a generic representation of behavioral biometrics from time series to 2D image color configuration and presents the proposed method for the classical approach with Machine Learning and the different tested Deep Learning models with the specifications and the impact of different parameters on our evaluation system. Section 3.4 draws the experimental protocol. Section 3.5 gives the experimental results as well as subsection 3.5.1 details the experimental results on benchmark datasets for identification in subsection 3.5.2 and authentication process in subsection 3.5.3 on both modalities. Section 3.6 gives the conclusions of this work and some perspectives.

3.2 Related works

Behavioral biometrics represented as time series have been seen as a measurement that characterizes how an individual interacts with his or her environment, and this interaction is captured by a device, such as eye movements, voice, signature, keystroke dynamics, or mouse handling. These measures can be used to improve computer security by enabling more reliable identification/authentication, but also to optimize the user experience by personalizing interfaces and improving performance.

3.2.1 Time series analysis

There are several studies that convert one-dimensional time series into a two-dimensional image representation, then apply a two-dimensional image-based feature extraction technique on this 2D image [Zhong and Deng, 2014, Sanchez Guinea et al., 2022, Sarkar et al., 2023]. Table 3.1 presents the relevant work on this topic. Most of this work focuses on user or activity recognition, but not on user authentication in time-series configurations.

TABLE 3.1: Review of time series analysis: signal-to-image transformation.

Paper	**Approach**	**Method**	**Activity**	**Input Source**	**Accuracy**	**EER**
[Sarkar et al., 2023]	Activity recognition	spatial attention-aided CNN, KNN	Standing, sitting, laying, walking, walking downstairs, walking upstairs, jogging, cycling, running, relaxing, knees bending	Smartphone	[97.72 – 99.90]%	-

[Sanchez Guinea et al., 2022]	Activity recognition	CNN	Standing, sitting, laying, walking, walking downstairs, walking upstairs	Smartphone	99.30%	-
[Sun et al., 2015]	Action recognition	GP-based & k-NN	golf swing (back, front, side), kicking (front, side), riding horse, run, skate-boarding, swing bench, swing (side), and walk	Virtual camera	[86.90 – 88.50%]	-
[Zhong and Deng, 2014]	User identification	I-vector	Gait	Mobile devices	[67.5 – 85.0]%	[06.80 – 08.90]%
[Körner and Denzler, 2013]	Multi-view action recognition	Gaussian process + Histogram intersection kernel	Appearance of dynamic systems captured from different viewpoints	Sony AIBO robot dogs (6)	79.00%	-
[Junejo et al., 2008]	Action recognition	Nearest Neighbour Classifier & SVM	Bend, jack, jump, pjump, run, side, skip walk, wave	Virtual camera (6)	[90.50 – 95.70]%	-

Our study aims to highlight the relevance of our approach by applying it to the analysis of human activity using smartphones and keystroke dynamics on a laptop. Next, we draw up a list of research dealing with these biometric modalities in order to position our contribution in relation to this established knowledge.

3.2.2 Human activities

The applications of user identification are extensive, including logical access control, supervision, and more. The recent advancements in AI have led to a growing interest among researchers in novel research aims, such as object recognition, environmental learning, time series analysis, and predicting future sequences [Rasekh et al., 2014]. Machine and Deep Learning, which have a wide range of applications in speech recognition, language modeling, video processing, and time series analysis, have captured significant attention from AI researchers. Within this fascinating AI domain, one challenging problem is Human Activity Recognition (HAR), which holds promise for eldercare and childcare, when combined with technologies like IoT. Table 3.2 provides a summary of the main research works in the literature discussed below.

TABLE 3.2: Overview of activity recognition based on classical Machine Learning approaches. k-NN : k-Nearest Neighbor; SVM : Support Vector Machine; RF : Random Forest; MLP : Multi-Layer Perceptron; GMM : Gaussian mixture model; KF : Kalman Filter [Al Machot et al., 2020]

Paper	Approach	Method	Activity	Input Source	Performance
[Jaouedi et al., 2020]	Hybrid Deep Learning for activity and action recognition	GMM, KF, Gated Recurrent Unit	Walking, jogging, running, boxing, hand-waving, hand-clapping	Video	96.3%
[Antón et al., 2019]	Infer high-level rules for noninvasive ambient that help to anticipate abnormal activities	RF	Abnormal activities: agitation, alteration, screams, verbal aggression, physical aggression and inappropriate behavior	Ambient sensors	98.0%

[Alex et al., 2018]	Comparison study to classify human activities	SVM, MLP, RF, Naive Bayes	Sleeping, eating, walking, falling, talking on the phone	Image	86.0%
[Shahmohammadi et al., 2017]	Active learning to recognize human activity using Smartwatch	RF, Extra Trees, Naive Bayes, Logistic Regression, SVM	Running, walking, standing, sitting, lying down	Smartwatch	93.3%
[Anguita et al., 2013]	Recognizing human activity using smartphone sensors	Quadratic, k-NN, ANN, SVM	Walking upstairs, downstair	Smartphone	84.4%

It compares different classification techniques, different activities, different sources of input, and finally the best performance that was obtained using a particular classifier.

Biometrics have been widely proposed as a means of continuous user authentication in various studies [Sitová et al., 2015, Patel et al., 2016, Zhang, 2019, Giorgi et al., 2021, Mekruksavanich and Jitpattanakul, 2021]. In the field of continuous authentication, inertial data is used to determine the motion, orientation, and position of a device in the surrounding environment. Methods that use this type of data for non-intrusive authentication employ user behavioral features such as gait, touch screen operations, hand gestures, keyboard patterns, speech, or signature movements to generate behavioral features [Mekruksavanich and Jitpattanakul, 2021].

Zheng et al. [Zheng et al., 2014] were trailblazers in amassing a substantial dataset for continuous authentication and employing a one-class distance-based classifier. Their approach involved utilizing inertial data from the device's accelerometer and gyroscope, along with touchscreen, acceleration, pressure, touch area size, and time frame information between interactions. By developing user profiles based on how individuals held their smartphones when entering their PIN numbers, they achieved an impressive identification of either the genuine owner or an impostor, with an EER of up to 3.6%.

Trojahn et al. [Trojahn and Ortmeier, 2013], on the other hand, explored the application of Deep Learning techniques for smartphone user authentication based on data collected during repeated password entry. They utilized various models, including the multilayer perceptron (MLP) [Pal and Mitra, 1992], Bayesian Net classifiers [Kohavi et al., 1996], and Naïve Bayes [Neverova et al., 2016], to classify users effectively.

Additionally, De Marsico et al. proposed an effective procedure for normalizing signals from smartphone accelerometers in [De Marsico et al., 2016]. The authors demonstrated that normalization had a positive impact on matching data from the same device, particularly in the context of gait recognition.

TABLE 3.3: Human activity aims.

HAR tasks
Basic Activity Recognition
Daily Activity Recognition
Unusual Event Recognition
Biometric Subject Identification
Prediction of Energy Expenditures
Biometric Subject Verification/Authentication

Table 3.3 presents the various aims of human activity, which include identification, authentication, and soft biometrics, depending on the specific case, whether continuous or static [Chen et al., 2021, Piugie et al., 2021]. Biometric solutions using hand movement often rely on reference data, such as typing style, to verify new samples. For identification and authentication, the reference data corresponds to a specific user's typing style, while for soft biometrics, it represents a group of users' typing styles, such as male, female, or left/right-handed. This reference data is crucial for matching or verifying a user's identity from a given sample [Migdal, 2019a]. Our focus lies in biometric verification based on human activity data. We are focusing on the biometric authentication of individuals based on human activity data.

TABLE 3.4: Overview of user activity identification/authentication from the state of art.

Paper	**Approach**	**Method**	**Activity**	**Input Source**	**Accuracy**	**EER**
[Sanchez Guinea et al., 2022]	Activity recognition	CNN	Standing, sitting, laying, walking, walking downstairs, walking upstairs	Smartphone	99.30%	-
[Sarkar et al., 2022]	Activity recognition	Spatial Attention-aided CNN	Standing, sitting, laying, walking, walking downstairs, walking upstairs	Smartphone	99.45%	-
[Parkinson et al., 2021]	User verification	Manhattan distance	Hand movement	Keyboard	[89.00% – 94.00%]	[06.00% – 11.00%]
[Gao et al., 2020]	User identification	SVM	Pose estimation	GUI	74.35%	-
[Jaouedi et al., 2020]	Hybrid Deep Learning for activity and action recognition	GMM, KF, Gated Recurrent Unit	Walking, jogging, running, boxing, hand-waving, hand-clapping	Video	96.3%	-

[Antón et al., 2019]	Infer high-level rules for noninvasive ambient that help to anticipate abnormal activities	RF	Abnormal activities: agitation, alteration, screams, verbal aggression, physical aggression and inappropriate behavior	Ambient sensors	98.0%	-
[Zhang, 2019]	Learning human identity from motion patterns	Dense Clockwork RNN	Walking	Smartphone	93.02%	18.17%
[Barra et al., 2019]	Gender recognition	SVC, RF, AdaBoost, k-NN	Looking and avoid the camera in motion	Video	[68.10% – 82.50%]	-
[Alex et al., 2018]	Comparison study to classify human activities	SVM, MLP, RF, Naive Bayes	Sleeping, eating, walking, falling, talking on the phone	Image	86.0%	-
[Marsico and Mecca, 2017]	Action recognition	DTW	Gait	Smartphone	[83.00% – 93.00%]	[0.09% – 0.10%]
[Shahmohammadi et al., 2017]	Active learning to recognize human activity using Smartwatch	RF, Extra Trees, Naive Bayes, Logistic Regression, SVM	Running, walking, standing, sitting, lying down	Smartwatch	93.3%	-

[Patel et al., 2016]	Continuous user authentication	Ten different classifier	Walking, sitting	Mobile devices	-	07.50%
[Zareen and Jabin, 2016]	User verification	HMM	25 users, 500 signatures	Samsung Galaxy Note	-	06.20%
[Sun et al., 2015]	Action recognition	GP-based & k-NN	golf swing (back, front, side), kicking (front, side), riding horse, run, skateboarding, swing bench, swing (side), and walk	Virtual camera	[86.90% – 88.50%]	-
[Zhong et al., 2015]	Pace independent mobile gait biometrics	Nearest neighbor	Walking	Mobile	-	7.22%
[Zhong and Deng, 2014]	User identification	I-vector	Gait	Mobile devices	[67.5% – 85.0%]	[06.80% – 08.90%]
[Körner and Denzler, 2013]	Multi-view action recognition	Gaussian process + Histogram intersection kernel	Appearance of dynamic systems captured from different viewpoints	Sony AIBO robot dogs (6)	79.00%	-
[Anguita et al., 2013]	Recognizing human activity using smartphone sensors	Quadratic, k-NN, ANN, SVM	Walking upstairs, downstair	Smartphone	84.4%	-

[Muaaz and Mayrhofer, 2013]	Gait recognition, analysis of approaches	SVM	Walking	Cell phone	-	33.30%
[Thang et al., 2012]	Gait identification using accelerometer	SVM	Walking	Mobile phone	92.7%	-
[Junejo et al., 2008]	Action recognition	Nearest Neighbour Classifier & SVM	Bend, jack, jump, pjump, run, side, skip walk, wave	Virtual camera (6)	[90.50% – 95.70%]	-
[Gafurov et al., 2006]	User verification	Histogram similarity and Cycle length	Gait	Mobile devices	-	[05.00% – 09.00%]
[Mantyjarvi et al., 2005]	Identifying users from gait pattern	Correlation coefficients	walking	Smartphone	[72% – 88%]	7%

3.2.3 Keystroke dynamics

Most of the work in the literature is based on conventional machine learning for user authentication (based on the behavioral biometric modality). Although the proposed algorithms are excellent, the results can be improved. Therefore, this research proposed an image architecture (for a chosen behavioral biometric modality) and a deep learning authentication process using neural networks for user authentication based on a passphrase.

Keystroke dynamics can be used for different goals (identification, authentication, soft biometrics) in different cases (free text, fixed text, same-text). Like any biometric solution, keystroke dynamics systems require sets of prior knowledge (references) that are used to verify the newly acquired data (sample). For identification and authentication, a reference describes the typing style of a specific user, while for soft biometrics, a reference describes the typing style of a set of users (e.g., male, female, left/right-handed). The references are

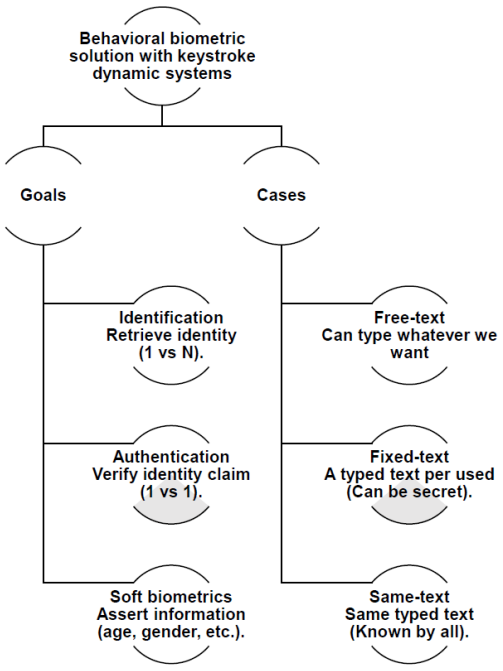


FIGURE 3.3: Keystroke dynamics usages [Migdal, 2019a].

then used to retrieve, or verify, the identity of the user who typed from a sample [Migdal, 2019a]. This is depicted in Figure 3.3.

Keystroke typing dynamics allows profiling users (identification, authentication, gender recognition, profiling) by analyzing the way a user is typing on a keyboard for example when surfing on the Internet. Keystroke dynamics was first used in 1975 [Spillane, 1975] and the basic idea was to use a keyboard to automatically identify individuals. In the preliminary report dressed by Gaines et al. [Gaines et al., 1980], seven secretaries typed several paragraphs of text and the researchers showed that it was possible to differentiate users by their typing habits [Migdal and Rosenberger, 2019].

Table 3.5 gives an overview of keystroke dynamics relative works on user identification context.

TABLE 3.5: An overview of keystroke dynamics: relative studies, performance metrics in controlled environments, and static vs. dynamic types [Banerjee and Woodard, 2012]

Study	Features	Classification	Subjects	Samples	Identification Rate
[Samura and Nishimura, 2009]	Latency, Key hold time	Euclidean dist.	112	-	90.7%
[Lv et al., 2008]	Key Pressure	Statistical classifiers	50	3000	6.6%
[Rybnik et al., 2008]	Latency, Key hold time	Statistical	37	-	72.97%
[Jin et al., 2008]	Latency	Statistical	11	-	76%
[Bergadano et al., 2003]	Latency, Trigraph/N-graph	Distance measure	40	364	90%

TABLE 3.6: An overview of user authentication in keystroke dynamics: Neural Network-based approaches [Banerjee and Woodard, 2012]

Study	Features	Classification	Subjects	Samples	EER
[Andrean et al., 2020]	Latency, Trigraph/N-graph	MLP	51	400	16.14%
[Alpar, 2017]	-	Gauss-newton based neural network	13	780	4.1%
[Harun et al., 2010]	Latency	NN, dist. classifier	15	150	22.9%
[Revett et al., 2007]	Latency, Trigraph/N-graph	Specht Probabilistic NN	50 -		4%

Keystroke dynamics is a two-factor authentication scheme as we combine the knowledge of a password and the way of typing. In case of an attack, it can be revoked by changing the password. Nevertheless, some studies showed it is possible to profile users on the Internet (gender recognition, age category) [Idrus et al., 2014] without the consent or awareness

of the users [Migdal and Rosenberger, 2019]. Many studies have shown that it is possible to authenticate an individual by typing on a mobile device or keyboard. Table 3.6 lists the main works undertaken by researchers to develop neural network-based authentication systems. We can note that these works are tested on small databases. In this study, we want to investigate how recent Deep Learning methods can improve these results on a more important dataset.

Authentication factors can be organized into three categories depending on how users authenticate themselves on a system. It is about knowledge (e.g., a password, a personal identification number, etc.), token (for example a card access sent to a cellphone), and biometrics (physical biometrics, behavioral biometrics) [Bhana and Flowerday, 2020]. In traditional authentication systems, authentication tokens are usually presented as the forms of what we know (e.g., user name and password), what we have (e.g., smartcard or key token), and what we are (e.g., biometrics derived from physiological signals or user behaviors) [Yeh et al., 2018].

These two behavioral biometrics modalities have been studied in the literature but not in a constrained context (i.e., we consider that we know what a person types on the keyboard, or that we know his activity) and a realistic context (i.e., we consider that we do not know what a person types on the keyboard, or that we do not know his activity). We would like to analyze how efficient can be a generic approach for user identification/authentication in such contexts.

3.3 Proposed architecture

The distribution of our contribution is shown in Figure 4.13. It comprises 4 important components, as follows:

1. Transformation of time series into signal-to-image representations
2. Features generations/extractions
3. Classical machine learning and deep learning (for transfer methods)

4. Identification and authentication (for applications)

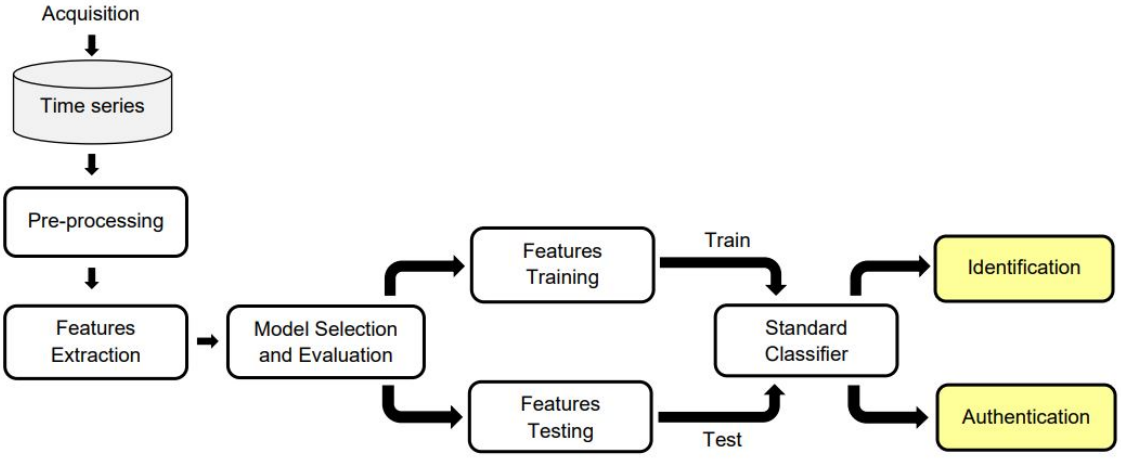


FIGURE 3.4: General overview of our proposed generic system.

3.3.1 Features generation

Time series analysis in the frequency domain plays an essential role in signal processing.

The same is true for image analysis in the frequency domain, which plays a key role in computer vision and was even part of the standard pipeline in the early days of Deep Learning [Vasconcelos et al., 2021]. We propose a method by transforming the time series (behavioral biometric signal) into an image, *i.e.*, we convert the behavioral biometric vector (represented as time series) of size $1 \times m$, into a matrix of size $n \times n$ such that: $m = \frac{n(n-1)}{2}$.

This is done though *squareform()* [2] function in MatLab. The matrix is displayed with *imagesc()* [3] function in MatLab which shows an image with scaled colors. We finally have a 2D color image in RGB format.

The time series to image representation is done through the formal representative equation:

- Let X be a square n-by-n symmetric distance matrix,
  $v = squareform(X)$ returns a $\frac{n(n-1)}{2}$ (i.e. binomial coefficient n choose 2, $\binom{n}{2}$) sized

[2] `https://fr.mathworks.com/help/stats/squareform.html`
[3] `https://fr.mathworks.com/help/matlab/ref/imagesc.html`

vector v where

$$v\left[\binom{n}{2} - \binom{n\ -\ i}{2} + (j - i - 1)\right] \tag{3.1}$$

is the distance between distinct points i and j. If X is non-square or asymmetric, an error is raised.

- Let v a $\frac{n(n-1)}{2}$ sized vector for some integer $n \geq 1$ encoding distances,
  $X = squareform(v)$ returns a n-by-n distance matrix X. The $X[i, j]$ and $X[j, i]$ values are set to:

$$v\left[\binom{n}{2} - \binom{n\ -\ i}{2} + (j - i - 1)\right] \tag{3.2}$$

  and all diagonal elements are zero.

- Example:
  Let v a vector define as v = (1 2 3 4 5 6)
  The squareform of vector v is matrix X.

$$squareform(v) = \mathrm{X} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 4 & 5 \\ 2 & 4 & 0 & 6 \\ 3 & 5 & 6 & 0 \end{pmatrix}$$

This transformation is done on each trial sub-database separately and on the fusion of sub-databases in order to build a new database of images. 70% of the obtained images were used for training (enrollment) and the remaining 30% were used for validation (verification) on the different deep models used for user classification and extraction of features. The matrix representation for user's behavioral typing time series transformation is illustrated in Figure 3.5. This figure presents detailed step-by-step instructions for creating a dataset containing raw parameters computed using statistical tools. The built dataset is used for signal-to-image processing. We finally have a 2D image in RGB format as depicted in Figure 3.6.

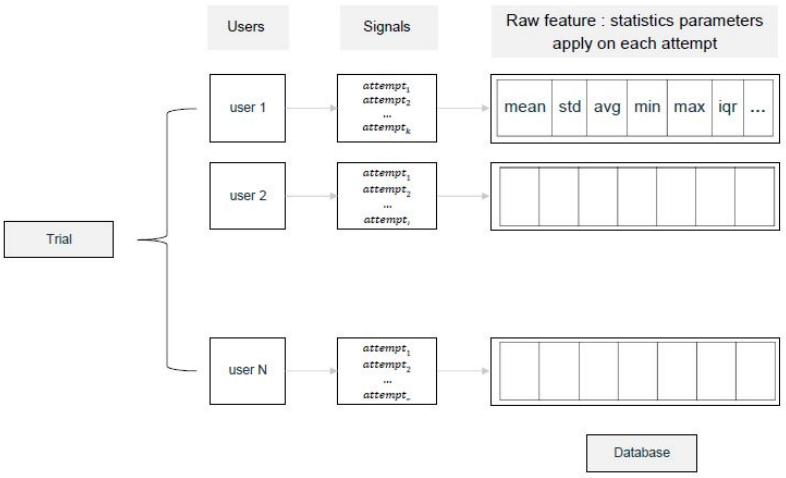


FIGURE 3.5: Diagram of the raw characteristic of the database based on statistics.

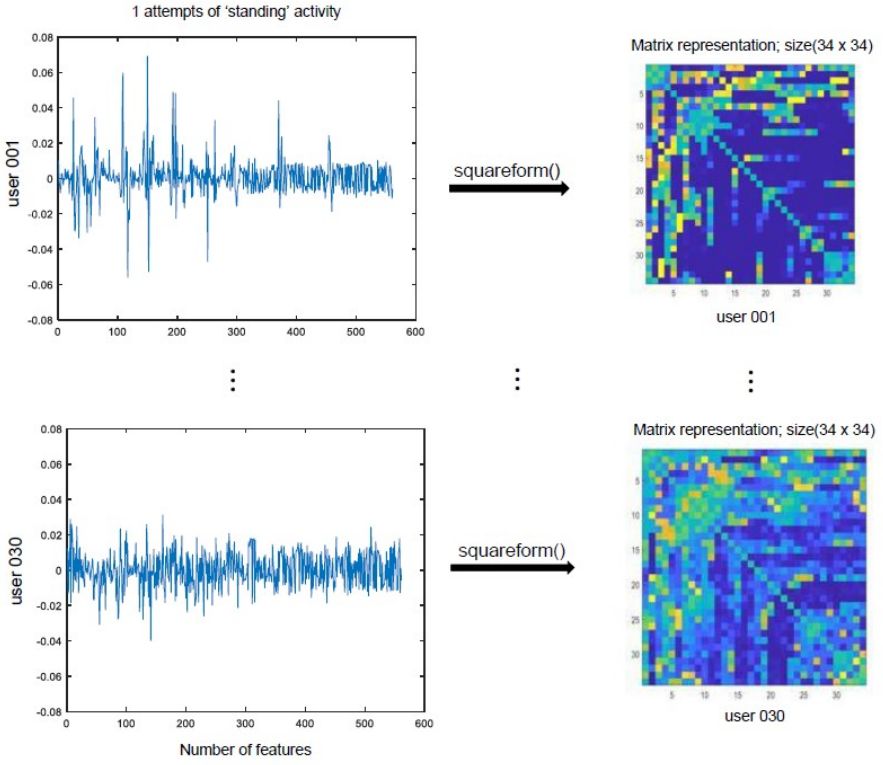


FIGURE 3.6: Examples of the obtained results when the signal-to-image transformation is applied.

3.3.2 Identification

In this work, we consider behavioral biometric data represented as time series. We want to assess the capability of Machine Learning methods to obtain good performance on

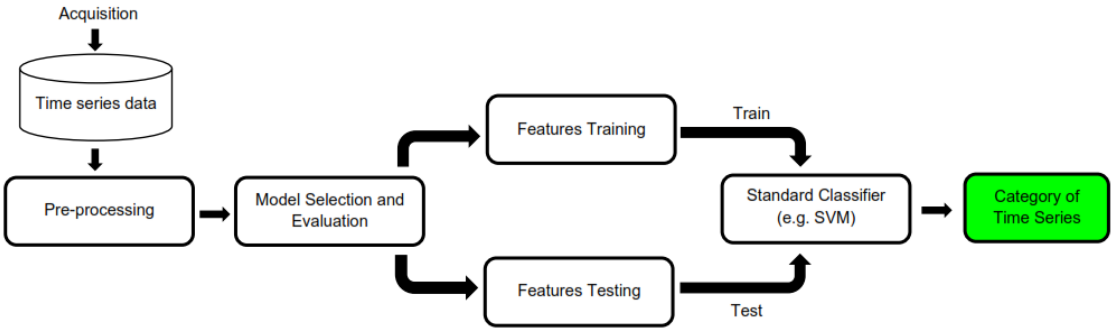


FIGURE 3.7: Architecture of the identification system.

user identification. We use two main approaches. The first approach consists of using classical classifiers by using raw data as input. We considered the following Machine Learning classifiers: SVM (Support Vector Machine), NN (Neural Networks), RF (Random Forest), AdaBoost (Adaptive Boosting), LR (Linear Regression), Naive Bayes, k-NN (k-Nearest Networks), and Stacking (fusion of classifiers). Second, we intend to test Deep Learning techniques that consist of optimizing the representation of raw data to enhance user identification. We considered the following architectures: FCN and ResNet used in [Fawaz et al., 2019]. We used Machine Learning and Deep Learning techniques because they are tried-and-tested techniques that have demonstrated their ability to solve binary and multi-class classification problems.

We describe the proposed generic system that can be applied to any behavioral biometrics modalities in Figure 3.7 for identification purposes. It is composed of different steps namely data acquisition, data preprocessing, feature representation, and machine learning/identification. The step-by-step instructions are to verify individual identities through our framework starting from the signal (computing the time series signal into a color image). The transformation is performed on each attempt of each user for each human activity.

We use the Orange data mining software. It is an open-source data visualization, Machine Learning, and Data Mining toolkit [Demšar et al., 2013]. It features a visual programming front-end for explorative data analysis and interactive data visualization, and can also be used as a Python library. The software was developed by the University of Ljubljana under GNU General Public License in 1997.

We designed a data workflow composed of widgets (data processing unit) with Orange as depicted by Figure 3.8. This workflow can be used for any behavioral biometrics data and generates performance metrics. Import and preprocessing database sub-workflow is illustrated in *Block 1*. *Block 2* represents seven widgets associated with the following 7 classifiers: SVM, NN, RF, AdaBoost, Logistic Regression, Naive Bayes, and k-NN [Demšar et al., 2013, Subasi et al., 2020].

Using the stacking widget in *Block 3* permits the fusion of the different classifiers. In order to evaluate the performance of the defined workflow, it is preprocessed by the Test and Score widget. We use *Block 4* for the evaluation of this system by computing the confusion matrix and ROC curve. *Block 5* allows a visual inspection of the obtained predictions.

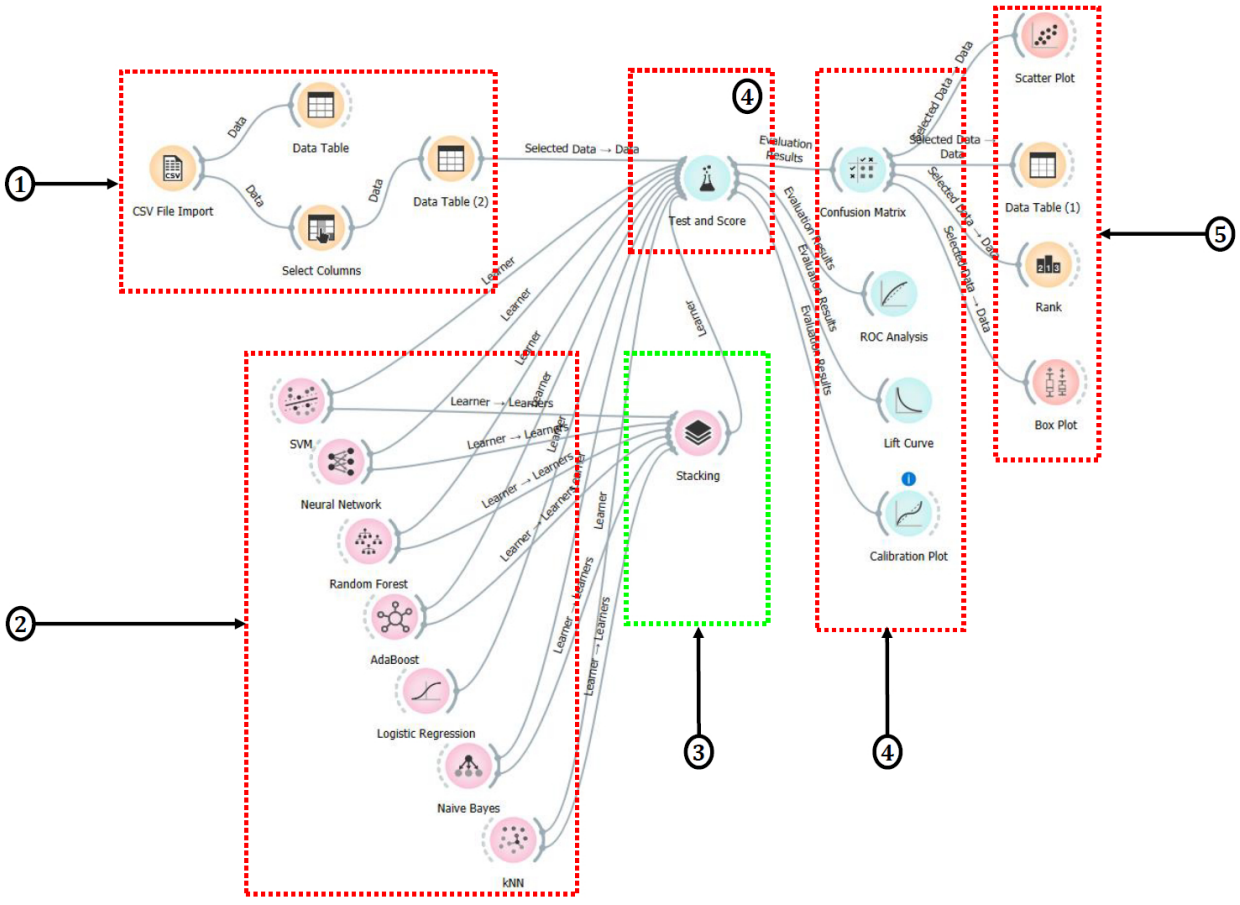


FIGURE 3.8: Global workflow for user identification from behavioral data.

3.3.3 Authentication

We describe the proposed generic system that can be applied to any behavioral biometric modalities in Figure 3.9. It is composed of different steps namely data collection, feature representation, and deep learning/verification.

Deep learning algorithms have been used in the last decade in several fields and are becoming more and more widespread [Maheshwary et al., 2017, Aversano et al., 2021]. One advantage of using such approaches relies on its capabilities to provide relevant features at deeper layers which can be used as feature vectors by any dissimilarity measure. In this work, we generate deep feature vectors by transfer learning from different deep networks namely ResNet-101, DarkNet-53, GoogLeNet, ShuffleNet, DenseNet-201, and SqueeseNet. In the literature, these models were firstly pre-trained on the ImageNet dataset [4] and they are the most recent successful Deep Learning architectures for image classification [Azizi et al., 2023] and can be used for an authentication context since authentication can be considered as the result of a binary classification problem (genuine or impostor).

