# Action Invariant IMU-Gait for Continuous Authentication

Sanka Rasnayaka
School of Computing,
National University of Singapore
Email: sanka@nus.edu.sg

Terence Sim
School of Computing,
National University of Singapore
Email: tsim@comp.nus.edu.sg

## Abstract

*Continuous Authentication (CA) is proposed as an alternative authentication scheme for modern personal devices. Gait is a suitable biometric for CA due to its availability and low resource requirements. However, the drastic change in the gait pattern with changes in actions (such as walking, running or climbing stairs) and changes in terrain (such as walking on a flat surface or down an incline) makes it challenging to deploy in a real-world CA system.*

*We show that standard gait features are influenced by different actions. The gait pattern of an action is also influenced by the actions performed before and immediately after. Therefore, gait features are usually not robust to these action variations. We propose action invariant gait features to address this robustness issue. Our proposed method learns action invariant gait features utilizing a Siamese Network architecture with triplet loss and a unique triplet mining protocol. Our evaluations highlight that our action invariant features are robust to pre and post action impacts and real world action variations. These features allows for a CA system to be enrolled using a single action (walk) and be used across multiple different actions encountered throughout the day.*

## 1. Introduction

Recently there has been a rapid increase in personal devices such as mobile phones and wearable devices such as smart watches, fitness trackers and foot pods. This has enabled many security critical applications to be accessed through these devices. Banking, e-commerce, financial markets and health care are examples of security critical applications which are now accessible through these personal devices.

The increase in the usage and the subsequent increase in the required security level has highlighted the inadequateness of traditional authentication schemes and the requirement of Continuous Authentication [25]. CA systems rely on passive biometrics to repeatedly authenticate the user. This process allows the mobile device to determine the continued presence of the legitimate user, allowing CA to provide a digital identity at any point in time.

One such passive biometric is Gait [17, 1]. We define gait as the movement pattern of a person's ambulation. The gait biometric can be divided into two broad areas based on how it is captured: (1) Visual gait and (2) On-body gait. Figure 1 compares the standard processing pipeline of on-body gait with visual gait. In the scope of this work, we are only interested in **on-body gait** to which we will simply refer as *gait* from here on-wards.

Gait can be captured using Inertial Measurement Units (IMU) available in modern devices. IMUs are ideal for continuous usage due to their low demand on resources like power and memory [24]. Therefore, more data can be logged while the user wears or carries devices with IMUs in their daily life. This data has engendered many interesting applications such as activity recognition, step tracking, health monitoring, gait-based authentication and Continuous Authentication [14, 22].

Most of the existing work on gait focus on the walk pattern of a person. However, in a typical day, a person would go through different physical activities, such as walking, running, climbing stairs, etc. We will refer to these physical activities as *actions* in the context of this paper. Moreover, there can be a residual effect from the previous action and an anticipatory effect from the next action towards a persons walk.

A CA system should be robust to these variations and should be able to enroll a new user with only one or a few enrollment actions. To this end, we propose an action invariant feature space which enables the system to learn person specific features using a single action enrollment.

Key contributions of this work:

- We highlight the effect of different actions and pre and post action variations towards a gait pattern. A gait system evaluated on curated datasets lead to an over-estimate of accuracy values which are not attainable in a real-world scenario.

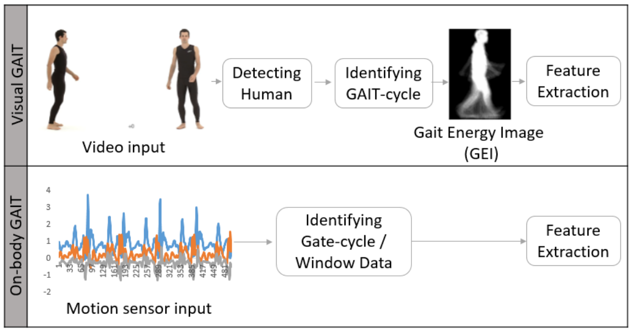



Figure 1. Visual gait vs On-body gait

- We propose an action invariant feature space which is able to differentiate between users regardless of actions performed.
- Our extensive evaluations using traditional accuracy measurements as well as CA specific measures showcase the robustness and invariability of the learnt features. Our evaluations also look into single action enrollment, as well as imposter detection efficacy.

## 2. Related Work

### 2.1. On-body Gait Biometric

The prominent approaches for on-body gait is summarized here.

**Hand Crafted Features** looks at different phases in a gait cycle (swing, stance). Characteristics at each phase are used as features for gait biometric. [31, 11, 6]

**Statistical Features** focus on using windowed cropped data or gait cycle data to extract statistical features like mean, median, max, range and variance values as features. [2, 12, 15, 30]

**Fourier Transform** considers the gait pattern as a continuous signal to obtain its dominant frequency components as features. [32, 17]

**Convolutional Neural Networks** are utilized by converting gait signals into an image. Various different ways for converting the time-series signal into an image has been proposed with differing success. [5, 10, 34, 21]

**Recurrent Neural Networks(RNNs)/Long Short-Term Memory (LSTM)** is the current state-of-the-art. These models are able to extract temporal features from the gait signal. [8, 37, 33]

In our work we utilize LSTM features in our final method. We also implement statistical feature based methods as a comparison of prior methods.

