

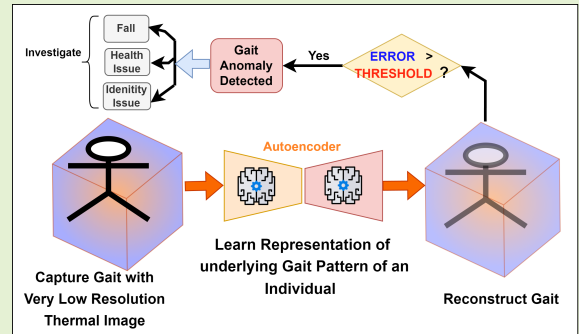


Gait Anomaly Detection with Low Cost and Low Resolution Infrared Sensor Arrays

Farbod Zorriassatine , Abdallah Naser , and Ahmad Lotfi 

Abstract—Detecting anomalies in human gait could be used as indicators of human fall risk or other underlying health or psychological issues. This would require collecting reliable gait data. However, collecting human abnormal gait data is very challenging compared to data gathered from normal daily activities mainly because the former are relatively scarce and may exhibit an unmanageable variability with unpredictable combinations of distorted gait patterns. Recently, it was proposed that privacy concerns due to potential misuses of recorded gait images can be alleviated by using the thermal images captured by the low-resolution and low-cost thermal sensor arrays (TSAs). Therefore, to resolve the privacy concerns and data scarcity simultaneously, this paper proposes a Gait Anomaly Detection (GAD), to be created as a one-class classification (OCC) model and implemented as a reconstruction-based autoencoder (AE), while using TSAs to capture the input data. The data scarcity is conveniently addressed since this GAD design, needs only the plentiful ‘normal’ gait of one person of interest (POI) to build its base model. AE’s were deployed since they learn the intricacies of normal gait patterns, with anomaly threshold placed on the reconstruction errors of the training data. The high performance in detecting specific classes of POI’s gait anomalies, achieving impressive mean values across five critical classification metrics, including F1-score (95.26%), accuracy (96.20%), precision (92.76%), recall (97.92%), and specificity (95.00%), demonstrates the model’s feasibility and practicality. The proposed framework can facilitate independent living among the older adults as an individualised data-efficient, privacy-safe, and low-cost approach to GAD.

Index Terms—Thermal sensor array, anomaly detection, deep learning, autoencoders, gait analysis, data scarcity, data privacy



I. INTRODUCTION

GAIT analysis (GA), is a highly multidisciplinary field. GA can be used for detecting changes in gait for enabling early identification of important hazards such as fall occurrences or their increased risks [1], [2], as well as other health issues including musculoskeletal and neurological abnormalities, [3], [4]. Numerous methods are used for gait analysis including qualitative techniques comprising observational (physical and visual examination) and subjective clinical assessment. Quantitative GA methods include kinematic analysis, kinetic analysis, electromyography, deep learning (DL) and other machine learning (ML) driven techniques [5]. Although these methods can be complementary, they differ in many aspects such as simplicity, validity, reliability, cost, availability, responsiveness, and scalability [6]. GA can be implemented with two main types of sensors to capture gait data and

patterns: external sensors (including vision-based systems [7]), and wearable sensors [5], [8]. The focus here will be on vision-based sensors. Despite the reported success of modern quantitative GA research and practices, many challenges exist which include difficulty of detailed clinical interpretation of detected gait problems, gait data distortion due to changes in view point, and ‘noise’ caused by environmental factors such as terrain obstacles or occlusion affecting the quality and reliability of gait data [5]. The majority of GA applications are confined to research institutions and not leveraged sufficiently in clinical settings [6]. Particularly, with respect to ML based approaches to GA, some of the main challenges ([9], [10]), are privacy concerns (inadvertently allowing misuse of personal identity information [11]), and gait complexity and variability (caused by complex interplay of major parts of the nervous, musculoskeletal and cardiorespiratory systems influenced by multitude of factors like age, personality, mood and socio-cultural [12]). A considerable proportion of the abnormal gait patterns either collected or simulated by using individuals replicating gait abnormalities lack adequate resemblance to the wide range of actual anomalies experienced by people in the real-world [13]. The effectiveness of DL and other ML-based methodologies for GA, heavily depends on vast

Farbod Zorriassatine is a research fellow within the Computational Intelligence and Applications Research Group, School Of Science and Technology, Nottingham Trent University, NG11 8NS, United Kingdom (email: farbod.zorriassatine@ntu.ac.uk).

Abdallah Naser is a lecturer within the Department of Computer Science at Nottingham Trent University, Nottingham, NG11 8NS, United Kingdom (email: abdallah.naser@ntu.ac.uk).

Ahmad Lotfi is the Head of Computational Intelligence and Applications Research Group at Nottingham Trent University, Nottingham, NG11 8NS, United Kingdom (email: ahmad.lotfi@ntu.ac.uk).

quantities of reliable data for training purposes. However, lack of such data that can truly reflect gait variability has given rise to another important challenge in GA, i.e. data scarcity [14]. The standard solutions to data scarcity such as data augmentation using Generative Adversarial Networks (GANs) and attention mechanism [15], are claimed to produce gait sequences that are not consistent with the biomechanical constraints of human walking [14]. As a viable alternative, the use of physics-based simulators to synthesise biomechanically plausible walking sequences has been proposed [14]. As for the privacy challenge to GA, conventional visible-light cameras which are very popular in GA research [7], are not appropriate for indoor environments (e.g., homes or nursing homes). To overcome this, [11] recently demonstrated Thermal Sensor Arrays (TSAs) as practical alternatives to the conventional cameras and high resolution infrared imaging for analysing gait data, and in particular for fall detection. Despite their reported success, TSA-based systems need to reduce their false-positive errors, although as a possible solution a human-in-the-loop has been proposed [11].

Anomaly detection (aka novelty detection and abnormality detection), is a method for raising alarms when a phenomenon of interest exhibits behaviour that is unusual [16]. GA aims to assess and quantify gait patterns to specifically provide diagnosis of the type of problem, while a related technique called Gait Anomaly Detection (GAD) can be used to detect when the healthy (or normal) state of an individual has significantly deviated from their gait norm [4] [17] [18] [19]. Although GAD can indicate that a person may no longer be in a safe condition to maintain their normal gait patterns, it typically does not provide specific diagnostics. However, as a critical early warning system, GAD is used to raise alarms regarding the safety or health of the individuals concerned. After detecting a gait anomaly, depending on the situation, investigative, remedial or rescue measures can be taken by gait specialists, healthcare staff or family members to help the person in trouble.

The challenges of data variability and scarcity can lead to increases in false diagnoses in GA. Therefore, this article aims to explore whether it is possible to address these challenges within a GAD solution to achieve a high classification performance with low false positives and negatives realistically and efficiently. The proposed GAD system, incorporated TSA sensors to also address the privacy concerns, was created as a one-class classification (OCC) [20] model and successfully implemented as a reconstruction-based Autoencoder (AE) [21]. The reference model was trained on examples of normal gait, captured by the TSA sensor from only one person. Three input features including the original TSA recorded thermal data, and two features of pixel motion analysis using Optical Flow (i.e. speed and direction) [22], were included in the gait data. Gait anomalies, presented as three classes of simulated abnormalities for the same person, were successfully detected. The contribution of the proposed novel framework in achieving successful GAD can be summarised as:

- 1) efficiently creating an anomaly detection base model from very low resolution thermal images and with only several short normal-gait footage of one person of

interest (POI), as opposed to requiring gait examples (normally captured with much higher image resolution) from a large number of people or public data repositories

- 2) providing an effective GAD solution to the challenges of data scarcity and variability while maintaining low false negative and false positive rates
- 3) providing the first successful reported use of TSAs within a semi-supervised AE for a privacy preserving GAD and facilitating adaptability to real-world settings
- 4) demonstrating the strength of standard AEs to capture essential aspects of gait pattern even without needing additional features such as Optical Flow's speed and direction components

The remainder of this paper is organised as follows. In Section II, a brief overview of related work on vision-based GAD with emphasis on deep learning methods is provided. In Section III, the research methodology is described in detail, including gait data acquisition, architecture of the AE, choice of the loss function, and training procedure. Section IV presents three GAD experiments and their setups, results, and analysis investigating the effects of training size, window size, and type of input feature. In Section V, the success of the proposed framework is described while outlining limitations of the research and avenues for future work.