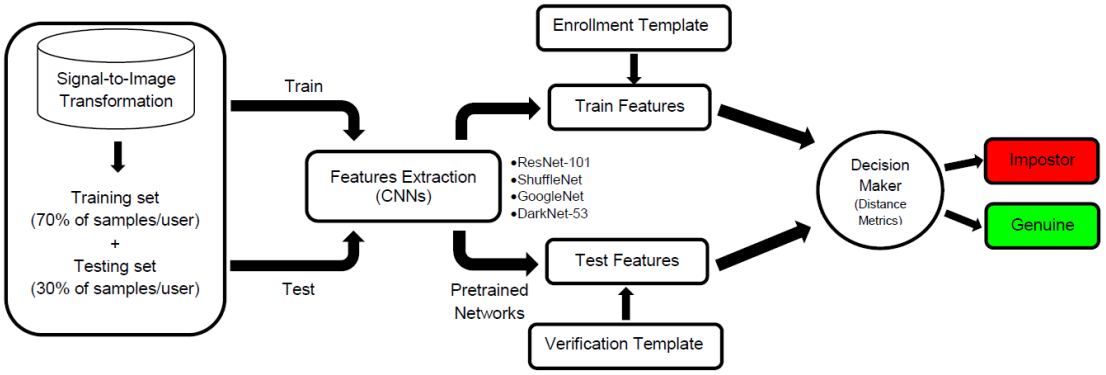


FIGURE 3.9: Architecture of the authentication system.

Table 3.7 summarizes the architecture and the optimization hyper-parameters for the used deep networks where the network depth is defined as the largest number of sequential convolutional or fully connected layers on a path from the input layer to the output layer. Figure 3.10 reports top-1 one-crop accuracy versus the number of operations required for a single forward pass in the most popular neural network architectures as an illustration. The inputs to all these networks are RGB images. We used these convolutional networks

[4] `https://image-net.org`

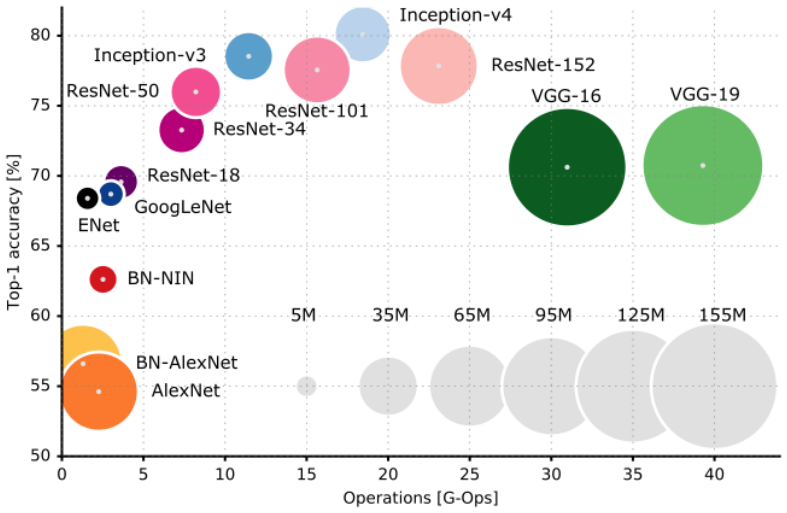


FIGURE 3.10: Illustration of Deep Learning architectures (source: `https://www.topbots.com/a-brief-history-of-neural-network-architectures/`).

to build a features vector as output which is then compared with is used as a reference/test template.

3.3.4 Matching scores

Deep architectures as previously explained generate feature vectors that can be used as reference/test templates. We need a matching algorithm to compare them and make the verification decision. Many distance metrics can be used to compute a distance score [Migdal, 2019b] between a reference (x_s) and a sample (x_t) such as:

- The Minkowski distance

$$d = \sum_{j=1}^{n} |x_{sj} - x'_{tj}| \tag{3.3}$$

- The Euclidean distance

$$d^2 = (x_s - x_t)(x_s - x_t)' \tag{3.4}$$

- The Cosine distance

$$d = 1 - \frac{x_s x'_t}{\sqrt{(x_s x'_s)(x_t x'_t)}} \tag{3.5}$$

Once we obtain a matching score, we decide if the user is authenticated by a simple thresholding approach (accept when the score is upper a given threshold).

To illustrate the interest of the proposed architectures, we evaluate their performance on two behavioral biometrics modalities.

TABLE 3.7: Architectures and optimizations hyper-parameters for the Deep Learning approaches

Models	#Layers	#Depth	Image Input Size	Activate	Normalize	Algorithm	Loss	#Epochs	#Batch	#Learning rate
ResNet-101	347	101	224-by-224	ReLU	Batch	SGDM	cross-entropy	300	10	0.001
DarkNet-53	184	53	256-by-256	ReLU	Batch	SGDM	cross-entropy	300	10	0.001
GoogleNet	144	22	224-by-224	ReLU	Batch	SGDM	cross-entropy	300	10	0.001
ShuffleNet	172	50	224-by-224	ReLU	Batch	SGDM	cross-entropy	300	10	0.001
DenseNet-201	708	201	224-by-224	ReLU	Batch	SGDM	cross-entropy	300	10	0.001
SqueezeNet	68	18	227-by-227	ReLU	Batch	SGDM	cross-entropy	300	10	0.001

3.4 Experimental protocol

We draw in this part the experimental protocol we follow in this work. We detail the used biometric dataset and the performance metrics.

3.4.1 Datasets description

3.4.1.1 UCI-HAR database

We use in this work the UCI-HAR database [Anguita et al., 2013] which was collected with data from 30 people aged between 19 and 48 years. Each person performed 6 physical activities such as *sitting*, *standing*, *laying walking*, *walking upstairs* and *walking downstairs*. The data were collected from a *Samsung Galaxy S II* mobile phone handset using the accelerometer and gyroscope (3-axial raw signals with *tAcc-XYZ* and *tGyro-XYZ*) sensors at a frequency of 50Hz. The collection was obtained with the smartphone located at the user's waist. All steps of data collection were recorded and the data was manually labeled. UCI-HAR contains $10,299$ samples.

Table 3.8 presents the activities, the abbreviation of each activity, the proportion of activity samples, and their descriptions. For each signal of each activity (trials), the signal-to-image transformation (as mentioned in section 3.3) is applied to obtain a 2D color image. To

TABLE 3.8: Activities, sample number of each activity and their descriptions on UCI-HAR dataset [Anguita et al., 2013].

Activity	Abbreviation	No. of Samples	Each Human Activity ratio	Description
Laying	lyx	1722	16.72%	Subject sleeps or lies down on a bed
Sitting	six	1544	14.99%	Subject sits on a chair either working or resting
Standing	stx	1406	13.65%	Subject stands and talks to someone
Walking	wlx	1777	17.25%	Subject goes down multiple flights
Walking Downstair	wdn	1906	18.51%	Subject goes down multiple flights
Walking Upstairs	wup	1944	18.88%	Subject goes up multiple flights

the best of our knowledge, such transformation with the *squareform()* function applied to the UCI-HAR dataset does not exist in the literature up to now. Among the transformed samples of each user, 70% out of 100% samples (attempts per subject) are used for the training set, and the remaining 30% for the validating set.

3.4.1.2 GREYC-NISLAB database

The GREYC-NISLAB dataset [Syed Idrus et al., 2013] for keystroke dynamics is constituted of five passwords entered by 110 users. There were 10 samples per password per user for each way of typing. The best password is a sentence according to experts [Idrus et al., 2013]. In total, we have $5,500 = 110 \times 10 \times 5$ keystroke dynamics samples, we believe this dataset is significant.

TABLE 3.9: Description of passphrases used in the GREYC-NISLAD dataset.

Password	Description	Size	Features
P1	leonardo dicaprio	17-char	64
P2	the rolling stones	18-char	68
P3	michael schumacher	18-char	68
P4	red hot chilli peppers	22-char	84
P5	united states of america	24-char	92
P_T	fusion of features (P1+P2+P3+P4+P5)	99-char	376

For keystroke dynamics modalities, 5 passphrases were presented to users as shown in Table 3.9, which are between 17 and 24 characters (including spaces) long, chosen from some of the well-known or popular names or artists (known both in France and Norway), denoted P1 to P5. The GREYC Keystroke software has been used to capture biometrics data. P_T denotes the fusion of the 5 passwords (fusion of features) [Piugie et al., 2021]. Keystroke dynamics databases are very tedious to realize. One of the biggest advantages

of using the GREYC-NISLAB database is that we have several passwords for the same users. To the best of our knowledge, such a database does not exist in the literature up to now.

3.4.2 Performance metrics used for identification

The performance evaluation metrics used in this work for the identification task are CA, P, R, AUC, and CMC where T_P, T_N, F_P, and F_N are respectively true positive, true negative, false positive, and false negative. We provide in the following the definitions of all these metrics:

- Classification Accuracy (CA)

  For a given test dataset, the ratio of the number of samples correctly classified to the right user by the classifier compared to the total number of samples. This metric (also called rank 1 accuracy) formula is given by equation 3.6.

$$CA = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{3.6}$$

- Precision (P)

  The ratio of the number of correctly identified positive samples (corresponding to the right individual) to the total number of samples identified as positive in the identified sample. The computation formula is given by equation 3.7.

$$P = \frac{T_P}{T_P + F_P} \tag{3.7}$$

  The precision score can be the number of correct predictions made divided by the total number of predictions made.

- Recall (R)

  The recall rate is the ratio of the number of positively identified individuals correctly identified to the total number of positive samples in the total sample used. The recall

rate is how many positive samples are identified. The formula drawn by 3.8.

$$R = \frac{T_P}{T_P + F_N} \tag{3.8}$$

- Area Under the ROC Curve (AUC)

  The AUC computes the area under the ROC curve when plotting the precision versus the recall value. It should be as high as possible (the maximal value is 100% or 1).

- Cumulative Match Characteristic (CMC) Curve

  The CMC curve is a method of showing the measured accuracy performance of a biometric system operating within an identification task. Templates are compared and ranked based on their similarity. The CMC indicates how often the biometric subject template appears in the ranks (1, 5, 10, 100, etc.) based on the match rate. A CMC compares the rank (1, 5, 10, 100 etc.) versus identification rate.

The authentication/verification stage involves acquiring and processing raw data to create a biometric template, which is then compared to reference templates in the dataset. A matching algorithm is used to determine the similarity between the biometric sample and existing reference templates. Scores are calculated based on features extracted from deep networks, and three distance metrics described in subsection 3.3.4 are applied to evaluate the degree of similarity between the activity of each user.

Two important error rates are used to assess the performance of a biometric authentication system according to ISO19795 [19795-1, 2021]: 1) False Match Rate (FMR) and 2) False Non-Match Rate (FNMR):

- The FMR is the proportion of a specified set of completed non-mated comparison trials that result in a comparison decision of *match*,
- The FNMR is the proportion of completed mated comparison trials that result in a comparison decision of *non-match*.

The Equal Error Rate (EER) is obtained when the biometric decision threshold is set to have the FMR value equal to the FNMR one as illustrated in Figure 3.11. It can be

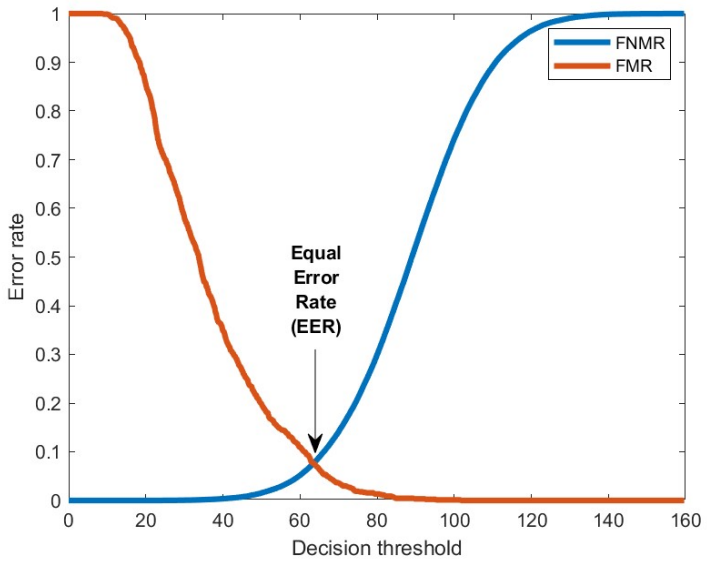


FIGURE 3.11: Relationship between FMR, FNMR and EER (source [Piugie et al., 2022]).

TABLE 3.10: Optimization's hyperparameters for the Deep Learning approaches

Methods	Algorithm	Valid	Loss	Epochs	Batch	Learning rate
FCN	Adam	$\text{Split}_{70\%}$	Entropy	250	10	0.001
ResNet	Adam	$\text{Split}_{70\%}$	Entropy	250	10	0.001

seen as a compromise between usability and security. The goal of a matching algorithm is to minimize this value. The lower the value of EER, the better the performance of the authentication system is. This error rate is the most commonly used in the literature to evaluate the performance of biometric systems. In this work, we evaluate the performance of the proposed architecture in terms of EER. The performance rate used for user identification is the Classification Accuracy (CA) where the formula is given by Equation 3.6.

3.4.3 Pre-trained models

As previously mentioned, we used pre-trained models to experiment with user authentication. We compare the different architectures used. These are the following networks: ResNet-101, ShuffleNet, GoogleNet, DenseNet-201, SqueezeNet, and DarkNet-53 for authentication [Piugie et al., 2022].

TABLE 3.11: Models parameters for the classical approach

Model	**Parameters**	**Regression loss / Activate**	**Optimization Parameters**	**Maximal number of iterations**	**Regularization**
LR	–	–	–	–	Ridge (L2)
SVM	Cost: 1	ϵ: 0.1	Kernel: RBF	–	–
k-NN	No. of neighbors: 5	Metric: Euclidean	Weight: Uniform	–	–
AdaBoost	No. of estimators: 50	Learning rate: 1.0	Regression loss function: Linear	–	–
RF	No. of trees: 10	–	–	–	–
NN	Neurons in hidden layers: 200	ReLU	solver: Adam	Max_iter: 500	α : 0.0001
Naive Bayes	–	–	–	–	–
Stack	–	–	–	–	–

3.4.4 Classifiers parameters for identification

TABLE 3.13: Architecture's hyperparameters for the Deep Learning approaches

Methods	**#Layers**	**#Conv**	**#Invar**	**Normalize**	**Pooling**	**Feature**	**Activate**	**Regularize**
FCN	5	3	4	Batch	None	GAP	ReLU	None
ResNet	11	9	10	Batch	None	GAP	ReLU	Dropout

In this subsection, we list the different parameters used through the learning methods. Table 3.11 gives model parameters defined in Orange for the classic basic approach. Tables 3.13 and 3.10 show respectively the architecture and the optimization hyperparameters for the Deep Learning approaches. A model checkpoint procedure was performed either on the training set or a 30% validation set (split from the 70% training set). This means that if the model is trained for 250 epochs, the best one on the validation set (or the train set) loss is chosen for evaluation. This characteristic is included in Table 3.10 under the *valid* column. In addition to the model checkpoint procedure), models in Table 3.13 were initialized randomly using Glorot [Glorot and Bengio, 2010] which is a uniform initialization method. Models were optimized using a variant of Stochastic Gradient Descent (SGD) such as Adam [Kingma and Ba, 2014] and AdaDelta [Zeiler, 2012].

We add that for FCN and ResNet proposed in [Wang et al., 2017], the learning rate was reduced by a factor of 0.5 each time the model whose training loss has not improved for 50 consecutive epochs (with a minimum value equal to 0.0001). One final note is that we have no way of controlling the fact that those described architectures might have been overfitted

for the UCI-HAR and GREYC-NISLAB archive and designed empirically to achieve a high performance, which is always a risk when comparing classifiers on a benchmark [Bagnall et al., 2017]. We, therefore, think that challenges where only the training data is publicly available and the testing data are held by the challenge organizer for evaluation might help in mitigating this problem.

3.5 Experimental results

3.5.1 User identification

All experiments in this research work were conducted in the same environment which is composed of: Windows 10 Pro operating system, Intel(R) Core (TM) CPU @ 1.8GHz, 8GB RAM, and TensorFlow 2.2.0 - G.P.U. on Python 3.8.2. Then, the system was tested using the following datasets:

- UCI-HAR database,
- GREYC-NISLAB database.

3.5.1.1 Classical machine learning

The models used in the classical approach and developed for identification in our system proposed in Figure 3.8 are evaluated in terms of AUC, CA, P, and R on the test set (30% of user data in the database).

Table 3.14 gives classification results from UCI-HAR database for the human activity modalities. Table 3.15 gives classification (best models) results from GREYC-NISLAB databases for keystroke dynamics modalities. On this last modalities, the 7 other classical Machine Learning models were also trained, but the *Stacking* model provided the best identification score for each sub-dataset P1, P2, P3, P4, and P5. We can see in these two tables that the fusion of classifiers (Stack) provides very good results on both datasets.

In Table 3.15, the best rate is achieved by merging the features of the sub-bases (P1, P2, P3, P4, and P5), and the best model is also the *Stacking* (that is taking a last estimator to learn and predict the final result according to the seven previous predictions, which further improves the general performance).

TABLE 3.14: User identification performance metrics with Orange workflow on HAR dataset from human activities.

Model	**AUC (%)**	**CA (%)**	**P (%)**	**R (%)**
Stack	**99.65**	**93.89**	**93.90**	**93.89**
Neural Networks	98.97	87.75	87.73	87.75
Random Forest	98.21	85.78	85.89	85.78
kNN	97.56	81.40	82.07	81.40
AdaBoost	89.33	81.06	81.25	81.06
SVM	96.57	78.45	80.22	78.45
Logistic Regression	96.76	78.38	78.40	78.38
Naive Bayes	79.24	41.33	48.49	41.33

TABLE 3.15: User identification performance metrics with Orange workflow on GREYC-NISLAB from keystroke dynamics.

Password database	**Model**	**AUC (%)**	**CA (%)**	**P (%)**	**R (%)**
P1	Stack	96.22	63.09	63.67	63.10
P2	Stack	99.08	69.73	72.15	69.73
P3	Stack	98.49	63.91	66.10	63.91
P4	Stack	99.22	77.73	79.64	77.73
P5	Stack	98.56	83.73	84.30	83.73
P_T	**Stack**	**99.99**	**98.10**	**98.3**	**98.10**

After merging the five passwords, Table 3.16 gives the results of classification accuracy when identifying a person with knowledge of how they type on a keyboard.

TABLE 3.16: User identification (based on user knowledge) performance with Orange workflow on GREYC-NISLAB dataset.

Targets	**CA(%)**
Subject	98.18
Handedness	99.27
Gender	88.73
Age	70.73

We can therefore identify a person knowing his/her typing style, the type of hand used, his/her gender, and his/her age with a clear classification accuracy of 98.18%, 99.27%

88.73% and 70.73% respectively by using the Classical Machine Learning workflow we developed as related in Table 3.16.

3.5.1.2 Deep Learning techniques

The used models for the Deep Learning approach are evaluated in terms of CA, P, and R. The performance metrics for both modalities are provided in Tables 3.17. We see that from our two behavioral modalities, ResNet deep classifier performs better than FCN on user identification. The obtained results are worst than classical Machine Learning (the datasets are probably not enough large to obtain better results).

TABLE 3.17: UCI-HAR and GREYC-NISLAB deep performance metrics

Dataset	**Classifier name**	**CA (%)**	**P (%)**	**R (%)**
HAR	ResNet	87.05	87.20	86.73
	FCN	68.58	80.09	68.24
P_T	ResNet	80.30	82.94	82.23
	FCN	76.06	78.95	79.01

3.5.1.3 Discussion

This section presented a study on how it is easy to define a user identification method given behavioral biometrics data.

The UCI behavior recognition dataset is collected by measuring the six daily behaviors of the 30 participants. The experiment uses a three-axis embedded accelerometer and a gyroscope operating at 50 Hz. The three component values of the accelerometer and the gyroscope are obtained separately, and the data dimension is 561. The tested behaviors of the participants included *walking*, *walking upstairs*, *walking downstairs*, *sitting*, *standing*, and *laying*.

Obtained results in Table 3.14 for user identification through his/her activity shows us that the stacking model performs well with a precision score of 93.90% than the seven other models in Orange workflow. As depicted in Table 3.15, we have the best precision

with the staking model, and this with all the keystroke P1 (63.67%), P2 (72.15%), P3 (66.10%), P4 (79.64%) and P5 (84.30%) databases.

With a fusion of features (P_T), we obtain 98.30% score of precision. This means that the larger our behavioral database, the higher the classification score. So far considering these model comparisons on different behavioral biometrics databases, it comes out that the stacking model performs well. This shows that methods evaluated in this work are highly successful in identifying users from behavioral biometrics data.

Another study that used deep Neural Networks notably ResNet and FCN achieved a precision rate of 87.20%, 80.09% respectively for HAR dataset and 82.94%, 78.95% for the fusion of dynamic keystrokes feature P_T dataset. In view of these comparisons, we say that ResNet performs more than FCN for both behavioral modalities in detecting the user than the other models described in [Fawaz et al., 2019].

We tested this work on complicated databases because the idea was to recognize a person just from his/her behavior. However, in the literature, there is few work on the identification of individuals using the dynamics keystroke with low classification accuracy.

TABLE 3.18: HAR and keystroke rank scoring

Dataset	**rank1 (%)**	**rank2 (%)**	**rank3 (%)**
HAR	**90.34**	**94.06**	**95.61**
PT	**98.09**	**98.73**	**99.00**
P1	63.36	72.64	77.73
P2	71.00	79.73	83.45
P3	66.18	73.27	77.18
P4	76.64	83.09	86.18
P5	83.27	89.73	91.00

We compute the rank 3 identification accuracy in order to know how many times the individual has been identified with the 3 most likely. The obtained results conducted on the two datasets show that the classification rate obtained by the best Machine Learning model (Stacking) is 93.90% for human activity and 98.10% for the fusion of the keystroke dynamics features. Using the cumulative matching characteristic (CMC) curve, we show that for our best model (Stacking), an individual appears in rank 3 (the three most likely) for a match rate of 95.61% for human identity recognition and a matching rate of 99.00% for the fusion of keystroke dynamics identification.

Interpreted by Table 3.18, Figures 3.12 and 3.13 are used to compare the performance of the biometric identification system. The depicted curves represent the values of the identification rank and the probabilities of a correct identification less than or equal to these values, respectively on the x-axis and y-axis.

Figure 3.12 allows us to observe that the probability of identifying one person is 90.34%, the probability of identifying 2 persons is 94.06% and the probability of identifying 3 persons is 95.61% in the UCI-HAR database. Figure 3.13 is interpreted by Table 3.18. The database fusioned P_T gives 99.00% as rank 3 score. The fourth column of the Table 3.18 gives the rank 3 for each keystroke P1 to P5.

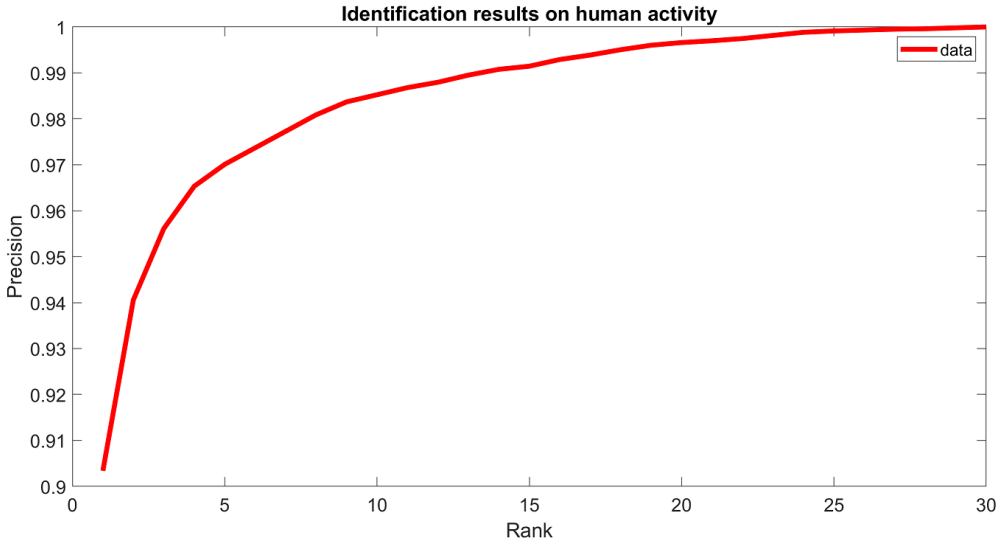


FIGURE 3.12: CMC **HAR** curve of Stacking model in Orange workflow

Ensuring strong security for identification and keeping privacy is key when developing and deploying any new technology. This work shows that Machine Learning solutions, specifically Stacking, can be a reliable help for Identification and can be used through multi-identification factors solutions to reach a very high level of confidence. Identification with two factors might be appropriate in some circumstances, but too much in others. Behavioral with Machine Learning and Deep Learning enables multimodality authentication without increasing the burden on the user. Moreover, it is essential to build confidence and trust, especially for technologies that process our personal data. Machine Learning and Deep Learning using behavioral for identification can be a key to success but also a way to classify individuals into targets or groups, without the consent of the users.

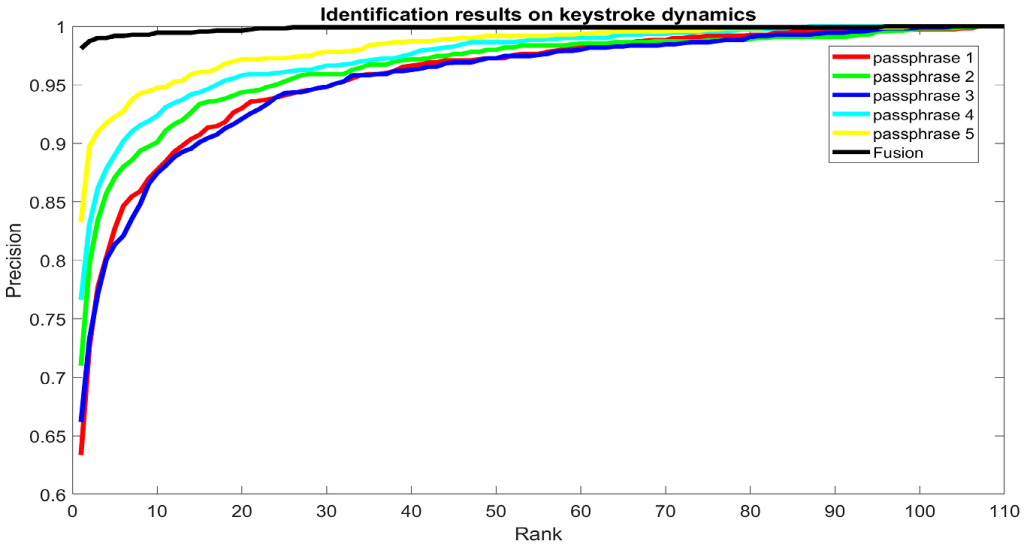


FIGURE 3.13: CMC **Keystroke Dynamics** curve of Stacking model in Orange workflow

To summarize, this part proposes an approach based on multimodal behavioral biometrics for user identification, using both Machine Learning (classical and deep). Two behavioral modalities are studied: human activities on smartphones and keystroke dynamics on laptops. The results show that classical Machine Learning and Deep Learning achieve state-of-the-art performance in time series classification. However, the use of behavioral biometrics raises privacy concerns, as it enables users to be identified as they browse the Internet. Despite this, these frictionless solutions offer opportunities to enhance the user experience and will be explored in future work on authentication.

We present the authentication results obtained for the two studied modalities.

3.5.2 User activity authentication

In this section, we present the experimental results we obtained when using human activities data. We structure them by addressing some questions concerning the performance of the proposed method on such behavioral biometric datasets. Note that we considered 70% of user samples (attempts per subject) for the learning phase and 30% for the testing one.

3.5.2.1 Which performance can we expect on a larger dataset ?

First, we consider all the six activity sub-datasets defined as the UCI-HAR dataset (**Fusion of features**). Table 3.19 draws the obtained results for the user verification task considering the three distances presented in subsection 3.3.4. One can observe that whatever the tested deep architecture, the distance minimizing the EER value is the *Cosine* one. In the rest of the human activity analysis, all results are given considering this distance.

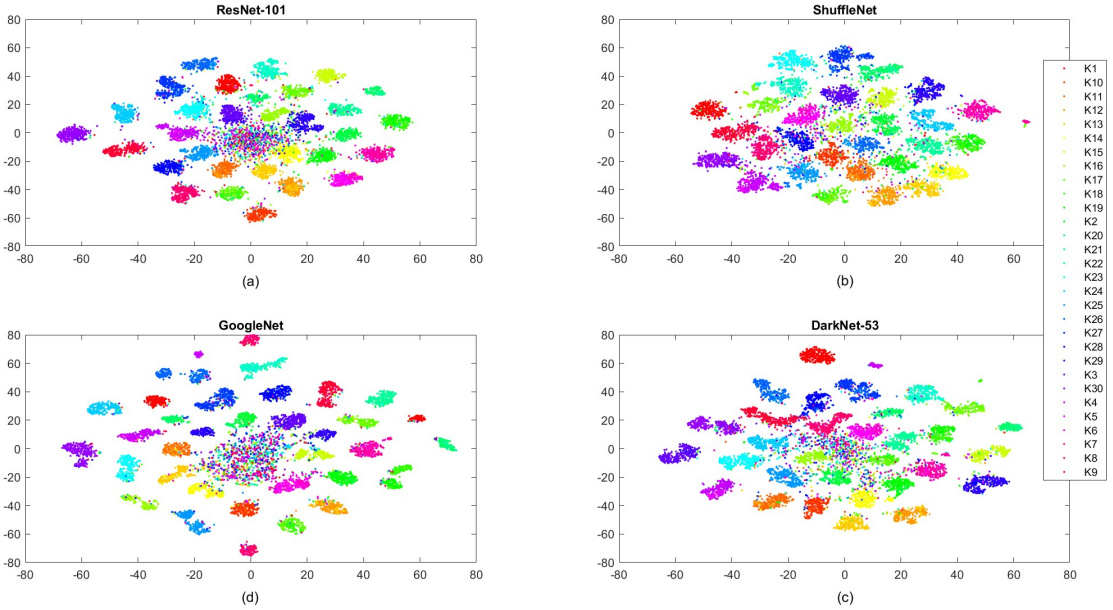


FIGURE 3.14: Visual inspection of deep features projection from (a) ResNet-101, (b) ShuffleNet, (c) DarkNet-53 and (d) GoogleNet.

TABLE 3.19: EER value on HAR dataset for the three tested distances.

Models	$EER_{mananthan}$	$EER_{euclidean}$	EER_{cosine}
ResNet-101	22.69%	17.71%	**12.48%**
ShuffleNet	14.77%	14.63%	**11.57%**
GoogleNet	14.88%	14.56%	**13.52%**
DarkNet-53	17.46%	14.31%	**11.72%**

We visually inspected our four best deep networks by performing a feature projection through the t-SNE (t-Distributed Stochastic Neighbor Embedding) function [5] as shown in Figure 3.14 for each architecture. We observe that the deep features projection forms a nearly distant cluster in ShuffleNet ($EER = 11.57\%$) than DarkNet-53 ($EER = 11.72\%$),

[5] https://scikit-learn.org/stable/modules/generated/sklearn.manifold.TSNE.html

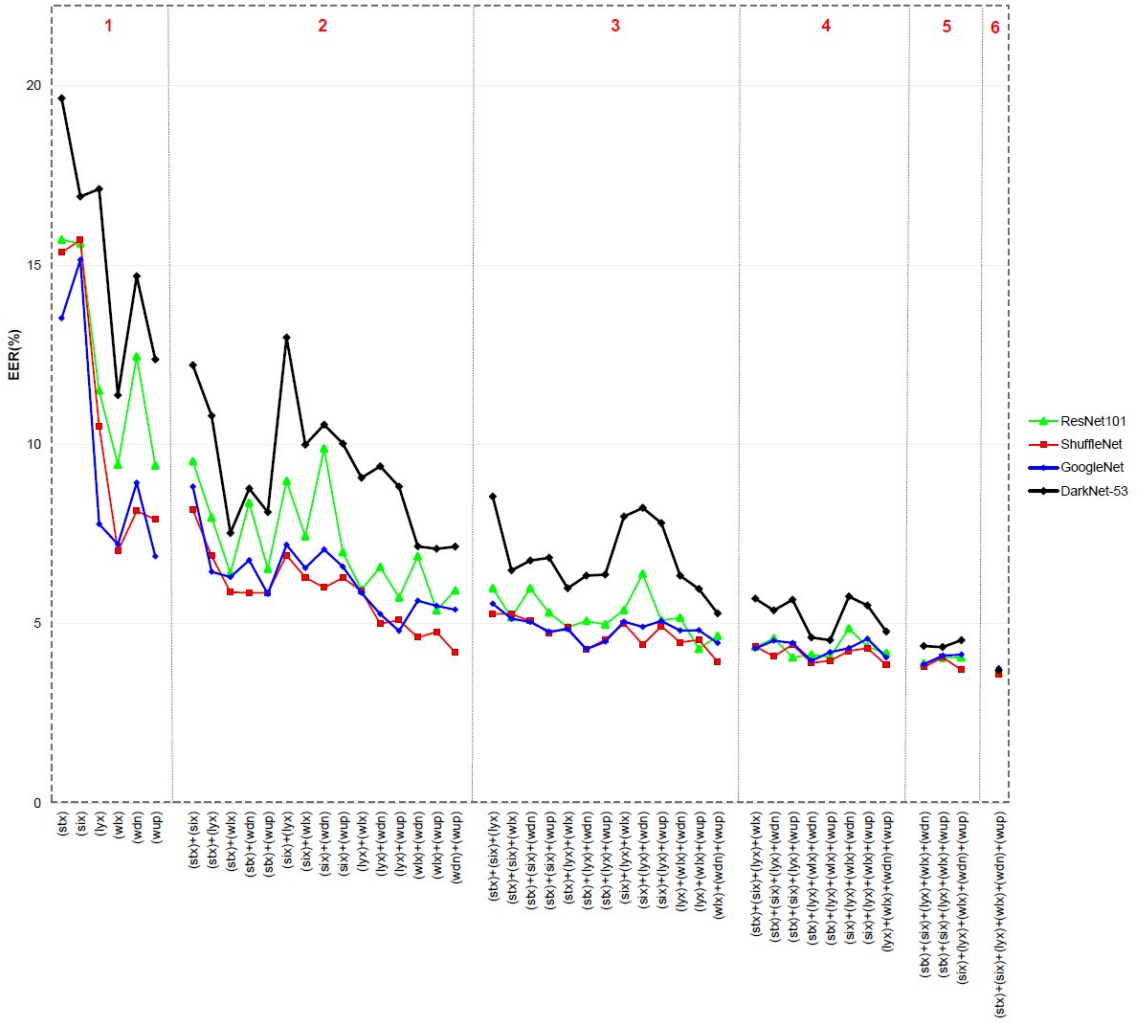


FIGURE 3.15: EER rate on deep architectures for the multi-instance biometric system. In block 1, we have (stx), (six), (lyx), (wlx), (wdn) and (wup) activities. In block 2, we have the fusion of inter and intra-class scores for all the combinations of the two activity pairs. In block 3, a combination of all the couple of three activities possible. In block 4, a combination of all the couple of four activities possible. In Block 5, a combination of all the couple of five activities, and in Block 6 {(stx)+(six)+(lyx)+(wlx)+(wdn)+(wup)}.

