HotFoot: Foot-Based User Identification Using Thermal Imaging

Alia Saad[1], Kian Izadi[3], Anam Ahmad Khan[4], Pascal Knierim[2,5], Stefan Schneegass[1], Florian Alt[2], Yomna Abdelrahman[2,6]
[1]University of Duisburg-Essen, Essen, Germany, {firstname.lastname}@uni-due.de
[2]University of the Bundeswehr Munich, Munich, Germany, {firstname.lastname}@unibw.de
[3]LMU Munich, Munich, Germany, {firstname.lastname}@campus.lmu.de
[4]National University of Science and Technology, Islamabad, Pakistan
[5]University of Innsbruck, Innsbruck, Austria
[6]European Universities in Egypt, Cairo, Egypt

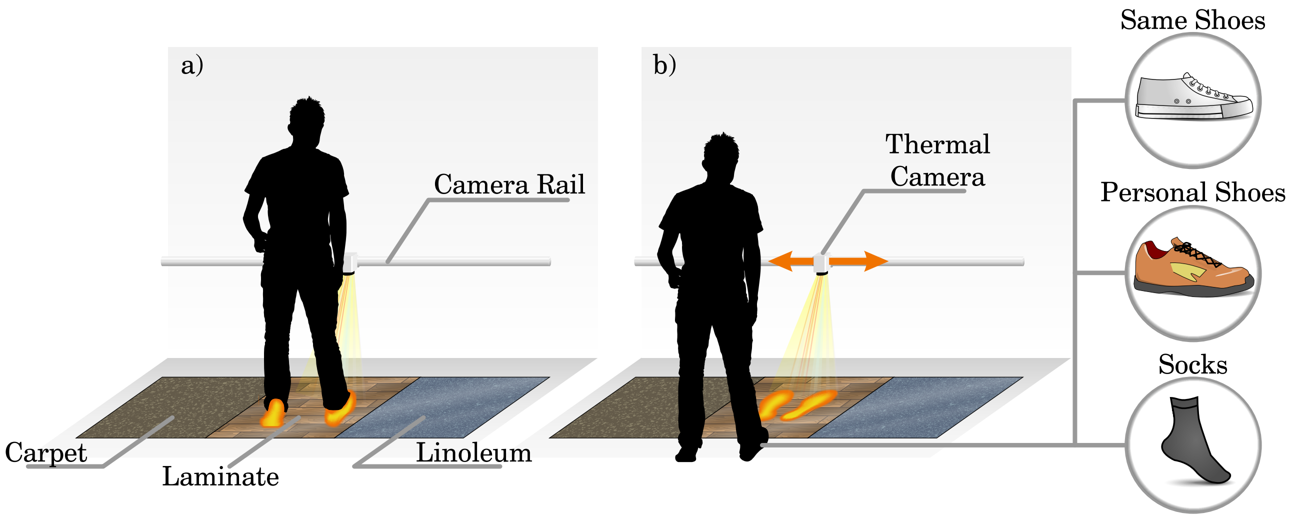

Figure 1: In this work, we explore the use of thermal cameras to capture a user's foot for implicit identification. We investigate the influence of different floor types (carpet, laminate, and linoleum) as well as the influence of footwear (socks, own shoes, standard shoes). With AUC scores up to 98.9%, we show that feet's thermal features can be used as biometrics.

ABSTRACT

We propose a novel method for seamlessly identifying users by combining thermal and visible feet features. While it is known that users' feet have unique characteristics, these have so far been underutilized for biometric identification, as observing those features often requires the removal of shoes and socks. As thermal cameras are becoming ubiquitous, we foresee a new form of identification, using feet features and heat traces to reconstruct the footprint even while wearing shoes or socks. We collected a dataset of users' feet ($N = 21$), wearing three types of footwear (personal shoes, standard shoes, and socks) on three floor types (carpet, laminate, and linoleum). By combining visual and thermal features, an AUC between 91.1% and 98.9%, depending on floor type and shoe type can be achieved, with personal shoes on linoleum floor performing best. Our findings demonstrate the potential of thermal imaging for continuous and unobtrusive user identification.

CCS CONCEPTS

• **Human-centered computing** → *Ubiqu. and mobile computing*.

KEYWORDS

Footprint, User identification, Thermal Imaging

ACM Reference Format:
Alia Saad[1], Kian Izadi[3], Anam Ahmad Khan[4], Pascal Knierim[2,5], Stefan Schneegass[1],, Florian Alt[2], Yomna Abdelrahman[2,6]. 2023. HotFoot: Foot-Based User Identification Using Thermal Imaging. In *CHI Conference on Human Factors in Computing Systems (CHI '23), April 23-April 28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3544548.3580924

CHI '23, April 23-April 28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3580924

1 INTRODUCTION

The omnipresent use of the term 'smart environment' suggests that Mark Weiser's vision of a world in which computers weave themselves into the fabric of everyday life [75] has become true –

a vision in which environments adapt to the users and their context. However, taking a closer look at what is today considered a smart environment – at least from a commercial perspective – is disillusioning. Technologies that turn on/off lights and heating or open/close windows, depending on whether users are at home, build on knowledge about user presence rather than on user identity.

As a result of this, truly personalized applications for smart environments often remain concepts, yet exciting ones. Consider the simple examples of an environment that adapts to an individual user, for example, by playing the user's favorite music or automatically turning on the user's favorite TV show. Or an environment using knowledge of the identity of a user for security purposes, triggering an alarm as a non-legitimate user is detected.

A major reason for which personalization today is rarely seen in smart environments is that the deployment of technology to identify users is a non-trivial task. Generally, approaches should work implicitly, that is without the need for interaction by users, as well as continuously [12]. This is difficult with technologies, such as fingerprint sensors. Approaches for implicit identification include augmenting floors with sensors capable of identifying users [50] or the use of cameras to enable identification based on face recognition or gait [14]. However, also such solutions come with challenges: They are often costly due to the large number of sensors needed. And the placement of the sensors to optimally capture the information required for identification is difficult. Consider a camera to identify users from the characteristics of the face. This would require a set of cameras placed around a space to continuously capture the face of the user.

In this paper, we propose and explore a novel solution to this challenge, which is the use of *thermal imaging to identify users from the thermal characteristics of their feet and the heat traces left behind.* We demonstrate that using a thermal imaging camera deployed at a top view setup allows for seamlessly identifying users of a smart environment (cf. Figure 1). As part of our investigation, we explore how the type of footwear, as well as the type of floor, influence our approach, so as to demonstrate the feasibility of the approach in different settings (cf. Section 3).

Our research approach is as follows: First, we built a detection system, to capture the data necessary to identify users. Second, we collect a data set of thermal images under realistic conditions: we capture data from three floor types (carpet, wood, and linoleum) and from users wearing different footwear (socks, personal shoes, standard shoes). Third, we built predictive models, demonstrating how floor types and footwear influence identification accuracy.

Our results show that by using visual features only, an accuracy of 73.1% to 84.1% can be achieved. In contrast, the use of thermal features yields accuracies of 89.1% to 98.8%. Combining visual and thermal features further improves accuracy: a classifier identifying users independent of the footwear and floor type can predict users' identity with 91.7%. In the best case (personal shoes on a linoleum floor) an accuracy of up to 98.9% can be achieved.

We release both our system implementation and our dataset as open source for researchers to build upon, replicate, and extend our work[1]. We hope our system provides an easy-to-build and deploy solution for smart environment applications that rely upon identifying users.

