Pressure-Based Gait Recognition: Feature Extraction Techniques for Footwear-Agnostic Identity Recognition

by

Ala Salehi

Bachelor of Science in Electrical Engineering, University of Science and Technology, 2020

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science in Engineering

In the Graduate Academic Unit of Electrical and Computer Engineering

Supervisor:	Erik Scheme, PhD, Electrical and Computer Engineering
Examining Board:	Jon Sensinger, PhD, Electrical and Computer Engineering
	Saqib Hakak, PhD, Computer Science

This thesis is accepted by the

Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

December, 2024

Abstract

This research explores the development of a robust pressure-based gait recognition system, with a focus on reducing the impact of changes in footwear. Using two datasets; CASIA-D and a newly collected UNB dataset, we compare traditional and deep learning methods, including two novel architectures: UMAPNet for spatial feature learning and FootPart, a comprehensive spatiotemporal model. FootPart integrates local spatial partitioning with dynamic temporal modelling, achieving significant improvements in both closed-set and open-set verification tasks. Results show that FootPart maintains high accuracy under variable conditions, outperforming baseline models in identification tasks and demonstrating resilience to unseen footwear. This work underscores the importance of detailed spatial and temporal features in robust gait recognition, with implications for security, healthcare, and smart environments.

Dedication

To my mother, whose unwavering faith in me has been my greatest source of strength; to my father, whose resilience is my constant inspiration; and to my siblings, whose encouragement lifts me in the hardest moments.

Though distance separates us, your love has bridged every mile, and your belief has accompanied me every step of the way. This journey would not have been possible without each of you in my heart.

Acknowledgements

I would like to acknowledge the following individuals, each of whom has been instrumental to the completion of this dissertation and to my growth throughout this journey:

My supervisor, Dr. Erik Scheme. Your guidance and support have been invaluable at every stage of this research. Thank you for your mentorship and for believing in me when I doubted myself. Your encouragement and insights have left a lasting impact.

My co-authors and collaborators. Research is a collaborative effort, and I am grateful for the opportunity to work alongside such talented and inspiring individuals. Special thanks to Dr. Angkoon Phinyomark and Dr. Aaron Tabor, whose thoughtful feedback throughout the thesis process greatly strengthened my work. Your expertise and dedication have been invaluable.

My colleagues at the Institute of Biomedical Engineering. Thank you for creating a friendly and dynamic environment where we could support each other's growth, both professionally and personally.

Finally, *the funding and project partners who made this project possible*, including CyberNB, Knowledge Park, Stepscan Technologies, the New Brunswick Innovation Foundation, the Atlantic Canada Opportunities Agency, and the Natural Sciences and Engineering Research Council of Canada (NSERC) Alliance grants program.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The term gait refers to the complex, coordinated manner in which people walk, characterized by the cyclic movement pattern of the limbs and body. The science of capturing and analyzing these patterns is called gait analysis. In healthcare, gait can be compared against an established population-level "healthy norm" to test for abnormalities or monitor for changes that may serve as an early marker of mobility and cognitive issues. Alternatively, characteristics of the gait signal (i.e., called *features* in this thesis) can be used to uniquely identify individuals from a larger group. Biometrics (i.e., a physiological and/or behavioral trait that uniquely identifies someone) can be applied in a wide variety of applications where secure access is essential, but gait may offer benefits when others forms of biometric authentication are infeasible or inappropriate, such as when physical contact (e.g., fingerprints) is undesirable due to hygiene concerns, or where facial recognition may be compromised by environmental conditions (e.g., low lighting or when individuals are wearing masks) in locations like hospitals, airports, or outdoor settings.

Popular authentication methods like passwords, swipe cards, and signatures are

relatively easy to guess, steal, or imitate, whereas physical biometrics like iris, fingerprint, palm, and facial recognition generally provide more robust security [4]. These approaches can be more invasive, however, requiring active user cooperation (e.g. scanning with a sensor), and are confounded by accessories like glasses or masks, privacy concerns, and potential biases. These inherent limitations motivate the design of new non-invasive biometric systems that are easy and robust. Fortunately, gait has emerged as a promising biometric modality that may be less invasive and can be captured as users approach an access point.

Vision-based gait recognition, which involves identifying individuals based on their walking patterns as captured through camera systems, has dominated the research landscape in recent years because of the ease of use and prevalence of camera-based capture systems. While convenient, vision methods have several known limitations - poor performance in low visibility conditions, occlusion issues when tracking individuals, and privacy/bias concerns akin to facial recognition. As an alternative, pressure-based gait recognition emerged, based on specialized flooring that captures the distribution of forces at the feet during gait, providing both spatial (e.g., foot shape) and temporal (e.g., cadence) information [38]. Pressure signals can be acquired even in complete darkness and are not susceptible to occlusion when multiple subjects are in view [95]. Research has shown that pressure-based gait recognition can even outperform other biometric modalities [15]. However, several challenges still hinder its real-world application. Factors such as load carrying, changes in walking speed, and variations in footwear have been shown to degrade recognition performance, as these conditions affect the measured gait signals. Additionally, pressure-based systems require sensor installation in the environment, which may not be desirable in all situations. Despite these challenges, pressure-based gait recognition holds great potential, warranting further research.

Of the challenges faced by pressure-based gait recognition systems, *footwear* is

particularly problematic because it influences *both* the physiological and behavioural properties of the underlying signal. For example, footwear-agnostic identity recognition is challenging because the shoe's soles (i.e., the interface through which the system captures information) come in various shapes, sizes, and patterns. To be robust, recognition systems must be designed to infer identity in a way that is resilient to these physical differences. Footwear is also known to influence *behavioural* gait patterns (e.g., individuals walk differently when wearing loose flip-flops vs heavy boots [28, 101]). This suggests that footwear likely *modifies* the identifying properties of the underlying signal, making it difficult to infer generalizations across shoe types. To explore this footwear challenge, this thesis explores two variants of footwear-agnostic gait recognition: a) *closed-set* or within-domain, and b) *open-set* or cross-domain recognition. Closed-set recognition assumes that the system is only responsible for recognizing participant-footwear pairings that were encountered during training. This simplification represents the case when user may be required to train the system with all footwear that they may use, such as in a private school or military installation where users must wear common uniforms or standard-issue clothing. It may also be possible to extend the applicable use cases for closed-set recognition by requiring a short supervised registration sequence when a new set of footwear is encountered. However, open-set recognition case makes no assumptions about encountered footwear - the system is expected to perform in all cases, irrespective of whether or not the presented shoe was encountered during system training (i.e., the system is expected to recognize an open set of footwear categories). This problem variant is much more difficult but is more desirable in many cases, such as office buildings or border security, where a large number of intermittent users would be less motivated to retrain the system with every pair of shoes they own.

In both scenarios, *feature extraction* plays a critical role in extracting relevant information from the pressure data samples to enable robust, efficient, and resilient

gait recognition. High-resolution pressure sensor data can be represented as a three-dimensional (3D) array, incorporating height and width (dictated by sensor size and resolution) over time. Relevant features from this data can be spatial (pressure images), temporal (pressure time series), or spatiotemporal (combining spatial and temporal aspects). Extracting discriminative features is crucial for accurate gait recognition, but traditional machine learning (ML) approaches often struggle with the high dimensionality and variability of gait data. Deep learning (DL) methods, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), may better capture complex patterns but typically require large datasets. In response to this limitation, Feature Learning Networks (FLN) offer a more structured approach, integrating conventional feature extraction within DL frameworks.

Because footwear affects both *spatial* (e.g., sole shape and design) and *temporal* (e.g., cadence) aspects of gait [19], this thesis explores these various representations of the gait signal in efforts to overcome the footwear challenge. First, known *spatial* and *temporal* features are evaluated with current ML models, investigating the impact of footwear and the potential role of feature extraction and dimensionality reduction. FLNs are then explored given the limited amount of data available for each user, and a deep *temporal* learning approach is explored using Long Short Term Memory (LSTM) to more explicitly model the temporal dynamics of gait. Building on the unique benefits afforded by spatial and temporal features, fusions of both information representations are explored through *spatiotemporal* features, evaluating new models inspired by similar challenges faced in the larger vision-based gait recognition field, and translating them to the pressure-based scenario. Finally, a novel multi-modal FootPart LSTM network is introduced, which partitions the pressure footstep into regions and computes Dynamic Time Warping (DTW) based distances between them to improve robustness to footwear variations.

1.2 Objectives

The primary objective of this research is to explore how different feature extraction techniques can improve footwear-agnostic pressure-based gait recognition. More specifically, this research aims to:

- Aim 1: Explore and compare various feature learning and extraction techniques to improve the closed-set user recognition performance for known footwear.
- Aim 2: Explore the generalization of these approaches to the more challenging open-set use case, where models are evaluated using previously unseen footwear.

1.3 Thesis Structure

The thesis is organized as follows: Chapter 2 provides background information and a review of relevant literature on gait recognition. Chapter 3 details the data collection process and describes the datasets used. Chapter 4 focuses on conventional feature extraction and dimensionality reduction approaches to solve the recognition problem, before exploring more advanced FLNs and introducing a novel UMAPNet architecture. Chapter 5 investigates deep temporal and spatiotemporal approaches, adopts and modifies a partitioning approach inspired by vision-based gait recognition, and introduces a novel multi-variate spatiotemporal network with similarity distance metrics. Chapter 6 then explores generalizing these results and models to a brand new, unseen dataset, and the challenging recognition problem. Finally, Chapter 7 discusses the research findings, implications, and potential future directions.

Chapter 2

Background and Related Works

2.1 Gait Biometrics

Gait biometrics refers to the unique identification of individuals based on their walking patterns, encompassing a combination of physiological traits (e.g., body structure, weight distribution) and behavioral characteristics (e.g., cadence, stride length, joint movement). The process of recognizing someone can be performed through one of two main approaches, both of which are commonplace in security-related applications; *verification* or *identification*. Verification, also known as *authentication* in some fields of security and access control, involves the process of confirming an individual's claimed identity by comparing their features against a stored template or database, whereas identification involves assigning a label to an individual (without a claimed identity) [90]. This dissertation explores both of these security-related tasks and uses the term *gait recognition*, or *user recognition* more generally, to describe both *verification* and/or *identification*.

The use of knowledge-based methods of recognition such as passwords and tokens are now commonplace, but these can easily be forgotten, forged, and/or stolen[96]. Biometric authentication systems have, therefore, been widely used as a more secure

alternative (i.e., difficult to forge and/or steal) [47, 36]. Some examples of physiological biometrics include fingerprints, iris patterns [24, 14], and facial features. More behavioural biometrics encompass examples such as voice/speech patterns and keystroke dynamics [1, 97].

Research suggests that both genetic and environmental factors contribute to the development of unique gait patterns, influenced by characteristics such as body structure, weight distribution, neurological pathways, health conditions, and habitual tendencies [102, 26, 64, 82, 7, 83]. This complex spatiotemporal pattern is highly resistant to imitation, making gait appealing for secure biometric systems. These attributes, coupled with constraints faced by other biometric methods, have motivated further exploration of the performance limits of gait recognition technology [106, 93, 33].

2.1.1 Video Based Gait Recognition

The field of gait recognition began predominantly with vision-based approaches that capture and analyze movement patterns from video footage [50]. Vision-based methods gained popularity for many reasons, including the widespread deployment and accessibility of cameras as well as the coordinated creation of large datasets collected to support research and development [76]. These systems typically employ computer vision techniques to extract gait features from the video data and match them against a database of known gait signatures [91]. The typical pipeline for video-based gait recognition begins with capturing video footage of an individual walking, followed by preprocessing steps such as background subtraction and silhouette extraction to isolate the person's movement. The processed video is then analyzed to extract distinctive gait features, which are compared against a database of known gait patterns for identification or verification. Two main feature extraction approaches used in vision-based gait recognition are Gait Energy Images (GEIs) and skeleton-based methods. Each method is briefly described below.

A GEI is a two-dimensional (2D) depiction (i.e., image) of the gait cycle. The image is created by first centering the observed individual in the video feed (e.g., often achieved by locking a landmark such as the hip bones to the center of the video frame), binarizing each video frame to create a silhouette mask (i.e., 0s denote background, 1s denote the observed individual), and then averaging each resulting pixel time series to construct a "flattened" 2-dimensional image [41]. Regions of the resulting GEI image with higher intensity values denote regions that were physically occupied for longer periods of the gait cycle. This compact representation has been effective in various gait recognition tasks (e.g., [104, 100, 61]), benefiting from its simplicity and robustness to noise. GEIs offer a straightforward and efficient way to capture gait patterns, making them suitable for scenarios with limited computational resources. However, the classification of GEIs may struggle due to variations in appearance, viewing angle, and clothing [3]. Figure 2.1 illustrates example GEI images obtained from different viewing angles to demonstrate this effect.

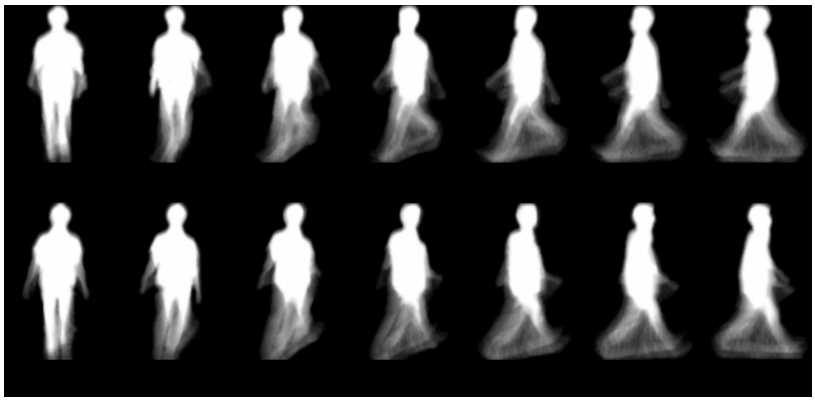

Figure 2.1: An example of a GEI image obtained from different viewing angles [42].

In contrast, skeleton-based methods extract the underlying joint structure of the recorded individual by tracking key joint positions (e.g., hips, knees, shoulders, elbows). The extracted "skeleton graph", as shown in Figure 2.2, can then be further analyzed using two main approaches: a) model-based methods, where established gait kinematics create a simplified representation of the gait dynamics for easier analysis

of limb movements and joint angles [45], and b) data-driven approaches, where the "skeleton graphs" are used to train ML or deep neural network. Recent advancements in DL have improved skeleton-based gait recognition by leveraging CNNs and RNNs to enhance feature extraction and temporal modelling [105]. Skeleton-based methods provide a more detailed and flexible representation of gait, but they require accurate pose estimation, which can be computationally intensive, and still suffer from similar challenges like clothing and viewing angle [38].

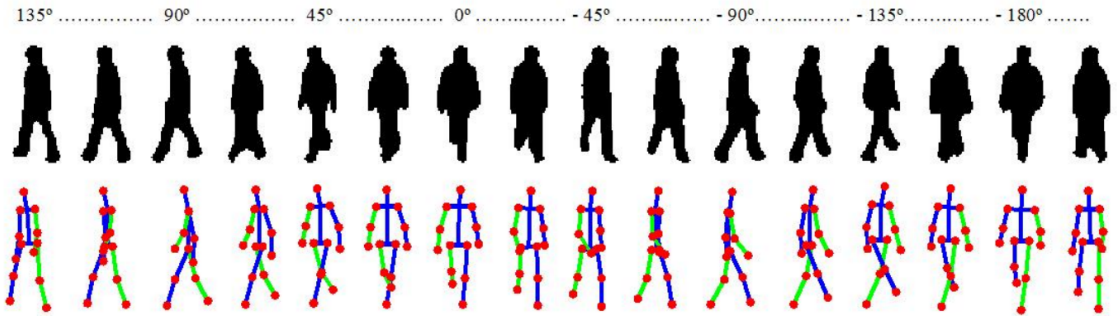

Figure 2.2: An example (up) Silhouettes and (down) Skeletons in different viewing angles [42].

Irrespective of the approach employed, a strength of vision-based gait recognition approaches, in general, is that the information about the full gait cycle is available for analysis (e.g., swing phase, arm swing). This comprehensive view allows for detailed extraction of features from the entire body's motion continuously during gait. In contrast, the pressure-based gait systems explore in this work only directly capture information about foot contact dynamics, such as pressure distribution and timing of steps. This restricts the scope of analysis to lower limb movements and ground contact patterns, posing a challenge in capturing the complete spectrum of gait characteristics observable in vision-based methods. However, this perceived deficit may also help to overcome some of the key challenges of video-based gait recognition.

The very use of video makes the approach sensitive to variations in environmental conditions, viewing angle, visual occlusion of the subject by objects or other people [107], loose clothing or accessories (like scarves and hats), changes in clothing or car-

rying conditions [70], and image resolution [77]. Additionally, camera-based methods do not work at all in some situations, such as in darkness, haze, and smokey environments where the pictures are unclear [34]. This sensitivity to visual variations – particularly changes in clothing (often referred to as poor 'cross-clothing' performance [54]) – poses a challenge similar to an issue faced in pressure-based gait systems, where variations in footwear may significantly impact the recognition performance. In both cases, external factors unrelated to the underlying gait mechanics introduce variability, making it harder to achieve consistent and accurate recognition. Furthermore, the capture of an individual's likeness via video, whether for facial or gait recognition, has raised concerns about privacy and data use [88]. Extensive research has also shown that computer vision systems exhibit racial and gender biases, a broader issue in the field [9]. Although video-based gait recognition offers comprehensive gait analysis and benefits from widespread camera infrastructure, its scalability is often hindered by its sensitivity to environmental variations and the computational demands of processing high-resolution video data, limiting its effectiveness in certain real-world scenarios [103].

2.1.2 Pressure Based Gait Recognition

In contrast to video-based approaches, pressure-based gait recognition employs specialized mats or arrays of pressure sensors embedded in the floor to capture pressure patterns generated by the foot while walking [51]. The signal captured contains both *spatial* information, such as the shape and pressure distribution of the foot, and *temporal* information about how the foot movement progresses throughout the gait cycle with each footstep. Like video-based approaches, the rich spatiotemporal biometric information encoded in this signal can also be leveraged for person identification [88].

Despite its relative infancy, pressure-based gait analysis offers several potential advantages over video-based gait approaches. First, pressure-based approaches are

not as sensitive to signal distortion as video - pressure signals can just as readily be captured in complete darkness. Similarly, even in crowded environments, sensor occlusion is unlikely in pressure-based capture systems (e.g., unless the person steps on top of something on the floor) [40]. Preprocessing is often more straightforward in pressure-based systems because there is no background image (the "background" is simply a no-pressure reading) and the generated patterns do not overlap spatiotemporally (unless users are literally walking over the top of one another). Individual gait patterns can therefore be more easily segmented and distinguished, even in the presence of multiple subjects. However, unlike video, the only information available to a pressure-based gait recognition system is the signal derived from the underfoot contact with the instrumented flooring. This means that some portions of the gait cycle are not directly observable, such as the swing phase for any given foot. Despite this, research has shown that it is possible to perform person identification using only these directly observed portions of the gait cycle [88].

Although early results using pressure-based gait recognition are promising, research is needed to address known challenges. For example, factors such as walking speed and load carrying (e.g., backpack, briefcase) have been shown to degrade recognition performance [73], as can small training sets [48]. One particularly challenging factor, explored in this research, is the impact of different footwear conditions [62]. Footwear presents a significant challenge in pressure-based gait recognition because it not only influences the *spatial* properties of the signal (e.g., the shape and pattern of shoe soles), but it also influences its *temporal* dynamics (e.g., consider walking in loose flip-flops versus heavy boots [28, 101]).

2.1.3 Related Works: Performance of Pressure-Based Gait Recognition Systems

Pressure-based gait recognition has garnered increasing research attention, with performance varying based on feature selection, dataset characteristics, and footwear conditions. This subsection reviews key studies, focusing on their findings and their relevance to the objectives of this dissertation.