ResNet-101 ($EER = 12.48\%$) and GoogleNet ($EER = 13.52\%$). This result is correlated with the fact that ShuffleNet performs better than other networks in terms of EER value.

3.5.2.2 How well can we perform each activity separately?

In this section, we consider each activity separately as shown in Table 3.8 to generate six sub-datasets among others laying (1722 samples for the 30 subjects), sitting (1544 samples for the 30 subjects), standing (1406 samples for the 30 subjects), walking (1777 samples for

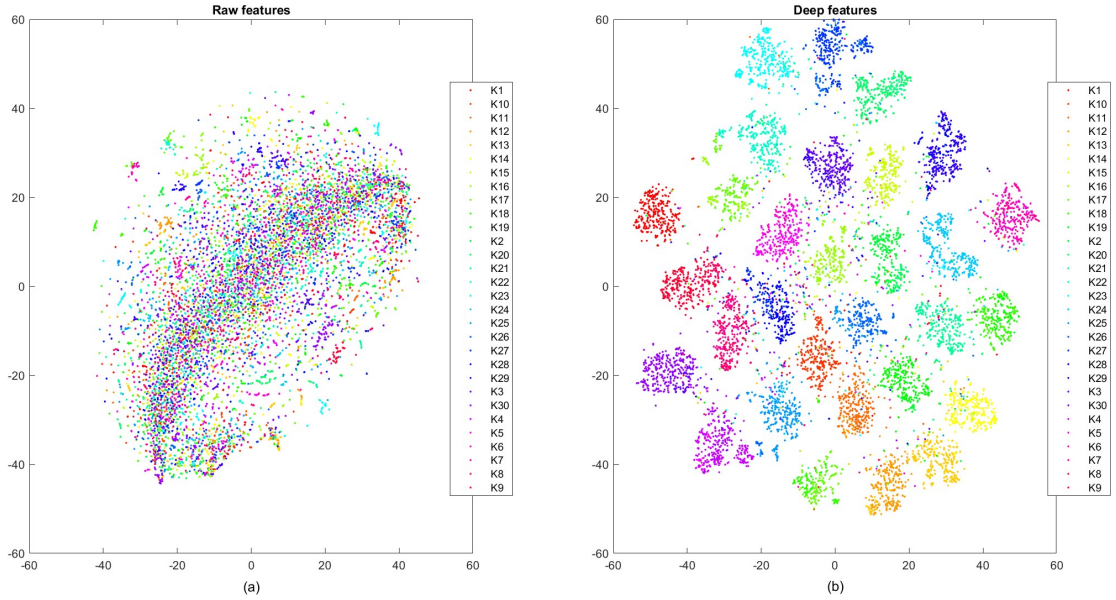


FIGURE 3.16: t-SNE projection of (a) raw features and (b) deep features extracted from the top-performing method (ShuffleNet). The x-axis corresponds to dimension 1, while the y-axis corresponds to dimension 2.

the 30 subjects), walking downstairs (1906 samples for the 30 subjects), walking upstairs (1944 samples for the 30 subjects). We illustrate the four architectures among the six (namely ResNet-101, ShuffleNet, GoogleNet, and DarkNet-53) and we draw the model on each sub-dataset separately.

This is illustrated by the block 1 in Figure 3.15. The best model among the four deep networks for each activity in terms of EER value are: standing (GoogleNet=13.52%), sitting (GoogleNet=15.15%), laying (GoogleNet=07.78%), walking (ShuffleNet=07.02%), walking downstairs (ShuffleNet=08.14%) and walking upstairs (GoogleNet=06.88%). Here, we try to verify one user among the 30 users based on their activities separately.

We note that we do not have the same performance from one activity to another. So, using 70% of samples (attempts per subject) for the generation of the reference template does not provide exceptional results (with an EER value between 06.88% to 19.65%) as shown in the block 1 in Figure 3.15. Obviously, if we had more samples per subject (or by combining activities), we could expect to obtain a better performance.

3.5.2.3 What performance can be achieved if the user performs more than one activity ?

In this part, it is assumed that a person achieved more than one activity to authenticate himself/herself. We merge by summing the legitimate and impostor scores considering the number of samples (activity attempt) per user (**Fusion of scores**). Table 3.20 shows the obtained results if we used all the six activities (*i.e.*, simulating a user achieving 6 activities to be authenticated). ShuffleNet comes out as the best method with an EER score of 03.58% as presented in Table 3.20. ShuffleNet is ahead of ResNet-101 (03.63%), GoogleNet (03.76%) and DarkNet-53 (03.70%).

TABLE 3.20: Performance evaluation on the multi-instance biometric system by fusion of features and scores level on UCI-HAR dataset.

Models (EER_{cosine})	**Fusion of features**	**Fusion of scores**
ResNet-101	12.48%	3.63%
ShuffleNet	**11.57%**	**3.58%**
GoogleNet	13.52%	3.76%
DarkNet-53	11.72%	3.70%

To complete these results, we studied the obtained performance versus the number of activities (*laying*, *sitting*, *standing*, *walking*, *walking downstairs* and *walking upstairs*) achieved by a user in a multi-instance context. Figure 3.15 highlights the EER value obtained for each case:

- If we use 2 activities, we obtain an EER value between $[04.20\% - 12.98\%]$ illustrated by block 2 in Figure 3.15.
- If we have 3 activities, we have an EER value between $[03.94\% - 08.55\%]$ represented in block 3.
- If we use 4 activities, we have an EER value around $[03.85\% - 05.71\%]$ depicted by block 4.
- If we use 5 activities, we have an EER value around $[03.72\% - 04.54\%]$ shown by block 5.

- If we use 6 activities (*laying* + *sitting* + *standing* + *walking* + *walking downstairs* + *walking upstairs*), we get an EER value around $[03.58\% - 03.76\%]$ depicted by block 5. This shows we can decrease easily the EER value for this kind of authentication.

We find that the value of EER obtained by fusion of the scores of each activity decreases for all architectures. It also appears from this work that the more information we have, the better the performance can be, this is not surprising. With a more extensive database, we could expect to get better results (*i.e.* with an EER value very close to 0%) by increasing the number of samples per user [Piugie et al., 2022].

TABLE 3.21: Comparison with other published works on user activity (target = activities).

Dataset	**Author/S (ref)**	**Years**	**Classifiers**	**Accuracy**	**EER**
UCI-HAR (target = activities)	Sanchez *et al.* [Sanchez Guinea et al., 2022]	2022	CNN	99.30%	-
UCI-HAR (target = activities)	Sarkar *et al.* [Sarkar et al., 2022]	2022	Spatial Attention-aided CNN	99.45%	-
UCF sports (target = actions)	Chuan *et al.* [Sun et al., 2015]	2015	GP-based & k-NN	[86.90% – 88.50%]	-
Naturalistic McGill University gait dataset and Osaka University gait dataset	Zhong *et al.* [Zhong and Deng, 2014]	2014	I-vector	[67.5% – 85.0%]	[06.80% – 08.90%]
IXMAS (target = actions)	Korner *et al.* [Körner and Denzler, 2013]	2013	Gaussian process + Histogram intersection kernel	79.00%	-

CMU mocap (target = actions)	Junejo *et al.* [Junejo et al., 2008]	2008	Nearest Neighbour Classifier & SVM	[90.50% – 95.70%]	-

3.5.2.4 Discussion

Due to their ability to perform sensitive operations like mobile banking, communication, and personal data storage, smartphones have become a crucial part of daily life. This has led to a greater need for secure authentication methods to protect critical information from unauthorized access. [Mekruksavanich and Jitpattanakul, 2021].

The purpose of this work is to analyze information from user activity in order to authenticate himself/herself. A comparative analysis of the four best architectures on the UCI-HAR dataset allows us to identify the best performance for continuous authentication. From Table 3.19, we observe that the best results are provided by the ShuffleNet architecture with an EER value equal to 11.57%. Figure 3.16 shows a visual inspection of features (raw versus deep features) projection for ShuffleNet. It shows clearly the good separability of the deep features as depicted in Figure 3.16(b). Multi-instance systems intend to capture samples of two or more different instances of the same biometric characteristics. Table 3.20 shows that for authentication performed on human activity, the best verification scores are obtained on the fusion of scores (EER = 03.58%) as opposed to the fusion of features (EER = 11.57%) on ShuffleNet among the four different deep neural network architectures.

In the literature, several works as shown in Table 3.4 and Table 3.23 have been carried out only on the recognition of activities (where the target is: standing, sitting, laying, walking, walking downstairs, walking upstairs) from the UCI-HAR dataset. Among these works, there are several studies that convert the 1D time series into a 2D image representation and then apply 2D image-based feature extraction technique [Sanchez Guinea et al., 2022, Sarkar et al., 2022, Junejo et al., 2008]. These transformations are not reversible and the related works are typically based on activity or action recognition. However, this work focuses on activity-based user verification.

TABLE 3.22: Comparison with other published works on user activity (target = users).

Dataset	Author/S (ref)	Years	Classifiers	EER
UCI-HAR (target = users)	**[Wandji Piugie et al., 2023]**	**2023**	**ShuffleNet**	**03.57%**
UCI-HAR (target = users)	Mekruksavanich *et al.* [Mekruksavanich and Jitpattanakul, 2021]	2021	DeepConvLSTM	5.10%
Touch gestures data	Patel *et al.* [Patel et al., 2016]	2016	Ten classifiers	07.50%
WISDM	Zhang *et al.* [Zhang, 2019]	2019	Dense Clockwork RNN	18.17%
Gait signal data	Mantyjarvi *et al.* [Mantyjarvi et al., 2005]	2005	Correlation coefficients	07%
Biometric gait data	Muaazz *et al.* [Muaaz and Mayrhofer, 2013]	2013	SVM	33.30%
Mobile gait data	Zhong *et al.* [Zhong et al., 2015]	2015	Nearest neighbor	07.22%

We can compare our results with research works that have been performed on the UCI-HAR dataset in Table 3.22. Mekruksavanich *et al.* [Mekruksavanich and Jitpattanakul, 2021] in 2021 work on Deep Learning approaches for continuous authentication based on activity patterns using mobile sensing. They had obtained for each distinct activity an EER score of 5.10% with the DeepConvLSTM network. By merging the legitimate and impostor scores of each activity, we obtain an EER score of 03.58% with the ShuffleNet network. This means that during a verification scheme, the more activities a user performs, the better it can be authenticated by our framework. To the best of our knowledge, in the literature, there is no work addressing the fusion of scores on the basis of the UCI-HAR dataset.

TABLE 3.23: Comparison with other works on user activity (target = activities).

Dataset	**Author/S (ref)**	**Years**	**Methods**	**Classifiers**	**Accuracy**
UCI-HAR (target = activities)	Sanchez *et al.* [Sanchez Guinea et al., 2022]	2022	signal-to-image (Pattern-to-Pixel)	CNN	99.30%
UCI-HAR (target = activities)	Sarkar *et al.* [Sarkar et al., 2022]	2022	signal-to-image (Continuous Wavelet Transform)	Spatial Attention-aided CNN	99.45%
UCF sports (target = actions)	Chuan *et al.* [Sun et al., 2015]	2015	Joint Self-Similarity Volume (Joint-SSV)	GP-based & k-NN	[86.90% – 88.50%]
Naturalistic McGill University gait dataset and Osaka University gait dataset	Zhong *et al.* [Zhong and Deng, 2014]	2014	signal-to-image (Gait dynamics images)	I-vector	[67.5% – 85.0%]
IXMAS (target = actions)	Korner *et al.* [Körner and Denzler, 2013]	2013	signal-to-image : Multi-View SSM	Gaussian process + Histogram intersection kernel	79.00%
CMU mocap (target = actions)	Junejo *et al.* [Junejo et al., 2008]	2008	signal-to-image : self-similarity matrices (SSM)	Nearest Neighbour Classifier & SVM	[90.50% – 95.70%]

This part explores a new user authentication by analyzing their human activities, using Deep Learning classifiers on the UCI-HAR dataset. The results show that using combinations of sensor data achieves the lowest equal error rates (EER) for binary classification.

Another part of the study focuses on authenticating smartphone users using accelerometers, gyroscopes, and magnetometer sensors. The results indicate that the proposed new framework outperforms existing methods, with an improvement in equal error (EER). The major contributions of this study are the demonstration of the effectiveness of Deep Learning approaches in individual identity verification and the use of a signal-to-image transformation to improve authentication results. Future research will focus on improving the security of biometric continuous authentication systems by creating presentation attack tools, evaluating the quality of human activities with Deep Learning architectures, and using synthetic databases for laboratory evaluation.

3.5.3 Keystroke dynamics authentication

In this section, we present the results we obtained on keystroke dynamics data. We tried to structure them by answering some questions concerning the performance of the proposed method.

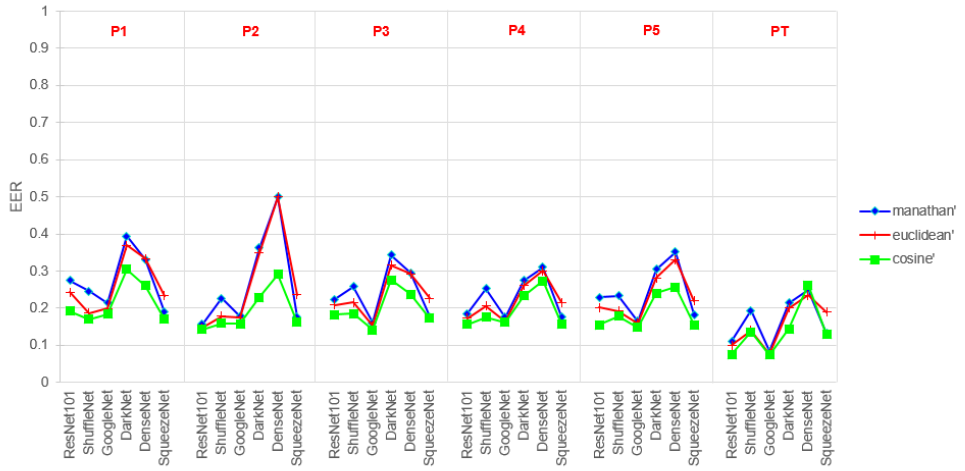
3.5.3.1 Which performance can we obtain on each dataset?

First, we consider a single password, i.e. we take each database separately to generate results. We used $1,100$ samples in total (110 users * 10 entries) taking 70% for the learning phase and 30% for the testing one.

We illustrate the six architectures among the seven (namely ResNet-101, ShuffleNet, DarkNet, GoogleNet DarkNet-53, DenseNet-201, and SqueezeNet) and we draw the model and the metric that offers the best performances on each database separately. This is illustrated by Figure 3.17.

We reveal the three architectures that offer the best performances on each database, namely ResNet-101, DarkNet, and GoogleNet in Table 3.24.

Figure 3.17 shows that the *Cosine* distance outperforms than *Minkowski* and *Eulidean*



.

FIGURE 3.17: EER (×100) rate on deep architectures for P1, P2, P3, P4, P5 and PT sub-databases

TABLE 3.24: Databases fusion P1, P2, P3, P4 and P5

Methods	**CA** (%)	**EER**$_{Minkowski}$(%)	**EER**$_{Euclidean}$(%)	**EER**$_{Cosine}$(%)
ResNet-101	74.85	10.01	09.98	**07.45**
ShuffleNet	72.12	19.97	19.78	**14.47**
GoogleNet	66.06	07.71	07.71	**07.45**
DarkNet-53	60.91	14.04	14.05	**13.60**
DenseNet-201	57.27	26.18	23.81	**23.46**
SqueezeNet	47.27	18.57	18.69	**16.10**

disance. The values of each method are represented in Table 3.24. This helps us in the following experiment, to only use the *Cosine* metrics distance (to have quicker computations).

We observe that GoogleNet offers the best performance with an EER value equal to 18.43% (P1), 14.20% (P3) and 14.80% (P5). Sometimes, ResNet-101 performs well with a EER value to 14.23% (P2) and 15.70% (P4). We can also note that we do not have the same performance from one password to another. Using 7 samples for the user reference template generation does not provide very good results (with an EER value between 14% to 18%). Obviously, if we had much more data we would expect to get better scores for a database.

TABLE 3.25: Performance evaluation on sub-databases separately P1, P2, P3, P4 and P5

Sub-databases	**Models**	**CA(%)**	EER_{cosine}
P1	ResNet-101	33.94	19.15
P1	DarkNet-53	19.39	30.50
P1	GoogleNet	25.76	**18.43**
P2	ResNet-101	40.61	**14.23**
P2	DarkNet-53	38.48	24.70
P2	GoogleNet	34.24	15.77
P3	ResNet-101	32.42	18.25
P3	DarkNet-53	36.06	24.19
P3	GoogleNet	35.45	**14.20**
P4	ResNet-101	46.36	**15.70**
P4	DarkNet-53	29.39	22.72
P4	GoogleNet	35.45	16.18
P5	ResNet-101	48.18	15.50
P5	DarkNet-53	37.88	23.96
P5	GoogleNet	40.91	**14.80**

3.5.3.2 Which performance can we obtain on a larger dataset?

In this section, we consider the 5 datasets as if biometric samples were from different users. We merge all the sub-databases (**fusion of features**), and we consider that we have 5×110 users. We also took 70% of data for the learning phase and 30% for the testing one. Table 3.24 draws the obtained results for keystroke dynamics authentication.

In a comparative analysis of the 7 trial architectures, the best performance for static authentication is provided by ResNet-101 ($EER = 07.45\%$), DarkNet-53 ($EER = 13.60\%$) and GoogleNet ($EER = 07.45\%$). We also note that the cosine distance metric provides a better EER value whatever the architecture used, except DenseNet-201 case. We keep the cosine distance in the rest of this work.

Performance is largely improved because we used more data that is generally requested for Deep Learning techniques. This illustrates the need of large keystroke dynamics (in terms of users and samples per user) to optimize the performance of Deep Learning methods on this biometric modality.

3.5.3.3 Which performance can we obtain if the user types more than one passphrase?

In this part, it is assumed that a person types more than one passphrase on the keyboard to authenticate himself/herself. We merge the inter-class and intra-class scores (**fusion of score**) considering the number of typed passwords. We also took 70% of data for the learning phase and 30% for the testing one. Table 3.27 draws the obtained results if we used the five typed passphrases (i.e. simulating a user typed the 5 passphrases to be authenticated). GoogleNet comes out as the best method with an EER score of 04.49% as presented in Table 3.26. GoogleNet is ahead of ResNet101 (06.70%) and ShuffleNet (07.34%).

TABLE 3.26: Performance evaluation on the multi-instance biometric system by fusion of features and scores level on P_T.

Models (EER_{cosine})	**Fusion of features**	**Fusion of scores**
ResNet-101	07.55%	06.70%
ShuffleNet	13.59%	07.34%
GoogleNet	**07.45%**	**04.89%**
DarkNet-53	14.96%	11.11%
DenseNet-201	26.18%	10.50%
SqueezeNet	12.87%	08.68%

TABLE 3.27: Performance with scores fusion (P1, P2, P3, P4 and P5).

Models	EER_{cosine}
ResNet-101	06.70%
DarkNet-53	10.97%
GoogleNet	**04.89%**

GoogleNet comes out as the best method with an EER score of 04.49% as presented in Table 3.27. GoogleNet is ahead of ResNet-101 (06.70%) and DarkNet (10.97%).

To complete these results, we studied the obtained performance versus the number of passphrases typed by a user. Figure 3.18 highlights the EER value obtained for each case.

- In the most classical case, if we use 2 inputs (*i.e.* login + password), we obtain an EER value between $[9.17\% - 22.95\%]$ illustrated by block 2 in Figure 3.18.

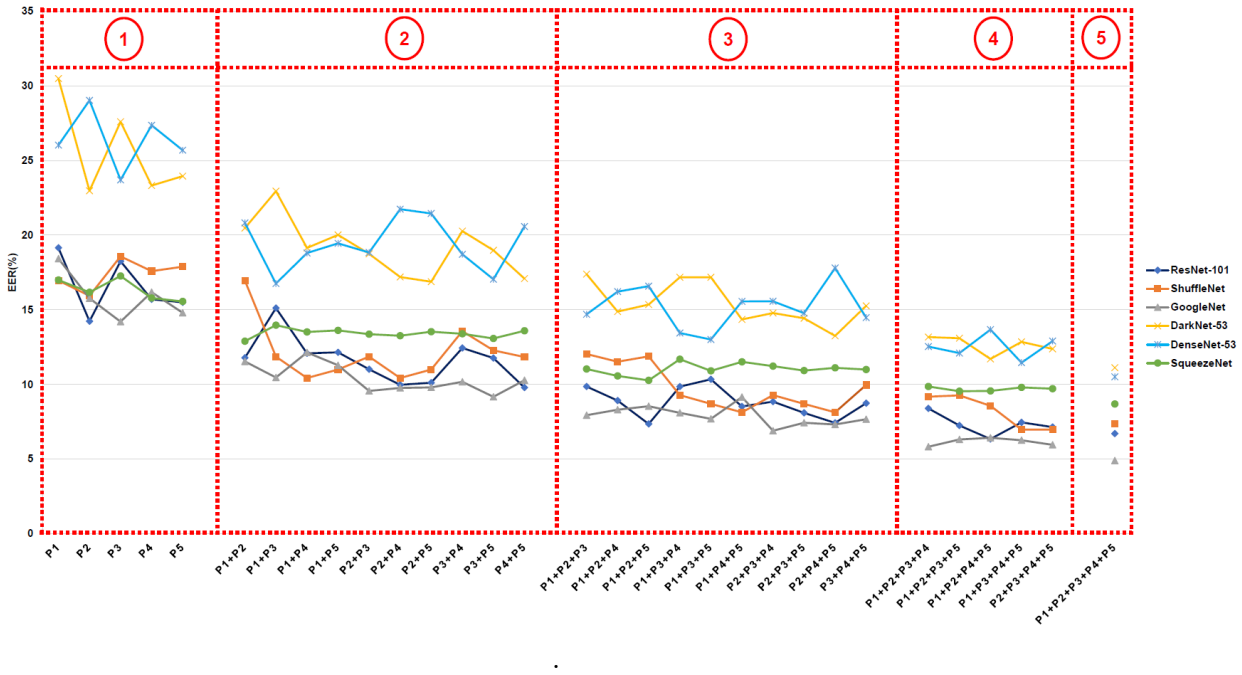


FIGURE 3.18: EER rate on deep architectures for the multi-instance biometric system. In block 1, we have P1, P2, P3, P4, and P5 sub-database. In block 2, we have the fusion of inter and intra-class scores from (P1+P2) to (P4+P5) sub-databases respectively. In block 3, (P1+P2+P3) to (P2+P3+P5). In block 4, (P1+P2+P3+P4) to (P2+P3+P4+P5) and in Block 5, (P1+P2+P3+P4+P5)

- If we have 3 inputs (*i.e.* login + password + secret question), we have an EER value between $[6.89\% - 17.80\%]$ represented by block 3.
- If we use 4 inputs (*i.e.* login + 2 passwords + secret question), we have an EER value around $[5.95\% - 13.67\%]$ depicted by block 4.
- If we use 5 inputs, we get an EER value around $[4.89\% - 11.11\%]$ depicted by block 5. Even if this scenario is less realistic, it shows we can easily decrease the EER value for this kind of authentication.

We note that the EER value obtained when merging the score of both databases decreases for each architecture. It also appears from this work that the more information we have, the better the performance can be, it is not surprising.

3.5.3.4 Discussion

Biometric keystroke authentication is totally based on the routines that the users have since they probably entered the same password numerous times. In this case, their typing styles become so unique and hard to imitate, which is also the core of the keystroke recognition systems [Alpar, 2017]. Neural networks have the advantage of being able to handle many parameters. However, they can be slow not only during training but also in the application phase. The purpose of this work is to analyze several pieces of information entered by a user in order to authenticate him/her.

If we focus on research works that have been performed on the GREYC-NISLAB database, we can compare our results. Idrus *et al.* [Idrus et al., 2015] obtained an EER value of 10.63% using a SVM-based method. They further improved the keystroke dynamics authentication accuracy from an EER value of 8.45%. Considering the same database, the proposed approach with GoogleNet performs better with an EER value of 04.89% (Table 3.28).

TABLE 3.28: Comparison with other published works in keystroke dynamics. EER values are reported (note some works have used non-representative datasets). For each reported work, different biometric samples are merged.

Database	Author/S (ref)	Years	Classifiers	EER
GREYC-NISLAB	**This work**	**2021**	**GoogleNet**	**4.89**%
GREYC-NISLAB	Idrus *et al.* [Idrus et al., 2015]	2015	SVM	[08.45% – 10.63%]
Clarkson II	Li *et al.* [Li et al., 2021]	2021	CNN & CNN-GRU	[07.55% – 07.74%]
Synthetic	Ayotte *et al.* [Ayotte et al., 2021a]	2021	SVM & MLP	[04.90% – 05.46%]
GREYC 2009 vs WEB GREYC	Mhenni *et al.* .[Mhenni et al., 2018]	2018	kNN	[06.61% – 07.08%]
GREYC Keystroke	Zhong *et al.* [Zhong and Deng, 2015]	2015	SVM	[08.45% – 10.65%]

To complete this comparison, we consider other works on different biometric databases. Of course, it is not possible to have a fair comparison but we give these values for illustration. When different keystroke dynamics samples are fused, Ayotte et al. [Ayotte et al., 2021a] (2021) obtained an EER value of 04.90% with the MLP method. Wahab et al. [Wahab et al., 2021] in 2021 obtained an EER value of [0% – 05.47%] with Manhattan distance. They used a different database than the GREYC-NISLAB one (and is private). Both of these works are based on content knowledge. This is not the case with our work because we place ourselves in an attack situation, that is to say, we consider that the attacker (passphrase situation) knows the password. We try to recognize and authenticate a person only by the way he/she types.

3.6 Conclusion

The aim of this study was to present a generic architecture that can be adapted to any behavioral biometric modality. The use of a signal-to-image representation to present behavioral data makes it possible to handle various types of behavioral biometric modalities.

This research work presents an approach based on multimodal behavioral biometrics for user identification and authentication, using Machine Learning. Human activities on smartphones and typing dynamics on laptops are studied as biometric modalities. Deep Learning achieves state-of-the-art performance in time series classification but raises privacy concerns. In the context of authentication, a method based on keystroke dynamics is proposed, offering advantages such as low cost, but also potential privacy issues. Finally, authentication based on human activities shows promising results using combinations of sensors, and the use of signal-image transformation improves performance.

In the next chapter, we focus on the use of synthetic behavioral biometric databases. Synthetic data could be used to better train deep learning architectures in order to enhance performance.

CHAPTER 4

Generation of Synthetic Behavioral Biometric Data

Summary

This chapter explores the generation of synthetic behavioral biometrics using generative adversarial networks (GANs), with an introduction to the importance of this approach. It dives into the principles and architecture of GANs, highlighting TimeGAN, a specialized methodology for synthetic signal generation. The evaluation of generation models, including the performance of TimeGAN, is examined in detail. Potential applications of synthetic data generation are discussed, and the chapter concludes with a discussion and overall synthesis of the presented concepts.

Keywords: Synthetic behavioral biometrics; GANs; TimeGAN; Performance.

4.1 Introduction

When it comes to data protection, cybersecurity plays an essential role in user authentication and access to personal and confidential information. A failure of the user authentication system can result in significant economic, social, and reputational damage [Clark et al., 2017, Bud, 2018]. As a result, authentication systems are becoming increasingly robust. In this context, biometrics plays a crucial role, as it offers a universal, unique, permanent over time, and measurable means of authenticating users [Matyáš and Říha, 2002].

In recent years, biometric-based user authentication systems have been widely used in a variety of scenarios, including airport scanners, banking, military access control, smartphones, and forensics, among others [Muley and Kute, 2018, Bud, 2018]. These systems, usually based on machine learning techniques, extract feature measurements and determine whether they match the characteristics of the user requesting access.

Despite the promising results it offers in a wide range of applications, behavioral biometrics is vulnerable to various forms of attacks [Marcel et al., 2014]. The most common attack against a user authentication system is known as a presentation attack (PA), where the adversary targets the biometric sensor that collects the individual's measurements, as outlined in [Ness, 2017].

Keystroke dynamics, a form of behavioral biometrics, for example, refers to the way a user types. Studies have shown that users can be distinguished and authenticated on the basis of their typing habits. However, it is often extremely difficult for another user or a robot to reproduce these typing patterns (except the keylogger attack), as mentioned [Eizaguirre-Peral et al., 2022].

Many prominent fields, including finance, medicine, meteorology, and geophysics, are among the main sources of the most relevant temporal data. With this growth in information comes an increasing demand for solving machine learning tasks such as classification, prediction, detection, and many others, to meet the challenges at hand. Factors such as the high cost of data collection or data quality make data acquisition difficult, if not impossible. The limited availability of such data has a significant impact on machine

learning performance and restricts the ability of models to operate effectively, as indicated by [Smith and Smith, 2020]. In the literature, the aspect of generating behavioral biometric data has not yet been thoroughly explored in the context of presentation attacks. Therefore, in this chapter, we propose to generate synthetic behavioral biometric data through adversarial networks as time series configurations for a presentation attack.

What are the criteria for an effective generative model for time series data? The temporal context presents a particular challenge for generative modeling. Not only must a model be able to capture the distributions of features at each instant in time, but it must also be able to capture the potentially complex dynamics of these variables through time. Specifically, when modeling multivariate sequential data $x_{1:T} = (x_1, ..., x_T)$, our goal is to accurately capture the conditional distribution $p(x_t|x_{1:t-1})$ of temporal transitions, as described in [Yoon et al., 2019].

Much research has focused on improving the temporal dynamics of autoregressive models for sequence prediction. They mainly address the problem of error accumulation during multi-stage sampling, by introducing different adaptations of the training time to better match the conditions of temporal evaluation [Bengio et al., 2015, Lamb et al., 2016, Bahdanau et al., 2016].

Autoregressive models explicitly factor the distribution of sequences into a product of conditionals $\prod_t p(X_t|X_{1:t-1})$. However, while this approach is useful in the context of forecasting, it remains fundamentally deterministic and cannot truly be called generative, in the sense that it is not possible to randomly sample new sequences without external conditioning. Furthermore, another approach has focused on the direct application of the Generative Adversarial Networks (GAN) framework to sequential data, in particular using recurrent networks to play the roles of generator and discriminator [Mogren, 2016, Esteban et al., 2017, Ramponi et al., 2018].

Despite its simplicity, the contradictory aim is to model $p(x_{1:T})$ directly without exploiting the autoregressive prior. It is vital to emphasize that adding only standard GAN loss to vector sequences may not be sufficient to ensure that the network effectively captures the progressive dependencies present in the training data [Yoon et al., 2019].