While most of these works show high accuracies for gait as a biometric there are few shortcomings when we consider gait for Continuous Authentication.

- *Limited real world actions:* Most studies work with the subjects performing a walk. While few studies look at other actions like jogging and stairs [15] or speed variations of the walk [17].
- *Curated actions:* Most studies are conducted by participants performing the actions in a curated manner. They would start the action from a stand still and end in a stand still. However, in the real world a person would be performing different actions before and after each action. There is a significant influence from these pre and post actions which affects the gait pattern.
- *No common dataset:* Most of the gait studies are done using independent datasets which make comparisons between methods difficult.

| | |
|---|---|
| Number of IDs | 53 |
| Actions | walk, run, incline up/down, stairs up/down |
| Sensors | 3 Accelerometers at 100Hz |
| Sensor Locations | Foot L/R, Pocket L/R, Wrist L/R, Back Pocket, Bag, Phone call |
| Geneder | 26 female, 27 male |
| Age | $\mu = 33yrs, \sigma = 10.43$ |

Table 1. NUS-IMU Gait Dataset Summary

In this study we address these shortcomings. The NUS-IMU Gait Dataset [26] (NUS-IGD) has multiple real-world actions performed in a sequence. We analyze the effect of pre and post actions using the continuous data-stream available in this dataset. Since this dataset is publicly available it can be used as a benchmark for future gait studies and comparisons. We evaluate our action invariant feature space against traditional statistical feature method as well as LSTM feature method inspired from previous work.

### 2.2. Continuous Authentication

Initial research on CA was started during desktop computer era by use of biometrics like mouse dynamics [28] and key stroke dynamics [3]. As the sensor rich personal devices came to be, CA research also took advantage of this shift by incorporating novel biometric modalities such as touch gestures [9, 36, 35], face [27, 4], voice [16] and gait [22, 7, 23].

As the resources available in personal devices like computing power and memory grew, Continuous Authentication research looked at combining multiple different modalities to increase security of the CA systems [29, 19, 14, 13].

However, the constant increase in modalities to achieve higher accuracies is not sustainable in real-world due to the limitations in resources such as battery life. The trade-off between these resources and the authentication accuracies was studied by Rasnayaka et. al [24] which highlights the importance of limiting resource consumption while maintaining an acceptable level of security. The results shows the suitability of gait due to its low resource strain. Therefore, we will be focusing on gait for Continuous Authentication in our work.

Many biometric research studies focus on False Acceptance Rate (FAR), False Reject Rate (FRR) and Equal Error Rate (EER) as evaluation metrics. However, when evaluating a CA system, the time aspect should also be considered. The time taken by the system to detect an imposter is a key security concern, similarly if the CA system logs out the legitimate user, that is a usability concern. To measure these time aspects in CA systems novel evaluation metrics have been proposed.

- Time to Correct Reject (TCR) [29]
- Usability [29]
- False Accept Worse Interval (FAWI) [20]
- False Reject Worse Interval (FRWI) [20]

We will make use of the traditional biometric authentication metrics as well as these novel CA authentication metrics to give a comprehensive performance evaluation of our proposed feature space.

## 3. Methodology

The main focus in this work is the use of gait for Continuous Authentication. Therefore, the task is verification of a legitimate user.

Unlike hard biometrics such as fingerprint, soft/behavioural biometrics such as gait changes drastically depending on the context. When the users behaviour changes from walking to running or climbing stairs the gait pattern will change drastically. Therefore, a residual effect from the pre action and an anticipatory effect from the post action is possible. The intuition is, a person walking soon after he was running a few steps might walk differently compared to a person walking from a stand-still due to the residual effect. We aim to analyze the level of impact this pre and post actions might have on gait features.

Next, we will study how gait performs with different actions by analyzing how well a model trained in one action generalizes for other actions. These will highlight the robustness issues arising due to the high variability of gait biometric in a real world setting. We propose action invariant features for a CA scenario to have a robust authentication system, where we cannot expect the user to enroll all possible actions.

### 3.1. Datasets

We make use of two datasets in this study.

**1. NUS IMU Gait Dataset (NUS-IGD) [26]:** A summary of the NUS IMU gait dataset is available in Table. 1. This dataset consists of gait data of 53 participants performing multiple actions and terrain variations (walk, run, incline-down walk, incline-up walk, climb upstairs and climb downstairs) as well as multiple sensor location variations (wrist, front pocket, foot, back pocket, phone call and side bag). The gait signals are three data-streams from the x, y, z axis of accelerometer values captured at 100 Hz. An analysis on privacy invasiveness along with more details about NUS-IGD is published by Rasnayaka et. al [26].

**2. Cross Device Compatibility Validation Dataset:** A small scale validation dataset with four devices namely an Android phone, an iPhone, EarPods and a smart watch was collected. The two phones were carried in the front pocket, the EarPods were worn on the ears and the smart watch was worn on the wrist to simulate there real world sensor locations. The data was collected in June of 2021 with 7 participants. The participants performed a level ground walk and two walk segments were recorded.

The iPhone and the smart watch records data at 100 Hz which is the same sampling rate as the data in NUS IMU gait dataset. The Android phone records data at 256 Hz, which will be resampled at 100 Hz before use. The EarPods record data at 20 Hz.

