Neural Computing and Applications (2024) 36:3817–3836
https://doi.org/10.1007/s00521-023-09390-3

ORIGINAL ARTICLE

Biometric systems for identification and verification scenarios using spatial footsteps components

Ayman Iskandar[1] · Marco Alfonse[1,2] · Mohamed Roushdy[3] · El-Sayed M. El-Horbaty[1]

Received: 17 August 2023 / Accepted: 13 December 2023 / Published online: 13 January 2024

Abstract

Humans are distinguished by their walking patterns; many approaches, including using various types of sensors, have been used to establish walking patterns as biometrics. By studying the distinguishing features of a person's footsteps, footstep recognition may be utilized in numerous security applications, such as managing access in protected locations or giving an additional layer of biometric verification for secure admittance into restricted regions. We proposed biometric systems for verifying and identifying a person by acquiring spatial foot pressure images from the values obtained from the piezoelectric sensors using the Swansea Foot Biometric Database, which contains 19,980 footstep signals from 127 users and is the most prominent open-source gait database available for footstep recognition. The images acquired are fed into the ConvNeXt model, which was trained using the transfer learning technique, using 16 stride footstep signals in each batch with an Adam optimizer and a learning rate of 0.0001, and using sparse categorical cross-entropy as the loss function. The proposed ConvNeXt model has been adjusted to acquire 512 feature vectors per image, and these feature vectors are used to train the logistic regression models. We propose two biometric systems. The first biometric system is based on training one logistic regression model as a classifier to identify 40 different users using 1600 signals for training, 6697 signals for validation, and 200 signals for evaluation. The second biometric system is based on training 40 logistic regression models, one for each user, to validate the user's authenticity, with a total number of 2363 training signals, 7077 validation signals, and 500 evaluation signals. Each of the 40 models has a 40-training signal per client and a different number of signals per imposter, a different number of signals for the validation that ranges between 8 and 650 signals, a 5-signal for an authenticated client, and a different number of signals per imposter for evaluation. Independent validation and evaluation sets are used to evaluate our systems. In the biometric identification system, we obtained an equal error rate of 15.30% and 21.72% for the validation and evaluation sets, while in the biometric verification system, we obtained an equal error rate of 6.97% and 10.25% for the validation and evaluation sets.

Keywords Biometrics · Logistic regression · ConvNeXt network · Transfer learning · Footsteps recognition

Abbreviations

SFootBD	Swansea foot biometric database
GRF	Ground reaction force
PCA	Principal component analysis

✉ Ayman Iskandar
Ayman_Adel@cis.asu.edu.eg

Marco Alfonse
marco_alfonse@cis.asu.edu.eg

Mohamed Roushdy
mohamed.roushdy@fue.edu.eg

El-Sayed M. El-Horbaty
shorbaty@cis.asu.edu.eg

[1] Computer Science Department, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt

[2] Laboratoire Interdisciplinaire de L'Université Française d'Égypte (UFEID LAB), Université Française d'Égypte, Cairo, Egypt

[3] Computer Science Department, Faculty of Computers and Information Technology, Future University in Egypt, Cairo, Egypt

SVM	Support vector machine
AP	Accumulated pressure
RBF	Radial basis function
CNN	Convolutional neural network
ANN	Artificial neural network
Res Net	Deep residual neural network
CFPI	Cumulative foot pressure images
DataL	Data left
DataR	Data right
RELU	Rectified linear units
FAR	False acceptance rate
FRR	False rejection rate
EER	Equal error rate
DET	Detection error tradeoff

1 Introduction

Biometrics is the measuring and study of a person's unique physical or behavioral characteristics [1, 2]. It is commonly used for identification and authentication purposes, as biometric traits are difficult to forge or replicate. Biometric systems collect and compare these traits to authenticate the identity of individuals. Some of the commonly used biometric traits are Fingerprint Recognition, Iris Recognition, Face Recognition, Voice Recognition, Gait Recognition, Footsteps Recognition, Retina Recognition, and DNA Matching [1, 2]. Biometric systems are used in various applications, including access control to secure facilities, authentication for electronic devices, border control, and immigration processes, time and attendance tracking, and law enforcement investigations [1, 2]. However, while adopting biometric systems to protect sensitive biometric data, it is critical to address privacy considerations as well as effective security measures.

In comparison with other known biometric characteristics, the identification of footsteps is a relatively recent technology and less widespread biometric system. While footstep recognition is a fascinating notion, there are some key distinctions between it and other biometric systems, for example:

- **In terms of uniqueness**: Footstep recognition is based on analyzing the distinct aspects of a person's gait or walking pattern [3]. Each person walks in a unique manner, which includes elements such as stride length, cadence, pressure distribution, and foot angles. Other biometric systems, such as iris, face, or fingerprint recognition, rely on distinct physical characteristics as well, but they capture different types of characteristics.
- **In terms of contact versus non-contact**: Footstep recognition is a non-contact biometric system, which means that no physical contact with the person being identified is required. Systems such as fingerprint or palmprint identification, on the other hand, require direct contact with the sensor to capture the biometric characteristic [1]. Non-contact biometrics might provide convenience and hygienic benefits.
- **In terms of environmental factors**: Footstep recognition can be affected by environmental factors such as the kind of surface, footwear, or changes in walking patterns induced by injury or fatigue. Other biometric systems like fingerprint or iris recognition are less affected by such factors and tend to be more stable and reliable.
- **In terms of user acceptance**: Walking-based biometrics may be less common for people who are used to using their hands or faces for identification.

Footstep and gait recognition are two similar but separate biometric systems that analyze a person's walking patterns. Footstep recognition is concerned with analyzing the features of footsteps, such as pressure distribution, foot angles, and foot strike timing [3]. The advantages of footstep recognition over gait recognition are as follows:

- Footstep recognition can provide more exact and thorough information about an individual's foot contact with the ground by focusing on this localized element. This level of specificity might be useful in situations where foot placement or footprint analysis is critical, such as forensic investigations or security systems.
- Footstep recognition often necessitates the use of specialized sensors or flooring systems developed expressly to capture the characteristics of footsteps, these sensors can be integrated into flooring materials or placed on the ground. As opposed to gait recognition which may require video cameras or motion capture systems to capture the full body's movement, footstep recognition can have a simpler and more targeted sensor configuration, in some cases, this can result in lower costs and easier deployment [3, 4].
- Footstep recognition can achieve higher accuracy in foot-based identification than gait recognition due to its focused study of footsteps. The distinctive features of a person's footsteps, such as pressure distribution and foot angles, can serve as a valid biometric attribute. This can be useful in applications requiring foot-based identification, such as secure access control or footprint forensic investigation.
- Gait recognition considers the full body's movements while walking, including arm swing, torso movement, and posture, because they are more influenced by clothing, environmental conditions, walking pace

fluctuations, changes in lighting or camera perspective, or ambient circumstances, these upper-body differences can offer additional complications and obstacles in gait detection [4]. Footstep recognition, on the other hand, concentrates entirely on the features of footsteps and is less affected by upper-body differences. In certain cases, this can lead to increased accuracy and reliability.

Footstep identification has numerous potentials uses in a variety of industries [3, 5]. Some of the most notable applications of footstep recognition are as follows:

- Footstep recognition can be used to safeguard access control systems in places where traditional identifying techniques such as keycards or passwords may be compromised. Footstep identification can provide an additional layer of biometric authentication for secure entrance into restricted places by analyzing the unique properties of an individual's footsteps.
- By analyzing footprints left at crime scenes, footstep recognition can help forensic investigators. Investigators can gather evidence and identify probable suspects by matching the characteristics of the footprints with known footstep profiles. Footstep recognition can assist in associating people with specific locations or activities.
- Footstep recognition can be used as part of an intrusion detection system in secure locations to identify and inform security personnel about unauthorized footsteps. When an unfamiliar or unauthorized person enters the premises, the system can detect and raise an alarm by analyzing the distinctive footstep patterns of authorized individuals.
- Footstep identification can help behavioral biometric systems track and analyze human motions. It can also be used to measure and analyze individuals' gait patterns, providing insights into mobility, health, and well-being. This data could be useful in healthcare monitoring, assisted living, and elderly care.
- Footstep recognition can be utilized as an input modality for human–computer interaction, especially in situations where hands or other body parts are busy or unavailable. Footstep recognition, for example, can enable gesture-based control of smart devices or virtual reality systems by recognizing certain foot movements.
- Footstep recognition can help athletes monitor and analyze their motions during training or competitions. Coaches and trainers can get insights into running tactics, stride patterns, and foot striking patterns by capturing and analyzing footstep characteristics, allowing them to optimize performance and prevent injuries.

It's worth mentioning that footstep identification is still a developing technology, with continuing study and development. While it has the potential for applications such as security or gait analysis, it is not as commonly used or standardized as other biometric systems. The biometric system selected is determined by the unique needs, environment, and level of accuracy required for the intended application. We proposed two biometric systems that used a fused algorithm of the ConvNeXt neural network architecture and a logistic regression classifier. These biometric systems should be able to recognize and validate the identities of forty different users, as well as distinguish between authenticated and impostor users.

The remainder of the paper is structured as follows: Sect. 2 discusses the related work of various approaches and algorithms applied to the Swansea Foot Biometric Database. Our proposed methodology for the biometric identification and biometric verification systems is described in Sect. 3. Section 4 presents the experimental results, analysis, and observations. Section 5 offers the conclusions and recommendations for future work.

2 Related work

Vera-Rodriguez et al. [6–9] proposed the biggest footstep database (SFootBD) to date to analyze footstep signals as a biometric, which contains around 9900 single strides from 127 people. They measured footstep pressure with a pair of floor mats, each with 88 piezoelectric sensors to record most of the gait cycle. The sensors were installed on a large-printed circuit board and positioned beneath a standard mat. The sensor layout geometry enables a compact arrangement with consistent intersensor separation. The database was partitioned into different sets. Most of the studies were conducted utilizing 40 user models, each with 40 signals to train, as well as validation and evaluation of test sets, other benchmarks were also considered during different experiments.

Their first experiment focused on signal time information analysis [6], employing three distinct time domain feature approaches: the ground reaction force (GRF), the spatial average of the sensors, and the upper and lower contour profiles of the time domain signal. These three features are merged at the feature level, and principal component analysis (PCA) is used to minimize the data dimensionality. Finally, support vector machines (SVM) are employed to perform the matching. The experimental methodology is intended to investigate the impact of the amount of data used in the reference models by imitating environments for potential extreme applications such as smart homes or border control scenarios. Previous footstep databases used in related studies had drawbacks in that they

randomized data time sequence in the trials, making the reference and test datasets are more comparable, and hence allowing for artificially favorable findings. The footstep database employed in this study, on the other hand, is the largest to date, with more than 120 people and nearly 20,000 signals for left and right footsteps separated, and it monitors the temporal difference between reference and test data. This increases the realism of the experimental environment and enables a more accurate evaluation of footstep identification systems. The findings obtained in the various settings for the case of the stride footstep for the fusion of the three feature techniques are in the range of 5–15% EER, which is better than prior research and obtained with a lot more realistic experimental setup.