In this chapter, we propose to create synthetic behavioral biometric data using Time Series Generative Adversarial Networks (TimeGAN), a framework for generating realistic time series data in various domains.

Firstly, in addition to unsupervised loss, we use the original data as a reference, explicitly encouraging the model to capture the progressive conditional distributions present in the data. This means that we use this original data to guide the model, going beyond the simple distinction between real and synthetic data. In this way, we can learn explicitly from the transition dynamics of real sequences.

Secondly, in the TimeGAN architecture, there is an integration network that establishes a reversible correspondence between features and latent representations, thus reducing the high dimensionality of the contradictory learning space. This approach capitalizes on the fact that the temporal dynamics of complex systems are often guided by a small number of lower-dimensional variation factors.

It is essential to note that supervised loss is minimized by simultaneously training the integration and generation networks. Thus, the latent space is not only used to optimize parameter efficiency but is specifically conditioned to facilitate the learning of temporal relationships by the generator.

Since the TimeGAN framework has been applied to mixed data, generating both static and temporal data, this approach stands out as the first to combine the flexibility of the unsupervised GAN framework with the control offered by supervised learning in autoregressive models. On the qualitative side, we use t-SNE [Van der Maaten and Hinton, 2008] to visualize how similar the generated distributions are to the original ones. On the quantitative side, we evaluate the ability of a posterior classifier to distinguish real from generated sequences. Furthermore, by applying the "training on synthetic data, testing on real data (TSTR)" framework [Esteban et al., 2017, Jordon et al., 2018] to the sequence prediction task, we measure the extent to which the generated data retains the predictive features of the original. Our results consistently and significantly demonstrate that TimeGAN improves the state-of-the-art in the generation of realistic time series (behavioral biometric data).

4.2 Related work

TimeGAN is a generative time series model that is trained consistently and simultaneously through a learned integration space, using supervised and unsupervised loss [Yoon et al., 2019]. This chapter focuses on the approach, which is positioned at the intersection of several research areas, combining the concepts of autoregressive models for sequence prediction, GAN-based methods for sequence generation, and time series representation learning.

The methods cited and listed in the literature are mainly supervised autoregressive models used for prediction or unsupervised GANs used for generation. Although other approaches exist for generating data by various means, these two methods predominate in the literature. This section focuses more on an in-depth review of non-autoregressive models.

Autoregressive models are of particular interest in the field of generative networks because of their stability, suitability for parallel training, speed of inference for new predictions, and lack of truncated backpropagation in time. They share similarities with recurrent models, which use input data to predict the next time step in the data [Smith and Smith, 2020]. Basically, autoregressive models make predictions by regressing the value of a time series on its previous values. Deep autoregressive models exploit neural networks to learn the function that predicts the future data of the time series. In short, these predictions can be interpreted as generating the distribution of the time series based on previous time steps.

Recurrent autoregressive networks, trained using maximum likelihood, can have significant prediction errors during multi-step sampling due to the difference between closed-loop training (i.e. conditioned by basic truths) and open-loop inference (i.e. conditioned by previous assumptions) [Yoon et al., 2019]. To solve this problem, several approaches have been developed, including Scheduled Sampling [Bengio et al., 2015], the Professor Forcing [Ganin et al., 2016], and Actor-Critic methods [Konda and Tsitsiklis, 1999].

Scheduled Sampling was initially proposed as a solution, where models are trained to generate outputs conditioned on a mixture of prior assumptions and reference data. Professor

Forcing is inspired by contradictory domain adaptation, involving the training of an auxiliary discriminator to distinguish between free hidden states and teacher-forced hidden states, thus promoting the convergence of network training and sampling dynamics [Lamb et al., 2016]. Finally, Actor-critic-based methods introduce a target-output-conditioned critic, trained to estimate future values of the next element that guides the actor's autonomous predictions [Bahdanau et al., 2016]. However, it's important to note that these methods are deterministic and don't allow explicit sampling from a learned distribution, contrary to our goal of generating synthetic behavioral biometrics data [Yoon et al., 2019].

Numerous studies have adopted the GAN framework for generating temporal data. Early approaches, such as C-RNN-GAN and RCGAN, used LSTM (Long Short Term Memory) networks to generate sequences based on recurrent data. These frameworks were then successfully applied to various domains, such as word processing, finance, biological signals, sensor data, smart grids and renewable energy scenarios. Recent work has also proposed the integration of timestamp information to manage irregular sampling.

In contrast to the TimeGAN [Yoon et al., 2019] approach, these methods mainly use binary adversarial feedback for learning, which may not effectively guarantee the capture of the temporal dynamics of the training data. With regard to representation learning in the context of time series, it generally focuses on efficiency for tasks such as prediction, forecasting, and classification. Other work explored the use of latent representations for pre-training, disentanglement and interpretability. In the static context, research has examined the benefits of combining autoencoders with adversarial learning for a variety of purposes, including learning similarity measures, efficient inference and improving generative capacity, which has also been applied to the generation of discrete structures.

However, the method proposed for TimeGAN presents a generalization to arbitrary time series data. It incorporates stochasticity at each time step and uses an integration network to identify a lower-dimensional space for the generative model, enabling learning of stepwise distributions and latent data dynamics.

This chapter presents the results obtained in the synthetic generation of behavioral biometric signals, including human activity and keystroke dynamics data, using TimeGAN.

This model proved to be the most suitable for the synthetic generation of temporal sequences, as suggested by the convincing findings obtained compared to other generative models [Brophy et al., 2023]. In light of our results, we identify promising potential in the use of TimeGAN to generate realistic synthetic data, particularly in the context of behavioral temporal sequences.

4.2.1 Generative Adversarial Networks

The introduction of GANs has considerably facilitated progress in the generation of synthetic data. These deep learning models typically consist of two neural networks: a generator and a discriminator. The generator G takes a random noise vector $z \in \mathbb{R}^r$ and aims to produce synthetic data similar to the distribution of the training data. In parallel, discriminator D attempts to determine whether the generated data is genuine or false. The generator seeks to maximize the discriminator's failure rate, while the discriminator seeks to minimize it. Figure 4.11 provides a concise illustration of GAN architecture and the game being played between these two neural networks. These two networks are engaged in a zero-sum game, defined by the value function V(G, D) as:

$$\min_G \max_D V(G,D) \;=\; \mathbb{E}_{x\sim P_{data}(x)}\left[\log D(x)\right] \;+\; \mathbb{E}_{z\sim P_z(z)}\left[\log(1 \;-\; D(G(z)))\right] \tag{4.1}$$

where $D(.)$ represents the probability that the data comes from real data rather than generated data [Goodfellow et al., 2014].

GANs belong to the generative model family and are an alternative method for producing synthetic data, not requiring specific domain expertise [Brophy et al., 2023]. They were initially introduced by Goodfellow et al [Goodfellow et al., 2014] in 2014, using a multilayer perceptron for both discriminator and generator.

In 2015, Radford et al [Radford et al., 2015] developed the Deep Convolutional Generative Antagonist Network (DCGAN) for synthetic image generation. Since then, researchers have continued to improve early GAN architectures, loss functions, and evaluation metrics, while exploring their potential real-world applications. To understand the scope of these

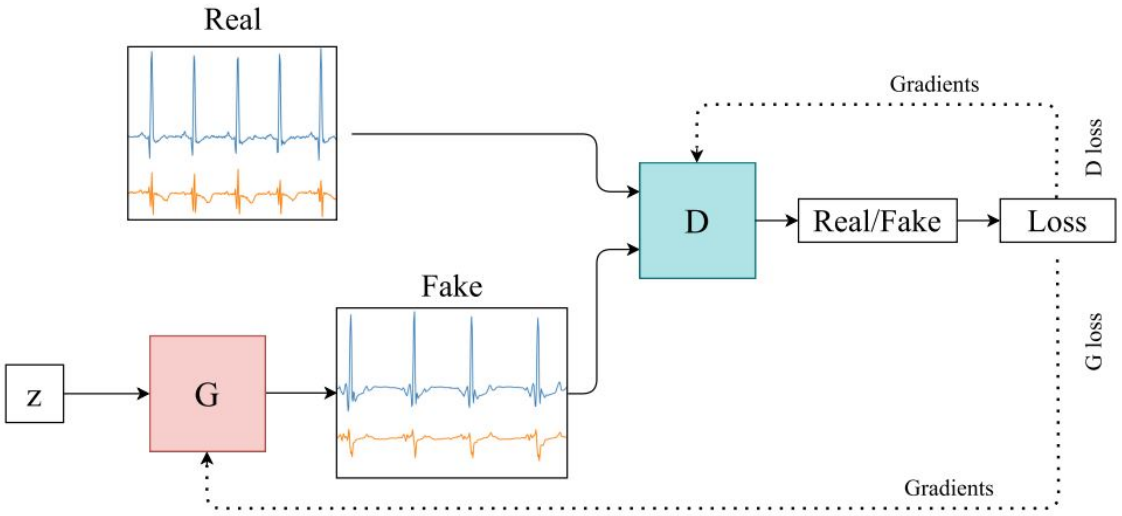


FIGURE 4.1: Generative adversarial network [Brophy et al., 2023].

efforts, it is crucial to grasp the initial limitations of the first architectures, loss functions, and evaluation metrics, as well as the challenges they pose.

4.2.1.1 GANs basic principles

Generative Adversarial Networks (GANs) are based on a fundamental principle: competition between two neural networks, the generator, and the discriminator. The generator learns to generate synthetic signals, while the discriminator tries to distinguish synthetic signals from real signals. This competition creates a dynamic learning loop, where the generator constantly seeks to fool the discriminator by producing increasingly realistic synthetic signals [Goodfellow et al., 2014].

The generator is responsible for generating synthetic signals from random input noise. Its aim is to produce signals that resemble as closely as possible the real signals present in the training data set.

The discriminator, meanwhile, is trained to distinguish the synthetic signals generated by the generator from the real signals in the training dataset. Its role is to learn how to correctly classify signals as real or synthetic. The learning process takes place iteratively, with the generator and discriminator facing each other in a min-max game.

During training, the generator and discriminator are updated alternately. The generator is optimized to maximize the probability of synthetic signals being classified as real by the discriminator, while the discriminator is optimized to minimize the probability of incorrect classification of both real and synthetic signals.

The discriminator, meanwhile, is trained to distinguish the synthetic signals generated by the generator from the real signals in the training dataset. Its role is to learn how to correctly classify signals as real or synthetic. The learning process takes place iteratively, with the generator and discriminator facing each other in a min-max game. The generator tries to fool the discriminator by producing increasingly realistic synthetic signals, while the discriminator tries to improve its ability to distinguish real from synthetic signals.

During training, the generator and discriminator are updated alternately. The generator is optimized to maximize the probability of synthetic signals being classified as real by the discriminator, while the discriminator is optimized to minimize the probability of incorrect classification of both real and synthetic signals.

4.2.1.2 General architecture of GANs

The structure of the generator and discriminator can vary according to the type of synthetic signal to be generated. Generators can be based on deep neural network architectures, such as convolutional neural networks (CNN) for images, recurrent neural networks (RNN) for text, or transposed convolutional neural networks (Transposed CNN) for audio. Similarly, discriminators are designed to be able to distinguish between the specific characteristics of real and synthetic signals.

4.2.1.3 Adversarial learning loss function

GAN learning is adversarial, i.e. the generator and discriminator improve each other's performance. A loss function is used. The generator loss function is based on the discriminator probability of success in classifying synthetic signals as real. The generator seeks to minimize this probability, which is equivalent to maximizing the probability that the synthetic signals are considered real. The loss function is a key element in GAN learning. It

measures the difference between the discriminator's classifications and the expected labels (real or synthetic).

Setting hyperparameters such as learning rate, number of iterations, batch size, and using regularization techniques to avoid overlearning is also crucial to the successful training of a GAN. Additional techniques, such as batch normalization, label smoothing, and the use of specific loss functions, can also be used to improve training stability.

For time series generation, several techniques can be used to improve synthetic signal generation with GANs. These include:

- The use of normalization layers, such as batch normalization, to help stabilize generator and discriminator learning.
- The introduction of regularization, such as L1 or L2 regularization, to control model complexity and avoid overlearning.
- Exploration of latent space, by manipulating input noise vectors to generate interesting variations in the synthetic signals produced.
- The use of transfer learning techniques, relying on pre-trained models for the generator or discriminator to speed up learning or improve the quality of the synthetic signals generated.
- The use of transfer learning techniques, relying on pre-trained models for the generator or discriminator to speed up learning or improve the quality of the synthetic signals generated.
- Adding regular constraints: Incorporating regular constraints, such as gradient penalty regularization (WGAN-GP) or Kullback-Leibler (KL) divergence, can help stabilize learning and improve the quality of synthetic signals.

By using these techniques and other advanced methods, it is possible to significantly improve the generation of synthetic signals by GANs and obtain more realistic and diverse results.

4.2.2 TimeGAN: methodology for generating synthetic signals

TimeGAN offers an innovative approach by merging the classical unsupervised learning principles of GAN with a more controllable supervised learning method [Yoon et al., 2019]. This combination of an unsupervised GAN network and a supervised AR model makes it possible to generate time series while preserving their temporal dynamics. An overview of TimeGAN's architecture can be found in Figure 4.12.

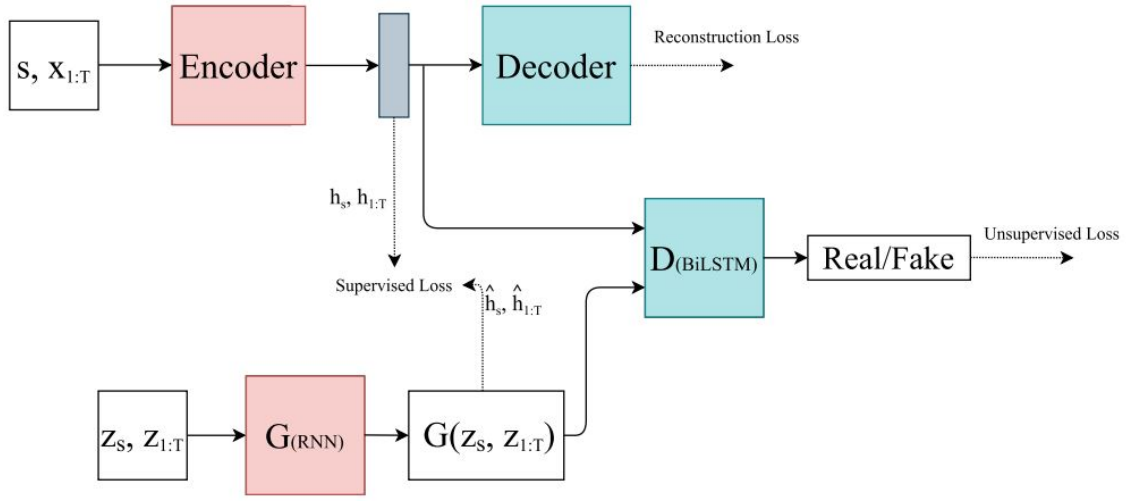


FIGURE 4.2: TimeGAN architecture.

- The input framework is considered to consist of two distinct components: a static feature, represented by the vector **s**, and a temporal feature, represented by the vector **x** at the encoder input.
- The generator takes as input a tuple of static and temporal feature vectors randomly generated from a previously defined distribution.
- Real ($\mathbf{h}$) and synthetic ($\hat{\mathbf{h}}$) latent codes are used to calculate the supervised loss element of this network.
- The discriminator evaluates the set of latent codes by classifying them as real (y) or synthetic ($\hat{y}$), and the operator $\sim$ indicates that the sample is either real or false.

The three losses used in TimeGAN are determined as follows:

1. The reconstruction loss function, which refers to the auto-encoder, compares how well was the reconstruction of the encoded data when compared to the original one. It is defined as follows:

$$L_{reconstruction} = \mathbb{E}_{s\,,\,x_{1:T\sim P}} \left[\|s - \tilde{s}\|_2 + \sum_t \|x_t - \tilde{x}_t\|_2 \right] \tag{4.2}$$

2. The unsupervised loss reflects the relation between the generator and discriminator networks (min-max game), and is defined as follows:

$$\begin{aligned} L_{uns} = \mathbb{E}_{s,x_{1:T\sim P}} \left[log(y_s) + \sum_t log(y_t) \right] + \\ \mathbb{E}_{s,x_{1:T\sim \hat{P}}} \left[log(1-\hat{y}s) + \sum_t log(1-\hat{y}_t) \right] \end{aligned} \tag{4.3}$$

3. The supervised loss, which is responsible to capture how well the generator approximates the next time step in the latent space, is defined as:

$$L_{supervised} = \mathbb{E}_{s\,,\,x_{1:T\sim P}} \left[\sum_t \|h_t - g_X(h_s, h_{t-1}, z_t)\|_2 \right] \tag{4.4}$$

The TimeGAN authors carried out experiments that involved the generation of various types of data. This included sine waves, daily historical data of Google's actions from 2004 to 2019, energy data from the UCI Appliances Energy Prediction dataset [Dua et al., 2017], as well as event data from the Private Lung Cancer Pathways dataset. They used a batch size of 128 and the Adam optimizer for training, with implementation details available online [1]. The authors demonstrated improvements over other time-series GANs, such as RCGAN, C-RNN-GAN, and WaveGAN.

Consequently, our work is based exclusively on TimeGAN.

[1] `https://github.com/jsyoon0823/TimeGAN`

4.3 Proposed architecture

In this section, we proposed a system that leans a behavioral modality, generates a synthetic dataset through the TimeGAN architectures, and evaluates the performance of the proposed system. This is done with subjective inspection (with t-SNE) and objective metrics (with the predictive score).

Figure 4.14 presents our proposed generic workflow, which integrates the temporal adversarial generator for the synthetic time-series data generation module, as well as the matching module. The aim of this workflow is to generate synthetic behavioral biometric data to simulate the impersonation of an authorized user as part of our authentication system described earlier in Chapter 3. This generic process has been designed to validate our proposed method.

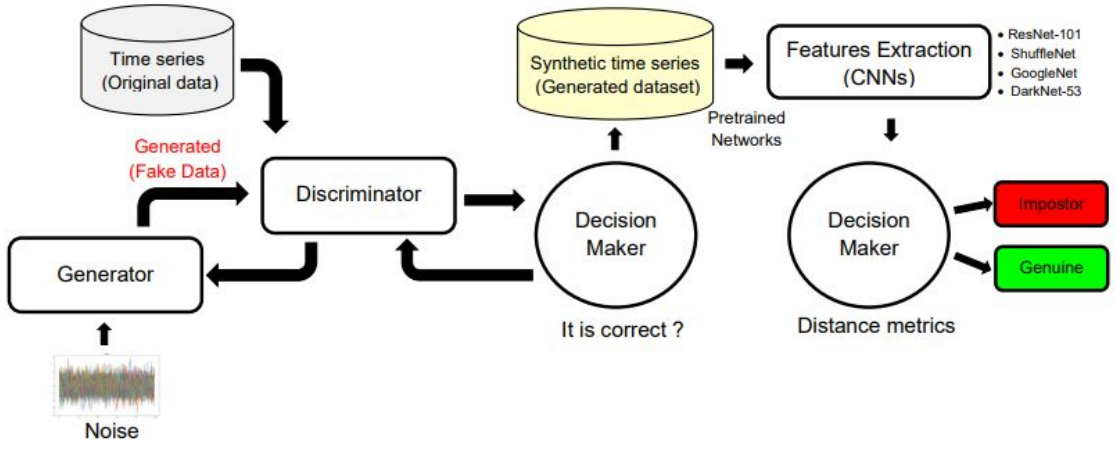


FIGURE 4.3: General overview of the evaluation attack system.

The process begins with the creation of a synthetic time series, which is then converted into an image representation through the signal-to-image process described in Chapter 3. The result obtained is a 2D color image. This resulting image is then exploited in convolutional neural networks (GoogleNet and ShuffleNet) to extract relevant features. For this feature extraction, we use a pre-trained network that has been previously (in Chapter 3) on the original dataset. Features are extracted at the last layer of the convolutional neural networks, enabling us to construct an output feature vector.

This feature vector is then compared with a reference/test model to assess the match between the synthetic biometric data and the authentic data. Once we have obtained a match score, we make an authentication decision. This decision is based on a simple threshold approach, where the user is *accepted* if the score exceeds a predefined threshold, or *rejected* if the score is below this given threshold.

Our generic workflow uses TimeGAN to generate synthetic behavioral biometric data, transforms it into actionable features via convolutional neural networks, and then compares it to a reference/test model to make an authentication decision based on defined thresholds.

The TimeGAN network parameters optimized for the dataset are shown in Table 4.1. For the configuration of the TimeGAN, in the RNN cell of the TimeGAN, we choose several modules separately, such as GRU, LSTM, or LSTM Layer Normalization (LSTM LN). By comparing the results provided by the three modules, we select the module that offers the best contribution to the reproduction of the original behavioral biometric data. We specify the hidden dimensions, the number of layers, the number of training iterations, and the number of samples in each batch. These parameters allow us to customize the model.

TABLE 4.1: TimeGAN network parameters

Parameter	**Option**
Module	'GRU', 'LSTM', 'LSTM LN'
Hidden dimensions	24
Number of layers	5
Iterations	10000
Batch size	128

4.4 Experimental protocol

We use the same databases as those used in Chapter 3, *i.e.* UCI-HAR for human activity and GREYC-NISLAB for keystroke dynamics.

4.4.1 Evaluation metrics

Assessing GANs is a complex task, and there is as yet no consensus among researchers on the most appropriate measures for evaluating their performance. Numerous measures

have been proposed in the literature [Borji, 2019], most of them are adapted to the field of computer vision. Efforts are continuing to arrive at an adequate evaluation of time-series GANs.

We classify evaluation measures into two categories: qualitative and quantitative. Qualitative evaluation refers to visual inspection by human observers, who inspect the samples generated by the GAN. However, this approach does not constitute a comprehensive assessment of GAN performance, due to the lack of appropriate objective evaluation measures.

Quantitative evaluation, on the other hand, relies on the use of metrics linked to statistical measures commonly used in time series analysis. These metrics include Pearson's correlation coefficient (PCC) defined by equation 4.5 (an example is illustrated in Figure 4.4), percent root mean square difference (PRD) defined by equation 4.6, root mean square error (RMSE) and mean square error (MSE) defined by equation 4.7, mean relative error (MRE) and mean absolute error (MAE) defined by equation 4.8. These are among the most frequently used measures for evaluating time series and are therefore considered appropriate performance measures for GANs. They reflect the fit between the training data and the synthetic data generated.

$$\text{PCC} = \frac{\sum_{i=1}^{N}(x_i - \tilde{x})(y_i - \tilde{y})}{\sqrt{(\sum_{i=1}^{N}((x_i - \tilde{x})^2 \sum_{i=1}^{N}((y_i - \tilde{y})^2}} \tag{4.5}$$

$$\text{PRD} = \sqrt{\frac{\sum_{i=1}^{N}((x_i - y_i)^2}{\sum_{i=1}^{N}((x_i)^2}} \tag{4.6}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}((x_i - y_i)^2} \tag{4.7}$$

$$\text{MRAE} = = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - y_i}{x_i - f_i}\right| \tag{4.8}$$

In these equations, x_i represents the actual value of the time series x at time/sample i, while y_i represents the generated value of the time series y at the same time/sample i. The symbols $\tilde{x}$ and $\tilde{y}$ denote the mean values of x and y respectively. A lower MSE value indicates better generation quality.

The parameter f_i is used to compute the MRAE (Mean Relative Absolute Error) for the forecast value at time i of a specified reference model. As a general rule, f_i can be chosen to be equal to y_{i-1} for non-seasonal time series, and y_{i-M} for seasonal time series, where M represents the seasonal period of x.

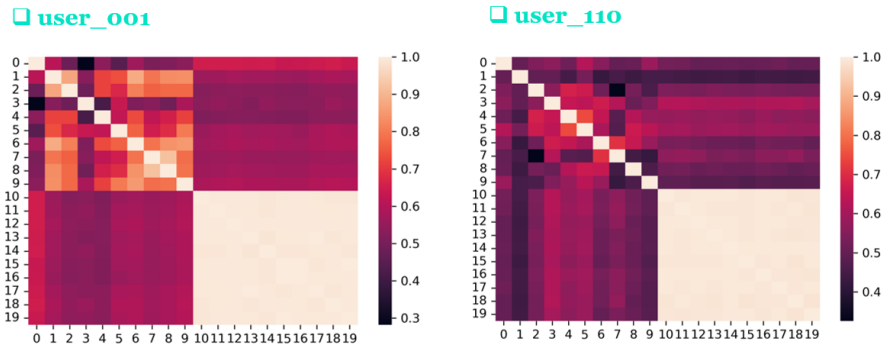


FIGURE 4.4: Pearson correlation values for time series (real versus synthetic).

Table 4.2 provides a review of the current state-of-the-art in GAN for time series, as well as new approaches to solving concrete problems involving time series data.

TABLE 4.2: Collection of GAN architectures, their applications, datasets used in their experiments, and evaluation criteria for assessing the quality of each respective GAN [Brophy et al., 2023].

Application	**GAN Architecture(s)**	**Dataset(s)**	**Evaluation Metrics**
Anomaly detection	LSTM-LSTM [Leangarun et al., 2018]; LSTM-(LSTM&CNN) [Zhu et al., 2019b]; LSTM-LSTM (MAD-GAN) [Li et al., 2019]	SET50, NYC taxi data, ECG, SWaT, WADI	Manipulated data used as a test set, ROC curve, precision, recall, F1, accuracy

Audio generation	C-RNN-NN [Mogren, 2016]; TGAN(variant) [Cheng et al., 2020]; RNN-FCN [Zhang et al., 2020]; DCGAN(variant) [Kolokolova et al., 2020]; CNN-CNN [Juvela et al., 2019]	Nottingham dataset, midi music files, MIR-1K, TheSession, speech	Human perception, polyphony, scale consistency, tone span, repetitions, NSDR, SIR, SAR, FD, t-SNE, distribution of notes
Financial time series generation/prediction	TimeGAN [Yoon et al., 2019]; SigCWGAN [Ni et al., 2020]; DAT-GAN [Sun et al., 2020]; QuantGAN [Wiese et al., 2020]	S&P 500 index(SPX), Dow Jones index (DJI), ETFs	Marginal distributions, dependencies, TSTR, Wasserstein distance, EM distance, DY metric, ACF score, leverage effect score, discriminative score, predictive score
Time series estimation/prediction	LSTM-NN [Li et al., 2020]; LSTM-CNN [Kaushik et al., 2020]; LSTM-MLP [Kaushik et al., 2020]	Meteorological data, Truven MarketScan dataset	RMSE, MAE, NS, WI, LMI
Time series imputation/repairing	MTS-GAN [Guo et al., 2019]; CNN-CNN [Qu et al., 2020]; DCGAN(variant) [Han et al., 2020]; AE-GRUI [Luo et al., 2019]; RGAN [Sun et al., 2018]; FCN-FCN [Chen et al., 2019]; GRUI-GRUI [Luo et al., 2018]	TEP, point machine, wind turbine data, PeMS, PhysioNet Challenge 2012, KDD CUP 2018, parking lot data	Visually, MMD, MAE, MSE, RMSE, MRE, spatial similarity, AUC score
Other time series generation	VAE-CNN [Parthasarathy et al., 2020]	Fixed length time series "vehicle and engine speed"	DTW, SSIM

Medical/Physiological generation	LSTM-LSTM [Abdelfattah et al., 2018, Esteban et al., 2017, Haradal et al., 2018, Harada et al., 2019, Nikolaidis et al., 2019, Wang et al., 2019]; LSTM-CNN [Brophy, 2020, Delaney et al., 2019]; BiLSTM-CNN [Zhu et al., 2019a]; BiGridLSTM-CNN [Hazra and Byun, 2020]; CNN-CNN [Fahimi et al., 2019, Hartmann et al., 2018]; AE-CNN [Pascual et al., 2020]; FCNN [Yi and Mak, 2019]	EEG, ECG, EHRs, PPG, EMG, speech, NAF, MNIST, synthetic sets	TSTR, MMD, reconstruction error, DTW, PCC, IS, FID, ED, S-WD, RMSE, MAE, FD, PRD, averaging samples, WA, UAR, MV-DTW

It is important to note that these metrics do not always exhaustively capture the perceived quality of synthetic signals. They provide an objective measure, but may not fully reflect subjective aspects such as aesthetics, intelligibility, or semantic coherence. Consequently, it is often necessary to combine these objective metrics with subjective assessments to obtain a more complete overall evaluation.

4.4.2 TimeGAN performance evaluation

To assess the quality of the data generated, we consider two criteria:

1. *Diversity*: samples must be distributed in such a way as to adequately represent the diversity of real data.

2. *Usefulness*: samples must be as relevant as real data when used for similar predictive purposes, i.e. when training with synthetic data or testing with real data.

A wide range of measures of evaluation are proposed by researchers to assess the performance of GANs including:

1. **Visualization**

   We apply the t-SNEanalyses to the original and synthetic datasets, flattening the temporal dimension. This allows us to visualize in two-dimensional space how closely the distribution of the generated samples resembles that of the original, thus providing a qualitative assessment of the diversity.

2. **Predictive score**

   For the sampled data to be useful, it must retain the predictive characteristics of the original. Specifically, we expect TimeGAN to excel at capturing conditional distributions over time. Therefore, using the synthetic dataset, we train a post-hoc sequence prediction model (by optimizing a two-layer LSTM) to predict next-step temporal vectors on each input sequence. We then evaluate this trained model on the original dataset. Performance is assessed using the Mean Absolute Error (MAE) metric. For event-based data, MAE is computed as the absolute value of the difference between 1 and the estimated probability of the event having occurred. This provides a quantitative assessment of the usefulness.It can be also be seen with Equation 4.7.

4.5 Experimental results

4.5.1 Statistical evaluation

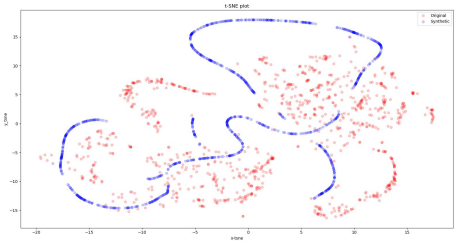
We evaluate our experimental results using visual inspection (through t-SNE analysis) and objective measurement (predictive score calculated from MAE).

Figure 4.5 illustrates the impact of synthetic behavioral biometric data generated from GREYC-NISLAB and HAR-UCI datasets by TimeGAN with GRU, LSTM, and LSTM

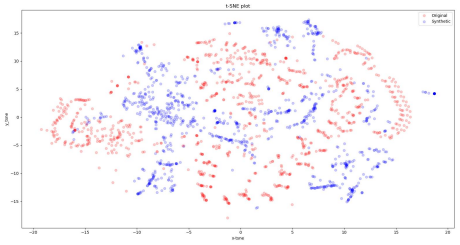
LN module. These data show a significantly higher level of overlap with the original data with GRU than other benchmark modules using the t-SNE method for visualization.

A significant observation in this graph reveals that the blue-colored samples (generated) are approximately similar to the red-colored samples (original) when using the GRU module.

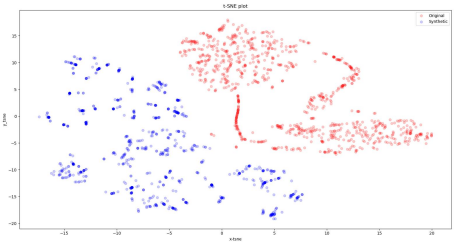
Figure 4.5a and 4.5d show that the synthetic datasets generated (respectively for UCI-HAR and GREYC-NISLAB datasets) by TimeGAN show a significantly higher overlap with the original data when the GRU module is used, compared to the other modules at the RNN cell (LSTM and LSTM LN), when using t-SNE for visualization.



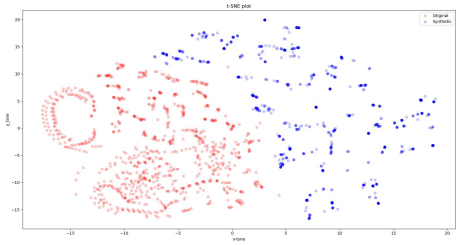
(A) UCI-HAR with GRU module



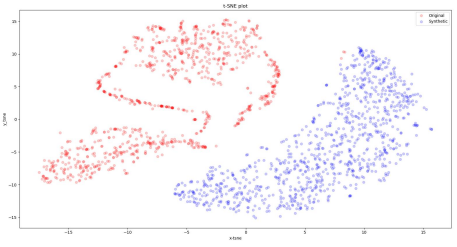
(B) GREYC-NISLAB with GRU module



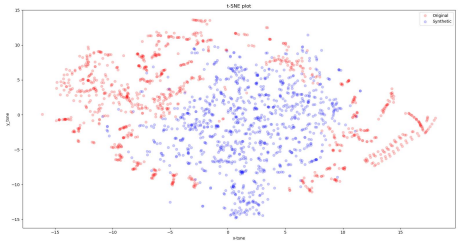
(C) UCI-HAR with LSTM module



(D) GREYC-NISLAB with LSTM module



(E) UCI-HAR with LSTM LN module



(F) GREYC-NISLAB with LSTM LN module

FIGURE 4.5: The first column corresponds to the UCI-HAR signals dataset (applied on GRU 4.5a, LSTM 4.5c, and LSTM LN 4.5e) and shows the t-SNE visualization. The second column, the t-SNE visualization of keystroke dynamics dataset signals applied on (GRU 4.5b, LSTM 4.5d, and LSTM LN 4.5f). The real dataset is in *red color*, and the synthetic dataset is in *blue*.