We make use of this dataset as a validation dataset to see if our model trained on the Axis AXIVITY AX3 data from NUS-IGD can be used without further fine-tuning for different devices.

### 3.2. Data preprocessing

Each input dataset stream from these datasets consists of x, y, z axis of accelerometer values. We use a 30ms window to create input data windows of size (30, 3). We use the raw acceleration values and do not perform any pre-processing on the data. In order to study each action type in isolation we will use the action based crops available in the NUS-IGD dataset. When we want to consider all the actions as a continuous stream we will use the raw data stream which is available in both datasets.

We will focus on one sensor location (left_foot) from the NUS-IGD for this study.

### 3.3. Siamese Architecture

The idea is to force the clusters of the same person with different actions to be pushed together while the clusters are pushed away from different people regardless of the action.

The Siamese model is created with a shared weight sister network to extract features from a triplet of samples. The triplet is of the form (Anchor, Positive, Negative).

Fig. 2 shows the Siamese Network architecture used. Fig. 3 shows the sister network architecture. We use a layered Long Short-Term Memory (LSTM) architecture for the

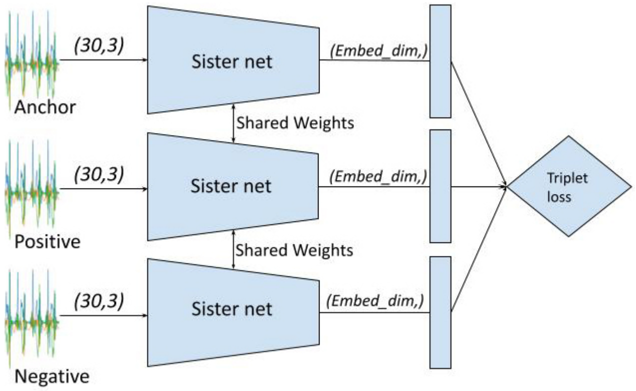



Figure 2. Siamese Network Architecture

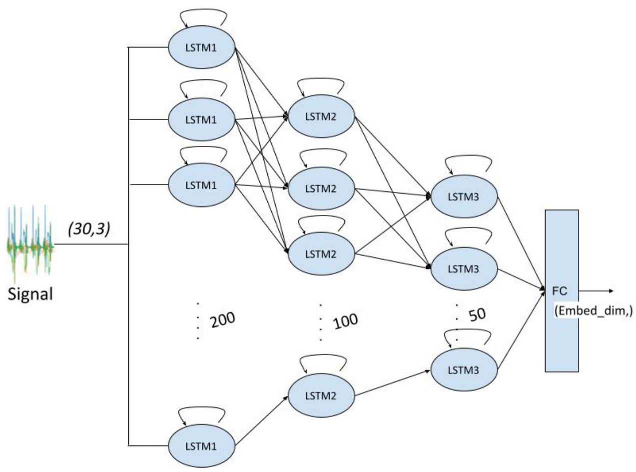



Figure 3. Sister Network Architecture

sister network. We use tanh activation and 48 dimensions as our embedding feature size.

### 3.3.1 Triplet loss

The triplet loss function is used to train the network which will push the positive sample towards the anchor while pushing the negative sample away. Triplet loss is calculated from the feature vectors of Anchor, Positive and Negative samples as shown below,

Let, $f_A, f_B, f_C$ = features of anchor, positive, negative

Triplet Loss(Anchor, Positive, Negative) =

$$\max\{(\|f_A - f_P\|^2 - \|f_A - f_N\|^2 + \alpha), 0\} \quad (1)$$

We use $\alpha = 0.2$ for the triplet loss calculation.

### 3.3.2 Offline triplet mining protocol

To learn action invariant gait features we need to ensure a proper distribution of actions in the triplets. We propose the following triplet mining scheme to achieve this.

| Positive | Negative | Samples |
|---|---|---|
| $p_i, a_y$ | $p_j(i \neq j), a_y$ | $\{\forall a_y \in A \vert a_y \neq a_x\}$ (5 triplets) |
| $p_i, a'_x$ (diff. window) | $p_j(i \neq j), a_y$ | $\{\forall a_y \in A \vert a_y \neq a_x\}$ (5 triplets) |
| $p_i, a_y$ | $p_j(i \neq j), a_x$ | $\{\forall a_y \in A \vert a_y \neq a_x\}$ (5 triplets) |
| $p_i, a'_x$ (diff. window) | $p_j(i \neq j), a_x$ | (1 triplet) |

Table 2. Triplet mining for anchor $(p_i, a_x)$

Actions, $A = \{$walk, run, incline down, incline up, stairs down, stairs up$\}$

We denote an anchor data window as $(p_i, a_x)$, where $p_i$ - $i$th person, $a_x \in A$ - the action performed in the window. Note, there are multiple windows for the same $(p_i, a_x)$ pair.

For every triplet, the positive sample should be from the same person, and the negative sample should be from a different person. While ensuring these, we combine same and different actions. When ever we need the same action we use a new window of the same action, for different actions we have a choice from set $A$, which results in 5 different options. This protocol shown in Table 2. By using this protocol we create 16 triplets for each anchor.