2.1.3.1 Barefoot Gait Recognition

Numerous studies have investigated gait recognition using barefoot datasets, which provide a controlled environment for capturing gait patterns. These studies typically report high recognition rates, as the absence of footwear reduces variability in pressure patterns caused by factors such as sole structure and cushioning. For instance, Pataky et al. [80] achieved a 95% recognition rate using locally linear embedding (LLE) for dimensionality reduction alongside handcrafted features like P100 and COP. Despite this success, the reliance on barefoot data limits the generalizability of these findings to real-world scenarios where individuals typically wear shoes. Similarly, Gafurov et al. [33] demonstrated a 92% accuracy rate in user identification by combining ground reaction force (GRF) and spatiotemporal features. While foundational for understanding temporal gait dynamics, these studies hold limited applicability to this work, which focuses on challenges posed by footwear variability in real-world scenarios.

2.1.3.2 Footwear-Specific Gait Recognition

Research targeting footwear-specific recognition systems has highlighted the complexities introduced by footwear variability and the necessity of robust spatiotemporal feature extraction methods. These studies are directly relevant to this work as they address the shod problem and unseen footwear conditions. For example, Kazemi et

al. [48] employed convolutional neural networks (CNNs) to analyze spatiotemporal features from pressure-sensitive flooring, achieving 85% accuracy. However, the performance dropped significantly to 60% in experiments that included unseen footwear types, illustrating the challenge of generalization across diverse footwear conditions. Similarly, Macdonald et al. [62] used a hybrid model combining principal component analysis (PCA) for dimensionality reduction with an SVM classifier, achieving 87% accuracy in recognizing individuals across three predefined shoe types. Although promising, this study did not address the more challenging open-set scenario involving unseen footwear types.

2.1.3.3 Mixed Conditions (Barefoot and Footwear)

Some studies have explored mixed conditions, analyzing both barefoot and footwear scenarios to understand the impact of footwear on recognition performance. For instance, Guru et al. [40] achieved a 75% recognition rate across barefoot and various footwear types using GRF and COP features. This study highlighted the accuracy degradation when transitioning from barefoot to shod conditions, emphasizing the need for robust models capable of handling variability. Wang et al. [102] utilized deep learning methods to analyze both barefoot and footwear conditions, reporting 88% accuracy for barefoot samples and 80% for mixed footwear conditions. Connor et al. [23] also investigated mixed conditions using a combination of handcrafted features and machine learning classifiers. Their work emphasized the importance of extracting both spatial and temporal characteristics to improve accuracy across varying conditions, achieving over 90.5% recognition performance in their experiments in identification using multiple footsteps. This study highlights the feasibility of hybrid approaches in addressing variability introduced by mixed conditions. The study attributed the challenges posed by footwear to temporal shifts in pressure signals caused by changes in gait dynamics, which aligns closely with the challenges

addressed in this thesis. It is worth noting that Connor et al. [23] leveraged soft biometrics, such as shoe size, to achieve such high accuracies. While impressive, the use of such metrics may risk poor generalization to larger datasets. For example in a small set of users, only one person may have a size 13 shoe, leading the model to rely heavily on that one feature (but a future second user with size 13 shoes may then be misclassified). Consequently, this thesis focuses on features that can be extracted or learned from the actual spatiotemporal information in gait. Future works could then re-introduce soft biometrics through late-fusion to further improve the absolute performance in based on the use case.

Ultimately, barefoot-only studies serve as foundational research but bypass the challenges of real-world gait recognition introduced by footwear variability. Studies that address footwear-specific or mixed conditions are more aligned with the objectives of this work, as they focus on mitigating the effects of external factors such as shoe type on recognition performance. This thesis extends existing work by employing advanced spatiotemporal feature extraction methods and evaluating models in both closed-set and open-set scenarios. By addressing the limitations of prior approaches, it aims to develop a generalizable framework for pressure-based gait recognition in real-world applications.

2.2 Extracting Information From Gait

The raw gait signals captured by high-resolution pressure sensors are inherently high-dimensional. For example, in the popular CASIA-D dataset (see Section 3.1), each footstep is captured as a 60x40x100 matrix of floating point pressure readings (i.e., 240k data points). Preprocessing and effective *feature extraction* are therefore critical to reducing data dimensionality and increasing information density for pressure-based gait recognition. In the context of ML, feature extraction is the process of selecting

relevant information and synthesizing new dimensions to better represent the data [27]. This is typically achieved through one of the three different feature extraction pipelines outlined in Figure 2.3: (a) Machine Learning using hand-crafted features, (b) Feature Learning Networks using structured convolutional features with prefixed or data-adapted filters, and (c) Deep Learning, where filters are learned during training. In the Machine Learning pipeline, hand-crafted features are extracted and passed into a machine learning classifier such as SVM. In the Feature Learning Network pipeline, convolutional features are extracted using prefixed or data-adapted filters, followed by nonlinearities and pooling. These features are then classified, again with an ML classifier such as an SVM. Finally, in the Deep Learning pipeline (i.e., right column), convolutional features are learned during training, followed by activation and pooling layers. The resulting feature vectors are classified using a fully connected layer.

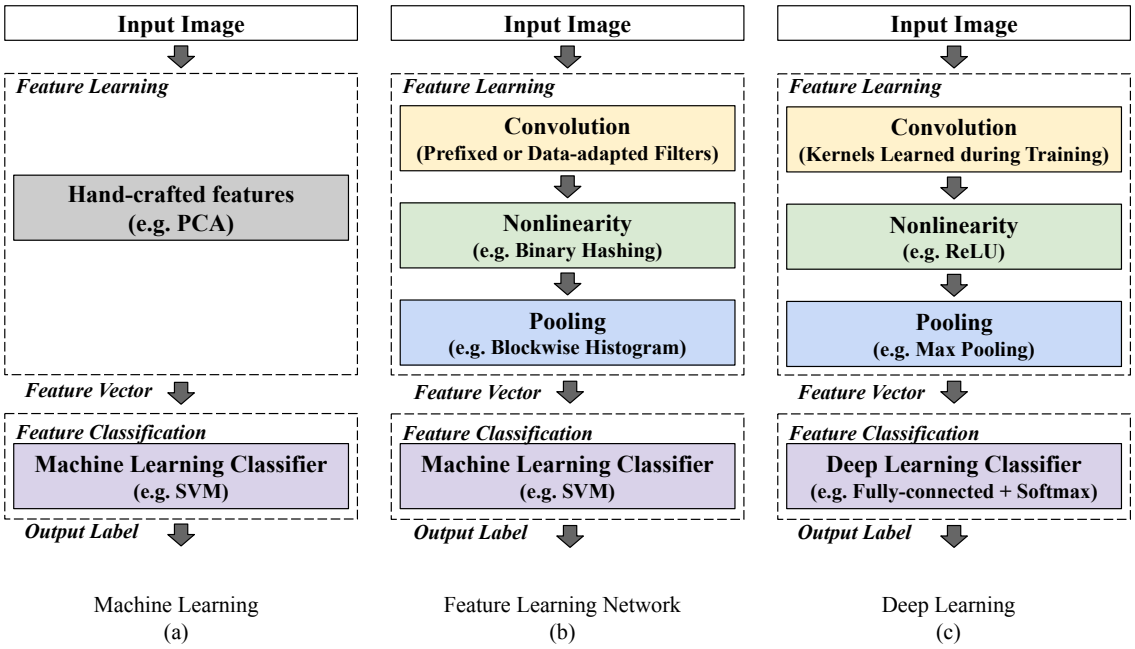

Figure 2.3: Overview of classification frameworks using spatial feature extraction models.

Importantly, the extraction of *relevant* information - in this context - refers to discriminative information that promotes class separability (i.e., the resulting system's ability to assign a class label to a given user correctly) rather than simply describ-

ing maximal data variance. Gait data must be preprocessed using feature extraction methods to isolate characteristics that best represent each individual's unique gait pattern [88]. This is a fundamental step in building a robust, efficient recognition system [81]. These approaches lead to three types of features: a) Hand-crafted features, b) Deep-learned features, and c) Feature learning. Each approach is described below.

Approach 1 - Hand-Crafted Features: Hand-crafted features are designed by human experts based on domain knowledge to capture relevant patterns in the data. In gait recognition, hand-crafted features such as *ground reaction force* (i.e., GRF - a time-series of sum force exerted on the floor throughout the footstep progression) and *peak pressure images* (i.e., P100 - an image depicting the peak pressure reading from each sensor) are common [69].

Hand-crafted features can be effective with smaller datasets and require less computational power because their design draws on broader knowledge than may be represented in a given training set. Although intuitive, these features are predetermined and may thus fail to capture the full complexity of the data. Nevertheless, this pre-selection of known features based on domain knowledge can reduce the amount of data required to inform subsequent ML models.

Figures 2.4, 2.5, and 2.6 show examples of the three most common hand-crafted features used in pressure-based gait analysis.

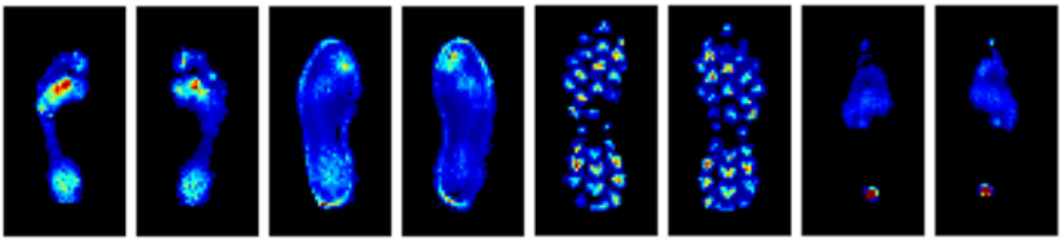

Figure 2.4: Examples of P100 pressure images for different pairs of shoes. Each subfigure represents the peak pressure values, highlighting the spatial distribution and intensity of the different pressure points due to footwear.

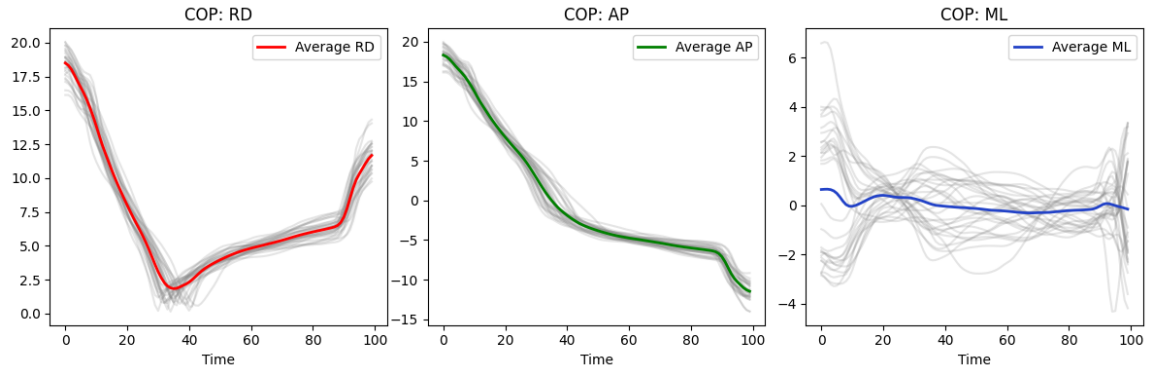

Figure 2.5: Example of the Center of Pressure (COP) time series for a representative user. The plot illustrates how the vertical, anterior-posterior, and medial-lateral force components vary over time during walking.

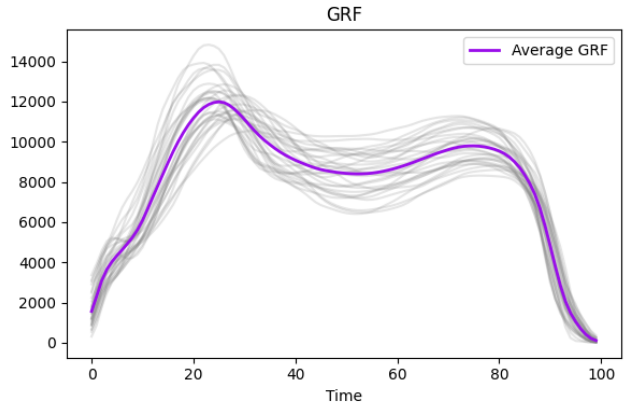

Figure 2.6: Example of the Ground Reaction Force (GRF) time series for a representative user. The plot illustrates the variations in vertical, anterior-posterior, and medial-lateral force components over time during walking.

Approach 2 - Deep-Learned Features: DL architectures such as CNNs have demonstrated promising results in gait recognition tasks [17]. CNNs typically consist of convolutional, pooling, and fully connected layers. The convolutional layers are responsible for extracting features while pooling layers help with dimensionality reduction. The fully connected layers then map the extracted features to the final classification output. Arguably, the most important building blocks of the CNN architecture are the convolutional layers, which consist of kernels (or filters). In traditional CNNs, the parameters of these kernels are carefully learned through back-propagation and gradient descent during training.

Deep networks, in general, are powerful because of their high capacity to identify discriminative patterns in data. However, one major limitation of DL approaches is their tendency to overfit to specific training data and failure to generalize - therefore, large labeled datasets are typically required to achieve optimal DL performance - *especially* when data samples exhibit high-dimensionality (e.g., recall the 240k data points representing each footstep). In real-world scenarios, gathering thousands of training samples per user is often impractical, which may limit the viability of DL approaches in commercial applications. Moreover, DL models can be computationally expensive, requiring substantial resources for training, which may make them less suitable for resource-constrained environments.

Approach 3 - Feature Learning: Feature learning is a hybrid approach that strives to strike a balance between the *hand crafted* and *deep learned* approaches. As with DL, it involves training a neural network; however, rather than starting with fully unconstrained parameters, feature learning aims to reduce the training search space by introducing constraints derived from common hand-crafted features and/or statistical properties of the data. These constraints are typically derived through structured feature extraction methods such as clustering or dimensionality reduction techniques. For example, PCANet, one of the first FLNs, incorporated Principal Components Analysis (PCA) within a DL framework [12].

By incorporating such structured methods, feature learning models can sometimes extract more meaningful representations from smaller datasets than standard DL approaches [8, 75, 109, 12, 39, 57].

2.3 Spatial, Temporal, and Spatiotemporal Representations of Gait Data

The pressure data captured from flooring sensors are commonly described as three-dimensional, with height, width, and time dimensions [79]. In the context of this thesis, this 3D representation - the "raw" footstep recording - is referred to as a *Pixel Time Series (PTS)* to emphasize the spatial and temporal nature of the recorded pressure values. Depending on the application, different information types can be extracted from this raw representation.

- *Spatial features:* The PTS is converted to an image or set of images, such as to support analysis through computer vision methods.
- *Temporal features:* The PTS is down-sampled spatially to a univariate or multivariate time series, such as to support analysis through established time series approaches.
- *Spatiotemporal features*: A combination of spatial and temporal information. Analysis can include combinations of computer vision and time series approaches.

2.3.1 Spatial Features for Gait Recognition

In the past, pressure-based gait analysis was primarily conducted using force plates, which provided a single measure or pressure vector. However, the advent of pressure-sensitive flooring has made it possible to capture a 2D pressure "image." This novel format naturally motivates the exploration of spatial and image-based features to capture the unique underfoot pressure patterns defined by the pixels in an individual's footsteps. Various spatial features such as P100, P90, P80, P70, P60, Peak Time Interval (PTI), and Mean Pressure Image (MPI) have been explored to represent

these patterns effectively [79], and yield tangible images of pressure distributions that can be readily interpreted.

- P100, P90, P80, P70, P60: These features represent the peak pressure values from each pixel time series at different percentiles. P100, for example, captures the maximum pressure value (See Figure **??**), while P90 captures the 90th percentile, and so on. These features are effective because they summarize the pressure distribution over time, highlighting pressure differences across the foot.

- Peak Time Interval (PTI): PTI captures the time interval during which the pressure exceeds a certain threshold, typically determined as a percentage of the maximum observed pressure across all samples (e.g., 50% or 75% of peak pressure). This feature helps in understanding the duration of high-pressure points during a gait cycle, which can be distinctive for individuals. Although this feature does capture temporal information, its use across many pixels enables a spatial comparison between them.

- Mean Pressure Image (MPI): MPI averages the pressure values over the entire gait cycle, providing a holistic view of the pressure distribution. This feature smooths out the temporal variations and highlights the overall pressure patterns.

Of these features, P100 has been shown to be effective in capturing relevant spatial characteristics [79], and it provides a robust representation of spatial features for gait recognition [80]. Consequently, P100 was chosen for use in this work.

Even when compressed using P100, when derived from high-resolution pressure sensors, the signal is inherently high-dimensional. This creates challenges for conventional ML approaches (e.g., through the "curse of dimensionality"), often requiring further processing through additional feature extraction or dimensionality reduction steps[111, 25, 49, 23]. For example, dimensionality reduction techniques such as locally linear embedding (LLE), kernel PCA, Laplacian eigenmaps, and normalized

spectral clustering with symmetric Laplacian have been explored for foot pressure-based identification [80].

2.3.2 Temporal Features for Gait Recognition

Temporal features have been widely explored for gait recognition because of the dynamic nature of gait. Conventional handcrafted temporal features, such as Center of Pressure (COP) and Ground Reaction Force (GRF), have been shown to be effective in recognizing gait patterns when individuals are barefoot or wearing known shoe types (within *closed set* problem) [79]. While it is perhaps intuitive that spatial information may be affected by footwear, certain types of shoes (such as those with different sole structures or cushioning) may also alter the temporal dynamics of gait [68]. Prior to this work, it was unclear how these features might perform in open-set scenarios where the system has not been trained on these variations. Consequently, Chapter 5 explores the quality and generalizability of such handcrafted temporal features.

Deep temporal learning algorithms have also gained popularity in recent years for analyzing gait data in biometrics. RNNs [46], including variants like Long Short-Term Memory [19] and Gated Recurrent Units (GRU) [29], have shown promising results in pressure-based gait analysis. These models are effective at capturing longer-term dependencies in the data, such as information about an earlier heel strike that may influence the subsequent toe-off [63].

Although many other deep models are emerging in time series analysis, the LSTM remains a common reference model for the potential of deep temporal learning [10, 6, 18]. Regardless of the specific implementation, deep temporal models learn by exploiting the temporal dependencies and patterns in the sequential data, leading to improved recognition performance [95].

2.3.3 Spatiotemporal Features for Gait Recognition

The collection of gait pressure distributions over time makes pressure measurements an inherently spatiotemporal signal. Consequently, spatiotemporal approaches may offer a more complete view of these dynamics by integrating spatial configurations (e.g., foot placement) with temporal dynamics (e.g., the rollover of the foot).

In vision-based gait recognition, two recent deep-learning techniques, GaitPart [30] and GaitSet [13], have demonstrated the effectiveness of spatiotemporal approaches. GaitPart explicitly partitions the gait video into different regions of the body, extracting spatial features from each part separately to enhance the focus on specific areas. GaitSet emphasizes different elements of the temporal gait cycle, capturing how gait evolves over time and extracting dynamic features across multiple frames.

These approaches were specifically developed to address challenges in vision-based recognition, such as changes in clothing conditions and/or variations in camera angle [52]. To overcome these challenges, GaitPart emphasizes those regions that remain relatively unchanged (e.g., the legs and torso regions are likely unaffected by the presence of a hat or jacket). This partitioning scheme is likely what makes GaitPart resilient to changes in external factors such as clothing or occlusion. This issue is somewhat analogous to the impact of different types of footwear in pressure-based systems. However, whereas different clothing may change a person's appearance in vision-based approaches (an issue not present in pressure-based approaches), different footwear may actually change a person's underlying walking patterns [43]. As previously introduced, individuals may walk differently in stiff boots versus loose sandals, causing variations in the spatial footprint as well as in the dynamics of the gait. In some cases, these variations may be localized to specific foot regions, suggesting that partitioning the footstep into regions (e.g., forefoot, midfoot, heel) may similarly help capture more robust features. Therefore, Chapter 5 explores the suitability of GaitPart for use in pressure-based gait recognition tasks and then introduces a novel

recognition model informed by its partitioning approach, but optimized for pressure-based recognition.