[1]http://tiny.cc/HotFoot

Contribution Statement. We advance the state-of-the-art in user identification through the following contributions (according to Wobbrock et al.'s classification of research contributions in HCI [76]). (1) We propose a novel approach for identifying users based on the thermal and visual features of their feet. (2) We provide a proof-of-concept implementation and collect a dataset, both of which we release as open source for future research. (3) We report on an exploration of the accuracy of our approach, investigating the influence of floor type and footwear.

2 BACKGROUND AND RELATED WORK

Our work draws from several strands of prior research, most importantly, foot-based interaction, biometric identification, including relevant ML techniques, and applications of thermal imaging.

2.1 Foot-Based Interaction

Different approaches exist that enable users to interact with their feet. An overview of work on foot-based interaction in the HCI community is provided by Velloso et al. [72]. Prior approaches vary in level of sophistication and regarding the used technology. A simple example is Vote-with-your-Feet [67]. Here, two tangible buttons on the ground enable users to cast a vote for a question shown on a public display. A second example is ShoeSense, an approach using a sensor on the shoe to identify gestures, for example, for controlling a music player [15]. Richter et al. proposed a similar solution to identify tabletop users based on their shoes, captured by top-view depth cameras attached to the tabletop [57]. Another work identified infants from footprints [13, 41]. While many other interaction examples exist, the identification of people based on foot characteristics has received little attention [29, 49, 50]. Approaches mainly focused on providing a proof of concept as to how individuals [10, 37] can be identified based on image processing as well as based on floor sensors [73].

2.2 Biometric Identification

Identifying people based on their feet is a biometric approach. We briefly introduce the fundamentals behind biometric identification approaches, provide some examples, and then explain how a biometrics-based system can be built.

2.2.1 Definition and Fundamentals. According to Jain et al., 'biometrics' refers to the process of identifying a person based on their physiological or behavioral traits [35]. They considered the identification based on biometric features as a novel and robust technique, capable to overcome limitations of traditional, possession-based (i.e. keys or tokens) and knowledge-based (i.e. passwords, PINs, patterns) authentication approaches. The authors defined four criteria to be fulfilled so that a physiological or behavioral trait qualifies as a biometric feature: they should be *universal* and *unique*, i.e. everyone should have the feature while at the same time being distinctive among people. Furthermore, they should be *permanent*, that is the feature should stay consistent over time, and *collectable*, that is being measurable quantitatively [35]. The efficiency of a practical biometric system is evaluated using three additional criteria: *performance*, that is high identification accuracy rates; *acceptability*, that is the willingness of people to adopt the system; and *circumvention*, that is the robustness of the system to spoofing attacks [36].

It is worth noting that (as opposed to knowledge or token-based schemes) biometric approaches provide a probability of users' identity. As a result, biometric identification systems rely on a threshold, beyond which the system considers the identity of an individual to be confirmed [45, 51]. A high threshold (for example, 99.9% accuracy) increases security but comes at the expense of users' being falsely rejected. Hence, what is an acceptable accuracy strongly depends on the use case. Whereas for authentication a higher accuracy is desirable, for use cases in which identification serves as a means to personalize content a lower accuracy might be acceptable.

2.2.2 Examples of Biometric Systems.

Much research exists on biometric identification based on physiological features. The most explored approach is recognition based on facial features [1, 34, 39]. Additionally, hand geometry, including palm [19, 27, 46, 82], finger [11, 18, 80], and finger-knuckle prints [48], as well as ear recognition [2] have been investigated. To improve accuracy, researchers adopted a multi-modal approach, combining two or more features, such as face and fingerprint [31], or face and handprint [61].

More recently, biometric approaches based on human behavior shifted into focus. Researchers primarily focused on the use of gait [58, 74], keystroke dynamics [16, 63, 83], touch gestures [47, 53], and mouse movements [64, 81].

Some research exists, that tried to identify users from foot characteristics. Existing research mainly provided a proof of concept as to how individuals can be identified based on image processing [10, 37] as well as based on floor sensors [73]. Both approaches differ in the capturing technique: the former approach captures data on feet characteristics through a sensor installed in a fixed position, whereas the latter approach requires multiple pressure sensors integrated into the floor to identify a person while walking [78].

2.2.3 Machine Learning Models for Biometric Identification.

All biometric-based identification systems follow a similar design approach [22, 56]. Initially, user data on the biometric feature of interest is acquired using sensors (cameras, inertial sensors, etc.). This data then serves as a basis for building a classifier, used later on to predict the identity of a user. To do so, in a preprocessing step, the most important features (i.e. those features most unique to single users) are identified and then used as input for training the classifier through machine learning (ML) or deep learning (DL). Ultimately, user data for identification is captured and tested against the classifier, returning a probability for the identity of the current user.

Technical advances in ML algorithms allowed the performance and robustness of biometric systems to be considerably enhanced [14, 38, 65]. For instance, Darwish et al. proposed a task-independent method that leveraged gaze data to build a predictive model for biometric identification [65]. The authors engineered a large set of eye movement features and built a random decision forest model that successfully identified people with an average accuracy of 88%. Due to their ability to learn features from heterogeneous datasets, deep learning models are now extensively used for biometric authentication and verification [70]. For instance, multiple attempts have been made to train convolution and recurrent neural networks to learn features and develop models that can distinguish between large numbers of individuals by leveraging physical (e.g., facial images [66], finger [68] and palm prints [23]) and behavioral (e.g., gait [79] and keystrokes [69]) identifiers.

2.3 Thermal Imaging

2.3.1 Functionality.

Thermal cameras capture the far-infrared spectrum (i.e. wavelengths between 7.5 and 13 μm). This enables capturing the heat map of the camera's field of view. As thermal cameras operate in a different spectrum than RGB cameras, they are capable of capturing distinctive properties of the spectrum. For instance, thermal cameras capture heat radiation and reflections [5, 8, 62], which can then be visualized using a false color mapping.

2.3.2 Research in HCI.

Thermal imaging has been used in several application areas in HCI. For example, prior work utilized thermal reflection to capture users' hand movements and thus enable gesture interaction [6, 26, 62]. As thermal imaging works independently of lighting conditions, it has been used as an alternative for detecting faces and hands as well as their properties based on RGB data. Thermal imaging provides information about the observed skin temperature, which can be used to infer the physiological and cognitive state of users [4, 7, 32, 33]. Another thermal property captured by thermal imaging is heat transfer. Thermal imaging is capable of detecting actions even after being performed. When a user interacts with a surface, the heat is transferred from the user's hand to the surface, leaving behind a heat/cold trace taking time to decay. Thus, the trace is recognizable by a thermal camera even after interaction takes place. Heat traces have been utilized for input detection [40, 62], device state detection [55], reconstructing PIN and authentication patterns [3], extracting veins patterns [25], as well as for forensic identification based on thermal hand print [19]. One example where thermal imaging has been used to identify people from behavior is the work of Cho et al. [19]. The authors applied the Heat-Earth Mover's Distance (HEMD) similarity metric, to identify people based on their handprints' thermal images.

2.3.3 Foot-related Thermal Imaging Research.

A few examples exist, where thermal cameras have been used in research related to human feet and walking behavior. One example is research in law enforcement and forensics [20, 77], where heat traces left from shoes have been used to verify how much time a person has spent in a room. In medicine, researchers captured the temperature of diabetic patients' feet to assess their health status [54].

2.4 Summary

We learn that human feet have been at the focus of HCI research with a main focus on interaction, showing that human feet are a rich source of information. At the same time, thermal imaging has several applications in HCI, though research on identifying humans from thermal imaging data is scarce. In our work, we explore user identification as an application area that fuses knowledge from the aforementioned areas of research. Hereby we draw from a third strand or prior work, that is, work on biometric identification. The ability to identify people from different physiological and behavioral traits motivated our attempt to identify users from their feet thermal properties.