## 3.4. Experimental Setup

To simulate verification we first divide the dataset into training and testing (43-10). Next, the training set is used to train the feature extractor. The testing set is again split for each person into enrollment and verification in time axis. We use enrollment window = $min\{$400 windows, total time/2$\}$ to determine the enrollment time.

We conduct the following experiments,

**1. Isolated Actions:** We train individual LSTM model inspired by previous studies for each action to simulate a scenario where all actions are known beforehand. We evaluate the performance of individual action models with our triplet model.

**2. Pre and Post Actions:** We train individual LSTM model and statistical feature based models inspired by previous work using walk data which starts and ends from a stand still. Then we evaluate these models with walk data with different pre and post actions. This highlights how the features used in previous work are susceptible to pre and post action variations.

**3. Continuous Authentication:** We use the entire data stream with multiple action to simulate a CA scenario. In this scenario all people will be enrolled using the first 400 windows. Which means all the people will be enrolled using a walk action. The windows following the enrollment period will be used for testing. This is illustrated in Fig. 4.

We will evaluate a standard LSTM model with the proposed features.

### 3.5. Evaluation

We calculate the average feature vector for the enrollment period $f_{enroll}$ and use the euclidean distance between the enrollment feature and test feature for authentication decisions as follows,

$$\text{if}(\|f_{enroll} - f_{test}\|^2 \leq threshold): \quad \text{Accept as legitimate user} \tag{2}$$

We use the False Acceptance Rate (FAR)/False Reject Rate (FRR) and Equal Error Rate (EER) calculations for basic comparisons with standard feature methods. This allows us to show that our method shows comparable results in standard classification settings.

Next, we focus on Continuous Authentication specific measures.

- **Usability [29]:** We calculate the usability ($U_{p_i}$) of person $p_i$ by looking at the complete input as,

$$\text{U}_{p_i} = \frac{\text{windows where } (distance \leq threshold)}{\text{total number of windows}} \tag{3}$$

  The overall usability of the model is calculated as the mean of $U_{p_i}$ for all the test subjects.

- **Time to Correct Reject (TCR) [29]:** We simulate a sudden switch from the legitimate user to an imposter by concatenating two walk patterns from $p_l$ (legitimate) and $p_i$ (imposter), this is illustrated in Fig. 5.

  The red vertical line represents the time $t_s$ where the switch to the imposter happens. TCR for this legitimate and imposter pair is calculated as follows,

$$TCR(p_l, p_i) = t_{reject} - t_s \tag{4}$$

  Here $t_{reject}$ is the first timestamp where the imposter is rejected access to the system.

  We have nine imposters for each test case ($p_l$). By averaging across all imposters we measure the $TCR_{p_l}$ (TCR for person $p_l$). The overall TCR of the model is calculated as the mean of $TCR_{p_l}$ values for all the test subjects.

- **False Reject Worse Interval (FRWI) [20]:** The input and enrollment is done similar to what is shown in Fig. 4. FRWI for each test subject ($p_l$) is,

$$FRWI_{p_l} = \text{longest time window where } (\|f_{enroll} - f_{test}\|^2 > threshold) \tag{5}$$

  Overall FRWI for the model is calculated by taking the mean over all test subjects.

- **False Accept Worse Interval (FAWI) [20]:** We calculate the enrollment feature of person $p_i$ using the first 400 windows of his walk data. Then another user $p_l$'s complete data stream is used for measuring FAWI between the $p_i$ and $p_l$ as follows,

$$FAWI(p_l, p_i) = \text{longest time window where } (\|f_{pi-enroll} - f_{pl-test}\|^2 \leq threshold) \tag{6}$$

  By averaging across all $p_i$ we measure the $FAWI_{p_l}$ (FAWI for person $p_l$). The overall FAWI of the model is calculated as the mean of $FAWI_{p_l}$ values for all the test subjects.

## 4. Results

First we train a set of specialized models for each sensor location and action pair. We use the LSTM architecture proposed in [26]. We measure the FAR and FRR for each LSTM model and report (1 - EER) in the Table 3. We can see the performance is very high with every sensor location and action pair performing over 80%. We can see that if each action and sensor location is known before hand we can provide a higher level of accuracy.

Next, we focus on the impact different actions and sensor locations will have towards this accuracy level. We will highlight the results for these individual LSTM models as well as our action invariant triplet loss model.

### 4.1. Isolated Actions

We first compare the proposed triplet loss model with a set of specialized models trained for each action separately.

We train a different model for each action using the LSTM architecture proposed in [26]. We measure the FAR and FRR for each LSTM model and report (1 - EER). The 6 different models trained for each action is compared with the single triplet loss model in Table. 4

The triplet loss model shows comparable results with the specialized LSTM models trained on each action. The triplet loss model out performs the individual LSTMs for stairs up and down actions.