Chapter 3

Dataset and Preprocessing

3.1 Datasets

Two different datasets were used in this work. Chapters 4 and 5 present approaches developed using the openly available CASIA-D dataset. In Chapter 6, the generalization of the findings is evaluated using a custom UNB dataset that was collected in parallel with this research. Both datasets are introduced, below.

3.1.1 CASIA-D

CASIA-D is a public gait dataset widely adopted in gait biometrics research (e.g., [23, 32, 92, 112]). It consists of pressure-based gait data recorded using a RSScan Footscan device consisting of a 255-by-64 grid array of pressure sensors with 5mm $\times$ 5mm spatial resolution and 100Hz sampling rate. During collection, participants walked across this grid capturing 3-4 footsteps per pass. The dataset consists of a series of rectangular footstep data records extracted from the larger pressure grid. Two versions of each footstep are provided: a) *raw* - no post-processing applied, and b) *aligned* - de-correlation and registration post-processing applied (see [112] for details). In both cases, each footstep is represented by a 60$\times$40 pixels (30$\times$20cm) $\times$

100 (time increments) matrix. The time dimension represents absolute clock time for raw footsteps (i.e., padded to 1sec total, in 10ms increments), whereas it represents normalized percentiles in the aligned samples (i.e., representing the duration of total foot contact in 1% increments). The spatial dimensions represent absolute distance in both the raw and aligned versions. In this thesis, the aligned footsteps were used for all analyses, with even further preprocessing applied as described below (see Preprocessing Section 3.2).

Data were collected while participants walked in one of three footwear conditions: *barefoot* and while wearing one of two distinct pairs of personal footwear (i.e., *shod*). All 97 participants included in the dataset completed the barefoot walking trails. However, only a subset of participants completed the shod trails. This work makes use of the subset of 14 participants who had at least 30 captured footsteps in each of the three footwear conditions, yielding a total of 90 samples per participant. Figure 3.1 shows examples of samples taken from the CASIA-D dataset.

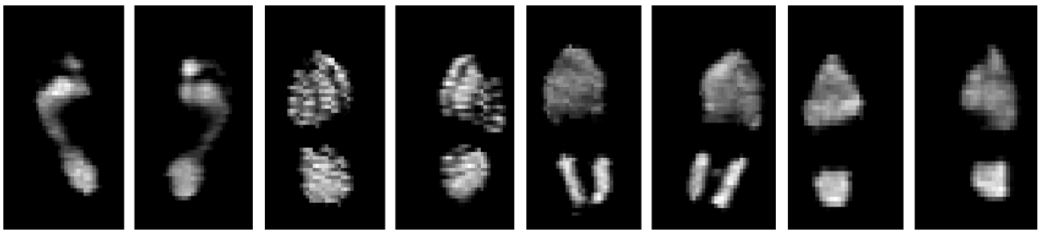

Figure 3.1: CASIA D Data Samples.

3.1.2 UNB Dataset

The UNB dataset was collected using pressure-sensitive flooring produced by Stepscan Technologies Inc., and installed in the Health Technologies Lab at UNB [62]. Over the course of three years, gait data were recorded from more than 150 participants in a controlled environment. All participants provided informed consent as approved by UNB's Research Ethics Board (REB #2021-071).

The data were captured using a 6x2 grid of pressure sensing tiles, each consisting of a 120x120 grid of 16-bit pressure sensors at 5mm x 5mm resolution and sampled at 100Hz. The resulting walkway measured 3.6m x 1.2m (720 x 240 pixels). Each trial lasted 90 seconds, during which participants walked back and forth across the walkway. Four different footwear conditions were recorded for each participant: bare-foot, a common shoe (Adidas Grand Court 2.0), and two pairs of each participant's personal, self-selected, footwear.

As with CASIA-D, the footstep records were provided in two formats: *raw* and *aligned*. The raw data represent the unprocessed pressure readings from the tiles, while the aligned data underwent post-processing for de-correlation and registration, similar to the CASIA-D dataset, to reduce variability in foot placement and angle. Each segmented footstep is represented as a three-dimensional (3D) pixel time-series matrix with dimensions of 64x36 pixels (spatial) by 100-time increments (temporal).

This study uses a normal/self-paced walking speed, which was found to be most consistent with the CASIA-D dataset, for comparison and analysis. There are other walking conditions (e.g., slow, fast) in the broader UNB dataset, but these are outside the scope of this research. Figure 3.2 shows examples of samples taken from the UNB dataset.

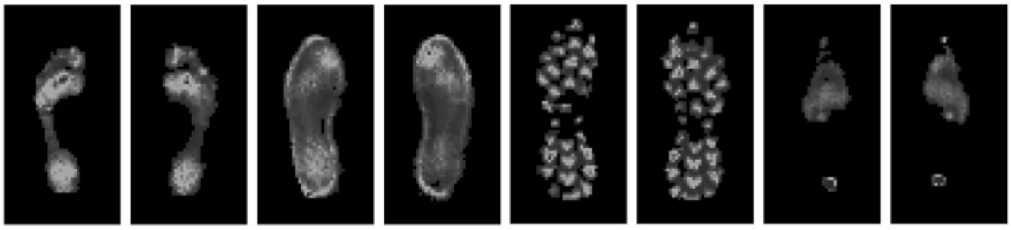

Figure 3.2: UNB Data Samples.

3.2 Preprocessing

The preprocessing steps described in the following subsections were applied consistently to each of the footstep records in both the CASIA-D and UNB datasets. Figure 3.3 provides a block-diagram summary of the preprocessing pipeline.

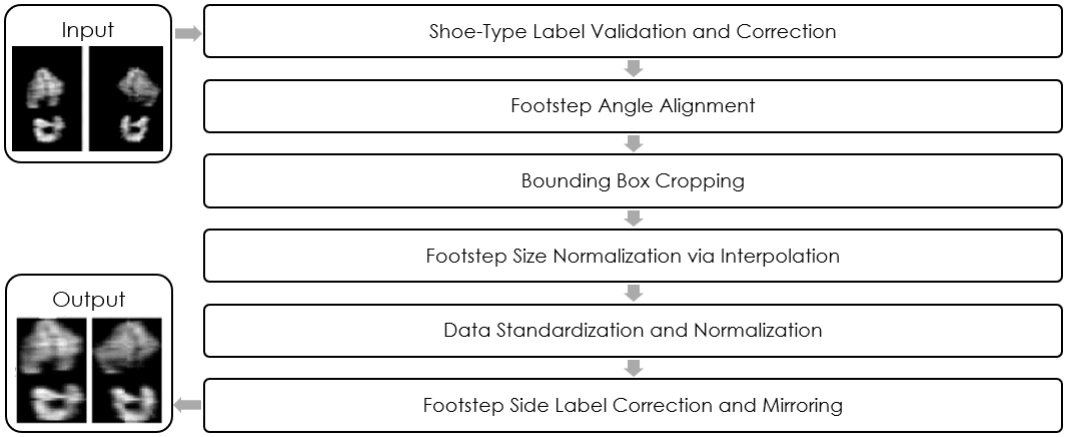

Figure 3.3: Preprocessing framework.

3.2.1 Shoe-Type Label Correction

Through visual inspection of the CASIA-D dataset, several labeling errors were identified and corrected (e.g., incorrectly labeled barefoot samples as shod). This step was necessary to ensure that samples were correctly assigned within the three footwear categories. Similarly, the UNB dataset underwent two rounds of manual inspection to ensure label correctness across all collected samples.

3.2.2 Angle Alignment

Despite the pre-existing alignment procedure applied to the CASIA-D records, the aligned samples still exhibited large regions of zero pressure and residual foot rotation angles. Both of these characteristics presented challenges for some approaches in this work (e.g., see Chapter 5.3). Therefore, each footstep was further cropped

to minimize zero-padding and rotated to correct for residual angle. Cropping was automated by computing a minimum area bounding box around the P100 representation of each footstep, and the entire 3D footstep matrix was cropped accordingly. Rotation corrections were performed by estimating a residual angle for each sample. This was achieved using Principal Component Analysis (PCA) on the P100 representation, where the primary component of the foot pressure distribution provided the orientation axis. The angle of this axis relative to the horizontal was then used to align each 3D footstep data sample. This step was critical, as some of the proposed approaches rely on footstep angle, and residual angle biases could negatively impact model performance (see Chapter 7).

3.2.3 Bounding Box Cropping

Each footstep image was then cropped to a new minimum area bounding box by repeatedly removing zero-only border rows and columns until the first non-zero value was encountered. This ensured that each cropped footstep retained only relevant pressure data.

3.2.4 Footstep Interpolation

After cropping, each footstep matrix varied in size depending on the participant's foot dimensions and pressure distribution. To avoid size-related classification bias, an interpolation process was applied to standardize the dimensions of each footstep. Although size can be used as a feature, it is ill-advised for smaller datasets because of the likelihood of falsely accepting new users with similar foot sizes [67, 74]. In both datasets, the largest footstep was selected as the target dimension for interpolation to avoid down-sampling artifacts. Thus, in the CASIA-D dataset, footsteps were interpolated to a common dimension of 40x28 pixels, while in the UNB dataset, footsteps were resized to 64x36 pixels. This interpolation ensured that all footsteps

had consistent dimensions across both datasets.

3.2.5 Data Standardization and Normalization

The samples were standardized to a mean of zero and a standard deviation of one, based on scaling parameters calculated from the training set. These parameters were then consistently applied to both the training and test sets to avoid information leakage from the test into the training process. Standardization prevents features with larger ranges, but not necessarily more information, from dominating the learning process. Furthermore, standardization can improve computational efficiency by speeding up convergence during model training, as machine learning algorithms such as gradient descent perform better when input data are normalized.

3.2.6 Footstep Side Label Correction and Mirroring

The classification methodology developed by MacDonald et al. [62] was used to verify the accuracy of the left vs right foot labels in both datasets, with manual corrections made for any labeling errors. Consistent with prior gait recognition research (e.g., [98]), all right foot samples were flipped horizontally to align them with left foot samples. This normalization step allowed all footsteps to be processed consistently within a single classification model.

3.3 Implementation

As described in Chapter 2, the gait recognition problem can be framed in different ways. Recall that a *verification* system is designed to confirm a claimed identity, whereas an *identification* system is designed to determine identity outright. Similarly, a *closed-set* system is only expected to perform on previously encountered user-footwear data pairings, whereas an *open-set* system is expected to generalize to

previously unseen pairings. For conciseness and because it is the focus of the larger project, Chapters 4 and 5 present verification results performed on CASIA-d (i.e., both open- and closed-set variants). In Chapter 6, the generalizability of the work is demonstrated by applying the developed models to both the UNB dataset and the identification problem.

3.3.1 Constructing a Verification System

For the verification models in this work, the system was tasked with distinguishing between genuine and imposter samples by selecting a decision threshold derived from the training data. Specifically, the threshold is optimized by balancing the False Acceptance Rate (FAR), the percentage of imposters incorrectly classified as genuine users, and False Rejection Rate (FRR), the percentage of genuine users mistakenly classified as imposters. Using an alpha parameter set to 0.5 to give equal weight to both types of errors, the system seeks the threshold that minimizes the combined error rate:

$$\text{Combined Error} = \alpha \times \text{FAR} + (1 - \alpha) \times \text{FRR}$$

where $\alpha = 0.5$. This balanced approach finds a threshold that tries to differentiate between genuine users and imposters without biasing toward either type of error.

3.3.2 Data Use in the Open- and Closed-set Scenarios

To provide a fair comparison, a common evaluation framework was used to assess the performance of the different system architectures presented in this work. Although verification performance is the primary outcome measure considered in this work, Chapter 6 explores the use of each proposed system for the identification task to explore the generalizability of the findings of the work to this more challenging problem.

Closed-set Evaluation

Closed-set (or *within-domain*) evaluation refers to the scenario when the performance of a gait recognition system is evaluated using only footwear observed during training (including barefoot and all types of shoes). As mentioned in section 2, this does not mean that only barefoot or a specific shoe type was considered - rather, it means that all footwear conditions were presented to the system during training. Although a barefoot-only requirement may be practical in some cases (e.g. forced removal of footwear in an airport), many use cases would benefit from the ability to also recognize users when they are wearing shoes. Therefore, in this approach, the test set contained samples from all included footwear conditions, not exclusively barefoot samples. This approach allows the system's performance to be assessed comprehensively across different types of footwear, reflecting real-world scenarios where users may wear a variety of shoes.

Open-set Evaluation

Open set (or *cross-domain*) evaluation explores how the biometric system performs when users wear new/unfamiliar footwear that was not previously presented to the system during training. In this scenario, train-test splits are formed by footwear category using a leave-one-footwear-out approach - in each fold of the evaluation, one footwear condition for each participant is withheld from training to form the test set. In this case, the models encounter entirely new data and must extract features from these out-of-domain, unfamiliar samples. Consequently, this task is considered much more complex.

Furthermore, since each participant wore two pairs of their own personal shoes, the open-set problem is not simply a matter of generalizing system performance to a *single* new footwear condition/domain, but instead of generalizing to $N <= n\_participant$

domains, with each participant's unseen footwear condition potentially being a completely novel and distinct domain generalization task.

Data Use

Each participant in the CASIA-D dataset has barefoot, Shoe 1, and Shoe 2 samples. Therefore, the open set evaluation framework is divided into two folds: Fold 1 (Training Set = Barefoot and Shoe 1; Testing Set = Shoe 2) and Fold 2 (Training Set = Barefoot and Shoe 2; Testing Set = Shoe 1)[1]. In this way, the train test ratio was 66.7% and 33.3%. To maintain a fair comparison between the two evaluation pipelines, for the closed-set evaluation, the train and test were randomly split with a ratio of 66.7% and 33.3%.

Accordingly, the differences in performance observed in closed vs. open set evaluation results are due to the presentation of the unknown footwear during testing. Figure 3.4 illustrates the various combinations of results shown at various points throughout this dissertation.

3.3.3 Evaluation Metrics

Performance was evaluated using balanced accuracy (BACC), the arithmetic mean of sensitivity and specificity. Balanced accuracy was selected, in part, because of its robustness to imbalance in the class label distributions (e.g., negative samples often greatly outnumber positive samples in the verification problem framing).

Comparisons between the closed-set and open-set performances were conducted using paired t-tests (two-tailed) to determine if there were statistically significant differences between the two conditions. A significance threshold of $p < 0.05$ was applied. Both p-values and Cohen's D effect size are reported. Cohen established

[1]Note that the third possible fold (i.e., Training Set = Shoe 1 and Shoe 2; Testing Set = Barefoot) was intentionally omitted as it is less desirable in practice (e.g., most people don't go to work in their bare feet).

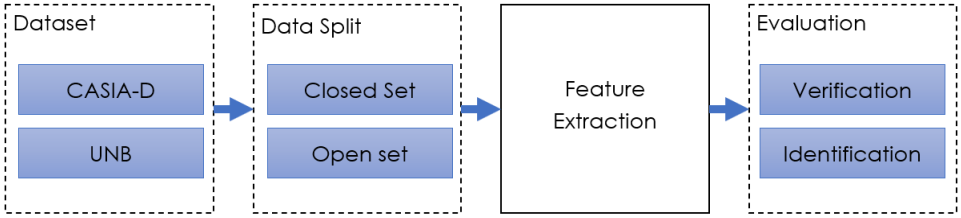

Figure 3.4: Evaluation framework

that values of D above 0.2, 0.5, and 0.8 can be considered to be small, medium, or large, respectively [22].

3.3.4 Programming Language, Libraries, and Resources

All models were implemented in Python. `Scikit-Learn` and `SciPy` were used for conventional ML models. `Pytorch` and `Tensorflow` were used for DL implementations. Packages such as `Numpy` and `Pandas` were used for numerical analysis, `Seaborn` for statistical analysis, and `Matplotlib` and `Plotly` for visualizations of the results. Results were computed using a Lambda Scalar 4U Server with 4 NVIDIA A1000 GPUs.

Chapter 4

Conventional Feature Extraction & Feature Learning

This chapter describes the performance of the verification system when using the raw footstep data or one of three common hand-crafted feature representations, classified using an SVM classifier. This analysis serves as a baseline against which other models may be compared. Additionally, the use of 9 different conventional dimensionality reduction algorithms is explored. These methods seek to compress the dimensionality of the data while preserving the information content and have been shown to enhance gait recognition performance [80].

Finally, three feature learning networks (FLNs) are presented. Two of these, PCANet and PCANet+, were adopted from established literature whereas the other, UMAPNet, is a new contribution of this dissertation. The results will demonstrate that UMAPNet significantly outperforms the existing FLN methods by effectively capturing spatial features for gait recognition.

4.1 Establishing Baseline Performance

Baseline performance was established by training separate instances of an SVM classifier model on four feature representations (i.e., the raw pixel time series (PTS), P100, COP, and GRF, as defined in Chapter 2). These results are labeled as *no dimensionality reduction - NDR* and accompany the remaining results presented below.

4.2 Dimensionality Reduction Techniques

The goal of feature extraction is to improve discriminative information density (i.e., the ability to discern between samples from each class/category). However, feature extraction generally leverages domain knowledge, such as understanding the physical characteristics of gait patterns, to identify features that are likely to be informative for classification tasks. Temporal features like GRF and COP have emerged from decades of research, and reflect the dynamic interaction between the foot and the ground during walking. From these signals, more detailed metrics can be derived, such as "time to peak heel loading" or "center of pressure trajectory," which provide deeper insights into individual gait patterns. However, in some cases, not all information may be captured from these engineered features or their information density may be low. Therefore, a common approach is to explicitly leverage dimensionality reduction techniques that aim to improve information density by identifying a reduced representation of the data or features that retain as much information as possible. Although the resulting feature set is not necessarily optimal for class discrimination, in practice, the use of these techniques often leads to improved system performance.

Each combination of the four feature representations and the nine dimensionality reduction techniques explored in this work was used to train and evaluate a gait recognition model.

- **Independent Component Analysis (ICA)**: ICA has been used to extract

independent features from pressure information for more accurate gait recognition [60]. The optimal number of ICA components in this work was empirically determined to be 70, providing a balance between capturing significant independent sources and maintaining computational efficiency.

- **Principal Component Analysis (PCA)**: PCA transforms the data into orthogonal components, retaining most variability. It has been shown to provide the best recognition quality with SVM classifiers for higher-dimensional feature spaces in gait recognition [58]. The number of principal components in this work was chosen to retain at least 95% of the variance.

- **Locally Linear Embedding (LLE)**: LLE is a non-linear technique that preserves local relationships between data points, which can be useful for capturing the local structure of plantar pressure data [53]. The number of embedding dimensions was chosen as 70 to preserve the nearest neighbor relationships of the data as well as the previous optimum value of LLE for the pressure-based gait data [80].

- **Modified Locally Linear Embedding (MLLE)**: MLLE is an extension of LLE that aims to improve robustness and accuracy by better handling noise and outliers in the plantar pressure data [110]. The number of dimensions was chosen to be 70, to be consistent with LLE.

- **Factor Analysis (FA)**: FA is used to identify underlying factors contributing to gait patterns from the observed, correlated variables in plantar pressure data [2]. The number of factors, 40, was determined based on retaining 95% of the variance.

- **Self-organizing maps (SOM)**: SOM is an artificial neural network that produces a low-dimensional, discretized representation of the input space, which

can be used to visualize and cluster plantar pressure data for gait recognition [11]. The map size was chosen to balance between resolution and generalization, yielding a grid size of 10x10.

- **Isometric Mapping (Isomap)**: Isomap is a non-linear technique that preserves geodesic distances between data points, effective for data that lies on a curved manifold, such as the complex patterns in plantar pressure data [87]. The number of dimensions was chosen empirically to be 10 neighbors and 50 dimensions to capture the intrinsic dimensionality of the data.