3 HOTFOOT: CONCEPT & METHODOLOGY

Prior work demonstrated that people can generally be identified from the geometric features of their feet [50, 71]. At the same time, this approach is limited in that it requires users to take off shoes, making it impractical for use in many everyday life scenarios. With thermal imaging, additional knowledge on users' feet becomes available: (a) the thermal radiation from the feet and (b) the heat traces left behind on the floor. In the following, we introduce use cases and explain the concept behind our work from a technical perspective (which features are available and can be used for identification), and present the main questions driving our research.

3.1 Use Cases

We envision an environment equipped with top-mounted thermal cameras (e.g., on the ceiling integrated with smoke detectors). These can capture both users' feet and heat traces. This data can then serve as input for an identification system.

3.1.1 Application Scenarios. Multiple applications can benefit from this light-independent, seamless, and affordable identification method. In a *smart home* environment, a user could be identified based on their feet's thermal characteristics and, accordingly, the smart home could adapt the settings to personal preferences. Furthermore, our work could be used in *(semi-) public places*, such as offices or airports. Applications could show personalized way finding information. As the approach does not require hand interaction, users could carry objects and still be identified based on their foot biometrics. Also, the approach could be used for continuous, implicit *authentication*, for example, in a work setting. The advantage is that users do not need to explicitly engage with an authentication task, such as entering a password or PIN. We explore the following settings:

Identifying Users Wearing the Same Shoes in Environments with Consistent Floor Type We explore how well our approach works in a setting in which users wear the same shoes and in which the floor consists of consistent material. An example would be the airport scenario.

Identifying Users Wearing the Same Shoes in Environments with Different Floor Types We explore settings in which users wear the same shoes, but floor types differ. An example is a work setting where the floor in different rooms consists of different materials.

Identifying Users Wearing Different Shoes in Environments with Consistent Floor Type We look at settings where the same type of floor is used across rooms, but where people change shoes. Examples could be at home, where people may wear street shoes in the entrance area and socks or slippers in other parts of the home. Similarly, in work settings, people may be required to change footwear (e.g., hospital clogs). Another example of this setting is cases in which users revisit the same location with different shoes.

Identifying Users Wearing Different Shoes in Environments with Different Floor Type Finally, environments exist where people wear different footwear and which consist of different floor types. Examples could be again home or work environments where people use different types of shoes but with different types of floors.

3.2 Feature Selection

A thermal camera captures temperatures as thermal and visual data, allowing thermal and visual features to be derived. In the following, we explain how different features are derived from the thermal imaging data.

3.2.1 Visual Features. A thermal camera generates false color images. We extract the visual features by analyzing the recorded video frames. We use open CV libraries to extract those features. We compared different, commonly used feature extractors, mainly Speed-Up Robust Features (SURF) [17] and Scale-Invariant Feature Transform (SIFT) [42]. The latter one yielded better results for our data.

3.2.2 Thermal Features. We also use heat traces and thermal features of the foot. Heat traces emerge from heat transfer, a phenomenon when two objects (feet and floor) come in contact. The amount of heat transferred relies on the surfaces' thermal properties, commonly referred to as thermal contact conductance [21].

As an object gets in touch with a surface, heat is transmitted and absorbed by the surface, causing a temperature change at the point of contact, leading to a heat trace left on the surface (i.e. footprint). The amount of the transferred heat influences the quality of the footprint (i.e. the thermal and visual features) and, hence, the identification performance. The amount of heat transferred $T_{contact}$ depends on the temperature of the contact objects ($T_{Footwear}$ and T_{Floor}) and their *thermal penetration coefficient (b)*, where b is the product of thermal conductivity (K), thermal density (P), and specific heat capacity (C) [52]. ($T_{contact}$) is quantified as follows:

$$T_{contact} = \frac{b_{Footwear}T_{Footwear} + b_{Floor}T_{Floor}}{b_{Footwear} + b_{Floor}} \tag{1}$$

$$b = \sqrt{K.P.C} \tag{2}$$

Different footwear and floor types have different thermal properties which influence the amount of heat transferred $T_{contact}$ (as depicted in Equation 1). For instance, *linoleum*, *carpet*, and *laminate* have different thermal conductivity [9]. These differences will influence the thermal penetration coefficient which in turn will influence the $T_{contact}$. In other words, the heat trace as well as the heat signature of the users' feet will be different. As this difference is likely to influence identification performance, we investigate the influence of the different floor types and footwear on identification accuracy.

The thermal data are exported from the camera, representing the non-visual temperature values such as min, max, and average temperatures. As thermal cameras are insensitive to color variations, the identification is not affected by different factors such as skin, socks, or shoe colors. Additionally, thermal imaging is contactless and non-invasive. Users are not required to take off shoes or socks.

We define six regions of interest (ROIs): heel, middle, and toe areas for both feet. These regions are selected based on prior work on classifying the feet into the abovementioned regions [60]. Additionally, feet have different shapes in nature. We expect them to exhibit different thermal characteristics upon walking. Thermal information is derived from the temperature data for each ROI. Thermal and visual features are determined for both the feet (i.e., as users appear) and the footprints (i.e., after leaving).

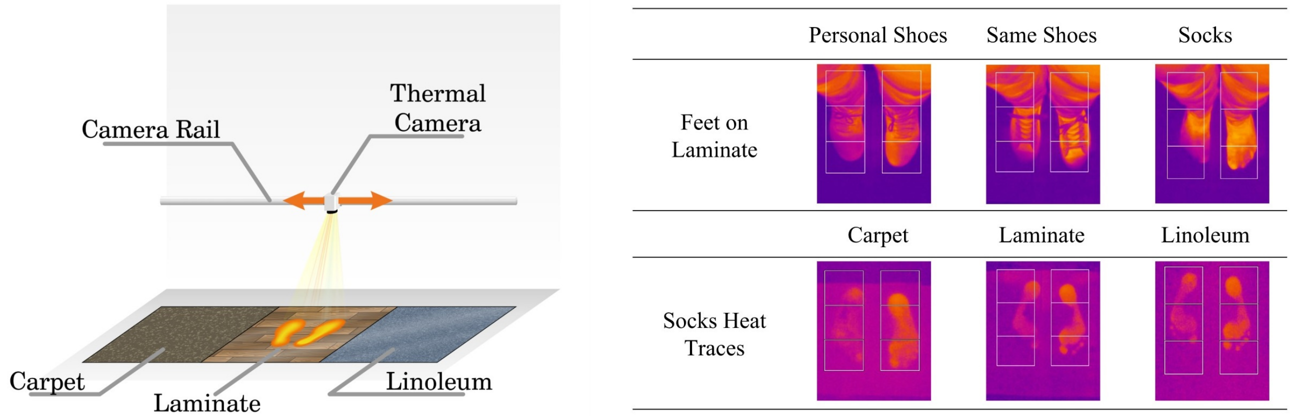

Figure 2: Left: Experimental setup consisting of a thermal camera facing the three different floors. Right: Sample of the ROIs of the feet and traces of one participant in different shoes and on Laminate being monitored by the thermal camera (RightFootHeel, RightFootMid, RightFootToes, LeftFootHeel, LeftFootMid, and LeftFootToes).

3.3 Research Questions & Research Approach

Our investigation is guided by the following research questions:

- **RQ1 –** How accurately can users be identified based on foot features captured by a thermal camera?
- **RQ2 –** How do different floor types (i.e. thermal properties) influence the identification performance?
- **RQ3 –** How does footwear influence identification performance?