Their second strategy was to extract biometric information from footstep signal pressure distribution as well as spatial domain and generate a 3D image of the pressure distribution along the spatial domain [7]. Because the time domain features are ignored in this research, a single value of pressure is derived through the integration of signals across the time axis for each sensor of the mat. It is important to note that, due to the differential nature of the piezoelectric sensors' footstep signals, a simple integration of the signal over time would provide a value close to zero. To address this, the integration is performed across each sensor's associated ground reaction force signal (GRF_i). This yields the accumulated pressure (AP_i), which is a measure used for studying the pressure distribution across the spatial domain of the signals. The signals were aligned to a single temporal position during the preprocessing stage by applying an energy detector across the 88 sensors of the signals to obtain the starting point of each footstep to align the signals to a common time position. The images were then smoothed using a Gaussian filter and the toe and heel regions were aligned and rotated based on the pressure points. The rows of the generated image, which has a resolution of 280 × 420 pixels, are concatenated to generate a feature vector with a range of 117,600. Principle component analysis (PCA) is also used to minimize data dimensionality, maintaining more than 96% of the original data by using the first 140 principal components. The feature vector for the stride (right and left) footstep is composed of the concatenation of the 140 component feature vectors for the right and left foot, as well as the stride's relative angle and length, for a total of 282 characteristics. In terms of the classifier, a support vector machine (SVM) with a radial basis function (RBF) as the kernel was used. The results reveal that the suggested methodology achieves high accuracy in identifying individuals based on their footstep signals, with EERs ranging from 6 to 10% in diverse scenarios. The results also suggest that including signals with high heels in the training data considerably improves the system's performance.

Another experiment was conducted to recognize the footsteps through the use of time, space, and a combination of the two [8]. They fused the features obtained in the above two experiments [6, 7]. The two domains perform very similarly, with equal error rates ranging from 5 to 15% for each domain and 2.5 to 10% for their fusion using the support vector machine algorithm, depending on the number of used signals ranging from 40 to 500 and whether the footstep is single or stride footstep.

Costilla Reyes et al. [10, 11] conducted other two experiments on the SFootBD database, which contains 127 people and over 20,000 valid footstep signals (i.e., 10,000 stride signals). For the first experiment there were forty clients in the training set, each with forty stride footstep signals, and eighty-seven impostor subjects [10]. There were 7077 samples in the validation dataset and 550 samples in the evaluation dataset. They proposed using a Convolutional Neural Network (CNN), which is utilized to extract features for the full spatial database. After the CNN has been trained, the training, validation, and testing datasets are input into the network, and the features created by the CNN are extracted at the last layer before softmax activation. This returns a 127-feature vector for each sample in the dataset. The 127-length feature vector generated by the CNN model per dataset sample is then used in a biometric verification scenario with a one-vs.-one linear SVM model. The models were trained in the training set for each user and then tested in the validation and evaluation datasets. The paper's reported equal error rate (EER) range is as follows: The EER for the left and right footstep models in the validation dataset was 14.76% and 14.23%, respectively. When compared to single footsteps, the stride footstep model demonstrated a significant EER improvement of 9.392%. By assessing the trained model on the evaluation dataset, an EER of 21.30% was obtained for the left footstep, an EER of 20.23% for the right footstep, and an EER of 13.86% for the stride footstep.

For the second experiment, they suggested a deep residual ANN based on the ResNet architecture [11]. Based on the fine-grained diversity of footsteps, their solution contains artificial intelligence that can distinguish between authenticated users and biometric system imposters using SFootBD. The data's raw spatial and temporal properties are represented in a variety of ways. For the spatial component, each footstep frame is molded into a two-dimensional matrix, with sensors formed by the merging of the two mats and pixels derived from the accumulated pressure. To maximize data variability versus training time, the temporal component representation selects frames that correspond to intervals of heel striking, flat foot, and heel-off. The authors carried out three separate studies based on three different circumstances. The three experimental settings altered the number of users and impostors by altering

the quantity of footstep data used to train the machine-learning model. Two ResNet models were trained, one for spatial component features and one for temporal component features. The trained ResNet models were then used as feature extractors to feed linear SVM classifiers. A single class was assigned to the imposter dataset. They applied their work to three benchmarks that were constructed to simulate three different scenarios (airport, workplace, and smart home). For the first benchmark which has the smallest footstep dataset and simulates a scenario involving an airport security checkpoint, included 40 footstep samples from 40 users and 763 imposter footstep samples, and the optimum model yields 7.10% and 10.50% EER for the validation and evaluation sets, respectively. The second benchmark was for the medium-sized training dataset, which consisted of 200 stride footsteps of 15 users and 2697 imposter footstep signals simulating a working environment scenario. EERs of 2.80% and 4.90% were obtained from the validation and evaluation sets, respectively. As for the third benchmark, which is based on a home-based scenario, it contains 500 footstep samples from 5 individuals as well as 5603 imposter footstep signals. It obtained the best results in comparison with the other two scenarios; the validation and evaluation sets had 0.70% EER and 1.70% EER, respectively.

Table 1 summarizes the related work in terms of the used deep learning and/or machine-learning algorithms, the used approach of each experiment, the features used in each experiment, and the hypotheses assumed in each experiment.

3 Proposed methodology

Two biometric systems were proposed, one for an identification scenario and the other for a verification scenario. The biometric identification system is to recognize the submitted biometric footstep signals between forty different users using a single logistic regression model. On the other hand, the biometric verification system is to validate the claimed identity of the user using forty different logistic regression models, one for each authenticated user. Figure 1 presents a high-level architecture of our proposed systems' components, which consist of five phases (enrollment, feature extraction, template creation, comparison and matching, and decision and authentication).

We used the ConvNeXt neural network as our template creation model to create a mathematical template that represents the biometric data of the footstep signals, and then we fed the output of the ConvNeXt network to the logistic regression models in both of our proposed systems. Figure 2 presents our flow for designing our biometric identification and verification systems.

3.1 Enrollment phase

The initial step in a biometric identification and verification system is to capture the biometric data of individuals who will be enrolled in the system. This involves collecting samples of their biometric features, such as fingerprints, face images, iris patterns, voice recordings, or behavioral traits like signature, keystroke dynamics or footstep pressure signals. The collected data are then stored in a database along with the associated identity information. In our case, we used the footstep pressure signals from the Swansea Foot Biometric Database (SFootBD).

The Swansea Foot Biometric Database (SFootBD) is one of the largest open-source gait databases available for footstep recognition [6–11]. The data are collected using two sensor mats of 45 CM × 35 CM; each mat contains eighty-eight piezoelectric sensors to capture two successive footstep signals over 2200 time samples for each footstep signal. To create a realistic signal, the participants were advised to walk at a steady pace for a few meters before walking on the sensor mats. The sensors provide a differential voltage output based on the pressure applied to them. The data were gathered in several sessions to provide more realistic samples of each person. Also, different conditions were applied, including wearing different shoes, barefoot, carrying extra weight by carrying a bag, wearing high heels, and walking at different speeds. The database also contains three other biometric modes: speech, face, and gait. These different biometric modes aided in the labeling of footstep signals, as the collection was an unsupervised process.

The data contain the footstep signals of 127 people, with a total of 19,980 footstep signals for single right and left footstep signals and 9900 for stride footstep signals. Three benchmark datasets can be constructed from the SFootBD database: the first contains 40 users and 87 imposters; the second contains 15 users and 112 imposters; and the third contains 5 users and 122 imposters. The three benchmarks represent airport security checkpoints, workplace scenarios, and home scenarios, respectively. Where each benchmark mimics a real-world scenario, the airport security checkpoint has the smallest amount of data considered per person with the largest number of people, the workplace scenario has a medium amount of data per person with a medium number of people, and the home scenario has the largest amount of data per person with a small number of people. Our work is conducted under the first benchmark (Airport Security Checkpoint) because it is the baseline benchmark and the most challenging one because of the small amount of training data and the large number of available users when compared to the other two benchmarks.

Table 1 Comparison of the related work

References	Algorithm	Approaches	Hypothesis test
Vera-Rodriguez et al. 2010 [6]	PCA and SVM	Three distinct time domain features were extracted: the Ground Reaction Force (GRF), the spatial average of the sensors, and the upper and lower contour profiles of the time domain signal. PCA is used for dimensionality reduction, and SVM is used as a classifier	The authors represented two security scenarios (smart homes and access control). For the smart home, they provided a large quantity of data for a small number of people; in contrast, for the access control scenario, a limited number of data was offered for many people to mimic border control, for example
Vera-Rodriguez et al. 2011 [7]	PCA and SVM	Calculating the accumulated pressure to obtain pressure images. PCA is used for dimensionality reduction, and SVM is used as a classifier	The authors assumed the same hypothesis as the previous experiment, with two security scenarios for smart homes and access control
Vera-Rodriguez et al. 2013 [8]	PCA and SVM	Combining the features of the previous two experiments	Testing the quantity of data is a key aspect of training and testing the biometric system. As the number of clients increases, the number of signals decreases Building models for too many persons that range from 75 to 5 persons, representing a variety of security scenarios
Costilla Reyes et al. 2016 [10]	CNN and SVM	Using CNN as a feature extractor and training 40 one-vs. one SVM models, a model per client	The authors represented the baseline benchmark for airport checkpoints, which has 40 persons and 87 imposters, using all the dataset in the experiment with 2363 training samples, 7077 validation samples, and 550 evaluation samples
Costilla Reyes et al. 2019 [11]	ResNet and SVM	Using two ResNet models as feature extractors, one for the spatial raw component and the other for the temporal raw component, and training two one-Vs.-one SVM models, one for each component, then fusing the score to obtain the results	Three different benchmarks were proposed to simulate three different security scenarios (airport checkpoints, workplaces, and smart homes). The three proposed benchmarks were 40 persons, 15 persons, and 5 persons, respectively, where the number of footstep signals is larger for the smaller benchmark

3.2 Feature extraction phase

The SFootBD includes four index files: a list of imposters, a list of authenticated clients, a list of validation tests, and a list of evaluation tests. Each index file has the id of the person and the corresponding MATLAB file title, our study is conducted under the baseline benchmark (Airport Security Checkpoint) which has forty authenticated clients and eighty-seven imposter users with a relatively small training set. The used dataset is distributed as shown in Table 2.

As shown in Table 2, authenticated users have the largest number of signals, while the remaining signals are considered imposter users. Validation and evaluation are two different test sets, the main difference between the two sets is that the validation set is an unbalanced test set used to tune the model weights, while the evaluation set is unseen balanced data with every client having five signals to measure the generalizability of the proposed model. The SFootBD dataset also provides a very realistic approach where the dataset was collected over eighteen months; therefore, the training data are the first collected data, and the evaluation data are the last collected data, which reflects an actual life scenario rather than the traditional related work that divides the data into training and test sets.