After visually inspecting the data using t-SNE, we used the predictive score score as an objective metric. The results are shown in Figure 4.5. In this figure, KD refers to GREYC-NISLAB and stands for Keystroke Dynamics and HAR refers to the UCI-HAR dataset and stands for Human Activity Recognition dataset. We have a less *predictive score* with the KD from the GREYC-NISLAB dataset with a value of 0.037 on the GRU module than others modules as presented in Figure 4.6.

As shown in Figure 4.6, TimeGAN consistently generates better quality synthetic data than the benchmarks, based on the predictive scores (mean absolute error) for both behavioral datasets (UCI-HAR and GREYC-NISLSAB dataset). Significantly, TimeGAN's predictive scores almost match those of the original datasets themselves. When the predictive score is low, the similarity between the synthetic data representations is important.

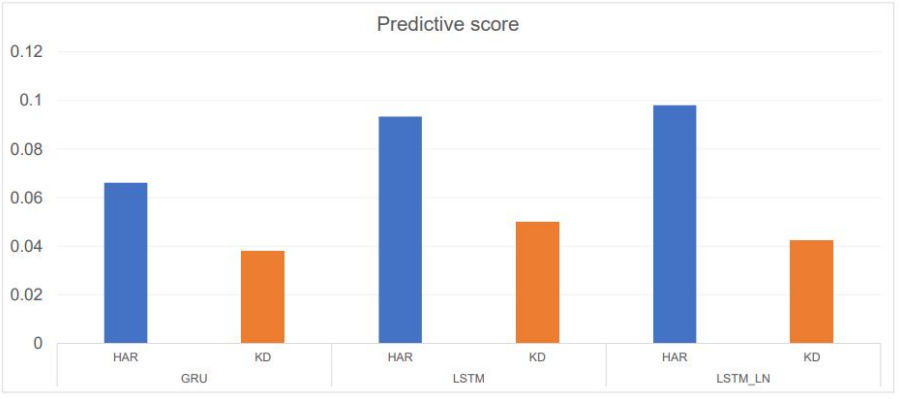


FIGURE 4.6: Time GAN with predictive score (MAE): train on synthetic, test on real.

4.5.2 Performance evaluation

With regard to the results obtained for the authentication process on our two behavioral modality datasets, as presented in Chapter 3, it is not necessary to use all the architectures described in this chapter, because:

- GoogleNet offered better results ($EER = 04.89\%$) on the real GREYC-NISLAB
- ShuffleNet offered better results ($EER = 03.57\%$) on the real UCI-HAR

In order to remain consistent, we only retained the best architectures for evaluation on synthetic bases:

- GoogleNet on the synthetic GREYC-NISLAB
- ShuffleNet on the synthetic UCI-HAR base

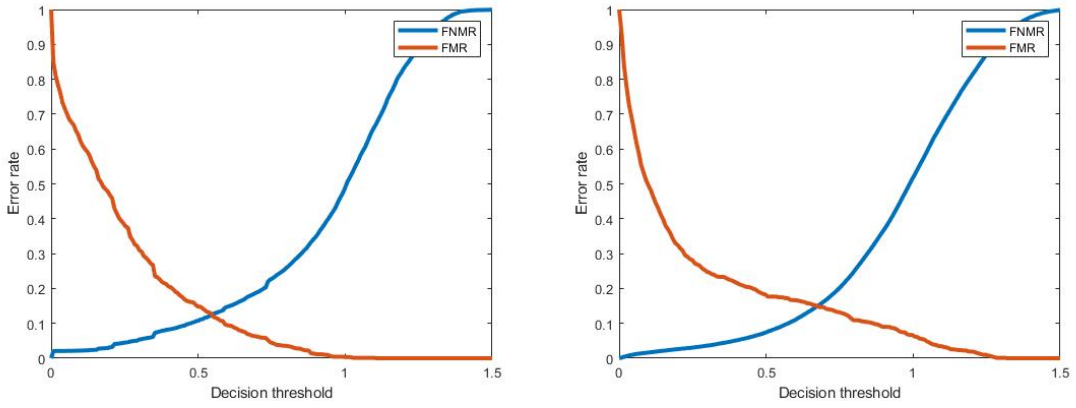
We use the synthetic dataset on the keystroke dynamics modalities obtained with the GRU module for this performance analysis since it performs better than LSTM and LSTM LN modules for each behavioral biometric data used in the context of the synthetic time series generation as depicted in Figure 4.6.

The performance analysis of the synthetic behavioral biometric modalities (keystroke dynamics) is given by Figure 4.7a. The given performance of the generated synthetic keystroke dynamics data from the GREYC-NISLAB dataset extracted with GoogleNet architectures draws an EER score equal to 12.48%, Figure 4.7a.

The EER score value obtained is based on the feature generation models of Chapter 3.

For the UCI-HAR synthetic dataset:

- The given results in Figure 4.7a are validated on the pre-trained networks (GoogleNet) (real dataset), with an $EER = 12.92\%$



(A) GREYC-NISLAB on GoogleNet: EER= 12.48% (B) UCI-HAR on ShuffleNet: EER= 14.92%

FIGURE 4.7: Performance evaluation of the synthesis GREYC-NISLAB with GoogleNet 4.7a; and performance evaluation of the synthesis UCI-HAR with ShuffleNet 4.7b

- The given results in Figure 4.7b are validated on the pre-trained networks (ShuffleNet) (real dataset), with an $EER = 14.92\%$

TABLE 4.3: Performance metrics comparison.

Dataset	Classifier name	Type of data	EER
UCI-HAR	ShuffeNet	Real Synthetic	03.57% 14.92%
GREYC-NISLAB	GoogleNet	Real Synthetic	04.89% 12.48%

Table 4.3 draws the performance metrics comparison between realistic and synthetic UCI-HAR and GREYC-NISLAB datasets.

4.5.3 Discussion

The TimeGAN is characterized by the supervised loss, embedding networks, and the joint training scheme.

The *predictive score* used as an objective metric helps to evaluate the generated data. We achieved the capacity of our model by carrying out the training on the synthetic dataset, followed by a test phase on the real dataset. Specifically, we use the Post-hoc RNN architecture to classify original data and synthetic data and to predict one-step ahead (last feature) [Yoon et al., 2019]. We evaluate the performance by computing the mean absolute error (MAE). To interpret these results, when the *predictive score* approaches zero, this indicates that the synthetic signals generated are significantly similar to the real ones.

The generation of synthetic signals also raises ethical and legal issues that require particular attention. These synthetic signals can be used in potentially sensitive contexts, such as media manipulation, the creation of falsified content, or the collection of personal data.

It is imperative to ensure that the use of synthetic behavior respects the rights and privacy of individuals, as well as applicable regulations. This means ensuring that synthetic behavioral data are not used for malicious purposes, securing the personal data used in

the generation process, and putting in place measures to prevent the dissemination of misleading or falsified content.

In general TimeGAN used for synthetic behavioral data generation offers a wide range of possibilities for emerging applications. Examples of potential applications include data augmentation, simulation, healthcare, multimedia content production through signal generation, and the use of synthetic signals in virtual reality.

The limitations of generating synthetic behavioral data with TimeGAN include convergence and stability problems, as well as the difficulty of controlling the generation and characteristics of synthetic data.

4.6 Conclusion

This chapter offers an in-depth exploration of synthetic signal generation using TimeGAN, covering key topics such as the theoretical foundations of GANs, methodology for synthetic signal generation, model variations, evaluation techniques, and associated limitations. The generation of synthetic behavioral signals using TimeGAN is presented as a crucial tool with wide applications in various fields, offering solutions to data limitations, enabling the simulation of complex scenarios, facilitating the creation of large datasets, and improving control over data generation. However, challenges in terms of convergence, stability, and ethical considerations persist, requiring ongoing research efforts. This chapter highlights the importance of interdisciplinary collaboration to exploit the full potential of the technology and calls for continued exploration of emerging applications, paving the way for new advances and opportunities in synthetic signal generation with GANs.

In summary, the use of TimeGAN for synthetic signal generation offers vast potential in many fields. It can be used to expand datasets, simulate complex scenarios, and meet specific needs. However, further research is needed to overcome the challenges and fully exploit this technology. By pursuing these efforts, we can pave the way for new advances, new applications, and new opportunities in the field of synthetic signal generation, particularly in behavioral biometrics.

Conclusion and Future works

This Ph.D. thesis has explored the field of behavioral biometrics in-depth, focusing on system evaluation and certification, description of behavioral biometric modalities, transactional applications of behavioral biometrics data, and the generation of synthetic data from behavioral biometrics data. First of all, behavioral biometrics has considerable potential in authentication and security, offering advantages such as non-intrusiveness and the ability to capture the unique characteristics of each individual. However, it also raises challenges, particularly with regard to high-quality data collection, privacy protection, and security against presentation attacks.

In the first chapter, we set the context for the certification of behavioral biometric systems. With regard to certification, we examined standards such as FIDO Alliance and ISO/IEC, which provide valuable guidelines for assessing and guaranteeing the quality and security of behavioral biometric systems. These standards play a crucial role in the development and adoption of these systems.

The second chapter offered an in-depth description of behavioral biometrics, exploring existing models, the advantages and disadvantages of different modalities, as well as ethical and legal implications. We also followed recent developments in this constantly evolving field.

The third chapter focused on transactional applications of behavioral biometrics, proposing an architecture for identification with various machine learning methods and authentication (ResNet-101, ShuffleNet, GoogleNet, and DarkNet-53) on the GREYCNIS-LAB (for keystroke dynamics) and UCI-HAR (for human activity) benchmark dataset for authentication applications. based on this behavioral modality. We presented experimental results relevant to these applications, highlighting performance and challenges. The results showed that using a combination of motion sensor data resulted in the lowest Equal Error Rate (EER) for binary classification. These experimental results demonstrated the feasibility of these approaches, paving the way for broader applications in the field of IT security. Since in the state of the art, results are given for the activity classification, our second contribution is the use of another signal-to-image transformation (a bijective transformation) of the input data, which leads to improved authentication results.

The fourth chapter introduced the generation of synthetic behavioral biometrics using generative adversarial networks (GANs). We explained the fundamental principles of GANs and detailed the TimeGAN methodology for the synthetic signal generation of behavioral biometric data. We also evaluated the performance of these models and examined their potential applications.

Contributions

The aim of this Ph.D. thesis was to propose a generic method for analyzing behavioral biometrics by presenting a generic architecture that can be adapted to any behavioral biometric modality. The use of a signal-to-image representation to present behavioral data makes it possible to handle various types of behavioral biometric modalities.

We began by proposing a generic user identification process for behavioral biometrics. Secondly, we proposed a signal-to-image transformation for the behavioral biometrics data transformation. Thirdly, we proposed a generic workflow for user authentication in the context of a certification system, and to conclude, we set out an architecture that authenticates a user through his synthetic behavioral biometric data. To achieve this, we use the

most popular temporal adversary generators (TimeGAN) to create synthetic behavioral biometrics, which are then used to impersonate an authorized user.

Future Works

For future research, we plan to add psychological features such as the user's emotions when achieving an activity for the training and testing process to improve accuracy since emotional states could be identified from his or her input behavioral style. We plan to explore biases in behavioral modalities related to gender, age, hand, and ethnicity.

Although this Ph.D. thesis has covered many aspects of behavioral biometrics, it is essential to note that this field is constantly growing and offers many research opportunities to explore in the future. These perspectives include privacy, and investigating approaches to guaranteeing user confidentiality when using behavioral biometric data while addressing the biases associated with the modality. In addition, multimodal integration needs to be further explored to improve the accuracy and reliability of biometric systems. Security against presentation attacks remains a major challenge, requiring in-depth research into advanced attack detection and resistance to adversarial attacks. Contribution to the development of specific standards and regulations for behavioral biometrics, ensuring compliance with security and data protection requirements, is also essential.

Compared to other biometric modalities such as fingerprint, there is no quality measurement for behavioral biometric data, which could be interesting to develop for the certification of behavioral biometric systems.

Finally, the continued evaluation and certification of behavioral biometrics systems is necessary to guarantee their long-term performance. In conclusion, behavioral biometrics is a dynamic field that promises significant advances, and future work will play a key role in exploiting its full potential while meeting the challenges ahead.

French synthesis

This short part provides a synthesis of this thesis in French.

Position de la thèse

La prolifération de la technologie biométrique a ouvert la voie à des méthodes d'identification et d'authentification innovantes et sûres. Cette technologie comprend un large éventail de techniques et de méthodologies permettant de reconnaître et de vérifier les individus sur la base de leurs caractéristiques physiologiques ou comportementales uniques [Piugie et al., 2022]. Dans notre travail, nous examinons les processus d'enrôlement, de vérification et d'identification, ainsi que les aspects architecturaux des systèmes biométriques. Une compréhension plus approfondie de ces éléments est essentielle pour saisir les nuances de la certification biométrique comportementale. La biométrie comportementale, un sous-ensemble de la biométrie, met l'accent sur les traits comportementaux distinctifs des individus, telles que la dynamique de la frappe au clavier, la reconnaissance de la démarche et autres modalités comportementales illustrées par la Figure 4.8. La motivation qui sous-tend l'utilisation de la biométrie comportementale réside dans son potentiel à offrir une sécurité, une praticité et une polyvalence accrues dans un grand nombre de domaines. À mesure que la technologie progresse, il devient essentiel de comprendre la certification de ces systèmes.

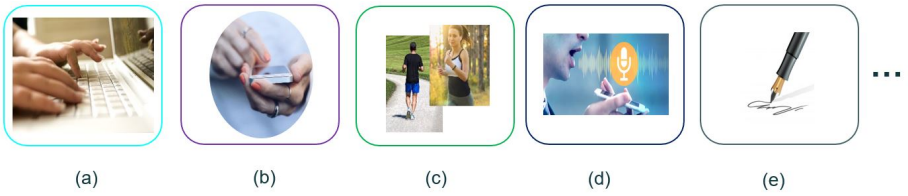


FIGURE 4.8: Caractéristiques comportementales : (a) dynamique de frappe au clavier, (b) écran tactile, (c) démarche ou activité humaine, (d) voix, et (e) signature.

Certification des systèmes de biométrie comportementale

La certification des systèmes de biométrie comportementale est un aspect crucial dans le domaine de la technologie biométrique, car elle garantit la précision et la fiabilité de ces systèmes pour diverses applications. Ce travail examine le contexte, la motivation, les méthodes d'évaluation et les normes associées à la certification de la biométrie comportementale, afin de souligner l'importance de ce domaine.

L'évaluation des performances des systèmes biométriques est une étape fondamentale de la certification. Nous examinons les méthodes générales utilisées pour évaluer l'efficacité des systèmes de biométrie comportementale. Il s'agit notamment d'évaluer les facteurs affectant les performances du système, de mesurer la qualité des modèles biométriques comportementaux et d'explorer les techniques de génération de bases d'attaques logiques.

L'alliance FIDO est une organisation leader dans le domaine de l'authentification dont la mission déclarée est de développer et de promouvoir des normes d'authentification qui contribuent à réduire la dépendance excessive du monde aux mots de passe . Elle a défini des critères de certification pour les systèmes biométriques. Nous examinons également les normes de certification biométrique de FIDO, en donnant un aperçu de la mesure des performances et de la présentation des critères de détection des attaques. La compréhension de ces normes est essentielle, car elles influencent la certification des systèmes biométriques comportementaux.

TABLE 4.4: Exigences biométriques par niveau (BioLevel) [Schuckers et al., 2023]. **O** : obligatoire à ; **F** : Facultatif à.

	BioLevel 1	**BioLevel 1+**	**BioLevel 2**	**BioLevel 2+**
#Sujets pour FAR/FRR	25	245	25	245
#Sujets pour PAD	15	15	15	15
FAR testé en laboratoire	1%	.01%	1%	.01%
FRR testé en laboratoire	7%	5%	7%	5%
IAPAR testé en laboratoire	15%	15%	7%	7%
#Instruments A/B	6/8	6/8	6/8	6/8
#IAPAR Sujets	15	15	15	15
Self Attestation FAR	**O** <= $1/10k$	**F** <= $1/10k$	**O** <= $1/10k$	**F** <= $1/10k$
Self Attestation FRR	**O** <= 5%	**F** <= 5%	**O** <= 5%	**F** <= 5%

Les critères de certification des composants biométriques FIDO sont répertoriés dans le Tableau 4.4. Sauf indication contraire, toutes ces exigences sont obligatoires pour la certification. Il existe deux niveaux de certification, chacun ayant des seuils différents pour la métrique **IAPAR** dans l'évaluation de la détection des attaques par présentation (PAD). Pour le reste, la procédure de test demeure identique pour les deux niveaux.

Une autre norme importante dans le domaine de la biométrie comportementale est la norme ISO/IEC 39794-17. Nous explorons la description, les modèles de données de reconnaissance de la marche et le flux de données de reconnaissance de la marche tels qu'ils

sont décrits dans cette norme. Il est essentiel de comprendre cette norme pour aligner les systèmes de biométrie comportementale sur les références mondiales.

Objectif

L'objectif globale de cette thèse est triple. Premièrement, nous cherchons à évaluer les différents facteurs qui affectent la performance des systèmes de biométrie comportementale. Deuxièmement, nous cherchons à établir des méthodes pour mesurer la qualité des modèles biométriques comportementaux. Enfin, nous cherchons à explorer la génération de bases d'attaques logiques pour améliorer la sécurité des systèmes de biométrie comportementale. Ces objectifs constituent le cœur de notre recherche et guident notre exploration tout au long de ces travaux de recherche.

L'objectif spécifique est d'évaluer les facteurs influençant la performance des systèmes comportementaux, de définir la qualité des modèles biométriques comportementaux, et de générer des bases d'attaques logiques pour ces systèmes. Ces éléments contribuent à la compréhension globale de l'évaluation des systèmes biométriques comportementaux et à leur utilisation sécurisée dans divers domaines.

Dans le domaine de la biométrie, les systèmes multimodaux qui combinent plusieurs caractéristiques biométriques pour l'authentification gagnent en importance. Nous discutons de la fusion au niveau de la décision, de la fusion au niveau du score, de la fusion au niveau de la caractéristique et de la fusion au niveau de l'échantillon, qui sont tous des aspects cruciaux du test des implémentations biométriques multimodales. La compréhension de ces techniques de fusion est essentielle pour obtenir des systèmes biométriques robustes et fiables. Nous donnons un aperçu des sujets essentiels qui sont couverts dans notre étude complète sur la certification des systèmes de biométrie comportementale. Nous examinons de plus près chacun de ces domaines, en mettant en lumière les complexités et l'importance de la biométrie comportementale dans le monde moderne de l'authentification et de l'identification.

Ce travail est une Thèse de doctorat CIFRE [2] (collaboration avec une entreprise) entre l'équipe Biométrie de Fime SAS et l'équipe SAFE [3] du laboratoire GREYC de l'ENSICAEN et de l'Université de Caen Normandie. Fime SAS est une société de paiement par carte et de tests biométriques (Facial, Fingerprint, Iris, Voice, Palm Vein systems) pour les certifications en France. C'est l'un des principaux produits de test au monde pour une large gamme de clients et de technologies avec un savoir-faire acquis depuis plus de 20 ans. Le GREYC est un laboratoire de recherche en informatique dédié à la modélisation, à la recherche méthodologique et aux applications pratiques dans le domaine des sciences numériques. Le GREYC est reconnu pour ses contributions originales, ses développements matériels et logiciels, ses expériences validées, ainsi que ses collaborations interdisciplinaires dans les domaines des sciences humaines et sociales et de l'interaction entre l'informatique, les mathématiques et les sciences de l'ingénieur. Il est également reconnu pour ses réalisations concrètes grâce à de solides partenariats académiques et industriels.

Plan de la thèse

Dans le chapitre 1, nous posons les bases en examinant le contexte essentiel de la biométrie comportementale dans le cadre de la certification. Nous commençons par établir le contexte de la thèse en mettant en lumière l'importance de la certification des systèmes biométriques. Nous examinons les spécificités de la technologie biométrique, notamment l'enrôlement, la vérification et l'identification, ainsi que les aspects architecturaux des systèmes biométriques. Nous abordons également l'évaluation des systèmes biométriques, en décrivant les méthodes générales utilisées pour évaluer leur efficacité. En outre, nous explorons les normes et les critères établis par des organisations telles que FIDO Alliance et ISO/IEC $39794-17$, en mettant en lumière leur rôle dans la certification de la biométrie comportementale.

[2] CIFRE signifie Convention Industrielles de Formation par la REcherche. La recherche entreprise par un boursier CIFRE s'inscrit dans le cadre d'un partenariat public et privé entre une entreprise française et un laboratoire et est formulée par les deux parties.

[3] Normandie Université, UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France

Technologie biométrique

Les systèmes biométriques fonctionnent selon trois modes : l'enrôlement, la vérification/authentification et l'identification :

- **Enrôlement**

  L'enrôlement est la première phase de tout système biométrique. C'est l'étape au cours de laquelle un utilisateur est enregistré dans le système pour la première fois. Elle est commune à la vérification et à l'identification. Lors de l'enrôlement, la caractéristique biométrique est mesurée à l'aide d'un capteur biométrique afin d'en extraire une représentation numérique. Cette représentation est ensuite réduite, à l'aide d'un algorithme d'extraction bien défini, afin de diminuer la quantité de données à stocker, ce qui facilite la vérification et l'identification. En fonction de l'application et du niveau de sécurité requis, le modèle biométrique sélectionné est protégé (crypté, par exemple) et stocké soit dans une base de données centrale, soit dans un élément personnel propre à chaque individu ;

- **Vérification**

  La vérification de l'identité consiste à vérifier si la personne qui utilise le système est bien celle qu'elle prétend être. Le système compare les informations biométriques acquises avec le modèle de référence biométrique correspondant stocké dans la base de données, ce que l'on appelle un contre un (1 contre 1). Dans ce cas, le système ne renvoie qu'une décision binaire (oui ou non), qui peut être pondérée. Le processus de vérification peut être formalisé comme suit : étant donné le vecteur d'entrée C_U définissant les caractéristiques biométriques de l'utilisateur U extraites par le système, et M_U son modèle biométrique stocké dans la base de données, le système renvoie une valeur booléenne à la suite du calcul de la fonction f définie par :

$$f(C_U, M_U) = \begin{cases} 1 & \text{si} \quad S(C_U, M_U) \geq \theta \\ 0 & \text{sinon} \end{cases} \tag{4.9}$$

où S est la fonction de similarité définissant la correspondance entre les deux vecteurs biométriques, et θ est le seuil de décision à partir duquel les deux vecteurs sont considérés comme identiques ;

- **Identification**

  En mode identification, le système biométrique détermine l'utilisateur d'un individu inconnu à partir d'une base de données d'identités, ce que l'on appelle un test "un pour tous" (1 vs N). Dans ce cas, le système peut soit attribuer à l'individu inconnu l'utilisateur correspondant au profil le plus proche trouvé dans la base de données (ou une liste de profils proches), soit rejeter l'individu. Le processus d'identification peut être formalisé comme suit : Étant donné un vecteur d'entrée C_U définissant les caractéristiques biométriques extraites par le système lorsqu'un utilisateur U se présente à lui, l'identification revient à déterminer l'identité de I_t, $t \in \{0, 1, \ldots, N\}$ où $I_1, \ldots, I_N$ sont les identités des utilisateurs précédemment inscrits dans le système, et I_0 indique un utilisateur inconnu. La fonction d'identification f peut donc être définie comme suit :

$$f(C_U) = \begin{cases} I_k & \text{si} \quad \max_{1 \leq k \leq N} S(C_U, M_k) \geq \theta \\ I_0 & \text{sinon} \end{cases} \tag{4.10}$$

  où M_k est le modèle biométrique correspondant à l'identité I_k, S est la fonction de similarité et θ est le seuil de décision.

Évaluation des systèmes biométriques : méthode générale

L'objectif de l'évaluation des systèmes biométriques est de réduire les limitations (performances, limitations culturelles, vulnérabilité à des attaques spécifiques). L'évaluation de ces systèmes est généralement effectuée selon trois aspects, comme le montre la figure 4.9.

Afin de comparer différents systèmes biométriques, qu'ils appartiennent à la même modalité ou à des modalités différentes, il est essentiel de les évaluer. De nombreuses études ont déjà abordé l'évaluation des systèmes biométriques [Conklin et al., 2004, Theofanos et al.,

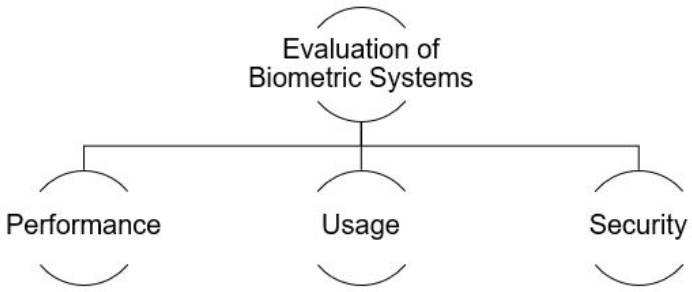


FIGURE 4.9: Aspects de l'évaluation des systèmes biométriques.

2008, Busch, 2023], et l'objectif de ce chapitre 1 est de présenter les méthodologies couramment utilisées. Lors de l'évaluation d'un système biométrique, trois aspects principaux doivent être pris en compte :

1. **Performance** [ISO 19795-1, 2021]. Il vise à quantifier diverses mesures statistiques concernant les performances du système, telles que le taux d'égal erreur (en anglais EER), le taux de fausse acceptation (en anglais FAR), le taux de faux rejet (en anglais FRR), l'échec à l'inscription (en anglais FTE), l'échec à l'acquisition (en anglais FTA) et les courbes ROC [ISO 19795-1, 2021, Busch, 2023]).

2. **Acceptabilité**, qui mesure l'acceptabilité et la satisfaction des utilisateurs lors de l'utilisation des systèmes biométriques [Theofanos et al., 2008]. Il s'agit davantage de la perception et de l'adhésion des individus, fournissant des informations sur ces aspects plutôt que sur les performances du système en matière d'erreurs.

3. **Sécurité**, qui mesure la robustesse d'un système biométrique (capteur et algorithmes) contre la fraude ISO [19792, 2008]. Elle évalue le niveau de sécurité du système en mesurant le nombre potentiel de fraudes qu'un imposteur ou un dissimulateur peut commettre.

Où :

- FAR (False Acceptance Rate) représente le pourcentage d'imposteurs acceptés à tort par le système.

- FRR (False Rejection Rate) représente le pourcentage d'utilisateurs rejetés à tort.
- EER (Equal Error Rate) représente le taux d'erreur correspondant à un réglage du seuil de décision du système biométrique de manière à ce que la valeur FAR soit égale à FRR.
- La courbe ROC est utilisée pour représenter l'efficacité d'un système biométrique. Elle représente l'évolution de la FAR en fonction du FRR.

Dans un monde idéal, un système parfait a un score d'$EER = 0$. Dans la pratique, c'est presque impossible car il est compliqué d'obtenir un FAR et un FRR proches de zéro étant donné la variabilité intrinsèque de la capture des données biométriques. Par conséquent, un FRR et un FAR faibles sont synonymes d'une meilleure sécurité.

Le chapitre 2 se focalise sur la biométrie comportementale. Tout d'abord, nous retraçons l'évolution des solutions biométriques représentée par la Figure 4.10, en soulignant la transition vers des méthodes plus subtiles, basées sur le comportement.

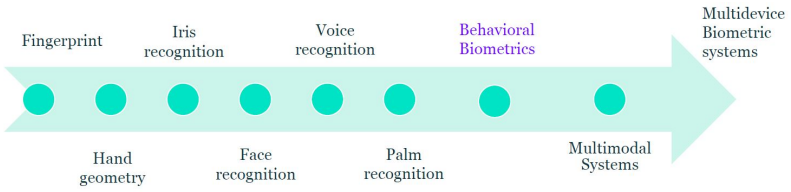


FIGURE 4.10: Évolution des solutions biométriques.

Nous examinons les objectifs et les problèmes liés à la collecte de données biométriques comportementales, en soulignant les défis en termes de respect de la vie privée et de conformité réglementaire. Cette analyse est complétée par une exploration approfondie des principales modalités de la biométrie comportementale, de ses avantages et d'une revue de la littérature. Nous abordons également la collecte et l'analyse des données biométriques comportementales, en examinant les méthodes de collecte, les types de capteurs utilisés et les précautions nécessaires pour garantir la qualité et la sécurité des données. Enfin, nous explorons les applications actuelles et les tendances émergentes de la biométrie comportementale, en examinant son rôle dans la sécurité et la santé, tout en identifiant les

limites et les risques potentiels associés à son utilisation. Nous concluons ce chapitre en soulignant les développements récents qui façonnent ce domaine en constante évolution.

Ce chapitre 2 souligne que la biométrie comportementale est une information personnelle collectée à partir des comportements individuels, tels que la manière de saisir un mot de passe ou de marcher. Ces données sont de plus en plus utilisées pour l'identification et l'authentification, la détection des fraudes et la surveillance de la santé.

Les principales caractéristiques de la biométrie comportementale sont sa nature unique et individuelle, sa capacité à être collectée de manière non intrusive et continue, et sa capacité à être utilisée en combinaison avec d'autres facteurs biométriques pour améliorer la précision de l'identification et de l'authentification.

La biométrie comportementale peut améliorer la sécurité des applications, notamment l'authentification des utilisateurs et la détection des intrusions, avec un impact minimal sur les utilisateurs. Toutefois, son efficacité dépend de la méthode de mise en œuvre, comme la dynamique de la frappe qui est influencée par le type de clavier. Les systèmes multimodaux bénéficient davantage de la biométrie comportementale, qui implique l'utilisation simultanée de plusieurs types de systèmes biométriques, que les systèmes unimodaux qui reposent sur un seul type de biométrie. Selon [Saeed, 2016], les attaques par usurpation multiples peuvent constituer une menace pour la sécurité de la biométrie comportementale.

Le chapitre 3 explore les applications transactionnelles de la biométrie comportementale en mettant l'accent sur l'identification et l'authentification. Nous donnons une vue d'ensemble de ce segment, préparant le terrain pour notre exploration approfondie. Il présente les subtilités des travaux connexes dans des domaines tels que l'analyse des séries temporelles, les activités humaines et la dynamique de la frappe au clavier. Nous présentons l'architecture proposée, y compris les algorithmes de génération, d'identification, et de mise en correspondance des caractéristiques. Nous décrivons ensuite le protocole expérimental, en détaillant les ensembles de données utilisés, les mesures de performance, les modèles pré-entraînés et les paramètres des classificateurs. Le chapitre est ensuite consacré à

l'identification de l'utilisateur, en abordant les techniques classiques d'apprentissage automatique et d'apprentissage profond, et en engageant des discussions autour de ces méthodologies. L'authentification de l'activité de l'utilisateur est étudiée, avec un examen des attentes en matière de performances, des performances pour une seule activité et des scénarios multi-activités. Enfin, l'authentification par clé d'utilisateur est abordée, avec une analyse des performances sur différents ensembles de données et scénarios.

L'objectif de chapitre est de présenter une architecture générique qui peut être adaptée à n'importe quelle modalité biométrique comportementale. L'utilisation d'une représentation signal-image pour présenter les données comportementales permet de traiter différents types de modalités biométriques comportementales.

En conclusion, ce chaptire présente une approche basée sur la biométrie comportementale multimodale pour l'identification et l'authentification des utilisateurs, en utilisant l'apprentissage automatique. Les activités humaines sur les smartphones et la dynamique de frappe sur les ordinateurs portables sont étudiées en tant que modalités biométriques. L'apprentissage profond permet d'obtenir des performances de pointe dans la classification des séries temporelles, mais pose des problèmes de protection de la vie privée. Dans le contexte de l'authentification, une méthode basée sur la dynamique de la frappe est proposée, offrant des avantages tels qu'un faible coût, mais aussi des problèmes potentiels de confidentialité. Enfin, l'authentification basée sur les activités humaines montre des résultats prometteurs en utilisant des combinaisons de capteurs, et l'utilisation de la transformation signal-image améliore les performances.

Le chapitre 4 aborde le domaine de la génération de données biométriques comportementales synthétiques à l'aide de réseaux adverses génératifs (GAN). Nous donnons un aperçu de cette approche, suivi d'une introduction qui contextualise l'importance de la génération de données synthétiques car l'introduction des réseaux de type GAN a considérablement facilité les progrès dans la génération de données synthétiques. Ces modèles d'apprentissage profond se composent généralement de deux réseaux neuronaux : un générateur et un discriminateur. Le générateur G prend un vecteur de bruit aléatoire $z \in \mathbb{R}^r$ et vise à produire des données synthétiques similaires à la distribution des données d'apprentissage.