II. RELATED WORK

In GA applications, when using vision-based monitoring, the footage of gait is recorded as videos or a series of still image frames for further detailed analysis of the walking patterns. Using ML approaches, the vision-based GAD had been applied in different ways, mainly including: detecting when the healthy (or normal) state of an individual has deviated - such as falls or mental disorders [18]; or providing biometric authentication relying on gait as a unique signature [23]. In order to analyse the captured data, various types of features are extracted, either manually and explicitly by computer vision experts e.g. by capturing and representing the motion and energy distribution of individuals during walking sequences [24], pose estimation or silhouette segmentation of the person from the video frames [25]; or by using inherent and implicit techniques based on Deep Learning (DL) algorithms [23]. DL models are used because of their strengths in dealing with high-dimensionality in gait data [26].

The DL structures reported included Convolutional Neural Networks (CNNs) using both vision-based [27], and non-vision data such as [28] [4], Recurrent Neural Networks (RNNs) [29] and some of the more sophisticated versions of their Long Short-Term Memory (LSTMs) which have the ability to retain long-term dependencies in sequential data, making them suitable for tasks involving time series data [11], and transformer models [30]. Variants of CNNs in integration with Transfer Learning (TL) techniques can learn the spatial and temporal aspects of gait [26], [31]. Some GAD studies showed that their Support Vector Machines (SVM) [17] and K-nearest neighbours (KNN) [32] had displayed a significantly higher accuracy than other CNN and LSTM models [33].

There seems to be no universal consensus regarding the superiority of any one particular ML model. For example,

TABLE I: Comparison of a selection of Gait Anomaly Detection methodologies including wearable, and vision-based sensors for generally improving healthcare of the elderly population. The entries are organised chronologically starting from year 2018 to 2023 and presently including our current paper. New abbreviation used: Random Forest Classifier (RFC), Decision Tree Classifier (DTC), and Extra Tree Classifier (ETC)

Article	Input features Used	Model Deployed	Performance Accuracy	Baseline	Classification
[4]	Accelerometer data from subject's shoes	LSTM, CNN and AE	CNN-based reduction: training accuracy 94% & testing 95%. CNN & AE accuracy 89 – 95%.	Public	Supervised
[29]	Accelerometric and gyroscopic data, video data, time-synchronised from a smartphone	RNN / CNN	Achieved 100% accuracy in detecting anomalous gaits. SVM classifier: Attained 98.077% classification accuracy on the test	Individual	Supervised
[32]	Statistical features related to multivariate time series electromyography (EMG)	KNN algorithm and random forest classifier	Accuracy varied from 21% to 97% for different gait patterns	Public	Supervised
[8]	Tri-axial accelerometer and gyroscope data.	KNN, RFC, DTC, ETC	KNN achieved 97.03% accuracy, RFC achieved 94.95% accuracy. Normal cases: KNN - 48.92%, RFC - 48.27%. Abnormal cases: KNN - 48.10%, RFC - 46.67%.	Public	Supervised
[27]	Depth images of gait sequences captured using Kinect sensor. Foreground subject images generated through image segment	2D-CNN and 3D-CNN	2D-CNN 94.3% for pathological , 3D-CNN accuracy: 95%	Public	Supervised
[18]	Ground Reaction Force (GRF) data from wearable sensors, Time series GRF amplitude values of different axes, Metadata information	mainly CatBoost	accuracy 96%, F1 score 95%.	Public	Supervised
[17]	x, y, z linear accelerations, xx, xy, xz angular speed, Lateromedial and anteroposterior angles, normalized force	One class -SVM	87.5% accuracy in healthy, 82.5% in MS patients	Individual	Semi-supervised
[19]	Accelerometer and gyroscope signals in 3 axes	Matrix-profile	widely varying depending on the class of anomaly detected	Individual	Unsupervised
[28]	Acceleration, gyroscope, and orientation components , Age, weight, and height data	CNN	Accuracy ranged from 88% – 91%. Normal vs. abnormal cases accuracy not explicitly mentioned in contexts.	Public	Supervised
Our Work	Raw body temperature values from thermal sensor arrays, plus extracted Optical Flow's pixel speed and direction	Reconstruction based AE	Minimum 95% accuracy in both normal and abnormal cases, also across F1-score, precision, and recall metrics	Individual	Semi-supervised

another related study reported using LSTM with 99.7% accuracy [11]. Other examples have deployed ensembles of various models such as One-Class Support Vector Machine together with isolation forest, robust covariance estimator and local outlier factor for GAD and reported an accuracy of 98% [34]. These differences may be well explained by variability between the gait problems addressed, type of deployed vision-sensor, preprocessing and feature extraction method, datasets, and model composition and structure, among other experimental settings that may have been different. In addition, not all the related studies shared the same performance metrics as some never reported accuracy and relied on other reliable criteria such as F1-score that combines the precision and recall scores [5]. Therefore, a balanced and objective comparison to identify the best GAD practice is indeed difficult. In Table I a selection of various important papers discussing GAD are compared. The table demonstrates the diversity of the deployed: type of input features (e.g. accelerometer and vision sensors), models (e.g. LSTMs, CNNs, SVMs and AEs), accuracy performance and other metrics, and type of classification approach (e.g. supervised or unsupervised). Table I also highlights a key difference: how GAD approaches establish their normal baseline references, either from a single individual or aggregated population data in public datasets.

In general, there are at least three main types of anomalies comprising point, collective, and contextual anomalies [35].

Furthermore, among the major types of anomalies occurring during GAD, anatomical, biomechanical, and physical ones are also discussed. Contribution of disorders such as neurological or musculoskeletal factors, further complicate distinguishing between the type and cause of gait anomalies identified by GAD systems [7].

An ML-based GAD can be implemented in three different ways in terms of type of learning supervised [27] [11], unsupervised [19], and semi-supervised learning [17]. The type of learning used may be detrimental to its success in overcoming the challenge of gait complexity and variability. It may be argued that supervised learning models commonly found in the research literature, are limited for addressing this challenge. This is because abnormal gait patterns are relatively scarce in the general population (i.e. amongst people of less than 80 years old), while normal gait patterns are plentiful [12]. Successful supervised learning (SL) classification models require a balanced representation of all the classes involved. In the case of GAD, this means having same size sets of examples from the normal gait as well as all the other abnormal gait classes. However, abnormal gait patterns can manifest themselves with an unmanageably substantial number of diverse and unpredictable combinations of distorted normal gait patterns. This is because the normal gait is the result of interaction of various body parts engaged together and when one or more of these fail to function as usual, a gait

abnormality occurs [36]. Consequently, it is impossible to gather balanced datasets with real-world abnormal examples. Especially, the datasets created with healthy people acting as if they have abnormal gait conditions, fail to provide sufficiently comprehensive representations of all the possible abnormal gait patterns for use in machine learning models. More specifically, GAD in the context of an older adult or a person with congenital gait anomaly, can easily lead to increased false positives due to a lack of individualisation of reference normal gait patterns [17]. Therefore, examining alternatives to supervised learning, i.e. the unsupervised or semi-supervised approaches, due to their potential for avoiding training on examples from large sets of anomalous gait patterns deserves great attention.

Conventionally, AE models as particular cases of feedforward Deep Neural Networks, attempt to copy (or reconstruct) their input to their output via an internal latent representation referred to as embeddings. AEs are frequently used for video anomaly detection [37]. AEs have been successfully applied in GAD as they can uncover the underlying structure of the data [38], for detecting gait sequence anomalies in [21], [39], and abnormal gait patterns based on a number of multi-modal bio-signals [38]. Sparse AEs use the sparsity constraint rule to find subtle patterns in the data by ensuring only a few neurons are active at any one time [40]. AEs are known in general for their ability to eliminate the need for manual feature engineering. AEs capture the 2D structure of gait patterns and aggregate the spatial and temporal features. By using the reconstruction error information and setting a threshold on their distribution related to the training with samples from only the normal class, these AEs, effectively, provide an anomaly detection model.