To answer these questions, the following research approach, commonly applied in biometrics research, is employed: first, we collect a *data set* of thermal images under different conditions. Second, we build four different *classifiers*, according to the settings described above. Third, we evaluate the *performance* of the classifiers.

4 STEP 1: COLLECTING THERMAL IMAGES FOR FOOT-BASED USER IDENTIFICATION

We collected a dataset in a controlled environment to investigate how accurately users can be identified based on their feet using visual and thermal features. We consider different floor types and footwear, to understand the influence on identification accuracy.

4.1 Experimental Design

The data collection follows a repeated measure within-subject design, i.e., data from participants were collected in all conditions. The two independent variables are Floor Type and Footwear with three levels each. For the Floor Type, we chose to collect data for three of the most commonly used flooring types [30]: *carpet*, *laminate*, and *linoleum*. In addition, each floor type has different thermal properties (cf. section 3.2.2). Footwear comprises *socks*, *personal shoes*, and white, low-cut sneakers (*standard shoes*) provided in fitting sizes. For each of the 3×3 conditions, we recorded thermal videos of participants' feet and the resulting heat trace per Floor Type and per Footwear. Each condition was captured three times,allowing us later to use two samples per user and condition for training and the remaining sample for testing.

4.2 Apparatus

We used the Optris PI450 thermal camera[2], along with the Optris PI Connect software[3] for data extraction. The camera was mounted on a rail 1 m above the floor (cf. Figure 2–left), facing different flooring types that are already placed side-by-side. The camera has an optical resolution of 382×288 pixels, a frame rate of 80 Hz, and captures temperatures between -20°C and 900°C, with thermal sensitivity of 0.04°C. The thermal camera is connected to a 14" windows operated laptop via USB for power supply and data transfer.

The Optris PI connect software used has a built-in automated annotation function, using the so-called measure areas[4] of 75×65 pixels. We annotated the regions of interest including RightFootHeel, RightFootMid, RightFootToes, LeftFootHeel, LeftFootMid, and LeftFootToes (cf. Figure 2–right). Using the Optris PI connect built-in save option, we saved the temperature values of the annotated regions in CSV files, corresponding to a participant per condition.

In each recording session, we created two files: a data file containing the thermal data of the six feet regions, along with the corresponding timestamps, and the thermal imaging feed stored as Radiometric Audio Video Interleave.

4.3 Participants and Procedure

We invited 21 participants (8 females, 13 males) via social networks and mailing lists. The age ranged from 18 to 58 years (M = 27.45, SD = 8.67). We also recorded participants' weight, height, and shoe size (cf. Macdonald et al. [44]). Participants' weights ranged from 53 to 89 kg (M = 69.31, SD = 10.51). Height ranged from 155 to 189 cm (M = 174.22, SD = 7.89). Participants' shoe sizes varied between 37 and 46 (M = 41.5, SD = 2.55, European size system). All participants

[2]https://www.optris.global/thermal-imager-optris-pi-400i-pi-450i, last accessed: March 14, 2023
[3]https://www.optris.global/downloads-software, last accessed: March 14, 2023
[4]https://www.optris.com/software-tutorial-pix-connect-measure-areas, last accessed: March 14, 2023

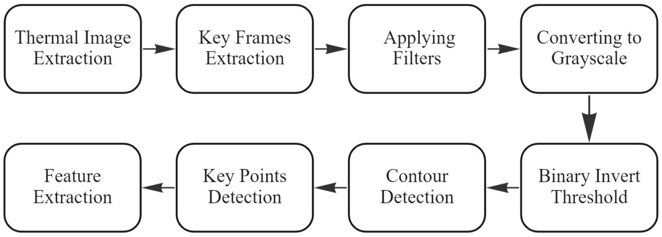

Figure 3: Visual Feature Extraction Pipeline

were healthy individuals with neither amputated limbs nor any chronic disease influencing their body or feet temperature.

We asked participants to bring socks and their everyday shoes. The data collection was conducted in a lab space[5]. After welcoming participants, we asked them to fill out a consent form and a demographic questionnaire. Each recording session started with personal socks, followed by personal shoes and standard shoes.

Each session began with capturing the floor. Participants were asked to step on and off three predefined spots on each floor type, to ensure that the feet are placed correctly for the automatic annotation of the regions of interest. Furthermore, this allowed us to parallelize the recordings without risking heat being transferred by mistake, affecting consecutive recordings. For each pre-defined spot, participants were asked to step on the marked spot in the field of view (FOV) of the camera and immediately step out, to capture realistic behavior, as shown in Figure 2–left. The camera kept recording the scene after both feet were out of view for 3 minutes to capture the feet's heat trace. The process was conducted for all floor types, three times with socks only, three times with participants wearing their own shoes, and three times with participants wearing the standard shoe.

For each participant, we captured 27 recordings (3 floor types $\times$ 3 footwear $\times$ 3 repetitions). The data acquisition resulted in a total number of 567 recordings (21 participants $\times$ 27 recordings per participant). The average duration of the study was 60 minutes in a maintained room temperature of $24°C$. Participants were compensated with 10 EUR.

4.4 Data Processing

We performed a series of prepossessing steps to extract the foot features for each recording. First, we marked the step-in frame, i.e. the frame where the participant's foot appeared in the camera's FOV. To ensure consistent data capturing and analysis across participants, we systematically identified the temperature change threshold for automatic detection of the step-in time. A minimum threshold of $1.5°C$, in 7 consecutive frames showed the best results. Similarly, the step-out frame is marked, when the temperature significantly decreases over the course of 7 consecutive frames. All detected temperatures before the step-in are discarded.

For the visual analysis, additional pre-processing steps were performed for each frame. We used the OpenCV library[6] for image processing and feature extraction as follows:

[5]All local health and hygiene regulations to prevent the spread of COVID-19 and create a safe study environment were implemented.
[6]https://opencv.org/, last accessed: March 14, 2023

(1) **Noise filtering:** We applied a 5×5 median filter to smooth the image and converted the output to gray-scale.
(2) **Background Subtraction:** We applied an inverse binary. Resulting images were used to find feet contours.
(3) **Features Extraction:** We used the Scale-Invariant Feature Transform (SIFT) algorithm [42], to extract the key points and descriptors, as shown in Figure 3.

4.4.1 Thermal Features. The thermal model is based on the thermal information extracted from the data files. As depicted in Figure 2–right, we defined three regions of interest for each foot, representing the toes, middle and heel areas, resulting in a total of six different regions per recording session. We extracted seven different thermal features, as shown in Table 1: the maximum temperature value, the minimum temperature value, the difference between the maximum and minimum values, the arithmetic mean and standard deviation of the whole foot's temperature between the step-in and step-out frames, the temperature distribution, i.e. (the weight of the region temperature among all 6 regions, computed as the average of the region over the average of both feet; the 6 regions distributions add up to a total of 1), the maximum temperature detected in all six regions, and the time required for the thermal traces to decay over time [3]. We chose to collect both feet and resulting heat traces in addition to the decay time. This allowed us to obtain a comprehensive understanding of possible thermal features, suitable to be adopted in various use case scenarios.

4.4.2 Visual Features. The second model is based on the visual features extracted from the thermal video frames. In each video, we use four frames with both participants' feet visible in the camera FOV, to generate the visual biometric features. Every frame was counted as an individual instance. This led to a total of 108 (4 frames $\times$ 3 repetitions $\times$ 3 floor type $\times$ 3 footwear) different frames to identify an individual, and 2268 frames for all 21 participants.