While this analysis provides an upper bound accuracy level, it is unrealistic to expect each action to be labeled and

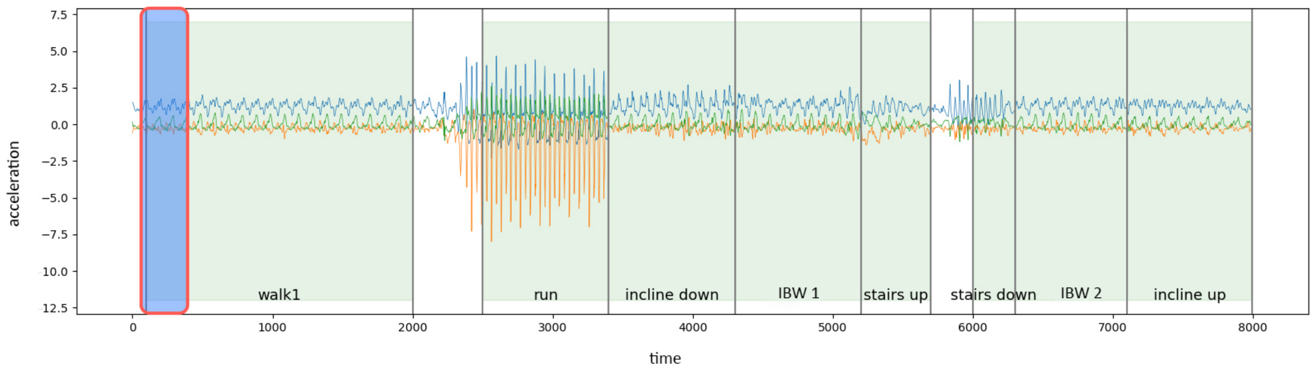



Figure 4. Complete input stream, the enrollment phase is highlighted in the blue box

| | Pocket | Foot | Wrist | Back Pocket | Bag | Phone Call |
|---|---|---|---|---|---|---|
| **Walk** | 0.906 | 0.956 | 1.0 | 0.989 | 0.956 | 0.978 |
| **Run** | 0.900 | 0.922 | 0.994 | - | - | - |
| **Incline Up** | 0.950 | 0.994 | 0.994 | - | - | - |
| **Incline Down** | 0.950 | 0.983 | 0.978 | - | - | - |
| **Stairs Up** | 0.933 | 0.833 | 0.956 | - | - | - |
| **Stairs Down** | 0.883 | 0.889 | 0.900 | - | - | - |

Table 3. Verification results for different actions and sensor locations (accuracy values reported)

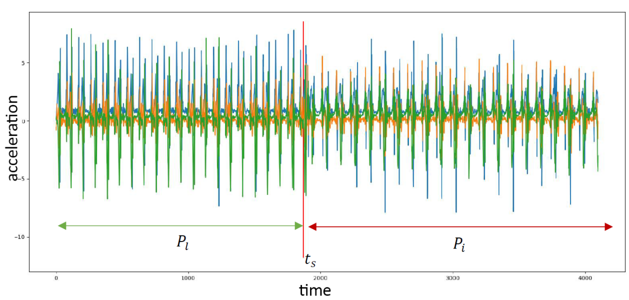



Figure 5. Simulated legitimate user to imposter switch

| Action | 6 individual LSTMs | Triplet loss model |
|---|---|---|
| Walk | **0.956** | 0.938 |
| Run | **0.922** | 0.894 |
| Incline Up | **0.994** | 0.961 |
| Incline Down | **0.983** | 0.933 |
| Stairs Up | 0.833 | **0.950** |
| Staris Down | 0.899 | **0.916** |

Table 4. Comparison of (1-EER) between individual LSTMs for each action and the triplet loss model. (Higher is better)

enrolled in a real world use case. Therefore, we focus on training a single model which learns action invariant personal features.

### 4.2. Pre and Post Actions

We train a similar LSTM model and a statistical feature based model from [26] using the initial walk data. The initial walk starts and ends at a stand still.

| Method | Walk | IBW 1 | IBW 2 |
|---|---|---|---|
| Stat | 0.937 | 0.325 | 0.660 |
| LSTM classifier | **0.955** | 0.348 | 0.680 |
| Proposed method | 0.938 | **0.933** | **0.938** |

Table 5. Comparison of (1-EER) between individual LSTMs for each action and the triplet loss model. (Higher is better)

Next we evaluate the models using two other walks available in the NUS-IGD. Namely,

1. In-between walk 1 (IBW 1) - starts after walking down an incline, finishes at the beginning of a stair climb.
2. In-between walk 2 (IBW 2) - starts after climbing down stairs, finishes at the beginning of an incline up walk.

In this evaluation setting, the only difference between the training set and the two evaluation sets is the pre and post actions, while the current action and terrain is fixed to a walk on level ground.

Table 5 summarizes the results obtained by the standard methods and the triplet loss method.

Looking at the results, even though all three actions are walking on a level floor, when the pre action and post action are different we can see the gait performance of the stat model and LSTM model drops drastically.

Table 5 shows that both IBW 1 and IBW 2 accuracy drops drastically for stat and LSTM methods ($\sim$30% and

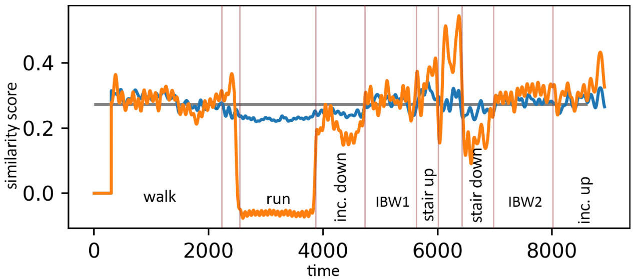


Figure 6. Similarity score throughout the entire data stream. Blue - proposed triplet model, Orange - contrastive model. (Lower variance is better)

$\sim$65%). However, Triplet method maintains a similar accuracy ($\sim$93%). This shows that triplet features show better intra-person separation which is robust to different pre and post action influences.