Parallèlement, le discriminateur D tente de déterminer si les données générées sont authentiques ou fausses. Le générateur cherche à maximiser le taux d'échec du discriminateur, tandis que le discriminateur cherche à le minimiser. La figure 4.11 illustre de manière concise l'architecture du GAN et le jeu auquel se livrent ces deux réseaux neuronaux. Ces deux réseaux sont engagés dans un jeu à somme nulle, défini par la fonction de valeur V(G, D) comme suit :

$$\min_G \max_D V(G,D) \;=\; \mathbb{E}_{x\sim P_{data}(x)}\left[\log D(x)\right] \;+\; \mathbb{E}_{z\sim P_z(z)}\left[\log(1 \;-\; D(G(z)))\right] \qquad (4.11)$$

où $D(.)$ représente la probabilité que les données proviennent de données réelles plutôt que de données générées [Goodfellow et al., 2014].

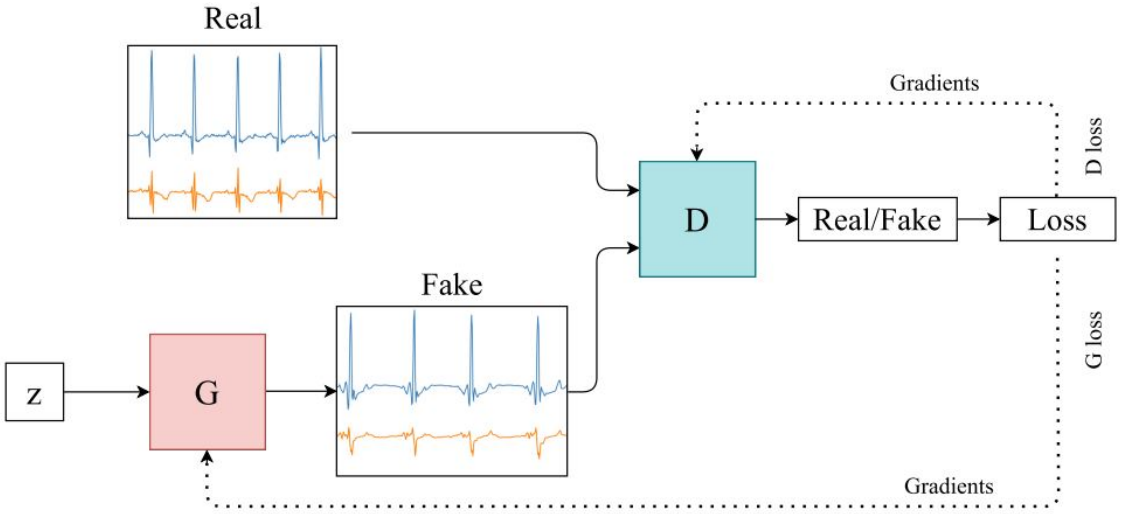


FIGURE 4.11: Réseaux adverses génératifs [Brophy et al., 2023].

Les GANs appartiennent à la famille des modèles génératifs et sont une méthode alternative pour produire des données synthétiques, ne nécessitant pas d'expertise spécifique dans le domaine [Brophy et al., 2023]. Ils ont été initialement introduits par Goodfellow et al [Goodfellow et al., 2014] en 2014, en utilisant un perceptron multicouche à la fois comme discriminateur et comme générateur.

Par la suite, dans ce dernier chapitre, nous passons en revue les travaux connexes dans le domaine et approfondissons les principes et l'architecture des réseaux de type GAN. Nous étudions spécifiquement le réseau TimeGAN représenté par la Figure 4.12, une

méthodologie spécialisée pour la génération de signaux synthétiques, y compris son processus d'apprentissage et les techniques permettant d'améliorer la génération de signaux.

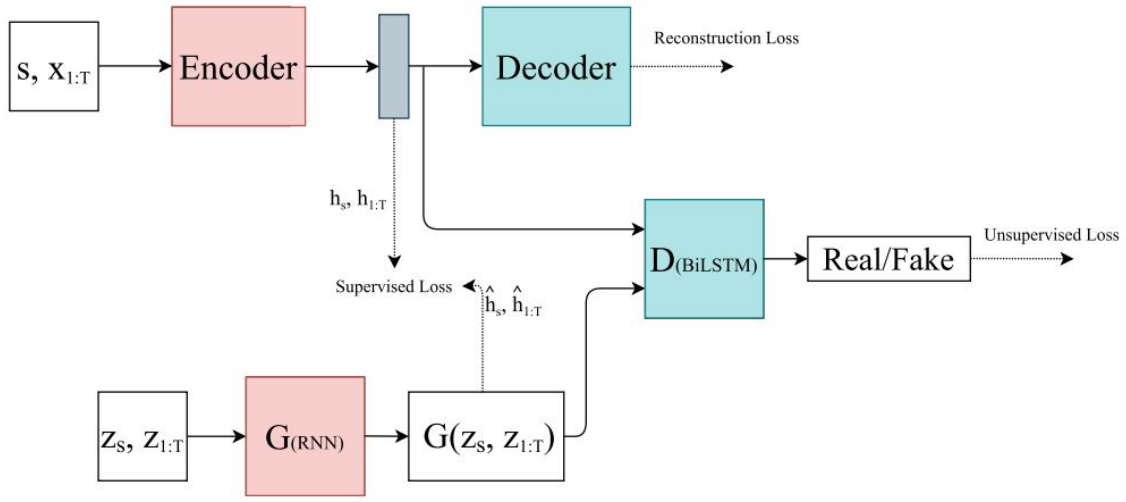


FIGURE 4.12: Architecture du TimeGAN.

- L'entrée est considéré comme composé de deux éléments distincts : une caractéristique statique, représentée par le vecteur **s**, et une caractéristique temporelle, représentée par le vecteur **x** à l'entrée du codeur.
- Le générateur prend en entrée un tuple de vecteurs de caractéristiques statiques et temporelles générés aléatoirement à partir d'une distribution préalablement définie.
- Les codes latents réels ($\mathbf{h}$) et synthétiques ($\hat{\mathbf{h}}$) sont utilisés pour calculer l'élément de perte supervisée de ce réseau.
- Le discriminateur évalue l'ensemble des codes latents en les classant comme réels (y) ou synthétiques ($\hat{y}$), et l'opérateur $\sim$ indique que l'échantillon est soit réel, soit faux.

Une section importante est consacrée aux mesures d'évaluation des modèles utilisés pour générer des signaux synthétiques, y compris l'évaluation des performances de TimeGAN. Les applications de la génération de signaux synthétiques sont discutées, et le chapitre se termine par une discussion et une conclusion générale.

Contributions de la thèse

Une première contribution de cette thèse de doctorat réside dans la proposition d'une méthode générique d'analyse de la biométrie comportementale, destinée à des applications telles que la dynamique de frappe au clavier et les activités humaines. Nous évaluons l'efficacité des techniques classiques d'apprentissage automatique pour l'identification, tout en examinant les méthodes d'apprentissage profond pour l'authentification des utilisateurs en se basant sur leurs comportements, avec une emphase sur l'activité humaine sur les smartphones et la dynamique de frappe au clavier sur les ordinateurs portables.

Le pipeline complet est présenté dans la Figure 4.13. Il englobe quatre composants essentiels, à savoir :

1. La transformation de séries temporelles en représentations signal-image
2. Générations/extractions de caractéristiques
3. Les apprentissages machine et profond (pour les méthodes de transfert)
4. L'identification et l'authentification (pour les applications)

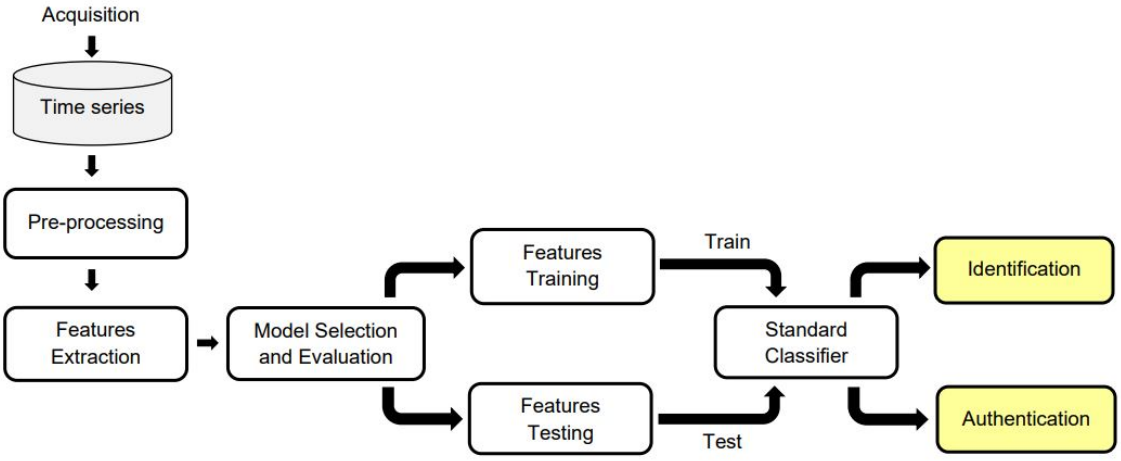


FIGURE 4.13: Vue d'ensemble de notre système générique proposé.

Cette thèse de doctorat propose également une deuxième contribution qui est focalisée sur le traitement des données biométriques comportementales considérées comme des

séries temporelles. Cette méthode permet d'obtenir des résultats très différents de ceux déjà disponibles. Le traitement des séries temporelles consiste à transformer les données biométriques comportementales brutes en des images couleur 2D. Ce processus de transformation conserve toutes les caractéristiques du signal comportemental. La série temporelle ne reçoit aucune opération de filtrage avec cette transformation et la méthode est réversible.

Cette transformation du signal en image nous permet d'utiliser les réseaux convolutionnels 2D pour construire des vecteurs de caractéristiques profondes efficaces. Cela nous permet de comparer ces vecteurs de caractéristiques aux vecteurs du modèle de référence pour calculer la métrique de performance.

Nous évaluons les performances du système d'authentification en termes de Taux d'Égal Erreur (TÉE) sur des ensembles de données de référence et nous montrons l'efficacité de l'approche.

Cependant, certaines méthodes d'authentification traditionnelles se sont révélées insuffisantes pour assurer une protection adéquate des données, d'où l'importance croissante de la biométrie comportementale. Malgré des résultats prometteurs et un large éventail d'applications, les systèmes biométriques restent vulnérables aux attaques malveillantes, en particulier aux attaques par présentation.

C'est pourquoi, dans cette thèse de doctorat, comme troisième contribution, nous avons entrepris de déployer une attaque de présentation contre un système d'authentification basé sur la biométrie comportementale. Notre approche consiste à utiliser les générateurs adversaires temporels les plus populaires (TimeGAN) pour créer des données biométriques comportementales synthétiques, qui pourraient être utilisées pour usurper l'identité d'un utilisateur autorisé. Ces données synthétiques sont générées tout en préservant la dynamique temporelle, ce qui signifie que les nouvelles séquences respectent les relations originales entre les variables au fil du temps.

Notre contribution sur la génération des données biométriques comportementales é-tiques est illustrée par la Figure 4.14. Cette figure montre notre proposition de la chaine

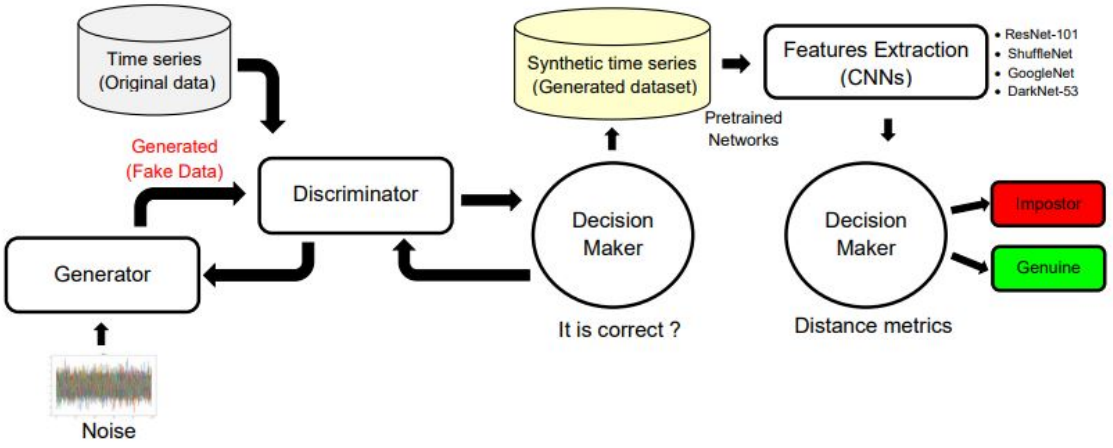


FIGURE 4.14: Vue d'ensemble du système d'évaluation des attaques.

de traitement générique, qui intègre le générateur adversarial temporel pour le module de génération de données de séries temporelles synthétiques, ainsi que le module d'appariement. L'objectif est de générer des données biométriques comportementales synthétiques afin de simuler de l'identité d'un utilisateur autorisé dans le cadre de notre système d'authentification décrit précédemment dans le chapitre 3. Ce processus générique a été conçu pour valider la méthode que nous proposons.

En outre, un système d'authentification a été mis en place pour évaluer l'efficacité des données générées. Les résultats obtenus, ainsi qu'une inspection visuelle, indiquent que TimeGAN peut effectivement générer des modèles comportementaux qui peuvent être utilisés pour tromper et par conséquent tester les comportements des utilisateurs.

Conclusion

Cette thèse de doctorat a exploré en profondeur le domaine de la biométrie comportementale, en se concentrant sur l'évaluation et la certification des systèmes, la description des modalités de la biométrie comportementale, les applications transactionnelles des données biométriques comportementales et la génération de données synthétiques à partir des données biométriques comportementales. Tout d'abord, la biométrie comportementale présente un potentiel considérable en matière d'authentification et de sécurité, car elle offre des avantages tels que l'absence d'intrusion et la capacité de capturer les caractéristiques

uniques de chaque individu. Cependant, elle soulève également des défis, notamment en ce qui concerne la collecte de données de haute qualité, la protection de la vie privée et la sécurité contre les attaques par présentation.

Dans le premier chapitre, nous avons défini le contexte de la certification des systèmes de biométrie comportementale. En ce qui concerne la certification, nous avons examiné des normes telles que FIDO Alliance et ISO/IEC, qui fournissent des lignes directrices précieuses pour évaluer et garantir la qualité et la sécurité des systèmes de biométrie comportementale. Ces normes jouent un rôle crucial dans le développement et l'adoption de ces systèmes.

Le deuxième chapitre propose une description approfondie de la biométrie comportementale, en explorant les modèles existants, les avantages et les inconvénients des différentes modalités, ainsi que les implications éthiques et juridiques. Nous avons également suivi les développements récents dans ce domaine en constante évolution.

Le troisième chapitre s'est concentré sur les applications transactionnelles de la biométrie comportementale, en proposant une architecture pour l'identification avec différentes méthodes d'apprentissage automatique et d'authentification (ResNet-101, ShuffleNet, GoogleNet, et DarkNet-53) sur l'ensemble de données de référence GREYC-NISLAB (pour la dynamique de la frappe) et UCI-HAR (pour l'activité humaine) pour les applications d'authentification. basées sur cette modalité comportementale. Nous avons présenté des résultats expérimentaux pertinents pour ces applications, en soulignant les performances et les défis. Les résultats ont montré que l'utilisation d'une combinaison de données de capteurs de mouvement a permis d'obtenir le taux d'égal erreur (TEE) le plus bas pour la classification binaire. Ces résultats expérimentaux ont démontré la faisabilité de ces approches, ouvrant la voie à des applications plus larges dans le domaine de la sécurité informatique. Étant donné que, dans l'état actuel des connaissances, les résultats sont donnés pour la classification des activités, notre deuxième contribution est l'utilisation d'une autre transformation signal-image (une transformation bijective) des données d'entrée, qui permet d'améliorer les résultats de l'authentification.

Le quatrième chapitre a présenté la génération de données biométriques comportementales synthétiques à l'aide de réseaux adversaires génératifs (GAN). Nous avons expliqué

les principes fondamentaux des GANs et détaillé la méthodologie du TimeGAN pour la génération de signaux synthétiques de données biométriques comportementales. Nous avons également évalué la performance de ces modèles et examiné leurs applications potentielles.

Perspectives

Pour les recherches futures, nous prévoyons d'ajouter des caractéristiques psychologiques telles que les émotions de l'utilisateur lors de la réalisation d'une activité pour le processus de formation et de test afin d'améliorer la précision puisque les états émotionnels pourraient être identifiés à partir de son style comportemental d'entrée. Nous prévoyons d'explorer les biais dans les modalités comportementales liés au sexe, à l'âge, à la main et à l'origine ethnique.

Bien que cette thèse de doctorat ait couvert de nombreux aspects de la biométrie comportementale, il est essentiel de noter que ce domaine est en constante croissance et offre de nombreuses opportunités de recherche à explorer dans le futur. Ces perspectives comprennent la protection de la vie privée et l'étude d'approches visant à garantir la confidentialité de l'utilisateur lors de l'utilisation de données biométriques comportementales, tout en tenant compte des biais associés à la modalité. En outre, l'intégration multimodale doit être étudiée plus avant afin d'améliorer la précision et la fiabilité des systèmes biométriques. La sécurité contre les attaques de présentation reste un défi majeur, nécessitant des recherches approfondies sur la détection avancée des attaques et la résistance aux attaques adverses. Il est également essentiel de contribuer à l'élaboration de normes et de réglementations spécifiques pour la biométrie comportementale, afin de garantir le respect des exigences en matière de sécurité et de protection des données.

Par rapport à d'autres modalités biométriques telles que les empreintes digitales, il n'existe pas de mesure de la qualité des données biométriques comportementales, ce qui pourrait être intéressant à développer pour la certification des systèmes biométriques comportementaux.

Enfin, l'évaluation et la certification continues des systèmes de biométrie comportementale sont nécessaires pour garantir leur performance à long terme. En conclusion, la biométrie comportementale est un domaine dynamique qui promet des avancées significatives, et les travaux futurs joueront un rôle clé dans l'exploitation de son plein potentiel tout en relevant les défis à venir.

List of Publications

Journal

- **Yris Brice Wandji Piugie**, Christophe Charrier, Joël Di Manno, Christophe Rosenberger, "Deep Features Fusion for User Authentication Based on Human Activity," in *IET Biometrics Journal*, vol. 12, no. 4, pp. 222-234, July 2023. [Wandji Piugie et al., 2023] **(ranked Q2)**

International Conference

- **Yris Brice Wandji Piugie**, Joël Di Manno, Christophe Rosenberger, Christophe Charrier, "Keystroke Dynamics based User Authentication using Deep Learning Neural Networks," *International Conference on Cyberworlds (CW)*, Kanazawa, Japan, 2022, p. 220-227. [Piugie et al., 2022] **(ranked CORE B)**.

- **Yris Brice Wandji Piugie**, Joël Di Manno, Christophe Rosenberger, Christophe Charrier, "How Artificial Intelligence can be used for Behavioral Identification?" *International Conference on Cyberworlds (CW). IEEE*, Caen, France, 2021, pp. 246-253. [Piugie et al., 2021] **(ranked CORE B)**.

- Cyrius Nugier, Diane Leblanc-Albarel, Agathe Blaise, Simon Masson, Paul Huynh and **Yris Brice Wandji Piugie**, " An Upcycling Tokenization Method for Credit Card Numbers," *SECRYPT 2021-18th International Conference on Security and Cryptography*, Paris, France, 2021, pp. 1-12. [Nugier et al., 2021] **(rank CORE B)**.

Honors & Awards

- **Best Full paper award**
  International Conference CYBERWORLDS 2022 in Kanazawa, Japan
  with the paper entitled "*Keystroke Dynamics based User Authentication using Deep Learning Neural Networks*"

Poster

- Yris Brice Wandji Piugie, Joël Di Manno, Christophe Rosenberger, and Christophe Charrier. 2022. Keystroke dynamics-based user authentication using deep learning neural networks **(DFKI-INRIA summer school, Saarbruken-Germany)**.
- Yris Brice Wandji Piugie, Joël Di Manno, Christophe Rosenberger, and Christophe Charrier. 2021. How artificial intelligence can be used for behavioral identification? **At 2021 International Conference on Cyberworlds (CW), Caen-France**.

Bibliography

[19792, 2008] 19792, I. F. (2008). Information technology — security techniques – security evaluation of biometrics. Standard ISO/IEC FCD 19792, International Organization for Standardization, Geneva, CH.

[19795-1, 2021] 19795-1, I. (2021). Information technology — Biometric performance testing and reporting — part 1: Principles and framework. Standard ISO/IEC 19795-1:2021, International Organization for Standardization, Geneva, CH.

[Abdelfattah et al., 2018] Abdelfattah, S. M., Abdelrahman, G. M., and Wang, M. (2018). Augmenting the size of eeg datasets using generative adversarial networks. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6. IEEE.

[Acien et al., 2021] Acien, A., Morales, A., Monaco, J. V., Vera-Rodriguez, R., and Fierrez, J. (2021). Typenet: Deep learning keystroke biometrics. *arXiv preprint arXiv:2101.05570*.

[Ahmed and Traore, 2007] Ahmed, A. A. E. and Traore, I. (2007). A new biometric technology based on mouse dynamics. *IEEE Transactions on dependable and secure computing*, 4(3):165–179.

[Ahmed and Traore, 2010] Ahmed, A. A. E. and Traore, I. (2010). Mouse dynamics biometric technology. In *Behavioral Biometrics for Human Identification: Intelligent Applications*, pages 207–223. IGI Global.

[Aimarre, 1999] Aimarre, J. (1999). *Crise de foie*. Editions du Pénitent, Paris.

[Akber et al., 2023] Akber, S. M. A., Kazmi, S. N., Mohsin, S. M., and Szczesna, A. (2023). Deep learning-based motion style transfer tools, techniques and future challenges. *Sensors*, 23(5):2597.

[Al Machot et al., 2020] Al Machot, F., R Elkobaisi, M., and Kyamakya, K. (2020). Zero-shot human activity recognition using non-visual sensors. *Sensors*, 20(3):825.

[Albarel., 2021] Albarel., D. L. (2021). Tokenization. `https://github.com/DianeLeblancAlbarel/Tokenisation`.

[Albera et al., 2012] Albera, L., Kachenoura, A., Comon, P., Karfoul, A., Wendling, F., Senhadji, L., and Merlet, I. (2012). Ica-based eeg denoising: a comparative analysis of fifteen methods. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, 60(3):407–418.

[Alex et al., 2018] Alex, P. M. D., Ravikumar, A., Selvaraj, J., and Sahayadhas, A. (2018). Research on human activity identification based on image processing and artificial intelligence. *Int. J. Eng. Technol*, 7.

[Ali et al., 2011] Ali, H., Dargham, J., Ali, C., and Moung, E. G. (2011). Gait recognition using gait energy image. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 4(3):141–152.

[Almeida et al., 2007] Almeida, P. S., Baquero, C., Preguiça, N., and Hutchison, D. (2007). Scalable Bloom Filters. *Information Processing Letters*, 101(6):255 – 261.

[Alpar, 2017] Alpar, O. (2017). Frequency spectrograms for biometric keystroke authentication using neural network based classifier. *Knowledge-Based Systems*, 116:163–171.

[Alpaydin, 2020] Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.

[Alsaadi, 2021] Alsaadi, I. M. (2021). Study on most popular behavioral biometrics, advantages, disadvantages and recent applications: A review. *Int. J. Sci. Technol. Res*, 10:15–21.

[Alzantot et al., 2017] Alzantot, M., Chakraborty, S., and Srivastava, M. (2017). Sensegen: A deep learning architecture for synthetic sensor data generation. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 188–193. IEEE.

[Alzubaidi and Kalita, 2016] Alzubaidi, A. and Kalita, J. (2016). Authentication of smartphone users using behavioral biometrics. *IEEE Communications Surveys & Tutorials*, 18(3):1998–2026.

[Amin and Yan, 2010] Amin, M. A. and Yan, H. (2010). Gabor wavelets in behavioral biometrics. In *Behavioral biometrics for human identification: intelligent applications*, pages 121–150. IGI Global.

[Andrean et al., 2020] Andrean, A., Jayabalan, M., and Thiruchelvam, V. (2020). Keystroke dynamics based user authentication using deep multilayer perceptron. *International Journal of Machine Learning and Computing*, 10(1):134–139.

[Anguita et al., 2013] Anguita, D., Ghio, A., Oneto, L., Parra, X., and Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. In *Esann*, volume 3, page 3.

[Antal et al., 2021] Antal, M., Fejér, N., and Buza, K. (2021). Sapimouse: Mouse dynamics-based user authentication using deep feature learning. In *2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, pages 61–66. IEEE.

[Antón et al., 2019] Antón, M. Á., Ordieres-Meré, J., Saralegui, U., and Sun, S. (2019). Non-invasive ambient intelligence in real life: Dealing with noisy patterns to help older people. *Sensors*, 19(14):3113.

[Ashibani and Mahmoud, 2020] Ashibani, Y. and Mahmoud, Q. H. (2020). A multi-feature user authentication model based on mobile app interactions. *IEEE Access*, 8:96322–96339.

[Aversano et al., 2021] Aversano, L., Bernardi, M. L., Cimitile, M., and Pecori, R. (2021). Continuous authentication using deep neural networks ensemble on keystroke dynamics. *PeerJ Computer Science*.

[Ayotte et al., 2021a] Ayotte, B., Banavar, M. K., Hou, D., and Schuckers, S. (2021a). Group leakage overestimates performance: A case study in keystroke dynamics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1410–1417.

[Ayotte et al., 2021b] Ayotte, B., Banavar, M. K., Hou, D., and Schuckers, S. (2021b). Study of intra-and inter-user variance in password keystroke dynamics. In *ICISSP*, pages 467–474.

[Azizi et al., 2023] Azizi, S., Kornblith, S., Saharia, C., Norouzi, M., and Fleet, D. J. (2023). Synthetic data from diffusion models improves imagenet classification. *arXiv preprint arXiv:2304.08466*.

[Bagnall et al., 2017] Bagnall, A., Lines, J., Bostrom, A., Large, J., and Keogh, E. (2017). The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data mining and knowledge discovery*, 31(3):606–660.

[Bahdanau et al., 2016] Bahdanau, D., Brakel, P., Xu, K., Goyal, A., Lowe, R., Pineau, J., Courville, A., and Bengio, Y. (2016). An actor-critic algorithm for sequence prediction. *arXiv preprint arXiv:1607.07086*.

[Bailey et al., 2014] Bailey, K. O., Okolica, J. S., and Peterson, G. L. (2014). User identification and authentication using multi-modal behavioral biometrics. *Computers & Security*, 43:77–89.

[Banerjee and Woodard, 2012] Banerjee, S. P. and Woodard, D. L. (2012). Biometric authentication and identification using keystroke dynamics: A survey. *Journal of Pattern Recognition Research*, 7(1):116–139.

[Barra et al., 2019] Barra, P., Bisogni, C., Nappi, M., Freire-Obregón, D., and Castrillón-Santana, M. (2019). Gait analysis for gender classification in forensics. In *Dependability in Sensor, Cloud, and Big Data Systems and Applications: 5th International Conference, DependSys 2019, Guangzhou, China, November 12–15, 2019, Proceedings 5*, pages 180–190. Springer.

[Bausinger and Tabassi, 2011] Bausinger, O. and Tabassi, E. (2011). Fingerprint sample quality metric nfiq 2.0. *BIOSIG 2011–Proceedings of the Biometrics Special Interest Group*.

[Bellare et al., 2009] Bellare, M., Ristenpart, T., Rogaway, P., and Stegers, T. (2009). Format-Preserving Encryption. In Jacobson, M. J., Rijmen, V., and Safavi-Naini, R., editors, *Selected Areas in Cryptography*, pages 295–312. Springer Berlin Heidelberg.

[Belouchrani et al., 1997] Belouchrani, A., Abed-Meraim, K., Cardoso, J.-F., and Moulines, E. (1997). A blind source separation technique using second-order statistics. *IEEE Transactions on signal processing*, 45(2):434–444.

[BenAbdelkader et al., 2004] BenAbdelkader, C., Cutler, R. G., and Davis, L. S. (2004). Gait recognition using image self-similarity. *EURASIP Journal on Advances in Signal Processing*, 2004(4):1–14.

[Bengio et al., 2015] Bengio, S., Vinyals, O., Jaitly, N., and Shazeer, N. (2015). Scheduled sampling for sequence prediction with recurrent neural networks. *Advances in neural information processing systems*, 28.

[Bengio et al., 2009] Bengio, Y., Louradour, J., Collobert, R., and Weston, J. (2009). Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*, pages 41–48.

[Bergadano et al., 2003] Bergadano, F., Gunetti, D., and Picardi, C. (2003). Identity verification through dynamic keystroke analysis. *Intelligent Data Analysis*, 7(5):469–496.

[Bhana and Flowerday, 2020] Bhana, B. and Flowerday, S. (2020). Passphrase and keystroke dynamics authentication: Usable security. *Computers & Security*, 96:101925.

[Bhatnagar et al., 2013] Bhatnagar, M., Jain, R. K., and Khairnar, N. S. (2013). A survey on behavioral biometric techniques: mouse vs keyboard dynamics. *Int. J. Comput. Appl*, 975:8887.

[Bhattacharyya et al., 2009] Bhattacharyya, D., Ranjan, R., Alisherov, F., Choi, M., et al. (2009). Biometric authentication: A review. *International Journal of u-and e-Service, Science and Technology*, 2(3):13–28.

[Bhowmik et al., 2022a] Bhowmik, A., Sannigrahi, M., Chowdhury, D., and Das, D. (2022a). Ricecloud: A cloud integrated ensemble learning based rice leaf diseases prediction system. In *2022 IEEE 19th India Council International Conference (INDICON)*, pages 1–6. IEEE.

[Bhowmik et al., 2022b] Bhowmik, A., Sannigrahi, M., Chowdhury, D., Dwivedi, A. D., and Mukkamala, R. R. (2022b). Dbnex: Deep belief network and explainable ai based financial fraud detection. In *2022 IEEE International Conference on Big Data (Big Data)*, pages 3033–3042. IEEE.

[Bhowmik et al., 2023] Bhowmik, A., Sannigrahi, M., Dutta, P. K., and Bandyopadhyay, S. (2023). Using edge computing framework with the internet of things for intelligent vertical gardening. In *2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC)*, pages 1–6. IEEE.

[Bhowmik et al., 2022c] Bhowmik, A., Sannigrahi, M., Guha, P., Chowdhury, D., and Gill, S. S. (2022c). Dynamite: Dynamic aggregation of mutually-connected points based clustering algorithm for time series data. *Internet Technology Letters*, page e395.

[Bicon, 1901] Bicon, P. (1901). *De la memoire.* Serigraph, Paris.

[Bloom, 1970] Bloom, B. H. (1970). Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM*, 13:422–426.

[Bonomi et al., 2006] Bonomi, F., Mitzenmacher, M., Panigrahy, R., Singh, S., and Varghese, G. (2006). An Improved Construction for Counting Bloom Filters. In Azar, Y. and Erlebach, T., editors, *Algorithms – ESA 2006*, pages 684–695. Springer Berlin Heidelberg.

[Borga, 1998] Borga, M. (1998). *Learning multidimensional signal processing.* PhD thesis, Linköping University Electronic Press.

[Borga and Knutsson, 2001] Borga, M. and Knutsson, H. (2001). A canonical correlation approach to blind source separation. *Report LiU-IMT-EX-0062 Department of Biomedical Engineering, Linkping University.*

[Borji, 2019] Borji, A. (2019). Pros and cons of gan evaluation measures. *Computer vision and image understanding*, 179:41–65.

[Borowik et al., 2020] Borowik, P., Adamowicz, L., Tarakowski, R., Siwek, K., and Grzywacz, T. (2020). Odor detection using an e-nose with a reduced sensor array. *Sensors*, 20(12):3542.

[Breiman, 2001] Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

[Brophy, 2020] Brophy, E. (2020). Synthesis of dependent multichannel ecg using generative adversarial networks. In *Proceedings of the 29th ACM international conference on information & knowledge management*, pages 3229–3232.

[Brophy et al., 2023] Brophy, E., Wang, Z., She, Q., and Ward, T. (2023). Generative adversarial networks in time series: A systematic literature review. *ACM Computing Surveys*, 55(10):1–31.

[Brown and Rogers, 1993] Brown, M. and Rogers, S. J. (1993). User identification via keystroke characteristics of typed names using neural networks. *International Journal of Man-Machine Studies*, 39(6):999–1014.

[Bryant and Yarnold, 1995] Bryant, F. B. and Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis.