4.5 Limitations

Our work has several limitations. Firstly, we simulate the best-case scenario for identification (i.e. clean foot heat traces without overlaps from previous steps). We acknowledge that in real-world scenarios this might be different. We believe with more advanced pre-processing of the data, traces could be restored properly. Additionally, participants had only one type of personal shoes, stepped onto a single type of floor, and the process of stepping in and out of the camera ensured no overlapping. The aim of the study was to investigate the viability of thermal imaging and the captured features to identify users.

Secondly, we recorded data in a single session only. However, previous works showed that thermal prints collected in a single session were robust and yielded a robust identification [19].

Thirdly, we opted for SIFT features extraction as an initial exploration of our approach. Furthermore, we only considered a single set of hyperparameters for our classifiers. Future work could consider the latest feature extraction techniques as well as optimize hyperparameter values, so as to obtain a better understanding of how the accuracy of the proposed approach could be further enhanced.

Table 1: Description of the Thermal Features.

Thermal Feature	Description
Maximum Region Temperature	maximum temperature value of each foot region.
Minimum Region Temperature	minimum temperature value of each foot region.
Difference between Maximum and Minimum Temperatures	difference between the min and max values of each foot region
Mean and Standard deviation of Foot Temperature	overall average and the standard deviation temperature per foot.
Temperature Distribution	weight of the region temperature relative to all 6 regions
Maximum Foot Temperature	maximum temperature per foot
Decay time	time it takes for the heat trace to disappear.

Fourthly, we tested our approach with 21 people only. While future work could explore how the approach scales to larger populations, our exploration still demonstrates the utility of the approach for home settings or environments with a moderately sized user group. Lastly, we focus on individuals with two healthy feet, leaving an investigation of people with chronically cold feet or people with only one limb for future work.

5 STEP 2: CLASSIFICATION

We modeled the task of biometric identification as a multi-class classification problem, where a feature vector was fed into a classifier to predict one of the 21 classes that correspond to the identity of the participant. We built four classifiers that correspond to the settings described in subsection 3.1: (1) a *floor-dependent, footwear-dependent classifier*, trained on data of all participants using the same floor and footwear condition respectively; (2) a *floor-dependent, footwear-independent classifier*, trained on data from the same floor but data from different footwear conditions; (3) a *footwear-independent, floor-dependent classifier* trained on data from the same floor but different footwear conditions; and (4) a *footwear-independent, floor-independent* classifier trained on data from different floor and different footwear conditions.

To build classifiers capable of identifying users, we split the data of each user into a training set and a test set. We do not have a validation set, as we are not changing the hyperparameters of our classifiers [59]. We opted for leaving one-sample out rather than one-participant out, as we are classifying participants, not conditions. This was possible as for each condition, three samples have been collected per user. In line with prior work on biometric identification [24], we used two of the samples for training the classifier and one sample for testing.

To train the classifiers, we experimented with Logistic Regression, Extreme Gradient Boosting (XGBoost), and Random Forest models. We observed that the Random Forest model provided the best performance score across the testing folds. Therefore, we trained all our proposed classifiers with a Random Forest model on the dataset of 21 participants. We tuned the Random Forest model with 100 trees, a maximum tree depth of 80, 5 samples for splitting the internal node, 2 samples to be at a leaf node, and entropy as the splitting criteria for building the tree. To assess the performance of the classifiers, we used the Area Under the ROC Curve (AUC) and F_1-Score [28], similar to related work. The F_1-Score is the weighted average of precision and recall. The AUC score represents the ability of a classifier to distinguish between classes and provides a measure of performance across all classification thresholds. For both metrics, the score ranges from 0 to 1, with 0 being the lowest and 1 being the highest value the model can achieve regarding performance. In the remainder of this section, we provide more details on the four different classifiers. Later, Table 2 shows the results corresponding to footwear-dependent and floor-dependent classifiers, presented in 5.1. The values in the rows of Table 3 result from the classifiers explained in 5.2, 5.3, and 5.4, respectively.

5.1 Footwear-Dependent and Floor-Dependent Classifier

First, we built a classifier for the case in which users wear only one type of footwear and in which only one type of floor exists. This could be a typical office setting. This constrained setting is expected to yield the highest accuracy. To do so, we trained and evaluated the classifier 9 times using all features – once for each footwear/floor combination. For example, we trained the classifier on the data of the *socks only* footwear using *carpet* flooring and then evaluated the classifier on the data of the *socks only* footwear and *carpet* flooring. We split the data based on the number of samples of each participant for the specified conditions. For this purpose, we used leave-one sample-out cross-validation which ensured that for each participant the classifier was trained on two samples in the specific condition, and was tested on the third sample (the total number of samples per participant for the specified conditions was three).

5.2 Footwear-Dependent and Floor-Independent Classifier

Second, we built a classifier for cases in which users wear the same footwear but on different types of floors. This corresponds for example to the setting of an office, in which users over the day wear the same pair of shoes, but where there are different types of floors in the aisles, offices, and coffee kitchens. This classifier was built by training on the data of all participants with the same footwear conditions but different floor conditions. To do this, we trained and evaluated the classifier 3 times using all features – each time for specific footwear but different floor conditions. For example, we trained the classifier on the data of a particular footwear condition using *linoleum* and *laminate* floor condition and evaluated the classifier on the data of the same footwear but the *carpet* floor condition. The reported results, in the next sections, are averaged by footwear condition but split by the floor condition on which it was evaluated.

Table 2: Random forest model Performance for footwear-dependent floor-dependent conditions.

	Visual		Thermal		Visual & Thermal	
Condition	F_1-Score	AUC	F_1-Score	AUC	F_1-Score	AUC
Linoleum-Personal Shoes	21.7 ± 4.1%	81.6 ± 4.8%	61.1 ± 7.2%	98.8 ± 0.8%	86.8 ± 4.7%	98.9 ± 0.9%
Linoleum-Dedicated Shoes	21.8 ± 6.2%	82.8 ± 1.2%	63.9 ± 1.3%	98.2 ± 1.3%	70.2 ± 7.5%	97.1 ± 1.1%
Linoleum-Socks Only	30.8 ± 0.3%	84.1 ± 1.7%	46.1 ± 0.2%	95.8 ± 1.3%	63.1 ± 1.2%	95.1 ± 2.5%
Laminate-Personal Shoes	24.7 ± 3.1%	79.5 ± 0.6%	49.3 ± 0.9%	91.5 ± 1.3%	56.3 ± 0.1%	95.2 ± 1.6%
Laminate-Dedicated Shoes	20.1 ± 3.1%	81.1 ± 2.0%	46.5 ± 1.6%	89.1 ± 5.5%	51.9 ± 0.9%	89.6 ± 0.5%
Laminate-Socks Only	18.5 ± 0.5%	77.8 ± 0.5%	44.9 ± 1.2%	91.3 ± 0.8%	49.9 ± 0.7%	91.9 ± 0.2%
Carpet-Personal Shoes	21.3 ± 4.1%	78.9 ± 2.6%	74.8 ± 1.7%	97.9 ± 0.4%	74.2 ± 1.4%	97.2 ± 3.2%
Carpet-Dedicated Shoes	23.8 ± 7.0%	78.9 ± 2.2%	60.1 ± 1.4%	97.7 ± 0.4%	69.5 ± 0.8%	98.2 ± 1.1%
Carpet-Socks Only	21.3 ± 2.0%	73.1 ± 1.1%	40.4 ± 4.2%	94.1 ± 2.4%	55.1 ± 0.5%	96.0 ± 0.1%

Table 3: Random forest performance for classifiers that are either floor-independent (i.e., tested with one type of floor and trained with the other two types of floors), footwear-independent (i.e., tested with one type of footwear and trained with the other two types of footwear), or floor- and footwear-independent (i.e., tested with floor and footwear unknown to the classifier). The reported results are averaged by the dependent condition (e.g., footwear dependent - floor independent classifiers are averaged over all three types of footwear).