IBW 1 shows a higher drop than IBW 2 for both stat and LSTM methods. While the normal walk action is performed starting from a standstill, IBW 1 is started after walking down an incline, therefore there is momentum carried over which increases the persons speed and cadence. However, IBW starts after climbing down a stair case, therefore people tend to pause after climbing down the stairs and re-start their walk. Which is more similar to starting from a stand still. Therefore, IBW 1 shows a greater dip than IBW 2.

### 4.3. Continuous Authentication

For comparison purposes we train another Siamese network with the same sister network architecture but with a contrastive loss function and random test pairs. This will be used as a baseline for comparison.

The euclidean distance measure is converted to a similarity score ranging from -1 to 1. The objective of a Continuous Authentication system is to maintain a constant value for the similarity score regardless of the action. Which means a lower variation in the similarity score is desired.

Fig. 6 shows the similarity score throughout the entire walk for person id 53. We can clearly see how the triplet model maintains a lower variance on the similarity scores ($var = 0.38$). Where as the contrastive model show high variance ($var = 2.45$) specially when the actions change.

There is still a significant dip in similarity score for the triplet model when the action changes to a run. The change from walk to run is quite drastic making it harder for the model to extract person identification features which are common across these two actions. For all other actions the triplet model has been able to maintain a low variance in the similarity score.

This study shows how our triplet model is able to extract action invariant features. Further studies using CA evaluation matrices help reinforce this claim.

Table 6 shows the Usability, TCR, FRWI and FAWI for both triplet and contrastive models for comparison.

| Evaluation Metric | Contrastive | Triplet Model |
|---|---|---|
| Usability | 70.63% | **76.33%** |
| TCR | 35 ms | **30.8 ms** |
| FRWI | 3.7s | **2.8s** |
| FAWI | 28s | **15s** |

Table 6. Continuous Authentication Evaluation metrices. (Higher Usability scores are better; while lower TCR, FRWI, FAWI are better)

| Enrollment Action | Usability Score |
|---|---|
| Walk | 76.33 |
| Run | 66.04 |
| Incline Down | 82.75 |
| Incline Up | 74.06 |
| Stairs Down | 61.09 |
| Stairs Up | 70.23 |

Table 7. Enrolling with different actions. (Higher is better)

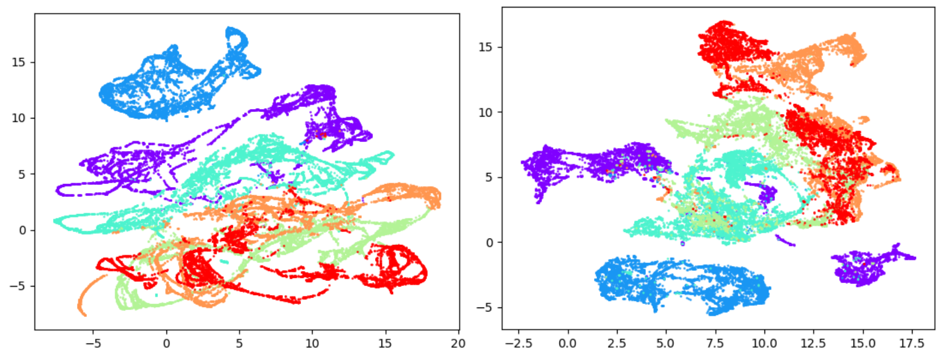
We can see that the proposed triplet model out performs the standard method for all measurements. The usability score improvement indicates that the triplet score method allows the legitimate user to be logged into the system with less false rejects. The TCR shows that an imposter will be locked out of the system within 30.8 ms when the triplet model is used, which is an improvement over the 35 ms in the contrastive model. This improvement is due to the better intra-person separation in the triplet feature space.

The FRWI and FAWI both are lower in the triplet model, however the window lengths are quite high (3s - 40s). The main reason is that, this measure is a worst case measure. The test scenarios where the model has failed will dominate the final score in FRWI and FAWI.

There is a trade-off between Usability and TCR and similarly a trade-off between FRWI and FAWI which can be adjust by changing the threshold value depending on the application scenario.

### 4.4. Single action enrollment

We study how different actions perform when used as the enrollment action. Table 7 shows how the performance changes when the enrollment action is changed. For this experiment we have changed the enrollment action and the usability score is calculated for the entire data stream with all available actions. The usability scores are above 70% for most actions while it drops lower for stairs down and run actions. This indicates our features show low inter-action variation for the same person. Thus, for easier enrollment in a practical CA system, walking is preferred.