We loaded the index files so we would be able to read the user's ID and its relevant MATLAB files. Each MATLAB file contains two matrices: the left footstep signals (DataL) and the right footstep signals (DataR). Each matrix has a dimension of 2200×88, where 2200 represents the time frames and 88 represents the sensors, so it shows the reading of each specific sensor at each of the 2200 time frames. Figures 3 and 4 show an example of the differential pressure taken straight from the 88 sensors from the left and right footstep signals plotted against the time frames.

After reading the 2 MATLAB matrices of all the files, we calculated the ground reaction force of all the files as shown in Eq. (1):

$$\mathrm{GRF}_i[t] = \sum_{t=0}^{T} (S_i[t]) \tag{1}$$

where T represents the time frame that we are standing at, t equals the start time frame, i represents the sensor, and the $S_i[t]$ represents the output signal of the ith piezoelectric sensor at time t. So, we calculated the ground reaction force which is a cumulative pressure for each time frame with all

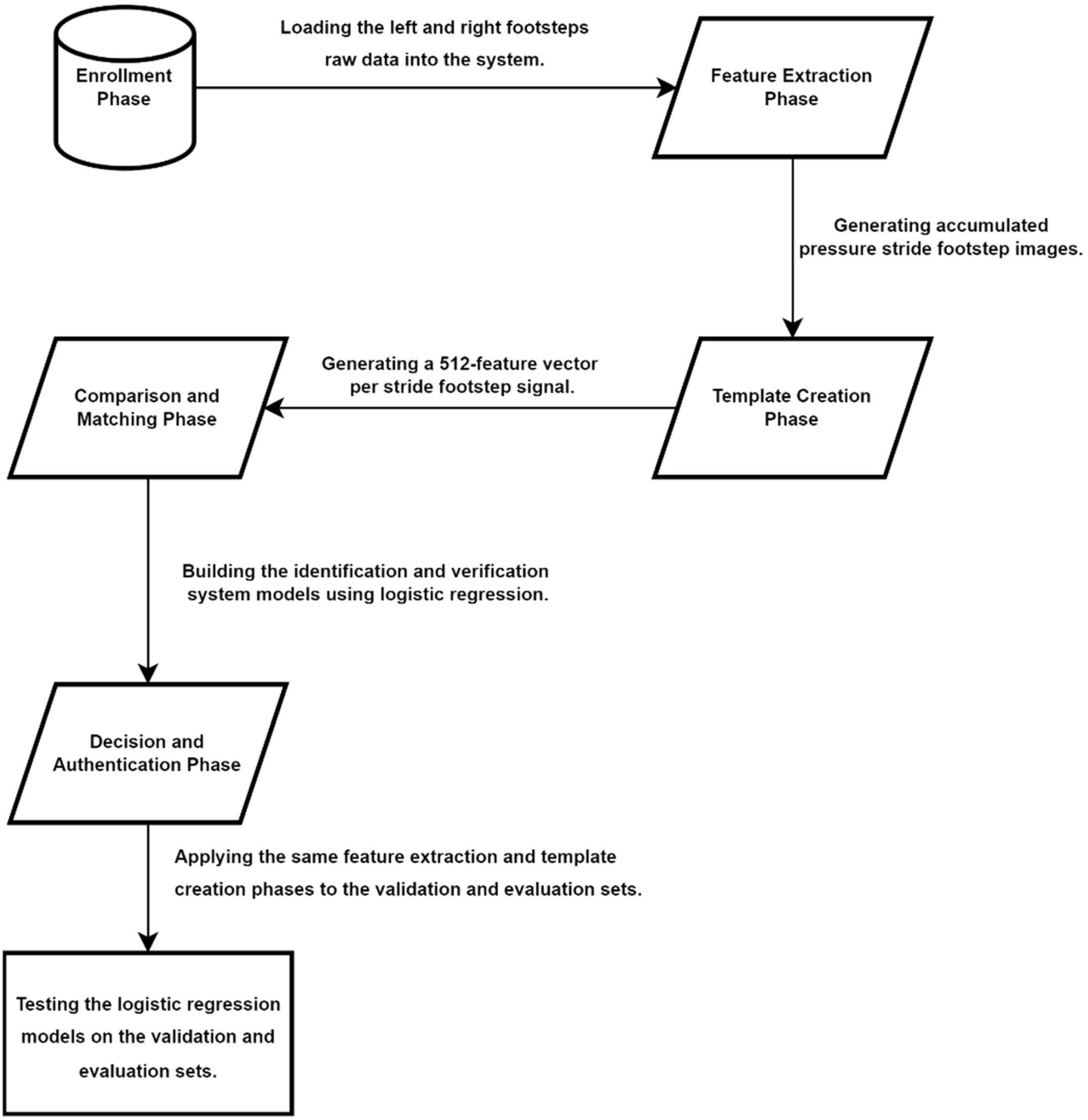

Fig. 1 The high-level architecture of our proposed systems' components

the previous times frames for the same footstep signal data file [7, 10, 11].

After calculating the ground reaction force of all the data signals, we calculated the accumulated pressure which is defined in Eq. (2):

$$\mathrm{AP}_i = \sum_{t=0}^{T_{\max}} (\mathrm{GRF}_i[t]) \tag{2}$$

The accumulated pressure is a cumulative summation for each sensor value of the calculated ground reaction force; the $T_{\max}$ was set to 1600 time frame because no signal has a larger time frame, which represents that the person has already walked out of the walking mat, so setting the maximum time frame to 1600 instead of 2200 to reduce the computational burden [7, 10, 11]. The output of this stage is a list that contains eighty-eight values; each

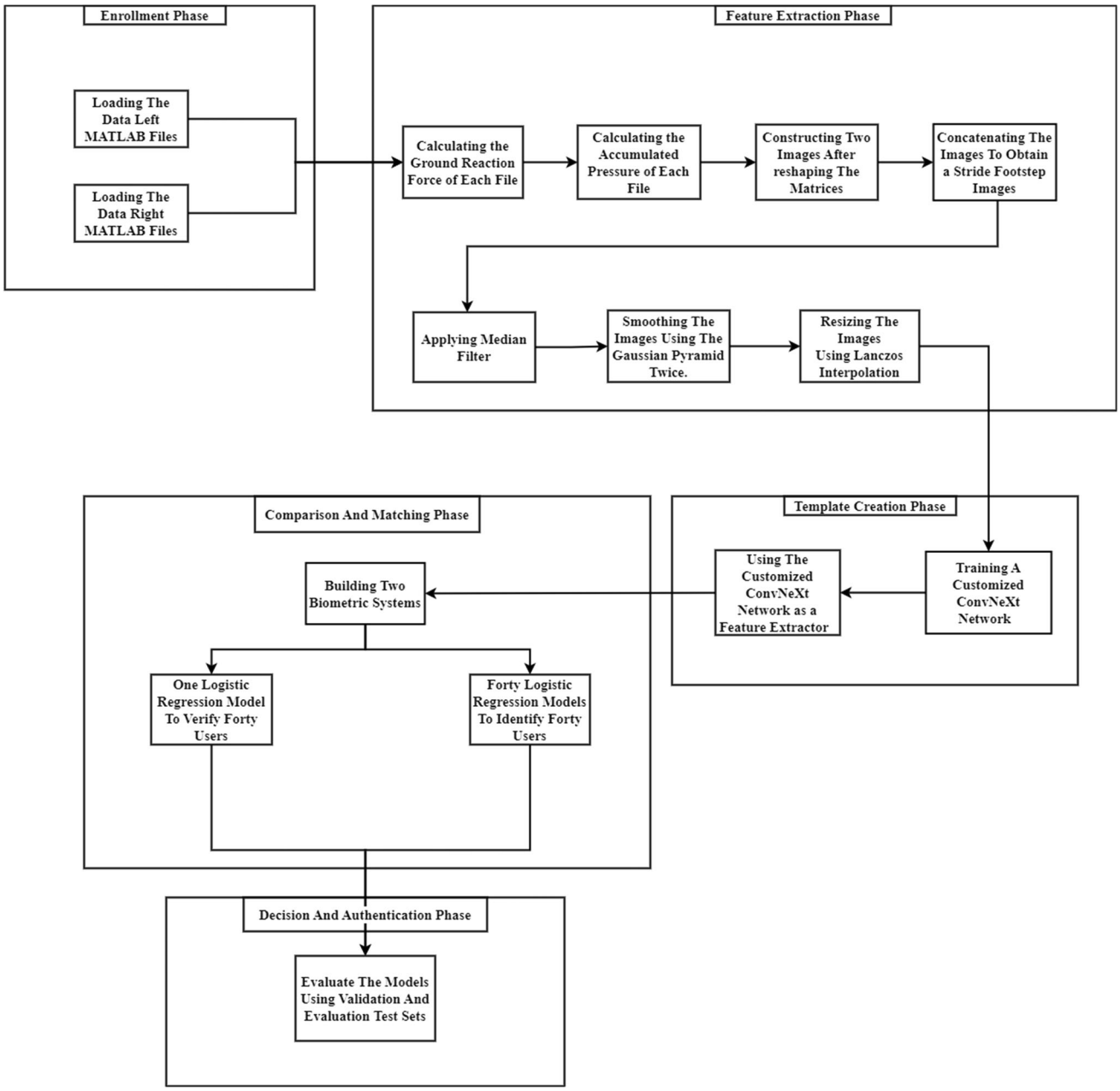

Fig. 2 The proposed systems' workflow

Table 2 SFootBD distribution

Dataset	Training	Validation	Evaluation
Signals of clients	1600	6697	200
Signals of imposter users	763	380	350
Total signals per set	2363	7077	550

value represents the accumulated pressure of one of the eighty-eight sensors.

After that, we added three zeros to each acquired data list, increasing the number of values in each list to ninety-one, which allowed us to reshape each list into a two-dimensional array with a size of 13 × 7 as explained in Fig. 5, and then the reshaped array can be presented as a cumulative footstep pressure image (CFPI) to represent each left and right footstep sample.

Figures 6 and 7 show a sample of the obtained images after reshaping the left and right footstep signals into a two-dimensional array of size 13 × 7.