Finally, as for using thermal imaging, important benefits of IR for gait analysis are reliability in silhouette extraction and the resilience towards complexity of background and variations in lighting [41]. However, some of the challenges include limitations caused by occlusion and the field of view [42]. Based on the motion sequence classification of human movements, a recent work successfully used the low resolution TSA data with supervised RNNs [11] to identify human falls from various other typical activities of daily living (ADL). A few other studies have harnessed DL alongside these low-cost, low-resolution infrared sensors to detect anomalies in gait patterns or for fall detection, for example see [43]. These TSAs using DL have been also used for counting people [44] and occupancy estimation [43]. A study using public datasets including high-resolution thermal images of 640×480 pixels deployed several variants of Deep Spatio-Temporal Convolutional Autoencoders (DSTCAEs) [45] by formulating fall detection as an anomaly detection problem. The latter reported that their proposed models outperform the convolutional LSTM-AEs for detecting falls.

III. THE PROPOSED APPROACH

As a pragmatic solution to implement GAD with visual sensors, autoencoders were identified as a reliable approach. Their function to reconstruct their input as output provides them with a powerful capacity to capture gait patterns' nuances

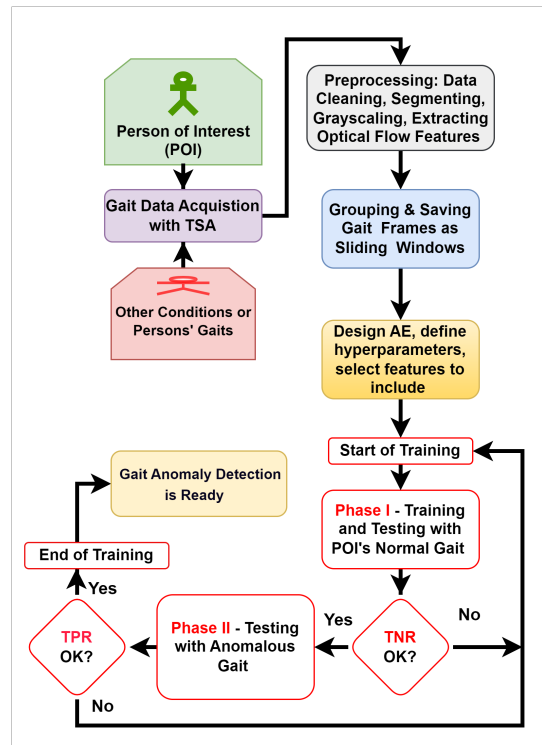


Fig. 1: Proposed framework for creating a personalised Gait Anomaly Detection system based on the current normal gait of the POI. TNR and TPR stand for True Negative Rate (specificity) and True Positive Rate (recall), respectively.

implicitly. This, in turn, enables AEs to manifest a reconstruction behaviour when receiving input that is different from what they have learnt as normal gait or the reference patterns. Gait anomalies should result in distinctively greater reconstruction errors that signal the likely presence of deviations from the learnt 'knowledge' of the AE. In Fig. 1, a schematic diagram of the proposed framework is shown which is further described below.

A. Thermal Sensor Arrays Signal Acquisition

The proposed framework uses TSAs for anomaly detection. This is mainly because of their preserving privacy, affordability, and being suitable for home use. Specifically, the MLX90640 sensor [46] with a -40 to 85°C operational temperature range, and an object temperature accuracy of 1°C precision across the full measurement scale was employed. The TSA measures the temperatures of objects in its field of view and returns the information as a 32×24 matrix. With a TSA sensor refresh rate adjustable between 0.5Hz and 64Hz , it was possible to determine a suitable value for detecting swift human movements. Specifically, high refresh rates ensure that even slow movements, often exhibited by elderly individuals, are captured more frequently across multiple frames [43], [47]. In the experimental setup depicted in Fig. 2, the TSA camera was placed at a height of approximately 1.7 meters, with a distance of 1.5 meters from a 2.5-meter walking path. With a TSA sampling frame rate set at 8Hz , each sampling session captured between 8 to 12 seconds of useful gait information

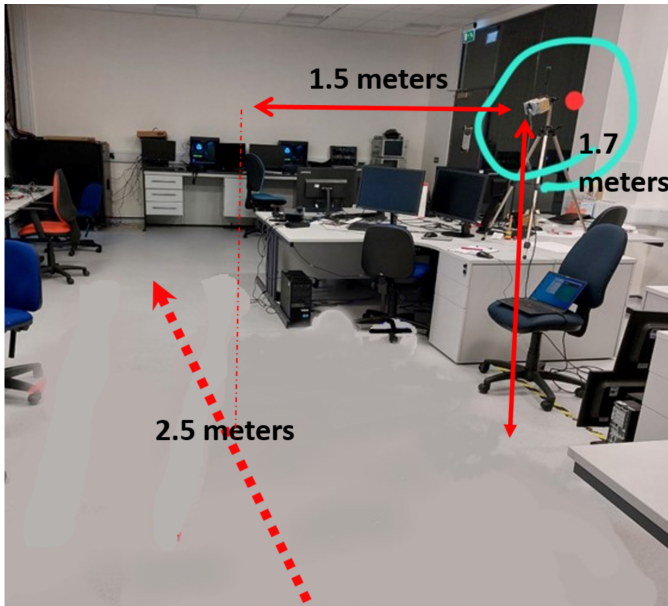


Fig. 2: The arrangement for recording gait showing TSA position and the subject walking path.

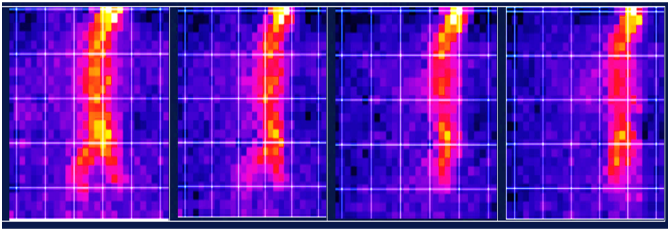


Fig. 3: Four consecutive heatmap frames of a person walking in front of the TSA camera.

while maintaining sufficiently swift playback of the subject's movement. The subject was instructed to walk at a 90° angle to the TSA, providing a lateral view.

Fig. 3 illustrates a sequence of four consecutive heat-map frames capturing the distinctive gait of a person in motion. These thermal snapshots offer a pixelated visual representation, demonstrating the evolving heat patterns as the individual walks, providing valuable insights into the dynamic thermal characteristics of their movement without any information to allow verification of human identity in the TSA's output.

B. Feature Preprocessing Strategies

In this section, a number of essential feature-related sub-processes to achieve feature selection, transformation and creation are provided. These sub-processes were crucial for improving the future performance of the anomaly detection model.

1) *Segmentation and Grayscaleing*: Adopting the approach described in [11], binary thresholding was used to convert the images into binary levels to isolate regions with different temperatures to better expose the human gait. First a vector level of threshold values was calculated using the 'multithresh()' function in MATLAB [48]. The thresholding segmented each of the frames, captured by TSA, into multiple

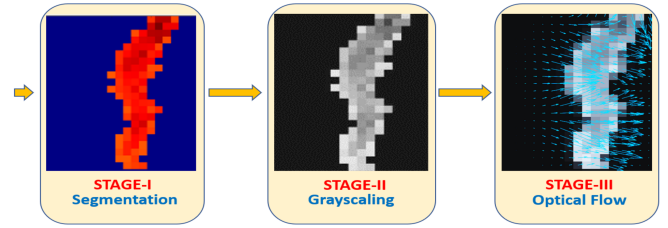


Fig. 4: Three consecutive processing stages - Segmentation, Grayscaleing and Optical Flow - were used to extract additional information to improve GAD performance.

regions using MATLAB's 'imquantize()' function [49]. The output of segmentation was initially a new image frame of the same size as the input image frame, where each pixel was assigned a value of either 1 or 2. Then, all pixel values less than or equal to 1, were further converted to 0, and otherwise to 1, essentially creating a binary mask. Here, a pixel value of '0' generally indicates the areas to ignore as the 'background', and a value of '1' or other non-zero values indicate the areas of interest, also known as 'foreground'. Lastly, all the original thermal pixels corresponding to the above binarised mask were split accordingly to remove the regions of image that corresponded to the lowest intensity values. The segmentation process divided the input thermal image frame into pixels that had either a strong correlation with a walking subject or none. This distinction was based on their temperature characteristics. The resulting transformation after applying a segmented binary mask to the original thermal data, is a heat-map where non-body parts appear as dark blue. A typical transformed heat-map is illustrated as Stage-I in Fig. 4. The next step, i.e. Grayscaleing, shown as Stage-II in Fig. 4, further simplified the representation of the segmented heat-map into shades of gray. This was done to allow a more efficient computation during the next stage, i.e. the Optical Flow analysis, which is described next.