- **t-Distributed Stochastic Neighbor Embedding (t-SNE)**: t-SNE is a non-linear technique well-suited for embedding high-dimensional data into a low-dimensional space for visualization, which can help in visualizing the complex structure of gait patterns [21]. A 3D embedding was used here.

- **Uniform Manifold Approximation and Projection (UMAP)**: UMAP is similar to t-SNE but often faster and better at preserving both local and global data structure, widely used for visualizing complex datasets, and can be particularly useful for pressure-based gait recognition [65]. The number of neighbors was set to 15 and the output dimensions to 50, chosen based on balancing computational efficiency and performance.

4.2.1 Implementation

Each of the 36 resulting data representations (i.e., 9 techniques x 4 features) was evaluated using four conventional classifiers: a Support Vector Machine (SVM), K Nearest Neighbours (KNN), a Decision Tree (DT), and a linear discriminate analysis (LDA). However, since SVM consistently outperformed all other models, only the SVM performance outcomes are reported in this dissertation.

The training set explained in Section 3.3.2 (i.e., two-thirds train; one-third test) was further divided into training and validation sets to select the optimum parameters of feature extraction methods. Consequently, a data split was adopted as follows: Train 50%, Validation 16.7%, and Test 33.3.%. The parameters for each dimensionality reduction technique, such as the number of components and neighbors, were optimized using a grid search method on the validation sets. To identify optimal parameters for each dimensionality reduction method, a pilot grid search was conducted using P100 data from three randomly selected users from the CASIA-D dataset. These optimal parameters were then consistently applied across all data representations to ensure a standardized approach following the evaluation protocol as explained in section 3.3.2.

4.2.2 Results

The balanced accuracies for closed-set and open-set verification are displayed in Figures 4.1 and 4.2, respectively. These figures compare the baseline performance (i.e., *no dimensionality reduction - NDR*) against the nine dimensionality reduction methods $\times$ four feature combinations.

Figure 4.7 presents the closed-set balanced accuracies (i.e., when all tested configurations of participant and footwear were provided during model training). Overall, PTS outperformed the P100, GRF and COP features, both as a baseline and when processed using all but one of the dimensionality reduction techniques. When applied to PTS, the UMAP, PCA, ICA, LLE, and MLLE approaches all outperformed the baseline (87.89%), each achieving over 94%. LLE-PTS, the top-performing model, reached 97.30% - significantly higher than all other approaches ($p < 0.01$, $d > 0.88$). These findings highlight the strength of non-linear techniques like LLE in preserving discriminative information in a closed-set scenario, which agrees with previous research [80].

Figure 4.8 presents the open-set balanced accuracies (i.e., when tested participant-footwear configurations were *not* provided during model training). Three main observations should be noted. First, performance is substantially lower than closed-set, with the best-performing combination (i.e., PCA-GRF) reaching a BACC of only 71.64 %. This reinforces the difficulty of footwear generalization. Second, in contrast to the closed-set, PTS is no longer the top-performing feature. Rather, GRF with PCA dimensionality reduction (i.e., GRF-PCA) was the top-performing configuration with 71.64% balanced accuracy, outperforming all other approaches, although not significantly. Interestingly, however, COP outperformed GRF in most of the dimensionality reduction methods, including the NDR case. These results suggest that temporal features like GRF and COP may be more robust to variations introduced by unseen footwear compared to spatial features like PTS. Third, despite the effectiveness of the various dimensionality reduction techniques in the closed-set tasks, many of these techniques did not generalize as well to the open-set problem. This suggests some element of over-fitting in the closed-set variant. Interestingly, however, PCA did not follow this broader pattern - with three of the four features performing better with PCA than with NDR. This consistent performance of PCA highlights its potential for improving generalization in open-set scenarios, motivating further exploration in the following section.

4.2.3 Discussion

Interestingly, the best-performing features and dimensionality reduction techniques differed between the closed-set and open-set problems. As expected, closed-set performance (i.e., 97.30%) was higher than open-set (i.e., 71.64%) - the 25% degradation reinforces the difficulty of the open-set problem. However, although PTS was the best closed-set feature with many dimensionality reduction techniques further improving performance, neither of these findings generalized to the open-set problem. Rather,

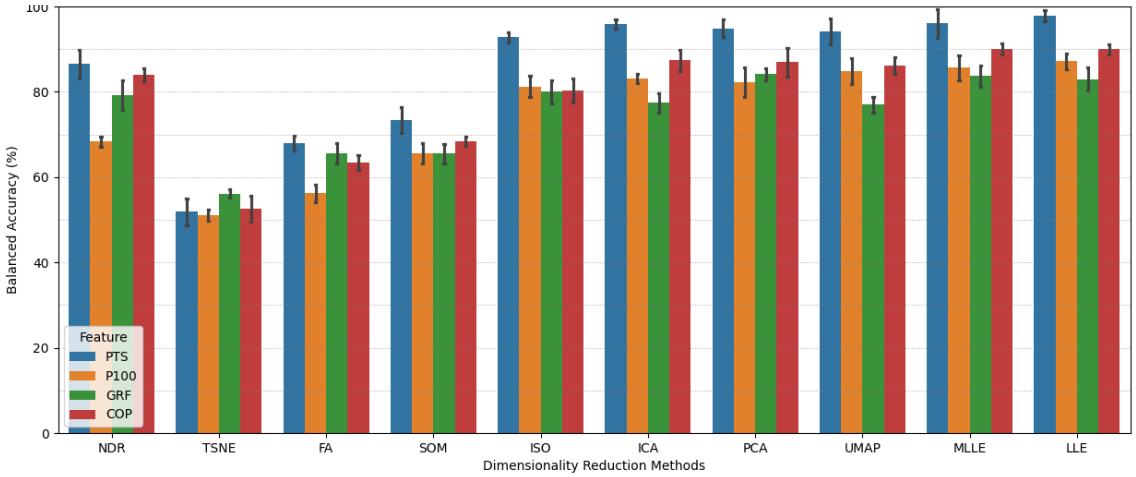

Figure 4.1: The closed-set BACC results when applying the various dimensionality reduction techniques on the pixel time series (PTS) or the P100, GRF, or COP features for person verification. NDR refers to the case when no dimensionality reduction was applied.

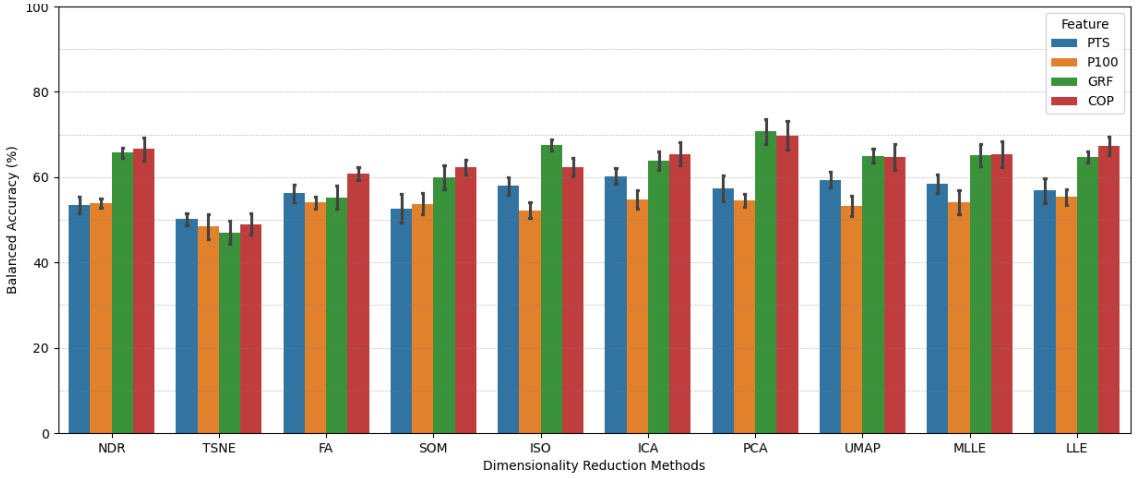

Figure 4.2: The open-set BACC results when applying the various dimensionality reduction techniques on the pixel time series (PTS) or the P100, GRF, or COP features for person verification. NDR refers to the case when no dimensionality reduction was applied.

the two temporal features (i.e., GRF and COP) were the best open-set performers, with the NDR approach significantly outperforming the dimensionality reduction techniques (p-val < 0.0121, cohen-d > 0.88) suggesting that dimensionality reduction led to overfitting and was less effective in generalizing to unseen footwear conditions.

The fact that PTS attained the highest closed-set performance suggests that the

user-discriminative information present in gait is not fully captured by any one of the features evaluated. This is further evidenced by the boost in performance attained through dimensionality reduction, particularly by methods like UMAP and LLE, which effectively preserve both local and global data structures. In contrast, linear techniques like PCA, while robust, may miss some finer spatial relationships. This motivates the exploration of other feature representations. One simple possibility is a novel feature created by concatenating those evaluated in this chapter into a single data structure, as such a fusion approach could further improve performance.

Also, an important consideration when interpreting these findings is that all dimensionality reduction technique model parameters were optimized for use with P100. Further performance improvements may be possible with other features through explicit optimization.

Nevertheless, the fact that both GRF and COP continued to perform well when extended to the open-set problem suggests that temporal features may be more resilient to variations in footwear [56, 78]. This is possibly due to their ability to filter out irrelevant spatial information, thereby focusing on gait characteristics that are more resilient to changes in footwear. However, it is important to note that some identifying information must also be encoded spatially. For example, although COP is often categorized as a temporal-only feature, in reality, it is spatiotemporal — capturing the evolution of the spatial position of pressure over time. This likely explains why COP outperformed GRF (i.e., which is purely temporal) in many cases, and motivates the exploration of approaches that may better extract this spatial information.

4.3 Feature Learning Networks

As introduced in Chapter 2, FLNs provide a hybrid approach that merges traditional feature extraction techniques with the power of DL. These networks aim to optimize the feature extraction process by capitalizing on domain knowledge incorporated into conventional dimensionality reduction methods, while also incorporating the learning capabilities of neural networks. In this section, we explore two prominent FLNs, PCANet, and its enhanced variant PCANet+, and introduce a novel FLN architecture developed in this dissertation research, UMAPNet.

Importantly, FLNs are often rooted in CNN architecture and are designed to consume 2D "image" input data. Since this is the case for the three FLNs investigated in this work, performance is only evaluated using the P100 feature. This therefore served as an exploration of whether more effective spatial information could be extracted.

4.3.1 PCANet & PCANet+

PCANet, introduced in 2015 [12], is a hybrid of conventional dimensionality reduction and DL that uses PCA to transform the input data into a set of orthogonal components, simplifying the input for further processing by the network. PCA is applied at each convolutional layer of the network, replacing the traditional learned filters with PCA filters. This approach is designed to extract and retain those components that best explain the variance in the data, hoping to capture useful information. The rest of the architecture of PCANet, as shown in Figure 4.3, then processes these extracted features through subsequent layers, including a hashing layer and block-wise histograms that help in producing a robust and compact feature representation for classification.

PCANet+ [57] builds on PCANet by incorporating better regularization techniques, the introduction of non-linearity after the PCA layers, and the use of multi-

scale approaches to capture features at different resolutions. These modifications, as shown in Figure 4.4, are intended to enable PCANet+ to better generalize across different datasets and variations in the input data.

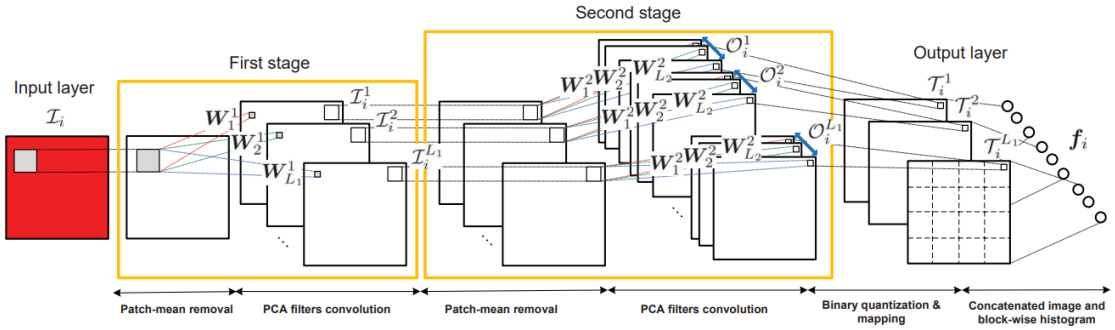

Figure 4.3: Architecture of PCANet as described in its original paper [12].

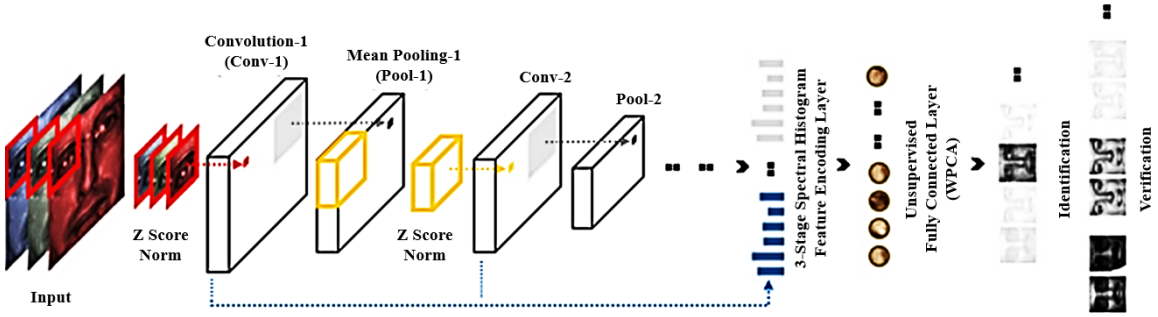

Figure 4.4: Architecture of PCANet+ as described in its original paper [57].

Both PCANet-based models have hyper-parameters that must be optimized. To perform this tuning, three random users were selected from the CASIA-D dataset as a pilot set, and a grid search was applied to find the optimum values. Both PCANet and PCANet+ use a patch size (k) and a number of filters (L). For the patch size, the range of $4 < k < 30$ was chosen, and for the number of filters, integer numbers of L from 1 to 8 were evaluated. In cases where more than one point yielded the maximum accuracy, those closest to the original authors' suggested parameters were chosen. The results yielded values of $k_1 = k_2 = 13$, and $L_1 = 5$ for both PCANet and PCANet+.

4.3.2 UMAPNet: A Novel Feature Learning Architecture

Unlike PCA, UMAP is a *nonlinear*, *graph-based* dimensionality reduction method. Its goal is to create a predefined k-dimensional weighted representation of the original data that minimizes the edgewise cross-entropy between the computed graph and the original data. UMAP was specifically targeted in this work because both the nonlinearity and graph-based components of the technique may be advantageous in gait recognition. For example, UMAP has the capacity to capture nonlinear relationships in pressure data while preserving both local and global structures. Similarly, graph-based approaches may better capture spatial information in pressure-based gait recognition. For these reasons, it was hypothesized that UMAPNet may outperform both PCANet variants.

The architecture of the proposed UMAPNet largely resembles that of PCANet. However, the architecture of PCANet was optimized for the PCA layer, so similar optimization was required for UMAP. The features were generated by the UMAP layer, learned from sampled $k \times k \times d$ patches from the input, normalized by subtracting the patch means from each pixel. Then, $k_1 \times k_2$ patches are uniformly sampled from a single activation map. The UMAP layer was trained until the UMAP function converged to a global minimum.

The UMAP transform results in a compact data representation. This reduces computational complexity during downstream processing (e.g., model training) by projecting the data into a lower-dimensional space without losing critical structures. This reduction is important when working with high-dimensional data, like foot pressure, where the computational cost of training DL models can be prohibitive [66]. Likely due to the fact that UMAP optimizes data into this new feature space, increasing the number of layers did not meaningfully improve performance. After the feature calculation layer, binary quantization and block histogram layers were added to the network. The final output feature vector was then classified using a classifier.

A diagram of the UMAPNet architecture is shown in Figure 4.5.

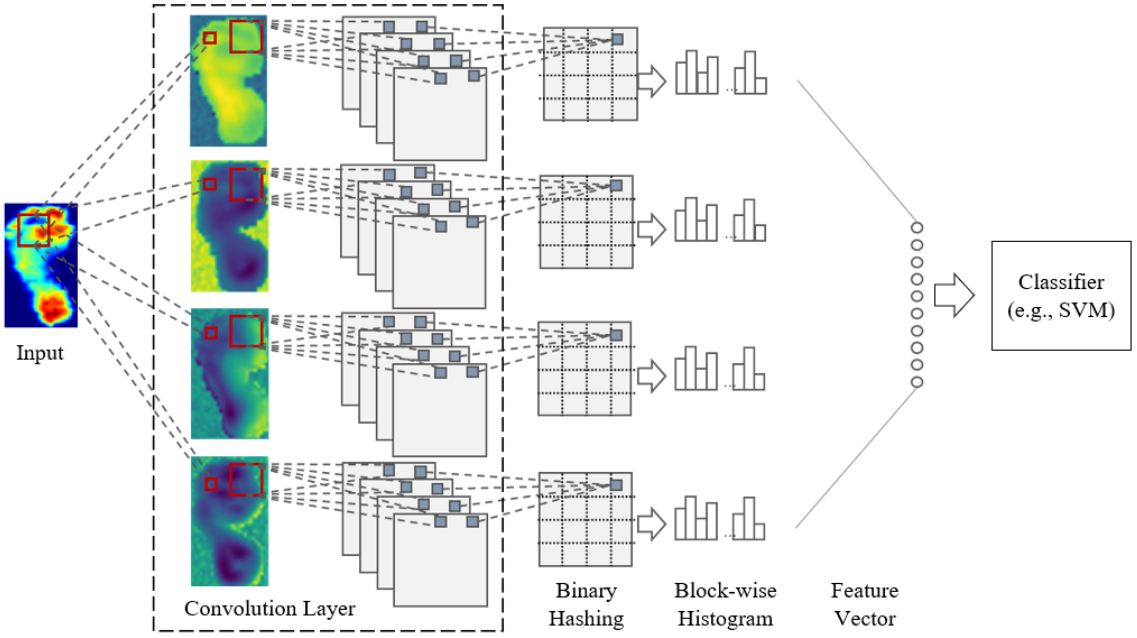

Figure 4.5: Architecture of the proposed UMAPNet. The UMAP layer generates features from sampled patches, followed by binary quantization and block histogram layers. The final output feature vector is classified using a classifier.

The optimization of the UMAPNet method requires a patch size (k) and a number of filter parameters (L). One user from the CASIA-D dataset was selected randomly as a case study to which a grid search was applied to find the optimum values. Odd values of k were tested for $1 < k < 30$. For the number of filters, L, integer numbers from 1 to 10 were chosen. The number of neighbors (n) and minimum distance (d) can also be optimized within the UMAP operation. Five equally spaced values of minimum distance between 0.2 and 1 were evaluated. Similarly, five values from 10 to 100 were used for the number of neighbors with the stepsize of 20. Figure 4.6 shows a visualization of the optimization curves observed from this process. Ultimately, values of $k_1 = k_2 = 11$, $L_1 = 5$ were determined for use with UMAPNet.

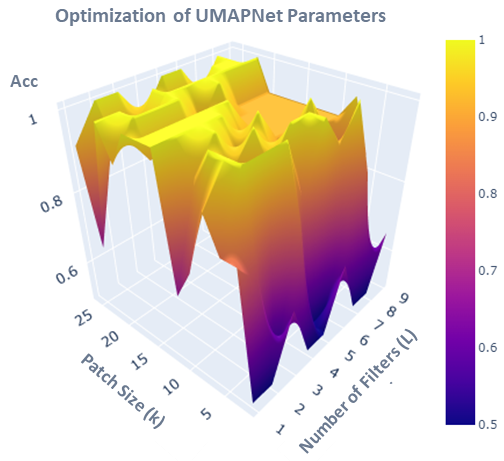

Figure 4.6: A surface plot of the optimization curves for the UMAPNet networks parameters.