We applied a median filter with a kernel of (1,3) to the left and right footstep images to eliminate the noise while preserving edges and fine details in the images [12]. The obtained left and right footstep images are then

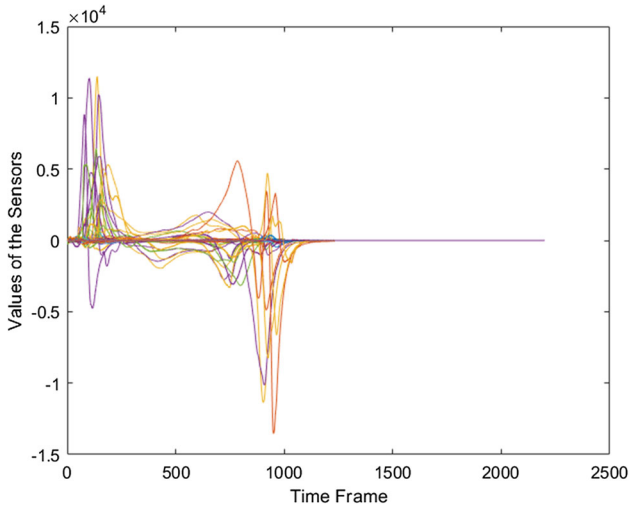

Fig. 3 Left footstep signal

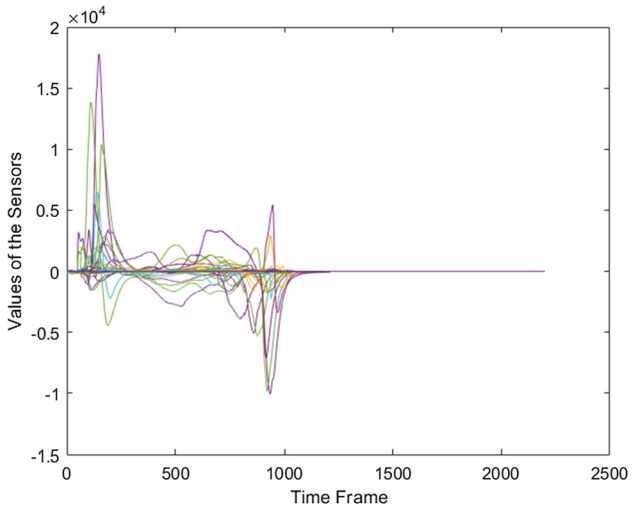

Fig. 4 Right footstep signal

concatenated to create a single stride footstep image with dimensions of 13 × 14, and then we used a minimax scaler to scale all the pixel values to range between 0 and 255 to ensure that all the images have the same scale and are within the color range.

To obtain a better resolution and smoother image, we used the Gaussian pyramid technique on the concatenated image [13]. The Gaussian pyramid was applied twice to create an image with dimensions of 52 × 56.

We examined many methods of interpolation before and after applying the Gaussian pyramid, such as nearest, bilinear, linear, spline, bicubic, quadric, Mitchell, and Lanczos [14]. We used Lanczos interpolation, which is a computationally intensive method that uses a sinc function as a windowed kernel to estimate pixel values, providing high-quality results with less aliasing [14]. Resized images are obtained with dimensions of 148 × 160 for each

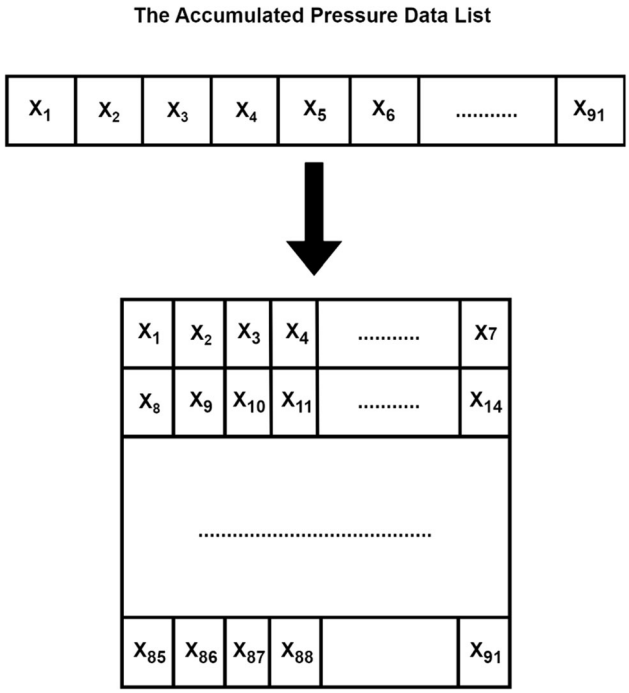

Fig. 5 Reshaping the accumulated pressure list into a 2-D array

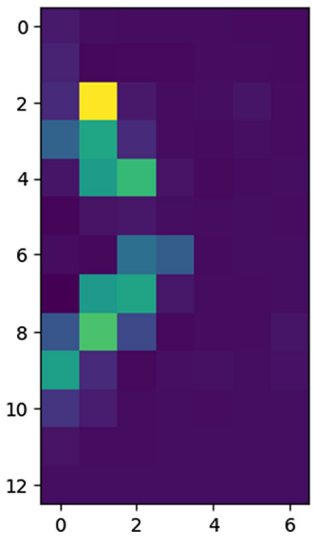

Fig. 6 Left footstep image

image. Figures 8 and 9 show the stride footstep image after concatenation, applying the minimax scaler, and applying the median filter.

As shown in Fig. 9, the median filter helped reduce the noise by removing the outlier high pixel values. Figures 10 and 11 show the stride footstep image after applying the Gaussian pyramid technique and the Lanczos interpolation that is used in resizing the image. The Lanczos interpolation made the image smoother with less aliasing.

Figures 12 and 13 represent different interpolation trials on the concatenated footstep stride image before and after

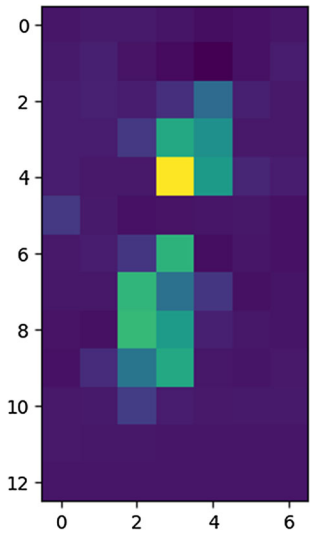

Fig. 7 Right footstep image

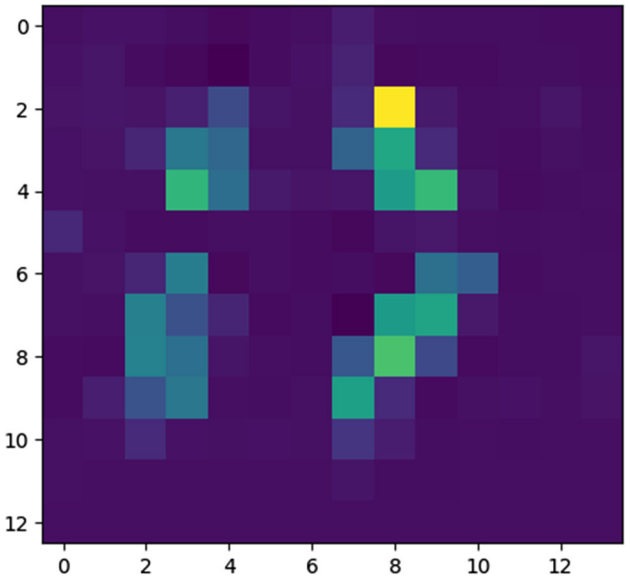

Fig. 8 Footstep image after concatenation, applying minimax scaler without applying median filter

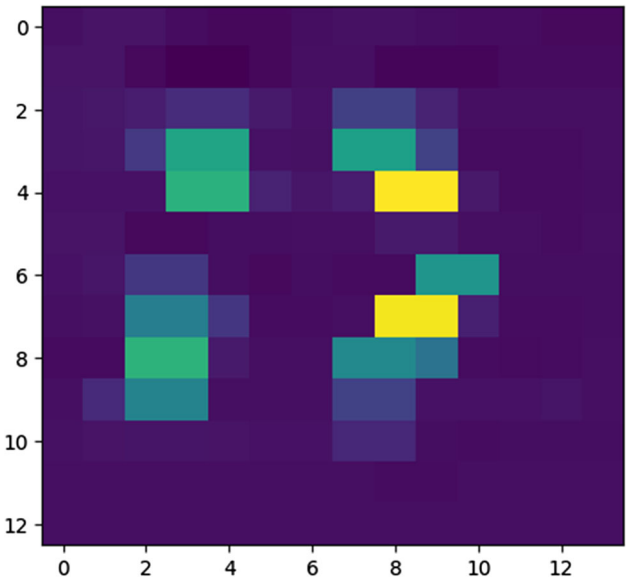

Fig. 9 Footstep image after concatenation, applying minimax scaler after applying median filter

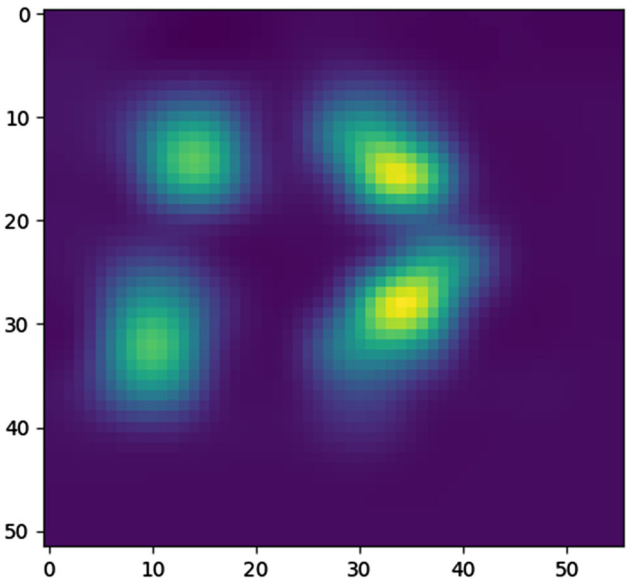

Fig. 10 Footstep image after applying Gaussian pyramid

the Gaussian pyramid technique is applied. In our experiment, we used Lanczos interpolation since it shows the primary features of the image without smoothing out the high-frequency details.

Figure 14 depicts an example of a stride footstep image computed straight from the accumulated pressure without firstly applying the ground reaction force equation. Because of the differential nature of the piezoelectric sensor footstep signals, simple integration of the signal across time would most likely provide a result close to zero. To remedy this, the integration is done across each sensor's associated ground reaction force signal (GRF_i), demonstrating the significance of first computing the ground reaction force before calculating the accumulative pressure.

Data augmentation techniques were applied to the constructed images, which represent some natural variations of the person's footsteps while he is walking on the floor sensor, like rotation or shifting on the stride footstep images. Figure 15 shows some samples of the images after data augmentation techniques were applied, where class (-1) represents the imposter class, and the rest represents authenticated users.

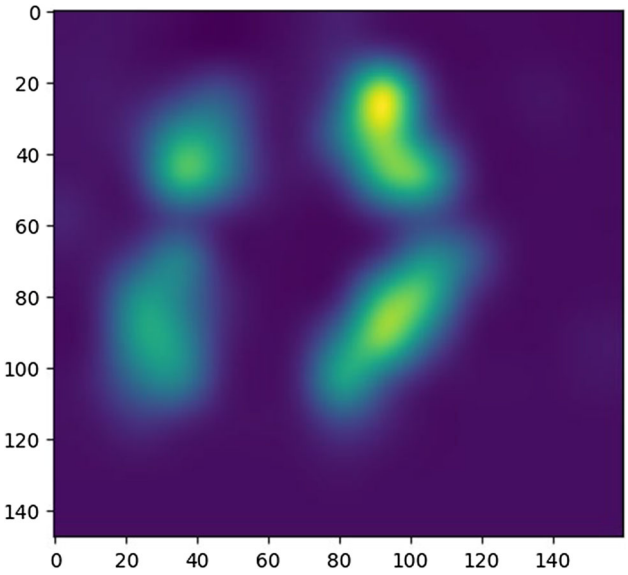

Fig. 11 Footstep image after applying Lanczos interpolation

3.3 Template creation phase

The extracted features are used to create a reference template, which is a mathematical representation of the individual's biometric data [1]. The template contains the essential information required for subsequent comparison and matching. We used transfer learning [15] to train our customized pre-trained ConvNeXt Network [16]. Then, we used the ConvNeXt network as a feature extractor to generate a vector of 512 values as our mathematical representation of the individual's biometric data.