		Visual		Thermal		Visual & Thermal	
Classifier	Testing Condition (Training Conditions)	F_1-Score	AUC	F_1-Score	AUC	F_1-Score	AUC
Footwear Dependent-Floor Independent	Laminate(Linoleum & Carpet)	20.7 ± 2.4%	76.4 ± 2.1%	50.4 ± 3.5%	92.9 ± 2.7%	55.4 ± 8.9%	91.9 ± 2.1%
	Linoleum (Laminate & Carpet)	23.4 ± 3.4%	80.7 ± 1.7%	63.6 ± 3.1%	96.1 ± 1.2%	70.3 ± 4.5%	96.5 ± 0.7%
	Carpet (Laminate & Linoleum)	18.2 ± 1.3%	74.7 ± 1.5%	57.7 ± 2.5%	94.9 ± 2.8%	57.1 ± 5.0%	96.1 ± 0.07%
Footwear Independent-Floor Dependent	Personal Shoes (Same Shoes & Socks)	22.6 ± 3.7%	80.1 ± 2.7%	61.7 ± 1.1%	96.1 ± 1.3%	69.4 ± 0.1%	96.1 ± 1.9%
	Same Shoes (Personal Shoes & Socks)	21.9 ± 5.4%	80.9 ± 1.8%	56.9 ± 1.4%	95.0 ± 2.4%	63.9 ± 0.8%	95.1 ± 2.4%
	Socks Only (Personal Shoes & Same Shoes)	23.5 ± 1.9%	78.4 ± 2.8%	43.8 ± 0.6%	93.7 ± 1.4%	56.1 ± 0.4%	94.3 ± 1.9%
Footwear Independent-Floor Independent		16.15 ± 2.6	72.9 ± 1.5%	47.6 ± 1.2%	91.6 ± 2.8%	51.2 ± 2.6%	91.7 ± 1.5%

5.3 Footwear-Independent and Floor-Dependent Classifier

Third, we built a classifier to reflect settings in which users would change their shoes but where the floor type would be consistent. This could be a workplace setting with the same floor being used everywhere but where for certain tasks users would switch their shoes (e.g., safety shoes or specific shoes to work in an operating theater). This classifier was built by training on the data of all participants with the same floor conditions but different footwear conditions. We trained and evaluated the classifier 3 times, each time for a specific floor condition but different footwear conditions. For instance, we trained the classifier on the data of a particular floor condition using *socks only* as well as *personal shoes* and evaluated the classifier on the data of the same floor but with the data from the *standard shoes* condition. The results in the next sections are averaged by floor type but split by the footwear condition on which it was evaluated.

5.4 Footwear-Independent and Floor-Independent Classifier

Finally, we built a classifier to reflect completely unconstrained settings with users wearing different shoes and using different types of floors. This could be a home environment in which the corridor, kitchen, and living room would have different types of flooring and where users could wear street shoes, slippers, or socks. This classifier was built by training on the data of all participants with different floor and different footwear conditions. We evaluated the condition-independent classifier by training it 3 times using leave-one footwear-out cross-validation three times – one for each floor condition. Each time, we trained it on the data of the two footwear and for the two-floor conditions and evaluated it on the data of the third-floor condition and from the last footwear. The reported results are averaged over all footwear and floor conditions.

6 STEP 3: EVALUATION

We first report and compare the performance of classifiers trained on thermal, visual, and the combination of both features. Furthermore, we explore the reasons why certain visual and thermal features have more impact on the performance of the built classifiers for the task of biometric identification. We report the F_1 Score and AUC to evaluate the classification performance of the models, shown in Table 2 and Table 3. Furthermore, we report the precision and accuracy scores of the built classifiers, reported in Tables 4 and 5.

6.1 Classification Performance per Condition

Our results show that the proposed approach of leveraging visual and thermal features for predicting the user's identity is generally feasible (see Tables 2 and 3). For all features (i.e., visual and thermal) footwear-dependent, floor-dependent classifier outperforms the other three classifiers with an average AUC score of 96% and F_1-Score of 64%. The high performance of this classifier is expected because it is trained and evaluated on the same footwear

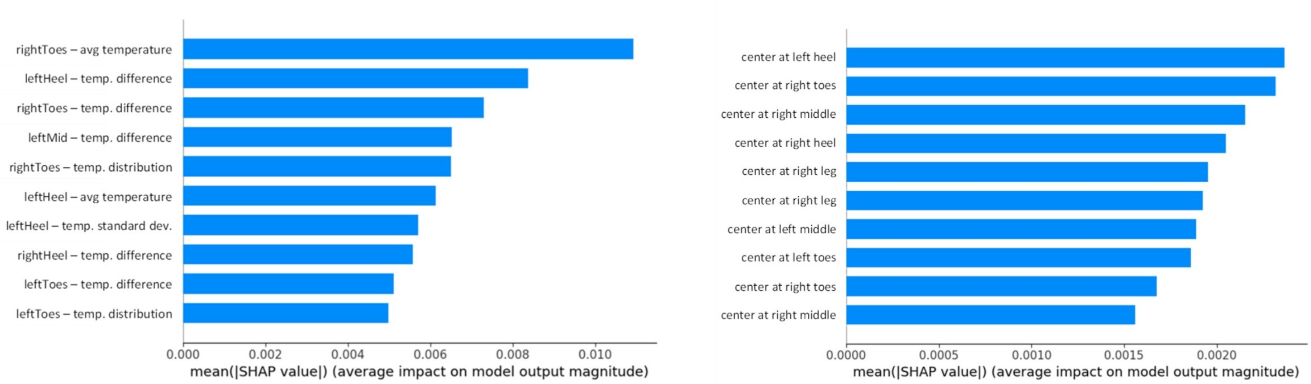

Figure 4: Left: Top ten important features for footwear-independent floor-independent model trained on thermal features. Right: Top ten important features for footwear-independent floor-independent model trained on visual features.

and floor condition. Hence, the classifier overfits a particular floor and footwear and gives high performance for that condition. In practice, the built classifier can be deployed in an identification system that works for a particular footwear and floor conditions.

Furthermore, we built *footwear-independent, floor-dependent* and *footwear-dependent, floor-independent* classifiers. For the entire feature set, both classifiers had similar performance, as the average AUC score for the classifiers was 95%. Lastly, to propose a generic and robust classifier (*footwear-independent, floor-independent* classifier) that could predict the identity of a user for any floor and footwear condition. Our results show that, although the built footwear-independent, floor-independent classifier, had a lower classification performance (AUC = 91.7%) compared to all other classifiers, it can still generalize reasonably well to any new floor and footwear and can predict new users’ identity.

6.2 Classification Performance per Feature Set

The results in Tables 2 and 3 suggest that for all four types of classifiers, a significant improvement in model performance is observed when trained on the combination of visual and thermal features compared to using the thermal and visual features independently. This increase in classification performance can be explained by the fact that both feature sets capture different characteristics of users’ feet: thermal features capture the heat traces of the foot regions and visual captures the structural property of the feet. Thus, when both feature sets are combined in a single model, the classifiers’ performance is enhanced substantially. For instance, we observe that for the footwear-independent floor-independent classifier, the prediction accuracy of the model increases the AUC score by 2% on average when trained on the combination of visual and thermal features compared to using just the thermal and feature set.