(a) Contrastive Model (b) Triplet Model

Figure 7. UMAP projected features for 6 people. Different colors represent different people (all their actions combined). The Triplet model shows greater separability among the colors.

| | EarPod | AppleWatch | iPhone | Android Phone |
|---|---|---|---|---|
| EER | 0.22 | 0.08 | 0.01 | 0.08 |

Table 8. Cross Device Evaluation Results

### 4.5. Action Invariance of the Proposed Features

To analyze the action invariance of the proposed feature space, we use a Uniform Manifold Approximation and Projection (UMAP) [18] for dimension reduction and plot the clustering in a 2D space. We aim to show the following characteristics of the action invariant features,

- Low inter-action variation for the same person
- High intra-person separation

Figure 7 shows the features for 6 different people, represented by different colors. All the actions for each person is combined together. We can see how the triplet model achieves a better person separation. This can be numerically re-enforced using the silhouette scores. (contrastive - 0.113, triplet 0.127). The tighter clusters provided by the proposed triplet loss model highlights how the triplet model learns action invariant person features with high intra-person separation.

### 4.6. Cross Device Compatibility

In this section we aim to evaluate our model on its robustness to changes in the personal device. We make use of the cross device compatibility validation data we have collected. We use the previously trained model and enroll each user by using a portion of data from the walk segment. Next, we use the remaining walk data and test against all possible imposters. In our case we have 6 imposter tests for each legitimate user.

Table 8 shows the final Equal Error Rates obtained by this experiment. It is clear that Apple Watch, iPhone and Android Phone have all performed with a high level of accuracy with the ERR being below 0.01. EarPod shows a relatively higher EER at 0.22. This is expected as the EarPods record the IMU data at a much lower frequency compared to our original dataset.

## 5. Conclusions and Future Work

We highlight the major variation in gait pattern with changes in actions and terrain. Our work is also the first to highlight the impact of pre and post action variations toward gait. Our evaluations show that current studies are susceptible to these variations, making them less robust in real world applications. These findings highlight the requirement of un-curated datasets and real world testing.

We introduce an action invariant feature space for gait, which can be used in a Continuous Authentication system with single action enrollment (like walk). The system can then be used to authenticate any action the user performs thereafter. This makes it practical to deploy in a CA system.

Evaluation of these features highlighted the robustness to changes in actions as well as pre and post action variations. The proposed features out performed standard statistical feature methods as well as a standard LSTM based methods in a real world dataset. The features learnt using the proposed triplet loss model also out performed in CA based usability and security measures.

Our cross device evaluations highlight that the features trained using the Axis AXIVITY AX3 sensor can be generalized to other device hardware. Therefore, the proposed features can be used to build CA systems that can be deployed in the real-world.

The results in our work can be reproduced and evaluated since the NUS-IGD is a publicly available dataset. Which will enable various extensions of this work. As an example, the triplet mining protocol proposed in our work can be adopted to learn features invariant to action as well as sensor location. That will be useful in a scenario where the device location with respect to the user can also change dynamically.

**Acknowledgement**

This research is supported by the National Research Foundation, Singapore under its Strategic Capability Research Centres Funding Initiative. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore.

## References

[1] H. J. Ailisto, M. Lindholm, J. Mantyjarvi, E. Vildjiounaite, and S.-M. Makela. Identifying people from gait pattern with accelerometers. In *Biometric Technology for Human Identification II*, volume 5779, pages 7–14. International Society for Optics and Photonics, 2005.

[2] A. Annadhorai, E. Guenterberg, J. Barnes, K. Haraga, and R. Jafari. Human identification by gait analysis. In *Proceedings of the 2nd International Workshop on Systems and Networking Support for Health Care and Assisted Living Environments*, pages 1–3, 2008.

[3] P. Bours. Continuous keystroke dynamics: A different perspective towards biometric evaluation. *Information Security Technical Report*, 17(1):36 – 43, 2012. Human Factors and Bio-metrics.

[4] D. Crouse, H. Han, D. Chandra, B. Barbello, and A. K. Jain. Continuous authentication of mobile user: Fusion of face image and inertial measurement unit data. In *2015 International Conference on Biometrics (ICB)*, pages 135–142, May 2015.

[5] O. Dehzangi, M. Taherisadr, and R. ChangalVala. Imu-based gait recognition using convolutional neural networks and multi-sensor fusion. *Sensors*, 17(12):2735, 2017.

[6] M. O. Derawi, P. Bours, and K. Holien. Improved cycle detection for accelerometer based gait authentication. In *2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 312–317, Oct 2010.

[7] M. O. Derawi, C. Nickel, P. Bours, and C. Busch. Unobtrusive user-authentication on mobile phones using biometric gait recognition. In *2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 306–311, Oct 2010.

[8] P. Fernandez-Lopez, J. Liu-Jimenez, K. Kiyokawa, Y. Wu, and R. Sanchez-Reillo. Recurrent neural network for inertial gait user recognition in smartphones. *Sensors*, 19(18):4054, 2019.

[9] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. *Trans. Info. For. Sec.*, 8(1):136–148, Jan. 2013.

[10] M. Gadaleta and M. Rossi. Idnet: Smartphone-based gait recognition with convolutional neural networks. *Pattern Recognition*, 74:25–37, 2018.

[11] D. Gafurov, E. Snekkenes, and P. Bours. Improved gait recognition performance using cycle matching. In *2010 IEEE 24th International Conference on Advanced Information Networking and Applications Workshops*, pages 836–841, April 2010.