The ConvNeXt model is trained for the biometric identification scenarios [16]. ConvNeXt [16] is a novel CNN architecture that is a modified ResNet-50 [17], which is changed by inspiration from Swin Transformer [18]. ConvNeXt has been widely deployed for a range of diverse applications, including object detection and semantic segmentation, and has been proved to perform well in a variety of computer vision tasks. Its modular form and good use of computational resources make it a promising architecture for real-world applications needing both precision and efficiency.

The model is trained on forty authenticated users, while a single class is assigned to the imposter footstep signals with a label of (−1). For our baseline benchmark, we load a pre-trained ConvNeXt model with ImageNet initialized weights. ImageNet is a very popular dataset that is used for object classification and detection problems [19]. It provides hundreds of objects and millions of images of different objects that anyone can come into contact with in real life. The ImageNet dataset has a variety of images of objects that vary between animals, birds, foods, laptops, fruits, and even more objects than the ones mentioned but it doesn't have images like the ones that we used here for our biometric systems because the ImageNet represents a real-life object while we are using generated heat map images

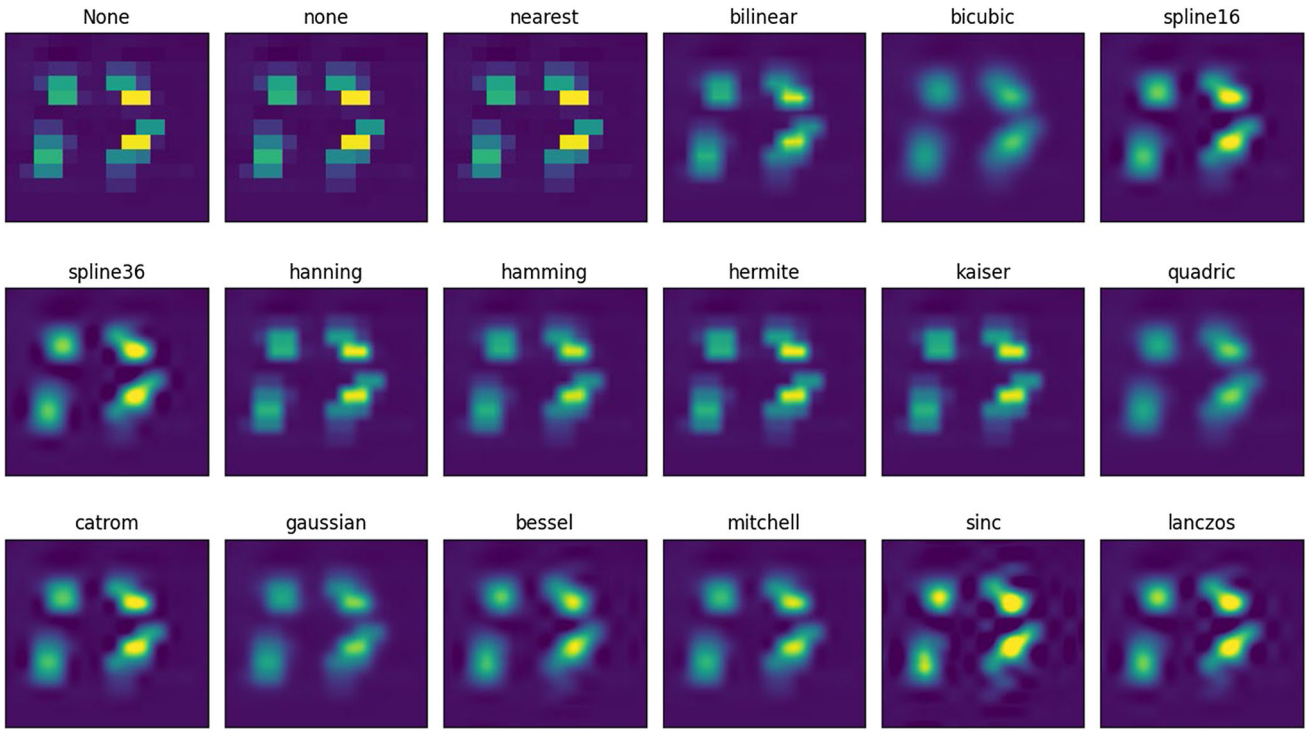

Fig. 12 Different interpolation techniques before applying the Gaussian pyramid

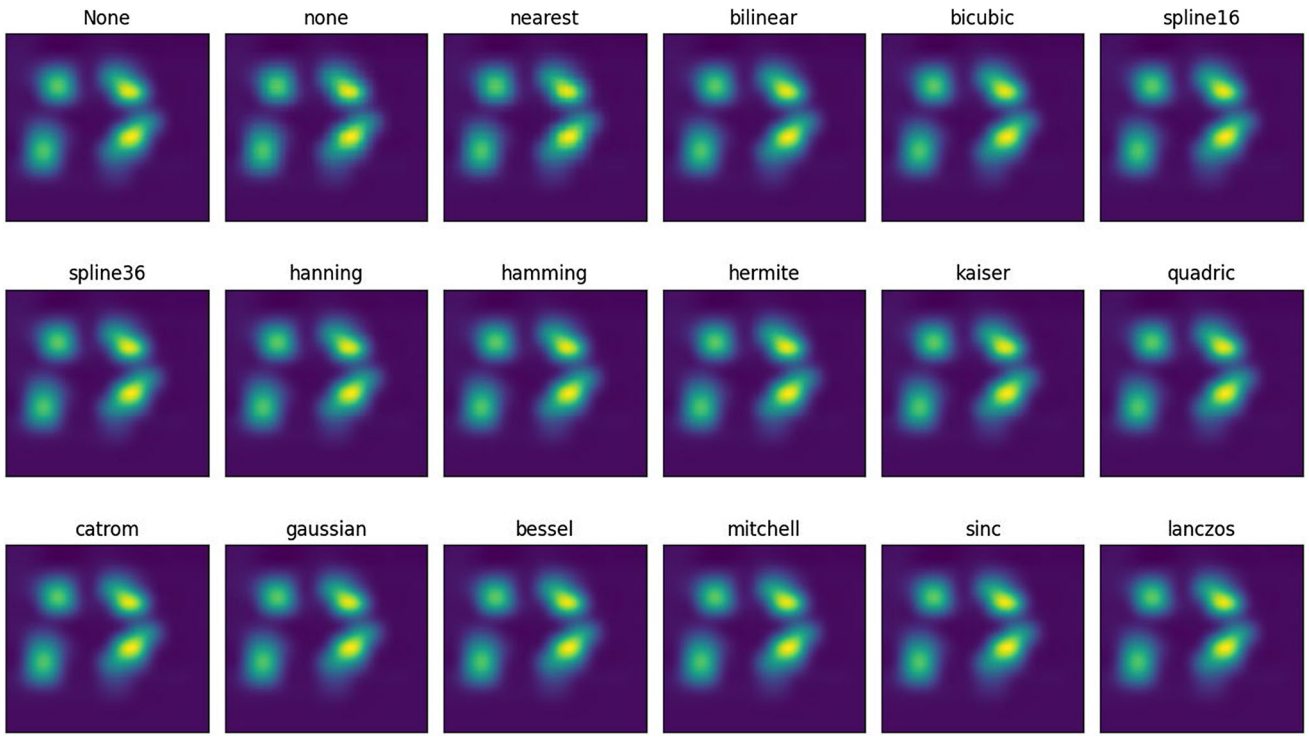

Fig. 13 Different interpolation techniques after applying the Gaussian pyramid

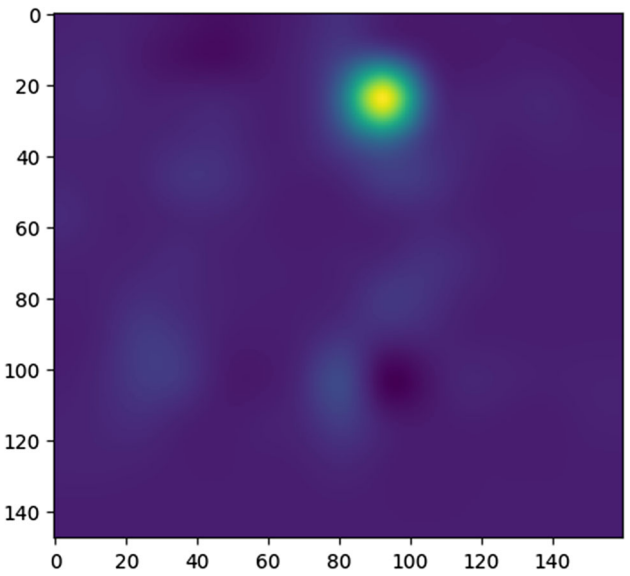

Fig. 14 Stride footstep image without calculating GRF

for our biometric systems that have been created using the accumulated pressure and processed as we discussed in the previous section. The fully connected layer is then removed from our proposed ConvNeXt model, a flattened layer is added to the architecture to flatten the output of the ConvNeXt architecture into a one-dimensional vector, and then a batch normalization layer is added to our model to normalize the output of the model into a standard Gaussian distribution with zero mean and unit variance to avoid internal covariate shift. Then the batch normalization layer output is passed to a fully connected layer with 512 neurons and RELU activation function; a dropout layer with a percentage of 20% is added to avoid overfitting by randomly eliminating some of the neurons; and finally, a fully connected layer with 41 neurons and SoftMax activation function is added to predict the output label of each sample. Figure 16 shows our proposed ConvNeXt network.

The model was trained using 16 stride footstep signals in each batch and with an Adam optimizer and a learning rate of 0.0001. Sparse categorical cross-entropy is used as our loss function because our data are mutually exclusive, with each sample belonging exactly to one class, which performs better than normal categorical cross-entropy, which consumes more time and memory and is typically used when the labels are in the form of one-hot encoded vectors [20]. If the validation loss did not improve for five epochs, an early stopping method was used to terminate the training process.