Furthermore, we observed that for all types of classifiers, the prediction accuracy of the models is much higher when trained on just the thermal features compared to using the visual feature alone (see Table 2 and Table 3). This observation suggests that if a single input modality is to be used for the biometric authentication tasks, thermal data input could more accurately predict the identity of the user than using the feet’s visual data.

6.3 Feature Importance

We used the SHapley Additive exPlanations (SHAP) algorithm [43] to investigate the importance of features on the performance of the footwear-independent floor-independent models trained on the visual-only and thermal-only feature sets. We chose this specific model for exploring feature importance, as it can be generalized to new floor and footwear conditions. The SHAP algorithm explains the output of a machine learning model by computing the contribution of each feature to its prediction. We use the SHapley values of the top 10 features obtained from this algorithm to plot feature importance on the model’s prediction. Features in the plots (Figure 4) are ordered by decreasing importance.

As observed in Figure 4–left, the classifier trained on thermal features relies heavily on the average temperature observed in various feet regions (e.g., heels and toes). Similarly, we observed that the feature capturing the difference between the maximum and minimum temperature occurring in various feet regions (e.g., heels and toes) also significantly contributed to correctly predicting the identity of participants. This result could be explained by the fact that the thermal temperature of participants’ feet can vary for different regions, leading to heterogeneous heat traces of the foot regions and, consequently, impacting the classifier’s decision to predict the identity of the participants. We also observed that vision features extracted from the heel, toes, and leg regions (e.g., *center at left heel* and *center at right middle*) impacts the model decision for biometric authentication (Figure 4–right). This observation suggests that the structural properties of participants’ toes, heels, and lower-leg region may be distinctive and can be used to predict user identity.

7 DISCUSSION

In this paper, we propose a novel approach, leveraging visual and thermal features for foot-based identification. Our approach analyses thermal images to extract distinct heat traces and structural properties of various regions of users’ feet. Based on these features, we built and evaluated classifiers that can accurately predict the identity of the users. Below, we discuss the most important observations from our exploration of the results.

7.1 Practical Relevance

According to Jain et al. [36], four factors qualify a human characteristic as a biometric trait. We reflect on how the use of the feet's thermal traces meets these criteria and briefly discuss how *HotFoot* can serve as a biometric identification system.

Regarding *universality*, our approach depends on analyzing humans' feet as a biological feature. This feature is available among the majority of the population. At the same time, it cannot be applied in cases where people rely on wheelchairs or lack limbs. The second factor is *distinctiveness*. We showed that thermal traces can be used to distinguish and, thus, identify people with reasonable accuracy. Furthermore, the proposed features can be measured quantitatively (i.e. are *collectable*). Unless an injury occurs, feet's thermal traces are invariant over time, meeting the *permanence* criterion. In summary, our approach is broadly applicable in cases where user identification is beneficial (for example, to personalize content). However, if the need for identification is a requirement (for example, constantly verifying users' identity in a secure environment), alternatives are needed in case users do not have the necessary features.

Jain et al. propose three metrics to assess the practicality of a biometric identification system [36]. The first criterion is *performance*, in terms of accuracy and speed. The results show that our solutions can achieve high accuracies, depending on the setting. Speed is comparable to other biometric identification systems, such as fingerprint or face identification. While building the predictive models requires time for training, comparing a user's traits against the predictive model is possible continuously and in real-time. Based on non-formal feedback from our participants, they consider the proposed identification system to be socially *acceptable*, and can imagine it to be incorporated in everyday situations. Finally, users cannot easily mimic the limb's heat traces of other users, thus making the approach robust with regard to *circumvention*. Future work should evaluate the resilience of the proposed approach against impersonation attacks in case it is sought to be used for security purposes. We can only speculate about how well the approach scales to a larger user number. A difference from other biometric approaches, though, is the influence of footwear. In populations with homogenous footwear, accuracy might be affected.

7.2 Classifiers Performance

Regarding *RQ1* on how accurately our approach performs when it comes to identifying people based on their thermal and visual features, our results show that vision- and thermal-based biometric identification classification is generally feasible, achieving an AUC score between 91.7% and 98.9%. The highest prediction accuracy can be achieved in situations with consistent floor types and people wearing the same footwear. This makes the approach suitable for work environments in which the objective is to make sure that only legitimate users are present. For cases in which people need to be identified while walking on different floor types and using different shoes, the accuracy is lower. In practice, this means that in about one out of 12 cases, the user is identified incorrectly. This might still be acceptable in cases where content is personalized to users.

7.3 Influence of Footwear and Floor Type

To answer *RQ2* and *RQ3*, we examined the effect of different footwear and floor types on identification accuracy. Floor type has a rather small effect on identification accuracy. We observed that *laminate* floor leads to a slightly lower accuracy compared to the AUC scores of *carpet* or *linoleum* floors. This could be explained by the observation that *laminate* surfaces are more strongly reflecting heat traces, affecting the accuracy for both thermal and visual features.

The effect is more pronounced for footwear. The *socks* condition performed worse than both the *personal* and *standard* shoes. An explanation for this could be that in case of wearing socks, the temperature of the foot is more strongly affected by the surrounding environment, leading to a more consistent temperature of feet across users, making individuals more difficult to discriminate.

8 CONCLUSION

In this work, we explored the use of thermal imaging to identify users based on the visual and thermal features of their feet. At the outset of our work, we identified different feature sets (visual, thermal, and the combination of thermal and visual features) and subsequently explored how accurately they allow the user to be identified while wearing different footwear as well as on different types of floors. We used the extracted features to train different classifiers, demonstrating that different use cases can be supported by our approach. Our classifiers achieved AUC scores up to 98.9%. We found that there is an increase in classification accuracy when using the combination of visual and thermal features as opposed to using visual or thermal features alone.

We see our work as proof of concept and hope other researchers use it as a point of departure as they create novel applications and user interfaces enabled by the ability to identify people seamlessly in smart environments. Our work enables a broad range of subsequent research, including but not limited to increasing the accuracy under different conditions, investigating threat models in security use cases, and understanding the challenges of real-world deployments.

ACKNOWLEDGMENTS

This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation – 425869382) and is part of Priority Program SPP2199 Scalable Interaction Paradigms for Pervasive Computing Environments. Furthermore, this research is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr in the context of the project Voice of Wisdom. dtec.bw is funded by the European Union – NextGenerationEU.

REFERENCES

[1] Andrea F Abate, Michele Nappi, Daniel Riccio, and Gabriele Sabatino. 2007. 2D and 3D face recognition: A survey. *Pattern recognition letters* 28, 14 (2007), 1885–1906.

[2] Ayman Abaza, Arun Ross, Christina Hebert, Mary Ann F Harrison, and Mark S Nixon. 2013. A survey on ear biometrics. *ACM computing surveys (CSUR)* 45, 2 (2013), 1–35.

[3] Yomna Abdelrahman, Mohamed Khamis, Stefan Schneegass, and Florian Alt. 2017. Stay cool! understanding thermal attacks on mobile-based user authentication. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3751–3763.

[4] Yomna Abdelrahman, Anam Ahmad Khan, Joshua Newn, Eduardo Velloso, Sherine Ashraf Safwat, James Bailey, Andreas Bulling, Frank Vetere, and Albrecht

Schmidt. 2019. Classifying attention types with thermal imaging and eye tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–27.

[5] Yomna Abdelrahman, Alireza Sahami Shirazi, Niels Henze, and Albrecht Schmidt. 2015. Investigation of Material Properties for Thermal Imaging-Based Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 15–18. https://doi.org/10.1145/2702123.2702290

[6] Yomna Abdelrahman and Albrecht Schmidt. 2018. Beyond the visible: Sensing with thermal imaging. *Interactions* 26, 1 (2018), 76–78.