[12] Z. Huan, X. Chen, S. Lv, and H. Geng. Gait recognition of acceleration sensor for smart phone based on multiple classifier fusion. *Mathematical Problems in Engineering*, 2019, 2019.

[13] R. Johnson, R. Murmuria, A. Stavrou, and V. Sritapan. Pairing continuous authentication with proactive platform hardening. In *Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on*, pages 88–90. IEEE, 2017.

[14] R. Kumar, V. V. Phoha, and A. Serwadda. Continuous authentication of smartphone users by fusing typing, swiping, and phone movement patterns. In *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–8. IEEE, 2016.

[15] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Cell phone-based biometric identification. In *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–7. IEEE, 2010.

[16] H. Lu, A. B. Brush, B. Priyantha, A. K. Karlson, and J. Liu. Speakersense: Energy efficient unobtrusive speaker identification on mobile phones. In *International conference on pervasive computing*, pages 188–205. Springer, 2011.

[17] J. Mantyjarvi, M. Lindholm, E. Vildjiounaite, S.-M. Makela, and H. Ailisto. Identifying users of portable devices from gait pattern with accelerometers. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 2, pages ii–973. IEEE, 2005.

[18] L. McInnes, J. Healy, and J. Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[19] S. Mondal and P. Bours. A study on continuous authentication using a combination of keystroke and mouse biometrics. *Neurocomputing*, 230:1–22, 2017.

[20] A. Mosenia, S. Sur-Kolay, A. Raghunathan, and N. K. Jha. Caba: continuous authentication based on bioaura. *IEEE Transactions on Computers*, 66(5):759–772, 2017.

[21] K.-T. Nguyen, T.-L. Vo-Tran, D.-T. Dinh, and M.-T. Tran. Gait recognition with multi-region size convolutional neural network for authentication with wearable sensors. In *International Conference on Future Data and Security Engineering*, pages 197–212. Springer, 2017.

[22] I. Papavasileiou, S. Smith, J. Bi, and S. Han. Gait-based continuous authentication using multimodal learning. In *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pages 290–291. IEEE, 2017.

[23] A. Primo, V. V. Phoha, R. Kumar, and A. Serwadda. Context-aware active authentication using smartphone accelerometer measurements. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on*, pages 98–105. IEEE, 2014.

[24] S. Rasnayaka, S. Saha, and T. Sim. Making the most of what you have! profiling biometric authentication on mobile devices. In *2019 International Conference on Biometrics (ICB)*, pages 1–7, 2019.

[25] S. Rasnayaka and T. Sim. Towards wider adoption of continuous authentication on mobile devices. In T. Bourlai,

P. Karampelas, and V. M. Patel, editors, *Securing Social Identity in Mobile Platforms: Technologies for Security, Privacy and Identity Management*, pages 235–262, Cham, 2020. Springer International Publishing.

[26] S. Rasnayaka and T. Sim. Your tattletale gait. In *International Joint Conference on Biometrics (IJCB)*, 2020.

[27] P. Samangouei, V. M. Patel, and R. Chellappa. Attribute-based continuous user authentication on mobile devices. In *2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–8, Sept 2015.

[28] C. Shen, Z. Cai, and X. Guan. Continuous authentication for mouse dynamics: A pattern-growth approach. In *Dependable Systems and Networks (DSN), 2012 42nd Annual IEEE/IFIP International Conference on*, pages 1–12. IEEE, 2012.

[29] T. Sim, S. Zhang, R. Janakiraman, and S. Kumar. Continuous verification using multimodal biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):687–700, April 2007.

[30] S. Sprager and M. B. Juric. An efficient hos-based gait authentication of accelerometer data. *IEEE transactions on information forensics and security*, 10(7):1486–1498, 2015.

[31] S. Terada, Y. Enomoto, D. Hanawa, and K. Oguchi. Performance of gait authentication using an acceleration sensor. In *2011 34th International Conference on Telecommunications and Signal Processing (TSP)*, pages 34–36. IEEE, 2011.

[32] H. M. Thang, V. Q. Viet, N. D. Thuc, and D. Choi. Gait identification using accelerometer on mobile phone. In *2012 International Conference on Control, Automation and Information Sciences (ICCAIS)*, pages 344–348. IEEE, 2012.

[33] L. Tran, T. Hoang, T. Nguyen, H. Kim, and D. Choi. Multi-model long short-term memory network for gait recognition using window-based data segment. *IEEE Access*, 9:23826–23839, 2021.

[34] W. Yuan and L. Zhang. Gait classification and identity authentication using cnn. In *Asian Simulation Conference*, pages 119–128. Springer, 2018.

[35] H. Zhang, V. M. Patel, M. Fathy, and R. Chellappa. Touch gesture-based active user authentication using dictionaries. In *Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision*, WACV '15, pages 207–214, Washington, DC, USA, 2015. IEEE Computer Society.

[36] X. Zhao, T. Feng, and W. Shi. Continuous mobile authentication using a novel graphic touch gesture feature. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–6, Sept 2013.

[37] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, and Q. Li. Deep learning-based gait recognition using smartphones in the wild. *IEEE Transactions on Information Forensics and Security*, 15:3197–3212, 2020.