4.3.3 Results

Performance results are presented for each FLN model using the P100 input. For context, baseline P100 and PTS performance are also included.

A consistent performance trend was observed across the three FLNs in both the closed-set and open-set problems. In both cases, UMAPNet (89.26% closed-set; 59.27% open-set) attained the highest performance, followed by PCANet+ (85.94% closed-set; 57.91% open-set), and finally PCANet (85.55% closed-set; 54.20% open-set). However, only the closed-set UMAPNet vs PCANet comparison reached statistical significance (p=0.05, d=0.9). Interestingly, while all three FLNs outperformed the baseline P100 in the closed-set (67.47% - $p < 0.03$, $d > 0.48$), none outperformed the open-set baseline performance (53.04% - $p > 0.08$).

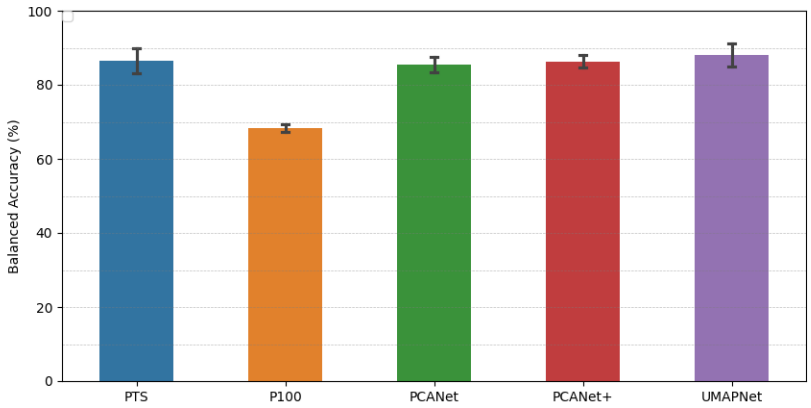

Figure 4.7: The closed-set balanced accuracies of the FLNs (PCANet, PCANet+, UMAPNet) compared to the baseline pixel time series and P100 feature representation for person verification.

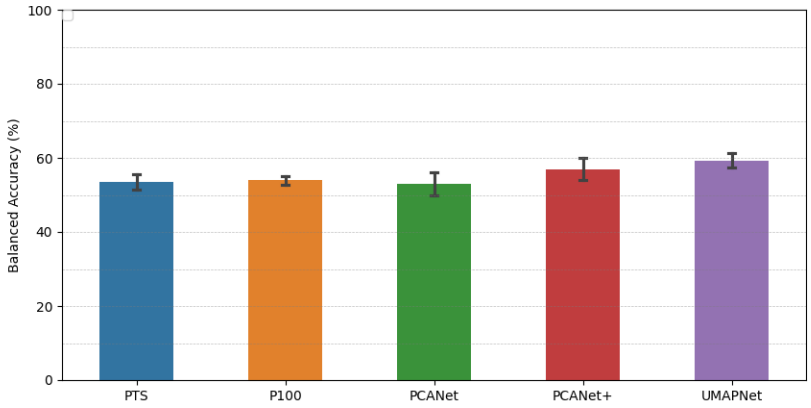

Figure 4.8: The open-set balanced accuracies of the FLNs (PCANet, PCANet+, UMAPNet) compared to the baseline PTS and P100 feature representations for person verification.

4.3.4 Discussion

Although the three FLNs investigated were able to improve gait recognition performance substantially over P100 alone, they ultimately did not meaningfully surpass the baseline performance of the PTS in the close-set problem. Furthermore, although UMAPNet was able to improve performance relative to both P100 and PTS on the open-set problem, it was not competitive with the GRF and COP features described previously (see Figure 4.2). One inherent limitation of these models is their requirement for 2D image formatted input data, which prevented them from drawing on

any temporal gait information in this context. It is likely that FLNs adapted for use with the other features (i.e., which *do* capture temporal gait information) could further improve performance. However, implementing feature learning models for PTS or spatiotemporal features poses architectural challenges - much of the research on feature-learning to date has been in the area of image classification and employs a CNN architecture that necessitates the 2D input shape (e.g., [16, 85, 86]). Despite this challenge, adapting FLNs for use with other forms of input data could be achieved through (at least) two different approaches: a) adapting FLN architecture to operate on other input data structures (e.g., [55]), and b) transforming the other features to comply with the current 2D image format requirements (e.g., converting time series to COP or GRF time series images [108]). Additionally, another solution is modifying the architecture to calculate 1D kernels and apply 1D convolution to extract temporal features[113, 5] or to work with 3D patches for PTS input. These modifications would enable more effective extraction for temporal dynamics.

It is interesting to compare the performance of the three FLNs evaluated. Although performance trends were as expected (i.e., UMAPNet > PCANet+ > PCANet), levels of significance were only detected between UMAPNet and PCANet in the closed-set problem. The marginally higher UMAPNet performance is likely due to its inherent non-linearity, which increases capacity over PCANet. Despite these architectural advantages, UMAPNet performance is still lower than baseline COP, highlighting the importance of temporally informed features in gait recognition. COP's ability to capture both spatial and temporal dynamics allows it to adapt more effectively to the unseen footwear, underscoring the need for future models to integrate spatiotemporal features more explicitly.

4.4 Summary

This chapter presented an analysis of 9 different dimensionality reduction techniques, 4 distinct features, and 3 FLNs. Collectively, these findings suggest that 1) spatial-only features are likely not suitable for open-set gait recognition, 2) conventional dimensionality reduction techniques are unable to achieve open-set generalization, likely due to their tendency to overfit when creating data projections, and 3) spatiotemporal information is likely needed to further improve gait recognition performance. These insights lay the groundwork for the exploration of spatiotemporal deep learning methods in the next chapter.

Chapter 5

Deep Spatiotemporal Learning

Three important insights were identified in Chapter 4, culminating in the hypothesis that data-informed spatiotemporal features will likely be needed to further improve open-set gait recognition performance. This chapter pursues this goal by exploring deep-learning methods that have the capacity to capture complex spatiotemporal patterns from the raw pixel time series signal. This will almost certainly improve closed-set performance, but it was unclear if this may allow the networks to more effectively generalize across different open-set footwear types. The three architectures explored in this chapter differ in their ability to model spatial and temporal aspects of gait:

1. **LSTM**: This model focuses exclusively on temporal dynamics, making it a purely temporal architecture. It processes sequential data to capture long-term temporal dependencies but does not incorporate spatial information about the pressure distribution across different foot regions. As such, it serves as a baseline model focusing solely on the improvements offered by more complex spatiotemporal approaches.

2. **GaitPart**: Originally developed for vision-based gait recognition, this architecture introduces spatial partitioning to divide input data into distinct regions

(e.g., heel, arch, toes). By integrating spatial partitioning with temporal modeling, GaitPart captures localized spatial features and their progression over time, making it a spatiotemporal architecture. However, it does not explicitly model dynamic relationships between different foot regions, which limits its ability to fully capture inter-regional interactions during the gait cycle.

3. **Footpart**: This novel architecture builds on the strengths of GaitPart by capturing both localized spatial patterns and the dynamic relationships between foot regions. FootPart combines detailed spatial partitioning with explicit inter-regional temporal modeling using Dynamic Time Warping (DTW) and a multi-modal LSTM network. This comprehensive spatiotemporal approach allows it to model the full complexity of gait patterns, including cross-regional dependencies, making it the most advanced architecture evaluated in this chapter.

5.1 LSTM: Deep Temporal Learning

The results presented so far have suggested that spatiotemporal deep learning networks may lead to improved gait recognition performance. To this end, an LSTM architecture was explored using the raw PTS as the input with the goal of capturing the full temporal dynamics of gait directly from the data source.

5.1.1 Model Architecture

Unlike earlier deep temporal models that learned by back-propagation through time, LSTMs are designed to capture long-term dependencies in time-series data by using memory cells, gates, and update mechanisms that control the flow of information over time. The general traits of LSTMs include their ability to mitigate the vanishing gradient problem and maintain information over extended sequences, making them effective for temporal pattern recognition.

In the specific LSTM model used here, the architecture consists of three layers of LSTM units, each with a hidden layer size of 512 units. The input size, dictated by the CASIA-D dataset, is 1176 (i.e., the X x Y x time shape of each PTS footstep record). After processing the temporal features through the LSTM layers, the features are passed to a fully connected neural network classifier, which makes the final classification decision.

5.1.2 Hyperparameter Tuning and Model Configuration

Grid search was used to optimize the following four hyperparameters of the LSTM model: a) Number of hidden units, b) Learning rate, c) Dropout rate, and d) Number of LSTM layers. Appropriate value ranges for these hyperparameters were determined through a combination of domain knowledge and empirical evidence. Grid search optimization resulted in the following configuration for the LSTM model: three LSTM layers, a hidden layer size of 512 units, an input size of 1176 (from CASIA-D dataset), a learning rate of 0.001, a mini-batch size of 16, and 20 training epochs.

A cross-entropy loss function was used for optimization, and model parameters were updated using the Adam optimizer. The choice of 20 epochs was empirically determined by monitoring the training and validation loss, ensuring sufficient learning while minimizing the risk of overfitting.

5.2 GaitPart: Adding Spatial Partitioning

Gait is inherently a spatiotemporal phenomenon, combining both spatial and temporal characteristics. Although LSTM is an effective method for capturing temporal dynamics, the LSTM model presented above does not learn any spatial relationships. Consequently, in this section, GaitPart is explored as a model that is designed to integrate both spatial and temporal information. Originally developed for vision-based

gait recognition to overcome challenges with clothing-induced variance, GaitPart partitions the image into vertically separated regions, allowing the model to focus on distinct spatial characteristics while learning to be resilient to changes in others (e.g., the model might learn to ignore the presence of a hat/scarf in the upper-most region of a gait video).

One potential motivation for its modified use in pressure-based gait analysis is its inherent spatiotemporal coupling. For instance, the natural heel-strike-through-toe-off progression of each step results in each spatial region being populated during a specific temporal period. By breaking down the pressure data into distinct foot regions, both localized spatial patterns and the temporal evolution of pressure distribution could be captured. This integration of spatial regions and temporal features was expected to enhance generalization, especially in open-set scenarios.

5.2.1 Model Architecture and Hyperparameters

In adapting this architecture for pressure-based gait recognition, the model partitions foot pressure data into different regions, such as the heel, arch, and toes, to capture spatial patterns. The input dimensions of GaitPart correspond to the spatial resolution of the foot pressure map (for each *pressure frame*[1]) and the temporal sequence of these frames, which is processed to generate a feature vector representing the gait signature. The network's key components include *convolutional layers* for spatial feature extraction from partitioned foot regions, *recurrent layers* for modelling temporal relationships between these regions, and *fully-connected layers* for aggregating the learned spatial-temporal features. Figure 5.1 provides an overview of GaitPart from its original paper.

[1]GaitPart operates on the 3D pixel time series (i.e., PTS) introduced to date. However, given its origins in video-based analysis, this 3D data structure is conceptualized as a "video" composed of a sequence of 2D "frames" (i.e., X and Y dimension) over time (i.e., temporal dimension). Each "frame" of this video depicts a snapshot of the pressure profile exerted on the tile grid at that point in time and is therefore referred to as a pressure frame.

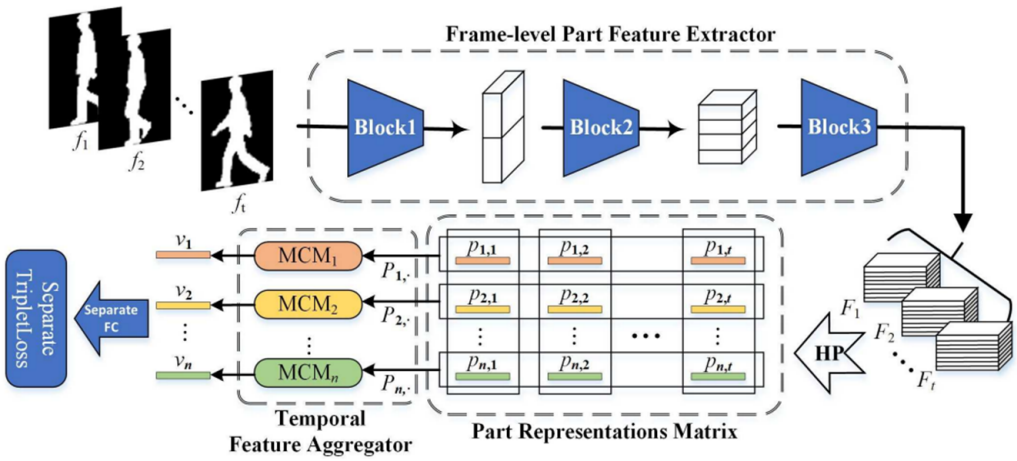

Figure 5.1: Architecture of the GaitPart model for vision-based gait recognition [31].

As with the LSTM model, hyperparameter tuning was necessary to optimize the performance of GaitPart. A grid search approach was employed to optimize the following key hyperparameters: *number of convolutional layers*, which determines how many layers the network uses to extract spatial features, *kernel sizes and strides*, which are adjusted to capture the characteristics of the pressure data better, *number of recurrent layers*, which model the temporal evolution of foot pressure data, and *learning rate and batch size*, which control how the network learns from the data during training and was tuned to maximize performance.

Appropriate value ranges for these hyperparameters were established based on domain knowledge and empirical evidence. Grid search optimization resulted in the configuration of GaitPart as shown in Table 5.1.

Table 5.1: Hyperparameters of GaitPart Implementation

Recurrent layers	2
Input size	1176 *
Learning rate	0.001
Batch size	16
Training epochs	20

* The spatial dimensions of the pressure frame

5.3 FootPart: Region-Based Spatiotemporal Learning

Although previous models such as GaitPart have shown that extracting features from distinct regions of the foot can enhance performance, no existing model fully captured the relationships between these regions across the gait cycle. To address this, FootPart was developed as a novel approach that integrates detailed spatial partitioning with explicit multivariate temporal modelling through LSTM and Dynamic Time Warping (DTW). By not only capturing spatial details within each region but also modelling the inter-regional dynamics, FootPart is designed to more fully explain the spatiotemporal evolution of the footstep. This added dimension of analyzing cross-regional temporal dependencies sets FootPart apart as a custom-built model for pressure-based gait recognition performance.

5.3.1 Model Architecture

Footpart employs a four-stage architecture, as illustrated in Figure 5.2. In the first stage, the foot is divided into distinct regions to capture detailed spatial pressure patterns. The second stage characterizes relationships between these regions using DTW to capture the temporal dependencies across regions. In the third stage, these relationship-based features are aggregated with each region's pressure time series (PTS) data in the Feature Aggregator Module (FAM). In the final stage, each region and its relationship to other regions is analyzed independently using a multimodal LSTM network to model spatiotemporal dynamics.

A detailed description of each stage of the Footpart pipeline is presented below.

Stage 1: Partitioning the footstep into spatial regions

As shown in Fig. 2.3, a sequence of pressure data formatted as a 3D matrix containing pressure information from 100 frames is fed into FootPart. The first module of the

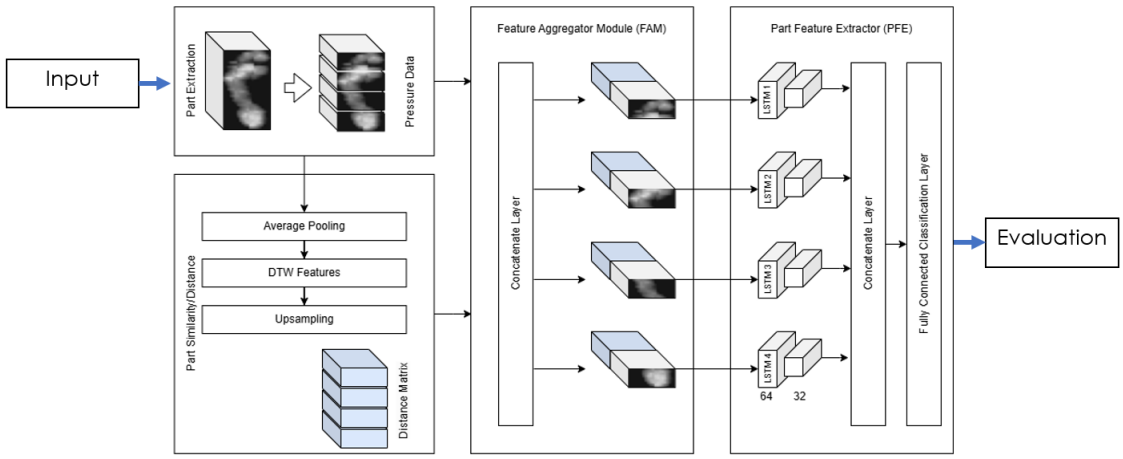

Figure 5.2: Proposed framework for Footpart, a novel deep network with explicit modelling of spatiotemporal dynamics between foot partitions.

pipeline is the Part Extractor, which slices input pressure data into parts.

A grid search of partition configurations ranging from 1×1 up to 7×7 yielded the best performance when using a partition size of 1×4. Using this partitioning, the input pressure 3D matrix was divided along the direction of gait into four main parts. This corroborates previous studies of pressure-based gait analysis which have also divided the foot into four segments (toes, forefoot, midfoot, and rearfoot) and have shown that the efficiency of evaluation of gait analysis was improved [84].

Stage 2: Characterizing Between-Region Relationships with DTW-based Similarity Score

DTW has been shown to be effective in assessing part similarity for foot pressure verification [72, 44]. However, prior studies did not leverage DTW in the context of explicitly modelling cross-regional temporal dependencies within a unified spatiotemporal framework. Unlike previous efforts, FootPart combines DTW-based temporal relationship modelling across foot regions with localized feature extraction, capturing the complex, multivariate interactions that occur throughout the gait cycle. In this stage, DTW is used to quantify temporal relationships between each partitioned foot region (by maintaining the warping path) rather than the traditional similarity

(defined by the summary warping cost). This method captures how patterns between different foot regions evolve over time.

This process is completed as follows. First, each region's PTS is converted to a 1D representation using average pooling. DTW then measures the temporal relationship between each pair of these sequences, accounting for potential variations in speed and duration through its warping abilities. A DTW similarity matrix is constructed, capturing the relationship between all possible pairs of regions within a 1x4 patch arrangement. These DTW matrix values are then upsampled proportionally to fit into a broader 2D matrix using nearest neighbor interpolation, which assigns the value of the closest neighboring pixel to each pixel in the resized image.

Stage 3: Consolidating Region-Specific Information with a Feature Aggregation Module

The Feature Aggregator Module (FAM) concatenates each region's temporal and spatial feature vectors, along with the corresponding DTW similarity matrix, into a Part Representation Matrix (PR-Matrix). Each PR matrix represents the spatiotemporal gait signature of a single region across time and its relationship with other parts.

Stage 4: Extraction of Features from Each Foot Region

Once each aggregated PR-Matrix is created, features are extracted from each partitioned region and its inter-regional relationships to capture localized gait information. This process is performed by the Part Feature Extractor (PFE), as shown in Figure 5.2. Within PFE, each PR-Matrix is processed by an LSTM network, which extracts both spatial features (e.g., peak pressure points, pressure distribution) and temporal features (e.g., duration of pressure contact) for each partitioned region. The final features extracted from each LSTM head for the different regions are concatenated and then passed through a fully connected layer (FCN) for classification. During training,

the LSTM and FCN are tuned to identify features that remain resilient across the various factors they are provided (e.g. the different footwear types including during training).