2) *Extracting the Optical Flow Components of Motion Analysis*: Optical Flow method was used to provide two extra features in addition to the 'raw' temperature data, namely: pixel speed and pixel direction as described below. As far as gait analysis is concerned, Optical Flow [22] as a fundamental computer vision technique allows tracking and detection of objects of interest, i.e. the walking person as in this paper. Optical Flow is used to break down the apparent movement of pixels in a picture into their various components i.e. speed and direction. All the following optical flow equations, fundamental to this section, are derived from the established work of [50]. The core assumption of Optical Flow is that the pixel intensities of an object remain constant over time. Mathematically, this can be expressed as:

$$I(x, y, t) = I(x + dx, y + dy, t + dt), \quad (1)$$

where (x, y) are the pixel coordinates in the first image, $(x + dx, y + dy)$ are the pixel coordinates in the second image, t is the time at the first frame, $t + dt$ is the time at the second frame, $I(x, y, t)$ is the pixel intensity at (x, y) at time t , $I(x + dx, y + dy, t + dt)$ is the pixel intensity at the new location at

the later time. This equation simply states that the intensity, I , of a particular pixel remains constant between consecutive frames. Taking the Taylor series expansion of the right side of the equation and keeping up to first order terms, results in:

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \left(\frac{dx}{dt}\right)Ix + \left(\frac{dy}{dt}\right)Iy + dt \cdot It. \quad (2)$$

where I_x and I_y are the spatial derivatives of the image at (x, y, t) , and I_t is the temporal derivative of the image at (x, y, t) . Setting the two sides of the equation equal to each other and rearranging terms, the Optical Flow equation is obtained:

$$\left(\frac{dx}{dt}\right)Ix + \left(\frac{dy}{dt}\right)Iy + It = 0 \quad (3)$$

or, in vector notation:

$$(u, v) \cdot \nabla(I) + It = 0 \quad (4)$$

where

$$(u, v) = \left(\frac{dx}{dt}, \frac{dy}{dt}\right) \quad (5)$$

is the Optical Flow (velocity) vector and $\nabla(I) = (I_x, I_y)$ is the gradient of the image. The above is Optical Flow's constraint equation. It relates pixel flow to image derivatives. More specifically, it relates the flow (velocity) of each pixel to the spatial and temporal derivatives of the image. This is under-determined, so additional constraints are needed, such as smoothness constraints or motion assumptions. Methods like Lucas-Kanade [51] or Horn-Schunck [50] can solve this equation. Here, only the Horn-Schunck method was used which assumes that the flow is smooth over the entire image, and solves the flow vectors in a global manner.

3) Features Saved for Anomaly Detection: For each recorded image frame, three distinct sets of data channels were preserved: temperature, speed, and direction. Speed and direction channels were initially derived from Optical Flow calculations and then masked as described in Section III-B.2. However, the original temperature values recorded for each frame were used without applying that mask, retaining all the original 'raw' temperature data for each individual pixel.

C. Sliding Windows and Data Integration

The input to a GAD model must include data associated with a long enough time interval to capture the gait pattern. Rather than analysing the entire video or sequences as a whole, windowing allows for focused examination within specific time intervals. Each window consists of a fixed number of consecutive frames. 'Sliding window' refers to scanning of temporal segments of a given video or sequence - to divide the longer contiguous footage into smaller sections of equal length, or windows for GAD processing purposes. This approach was similar to those of other researchers [52]. Thus, a window consists of a sequence of related consecutive frames from the same recording session of the POI's gait. Fig.5, provides an example of how the sliding window can extract windows with a fixed number of frames each. It must be noted

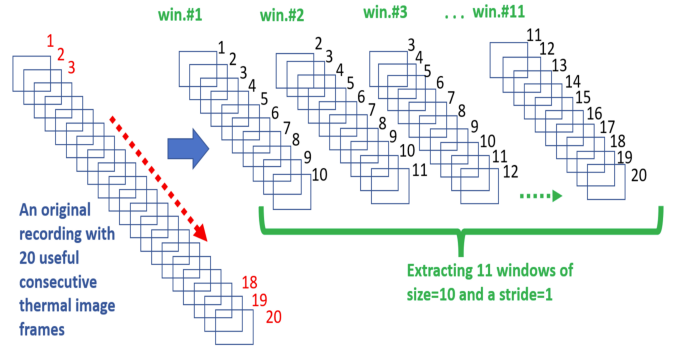


Fig. 5: An example of how efficiently the sliding window provides enough samples for use as input to the GAD model. As shown, from an initial short footage of only 20 frames recorded by TSA (shown on the far left of the illustration) a sliding-stride of 1 and window size of 10, a total of 11 training windows can be extracted.

that the possible number of extracted windows from a given recording decreases as the specified window size increases.

D. Designing the AE based on the Input Data

A fully connected stacked AE model consisting of five sections: an input layer, an encoder, the middle layer, a decoder and an output layer was deployed, see [21] for more typical details. In this model, all the image frames in each window are concatenated into a single input vector. The result was consequently a long, flattened vector of values. The size of the input layer, which must be equal to the size of the input vector, is then given as in (6).

$$S = F \cdot I \cdot P \quad (6)$$

S is the size of the input layer, F is the number of features used per pixel, I is the number of image frames in each window, and P is the number of pixels in each frame. The maximum value for F was 3 as any combination of temperature, pixel speed and pixel direction features could be used. As for the encoder the first layer had 512 nodes and a linear activation function, followed by the second layer with 256 neurons and Rectified Linear Unit (ReLU) activation function (AF). Then, 128 neurons were used in the middle layer (between the encoder and decoder) for the 'Bottleneck' section, with a ReLU AF. The use of ReLU in both the encoder and decoder layers contributed to sparsity because ReLUs output zero for negative inputs, effectively 'turning off' some neurons. Furthermore, each layer of the encoder used an $L1$ regularisation, which also encouraged sparsity in the encoded representations by placing sparsity on the hidden layer AFs. A noticeable AE performance improvement was achieved although a strong sparsity was avoided. This is attributed to using of rather low values of $L1$ (0.00001 at the input layer) and even smaller values at the remaining dense layers. The sparsity level could be tweaked for optimal balance between reconstruction accuracy and generalisation, with stronger sparsity constraint potentially leading to representations that were non-informative, while weaker sparsity

could lead to overfitting. The chosen sparsity levels were found via hyperparameter tuning, where the parameter was adjusted for the best overall performance on a validation set and confirming the choice by selecting the highest generalisation performance model on various test sets with anomalous and normal representations. The Decoder was designed symmetric to the encoder with two layers of 256 and 512 neurons respectively and using ReLU AFs. The output layer which had the same size as the input layer, used a Sigmoid AF to ensure the final output of AE lies in the same range as the input data i.e. a real value between 0 and 1.

E. Training and Testing the AE to Detect Gait Anomalies

The training data was recorded multiple times from only the current ‘normal’ gait of one POI in relatively healthy conditions, which is referred here as ‘P1’. Additionally, three different sets of simulated classes of gait anomalies, referred to as ‘P2’, ‘P3’, and ‘P4’ respectively, were ‘staged’ by the same POI and recorded. These gait-anomaly classes were not planned to mimic any already well-established gait disorders. ‘P2’, ‘P3’, and ‘P4’ were deliberately designed to be uniquely different from each other and from the normal gait of the selected POI, referred to as ‘P1’. Thus, ‘P1’ represents the normal gait of an ageing man with a slight stoop. ‘P2’s gait simulates the same individual walking with legs wide apart, creating a distinctly different stride pattern. ‘P3’, in contrast, depicts a faster-paced gait with legs positioned very close to each other, differing significantly from both ‘P1’ and ‘P2’. Finally, ‘P4’ involves walking with a much greater stoop than ‘P1’ and at a slower pace, introducing another unique variation. The simulation procedure was meticulously rehearsed multiple times before the final thermal recording to maintain consistency within each abnormal class. The sets ‘P2’ to ‘P4’ served as the anomalous gait test-sets for phase-II of the AE training. The total duration of footage for ‘P1’ was approximately 4 minutes spread across 26 recording session with an average session duration of 10 seconds. The number of extracted windows varied depending on the chosen window size. On the other hand, for the three test classes of abnormal gait—‘P2’, ‘P3’, and ‘P4’—the total duration of footage was only 2 minutes, sampled over several short 10-second sessions.