[Bud, 2018] Bud, A. (2018). Facing the future: The impact of apple faceid. *Biometric technology today*, 2018(1):5–7.

[Busch, 2023] Busch, C. (2023). Standards for biometric presentation attack detection. In *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection and Vulnerability Assessment*, pages 571–583. Springer.

[Cachin et al., 2017] Cachin, C., Camenisch, J., Freire-Stögbuchner, E., and Lehmann, A. (2017). Updatable Tokenization: Formal Definitions and Provably Secure Constructions. In Kiayias, A., editor, *Financial Cryptography and Data Security*, pages 59–75. Springer International Publishing.

[Cardoso and Souloumiac, 1993] Cardoso, J.-F. and Souloumiac, A. (1993). Blind beamforming for non-gaussian signals. In *IEE proceedings F (radar and signal processing)*, volume 140, pages 362–370. IET.

[CardRates, 2020] CardRates (2020). The Average Number of Credit Card Transactions Per Day & Year. `https://www.cardrates.com/advice/number-of-credit-card-transactions-per-day-year/`.

[Cavaro-Ménard et al., 2008] Cavaro-Ménard, C., Naït-Ali, A., Tanguy, J.-Y., Angelini, E., Le Bozec, C., and Le Jeune, J.-J. (2008). Specificities of physiological signals and medical images. *Compression of Biomedical Images and Signals*, pages 43–76.

[Çeker and Upadhyaya, 2017] Çeker, H. and Upadhyaya, S. (2017). Sensitivity analysis in keystroke dynamics using convolutional neural networks. In *2017 IEEE Workshop on Information Forensics and Security (WIFS)*, pages 1–6. IEEE.

[Chai et al., 2022] Chai, T., Li, A., Zhang, S., Li, Z., and Wang, Y. (2022). Lagrange motion analysis and view embeddings for improved gait recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20249–20258.

[Chang et al., 2021] Chang, H.-C., Li, J., Wu, C.-S., and Stamp, M. (2021). Machine learning and deep learning for fixed-text keystroke dynamics. *arXiv preprint arXiv:2107.00507*.

[Chen et al., 2019] Chen, Y., Lv, Y., and Wang, F.-Y. (2019). Traffic flow imputation using parallel data and generative adversarial networks. *IEEE Transactions on Intelligent Transportation Systems*, 21(4):1624–1630.

[Chen et al., 2018] Chen, Y., Wang, Y., Kirschen, D., and Zhang, B. (2018). Model-free renewable scenario generation using generative adversarial networks. *IEEE Transactions on Power Systems*, 33(3):3265–3275.

[Chen et al., 2021] Chen, Z., Cai, H., Jiang, L., Zou, W., Zhu, W., and Fei, X. (2021). Keystroke dynamics based user authentication and its application in online examination. In *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 649–654. IEEE.

[Cheng et al., 2020] Cheng, P.-S., Lai, C.-Y., Chang, C.-C., Chiou, S.-F., and Yang, Y.-C. (2020). A variant model of tgan for music generation. In *Proceedings of the 2020 asia service sciences and software engineering conference*, pages 40–45.

[Cherrier, 2021] Cherrier, E. (2021). *Authentification biométrique: comment (ré) concilier sécurité, utilisabilité et respect de la vie privée?* PhD thesis, Normandie Université.

[Chevalier et al., 2005] Chevalier, P., Albera, L., Ferréol, A., and Comon, P. (2005). On the virtual array concept for higher order array processing. *IEEE Transactions on Signal Processing*, 53(4):1254–1271.

[Choi et al., 2021] Choi, M., Lee, S., Jo, M., and Shin, J. S. (2021). Keystroke dynamics-based authentication using unique keypad. *Sensors*, 21(6):2242.

[Clark et al., 2017] Clark, G. W., Doran, M. V., and Andel, T. R. (2017). Cybersecurity issues in robotics. In *2017 IEEE conference on cognitive and computational aspects of situation management (CogSIMA)*, pages 1–5. IEEE.

[Comon and Jutten, 2010] Comon, P. and Jutten, C. (2010). *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press.

[Conklin et al., 2004] Conklin, A., Dietrich, G., and Walz, D. (2004). Password-based authentication: a system perspective. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*, pages 10–pp. IEEE.

[Crihan et al., 2023] Crihan, G., Crăciun, M., and Dumitriu, L. (2023). A comparative assessment of homomorphic encryption algorithms applied to biometric information. *Inventions*, 8(4):102.

[Cui et al., 2016] Cui, Z., Chen, W., and Chen, Y. (2016). Multi-scale convolutional neural networks for time series classification. *arXiv preprint arXiv:1603.06995*.

[Dai and Le, 2015] Dai, A. M. and Le, Q. V. (2015). Semi-supervised sequence learning. *Advances in neural information processing systems*, 28.

[De Marsico et al., 2016] De Marsico, M., De Pasquale, D., and Mecca, A. (2016). Embedded accelerometer signal normalization for cross-device gait recognition. In *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–5. IEEE.

[Delaney et al., 2019] Delaney, A. M., Brophy, E., and Ward, T. E. (2019). Synthesis of realistic ecg using generative adversarial networks. *arXiv preprint arXiv:1909.09150*.

[Demšar et al., 2013] Demšar, J., Curk, T., Erjavec, A., Črt Gorup, Hočevar, T., Milutinovič, M., Možina, M., Polajnar, M., Toplak, M., Starič, A., Štajdohar, M., Umek, L., Žagar, L., Žbontar, J., Žitnik, M., and Zupan, B. (2013). Orange: Data mining toolbox in python. *Journal of Machine Learning Research*, 14:2349–2353.

[Doan et al., 2002] Doan, A., Madhavan, J., Domingos, P., and Halevy, A. (2002). Learning to map between ontologies on the semantic web. In *Proceedings of the* 11^{th} *international conference on World Wide Web*, pages 662–673. ACM.

[Dua et al., 2017] Dua, D., Graff, C., et al. (2017). Uci machine learning repository.

[Dumoulin et al., 2016] Dumoulin, V., Belghazi, I., Poole, B., Mastropietro, O., Lamb, A., Arjovsky, M., and Courville, A. (2016). Adversarially learned inference. *arXiv preprint arXiv:1606.00704*.

[Durak and Vaudenay, 2017] Durak, F. B. and Vaudenay, S. (2017). Breaking the FF3 Format-Preserving Encryption Standard over Small Domains. In Katz, J. and Shacham, H., editors, *Advances in Cryptology – CRYPTO 2017*, pages 679–707, Cham. Springer International Publishing.

[Duzdevich et al., 2014] Duzdevich, D., Redding, S., and Greene, E. C. (2014). Dna dynamics and single-molecule biology. *Chemical reviews*, 114(6):3072–3086.

[Díaz-Santiago et al., 2014] Díaz-Santiago, S., Rodriguez-Henriquez, L. M., and Chakraborty, D. (2014). A cryptographic study of tokenization systems. In *11th International Conference on Security and Cryptography (SECRYPT)*, pages 1–6.

[Eizaguirre-Peral et al., 2022] Eizaguirre-Peral, I., Segurola-Gil, L., and Zola, F. (2022). Conditional generative adversarial network for keystroke presentation attack. *arXiv preprint arXiv:2212.08445*.

[El-Abed, 2011] El-Abed, M. (2011). *Évaluation de système biométrique*. PhD thesis, Université de Caen.

[El Zein and Kalakech, 2018] El Zein, D. and Kalakech, A. (2018). Feature selection for android keystroke dynamics. In *2018 International Arab Conference on Information Technology (ACIT)*, pages 1–6. IEEE.

[Engelsma and Jain, 2019] Engelsma, J. J. and Jain, A. K. (2019). Generalizing fingerprint spoof detector: Learning a one-class classifier. In *2019 International Conference on Biometrics (ICB)*, pages 1–8. IEEE.

[Esteban et al., 2017] Esteban, C., Hyland, S. L., and Rätsch, G. (2017). Real-valued (medical) time series generation with recurrent conditional gans. *arXiv preprint arXiv:1706.02633*.

[Estrela et al., 2020] Estrela, P. M. A. B., de Oliveira Albuquerque, R., Amaral, D. M., Giozza, W. F., Nze, G. D. A., and de Mendonça, F. L. L. (2020). Biotouch: a framework based on behavioral biometrics and location for continuous authentication on mobile banking applications. In *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–6. IEEE.

[Fabius and Van Amersfoort, 2014] Fabius, O. and Van Amersfoort, J. R. (2014). Variational recurrent auto-encoders. *arXiv preprint arXiv:1412.6581*.

[Fahimi et al., 2019] Fahimi, F., Zhang, Z., Goh, W. B., Ang, K. K., and Guan, C. (2019). Towards eeg generation using gans for bci applications. In *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pages 1–4. IEEE.

[Fan et al., 2014] Fan, B., Andersen, D. G., Kaminsky, M., and Mitzenmacher, M. D. (2014). Cuckoo filter: Practically better than bloom. In *Proceedings of the 10th ACM International on Conference on Emerging Networking Experiments and Technologies*, page 75–88. Association for Computing Machinery.

[Fawaz et al., 2019] Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., and Muller, P.-A. (2019). Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, 33(4):917–963.

[Feng, 2017] Feng, F. (2017). *Séparation aveugle de source: de l'instantané au convolutif.* PhD thesis, Universite paris sud.

[Févotte and Doncarli, 2004] Févotte, C. and Doncarli, C. (2004). Two contributions to blind source separation using time-frequency distributions. *IEEE Signal Processing Letters*, 11(3):386–389.

[Frank et al., 2012] Frank, M., Biedert, R., Ma, E., Martinovic, I., and Song, D. (2012). Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. *IEEE transactions on information forensics and security*, 8(1):136–148.

[Gafurov, 2007] Gafurov, D. (2007). A survey of biometric gait recognition: Approaches, security and challenges. In *Annual Norwegian computer science conference*, pages 19–21. Annual Norwegian Computer Science Conference Norway.

[Gafurov et al., 2006] Gafurov, D., Helkala, K., and Søndrol, T. (2006). Biometric gait authentication using accelerometer sensor. *J. comput.*, 1(7):51–59.

[Gaines et al., 1980] Gaines, R. S., Lisowski, W., Press, S. J., and Shapiro, N. (1980). Authentication by keystroke timing: Some preliminary results. Technical report, Rand Corp Santa Monica CA.

[Gamboa and Fred, 2004] Gamboa, H. and Fred, A. (2004). A behavioral biometric system based on human-computer interaction. In *Biometric Technology for Human Identification*, volume 5404, pages 381–392. International Society for Optics and Photonics.

[Ganin et al., 2016] Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., and Lempitsky, V. (2016). Domain-adversarial training of neural networks. *The journal of machine learning research*, 17(1):2096–2030.

[Gao et al., 2020] Gao, G., Yu, Y., Yang, J., Qi, G.-J., and Yang, M. (2020). Hierarchical deep cnn feature set-based representation learning for robust cross-resolution face recognition. *IEEE Transactions on Circuits and Systems for Video Technology*.

[Gao et al., 2010] Gao, J., Zheng, C., and Wang, P. (2010). Online removal of muscle artifact from electroencephalogram signals based on canonical correlation analysis. *Clinical EEG and neuroscience*, 41(1):53–59.

[George et al., 2019] George, A., Mostaani, Z., Geissenbuhler, D., Nikisins, O., Anjos, A., and Marcel, S. (2019). Biometric face presentation attack detection with multi-channel convolutional neural network. *IEEE Transactions on Information Forensics and Security*, 15:42–55.

[Georges and Dupuys, 1974] Georges, C. and Dupuys, E. (1974). La blondeur décolorée. histoire de la stupidité. *Revue des brunes*, 12:125–184.

[Ghosh et al., 2019] Ghosh, S., Hiware, K., Ganguly, N., Mitra, B., and De, P. (2019). Emotion detection from touch interactions during text entry on smartphones. *International Journal of Human-Computer Studies*, 130:47–57.

[Gibbs, 2010] Gibbs, M. D. (2010). Biometrics: body odor authentication perception and acceptance. *ACM Sigcas Computers and Society*, 40(4):16–24.

[Gilbarg and Trudinger, 2015] Gilbarg, D. and Trudinger, N. S. (2015). *Elliptic partial differential equations of second order*. Springer Publications.

[Giles et al., 2023] Giles, B., Peeling, P., Kovalchik, S., and Reid, M. (2023). Differentiating movement styles in professional tennis: A machine learning and hierarchical clustering approach. *European Journal of Sport Science*, 23(1):44–53.

[Giorgi et al., 2021] Giorgi, G., Saracino, A., and Martinelli, F. (2021). Using recurrent neural networks for continuous authentication through gait analysis. *Pattern Recognition Letters*, 147:157–163.

[Giot et al., 2015] Giot, R., Dorizzi, B., and Rosenberger, C. (2015). A review on the public benchmark databases for static keystroke dynamics. *Computers & Security*, 55:46–61.

[Giot et al., 2011] Giot, R., El-Abed, M., Hemery, B., and Rosenberger, C. (2011). Unconstrained keystroke dynamics authentication with shared secret. *Computers & security*, 30(6-7):427–445.

[Giot et al., 2009a] Giot, R., El-Abed, M., and Rosenberger, C. (2009a). Greyc keystroke: a benchmark for keystroke dynamics biometric systems. In *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–6. IEEE.

[Giot et al., 2009b] Giot, R., El-Abed, M., and Rosenberger, C. (2009b). Keystroke dynamics with low constraints svm based passphrase enrollment. In *2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pages 1–6. IEEE.

[Glorot and Bengio, 2010] Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings.

[Goethals and Zaki, 2003] Goethals, B. and Zaki, M. J. (2003). Workshop on frequent itemset mining implementations. In *Third IEEE International Conference on Data Mining Workshop on Frequent Itemset Mining Implementations*, pages 1–13.

[Goldberger et al., 2000] Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. (2000). Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220.

[González, 2023] González, N. (2023). Ksdsld—a tool for keystroke dynamics synthesis & liveness detection. *Software Impacts*, 15:100454.

[Goodfellow et al., 2014] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.

[Gunetti and Picardi, 2005] Gunetti, D. and Picardi, C. (2005). Keystroke analysis of free text. *ACM Transactions on Information and System Security (TISSEC)*, 8(3):312–347.

[Guo et al., 2019] Guo, Z., Wan, Y., and Ye, H. (2019). A data imputation method for multivariate time series based on generative adversarial network. *Neurocomputing*, 360:185–197.

[Halakou, 2013] Halakou, F. (2013). Feature selection in keystroke dynamics authentication systems. In *International Conference on Computer, Information Technology and Digital Media*.

[Halevi and Krawczyk, 1999] Halevi, S. and Krawczyk, H. (1999). Public-key cryptography and password protocols. *ACM Trans. Inf. Syst. Secur.*, 2(3):230–268.

[Han et al., 2020] Han, L., Zheng, K., Zhao, L., Wang, X., and Wen, H. (2020). Content-aware traffic data completion in its based on generative adversarial nets. *IEEE Transactions on Vehicular Technology*, 69(10):11950–11962.

[Harada et al., 2019] Harada, S., Hayashi, H., and Uchida, S. (2019). Biosignal generation and latent variable analysis with recurrent generative adversarial networks. *IEEE Access*, 7:144292–144302.

[Haradal et al., 2018] Haradal, S., Hayashi, H., and Uchida, S. (2018). Biosignal data augmentation based on generative adversarial networks. In *2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pages 368–371. IEEE.

[Hartmann et al., 2018] Hartmann, K. G., Schirrmeister, R. T., and Ball, T. (2018). Eeggan: Generative adversarial networks for electroencephalograhic (eeg) brain signals. *arXiv preprint arXiv:1806.01875*.

[Harun et al., 2010] Harun, N., Woo, W. L., and Dlay, S. (2010). Performance of keystroke biometrics authentication system using artificial neural network (ann) and distance classifier method. In *International Conference on Computer and Communication Engineering (ICCCE'10)*, pages 1–6. IEEE.

[Hazra and Byun, 2020] Hazra, D. and Byun, Y.-C. (2020). Synsiggan: Generative adversarial networks for synthetic biomedical signal generation. *Biology*, 9(12):441.

[He et al., 2016] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

[Hoang et al., 2018] Hoang, V. T., Tessaro, S., and Trieu, N. (2018). The Curse of Small Domains: New Attacks on Format-Preserving Encryption. In Shacham, H. and Boldyreva, A., editors, *Advances in Cryptology – CRYPTO 2018*, pages 221–251, Cham. Springer International Publishing.

[Hochet, 2003] Hochet, E. (2003). *Histoire d'un lac*. Van Dame, Paris.

[Hoerl and Kennard, 1970] Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: applications to nonorthogonal problems. *Technometrics*, 12(1):69–82.

[Hot and Delplanque, 2013] Hot, P. and Delplanque, S. (2013). *Electrophysiologie de la cognition*. Dunod.

[Hotelling, 1936] Hotelling, H. (1936). Relation between two sets of variates. *Biometrica*.

[Hsu et al., 2017] Hsu, W.-N., Zhang, Y., and Glass, J. (2017). Unsupervised learning of disentangled and interpretable representations from sequential data. *Advances in neural information processing systems*, 30.

[Hyvärinen et al., 2004] Hyvärinen, A., Karhunen, J., and Oja, E. (2004). *Independent component analysis*, volume 46. John Wiley & Sons.

[Idrus et al., 2013] Idrus, S. Z. S., Cherrier, E., Rosenberger, C., and Bours, P. (2013). Soft biometrics for keystroke dynamics. In *International Conference Image Analysis and Recognition*, pages 11–18. Springer.

[Idrus et al., 2014] Idrus, S. Z. S., Cherrier, E., Rosenberger, C., and Bours, P. (2014). Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords. *Computers & Security*, 45:147–155.

[Idrus et al., 2015] Idrus, S. Z. S., Cherrier, E., Rosenberger, C., Mondal, S., and Bours, P. (2015). Keystroke dynamics performance enhancement with soft biometrics. In *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2015)*, pages 1–7. IEEE.

[Imsand, 2008] Imsand, E. (2008). *Applications of GUI usage analysis*. PhD thesis.

[Inbavalli and Nandhini, 2014] Inbavalli, P. and Nandhini, G. (2014). Body odor as a biometric authentication. *International Journal of Computer Science and Information Technologies*, 5(5):6270–6274.

[International Organization for Standardization, 2017] International Organization for Standardization (2017). ISO/IEC 7812-1:2017 Identification cards – Identification of issuers – Part 1: Numbering system. Technical report, International Organization for Standardization. `https://www.iso.org/obp/ui/#iso:std:iso-iec:7812:-1:ed-5:v1:en`.

[Ioffe and Szegedy, 2015] Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR.

[ISO 19795-1, 2021] ISO 19795-1 (2021). ISO/IEC 19795-1:2021 Information technology — Biometric performance testing and reporting — Part 1: Principles and framework. Standard, International Organization for Standardization.

[ISO 19795-2, 2015] ISO 19795-2 (2015). ISO/IEC 19795-2:2007/Amd.1:2015(E) Information technology — Biometric performance testing and reporting — Part 2: Testing methodologies for technology and scenario evaluation . Standard, International Organization for Standardization.

[ISO 2382-37, 2022] ISO 2382-37 (2022). ISO/IEC 2382-37:2022 Information technology — Vocabulary — Part 37: Biometrics. Standard, International Organization for Standardization.

[ISO 24722, 2015] ISO 24722 (2015). ISO/IEC TR 24722:2015 Information technology — Biometrics — Multimodal and other multibiometric fusion. Standard, International Organization for Standardization.

[ISO 30107-3, 2023] ISO 30107-3 (2023). ISO/IEC 30107-3:2017 Information technology — Biometric presentation attack detection — Part 3: Testing and reporting. Standard, International Organization for Standardization.

[ISO 39794-17, 2021] ISO 39794-17 (2021). ISO/IEC 19795-1: 2021 Information technology — Extensible biometric data interchange formats — Part 17: Gait image sequence data. Standard, International Organization for Standardization.

[Jain and Kanhangad, 2017] Jain, A. and Kanhangad, V. (2017). Human activity classification in smartphones using accelerometer and gyroscope sensors. *IEEE Sensors Journal*, 18(3):1169–1177.

[Jain et al., 2004] Jain, A. K., Ross, A., and Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, 14(1):4–20.

[Jaouedi et al., 2020] Jaouedi, N., Boujnah, N., and Bouhlel, M. S. (2020). A new hybrid deep learning model for human action recognition. *Journal of King Saud University-Computer and Information Sciences*, 32(4):447–453.

[Jasper, 1958] Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.*, 10:370–375.

[Jin et al., 2008] Jin, Z., Teoh, A. B. J., Ong, T. S., and Tee, C. (2008). Typing dynamics biometric authentication through fuzzy logic. In *2008 International Symposium on Information Technology*, volume 3, pages 1–6. IEEE.

[Johnson et al., 2001] Johnson, D., Menezes, A., and Vanstone, S. (2001). The elliptic curve digital signature algorithm (ecdsa). 1(1):36–63.

[Jordon et al., 2018] Jordon, J., Yoon, J., and Van Der Schaar, M. (2018). Pate-gan: Generating synthetic data with differential privacy guarantees. In *International conference on learning representations*.

[Joyce and Gupta, 1990] Joyce, R. and Gupta, G. (1990). Identity authentication based on keystroke latencies. *Communications of the ACM*, 33(2):168–176.

[Junejo et al., 2008] Junejo, I. N., Dexter, E., Laptev, I., and Púrez, P. (2008). Cross-view action recognition from temporal self-similarities. In *European Conference on Computer Vision*, pages 293–306. Springer.

[Juvela et al., 2019] Juvela, L., Bollepalli, B., Yamagishi, J., and Alku, P. (2019). Waveform generation for text-to-speech synthesis using pitch-synchronous multi-scale generative adversarial networks. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6915–6919. IEEE.

[Kachenoura, 2006] Kachenoura, A. (2006). *Traitement aveugle de signaux biomédicaux*. PhD thesis, Université Rennes 1.

[Kantrop, 2005] Kantrop, J. (2005). *Le loup et ses proies*. Nuit Bleue, Paris.

[Kasprowski and Harezlak, 2018] Kasprowski, P. and Harezlak, K. (2018). Fusion of eye movement and mouse dynamics for reliable behavioral biometrics. *Pattern Analysis and Applications*, 21:91–103.

[Katiyar et al., 2013] Katiyar, R., Pathak, V. K., and Arya, K. (2013). A study on existing gait biometrics approaches and challenges. *International Journal of Computer Science Issues (IJCSI)*, 10(1):135.

[Katz and Sarnak, 1999] Katz, N. M. and Sarnak, P. (1999). *Random matrices, Frobenius eigenvalues, and monodromy*, volume 45. American Mathematical Soc.

[Kaushik et al., 2020] Kaushik, S., Choudhury, A., Natarajan, S., Pickett, L. A., and Dutt, V. (2020). Medicine expenditure prediction via a variance-based generative adversarial network. *IEEE Access*, 8:110947–110958.

[Kawulok et al., 2016] Kawulok, M., Celebi, E., and Smolka, B. (2016). *Advances in face detection and facial image analysis*. Springer.

[Khan et al., 2020] Khan, H., Hengartner, U., and Vogel, D. (2020). Mimicry attacks on smartphone keystroke authentication. *ACM Transactions on Privacy and Security (TOPS)*, 23(1):1–34.

[Khare et al., 2020] Khare, S., Sarkar, S., and Totaro, M. (2020). Comparison of sensor-based datasets for human activity recognition in wearable iot. In *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, pages 1–6. IEEE.

[Khatun et al., 2022] Khatun, M. A., Yousuf, M. A., Ahmed, S., Uddin, M. Z., Alyami, S. A., Al-Ashhab, S., Akhdar, H. F., Khan, A., Azad, A., and Moni, M. A. (2022).

Deep cnn-lstm with self-attention model for human activity recognition using wearable sensor. *IEEE Journal of Translational Engineering in Health and Medicine*, 10:1–16.

[Kim et al., 2020] Kim, D. I., Lee, S., and Shin, J. S. (2020). A new feature scoring method in keystroke dynamics-based user authentications. *IEEE Access*, 8:27901–27914.

[Kim et al., 2018] Kim, J., Kim, H., and Kang, P. (2018). Keystroke dynamics-based user authentication using freely typed text based on user-adaptive feature extraction and novelty detection. *Applied Soft Computing*, 62:1077–1087.

[Kingma and Ba, 2014] Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[Kohavi et al., 1996] Kohavi, R. et al. (1996). Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid. In *Kdd*, volume 96, pages 202–207.

[Kolokolova et al., 2020] Kolokolova, A., Billard, M., Bishop, R., Elsisy, M., Northcott, Z., Graves, L., Nagisetty, V., and Patey, H. (2020). Gans & reels: Creating irish music using a generative adversarial network. *arXiv preprint arXiv:2010.15772*.

[Konda and Tsitsiklis, 1999] Konda, V. and Tsitsiklis, J. (1999). Actor-critic algorithms. *Advances in neural information processing systems*, 12.

[Körner and Denzler, 2013] Körner, M. and Denzler, J. (2013). Temporal self-similarity for appearance-based action recognition in multi-view setups. In *International Conference on Computer Analysis of Images and Patterns*, pages 163–171. Springer.

[Kothari, 2004a] Kothari, C. R. (2004a). *Research methodology: Methods and techniques*. New Age International Publications.

[Kothari, 2004b] Kothari, C. R. (2004b). *Research methodology: Methods and techniques*. New Age International Publications.

[Krebs, Brian, 2019] Krebs, Brian (2019). A Month After 2 Million Customer Cards Sold Online, Buca di Beppo Parent Admits Breach. `https://krebsonsecurity.com/tag/davinci-breach/`.

[Kunze and Lukowicz, 2008] Kunze, K. and Lukowicz, P. (2008). Dealing with sensor displacement in motion-based onbody activity recognition systems. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 20–29.

[Lamb et al., 2016] Lamb, A. M., ALIAS PARTH GOYAL, A. G., Zhang, Y., Zhang, S., Courville, A. C., and Bengio, Y. (2016). Professor forcing: A new algorithm for training recurrent networks. *Advances in neural information processing systems*, 29.

[Larsen et al., 2016] Larsen, A. B. L., Sønderby, S. K., Larochelle, H., and Winther, O. (2016). Autoencoding beyond pixels using a learned similarity metric. In *International conference on machine learning*, pages 1558–1566. PMLR.

[Leangarun et al., 2018] Leangarun, T., Tangamchit, P., and Thajchayapong, S. (2018). Stock price manipulation detection using generative adversarial networks. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 2104–2111. IEEE.

[Lee and Mase, 2002] Lee, S.-W. and Mase, K. (2002). Activity and location recognition using wearable sensors. *IEEE pervasive computing*, 1(3):24–32.

[Li et al., 2023a] Li, C., Li, X., Chen, M., and Sun, X. (2023a). Deep learning and image recognition. In *2023 IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT)*, pages 557–562. IEEE.

[Li et al., 2019] Li, D., Chen, D., Jin, B., Shi, L., Goh, J., and Ng, S.-K. (2019). Madgan: Multivariate anomaly detection for time series data with generative adversarial networks. In *International conference on artificial neural networks*, pages 703–716. Springer.

[Li et al., 2021] Li, J., Chang, H.-C., and Stamp, M. (2021). Free-text keystroke dynamics for user authentication. *arXiv preprint arXiv:2107.07009*.

[Li et al., 2020] Li, Q., Hao, H., Zhao, Y., Geng, Q., Liu, G., Zhang, Y., and Yu, F. (2020). Gans-lstm model for soil temperature estimation from meteorological: a new approach. *IEEE Access*, 8:59427–59443.

[Li et al., 2023b] Li, Y.-K., Meng, Q.-H., Wang, Y.-X., and Hou, H.-R. (2023b). Mmfn: Emotion recognition by fusing touch gesture and facial expression information. *Expert Systems with Applications*, 228:120469.

[Lio et al., 2018] Lio, G., Thobois, S., Ballanger, B., Lau, B., and Boulinguez, P. (2018). Removing deep brain stimulation artifacts from the electroencephalogram: issues, recommendations and an open-source toolbox. *Clinical Neurophysiology*.

[Liu and Silverman, 2001] Liu, S. and Silverman, M. (2001). A practical guide to biometric security technology. *IT Professional*, 3(1):27–32.

[Liu et al., 2010] Liu, Z., Jia, C., Li, J., and Cheng, X. (2010). Format-preserving encryption for datetime. In *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*, volume 2, pages 201–205. IEEE.

[Lu et al., 2020] Lu, X., Zhang, S., Hui, P., and Lio, P. (2020). Continuous authentication by free-text keystroke based on cnn and rnn. *Computers & Security*, 96:101861.

[Luhn, 1960] Luhn, H. P. (1960). Computer For Verifying Numbers. US2950048A.

[Luo et al., 2018] Luo, Y., Cai, X., Zhang, Y., Xu, J., et al. (2018). Multivariate time series imputation with generative adversarial networks. *Advances in neural information processing systems*, 31.

[Luo et al., 2019] Luo, Y., Zhang, Y., Cai, X., and Yuan, X. (2019). E2gan: End-to-end generative adversarial network for multivariate time series imputation. In *Proceedings of the 28th international joint conference on artificial intelligence*, pages 3094–3100. AAAI Press Palo Alto, CA, USA.

[Lv et al., 2008] Lv, H.-R., Lin, Z.-L., Yin, W.-J., and Dong, J. (2008). Emotion recognition based on pressure sensor keyboards. In *2008 IEEE international conference on multimedia and expo*, pages 1089–1092. IEEE.

[Lytras et al., 2018] Lytras, M. D., Visvizi, A., Daniela, L., Sarirete, A., and Ordonez De Pablos, P. (2018). Social networks research for sustainable smart education. *Sustainability*, 10(9):2974.

[Lyu et al., 2018] Lyu, X., Hueser, M., Hyland, S. L., Zerveas, G., and Raetsch, G. (2018). Improving clinical predictions through unsupervised time series representation learning. *arXiv preprint arXiv:1812.00490*.

[Maedche and Staab, 2001] Maedche, A. and Staab, S. (2001). Ontology learning for the semantic web. *IEEE Intelligent systems*, 16(2):72–79.

[Maheshwary et al., 2017] Maheshwary, S., Ganguly, S., and Pudi, V. (2017). Deep secure: A fast and simple neural network based approach for user authentication and identification via keystroke dynamics. In *IWAISe: First International Workshop on Artificial Intelligence in Security*, volume 59.

[Mahmud et al., 2020] Mahmud, T., Sayyed, A. S., Fattah, S. A., and Kung, S.-Y. (2020). A novel multi-stage training approach for human activity recognition from multimodal wearable sensor data using deep neural network. *IEEE Sensors Journal*, 21(2):1715–1726.

[Makhzani et al., 2015] Makhzani, A., Shlens, J., Jaitly, N., Goodfellow, I., and Frey, B. (2015). Adversarial autoencoders. *arXiv preprint arXiv:1511.05644*.

[Manikantaa and Saranya, 2023] Manikantaa, C. and Saranya, G. (2023). Human odor security using e-nose check for updates v. anush kumar, cs manigandaa, s. dhanush hariharan. *Big Data and Cloud Computing: Select Proceedings of ICBCC 2022*, 1021:51.

[Mansfield, 2006] Mansfield, A. (2006). Information technology–biometric performance testing and reporting–part 1: Principles and framework. *ISO/IEC*, pages 19795–1.

[Mantyjarvi et al., 2005] Mantyjarvi, J., Lindholm, M., Vildjiounaite, E., Makela, S.-M., and Ailisto, H. (2005). Identifying users of portable devices from gait pattern with accelerometers. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 2, pages ii–973. IEEE.

[Marcel et al., 2019] Marcel, S., Nixon, M. S., Fierrez, J., and Evans, N. (2019). *Handbook of biometric anti-spoofing: Presentation attack detection*, volume 2. Springer.

[Marcel et al., 2014] Marcel, S., Nixon, M. S., and Li, S. Z. (2014). *Handbook of biometric anti-spoofing*, volume 1. Springer.

[Marsico and Mecca, 2017] Marsico, M. D. and Mecca, A. (2017). Biometric walk recognizer: gait recognition by a single smartphone accelerometer. *Multimedia Tools and Applications*, 76:4713–4745.

[Matyáš and Říha, 2002] Matyáš, V. and Říha, Z. (2002). Biometric authentication—security and usability. In *Advanced Communications and Multimedia Security: IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security September 26–27, 2002, Portorož, Slovenia*, pages 227–239. Springer.

[Mekruksavanich and Jitpattanakul, 2021] Mekruksavanich, S. and Jitpattanakul, A. (2021). Deep learning approaches for continuous authentication based on activity patterns using mobile sensing. *Sensors*, 21(22):7519.

[Mhenni et al., 2018] Mhenni, A., Cherrier, E., Rosenberger, C., and Amara, N. E. B. (2018). Towards a secured authentication based on an online double serial adaptive mechanism of users' keystroke dynamics. In *International Conference on Digital Society and eGovernments (ICDS)*.

[Migdal, 2019a] Migdal, D. (2019a). *Contributions to keystroke dynamics for privacy and security on the Internet*. Theses, Normandie Université.