To build our biometric identification and verification system, we used our trained ConvNeXt model as a feature extractor and compressor for the dataset used. We extracted the feature vector learned by the model before the final layer (the SoftMax layer). The intermediate layer generates a vector of size 512; therefore, the shape of our training sample will be (2363, 512), the validation sample (7077, 512), and the evaluation sample (550, 512), for each

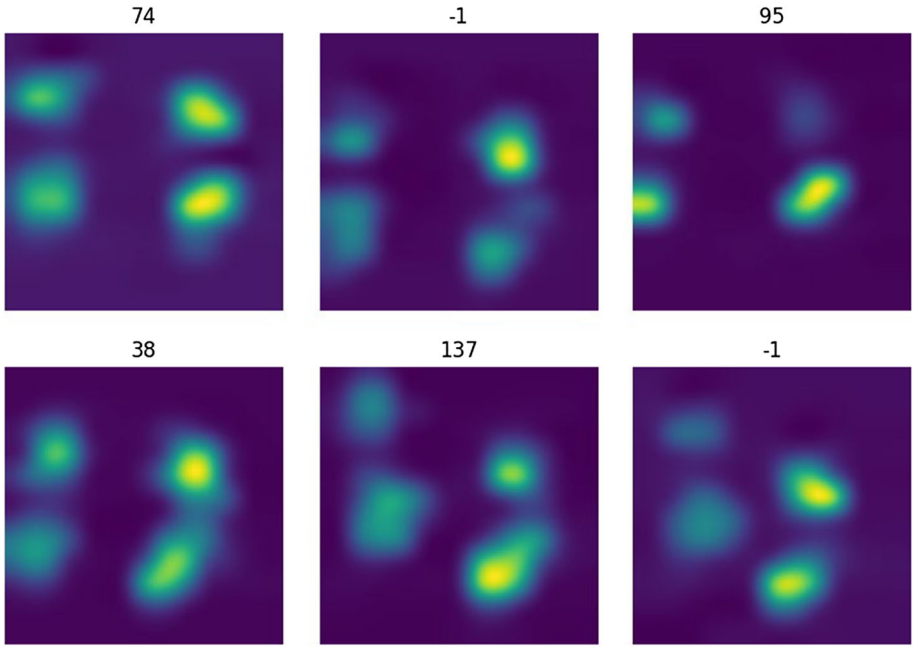

Fig. 15 Stride footstep images after applying data augmentation

sample having a vector of size 512. These extracted features from each set are then used to feed our machine-learning biometric identification and verification system.

The significance of this approach is that the ConvNeXt worked as a feature extractor and compressor, so we have a smaller data size to be able to train our machine-learning model with so many distinct features that it is able to identify the person from this set of features which help in saving time and computational resources while training and testing the machine-learning model, and the trained ConvNeXt model is saved with its obtained weights to be used it in a real-time biometric identification and verification scenario where the live footstep data can be fed to our saved model architecture and extract the feature vector that can be directly passed to the machine-learning model and be able to identify the individuals using only their footstep.

3.4 Comparison and matching

The extracted features from the presented sample are compared with the reference template stored in the database [1]. Various matching algorithms, such as correlation-based matching or machine-learning-based approaches, are employed to identify the similarity or dissimilarity between the provided sample and the stored template. We used logistic regression as our method for the comparison and matching phase [21, 22].

We conducted two separate experiments for biometric identification and verification scenarios. We picked logistic regression because it can handle huge datasets with large input features while requiring relatively little memory compared to more complex models, allowing us to train on limited computational resources [21, 22]. Furthermore, as compared to other models, logistic regression is less affected by irrelevant features and produces a probabilistic output in the form of estimated probabilities, which is important in decision-making scenarios.

The first experiment for the biometric identification scenario is conducted to train a single logistic regression model capable of differentiating between an imposter user and an authenticated user with identifying the identity of the authenticated user. The model was trained using the features obtained from the ConvNeXt model on only the forty authenticated users because the identity of the imposter is not important to recognize, and we cannot know all the imposter users in real-life scenarios, and it's only important to us to verify if the user is an authenticated user or an imposter, knowing his identity in the case of an authenticated user. The illustration of the biometric identification system is shown in Fig. 17.

The second experiment for the biometric verification scenario involves training forty logistic regression models, one for each user; each model was trained in the training set for each user using the features that were obtained from the ConvNeXt model to train the models. Samples that do not belong to the *i*th class of the *i*th model are handled as imposter samples throughout model training, validation, and evaluation. The illustration of the biometric verification system is shown in Fig. 18.

Other trials were also carried out using various machine-learning methods and approaches. We conducted these two experiments using different algorithms, for example: using multiple SVM kernel models, random forest tree, decision tree, K nearest neighbors, Gaussian processes, bagging

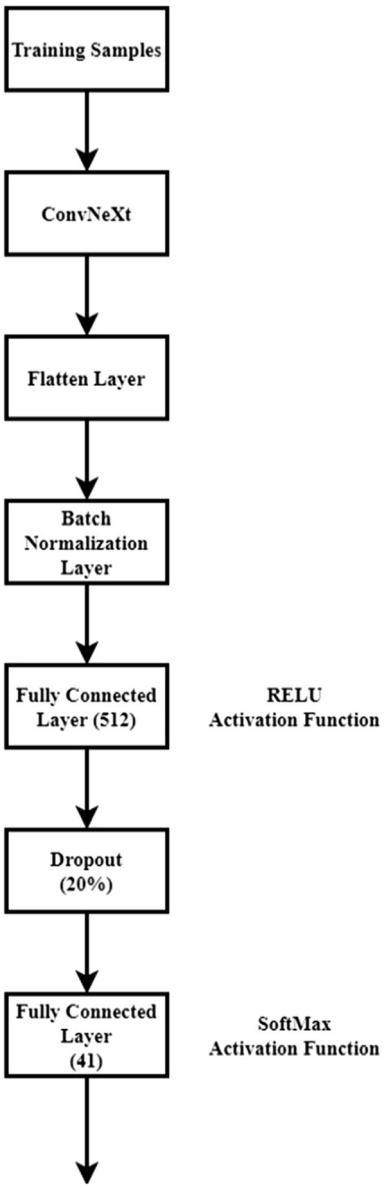

Fig. 16 Our customized ConvNeXt network

classifier, and ensemble learning technique [22–24]. The equal error rate of the suggested models was higher than the logistic regression model, where the results of the first biometric system were nearly between 18.85% and 16.32% for the validation set and 25.52% and 22.68% for the evaluation set, and for the second biometric system, the results were nearly between 10.17 and 8.56% for the validation set and 15.23% and 11.35% for the evaluation set. We chose logistic regression due to its probabilistic diversity between the outputs, which helped us distinguish between the users based on the probabilistic outcome of each class.

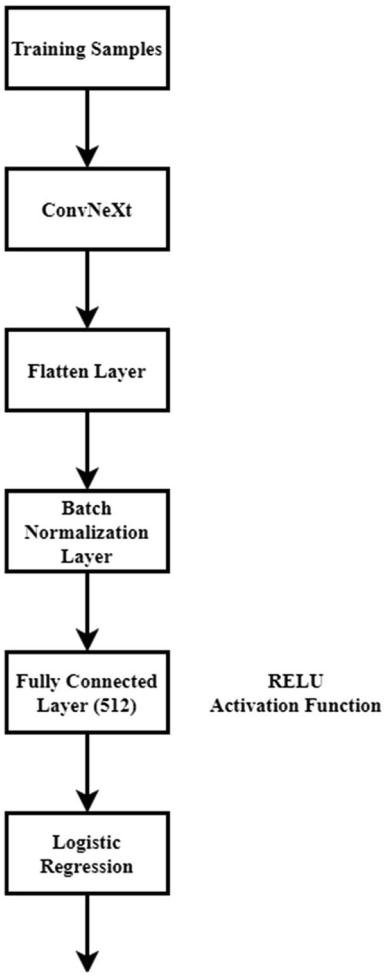

Fig. 17 Biometric identification system workflow

4 4. Decision and authentication

4.1 Evaluation metrics

The false rejection rate (FRR) assesses the system's ability to recognize legitimate users [1, 25, 26]. It is the percentage of times the system incorrectly rejects a user. The false acceptance rate (FAR) assesses the system's ability to detect imposters [1, 25, 26]. It is the percentage of times the system incorrectly accepts an imposter. FAR and FRR are two essential metrics that can be used to evaluate the performance of a biometric system. These rates are typically configured in software by adjusting the system's threshold value. The FAR and FRR have an influence on a biometric system's security level, this means that as you increase or decrease these rates, the number of usable authentication attempts will decrease or increase, respectively. A high FAR means that the system is more likely to incorrectly accept an unauthorized user, which can compromise the security of the system. A high FRR indicates

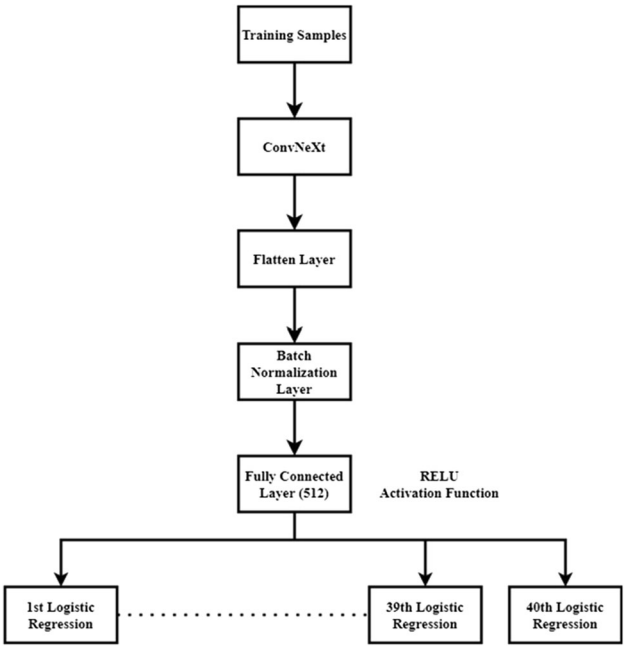

Fig. 18 Biometric verification system workflow

that the system is more likely to mistakenly reject a legitimate user, which can cause frustration and lead to a drop in productivity. A low FAR with a low FRR indicates a high-security level. When setting threshold values for a given system, it is important to find a balance between the FAR and the FRR. The choice of threshold should reflect the balance between security and usability. The number of FAR and FRR is inversely proportional. One will rise while the other will fall. The intersection of these two lines is known as the equal error rate (EER). This is the point at which the percentage of false acceptance errors and false rejection errors is equal.

We used the equal error rate (EER) as our evaluation metric [1, 27]. EER is a metric used in biometric security systems to measure the effectiveness of the system in identifying individuals correctly. EER is the point at which the FAR and FRR are equal. In other words, it is the threshold at which the system is equally likely to wrongly accept a non-matching individual as it is to wrongly reject a matching individual. EER is an important metric in biometric security systems as it provides a balance between security and convenience. The EER is defined as the crossover point on a graph that has both the FAR and the FRR curves plotted. The EER can also be calculated from a detection error tradeoff (DET) curve, which plots FAR versus FRR to measure the sensitivity and accuracy of a certain device [28, 29]. To compute the EER using the FRR/FAR crossover, the following is done: The FAR and FRR are computed and shown on the same graph for every specified threshold value ranging from 0 to 100, the EER is the point where the two curves cross. To calculate the DET of a biometric system, each corresponding FAR and FRR point is plotted, the EER is then found by extending a 45-degree line from the point of origin (0, 0), and where this line crosses the DET curve that is the EER, which occurs when the value of the FRR equals the value of the FAR on the curve are equals. The EER's strength is that it allows for a comparison of various biometric systems. This is due to the fact that not all biometric systems have the same threshold values. Thus, we can attempt to compare two biometric systems by comparing a normalized statistic such as the EER.