Due to the small and limited size of the original data recorded, the ratio according to which the ‘P1’ dataset was split between training, validation and test set varied according to the experiment, as described later in Section (IV).

As outlined in Fig. 1, the training process was divided between two phases: I and II. The first phase was where the actual GAD training occurred. This was a semi-supervised learning because although no class labels were given to the AE, the input examples were prepared from only one class (i.e. the normal gait of the POI). Therefore, this was contrary to unsupervised training where there is absolutely no separation between the normal and anomalous examples included in the training set. In Phase-I, the AE had to closely reconstruct, as its output, the normal-gait windows it received as input. Thus, as a binary classifier all the ‘normal’ gait data from the POI, i.e. ‘P1’, were treated as ‘Negative’. Next, if after testing

on gait windows from the P1’s test set, the True Negative Rate (TNR) - the same as specificity metric - was sufficiently high (ideally close to 95%), the training would proceed to Phase-II. Of course, it was not always possible to reach this specificity target if the model was not appropriate. The second stage aimed to evaluate the AE on the gait-anomaly test sets from ‘P2’, ‘P3’, and ‘P4’ to determine its suitability for later use as a finalised GAD model (see Fig. 1). Thus, any gait pattern sufficiently different from the ‘normal’ gait of the POI must be classified as ‘Positive’ and result in noticeably higher reconstruction errors. The reconstruction error which was evaluated in terms of the Mean Squared Error (MSE) is discussed in the literature [53], [54]. Throughout this study, a type of percentile-based cut-off on the distribution of MSE values of the reconstruction errors of the training set (which included examples from the normal class only) was deployed. The threshold line set at the 98th percentile of the distribution of the training set’s MSE values formed the anomaly limit or borderline between normal and anomalous gait samples.

IV. EXPERIMENTS AND DISCUSSION

The primary objectives of the experiments were to validate model’s feasibility and functionality, and to determine the most successful configuration for successfully implementing the proposed GAD system. Thus, different parameter settings of training size, window size, and the type of input feature were investigated. The novelty of the work meant that there were too many unknown influencing factors. Therefore, initial experiments focused on establishing a baseline understanding using a subset of data involving only ‘P1’ and ‘P4’, followed by a comprehensive evaluation with all four data sets, ‘P1’ to ‘P4’, to finalise the model configuration. Each set of experiments was conducted after first achieving a specific target performance at the end of Phase-I, to serve as the basis for comparison. Thus, all the experiments were conducted at the Phase-II of the training process (see Fig.1). During the Phase-I training, the batch size and number of epochs were set to 32 and 120, respectively. Adaptive Moment Estimation (ADAM) [55] was used as the optimiser with a fixed learning rate 0.005 to adjust the weights of the network to minimise the loss function. The AE model architecture was also kept identical to that described earlier in Section III-E. These measures were necessary to make the comparative tests within each of the three sets of experiments, as meaningful as possible.

For classification performance, numerous metrics are available (for example see [56]). We found F1-score, accuracy, precision, recall, and specificity most suitable for our objectives (see [3] for a detailed discussion), reducing the likelihood of biased performance conclusions. F1-score was particularly expected to stay as high as possible to obtain a good trade-off between the GAD’s recall and precision performance.

The training set was collected randomly from 55% of the extracted windows from the POI’s ‘P1’ set. About 8% and 37% of the remaining windows were allocated to the validation and test sets for ‘P1’ respectively. During Phase-II, only ‘P4’ was tested in the first and second experiments. In the final, most comprehensive experiment, the GAD model was tested against all three aforementioned anomaly classes.

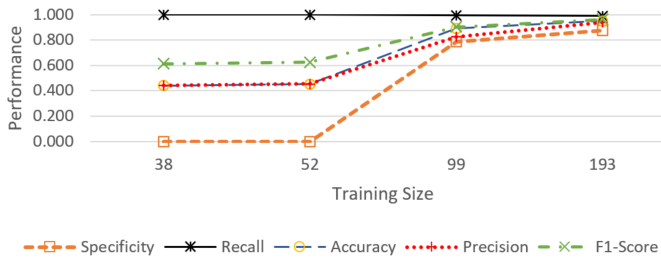


Fig. 6: Effect of training size on anomaly detection performance at a window size of 8, with ‘P1’ as ‘normal’ and only using ‘P4’ as ‘abnormal’.

For each of the following experiments, the associated discussion is provided within the same section as the experiment itself. This arrangement ensures relevance and efficiency by directly connecting the experiment’s description with its corresponding discussion.

A. Experiment-1: Investigating the Effect of Training Dataset Size

The intended users of this system, are expected to be primarily the older adults, many of which may not be able to conveniently participate in data acquisition efforts. This may be attributed to their often limiting physical and/or mental circumstances. Therefore, inspired by few-shot learning [57], in this experiment, the objective was also to determine the minimum training size for accomplishing GAD with the highest classification performance. Only the ‘raw’ temperature feature was included in each input frame. The window size was set to 8 frames to obtain a smooth flow of the gait pattern because the number of frames per second (FPS) of the TSA had been also set as 8 during all the initial gait recordings. This resulted in a total of 352 windows extracted by applying the sliding window process. Four different training sets with sizes of 38, 52, 99 and 193 windows were created for this investigation. The remaining windows from ‘P1’ set were allocated for validation and testing purposes in the Phase-I of the training.

At each training size, only the model with the highest recall results from a set of 15 separate runs was included for comparison with the other training sizes. The results in Fig. 6 clearly show that the largest training size of 193 windows produced superior performance across all the five metrics used, while the training size of 99 ranked second. It must be noted that although the recall metric showed impressive performance across all the training sizes, including for training sets with 52 and 38 windows, the specificity metric was zero for these smaller training sizes. This observation suggests under-fitting for the training sizes smaller than 99 windows. The results showed the effectiveness of the proposed approach for securing useful GAD outcomes with still only a small training size of 193 windows (despite being the largest in the above tested range). This was equivalent to only 2.2 minutes of total captured footage of the ‘normal’ gait. No previous study to provide a benchmark for comparison of data collection duration for GAD approaches was found. Investigating the

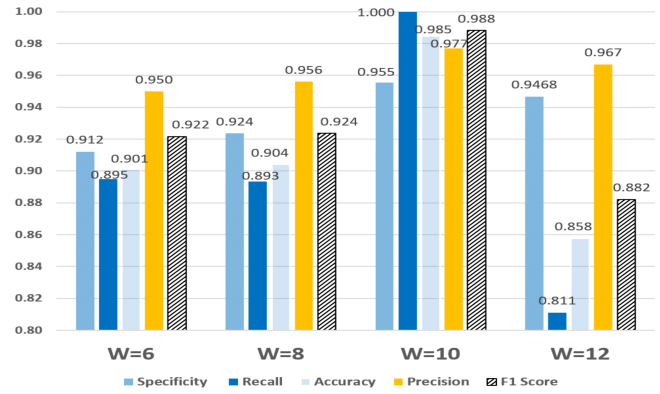


Fig. 7: Effect of different window sizes, on GAD across 5 performance metrics. At any window size, only the model with the highest F1-score from 15 separate runs was used for comparison.

effect of using window sizes larger than 193 was not possible since the existing TSA footage could not provide any more windows. However, it is recommended to examine larger window sizes, obtainable from longer initial recordings, for enhancing the results further.

B. Experiment-2: Investigating the Effect of Window Size

The objective of this experiment was to find the best input window size from the available training dataset. As described in the previous experiment, an initial window size of 8 was also used here to provide a smooth gait flow. However, other additional window size values of 6, 10 and 12 were also investigated. For each input frame, the corresponding values of ‘raw’ temperature, speed, and direction were used simultaneously. The performances included in Fig.7, represent only the models with the highest F1-score results from a set of 15 separate runs for each given window size. Only the ‘P4’ class was tested for anomalies. As seen from Fig.7, a window size of 10 showed the best results at or above 95.5% percent success rates across all the five metrics. The window size of 12 showed a marked degradation of overall performance suggesting that more data through larger window of TSA frames is not necessarily helpful.