5.3.2 Model Training Parameters

The optimization process and the configuration of the architecture were the same as in Section 5.1. A learning rate of 0.001 was used, and data were processed in mini-batches of 16 samples. The training process was conducted over 20 epochs, allowing the model to learn from the dataset's nuances. Cross-entropy loss was used for optimization, with model parameters updated through the Adam optimizer. This training setup enabled FootPart to balance learning efficiency with generalization performance effectively.

5.4 Results

Figures 5.3 and 5.4 show a summary of the closed-set and open-set performances of the three models introduced in this chapter. The baseline raw PTS (NDR-PTS), the COP feature with no dimensionality reduction (i.e. the best-performing open-set feature/dimensionality reduction combination), and UMAPNet (i.e., the top-performing FLN) performance are also provided for context.

In the closed-set problem (Figure 5.3), all three DL models significantly outperformed the baseline PTS, COP, and FLN networks ($p < 0.01$, $d > 0.92$). Although an increasing performance trend was observed across the three proposed DL models (i.e., LSTM (94.96%, GaitPart (96.15%), FootPart (98.89%)), ultimately no significant differences were detected between these models ($p > 0.12$).

A similar trend was observed in the open-set problem (Figure 5.4), again, with all three DL models outperforming baseline PTS and FLN networks ($p < 0.009$, $d >$

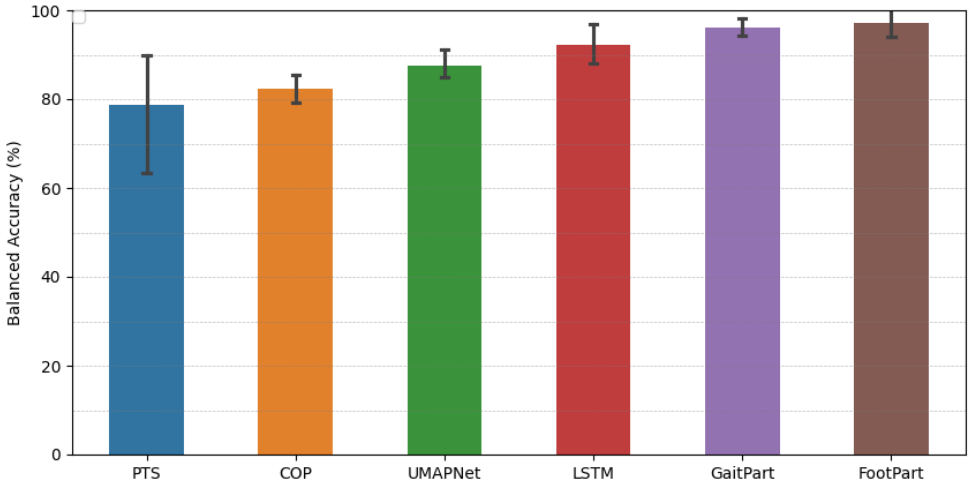

Figure 5.3: Balanced accuracies of Baseline, UMAPNet, GaitPart, and FootPart for the closed-set person verification task using the CASIA-D dataset.

1.2), but only GaitPart and FootPart outperformed COP ($p < 0.04$, $d > 1.17$). Interestingly, in this context, significant performance differences *were* detected between all three DL models, with FootPart being the top performer (73.65% - $p < 0.04$, $d > 1.00$), followed by GaitPart (71.64% - $p < 0.01$, $d > 1.50$), and finally LSTM (65.72%).

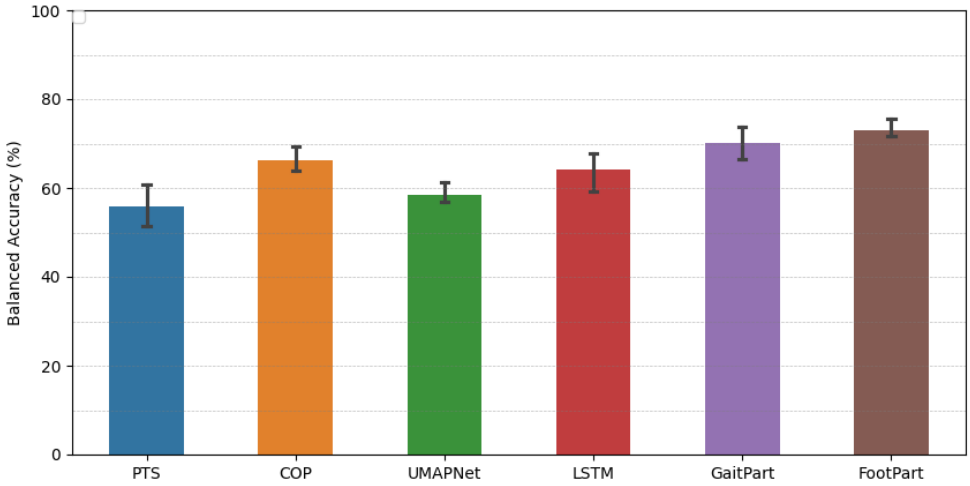

Figure 5.4: Balanced accuracies of Baseline, UMAPNet, GaitPart, and FootPart for open-set person verification task using the CASIA-D dataset.

5.5 Discussion

This chapter evaluated deep spatiotemporal learning techniques for gait recognition, emphasizing the benefit of employing spatiotemporal features. Discussion topics are organized around the three models introduced: LSTM, GaitPart, and FootPart.

5.5.1 LSTM for Temporal Feature Extraction

LSTM was included in this analysis as a representative DL baseline upon which to compare the other, gait-specific, DL models. Results demonstrate that LSTM was effective in its own right, outperforming all models presented in Chapter 4 in both closed-set scenarios (94.96%). This is likely because of the model's capacity to capture complex temporal dependencies in the gait patterns. In the open-set case (65.73%) LSTM outperformed PTS and UMAPNet, but not COP. This speaks to the ability of COP to capture robust temporal information and suggests that LSTM may require more data from each user in order to learn how to fully extract such information reliability.

5.5.2 GaitPart for Spatiotemporal Learning

Although the LSTM captures temporal dynamics, it does not consider the spatial aspects of gait. GaitPart was introduced to integrate both spatial and temporal features, addressing this limitation of LSTM. The results confirm the effectiveness of this approach, with GaitPart achieving a 96.15% accuracy in closed-set verification (Figure 5.3), outperforming LSTM by 5.1%. This performance boost underscores the importance of both spatial and temporal features for gait recognition.

The results also highlight GaitPart's improved ability to handle open-set scenarios, where it achieved 71.64% accuracy - a 5.9% improvement over LSTM (See Figure 5.4). This validates the hypothesis that spatiotemporal modelling can im-

prove robustness to variations, such as unseen footwear, by focusing on both the spatial configuration of pressure points and their temporal progression during the gait cycle. However, as GaitPart was designed for resilience to confounds in specific partitions in vision-based gait recognition, it was not intended to capture detailed inter-region relationships. This motivated the development of FootPart, which takes the spatiotemporal approach further.

5.5.3 FootPart for Enhanced Spatiotemporal Learning

FootPart addresses two key challenges left unresolved by GaitPart: 1) capturing the temporal evolution of pressure patterns within each foot region, and 2) modelling the relationships between different foot regions using DTW. The results demonstrate the effectiveness of this approach, as FootPart outperformed GaitPart in both closed-set (i.e., 96.15% vs. GaitPart's 98.89%) and open-set (i.e., 73.56% vs. GaitPart's 71.64%) performance. This improvement was likely a result of the two key architectural changes - the inclusion of*two-dimensional segmentation* and *DTW analysis* allowed FootPart to capture finer details of foot pressure distribution and their interactions, addressing the limitations of GaitPart's one-dimensional partitioning. This comprehensive modelling of spatiotemporal dynamics positions FootPart as a more robust solution for gait recognition, with an improved ability to handle variability caused by footwear.

5.6 Summary

The results from this chapter strongly support the initial motivation to explore deep learning methods for spatiotemporal feature extraction. The progression from LSTM's temporal focus to GaitPart's spatiotemporal integration, and finally to FootPart's detailed regional analysis, demonstrates a clear trend of improved performance

as more sophisticated modelling techniques were employed. While each method brings its own strengths, the overall findings suggest that robust gait recognition systems benefit significantly from capturing both spatial and temporal dynamics and interactions, with FootPart offering the most comprehensive approach. Further extensions to a verification system built on this model, such as a majority voting mechanism applied across multiple steps, could make it viable for high-security, real-world implementations.

In the next chapter, we further evaluate the robustness and scalability of the proposed models by testing their generalization to a completely different dataset and exploring their performance in the context of the identification task.

Chapter 6

Generalization of the Findings

The primary motivation for this chapter is to evaluate whether the proposed models can generalize effectively across datasets, ensuring robustness beyond a single controlled dataset like CASIA-D. While CASIA-D was critical for developing and testing the architectures under controlled conditions, real-world biometric systems often encounter scenarios where training and testing data are drawn from completely different sources. The inclusion of the UNB dataset provides an opportunity to test the scalability and adaptability of the proposed models in these more diverse and challenging scenarios.

The UNB dataset introduces unique challenges not present in CASIA-D, such as more varied demographics, footwear types, and experimental setups. While UNB maintains control over footwear by standardizing specific shoe types in its protocol, it also introduces variability through the inclusion of different footwear categories. These differences make it an essential complement to CASIA-D, allowing the models' robustness and scalability to be tested under more real-world variations. Additionally, while both datasets are controlled in indoor laboratory environments, the differences in collection methods, sensor platforms, and participant diversity provide a robust test for cross-dataset generalization. This cross-dataset evaluation is critical for demon-

strating the practical applicability of the proposed methods in diverse and unseen conditions.

In addition to testing cross-dataset generalization, this chapter also transitions from verification to identification tasks. While verification, or authentication, determines if an input matches a claimed identity, identification involves assigning an input to a specific individual from a larger group, making it inherently more complex and challenging. By addressing both of these tasks, this chapter highlights the versatility of the proposed models and their potential to handle both use cases in practical applications.

6.1 Generalization to the Novel UNB Dataset

Generalizing machine learning models to new and unseen datasets is important for validating their robustness and practical applicability. For this reason, the models were tested on the newly collected (and soon-to-be-released) UNB dataset. These data were collected at a different site, in a different country, by a different research team, using different sensing technology, and composed of different participants. This represented a completely unseen generalization challenge for the models, which had been developed without knowledge of this dataset.

All model configurations were held consistent with those used with CASIA-d, except for the LSTM architecture, which had to be adapted slightly for compatibility with the dimensionality of the UNB dataset records (i.e., 2304 dimensions per footstep record). Three LSTM layers were employed, each with a hidden layer size of 512 units, to classify the input data into one of two classes for verification.

6.1.1 Results

Surprisingly consistent performance trends were observed across the two datasets for both closed-set (Figure 6.1) and open-set (Figure 6.4) scenarios. For the closed-set problem, performance was consistently higher on CASIA-D than UNB (i.e., in all but one model), however, no significant differences were detected between the datasets. Interestingly, the opposite pattern was observed in the open-set comparison (i.e., performance was higher on the UNB dataset for all but 2 models). Although some significant differences were detected between specific algorithm-dataset pairings in the open-set case (i.e., for NDR-PTS, $p = 0.001$, $d = 0.33$; for UMAPNet, $p = 0.03$, $d = 0.58$), the model ranking was consistent across the datasets. Collectively, these results demonstrate the generalizability of the models introduced in this work to the novel UNB dataset.

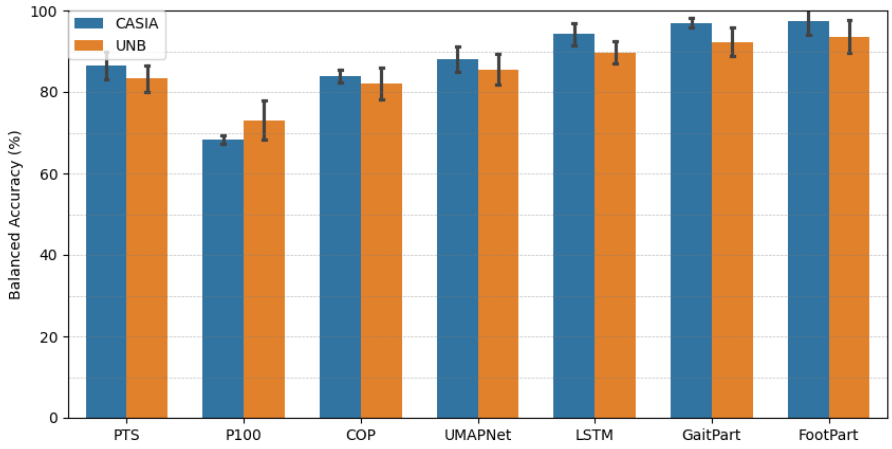

Figure 6.1: Comparison of the balanced accuracies of the various models for the closed-set verification task when using the CASIA-D and UNB datasets.

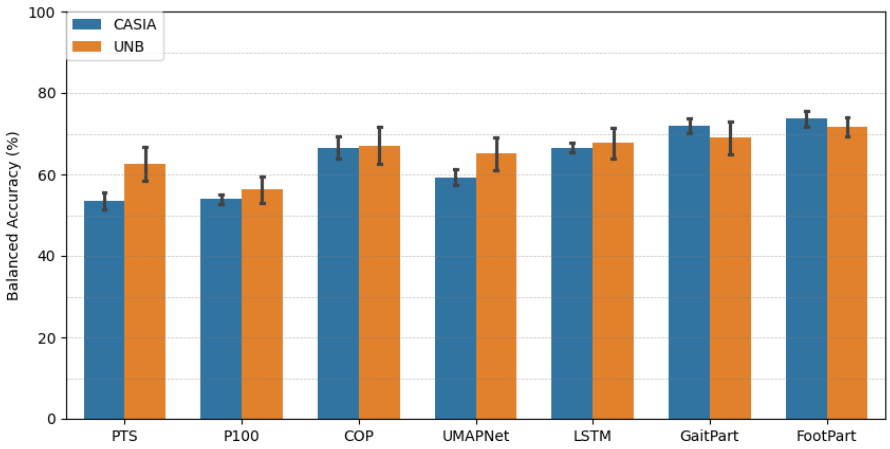

Figure 6.2: Comparison of the balanced accuracies of the various models for the open-set verification task when using the CASIA-D and UNB datasets.

6.1.2 Discussion

The performance of the developed models on the UNB dataset demonstrates their strong generalization capabilities, in both the closed-set and open-set scenarios. In the closed-set evaluation, the models achieved accuracy levels on the UNB dataset that were consistent with their performance on the CASIA-D dataset. Specifically, the FootPart model's accuracy of 96.14% on the UNB dataset aligns closely with its results on CASIA-D. It is natural that some differences in absolute performance may be observed due to the considerable differences between the datasets and natural variations between users. However, the strong consistency in trends across algorithms speaks to both the generalizability of the developed models and the rigor of the evaluations conducted in this work.

This finding extended to the open-set performance evaluation, where models must handle new and unseen data. In this open-set scenario, the balanced accuracy dropped slightly to 72.00% on the UNB dataset, compared to 73.57% on CASIA-D. Despite this marginal difference, the trends across models remained remarkably consistent, indicating that the improvements to the proposed models effectively capture essential features across domains in a generalizable way.

These results validate the robustness of the analytical approach used in this work and the models themselves. Additionally, they reinforce the quality and reliability of the new data being collected at UNB, suggesting that it can serve as a valuable resource for advancing research in gait analysis and biometric verification. The performance of these models across the CASIA-D and UNB datasets demonstrates their ability to generalize to different sensor configurations and users, which is promising for real-world applications where variability in sensor setups may occur.

6.2 Generalizing to the Identification Task

Verification and identification tasks, while both widely used in biometric applications, serve different purposes and present distinct challenges. Verification confirms if a person is who they claim to be, typically by comparing a presented identity to a stored reference. In contrast, identification requires determining an individual's identity from a pool of possible candidates, making it a more complex and demanding task (e.g. the problem scales with the number of possible users). The models developed in this work were initially tailored for verification, optimizing their ability to match or reject a claimed identity.

However, to fully assess the robustness and versatility of these models, this section explores how well they generalize to the identification task, where the complexity increases. The identification performance of the models was assessed using both the CASIA-D and UNB datasets. By doing so, we aimed to understand whether the insights and optimizations derived from the verification context can extend to identification, thereby broadening the potential applications of these models.

6.2.1 Results

Similar performance trends were observed across both datasets for the closed-set (Figure 6.3) and open-set (Figure 6.4) scenarios with only minor differences in performance between the datasets. For closed-set identification, FootPart achieved an accuracy of 95.14% on the UNB dataset and 96.61% on CASIA-D, showing only a slight drop from its verification performance (96.14% and 98.89%, respectively). Similarly, other models showed comparable performance, with GaitPart reaching 91.13% on UNB and 93.17% on CASIA-D.

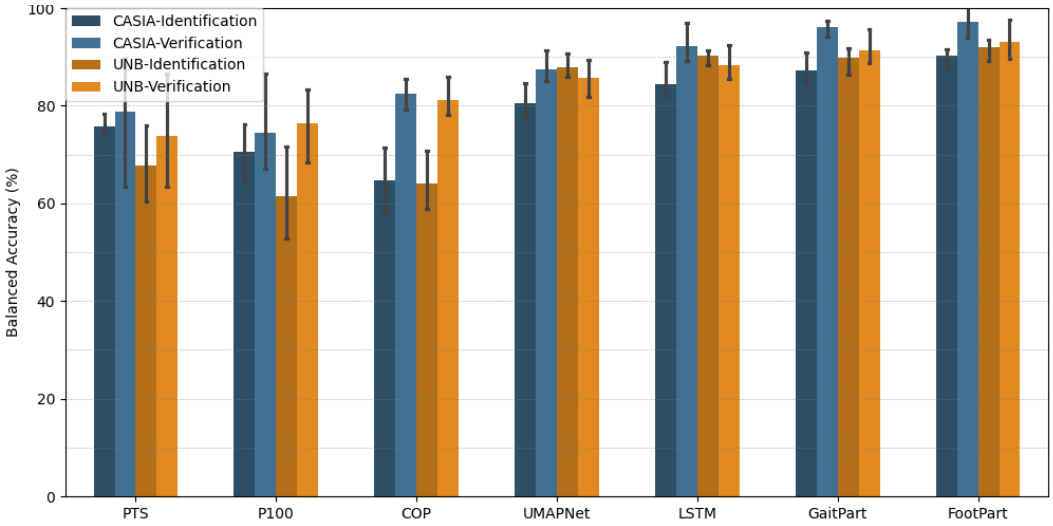

Figure 6.3: Balanced accuracies of Baseline, UMAPNet, GaitPart, and FootPart for the closed-set person identification using the CASIA-D and UNB datasets.

In the more challenging open-set scenario, despite overall much lower accuracies, FootPart still performed best, with a balanced accuracy of 52.57% on UNB and 50.44% on CASIA-D. GaitPart followed closely with 51.28% on UNB and 49.83% on CASIA-D. While FootPart's performance did not differ significantly from GaitPart ($p = 0.168$, $d = 0.894$), both models demonstrated a substantial advantage over other methods ($p < 0.0302$, $d > 1.335$). Notably, despite the lower open-set identification performance, the relative ranking of models remained consistent across datasets extending from verification to identification.