4.2 Learning environment

Open-source software was used to develop and assess the research experiments. The scientific computing libraries in Python were used. TensorFlow computing framework was used to create and train the ConvNeXt model. The Scikit-learn package was used to create the Logistic Regression models. For quick parallel optimization of the learning model, the ConvNeXt model was trained on an NVIDIA A100 SXM4 40 GB GPU. The logistic regression models used less memory and did not need a GPU. The training environment's hardware specs are shown in Table 3.

4.3 4.3 Results and discussion

The biometric verification system decides the authenticity of the claimed identity. If the presented sample matches the reference template above a certain threshold, the individual is considered verified and authenticated. If the match falls below the threshold, the individual may be rejected. In our proposed systems the probability returned from the logistic regression is compared to a threshold value to verify the authenticity of an individual.

We used the validation and evaluation sets to test both of our biometric systems. The fundamental distinction is that the validation set is used to fine-tune the ConvNeXt model's weights, whereas both sets are unknown to the logistic regression model. The validation set is a large set that contains unbalanced samples, whereas the evaluation set is a small set that contains balanced samples (5 samples per user).

To test our first biometric system, we divided the test sets into authenticated and imposter users. We tested the model on the imposter dataset, which we know does not contain any authenticated users, so all predictions will be incorrect because our model was only trained on the authenticated user training set, but we care about the prediction score of how many imposters will pass at each threshold and calculating the FAR by saving the number of

Table 3 Training environment specifications

Model	CPU	GPU	RAM
ConvNeXt	Intel® Xeon® Processor 2.20 GHz	NVIDIA A100 SXM4 40 GB GPU	83.5 GB
Logistic Regression	Intel® Xeon® Processor 2.20 GHz	Not Applicable	25.5 GB

the passed imposters at each threshold, where we define the threshold value from 0 to 100.

We used the authenticated users' test sets to calculate the FRR and ensured that the model predicted accurately. Then, for each threshold, we computed how many authenticated users would be rejected and saved the number of rejected authenticated users at each threshold, where we defined the threshold value from 0 to 100. We find the threshold at which the FAR equals the FRR to calculate the EER for each test set.

The FRR/FAR crossover curves for the validation and evaluation sets are shown in Figs. 19 and 20, respectively. The EER for the validation set is 15.30% at the threshold of 1.50%, while the EER for the evaluation set is 21.72% at the threshold of 1.05%, indicating that the ideal threshold for our biometric identification system will be between 1 and 1.5%. Because the validation set has 6697 authorized user samples and the evaluation set has only 200 authorized user samples, the validation set curves is smoother than the evaluation set curves, resulting in a smoother FRR curve on the validation compared to the FRR curve on the evaluation. Figure 21 illustrates the DET curve derived from the FAR and FRR of the validation and evaluation test sets to demonstrate our result on the crossover curves and summarize the results of the experiment in a single graph.

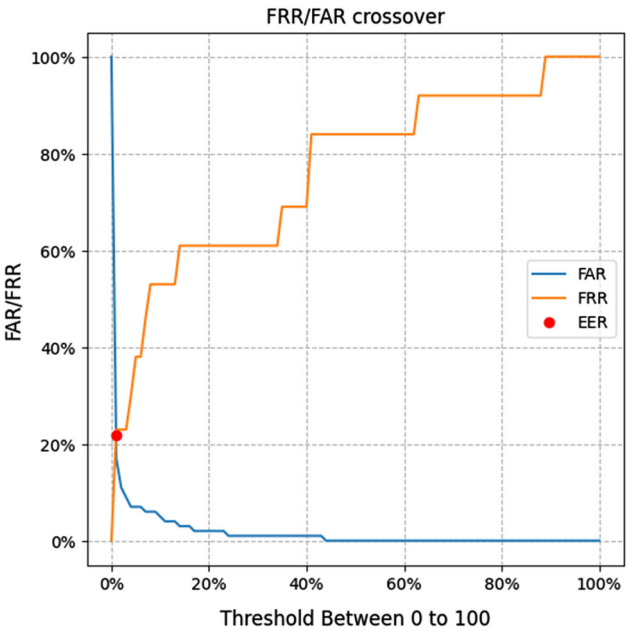

Fig. 20 Evaluation set EER for the biometric identification system

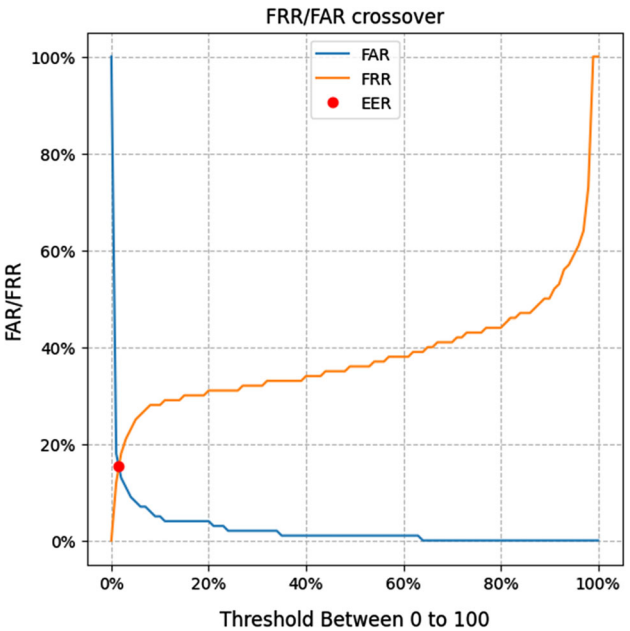

Fig. 19 Validation set EER for the biometric identification system

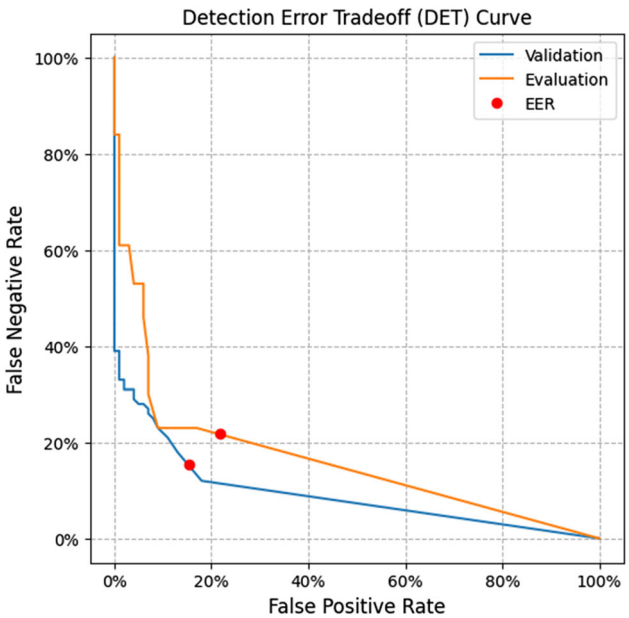

Fig. 21 DET curve for the biometric identification system

We cannot compare our identification system to earlier work because no one has created a single model to identify the forty clients in the baseline benchmark in the SFootBD; hence, we are the first to implement the idea of creating a single model to identify and verify the user based on his footsteps. Although our equal error rate is higher than some of the previous works, which we believe can be reduced in future work. Using only one model capable of identifying a person from his footsteps saves a significant amount of computational resources because training only one model that can be saved and deployed is more efficient and easier than training u models for u users.

To test the second biometric system, the returned probabilities scores returned by the logistic regression models are separated into authenticated and imposter scores to obtain the DET curve and obtain the equal error rate for the biometric verification system, as shown in Fig. 22. For this experiment, forty logistic regression models were trained for the forty users using the features obtained from the ConvNeXt model, and then the models were evaluated using the features obtained from the ConvNeXt model from the validation and evaluation sets. The stride footstep obtained an EER of 6.97% and 10.25% for the validation and evaluation test sets, respectively.

Table 4 compares the outcomes of two biometric identification and verification systems: one with a single classifier for identification, and one with multiple classifiers, one classifier per user for verification scenario.

Table 5 represents some of the best experimental results in the previous works on the main benchmark that we are working on, which considers forty users with forty samples of stride footstep signal per user. This is the most challenging benchmark because it has very few training samples compared to the number of users, which makes this scenario the closest to imitating a real-life security application.

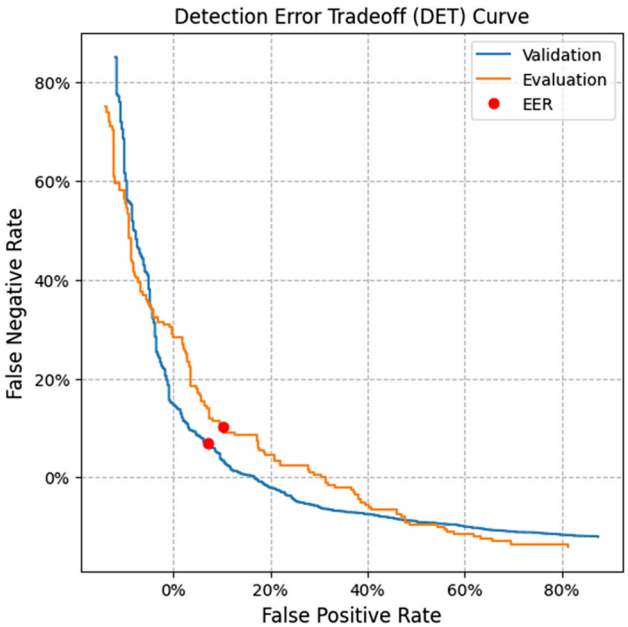

Fig. 22 DET curve for the biometric verification system

When compared to previous work, our proposed methodology is a major advance. In [6], using the same benchmark and a stride footstep, they obtained an EER of 10.51% for the validation set and 18.00% for the evaluation set by using the fusion of the time domain components (GRF, spatial average, and upper and lower contour). In [7], when using the spatial domain components on the same benchmark, they achieved an EER of 10.56% and 16.00% for the validation and evaluation sets, respectively, and for the fusion of the time and spatial domain components at the feature level [8], they obtained an EER of 7.20% for the validation set and 10.70% for the evaluation set using a stride footstep and the same proposed benchmark. Our results show an improvement with an EER of 6.97% and 10.25% for the validation and evaluation sets, respectively, using only the spatial domain components of the footsteps.

In comparison with [10], which obtained an equal error rate of 9.39% for the validation set and 13.86% for the evaluation set, our results show a significant improvement with an EER of 6.97% and 10.25% for the validation and evaluation test sets, and even better than [11], which obtained an EER of 7.10% and 10.50% for the validation and evaluation sets. It is worth noting that in [11], they used raw and processed features, as well as a feature fusion for the temporal and spatial components, whereas in the same experiment using only the processed spatial component by ResNet, they obtained an EER of 13.60% and 15.50% for the validation and evaluation sets, respectively, and using only the processed SVM features, they obtained an EER of 11.70% and 17.50% for the validation and evaluation sets, respectively. So, the novelty of our approach is represented by achieving nearly better results than [8, 11] using only the spatial footstep components without adding the temporal features of the footsteps.