[Migdal, 2019b] Migdal, D. (2019b). *Contributions to keystroke dynamics for privacy and security on the Internet*. PhD thesis, Normandie Université.

[Migdal and Rosenberger, 2019] Migdal, D. and Rosenberger, C. (2019). Statistical modeling of keystroke dynamics samples for the generation of synthetic datasets. *Future Generation Computer Systems*, 100:907–920.

[Mirza and Osindero, 2014] Mirza, M. and Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

[Mishra et al., 2023] Mishra, P., Choudhury, J. A., Kho, E., Basu, B., and Nandi, A. (2023). Speaker identification, differentiation and verification using deep learning for human machine interface. In *2023 Photonics & Electromagnetics Research Symposium (PIERS)*, pages 861–870. IEEE.

[Mogren, 2016] Mogren, O. (2016). C-rnn-gan: Continuous recurrent neural networks with adversarial training. *arXiv preprint arXiv:1611.09904*.

[Monaco, 2018] Monaco, V. (2018). Public keystroke dynamics datasets.

[Monaro et al., 2020] Monaro, M., Cannonito, E., Gamberini, L., and Sartori, G. (2020). Spotting faked 5 stars ratings in e-commerce using mouse dynamics. *Computers in Human Behavior*, 109:106348.

[Monrose and Rubin, 1997] Monrose, F. and Rubin, A. (1997). Authentication via keystroke dynamics. In *Proceedings of the 4th ACM Conference on Computer and Communications Security*, pages 48–56.

[Moore and Lopes, 1999] Moore, R. and Lopes, J. (1999). Paper templates. In *TEMPLATE'06, 1st International Conference on Template Production*. SCITEPRESS.

[Moskovitch et al., 2009] Moskovitch, R., Feher, C., Messerman, A., Kirschnick, N., Mustafic, T., Camtepe, A., Lohlein, B., Heister, U., Moller, S., Rokach, L., et al. (2009). Identity theft, computers and behavioral biometrics. In *2009 IEEE International Conference on Intelligence and Security Informatics*, pages 155–160. IEEE.

[Mowla et al., 2015] Mowla, M. R., Ng, S.-C., Zilany, M. S., and Paramesran, R. (2015). Artifacts-matched blind source separation and wavelet transform for multichannel eeg denoising. *Biomedical Signal Processing and Control*, 22:111–118.

[Muaaz and Mayrhofer, 2013] Muaaz, M. and Mayrhofer, R. (2013). An analysis of different approaches to gait recognition using cell phone based accelerometers. In *Proceedings*

of International Conference on Advances in Mobile Computing & Multimedia, pages 293–300.

[Muley and Kute, 2018] Muley, A. and Kute, V. (2018). Prospective solution to bank card system using fingerprint. In *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pages 898–902. IEEE.

[N. Owen and Shoemaker, 2008] N. Owen, W. and Shoemaker, E. (2008). Multi-factor authentication system.

[Ness, 2017] Ness, J. (2017). Presentation attack and detection in keystroke dynamics. Master's thesis, NTNU.

[Neverova et al., 2016] Neverova, N., Wolf, C., Lacey, G., Fridman, L., Chandra, D., Barbello, B., and Taylor, G. (2016). Learning human identity from motion patterns. *IEEE Access*, 4:1810–1820.

[Ni et al., 2020] Ni, H., Szpruch, L., Wiese, M., Liao, S., and Xiao, B. (2020). Conditional sig-wasserstein gans for time series generation. *arXiv preprint arXiv:2006.05421*.

[Nikolaidis et al., 2019] Nikolaidis, K., Kristiansen, S., Goebel, V., Plagemann, T., Liestøl, K., and Kankanhalli, M. (2019). Augmenting physiological time series data: A case study for sleep apnea detection. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 376–399. Springer.

[NIST, 2020] NIST (2020). Methods for Format-Preserving Encryption: NIST Requests Public Comments on Draft Special Publication 800-38G Revision 1. `https://www.nist.gov/news-events/news/2019/02/methods-format-preserving-encryption-nist-requests-public-comments-draft`.

[Nugier et al., 2021] Nugier, C., Leblanc-Albarel, D., Blaise, A., Masson, S., Huynh, P., and Piugie, Y. B. W. (2021). An upcycling tokenization method for credit card numbers. In *SECRYPT 2021-18th International Conference on Security and Cryptography*.

[Pal and Mitra, 1992] Pal, S. K. and Mitra, S. (1992). Multilayer perceptron, fuzzy sets, classifiaction.

[Papaioannou et al., 2023] Papaioannou, M., Pelekoudas-Oikonomou, F., Mantas, G., Serrelis, E., Rodriguez, J., and Fengou, M.-A. (2023). A survey on quantitative risk estimation approaches for secure and usable user authentication on smartphones. *Sensors*, 23(6):2979.

[Papamichail et al., 2019] Papamichail, M. D., Chatzidimitriou, K. C., Karanikiotis, T., Oikonomou, N.-C. I., Symeonidis, A. L., and Saripalle, S. K. (2019). Brainrun: A behavioral biometrics dataset towards continuous implicit authentication. *Data*, 4(2):60.

[Parkinson et al., 2021] Parkinson, S., Khan, S., Crampton, A., Xu, Q., Xie, W., Liu, N., and Dakin, K. (2021). Password policy characteristics and keystroke biometric authentication. *IET Biometrics*, 10(2):163–178.

[Parthasarathy et al., 2020] Parthasarathy, D., Bäckstrom, K., Henriksson, J., and Einarsdóttir, S. (2020). Controlled time series generation for automotive software-in-the-loop testing using gans. In *2020 IEEE International Conference On Artificial Intelligence Testing (AITest)*, pages 39–46. IEEE.

[Pascual et al., 2020] Pascual, D., Amirshahi, A., Aminifar, A., Atienza, D., Ryvlin, P., and Wattenhofer, R. (2020). Epilepsygan: Synthetic epileptic brain activities with privacy preservation. *IEEE Transactions on Biomedical Engineering*, 68(8):2435–2446.

[Patel et al., 2016] Patel, V. M., Chellappa, R., Chandra, D., and Barbello, B. (2016). Continuous user authentication on mobile devices: Recent progress and remaining challenges. *IEEE Signal Processing Magazine*, 33(4):49–61.

[Payment Card Industry, 2015] Payment Card Industry (2015). Tokenization Product Security Guidelines – Irreversible and Reversible Tokens. `https://www.pcisecuritystandards.org/documents/Tokenization_Product_Security_Guidelines.pdf`.

[Payment Card Industry, 2020] Payment Card Industry (2020). Payment Card Industry Security Standards. `https://www.pcisecuritystandards.org/pci_security/maintaining_payment_security`.

[Pérez-Cabo et al., 2019] Pérez-Cabo, D., Jiménez-Cabello, D., Costa-Pazo, A., and López-Sastre, R. J. (2019). Deep anomaly detection for generalized face anti-spoofing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 0–0.

[Pisani et al., 2019] Pisani, P. H., Mhenni, A., Giot, R., Cherrier, E., Poh, N., Ferreira de Carvalho, A. C. P. d. L., Rosenberger, C., and Amara, N. E. B. (2019). Adaptive biometric systems: Review and perspectives. *ACM Computing Surveys (CSUR)*, 52(5):1–38.

[Piugie et al., 2022] Piugie, Y. B. W., Di Manno, J., Rosenberger, C., and Charrier, C. (2022). Keystroke dynamics based user authentication using deep learning neural networks. In *2022 International Conference on Cyberworlds (CW)*, pages 220–227. IEEE.

[Piugie et al., 2021] Piugie, Y. B. W., Manno, J., Rosenberger, C., and Charrier, C. (2021). How artificial intelligence can be used for behavioral identification? In *2021 International Conference on Cyberworlds (CW)*.

[Piugie et al., 2019] Piugie, Y. B. W., Tchiotsop, D., Telem, A. N. K., and Ngouonkadi, E. B. M. (2019). Denoising of electroencephalographic signals by canonical correlation analysis and by second-order blind source separation. In *2019 IEEE AFRICON*, pages 1–8. IEEE.

[Plamondon and Srihari, 2000] Plamondon, R. and Srihari, S. N. (2000). Online and off-line handwriting recognition: a comprehensive survey. *IEEE Transactions on pattern analysis and machine intelligence*, 22(1):63–84.

[Qi et al., 2021] Qi, Y., Jia, W., and Gao, S. (2021). Emotion recognition based on piezoelectric keystroke dynamics and machine learning. In *2021 IEEE International Conference on Flexible and Printable Sensors and Systems (FLEPS)*, pages 1–4. IEEE.

[Qu et al., 2020] Qu, F., Liu, J., Ma, Y., Zang, D., and Fu, M. (2020). A novel wind turbine data imputation method with multiple optimizations based on gans. *Mechanical Systems and Signal Processing*, 139:106610.

[Radford et al., 2015] Radford, A., Metz, L., and Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

[Ramponi et al., 2018] Ramponi, G., Protopapas, P., Brambilla, M., and Janssen, R. (2018). T-cgan: Conditional generative adversarial network for data augmentation in noisy time series with irregular sampling. *arXiv preprint arXiv:1811.08295*.

[Rasekh et al., 2014] Rasekh, A., Chen, C.-A., and Lu, Y. (2014). Human activity recognition using smartphone. *arXiv preprint arXiv:1401.8212*.

[Ravanelli and Bengio, 2018] Ravanelli, M. and Bengio, Y. (2018). Speaker recognition from raw waveform with sincnet. In *2018 IEEE Spoken Language Technology Workshop (SLT)*, pages 1021–1028. IEEE.

[Rayani and Changder, 2023] Rayani, P. K. and Changder, S. (2023). Continuous user authentication on smartphone via behavioral biometrics: a survey. *Multimedia Tools and Applications*, 82(2):1633–1667.

[Revett et al., 2007] Revett, K., Gorunescu, F., Gorunescu, M., Ene, M., Magalhaes, S., and Santos, H. (2007). A machine learning approach to keystroke dynamics based

user authentication. *International Journal of Electronic Security and Digital Forensics*, 1(1):55–70.

[Ricci et al., 2019] Ricci, S., Ferreira, E., Menasche, D. S., Ziviani, A., Souza, J. E., and Vieira, A. B. (2019). Learning Blockchain Delays: A Queueing Theory Approach. *SIGMETRICS Perform. Eval. Rev.*, 46(3):122–125.

[Rosenberger, 2020] Rosenberger, C. (2020). Les identités numériques et l'authentification.

[Rosenblatt et al., 2014] Rosenblatt, M., Figliola, A., Paccosi, G., Serrano, E., and Rosso, O. A. (2014). A quantitative analysis of an eeg epileptic record based on multiresolutionwavelet coefficients. *Entropy*, 16(11):5976–6005.

[Roth et al., 2014] Roth, J., Liu, X., Ross, A., and Metaxas, D. (2014). Investigating the discriminative power of keystroke sound. *IEEE Transactions on Information Forensics and Security*, 10(2):333–345.

[Russakovsky et al., 2015] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252.

[Russell, 2023] Russell, S. (2023). Bypassing multi-factor authentication. *ITNOW*, 65(1):42–45.

[Rybnik et al., 2008] Rybnik, M., Tabedzki, M., and Saeed, K. (2008). A keystroke dynamics based system for user identification. In *2008 7th computer information systems and industrial management applications*, pages 225–230. IEEE.

[Saeed, 2016] Saeed, K. (2016). *New directions in behavioral biometrics*. CRC Press.

[Safieddine et al., 2012] Safieddine, D., Kachenoura, A., Albera, L., Birot, G., Karfoul, A., Pasnicu, A., Biraben, A., Wendling, F., Senhadji, L., and Merlet, I. (2012). Removal of muscle artifact from eeg data: comparison between stochastic (ica and cca) and deterministic (emd and wavelet-based) approaches. *EURASIP Journal on Advances in Signal Processing*, 2012(1):127.

[Samura and Nishimura, 2009] Samura, T. and Nishimura, H. (2009). Keystroke timing analysis for individual identification in japanese free text typing. In *2009 ICCAS-SICE*, pages 3166–3170. IEEE.

[Sanchez Guinea et al., 2022] Sanchez Guinea, A., Sarabchian, M., and Mühlhäuser, M. (2022). Improving wearable-based activity recognition using image representations. *Sensors*, 22(5):1840.

[Sarkar et al., 2022] Sarkar, A., Hossain, S., and Sarkar, R. (2022). Human activity recognition from sensor data using spatial attention-aided cnn with genetic algorithm. *Neural Computing and Applications*, pages 1–27.

[Sarkar et al., 2023] Sarkar, A., Hossain, S. S., and Sarkar, R. (2023). Human activity recognition from sensor data using spatial attention-aided cnn with genetic algorithm. *Neural Computing and Applications*, 35(7):5165–5191.

[Schuckers et al., 2019] Schuckers, S., Cannon, G., Tabassi, E., Karlsson, M., and Newton, E. (2019). Fido biometrics requirements. *Population*, 5(2-1):2–3.

[Schuckers et al., 2023] Schuckers, S., Cannon, G., Tekampe, N., Tabassi, E., Karlsson, M., and Newton, E. (2023). Fido biometrics requirements. *Population*, 5(2-1):2–3.

[Serrà et al., 2018] Serrà, J., Pascual, S., and Karatzoglou, A. (2018). Towards a universal neural network encoder for time series. In *CCIA*, pages 120–129.

[Sethi et al., 2023] Sethi, M., Kumar, M., and Jindal, M. (2023). Gender prediction system through behavioral biometric handwriting: a comprehensive review. *Soft Computing*, pages 1–21.

[Shadman et al., 2023] Shadman, R., Wahab, A. A., Manno, M., Lukaszewski, M., Hou, D., and Hussain, F. (2023). Keystroke dynamics: Concepts, techniques, and applications. *arXiv preprint arXiv:2303.04605*.

[Shahmohammadi et al., 2017] Shahmohammadi, F., Hosseini, A., King, C. E., and Sarrafzadeh, M. (2017). Smartwatch based activity recognition using active learning. In *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pages 321–329. IEEE.

[Sharma and Elmiligi, 2022] Sharma, M. and Elmiligi, H. (2022). Behavioral biometrics: Past, present and future. *Recent Advances in Biometrics*, page 69.

[Shen et al., 2010] Shen, C., Cai, Z., Guan, X., and Cai, J. (2010). A hypo-optimum feature selection strategy for mouse dynamics in continuous identity authentication and monitoring. In *2010 IEEE International Conference on Information Theory and Information Security*, pages 349–353. IEEE.

[Shi et al., 2023] Shi, Y., Wang, X., Zheng, K., and Cao, S. (2023). User authentication method based on keystroke dynamics and mouse dynamics using hda. *Multimedia Systems*, 29(2):653–668.

[Simonetto, 2018] Simonetto, L. (2018). Generating spiking time series with generative adversarial networks: an application on banking transactions. *MS thesis-Univ. of Amsterdam*.

[Sitová et al., 2015] Sitová, Z., Šeděnka, J., Yang, Q., Peng, G., Zhou, G., Gasti, P., and Balagani, K. S. (2015). Hmog: New behavioral biometric features for continuous authentication of smartphone users. *IEEE Transactions on Information Forensics and Security*, 11(5):877–892.

[Skansi, 2018] Skansi, S. (2018). *Introduction to Deep Learning: from logical calculus to artificial intelligence*. Springer.

[Smith, 1998] Smith, J. (1998). *The Book*. The publishing company, London, 2nd edition.

[Smith and Smith, 2020] Smith, K. E. and Smith, A. O. (2020). Conditional gan for timeseries generation. *arXiv preprint arXiv:2006.16477*.

[Spillane, 1975] Spillane, R. (1975). Keyboard apparatus for personal identification. *IBM Technical Disclosure Bulletin*, 17:3346.

[Srivastava et al., 2014] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958.

[Srivastava et al., 2015] Srivastava, N., Mansimov, E., and Salakhudinov, R. (2015). Unsupervised learning of video representations using lstms. In *International conference on machine learning*, pages 843–852. PMLR.

[Stiller et al.,] Stiller, B., Bocek, T., Hecht, F., Machado, G., Racz, P., and Waldburger, M. Protect users without frustrating them using ai-driven behavioral biometrics. Technical report, Technical report, 01 2010.

[Stragapede et al., 2023] Stragapede, G., Vera-Rodriguez, R., Tolosana, R., and Morales, A. (2023). Behavepassdb: Public database for mobile behavioral biometrics and benchmark evaluation. *Pattern Recognition*, 134:109089.

[Subasi et al., 2020] Subasi, A., Khateeb, K., Brahimi, T., and Sarirete, A. (2020). Human activity recognition using machine learning methods in a smart healthcare environment. In *Innovation in Health Informatics*, pages 123–144. Elsevier.

[Sun et al., 2015] Sun, C., Junejo, I. N., Tappen, M., and Foroosh, H. (2015). Exploring sparseness and self-similarity for action recognition. *IEEE Transactions on Image Processing*, 24(8):2488–2501.

[Sun et al., 2020] Sun, H., Deng, Z., Chen, H., and Parkes, D. C. (2020). Decision-aware conditional gans for time series data. *arXiv preprint arXiv:2009.12682*.

[Sun et al., 2018] Sun, Y., Peng, L., Li, H., and Sun, M. (2018). Exploration on spatiotemporal data repairing of parking lots based on recurrent gans. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 467–472. IEEE.

[Sussman, 2020] Sussman, B. (2020). 'BIGBADABOOM!' Carding Forum Selling Millions of Records from Wawa Stores Data Breach. `https://www.secureworldexpo.com/industry-news/carding-forum-wawa-data-breach-update`.

[Sweeney et al., 2012] Sweeney, K. T., Ayaz, H., Ward, T. E., Izzetoglu, M., McLoone, S. F., and Onaral, B. (2012). A methodology for validating artifact removal techniques for physiological signals. *IEEE transactions on information technology in biomedicine*, 16(5):918–926.

[Sweeney et al., 2013] Sweeney, K. T., McLoone, S. F., and Ward, T. E. (2013). The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique. *IEEE transactions on biomedical engineering*, 60(1):97–105.

[Syed Idrus et al., 2013] Syed Idrus, S. Z., Cherrier, E., Rosenberger, C., and Bours, P. (2013). Soft biometrics database: A benchmark for keystroke dynamics biometric systems. In *2013 International Conference of the BIOSIG Special Interest Group (BIOSIG)*, pages 1–8.

[Szegedy et al., 2015] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9.

[Tang et al., 2020] Tang, Y., Teng, Q., Zhang, L., Min, F., and He, J. (2020). Layer-wise training convolutional neural networks with smaller filters for human activity recognition using wearable sensors. *IEEE Sensors Journal*, 21(1):581–592.

[Taniguchi and Masuda, 2017] Taniguchi, T. and Masuda, T. (2017). Linear demixed domain multichannel nonnegative matrix factorization for speech enhancement. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 476–480. IEEE.

[Tanisaro and Heidemann, 2016] Tanisaro, P. and Heidemann, G. (2016). Time series classification using time warping invariant echo state networks. In *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 831–836. IEEE.

[Tarkoma et al., 2012] Tarkoma, S., Rothenberg, C. E., and Lagerspetz, E. (2012). Theory and Practice of Bloom Filters for Distributed Systems. *IEEE Communications Surveys Tutorials*, 14(1):131–155.

[Thales group, 2018] Thales group (2018). Card Data from 5M Customers Stolen in Data Breach at Saks Fifth Avenue, Lord & Taylor. `https://dis-blog.thalesgroup.com/security/2018/04/03/saksfifthavenuedatabreach/`.

[Thang et al., 2012] Thang, H. M., Viet, V. Q., Thuc, N. D., and Choi, D. (2012). Gait identification using accelerometer on mobile phone. In *2012 International Conference on Control, Automation and Information Sciences (ICCAIS)*, pages 344–348. IEEE.

[Theofanos et al., 2008] Theofanos, M. F., Stanton, B. C., and Wolfson, C. (2008). Usability and biometrics: Ensuring successful biometric systems.

[Thiruganam et al., 2010] Thiruganam, M., Anouncia, S. M., and Kantipudi, S. (2010). Automatic defect detection and counting in radiographic weldment images. *International Journal of Computer Applications*, 10(2):1–5.

[Tolosana et al., 2019] Tolosana, R., Gomez-Barrero, M., Busch, C., and Ortega-Garcia, J. (2019). Biometric presentation attack detection: Beyond the visible spectrum. *IEEE Transactions on Information Forensics and Security*, 15:1261–1275.

[Toosi and Akhaee, 2021] Toosi, R. and Akhaee, M. A. (2021). Time–frequency analysis of keystroke dynamics for user authentication. *Future Generation Computer Systems*, 115:438–447.

[Trojahn and Ortmeier, 2013] Trojahn, M. and Ortmeier, F. (2013). Toward mobile authentication with keystroke dynamics on mobile phones and tablets. In *2013 27th International Conference on Advanced Information Networking and Applications Workshops*, pages 697–702. IEEE.

[Tufek et al., 2019] Tufek, N., Yalcin, M., Altintas, M., Kalaoglu, F., Li, Y., and Bahadir, S. K. (2019). Human action recognition using deep learning methods on limited sensory data. *IEEE Sensors Journal*, 20(6):3101–3112.

[Turnip, 2015] Turnip, A. (2015). Comparison of ica-based jade and sobi methods eog artifacts removal. *Journal of Medical and Bioengineering*, 4(6).

[U.S. Department of Commerce, 2020] U.S. Department of Commerce (2020). Quarterly Retail E-commerce Sales. `https://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf`.

[Van der Maaten and Hinton, 2008] Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(11).

[Vasan et al., 2020] Vasan, D., Alazab, M., Wassan, S., Naeem, H., Safaei, B., and Zheng, Q. (2020). Imcfn: Image-based malware classification using fine-tuned convolutional neural network architecture. *Computer Networks*, 171:107138.

[Vasconcelos et al., 2021] Vasconcelos, C., Larochelle, H., Dumoulin, V., Romijnders, R., Roux, N. L., and Goroshin, R. (2021). Impact of aliasing on generalization in deep convolutional networks. *arXiv preprint arXiv:2108.03489*.

[Venkat and De Wilde, 2011] Venkat, I. and De Wilde, P. (2011). Robust gait recognition by learning and exploiting sub-gait characteristics. *International Journal of Computer Vision*, 91:7–23.

[Visvizi et al., 2020] Visvizi, A., Jussila, J., Lytras, M. D., and Ijäs, M. (2020). Tweeting and mining oecd-related microcontent in the post-truth era: a cloud-based app. *Computers in Human Behavior*, 107:105958.

[Voltage Security, 2012] Voltage Security (2012). Voltage secure stateless tokenization. `https://www.voltage.com/wp-content/uploads/Voltage_White_Paper_SecureData_SST_Data_Protection_and_PCI_Scope_Reduction_for_Todays_Businesses.pdf`.

[Wahab et al., 2021] Wahab, A. A., Hou, D., Schuckers, S., and Barbir, A. (2021). Utilizing keystroke dynamics as additional security measure to protect account recovery mechanism. In *ICISSP*, pages 33–42.

[Wandji Piugie et al., 2023] Wandji Piugie, Y. B., Charrier, C., Di Manno, J., and Rosenberger, C. (2023). Deep features fusion for user authentication based on human activity. *IET Biometrics*, 12(4):222–234.

[Wang et al., 2023] Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., et al. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972):47–60.

[Wang and Geng, 2009] Wang, L. and Geng, X. (2009). *Behavioral Biometrics for Human Identification: Intelligent Applications: Intelligent Applications*. IGI Global.

[Wang et al., 2019] Wang, L., Zhang, W., and He, X. (2019). Continuous patient-centric sequence generation via sequentially coupled adversarial learning. In *Database Systems for Advanced Applications: 24th International Conference, DASFAA 2019, Chiang Mai, Thailand, April 22–25, 2019, Proceedings, Part II 24*, pages 36–52. Springer.

[Wang et al., 2017] Wang, Z., Yan, W., and Oates, T. (2017). Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*, pages 1578–1585. IEEE.

[Whatman, 2020] Whatman, P. (2020). Credit card statistics 2020: 65+ facts for Europe, UK, and US. `https://blog.spendesk.com/en/credit-card-statistics-2020`.

[Wiese et al., 2020] Wiese, M., Knobloch, R., Korn, R., and Kretschmer, P. (2020). Quant gans: deep generation of financial time series. *Quantitative Finance*, 20(9):1419–1440.

[Williams and Zipser, 1989] Williams, R. J. and Zipser, D. (1989). A learning algorithm for continually running fully recurrent neural networks. *Neural computation*, 1(2):270–280.

[Wu et al., 2018] Wu, G., Wang, J., Zhang, Y., and Jiang, S. (2018). A continuous identity authentication scheme based on physiological and behavioral characteristics. *Sensors*, 18(1):179.

[Xia et al., 2020] Xia, K., Huang, J., and Wang, H. (2020). Lstm-cnn architecture for human activity recognition. *IEEE Access*, 8:56855–56866.

[Yampolskiy and Govindaraju, 2010a] Yampolskiy, R. V. and Govindaraju, V. (2010a). Game playing tactic as a behavioral biometric for human identification. In *Behavioral Biometrics for Human Identification: Intelligent Applications*, pages 385–413. IGI Global.

[Yampolskiy and Govindaraju, 2010b] Yampolskiy, R. V. and Govindaraju, V. (2010b). Taxonomy of behavioural biometrics. In *Behavioral Biometrics for Human Identification: Intelligent Applications*, pages 1–43. IGI Global.

[Yang et al., 2020] Yang, Y., Wang, J., Gao, Z., Huo, Y., and Qiu, X. (2020). Sri-xdfm: a service reliability inference method based on deep neural network. *INTELLIGENT AUTOMATION AND SOFT COMPUTING*, 26(6):1459–1475.

[Yeh et al., 2018] Yeh, K.-H., Su, C., Chiu, W., and Zhou, L. (2018). I walk, therefore i am: continuous user authentication with plantar biometrics. *IEEE Communications Magazine*, 56(2):150–157.

[Yi and Mak, 2019] Yi, L. and Mak, M.-W. (2019). Adversarial data augmentation network for speech emotion recognition. In *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 529–534. IEEE.

[Yingzhen and Mandt, 2018] Yingzhen, L. and Mandt, S. (2018). Disentangled sequential autoencoder. In *International Conference on Machine Learning*, pages 5670–5679. PMLR.

[Yohan et al., 2018] Yohan, M., Hanry, H., and Dion, D. (2018). Keystroke dynamic classification using machine learning forpassword authorization. In *3rd International Conference on Computer Science and Computational Intelligence, Procedia Computer Science*.

[Yoon et al., 2019] Yoon, J., Jarrett, D., and Van der Schaar, M. (2019). Time-series generative adversarial networks. *Advances in neural information processing systems*, 32.

[Zaky and Saxe, 2022] Zaky, K. and Saxe, D. H. (2022). Multi-factor authentication. *ID-Pro Body of Knowledge*, 1(10).

[Zareen and Jabin, 2016] Zareen, F. J. and Jabin, S. (2016). Authentic mobile-biometric signature verification system. *IET Biometrics*, 5(1):13–19.

[Zebin et al., 2019] Zebin, T., Scully, P. J., Peek, N., Casson, A. J., and Ozanyan, K. B. (2019). Design and implementation of a convolutional neural network on an edge computing smartphone for human activity recognition. *IEEE Access*, 7:133509–133520.

[Zeiler, 2012] Zeiler, M. D. (2012). Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*.

[Zhang et al., 2018] Zhang, C., Kuppannagari, S. R., Kannan, R., and Prasanna, V. K. (2018). Generative adversarial network for synthetic time series data generation in smart grids. In *2018 IEEE international conference on communications, control, and computing technologies for smart grids (SmartGridComm)*, pages 1–6. IEEE.

[Zhang et al., 2020] Zhang, H., Xiao, N., Liu, P., Wang, Z., and Tang, R. (2020). G-rnn-gan for singing voice separation. In *Proceedings of the 2020 5th International Conference on Multimedia Systems and Signal Processing*, pages 69–73.

[Zhang, 2019] Zhang, M. (2019). Gait activity authentication using lstm neural networks with smartphone sensors. In *2019 15th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN)*, pages 456–461. IEEE.

[Zhang et al., 2016] Zhang, Y., Gan, Z., and Carin, L. (2016). Generating text via adversarial training. In *NIPS workshop on Adversarial Training*, volume 21, pages 21–32. academia. edu.

[Zhao et al., 2017] Zhao, B., Lu, H., Chen, S., Liu, J., and Wu, D. (2017). Convolutional neural networks for time series classification. *Journal of Systems Engineering and Electronics*, 28(1):162–169.

[Zhao et al., 2018] Zhao, J., Kim, Y., Zhang, K., Rush, A., and LeCun, Y. (2018). Adversarially regularized autoencoders. In *International conference on machine learning*, pages 5902–5911. PMLR.

[Zhao et al., 2020] Zhao, X., Chen, S., Zhou, L., and Chen, Y. (2020). Sound source localization based on srp-phat spatial spectrum and deep neural network. *CMC-COMPUTERS MATERIALS & CONTINUA*, 64(1):253–271.

[Zheng et al., 2014] Zheng, N., Bai, K., Huang, H., and Wang, H. (2014). You are how you touch: User verification on smartphones via tapping behaviors. In *2014 IEEE 22nd International Conference on Network Protocols*, pages 221–232. IEEE.

[Zheng et al., 2011] Zheng, N., Paloski, A., and Wang, H. (2011). An efficient user verification system via mouse movements. In *Proceedings of the 18th ACM conference on Computer and communications security*, pages 139–150.

[Zheng et al., 2018] Zheng, P., Zheng, Z., Luo, X., Chen, X., and Liu, X. (2018). A Detailed and Real-Time Performance Monitoring Framework for Blockchain Systems. In *Proceedings of the 40th International Conference on Software Engineering: Software Engineering in Practice*, page 134–143.

[Zhong and Deng, 2014] Zhong, Y. and Deng, Y. (2014). Sensor orientation invariant mobile gait biometrics. In *IEEE international joint conference on biometrics*, pages 1–8. IEEE.

[Zhong and Deng, 2015] Zhong, Y. and Deng, Y. (2015). A survey on keystroke dynamics biometrics: approaches, advances, and evaluations. In *Recent Advances in User Authentication Using Keystroke Dynamics Biometrics*, number 1, pages 1–22. Science Gate Publishing.

[Zhong et al., 2015] Zhong, Y., Deng, Y., and Meltzner, G. (2015). Pace independent mobile gait biometrics. In *2015 IEEE 7th international conference on biometrics theory, applications and systems (BTAS)*, pages 1–8. IEEE.

[Zhou et al., 2016] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A. (2016). Learning deep features for discriminative localization. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2921–2929.

[Zhu et al., 2019a] Zhu, F., Ye, F., Fu, Y., Liu, Q., and Shen, B. (2019a). Electrocardiogram generation with a bidirectional lstm-cnn generative adversarial network. *Scientific reports*, 9(1):6734.

[Zhu et al., 2019b] Zhu, G., Zhao, H., Liu, H., and Sun, H. (2019b). A novel lstm-gan algorithm for time series anomaly detection. In *2019 prognostics and system health management conference (PHM-Qingdao)*, pages 1–6. IEEE.

Performance and Security Evaluation of Behavioral Biometric Systems

Behavioral biometrics enhances IT security and improves user experience by analyzing interactions. This Ph.D. thesis proposes a generic method based on the analysis of behavioral time series. It explores the use of traditional machine learning and deep learning techniques for user identification and authentication based on these behaviors.

In addition, the research carried out examines the vulnerability of behavioral biometric systems to presentation attacks. To this end, we use TimeGAN to generate synthetic behavioral biometric data capable of fooling authentication systems. These synthetic data preserve temporal characteristics, making it difficult to distinguish them from authentic data. The results highlight TimeGAN's ability to generate behavioral patterns that could be used to test authentication systems, raising questions about the robustness of such systems against malicious attacks.

Évaluation des performances et de la sécurité des systèmes biométriques comportementales

La biométrie comportementale est une approche prometteuse pour renforcer la sécurité des systèmes informatiques tout en améliorant l'expérience utilisateur grâce à l'analyse des interactions des utilisateurs. Cette thèse de doctorat propose une méthode générique basée sur l'analyse de séries temporelles comportementales. Elle explore l'utilisation de techniques d'apprentissage machine traditionnelles et d'apprentissage profond pour l'authentification des utilisateurs basée sur ces comportements.

En outre, nous examinons la vulnérabilité des systèmes biométriques comportementales aux attaques par présentation. Nous utilisons le TimeGAN pour générer des données biométriques synthétiques préservant les caractéristiques temporelles, rendant difficile leur distinction des données authentiques. Les résultats obtenus soulignent la capacité du TimeGAN à générer des modèles comportementaux pour tester les systèmes , remettant en question la robustesse de ces systèmes face aux attaques malveillantes.

Keywords: Behavioral Biometrics; Cybersecurity; Identification; Authentication; Time series to Image; Synthetic Behavioral Biometrics; Presentation Attack Instrument.