A major challenge of using the AE model for GAD, when using an input data obtained from very low resolution TSA infrared images was the apparent loss of detail. Another expected challenge was the possible loss of information when flattening the input data into a one-dimensional vector. The successful GAD performance, suggested that the sparsity driven by the use of $L1$ regularisation, ReLU activation functions, and AE’s inherent dimensionality reduction and reconstruction of the input may have overcome the above challenges to a great degree. However, it was also demonstrated here that using a specific number of random windows as input, each providing collection of unique frame sequences of gait pattern greatly improved the performance of GAD especially at a window size of 10. It is therefore concluded that the sliding windows must have also contributed to GAD by providing sufficiently

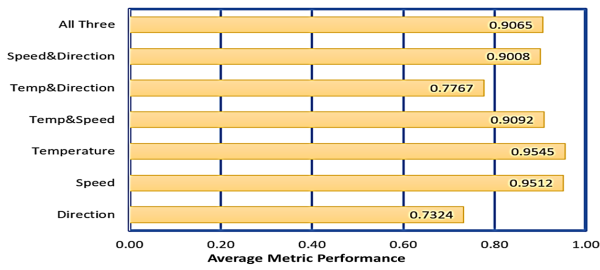


Fig. 8: Overall performance of the seven feature combinations. For each feature combination, this shows the average of the corresponding metric values across the three gait anomaly conditions: ‘P2’, ‘P3’, and ‘P4’.

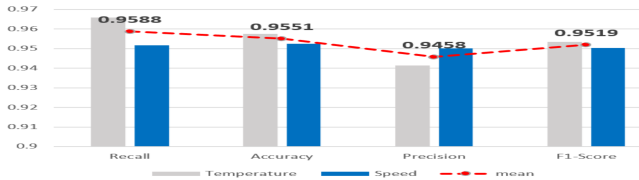


Fig. 9: Comparison of temperature with speed, including their averages across four metrics and averaged over all three gait anomaly conditions: ‘P2’, ‘P3’, and ‘P4’.

representative samples for effective learning of the hidden data structure of the gait patterns.

C. Experiment-3: The Impact of the Type and Composition of Input Features Used

The objective of the experiments here was to determine which form of input could provide the best GAD performance across all the three gait anomaly conditions: ‘P2’, ‘P3’, and ‘P4’. With the three initial features of ‘raw’ Temperature, and Optical Flow’s components of Speed and Direction, seven distinct feature combinations were possible: 1-All three features together, 2-Speed and Direction, 3-Temperature and Direction, 4-Temperature and Speed, 5-Temperature, 6-Speed, and 7-Direction. The previously discovered best performing window size of 10 and the associated training size of 166, and validation and test sets of 24 and 112 windows, respectively, were used. Only the models with a specificity of 95% were included in the comparisons. Fig. 8 shows that models based solely on either ‘raw’ Temperature (overall 95.45%), or Speed (overall 95.12%), outperformed all the GAD models that were based on other input feature combinations. The values reported are the averages of overall performance across the five metrics. Fig.9 shows the breakdown of GAD performances for ‘raw’ Temperature, Speed, and the averages of both of their corresponding performances. The specificity metric results were not included since they were maintained at the same value of approximately 95%.

To further examine GAD performance for various gait anomaly classes, Fig.10 provides a detailed breakdown of the results for each of the three gait anomaly conditions individually. These values are displayed separately for each of F1-score, Accuracy, Recall and Precision metrics. It can

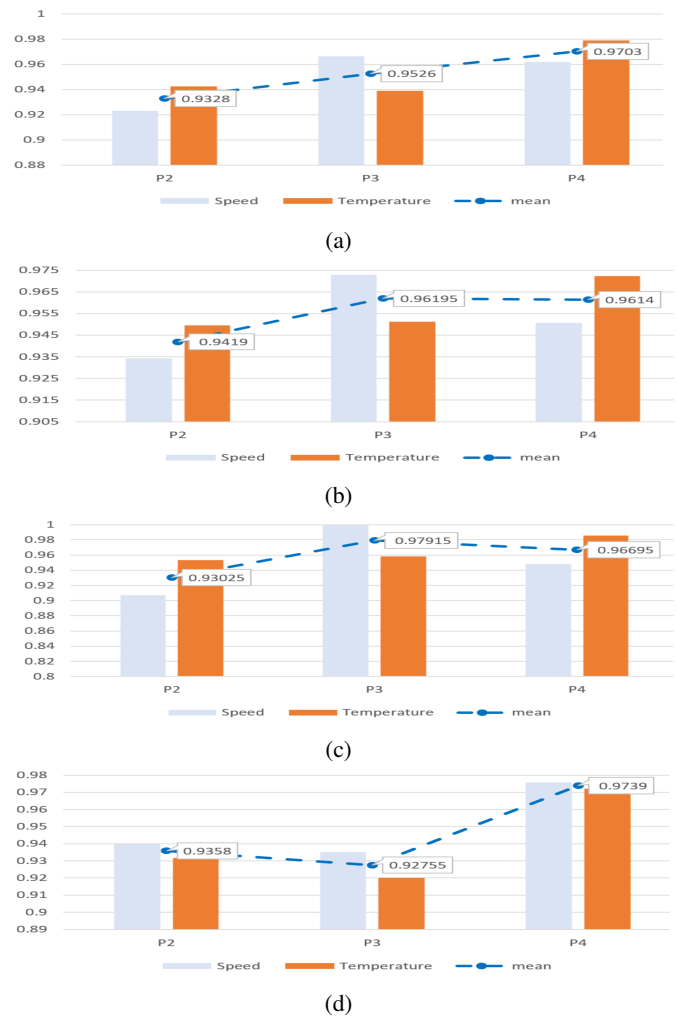


Fig. 10: Comparison of using temperature versus speed as input features for GAD, including their averages across the following metrics: (a) F1-score, (b) accuracy, (c) recall, and (d) precision.

be observed that even for a given ‘winning’ or dominant input feature type, the GAD performance was not uniform and varied depending on the type of anomaly class being detected. Overall, as evident from Fig.10, the ‘Raw’ Temperature feature provided better detection capability for ‘P2’ and ‘P4’ classes, while ‘Speed’ provided superior detection for the ‘P3’ group of gait anomalies. This implied that among the two dominant features, each input feature provided different detection capabilities depending on the type of anomaly encountered. As the actual type of future gait anomalies cannot be predicted, this means that an ensemble of different GAD models that each use different types of input features, should provide a more balanced overall detection capability than relying on only one type of ‘top-performing’ input feature.

V. CONCLUSION

Previous work to specifically confirm the feasibility of performing one-class semi-supervised GAD with very low resolution thermal images from TSA as input could not be

found. However, the proposed TSA-based model developed as a mildly sparse standard AE with fully connected layers, provided successful performance. This personalised solution with minimal training data, created from only less than 4 minutes total footage of one POI's normal gait allowed a convenient and efficient data collection suitable especially for the older adult population. Three simulated classes of the POI's gait anomalies were detected with approximately 95% success rate across all the five important classification metrics (as opposed to many of previously reported works which report high performance for one or two metrics only). The anomalies detected were contextual in relation to the normal gait of POI, as opposed to being global anomalies with respect to the larger human population. The proposed GAD solution significantly alleviates the gait data scarcity and data variability problems because it requires a short footage of just one person's normal gait.