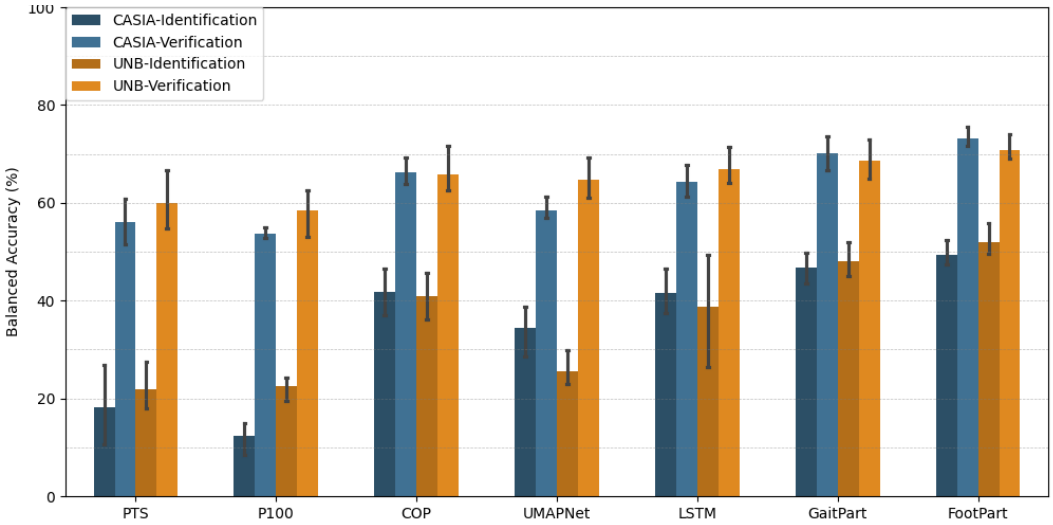

Figure 6.4: Balanced accuracies of Baseline, UMAPNet, GaitPart, and FootPart for the open-set person identification task using the CASIA-D and UNB datasets.

6.2.2 Discussion

A key observation of these results is the notable performance drop between the verification and identification tasks, particularly in open-set scenarios. This is expected because identification is inherently more challenging, requiring the system to differentiate between multiple potential candidates rather than confirming a single identity. Consistent with previously reported results [94, 71], the open-set identification task further introduces unseen conditions, such as different footwear, that add variability and increase task complexity.

The observed significant differences in performance between the three deep learning models ($p < 0.03$, $d > 0.74$) suggest that identification tasks may be more sensitive to model architecture. Models like FootPart, which incorporate both detailed spatial partitioning and temporal modelling using advanced methods like DTW, may improve identification because they capture fine-grained details and cross-region temporal dependencies more effectively than models focused on simpler spatiotemporal interactions (e.g., GaitPart). The increased complexity of identification tasks appears to better reveal these differences, particularly when dealing with diverse, unseen con-

ditions like the open-set problem. Further investigation into why certain models degrade less than others in these tasks could shed light on crucial factors for robustness in real-world deployments.

The results presented in this chapter underscore the generalization capacity of the proposed models across both verification and identification tasks. The proposed FootPart network, in particular, consistently outperformed baseline models across both tasks, demonstrating that optimizations designed for verification, such as capturing enhanced spatiotemporal features, also contribute to strong identification performance. Although identification accuracy was slightly lower, the model's performance remained competitive, suggesting that the core feature extraction methods are robust across various biometric tasks. Future work should explore how these results generalize when large sets of users are employed, as this can greatly affect the identification performance [48].

The consistency of performance across the CASIA-D and UNB datasets further supports the adaptability of the developed models to different sensor setups and data characteristics. The ability of the FootPart network to consistently outperform the other models, even when faced with increased variability in footwear and unseen conditions in the open-set problem, highlights its potential for real-world applications. In environments where variability in data collection methods, such as sensor types or user environments, is common, the robustness of FootPart could be a significant advantage over more conventional approaches.

Studies in vision-based gait recognition have highlighted the importance of capturing both spatial and temporal dynamics for robust performance, particularly when dealing with occlusions, clothing, and changes in viewpoint [99, 16, 30]. Given that FootPart outperformed the vision-based gait recognition inspired GaitPart, there may be value in evaluating its concepts of spatial partitioning with inter-regional temporal modelling with other modalities. Although there may be fewer such confounds

in pressure-based gait recognition, the explicit modelling of relationships between partitions helped FootPart reduce the impact of unseen footwear. It is worth future consideration whether this could help in other modalities of gait analysis.

Finally, although vision-based systems may still outperform pressure-based methods in certain controlled scenarios, the nature of pressure-based systems may make them particularly valuable for specific applications, such as indoor environments or where visual data is difficult to capture. The ability to incorporate multiple modalities, such as pressure and vision, could further enhance overall system robustness.

6.3 Summary

In this chapter, the generalization capabilities of the proposed models were assessed by evaluating their performance on a novel dataset (UNB) and a more complex biometric recognition task (identification). The results demonstrate that the proposed FootPart network consistently outperformed baseline models in both verification and identification tasks across both datasets. Although a performance drop was observed when transitioning from verification to identification (particularly in open-set conditions), as expected, the relative performance of the models held. These findings suggest that the spatiotemporal feature extraction methods employed, particularly in FootPart, are effective across different sensor configurations and unseen conditions.

Chapter 7

Discussion and Future Works

The research described in this dissertation had the goal of improving the performance of pressure-based gait recognition, with a particular focus on addressing the challenge of footwear variability. By exploring various machine learning and deep learning models, the thesis aimed not only to improve the closed-set recognition case, where models are tested using only known footwear (seen during training), but also the open-set case, wherein models must verify or identify previously unseen footwear. This chapter synthesizes the key findings from previous chapters, discussing the broader implications, contributions, limitations, and future directions for research and development. In doing so, it seeks to provide an understanding of how this work contributes to the growing field of biometric systems, and what potential lies ahead.

7.1 Implications of Results

The results presented in this research demonstrate that deep spatiotemporal learning, particularly through the use of purpose-built models like FootPart, can significantly improve gait recognition performance, even in challenging open-set scenarios. This suggests that the introduction of spatiotemporal features — capturing both curated pressure distribution and temporal dynamics — allows biometric systems to better

handle variability, such as changes in footwear. The implications extend beyond the security context, suggesting potential applications in healthcare (e.g., tracking mobility changes) and ergonomics, where understanding detailed gait patterns could support injury prevention or rehabilitation [88, 20].

7.2 Summary of the Contributions

This research identified several significant contributions to the field of gait recognition. It first establishes a baseline for performance using Conventional feature extraction methods, as discussed in Chapter 4, highlighting challenges related to the variability in footwear and limited training data. It was found that temporal features provided better generalization than spatial features when faced with changes in footwear. Furthermore, despite improving the close-set case, several dimensionality reduction techniques failed to sustain this benefit when tasked with generalizing to the open-set case. Advanced FLNs, such as the novel UMAPNet architecture, were introduced, but due in part to the use of spatial input features, largely suffered from similar challenges.

Chapter 5 further expanded on these contributions by exploring deep learning (DL) models like LSTMs, which were able to learn the temporal dynamics of gait. The further segmentation of the foot into spatial regions improved robustness by isolating the effects of footwear variations. The adaptation of vision-based models, such as GaitPart, demonstrated the value of leveraging successful spatiotemporal techniques from vision-based gait recognition. The introduction of the novel multi-modal FootPart network stands out as a key contribution. By combining local spatial information with temporal dynamics, FootPart significantly improved gait recognition accuracy, particularly in the challenging open-set scenario. The integration of Dynamic Time Warping (DTW) to explicitly model the temporal relationships between

parts of the foot was particularly effective in managing variations caused by footwear. Comparative analysis throughout the thesis revealed that while traditional methods establish a strong baseline, the DL approaches yielded superior performance, even though each participant had only modest amounts of training data. These contributions provide a robust framework for advancing pressure-based gait recognition, laying a foundation for future work in both verification and identification tasks.

This work also evaluated the performance of these models on a completely new dataset, showing remarkably consistent performance. Furthermore, similar consistency was shown when generalizing to the more challenging identification task. Although the absolute performance of the open-set identification task was substantially lower (as is known in the literature [37, 71, 94]), the relative performance of the various approaches persisted. This robustness to dataset and task lends credibility to the models, the methods employed in this work, and the results, and builds confidence in their real-world applicability.

7.3 Limitations and Future Works

Despite the advancements in pressure-based gait recognition provided in this work, several challenges remain. The FootPart model, while highly accurate, involves computationally intensive processes like DTW and deep recurrent networks, limiting real-time applicability. Future works should focus on optimizing model efficiency and reducing reliance on large labeled datasets by exploring semi-supervised and self-supervised learning approaches.

Generalization to unseen footwear and varying walking conditions remains challenging. While models like GaitPart and FootPart demonstrated improved robustness in open-set tests, further validation in diverse, real-world environments is essential. Additionally, addressing other covariates like walking speed and loading conditions

could improve model resilience, as early findings suggest DL-based models are more adaptive to these factors compared to traditional methods. While the fusion of conventional features (e.g., P100, GRF, and COP - which performed well throughout) into such deep architectures, could enhance recognition performance, caution must be taken to maintain this generalization.

Finally, it is worth noting that gait patterns are not confined to a single footstep (as explored in this work) but evolve over a series of steps. Capturing this inter-stride information may help the model to better understand the cyclic nature of gait, where variations in stride length, cadence, and/or step force over time could be indicative of a specific individual. However, although researchers have explored some of these patterns [35, 89], there are concerns that they may not generalize well when confounds such as footwear, walking speed, or turning are introduced [59]. Nevertheless, the concepts proposed in this work could be extended to this inter-stride context and provide similar benefits.

7.4 Conclusion

In conclusion, this research demonstrates that spatiotemporal deep learning models, specifically the novel FootPart architecture, offer a significant improvement in pressure-based gait recognition. The ability to generalize across both verification and identification tasks, even in open-set conditions, and across datasets highlights the potential for this work to be applied in real-world biometric security systems. The contributions of this thesis lay a strong foundation for future exploration into more efficient, scalable, and robust biometric systems, with a range of applications across security, healthcare, and beyond.

Bibliography

[1] Mohammed Abuhamad, Ahmed Abusnaina, DaeHun Nyang, and David Mohaisen. "Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey". In: *IEEE Internet of Things Journal* 8.1 (2020), pp. 65–84.

[2] Ilaria Arcolin, Marco Godi, and Stefano Corna. "Which model best assesses gait in healthy elderly? A confirmatory factor analysis of existing conceptual gait models". In: *Gait & Posture* 91 (2022), pp. 94–98.

[3] Sagor Chandro Bakchy, Md Rabiul Islam, M Rasel Mahmud, and Faisal Imran. "Human gait analysis using gait energy image". In: *arXiv preprintarXiv:2203.09549* (2022).

[4] Mohammadreza Hazhirpasand Barkadehi, Mehrbaksh Nilashi, Ali Zakeri Fardi, and Sarminah Samad. "Authentication systems: A literature review and classification". In: *Telematics and Informatics* 35.5 (2018), pp. 1491–1511.

[5] Shaorun Bian, Zhenshu Wang, Wenhao Song, and Xinhui Zhou. "Feature extraction and classification of time-varying power load characteristics based on PCANet and CNN+ Bi-LSTM algorithms". In: *Electric Power Systems Research* 217 (2023), p. 109149.

[6] Vladislava Bobić, Milica urić-Jovičić, Nataša Dragaševic-Mišković, Vladimir S Kostić, and Goran Kvaščev. "Comparison of Two Deep Learning Models

for the Recognition of Parkinson's Disease Gait Patterns". In: *2023 10th International Conference on Electrical, Electronic and Computing Engineering (IcETRAN)*. East Sarajevo, Bosnia and Herzegovina, 2023, pp. 1–5.

[7] Nidhi M. Bora, Gajendra V. Molke, and Hemant Munot. "Understanding Human Gait: A survey of Traits for Biometrics and Biomedical Applications". In: *2015 International Conference on Energy Systems and Applications*. Pune, India, 2015, pp. 723–728.

[8] Joan Bruna and Stephane Mallat. "Invariant Scattering Convolution Networks". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.8 (2013), pp. 1872–1886.

[9] Joy Buolamwini and Timnit Gebru. "Gender shades: Intersectional accuracy disparities in commercial gender classification". In: *Conference on fairness, accountability and transparency*. New York, USA, 2018, pp. 77–91.

[10] Entesar T Burges, Zakariya A Oraibi, and Ali Wali. "Gait Recognition Using Hybrid LSTM-CNN Deep Neural Networks". In: *Journal of Image and Graphics* 12.2 (2024).

[11] Rafael Caldas, Diego Rátiva, and Fernando Buarque de Lima Neto. "Clustering of Self-Organizing Maps as a means to support gait kinematics analysis and symmetry evaluation". In: *Medical Engineering and Physics* 62 (2018), pp. 46–52.

[12] Tsung-Han Chan, Kui Jia, Shenghua Gao, Jiwen Lu, Zinan Zeng, and Yi Ma. "PCANet: A Simple Deep Learning Baseline for Image Classification?" In: *IEEE Transactions on Image Processing* 24.12 (2015), pp. 5017–5032.

[13] Hanqing Chao, Yiwei He, Junping Zhang, and Jianfeng Feng. "Gaitset: Regarding gait as a set for cross-view gait recognition". In: *Proceedings of the*

AAAI Conference on Artificial Intelligence. Vol. 33. 01. Honolulu, USA, 2019, pp. 8126–8133.

[14] Mrunali Chaure and Amit Sahu. "Biometrics as a Security System-An Overview". In: 2015. URL: https://api.semanticscholar.org/CorpusID:15014688.

[15] Biao Chen, Chaoyang Chen, Jie Hu, Zain Sayeed, Jin Qi, Hussein F Darwiche, Bryan E Little, Shenna Lou, Muhammad Darwish, Christopher Foote, et al. "Computer vision and machine learning-based gait pattern recognition for flat fall prediction". In: *Sensors* 22.20 (2022), p. 7960.

[16] R Chhatrala and D Jadhav. "Gait recognition based on curvelet transform and PCANet". In: *Pattern Recognition and Image Analysis* 27 (2017), pp. 525–531.

[17] J. C. Chiou, Ching Yen, and Fei Wu. "Gait-based Biometric System Using Pressure Sensing Mats and Machine Learning Algorithms". In: *2023 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* (2023), pp. 1–6.

[18] Tzu-Chun Chiu, Tzung-Shi Chen, and Jing-Mei Lin. "Human gait recognition using LiDAR and deep learning technologies". In: *2022 IEEE International Conference on Consumer Electronics-Taiwan*. IEEE. Taipei, Taiwan, 2022, pp. 43–44.

[19] Ahnryul Choi, Hyunwoo Jung, Ki Young Lee, Sangsik Lee, and Joung Hwan Mun. "Machine learning approach to predict center of pressure trajectories in a complete gait cycle: a feedforward neural network vs. LSTM network". In: *Medical and Biological Engineering and Computing* 57 (2019), pp. 2693–2703.

[20] Hsi-Chiang Chou. "Remote monitoring and control smart floor for detecting falls and wandering patterns in people with dementia". In: *Journal of Intelligent & Fuzzy Systems* 36.2 (2019), pp. 1323–1331.

[21] Matthew C Cieslak, Ann M Castelfranco, Vittoria Roncalli, Petra H Lenz, and Daniel K Hartline. "t-Distributed Stochastic Neighbor Embedding (t-SNE): A tool for eco-physiological transcriptomic analysis". In: *Marine genomics* 51 (2020), p. 100723.

[22] Jacob Cohen. *Statistical power analysis for the behavioral sciences*. routledge, 2013.

[23] Patrick C Connor. "Comparing and Combining Underfoot Pressure Features for Shod and Unshod Gait Biometrics". In: *2015 IEEE International Symposium on Technologies for Homeland Security (HST)*. 2015, pp. 1–7.

[24] Vincenzo Conti, Carmelo Militello, and Salvatore Vitabile. "Biometric authentication overview: a fingerprint recognition sensor description". In: *Journal of Biosensors and Bioelectronics* 2 (2017).

[25] Kris Cuppens, Tom Saey, Eveline De Raeve, Inge den Herrewegen, Ingrid Knippels, Mario Broeckx, and Louis Peeraer. "Eigenfoot Decomposition of Plantar Pressure Images and Case Study of Feature Prediction of Two Modalities". In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Milan, Italy, 2015, pp. 767–770.

[26] Sonia Das, Sukadev Meher, and Upendra Kumar. "Analysis of Spatio Temporal Dynamic Patterns of Gait for Recognition". In: 2017.

[27] Bin Deng, Jiangtao Luo, Tianshi Gao, Zijian Cui, and Jiang Wang. "Multi-task learning: A solution of small sample size problem in floor-based gait recognition". In: *2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)* (2022), pp. 1–6.

[28] Hunter Derby, Harish Chander, Sachini NK Kodithuwakku Arachchige, Alana J Turner, Adam C Knight, Reuben Burch, Charles Freeman, Chip Wade, and John C Garner. "Occupational footwear design influences biomechanics and

physiology of human postural control and fall risk". In: *Applied Sciences* 13.1 (2022), p. 116.

[29] Rahul Dey and Fathi M Salem. "Gate-variants of gated recurrent unit (GRU) neural networks". In: *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*. IEEE. 2017, pp. 1597–1600.

[30] Chao Fan, Yunjie Peng, Chunshui Cao, Xu Liu, Saihui Hou, Jiannan Chi, Yongzhen Huang, Qing Li, and Zhiqiang He. "Gaitpart: Temporal part-based model for gait recognition". In: *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*. Seattle, USA, 2020, pp. 14225–14233.

[31] Chunxiao Fan, Xiaopeng Hong, Lei Tian, Yue Ming, Matti Pietikainen, and Guoying Zhao. "PCANet-II: When PCANet Meets the Second Order Pooling". In: *IEICE Transactions on Information and Systems* E101.D.8 (2018), pp. 2159–2162.

[32] Claudio Filipi Gonçalves dos Santos, Diego de Souza Oliveira, Leandro A. Passos, Rafael Gonçalves Pires, Daniel Felipe Silva Santos, Lucas Pascotti Valem, Thierry P. Moreira, Marcos Cleison S. Santana, Mateus Roder, Jo Paulo Papa, et al. "Gait recognition based on deep learning: a survey". In: *ACM Computing Surveys (CSUR)* 55.2 (2022), pp. 1–34.

[33] Davrondzhon Gafurov. "A survey of biometric gait recognition: Approaches, security and challenges". In: *Annual Norwegian Computer Science Conference*. Annual Norwegian Computer Science Conference Norway. Trondheim, Norway, 2007, pp. 19–21.

[34] Yao Ge, Wenda Li, Muhammad Farooq, Adnan Qayyum, Jingyan Wang, Zikang Chen, Jonathan Cooper, Muhammad Ali Imran, and Qammer Hussain Abbasi. "LoGait: LoRa Sensing System of Human Gait Recognition Using

Dynamic Time Warping". In: *IEEE Sensors Journal* 23 (2023), pp. 21687–21697.

[35] Jürgen T Geiger, Maximilian Kneißl, Björn W Schuller, and Gerhard Rigoll. "Acoustic gait-based person identification using hidden Markov models". In: *Proceedings of the 2014 Mapping Personality Traits Challenge and Workshop*. 2014, pp. 25–30.

[36] Muhammed Golec, Sukhpal Singh Gill, Rami Bahsoon, and Omer Farooq Rana. "BioSec: A Biometric Authentication Framework for Secure and Private Communication Among Edge Devices in IoT and Industry 4.0". In: *IEEE Consumer Electronics Magazine* 11 (2022), pp. 51–56.

[37] Yu Guan and Chang-Tsun Li. "A robust speed-invariant gait recognition system for walker and runner identification". In: *2013 International Conference on Biometrics (ICB)*. IEEE. Madrid, Spain, 2013, pp. 1–8.

[38] Pınar Güner Şahan, Suhap Şahin, and Fidan Kaya Gülağız. "A survey of appearance-based approaches for human gait recognition: techniques, challenges, and future directions". In: *The Journal of Supercomputing* (2024), pp. 1–38.

[39] Jun Guo, Ling Wang, Daiyin Zhu, Chang-Yu Hu, and Chen-Yan Xue. "Subspace Learning Network: An Efficient ConvNet for PolSAR Image Classification". In: *IEEE Geoscience and Remote Sensing Letters* 16.12 (2019), pp. 1849–1853.

[40] V. G. Manjunatha Guru and V. N. Kamalesh. "Vision based Human Gait Recognition System: Observations, Pragmatic Conditions and Datasets". In: *Indian Journal of Science and Technology* 8 (2015).