5 Conclusion and future work

In this research, we presented two biometric systems for identification and verification scenarios that used a fused algorithm of the ConvNeXt neural network architecture and a logistic regression classifier. We used the Swansea Foot Biometric Database, the most prominent open-source gait database available for footstep recognition. We used transfer learning on ConvNeXt to optimize computational power and time. ConvNeXt was trained on forty distinct users, each with forty samples of stride footsteps, and a single class was assigned to the imposter samples. We used

Table 4 Results of the two biometric systems

Model	Validation set	Evaluation set
ConvNeXt + One logistic regression	15.30%	21.72%
ConvNeXt + forty logistic regressions one classifier per user	6.97%	10.25%

Table 5 Biometric verification systems results for the related work on the main benchmark

References	Features	Model	Validation set EER	Evaluation set EER
Vera-Rodriguez et al. [6]	Temporal features	PCA and SVM	10.51%	18.00%
Vera-Rodriguez et al. [7]	Spatial features	PCA and SVM	10.56%	16.00%
Vera-Rodriguez et al. [8]	Temporal and spatial features	PCA and SVM	7.20%	10.70%
Costilla Reyes et al. [10]	Spatial features	CNN and SVM	9.39%	13.86%
Costilla Reyes et al. [11]	Raw spatial features	ResNet	16.30%	13.40%
Costilla Reyes et al. [11]	Processed spatial features	ResNet	13.60%	15.50%
Costilla Reyes et al. [11]	Processed SVM spatial features	SVM	11.70%	17.50%
Costilla Reyes et al. [11]	Processed SVM temporal and spatial features	SVM	8.00%	12.50%
Costilla Reyes et al. [11]	Raw and processed temporal and spatial features	ResNet	8.10%	10.70%
Costilla Reyes et al. [11]	Raw and processed and processed SVM temporal and spatial features	ResNet and SVM	7.10%	10.50%
Proposed biometric verification system	Spatial features	ConvNeXt and logistic regression	6.97%	10.25%

the ConvNeXt as a features extractor to extract a 512-feature vector per sample, and we used the extracted features vector to train and test our logistic regression models. The purpose of the first biometric experiment was to obtain an identification biometric system with one single model that can identify and verify the person based on his walking pattern, and the purpose of the second biometric experiment was to obtain a verification biometric system that consists of forty logistic regression models, one for each user, to obtain a biometric system that can verify the claimed identity of a person from his walking patterns. We proved that using only spatial foot pressure characteristics, our proposed models successfully differentiated between the users. In our first biometric system, we achieved an equal error rate of 15.30% and 21.72% for the validation and evaluation sets, and an equal error rate of 6.97% and 10.25% for the validation and evaluation sets in our second biometric system. In the future, we will extend our two proposed systems by fusing temporal and spatial information when training the ConvNeXt and logistic regression models, then we will propose an identification prototype system that makes use of a new walking mat equipped with piezoelectric sensors to collect a small dataset that can fit a workplace and retrain our suggested system. We will then test our model on the proposed walking mat to be able to continuously monitor our biometric identification system and check if the model is performing well or not yet in the production phase, so we can retrain our model if necessary.

Acknowledgements The authors would like to express their gratitude to Dr. Ruben Vera-Rodriguez for providing the Swansea Foot Biometric Database (SFootBD).

Author contributions Designing the algorithm AI, and MA: conducting the experiments AI: writing the manuscript AI: completing the analysis MA: revising the manuscript MA: proofreading MR and EE: write-reviewing, and editing MR, EE. All authors read and approved the final manuscript.

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB). This research received no external funding.

Availability of data and material The Swansea Foot Biometric Database (SFootBD) can be found at http://atvs.ii.uam.es/atvs/sfootbd.html.

Declarations

Conflict of interest All authors declare that they have no competing interest.

Ethical approval and Consent to participate The submitted work is original, and the manuscript has not been submitted to another journal for simultaneous consideration.

Consent for publication The authors declare that they consent to publish the article.

References

1. Unar JA, Seng WC, Abbasi A (2014) A review of biometric technology along with trends and prospects. Pattern Recognit 47(8):2673–2688. https://doi.org/10.1016/j.patcog.2014.01.016
2. Tistarelli M, Li SZ, Chellappa R (2012) Handbook of remote biometrics: for surveillance and security, Springer London. https://doi.org/10.1007/978-1-84882-385-3
3. Costilla-Reyes O, Vera-Rodriguez R, Alharthi AS, Yunas SU, Ozanyan KB (2020) Deep learning in gait analysis for security and healthcare. In: Pedrycz W, Chen SM (eds) Deep learning: algorithms and applications. Studies in computational intelligence, vol 865. Springer, Cham. https://doi.org/10.1007/978-3-030-31760-7_10
4. Singh JP, Jain S, Arora S, Singh UP (2021) A survey of behavioral biometric gait recognition: current success and future perspectives. Arch Comput Methods Eng 28:107–148. https://doi.org/10.1007/s11831-019-09375-3
5. Vera-Rodriguez R, Evans NWD, Mason JSD (2009) Footstep recognition. In: Li SZ, Jain A (eds) Encyclopedia of biometrics. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-73003-5_41
6. Vera-Rodriguez R, Mason JSD, Fierrez J, Ortega-Garcia J (2010) Analysis of time domain information for footstep recognition. In: Bebis G, et al. Advances in visual computing. ISVC 2010. Lecture notes in computer science, vol 6453. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-17289-2_47
7. Vera-Rodriguez R, Mason JSD, Fierrez J, Ortega-Garcia J (2011) Analysis of spatial domain information for footstep recognition. IET Comput Vis 5(6):380–388. https://doi.org/10.1049/iet-cvi.2010.0189
8. Vera-Rodriguez R, Mason JSD, Fierrez J, Ortega-Garcia J (2013) Comparative analysis and fusion of spatiotemporal information for footstep recognition. IEEE Trans Pattern Anal Mach Intell 35(4):823–834. https://doi.org/10.1109/TPAMI.2012.164
9. Vera-Rodriguez R, Lewis RP, Mason JSD, Evans NWD (2008) A large scale footstep database for biometric studies created using cross-biometrics for labelling, In: 2008 10th international conference on control, automation, robotics and vision, Hanoi, Vietnam, p 1361–1366, https://doi.org/10.1109/ICARCV.2008.4795721
10. Costilla-Reyes O, Vera-Rodriguez R, Scully P, Ozanyan KB (2016) Spatial footstep recognition by convolutional neural networks for biometric applications, IEEE SENSORS, Orlando, FL, USA, p 1–3, https://doi.org/10.1109/ICSENS.2016.7808890
11. Costilla-Reyes O, Vera-Rodriguez R, Scully P, Ozanyan KB (2019) Analysis of spatio-temporal representations for robust footstep recognition with deep residual neural networks. IEEE Trans Pattern Anal Mach Intell 41(2):285–296. https://doi.org/10.1109/TPAMI.2018.2799847
12. George G, Oommen RM, Shelly S, Philipose SS, Varghese AM (2018) A survey on various median filtering techniques for removal of impulse noise from digital image, In: 2018 conference on emerging devices and smart systems (ICEDSS), Tiruchengode, India, p 235–238, https://doi.org/10.1109/ICEDSS.2018.8544273
13. Gibson J, Hoontaek O (2020) Mutual information loss in pyramidal image processing. Information 11:6–322. https://doi.org/10.3390/info11060322
14. Getreuer P (2011) Linear methods for image interpolation, image processing on line, p 238–259, https://doi.org/10.5201/ipol.2011.g_lmii
15. Hosna A, Merry E, Gyalmo J, Alom Z, Aung Z, Azim MA (2022) Transfer learning: a friendly introduction. J Big Data 9:102. https://doi.org/10.1186/s40537-022-00652-w
16. Liu Z, Mao H, Wu C-Y, Feichtenhofer C, Darrell T, Xie S (2022) A ConvNet for the 2020s, 2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR), New Orleans, LA, USA, p 11966–11976, https://doi.org/10.1109/CVPR52688.2022.01167
17. He K, Zhang X, Ren S, Sun J (2021) Deep residual learning for image recognition, computer vision and pattern recognition, https://doi.org/10.48550/arXiv.1512.03385
18. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B (2021) Swin transformer: hierarchical vision transformer using shifted windows, computer vision and pattern recognition, https://doi.org/10.48550/arXiv.2103.14030
19. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) ImageNet large scale visual recognition challenge. Int J Comput Vis (IJCV) 115:211–252. https://doi.org/10.1007/s11263-015-0816-y
20. Ciampiconi L, Elwood A, Leonardi M, Mohamed A, Rozza A (2023) A survey and taxonomy of loss functions in machine learning. https://doi.org/10.48550/arXiv.2301.05579
21. Fernandes AAT, Figueiredo Filho DB, Rocha ECD, Nascimento WDS (2021) Read this paper if you want to learn logistic regression. Rev de soc e polit 28:006. https://doi.org/10.1590/1678-987320287406en
22. James G, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning: with applications in R 2nd edition, Springer New York, NY, p 197–209, https://doi.org/10.1007/978-1-4614-7138-7
23. Akkem Y, Biswas SK, Varanasi A (2023) Smart farming using artificial intelligence: a review. Eng Appl Artif Intell 120:105899. https://doi.org/10.1016/j.engappai.2023.105899
24. Akkem Y, Biswas SK, Varanasi A (2023) Smart farming monitoring using ML and MLOps. In: Hassanien AE, Castillo O, Anand S, Jaiswal A (eds) International conference on innovative computing and communications (ICICC 2023). Lecture notes in networks and systems, vol 703. Springer, Singapore. https://doi.org/10.1007/978-981-99-3315-0_51
25. Al-Banhawy N, Mohsen H, Ghali N (2020) Signature identification and verification systems: a comparative study on the online

and offline techniques. Future Comput Inform J 5(1):3. https://doi.org/10.54623/fue.fcij.5.1.3
26. Sivaram M, Ahamed AM, Yuvaraj D, Megala G, Porkodi V, Kandasamy M (2019) Biometric security and performance metrics: FAR, FER, CER, FRR, In: 2019 international conference on computational intelligence and knowledge economy (ICCIKE), Dubai, United Arab Emirates, p 770–772, https://doi.org/10.1109/ICCIKE47802.2019.9004275
27. El-Abed M, Charrier C, Rosenberger C (2012) Evaluation of biometric systems new trends and developments in biometrics. INTECH. https://doi.org/10.5772/52084
28. Martin A, Doddington G, Kamm T, Ordowski M, Przybocki M (1997) The DET curve in assessment of detection task performance. In: Proceeding of the European conference on speech communication and technology (EUROSPEECH), Rhodes, Greece, https://api.semanticscholar.org/CorpusID:9497630
29. Navratil J, Klusacek D (2007) On Linear DETs, In: 2007 IEEE international conference on acoustics, speech and signal processing—ICASSP '07, Honolulu, HI, USA, p 229–232, https://doi.org/10.1109/ICASSP.2007.367205