Finally, it is important to highlight the limitations of this study in order to pave the way for future researchers:

- Despite its potential as an early warning system, the semi-supervised GAD framework cannot resolve the challenge of positively identifying specific types of gait anomalies. A possible solution is to integrate diagnostic supervised-GAD models, which could complement our model.
- Due to the short length of the gait recordings, the suitability of larger window sizes could not be adequately examined. Larger window sizes may enhance the GAD model's ability to learn gait nuances and subtle changes to improve overall performance.
- The absence of comparative benchmark studies using low resolution thermal imaging, has made it challenging to evaluate and optimise our work. A potential solution is to repeat the experiments on a larger and more diverse group of targeted users (e.g. including at least 10 or more volunteer elderly individuals from differing ethnic, gender and age groups). This could lead to more personalised baseline models, refining the configuration of personalised GAD systems. As a result, various parameters, including training size, sliding window size, anomaly threshold, and the choice of AE type could be optimised. Additionally, considering ensembles of GAD modules with diverse sensor modalities and environmental conditions may improve performance.
- Creating a more comprehensive dataset with diverse and real anomalous gait patterns for each individual (as opposed to solely relying on a few simulated anomalous classes such as 'P2', 'P3', and 'P4'), could also provide a more realistic evaluation of GAD robustness and allow further optimisation. However, as discussed earlier, the scarcity of real anomalous examples remains a serious problem. Exploiting existing public gait datasets with higher resolution infrared images to extract lower resolution images (e.g., through average or max pooling) and simulating anomalous data with GANs are other options. Additionally, adapting and potentially retraining the model for different camera angles and whole-body images could enhance its robustness across various conditions and datasets.
- The definition of 'normal' gait for a given person must be periodically updated with the latest 'still-healthy' gait conditions to maintain high classification performance.
- In order to increase detection accuracy in the presence of challenges such as the distance between the human and the sensor, and occlusions caused by factors such as thick clothing or hot objects (e.g., a hot drink), various pre-processing techniques and the fusion of multiple TSAs have shown promising results [47] that deserve further exploration.

REFERENCES

- [1] T. Amundsen, M. Rossman, I. Ahmad, A. Clark, and M. Huber, "Fall risk assessment and visualization through gait analysis," *Smart Health*, vol. 25, p. 100284, Sept. 2022.
- [2] K. M. Sünnetci, M. Ordu, and A. Alkan, "Gait based human identification: a comparative analysis," *Computer Science*, no. Special, pp. 116–125, 2021.
- [3] D. Sethi, S. Bharti, and C. Prakash, "A comprehensive survey on gait analysis: History, parameters, approaches, pose estimation, and future work," in *Artificial Intelligence in Medicine*. Elsevier B.V, 2022-07-01, vol. 129.
- [4] G. Paragliola and A. Coronato, "Gait anomaly detection of subjects with parkinson's disease using a deep time series-based approach," *IEEE Access*, vol. 6, pp. 73 280–73 292, 2018.
- [5] A. Saboor, "Latest research trends in gait analysis using wearable sensors and machine learning: A systematic review," *IEEE Access*, vol. 8, pp. 167 830–167 864, 2020.
- [6] A. Hulleck, D. Mohan, N. Abdallah, M. Rich, and K. Khalaf, "Present and future of gait assessment in clinical practice: Towards the application of novel trends and technologies," *Frontiers in Medical Technology*, vol. 4, p. 901331, 2022.
- [7] Y. Sugiyama, K. Uno, and Y. Matsui, "Types of anomalies in two-dimensional videobased gait analysis in uncontrolled environments," *PLoS Comput Biol*, vol. 19, no. 1, p. e1009989, 2023-01.
- [8] S. Adhikary, R. Ghosh, and A. Ghosh, "Gait abnormality detection without clinical intervention using wearable sensors and machine learning," in *Innovations in Sustainable Energy and Technology: Proceedings of ISET 2020*. Springer, 2021, pp. 359–368.
- [9] P. Khera and N. Kumar, "Role of machine learning in gait analysis: a review," *Journal of Medical Engineering and Technology*, vol. 44, no. 8, pp. 441–467, 2020-11-16.
- [10] D. Nohelova, L. Bizovska, N. Vuillerme, and Z. Svoboda, "Gait variability and complexity during single and dual-task walking on different surfaces in outdoor environment," *Sensors*, vol. 21, no. 14, p. 4792, 2021.
- [11] A. Naser, A. Lotfi, M. Mwanje, and J. Zhong, "Privacy-preserving, thermal vision with human in the loop fall detection alert system," *IEEE Trans Hum Mach Syst*, vol. 53, no. 1, p. 164–175, 2022.
- [12] W. Pirker and R. Katzenschlager, "Gait disorders in adults and the elderly: A clinical guide," *Wien Klin Wochenschr*, vol. 129, no. 3–4, pp. 81–95, Feb 2017. [Online]. Available: <https://doi.org/10.1007/s00508-016-1096-4>
- [13] S. R. Simon, "Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems," *Journal of Biomechanics*, vol. 37, no. 12, pp. 1869–1880, 2004. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0021929004001228>
- [14] M. Chandrasekaran, J. Francik, and D. Makris, "Gait data augmentation using physics-based biomechanical simulation," *arXiv preprint arXiv:2307.08092*, 2023.
- [15] Y. Wang, Z. Li, X. Wang, H. Yu, W. Liao, and D. Arifoglu, "Human gait data augmentation and trajectory prediction for lower-limb rehabilitation robot control using gans and attention mechanism," *Machines*, vol. 9, no. 12, p. 367, 2021.
- [16] L. Ruff, "A unifying review of deep and shallow anomaly detection," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 756–795, 2021-05.
- [17] J. Otamendi, A. Zubizarreta, and E. Portillo, "Machine learning-based gait anomaly detection using a sensorized tip: an individualized approach," *Neural Comput Appl*, pp. 17 443–17 459, 2023-08.