[41] Jinguang Han and Bir Bhanu. "Individual recognition using gait energy image". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28.2 (2005), pp. 316–322.

[42] BingZhang Hu, Yu Guan, Yan Gao, Yang Long, Nicholas Lane, and Thomas Ploetz. "Robust cross-view gait recognition with evidence: A discriminant gait GAN (DiGGAN) approach". In: *arXiv preprint arXiv:1811.10493* (2018).

[43] Xiaohu Huang, Duowang Zhu, Hao Wang, Xinggang Wang, Bo Yang, Botao He, Wenyu Liu, and Bin Feng. "Context-sensitive temporal feature learning for gait recognition". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Montreal, Canada, 2021, pp. 12909–12918.

[44] Yeon-Keun Jeong and Kwang-Ryul Baek. "Asymmetric gait analysis using a DTW algorithm with combined gyroscope and pressure sensor". In: *Sensors* 21.11 (2021), p. 3750.

[45] Gunnar Johansson. "Visual perception of biological motion and a model for its analysis". In: *Perception and Psychophysics* 14 (1973), pp. 201–211.

[46] Kooksung Jun, Sanghyub Lee, Deok-Won Lee, and Mun Sang Kim. "Deep learning-based multimodal abnormal gait classification using a 3D skeleton and plantar foot pressure". In: *IEEE Access* 9 (2021), pp. 161576–161589.

[47] A. K. Kavitha. "Fusion of Iris and Retina Using Rank-Level Fusion Approach". In: *International Journal of Computer Technology and Applications*. 2011.

[48] Saeed Kazemi, Angkoon Phinyomark, and Erik J. Scheme. "Sample Size in Floor Sensor-Based Gait Recognition for Smart Home and Access Control Scenarios". In: *2023 IEEE Sensors Applications Symposium (SAS)* (2023), pp. 1–6.

[49] Rohit Khokher, Ram Chandra Singh, and Rahul Kumar. "Footprint Recognition with Principal Component Analysis and Independent Component Analysis". In: *Macromolecular Symposia* 347.1 (2015), pp. 16–26.

[50] Munish Kumar, Navdeep Singh, Ravinder Kumar, Shubham Goel, and Krishan Kumar. "Gait recognition based on vision systems: A systematic survey". In: *J. Vis. Commun. Image Represent.* 75 (2021), p. 103052.

[51] Robyn Larracy, Angkoon Phinyomark, and Erik J. Scheme. "Gait Representation: from Vision-Based to Floor Sensor-Based Gait Recognition". In: *2023 IEEE Sensors Applications Symposium (SAS)* (2023), pp. 1–6.

[52] Chao Li, Xin Min, Shouqian Sun, Wenqian Lin, and Zhichuan Tang. "DeepGait: A learning deep convolutional representation for view-invariant gait recognition using joint Bayesian". In: *Applied Sciences* 7.3 (2017), p. 210.

[53] Xuelong Li, Stephen Lin, Shuicheng Yan, and Dong Xu. "Discriminant locally linear embedding with high-order tensor data". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 38.2 (2008), pp. 342–352.

[54] Shengcai Liao, Yang Hu, Xiangyu Zhu, and Stan Z Li. "Person re-identification by local maximal occurrence representation and metric learning". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, pp. 2197–2206.

[55] Hui Liu and Zhihao Long. "An improved deep learning model for predicting stock market price time series". In: *Digital Signal Processing* 102 (2020), p. 102741.

[56] Tao Liu, Yoshio Inoue, and Kyoko Shibata. "A wearable ground reaction force sensor system and its application to the measurement of extrinsic gait variability". In: *Sensors* 10.11 (2010), pp. 10240–10255.

[57] Cheng-Yaw Low, Andrew Beng-Jin Teoh, and Kar-Ann Toh. "Stacking PCANet +: An Overly Simplified ConvNets Baseline for Face Recognition". In: *IEEE Signal Processing Letters* 24.11 (2017), pp. 1581–1585.

[58] Carlos A Lozano-Ortiz, Adriane MS Muniz, and Jurandir Nadal. "Human gait classification after lower limb fracture using Artificial Neural Networks and principal component analysis". In: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE. Buenos Aires, Argentina, 2010, pp. 1413–1416.

[59] Hsuan-Lun Lu, Mei-Ying Kuo, Chu-Fen Chang, Tung-Wu Lu, and Shih-Wun Hong. "Effects of gait speed on the body's center of mass motion relative to the center of pressure during over-ground walking". In: *Human Movement Science* 54 (2017), pp. 354–362.

[60] Jiwen Lu, Erhu Zhang, Zhigang Zhang, and Yanxue Xue. "Gait recognition using independent component analysis". In: *International Symposium on Neural Networks*. Springer. 2005, pp. 183–188.

[61] Zhuowen Lv, Xianglei Xing, Kejun Wang, and Donghai Guan. "Class Energy Image Analysis for Video Sensor-Based Gait Recognition: A Review". In: *Sensors* 15.1 (2015), pp. 932–964. ISSN: 1424-8220.

[62] Eve MacDonald, Robyn Larracy, Angkoon Phinyomark, and Erik Scheme. "Underfoot Pressure-based Left and Right Foot Classification Algorithms: The Impact of Footwear". In: *IEEE Access* (2023).

[63] Nemanja S Malinović, Bratislav B Predić, and Miloš Roganović. "Multilayer long short-term memory (LSTM) neural networks in time series analysis". In: *2020 55th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*. IEEE. Nis, Serbia, 2020, pp. 11–14.

[64] Hajra Masood and Humera Farooq. “Utilizing Spatio Temporal Gait Pattern and Quadratic SVM for Gait Recognition”. In: *Electronics* (2022).

[65] Leland McInnes, John Healy, and James Melville. “Umap: Uniform manifold approximation and projection for dimension reduction”. In: *arXiv preprint arXiv:1802.03426* (2018).

[66] Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. “UMAP: Uniform Manifold Approximation and Projection”. In: *Journal of Open Source Software* 3.29 (2018), p. 861.

[67] Q Mey, Yaodong Gu, Dong Sun, and Justin Fernandez. “How foot morphology changes influence shoe comfort and plantar pressure before and after long distance running?” In: *Acta of bioengineering and biomechanics* 20.2 (2018).

[68] Robert Miller, Natasha Kholgade Banerjee, and Sean Banerjee. “Temporal effects in motion behavior for virtual reality (vr) biometrics”. In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. Christ church, New Zealand, 2022, pp. 563–572.

[69] Jashila Nair Mogan, Chin Poo Lee, and Kian Ming Lim. “Advances in vision-based gait recognition: From handcrafted to deep learning”. In: *Sensors* 22.15 (2022), p. 5682.

[70] Jashila Nair Mogan, Chin Poo Lee, Kian Ming Lim, Mohammed Ali, and Ali Alqahtani. “Gait-CNN-ViT: Multi-model gait recognition with convolutional neural networks and vision transformer”. In: *Sensors* 23.8 (2023), p. 3809.

[71] Jucheol Moon, Jinwon Jung, Eunice Kang, and Sang-Il Choi. “Open set user identification using gait pattern analysis based on ensemble deep neural network”. In: *IEEE Sensors Journal* 22.17 (2022), pp. 16975–16984.

[72] Takafumi Mori and Hiroaki Kikuchi. “Robust Person Identification based on DTW Distance of Multiple-Joint Gait Pattern.” In: *International Conference on Information Systems Security and Privacy*. 2019, pp. 221–229.

[73] Serafeim P Moustakidis, John B Theocharis, and Giannis Giakas. “Subject Recognition Based on Ground Reaction Force Measurements of Gait Signals”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 38.6 (2008), pp. 1476–1485.

[74] Alvaro Muro-De-La-Herran, Begonya Garcia-Zapirain, and Amaia Mendez-Zorrilla. “Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications”. In: *Sensors* 14.2 (2014), pp. 3362–3394.

[75] Cong Jie Ng and Andrew Beng Jin Teoh. “DCTNet: A Simple Learning-free Approach for Face Recognition”. In: *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*. Hong Kong, China, 2015, pp. 761–768.

[76] Mayu Okumura, Haruyuki Iwama, Yasushi Makihara, and Yasushi Yagi. “Performance evaluation of vision-based gait recognition using a very large-scale gait database”. In: *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)* (2010), pp. 1–6.

[77] Anubha Parashar, Apoorva Parashar, Weiping Ding, Rajveer S Shekhawat, and Imad Rida. “Deep learning pipelines for recognition of gait biometrics with covariates: a comprehensive review”. In: *Artificial Intelligence Review* 56.8 (2023), pp. 8889–8953.

[78] Ji Su Park and Choong Hyun Kim. “Ground-reaction-force-based gait analysis and its application to gait disorder assessment: New indices for quantifying walking behavior”. In: *Sensors* 22.19 (2022), p. 7558.

[79] Todd C Pataky, Kerstin Bosch, Tingting Mu, Noël L W Keijsers, Veerle Segers, Dieter Rosenbaum, and John Y Goulermas. "An Anatomically Unbiased Foot Template for Inter-Subject Plantar Pressure Evaluation". In: *Gait & Posture* 33.3 (2011), pp. 418–422. ISSN: 0966-6362.

[80] Todd C Pataky, Tingting Mu, Kerstin Bosch, Dieter Rosenbaum, and John Y Goulermas. "Gait Recognition: Highly Unique Dynamic Plantar Pressure Patterns Among 104 Individuals". In: *Journal of The Royal Society Interface* 9.69 (2012), pp. 790–800.

[81] Lipeng Qin, Ming Guo, Kun Zhou, Jianqiang Sun, Xiangyong Chen, and Jianlong Qiu. "Gait Recognition Based on Two-Stream CNNs With Multisensor Progressive Feature Fusion". In: *IEEE Sensors Journal* 24 (2024), pp. 13676–13685.

[82] MV Ramakrishna, Sarikonda Harika, S Mukesh Chowdary, Thammarapu Pavan Kumar, Tokala Krishna Vamsi, and Medebalimi Adilakshmi. "Machine learning based gait recognition for human authentication". In: *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*. IEEE. Erode, India, 2023, pp. 1316–1322.

[83] Veenu Rani and Munish Kumar. "Human gait recognition: A systematic review". In: *Multimedia Tools and Applications* (2023), pp. 1–35.

[84] Jody L Riskowski, Thomas J Hagedorn, Alyssa B Dufour, and Marian T Hannan. "Associations of region-specific foot pain and foot biomechanics: the Framingham Foot Study". In: *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences* 70.10 (2015), pp. 1281–1288.

[85] Alex Roberts, Ala Salehi, Angkoon Phinyomark, and Erik Scheme. "Gait Recognition Using EigenfeetNet". In: *2023 IEEE Sensors Applications Symposium (SAS)*. IEEE. Ottawa, Canada, 2023, pp. 1–6.

[86] Ala Salehi, Alex Roberts, Angkoon Phinyomark, and Erik Scheme. "Feature learning networks for floor sensor-based gait recognition". In: *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE. Sydney, Australia, 2023, pp. 1–5.

[87] Oksana Samko, A David Marshall, and Paul L Rosin. "Selection of the optimal parameter value for the Isomap algorithm". In: *Pattern Recognition Letters* 27.9 (2006), pp. 968–979.

[88] Claudio dos Santos, Diego de Souza Oliveira, Leandro A. Passos, Rafael Gonç alves Pires, Daniel Felipe Silva Santos, Lucas Pascotti Valem, Thierry P. Moreira, Marcos Cleison S. Santana, Mateus Roder, Jo Paulo Papa, and Danilo Colombo. "Gait Recognition Based on Deep Learning: A Survey". In: *ACM Comput. Surv.* 55.2 (Jan. 2022). ISSN: 0360-0300.

[89] Leonor Saraiva, Mariana Rodrigues da Silva, Filipe Marques, Miguel Tavares da Silva, and Paulo Flores. "A review on foot-ground contact modeling strategies for human motion analysis". In: *Mechanism and Machine Theory* 177 (2022), p. 105046.

[90] Alireza Sepas-Moghaddam and Ali Etemad. "Deep gait recognition: A survey". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.1 (2022), pp. 264–284.

[91] Muhammad Sharif, Muhammad Attique Khan, Abdullah Alqahtani, Muhammad Nazir, Shtwai Alsubai, Adel Binbusayyis, and Robertas Damaeviius. "Deep Learning and Kurtosis-Controlled, Entropy-Based Framework for Human Gait Recognition Using Video Sequences". In: *Electronics* (2022).

[92] Chuanfu Shen, Shiqi Yu, Jilong Wang, George Q Huang, and Liang Wang. "A comprehensive survey on deep gait recognition: algorithms, datasets and challenges". In: *arXiv preprint arXiv:2206.13732* (2022).

[93] Jasvinder Pal Singh, Sanjeev Jain, Sakshi Arora, and Uday Pratap Singh. "A survey of behavioral biometric gait recognition: Current success and future perspectives". In: *Archives of Computational Methods in Engineering* 28 (2021), pp. 107–148.

[94] Yiyang Su, Minchul Kim, Feng Liu, Anil Jain, and Xiaoming Liu. "Open-Set Biometrics: Beyond Good Closed-Set Models". In: *arXiv preprint arXiv:2407.16133* (2024).

[95] Philippe Terrier. "Gait Recognition via Deep Learning of the Center-of-Pressure Trajectory". In: *Applied Sciences* 10.3 (2020). ISSN: 2076-3417.

[96] Nikolaos Theodorakis. "Secure and Privacy-Preserving User Authentication Using Biometrics". PhD thesis. University of Piraeus (Greece), 2018.

[97] Ka-Wing Tse and Kevin Hung. "User Behavioral Biometrics Identification on Mobile Platform using Multimodal Fusion of Keystroke and Swipe Dynamics and Recurrent Neural Network". In: *2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)* (2020), pp. 262–267.

[98] Ruben Vera-Rodriguez, Pedro Tome, Julian Fierrez, and Javier Ortega-Garcia. "Fusion of Footsteps and Face Biometrics on an Unsupervised and Uncontrolled Environment". In: *Sensing Technologies for Global Health, Military Medicine, Disaster Response, and Environmental Monitoring II; and Biometric Technology for Human Identification IX*. Ed. by Sárka O Southern, Arend H J Kolk, Kevin N Montgomery, Carl W Taylor, B V K Vijaya Kumar, Salil Prabhakar, and Arun A Ross. Vol. 8371. SPIE, 2012, 83711U.

[99] Ming Wang, Beibei Lin, Xianda Guo, Lincheng Li, Zheng Zhu, Jiande Sun, Shunli Zhang, Yu Liu, and Xin Yu. "Gaitstrip: Gait recognition via effective strip-based feature representations and multi-level framework". In: *Proceedings*

of the Asian Conference on Computer Vision. Macao, China, 2022, pp. 536–551.

[100] Xiuhui Wang and Wei Qi Yan. "Human gait recognition based on frame-by-frame gait energy images and convolutional long short-term memory". In: *International journal of neural systems* 30.01 (2020), p. 1950027.

[101] Yuan Wang, Hanhui Jiang, Lin Yu, Zixiang Gao, Wei Liu, Qichang Mei, and Yaodong Gu. "Understanding the role of children's footwear on children's feet and gait development: A systematic scoping review". In: *Healthcare*. Vol. 11. 10. MDPI. 2023, p. 1418.

[102] Zengbin Wang, Saihui Hou, Man Zhang, Xu Liu, Chunshui Cao, Yongzhen Huang, and Shibiao Xu. "LandmarkGait: Intrinsic Human Parsing for Gait Recognition". In: *Proceedings of the 31st ACM International Conference on Multimedia* (2023).

[103] Benjamin Wilson, Judy Hoffman, and Jamie Morgenstern. "Predictive inequity in object detection". In: *arXiv preprint arXiv:1902.11097* (2019).

[104] Chi Xu, Yasushi Makihara, Xiang Li, Yasushi Yagi, and Jianfeng Lu. "Speed-invariant gait recognition using single-support gait energy image". In: *Multimedia Tools and Applications* 78 (2019), pp. 26509–26536.

[105] Sijie Yan, Yuanjun Xiong, and Dahua Lin. "Spatial temporal graph convolutional networks for skeleton-based action recognition". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. 1. New Orleans, USA, 2018.

[106] Jieming Yang, Yuanyuan Liu, Yun Wu, PeiJin Yang, and Zhiying Liu. "Gait-Enhance:Robust Gait Recognition of Complex Walking Patterns Based on WiFi CSI". In: *2023 IEEE Smart World Congress (SWC)* (2023), pp. 1–9.

[107] Sang Soo Yeo, Seung-Ryong Rho, Hyung Joon Kim, Jibran Safdar, Umar Farooq Zia, and Mehr Yahya Durrani. "A Triplet-Branch Convolutional Neural Network for Part-Based Gait Recognition". In: *Comput. Syst. Sci. Eng.* 47 (2023), pp. 2027–2047.

[108] Gong Zhang, Yujuan Si, Di Wang, Weiyi Yang, and Yongjian Sun. "Automated detection of myocardial infarction using a gramian angular field and principal component analysis network". In: *IEEE Access* 7 (2019), pp. 171570–171583.

[109] Yongqing Zhang, Tianyu Geng, Xi Wu, Jiliu Zhou, and Dongrui Gao. "ICANet: A Simple Cascade Linear Convolution Network for Face Recognition". In: *EURASIP Journal on Image and Video Processing* 2018.1 (2018), p. 51. ISSN: 1687-5281.

[110] Zhenyue Zhang and Jing Wang. "MLLE: Modified locally linear embedding using multiple weights". In: *Advances in neural information processing systems* 19 (2006).

[111] Shuai Zheng, Kaiqi Huang, and Tieniu Tan. "Evaluation framework on translation-invariant representation for cumulative foot pressure image". In: *2011 18th IEEE International Conference on Image Processing*. Brussels, Belgium, 2011, pp. 201–204.

[112] Shuai Zheng, Kaiqi Huang, Tieniu Tan, and Dacheng Tao. "A Cascade Fusion Scheme for Gait and Cumulative Foot Pressure Image Recognition". In: *Pattern Recognition* 45.10 (2012), pp. 3603–3610. ISSN: 0031-3203.

[113] Hegui Zhu, Ruoyang Xiao, Jiapeng Zhang, Jinhai Liu, Chengqing Li, and Lianping Yang. "A driving behavior risk classification framework via the unbalanced time series samples". In: *IEEE Transactions on Instrumentation and Measurement* 71 (2022), pp. 1–12.

Vita

Candidate's full name: Ala Salehi Kasaei

Universities attended:

2022-2024	Master of Science in Electrical and Computer Engineering, University of New Brunswick
2016-2020	Bachelor of Science in Electrical Engineering, University of Science and Technology

Conference Presentations:

1. Salehi, A., Roberts, A., Phinyomark, A., Scheme, E. "Feature Learning Networks for Floor Sensor-based Gait Recognition", 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Sydney, Australia, Jul 24-27, 2023, DOI: 10.1109/EMBC40787.2023.10340596

2. Roberts, A., Salehi, A., Phinyomark, A., Scheme, E. "Gait Recognition Using EigenfeetNet", 2023 IEEE Sensors Applications Symposium, Ottawa, Canada, Jul 18-20, 2023, DOI: 10.1109/SAS58821.2023.10254099

Other Activities:

1. 1st Place Award, Image of Research Competition, *University of New Brunswick*, April 2024

2. Finalist, Paper Competition, *NB Health Research Week*, November 2023

3. Awarded the Bob and Joan Scott Scholarship in Biomedical Engineering, *University of New Brunswick*, December 2022

4. Leader, Google Developer Student Club, *University of New Brunswick*, September 2023-now

5. Leadership Degree, Economic Empowerment for Women, *Ignite Fredericton*, May 2022

6. Placed in the top 10%, International Mathematical Contest, *University of Waterloo*, September 2016

7. Ranked in the top 0.1% in Iran's National University Entrance Exam, July 2016