- [18] D. Jani, V. Vijayakumar, R. Parmar, M. Bohara, D. Garg, A. Ganatra, and K. Kotecha, "An efficient gait abnormality detection method based on classification," *J. Sens. Actuator Networks*, vol. 11, no. 3, p. 31, 2022.
- [19] B. Gerazov, E. Hadjieva, A. Krivošei, F. I. S. Sanchez, J. Rostovski, A. Kuusik, and M. Alam, "Matrix profile based anomaly detection in streaming gait data for fall prevention," in *2023 30th International Conference on Systems, Signals and Image Processing (IWSSIP)*. IEEE, 2023, pp. 1–5.
- [20] N. Seliya, A. Abdollah Zadeh, and T. M. Khoshgoftaar, "A literature review on one-class classification and its potential applications in big data," *Journal of Big Data*, vol. 8, no. 1, pp. 1–31, 2021.
- [21] P. Hasiec, A. Świtoński, H. Josiński, and K. Wojciechowski, "Anomaly detection of motion capture data based on the autoencoder approach," in *International Conference on Computational Science*. Springer, 2023, pp. 611–622.
- [22] E. Duman and O. Erdem, "Anomaly detection in videos using optical flow and convolutional autoencoder," *IEEE Access*, vol. 7, pp. 183 914–183 923, 2019.
- [23] J. Singh, S. Jain, S. Arora, and U. Singh, "A survey of behavioral biometric gait recognition: Current success and future perspectives," *Archives of Computational Methods in Engineering*, vol. 28, pp. 107–148, 2021.
- [24] D. M. Ahmed and B. S. Mahmood, "Survey: Gait recognition based on image energy," in *2023 IEEE 14th Control and System Graduate Research Colloquium (ICSGRC)*. IEEE, 2023, pp. 51–56.
- [25] Y. Liu, X. He, R. Wang, Q. Teng, R. Hu, L. Qing, Z. Wang, X. He, B. Yin, Y. Mou *et al.*, "Application of machine vision in classifying gait frailty among older adults," *Frontiers in Aging Neuroscience*, vol. 13, p. 757823, 2021.
- [26] C. G. D. Santos, "Gait recognition based on deep learning: A survey," *ACM Computing Surveys*, vol. 55, no. 2, 2023-03-01.
- [27] S. Jain and A. Nandy, "Human gait abnormality detection using low cost sensor technology," in *Computer Vision and Image Processing: 5th International Conference, CVIP 2020, Prayagraj, India, December 4-6, 2020, Revised Selected Papers, Part II 5*. Springer, 2021, pp. 330–340.
- [28] S. Shakya, A. Taparugssanagorn, and C. Silpasuwanchai, "Convolutional neural network-based low-powered wearable smart device for gait abnormality detection," *IoT*, vol. 4, no. 2, pp. 57–77, 2023.
- [29] R. Bonetto, M. Soldan, A. Lanaro, S. Milani, and M. Rossi, "Seq2seq rnn based gait anomaly detection from smartphone acquired multimodal motion data," *arXiv preprint arXiv:1911.08608*, 2019.
- [30] M. Morshed, T. Sultana, A. Alam, and Y. Lee, "Human action recognition: A taxonomy-based survey, updates, and opportunities," *Sensors*, vol. 23, no. 4. MDPI, Feb. 01, 2023.
- [31] A. Alharthi, S. Yunas, and K. Ozanyan, "Deep learning for monitoring of human gait: A review," *IEEE Sens J*, vol. 19, no. 21, pp. 9575–9591, 2019-11.
- [32] E. Ramirez, M. Wimmer, and M. Atzmueller, "A computational framework for interpretable anomaly detection and classification of multivariate time series with application to human gait data analysis," in *Artificial Intelligence in Medicine: Knowledge Representation and Transparent and Explainable Systems: AIME 2019 International Workshops, KR4HC/ProHealth and TEAAM, Poznan, Poland, June 26–29, 2019, Revised Selected Papers*. Springer, 2019, pp. 132–147.
- [33] B. Chen, C. Chen, J. Hu, Z. Sayeed, J. Qi, H. F. Darwiche, B. E. Little, S. Lou, M. Darwish, C. Foote *et al.*, "Computer vision and machine learning-based gait pattern recognition for flat fall prediction," *Sensors*, vol. 22, no. 20, p. 7960, 2022.
- [34] S. W. Yahaya, A. Lotfi, and M. Mahmud, "Detecting anomaly and its sources in activities of daily living," *SN Computer Science*, vol. 2, no. 1, p. 14, 2021.
- [35] A. Aldayri and W. Albattah, "Taxonomy of anomaly detection techniques in crowd scenes," *Sensors*, vol. 22, no. 16, 2022-08-01.
- [36] A. H. M. Atallah and O. De Jesus, "Gait disturbances," Aug 23 2023, available from: <https://www.ncbi.nlm.nih.gov/books/NBK560610/>. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK560610/>
- [37] B. Gayal and S. Patil, "Detection and localization of anomalies in video surveillance using novel optimization based deep convolutional neural network," *Multimed Tools Appl*, 2023-08.
- [38] O. Malik, "Deep autoencoder for identification of abnormal gait patterns based on multimodal biosignals," *International Journal of Computing and Digital Systems*, vol. 10, no. 1, pp. 1–8, 2021.
- [39] K. Jun, D. Lee, K. Lee, S. Lee, and M. Kim, "Feature extraction using an rnn autoencoder for skeleton-based abnormal gait recognition," *IEEE Access*, vol. 8, pp. 19 196–19 207, 2020.
- [40] A. Ng, "Cs294a lecture notes: Sparse autoencoder," *Lecture Notes: Sparse Autoencoder, Stanford, CA, USA: Stanford Univ. Press*, 2010.
- [41] Z. Xue, D. Ming, W. Song, B. Wan, and S. Jin, "Infrared gait recognition based on wavelet transform and support vector machine," *Pattern Recognit*, vol. 43, no. 8, pp. 2904–2910, 2010-08.
- [42] D. Nam and B. Ahn, "Gait analysis accuracy difference with different dimensions of flexible capacitance sensors," *Sensors*, vol. 21, no. 16, 2021-08.
- [43] A. Naser, A. Lotfi, and J. Zhong, "Adaptive thermal sensor array placement for human segmentation and occupancy estimation," *IEEE Sens J*, vol. 21, no. 2, pp. 1993–2002, 2021-01.
- [44] M. Bouazizi, C. Ye, and T. Ohtsuki, "Low-resolution infrared array sensor for counting and localizing people indoors: When low end technology meets cutting edge deep learning techniques," *Information (Switzerland)*, vol. 13, no. 3, 2022-03.
- [45] J. Nogas, S. S. Khan, and A. Mihailidis, "Deepfall: Non-invasive fall detection with deep spatio-temporal convolutional autoencoders," *Journal of Healthcare Informatics Research*, vol. 4, pp. 50–70, 2020.
- [46] "EVB90640-41 user manual." available: [Online]. Available: <https://www.mouser.com/en/product/EVB90640-41/Evaluation-Board-MLX90640-MLX90641>
- [47] A. Naser, A. Lotfi, and J. Zhong, "Multiple thermal sensor array fusion toward enabling privacy-preserving human monitoring applications," *IEEE Internet of Things Journal*, vol. 9, no. 17, pp. 16 677–16 688, 2022.
- [48] I. The MathWorks. (2023, September) Multilevel image thresholds using otsu's method. [Online]. Available: <https://www.mathworks.com/help/images/ref/multithresh.html>
- [49] —. (2023, September) Quantize image using specified quantization levels and output values. [Online]. Available: <https://www.mathworks.com/help/images/ref/imquantize.html>
- [50] B. K. Horn and B. G. Schunck, "Determining optical flow," *Artificial intelligence*, vol. 17, no. 1-3, pp. 185–203, 1981.
- [51] A. Bruhn, J. Weickert, and C. Schnörr, "Lucas/kanade meets horn/schunck: Combining local and global optic flow methods," *International journal of computer vision*, vol. 61, pp. 211–231, 2005.
- [52] Y. Kim, J. Yu, E. Lee, and Y. Kim, "Video anomaly detection using cross u-net and cascade sliding window," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, pp. 3273–3284, 2022-06.
- [53] I. Golan and R. El-Yaniv, "Deep anomaly detection using geometric transformations," 2018-05.
- [54] S. Bulusu, B. Kailkhura, B. Li, P. Varshney, and D. Song, "Anomalous example detection in deep learning: A survey," *IEEE Access*, vol. 8, pp. 132 330–132 347, 2020.
- [55] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [56] K. M. Sunnetci, E. Kaba, F. B. Celiker, and A. Alkan, "Deep network-based comprehensive parotid gland tumor detection," *Academic Radiology*, vol. 31, no. 1, pp. 157–167, 2024.
- [57] J. Moon, N. A. Le, N. H. Minaya, and S.-I. Choi, "Multimodal few-shot learning for gait recognition," *Applied Sciences*, vol. 10, no. 21, p. 7619, 2020.



Farbod Zorriassatine is a research fellow at Nottingham Trent University, UK. He graduated from University of Birmingham (UK), with a Bachelor in Mechanical Engineering. He obtained his MPhil degree in Cellular Manufacturing, and Ph.D. degree in application of Novelty Detection for Improving Multivariate Process Quality both at the University of Nottingham, UK. Before joining Nottingham Trent University (UK), he had been a researcher at the University of Nottingham (UK) and Loughborough University (UK) and taught as a university lecturer in three countries. His research interest centres on anomaly detection and machine learning, with experience in novelty detection for Statistical Process Control (SPC) and applications ranging from healthcare to predictive maintenance.



Abdallah Naser obtained his BSc in Computer Engineering from Cyprus International University with High Honours and his MSc in Information Security and Biometrics with Distinction from the University of Kent. He received his PhD in Privacy-preserving Human Behaviour Monitoring through Thermal Vision from Nottingham Trent University, UK. He is currently an academic at Nottingham Trent University, where he is also a member of the Computational Intelligence and Applications research group. Abdallah was

granted a patent for inventing a new way to secure online accounts using concepts of pattern recognition from the Palestinian patent office. His current research focuses on computational intelligence, sensor signal processing, and image processing towards enabling privacy-preserving human-centred applications.



Ahmad Lotfi (M'96-SM'08) received his BSc. and MTech. in control systems from Isfahan University of Technology, Iran and Indian Institute of Technology Delhi, India respectively. He received his PhD degree in Learning Fuzzy Systems from University of Queensland, Australia in 1995. He is currently a Professor of Computational Intelligence at Nottingham Trent University, where he is leading the research group in Computational Intelligence and Applications.

His research focuses on the identification of progressive changes in behaviour of older individuals suffering from Dementia or any other cognitive impairments. Accurate identification of progressive changes through utilisation of unobtrusive sensor network or robotics platform will enable carers (formal and informal) to intervene when deemed necessary. Research collaboration is established with world-leading researchers. He has worked in collaboration with many healthcare commercial organisations and end-users. He has received external funding from Innovate UK, EU and industrial companies to support his research. He has authored and co-authored over 200 scientific papers in the area of computational intelligence, internet of things, abnormal behaviour recognition and ambient intelligence in highly prestigious journals and international conferences. He has been invited as an Expert Evaluator and Panel Member for many European and International Research Programmes.