[7] Yomna Abdelrahman, Eduardo Velloso, Tilman Dingler, Albrecht Schmidt, and Frank Vetere. 2017. Cognitive Heat: Exploring the Usage of Thermal Imaging to Unobtrusively Estimate Cognitive Load. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3 (2017), 33:1–33:20. https://doi.org/10.1145/3130898

[8] Yomna Abdelrahman, Paweł W Woźniak, Pascal Knierim, Dominik Weber, Ken Pfeuffer, Niels Henze, Albrecht Schmidt, and Florian Alt. 2019. Exploring the domestication of thermal imaging. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*. 1–7.

[9] D Adelman. 1988. Some control strategies for radiant floor heating. *Radiant times* (1988), 4–5.

[10] Khamael Abbas Al-Dulaimi. 2013. Using feature extraction for human footprints recognition. *International Journal of Computer Applications* 64, 3 (2013).

[11] Mouad MH Ali, Vivek H Mahale, Pravin Yannawar, and AT Gaikwad. 2016. Overview of fingerprint recognition system. In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*. IEEE, 1334–1338.

[12] Florian Alt and Stefan Schneegass. 2022. Beyond Passwords—Challenges and Opportunities of Future Authentication. *IEEE Security & Privacy* 20, 1 (2022), 82–86. https://doi.org/10.1109/MSEC.2021.3127459

[13] VD Ambeth Kumar, S Malathi, R Venkatesan, K Ramalakshmi, K Vengatesan, Weiping Ding, and Abhishek Kumar. 2020. Exploration of an innovative geometric parameter based on performance enhancement for foot print recognition. *Journal of Intelligent & Fuzzy Systems* 38, 2 (2020), 2181–2196.

[14] Virginia Ortiz Andersson and Ricardo Matsumura Araujo. 2015. Person identification using anthropometric and gait data from kinect sensor. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

[15] Gilles Bailly, Jörg Müller, Michael Rohs, Daniel Wigdor, and Sven Kratz. 2012. ShoeSense: A New Perspective on Gestural Interaction and Wearable Applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1239–1248. https://doi.org/10.1145/2207676.2208576

[16] Salil P Banerjee and Damon L Woodard. 2012. Biometric authentication and identification using keystroke dynamics: A survey. *Journal of Pattern Recognition Research* 7, 1 (2012), 116–139.

[17] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. 2006. Surf: Speeded up robust features. In *European conference on computer vision*. Springer, 404–417.

[18] Kai Cao and Anil K Jain. 2018. Automated latent fingerprint recognition. *IEEE transactions on pattern analysis and machine intelligence* 41, 4 (2018), 788–800.

[19] Kun Woo Cho, Feng Lin, Chen Song, Xiaowei Xu, Fuxing Gu, and Wenyao Xu. 2016. Thermal handprint analysis for forensic identification using Heat-Earth Mover's Distance. In *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*. IEEE, 1–8.

[20] Justin HJ Chua, Adrian G Dyer, and Jair E Garcia. 2018. Hot Shoes in the Room: Authentication of Thermal Imaging for Quantitative Forensic Analysis. *Journal of Imaging* 4, 1 (2018), 21.

[21] MG Cooper, BB Mikic, and MM Yovanovich. 1969. Thermal contact conductance. *International Journal of heat and mass transfer* 12, 3 (1969), 279–300.

[22] Shaveta Dargan and Munish Kumar. 2020. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Systems with Applications* 143 (2020), 113114.

[23] Liu Dian and Sun Dongmei. 2016. Contactless palmprint recognition based on convolutional neural network. In *2016 IEEE 13th International Conference on Signal Processing (ICSP)*. IEEE, 1363–1367.

[24] Michal Dvořák, Martin Drahanskỳ, and Waleed H Abdulla. 2021. On the fly biometric identification system using hand-geometry. *IET Biometrics* 10, 3 (2021), 315–325.

[25] Sarah Faltaous, Jonathan Liebers, Yomna Abdelrahman, Florian Alt, and Stefan Schneegass. 2019. VPID: Towards Vein Pattern Identification Using Thermal Imaging. *i-com* 18, 3 (2019), 259–270.

[26] Sarah Faltaous, Mark Wittpoth, Yomna Abdelrahman, and Stefan Schneegass. 2022. HeatGoggles: Enabling Ubiquitous Touch Input through Head-Mounted Devices using Thermal Imaging. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*. 254–256.

[27] Lunke Fei, Guangming Lu, Wei Jia, Shaohua Teng, and David Zhang. 2018. Feature extraction methods for palmprint recognition: A survey and evaluation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 49, 2 (2018), 346–363.

[28] Qiong Gu, Li Zhu, and Zhihua Cai. 2009. Evaluation Measures of the Classification Performance of Imbalanced Data Sets. In *Computational Intelligence and Intelligent Systems*, Zhihua Cai, Zhenhua Li, Zhuo Kang, and Yong Liu (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 461–471. https://doi.org/10.1007/978-3-642-04962-0_53

[29] Anshu Gupta and Deepa Raj. 2019. Multimodal Biometrics for Personal Identification using Human Foot Prints: A Review. In *Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE)*.

[30] Mridula Gupta, Michael Bisesi, and Jiyoung Lee. 2017. Comparison of survivability of Staphylococcus aureus and spores of Aspergillus niger on commonly used floor materials. *American Journal of Infection Control* 45, 7 (2017), 717–722.

[31] Lin Hong and Anil Jain. 1998. Integrating faces and fingerprints for personal identification. *IEEE transactions on pattern analysis and machine intelligence* 20, 12 (1998), 1295–1307.

[32] Stephanos Ioannou, Sjoerd Ebisch, Tiziana Aureli, Daniela Bafunno, Helene Alexi Ioannides, Daniela Cardone, Barbara Manini, Gian Luca Romani, Vittorio Gallese, and Arcangelo Merla. 2013. The autonomic signature of guilt in children: a thermal infrared imaging study. *PloS one* 8, 11 (2013), e79440.

[33] Stephanos Ioannou, Vittorio Gallese, and Arcangelo Merla. 2014. Thermal infrared imaging in psychophysiology: potentialities and limits. *Psychophysiology* 51, 10 (2014), 951–963.

[34] Rabia Jafri and Hamid R Arabnia. 2009. A survey of face recognition techniques. *journal of information processing systems* 5, 2 (2009), 41–68.

[35] Anil Jain, Ruud Bolle, and Sharath Pankanti. 1996. *Introduction to Biometrics*. Springer US, Boston, MA, 1–41. https://doi.org/10.1007/0-306-47044-6_1

[36] Anil K Jain, Arun Ross, and Salil Prabhakar. 2004. An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology* 14, 1 (2004), 4–20.

[37] Rohit Khokher, Ram Chandra Singh, and Rahul Kumar. 2015. Footprint recognition with principal component analysis and independent component analysis. In *Macromolecular symposia*, Vol. 347. Wiley Online Library, 16–26.

[38] Song-Kyoo Kim, Chan Yeob Yeun, Ernesto Damiani, and Nai-Wei Lo. 2019. A machine learning framework for biometric authentication using electrocardiogram. *IEEE Access* 7 (2019), 94858–94868.

[39] Yassin Kortli, Maher Jridi, Ayman Al Falou, and Mohamed Atri. 2020. Face recognition systems: A survey. *Sensors* 20, 2 (2020), 342.

[40] Eric Larson, Gabe Cohn, Sidhant Gupta, Xiaofeng Ren, Beverly Harrison, Dieter Fox, and Shwetak Patel. 2011. HeatWave: Thermal Imaging for Surface User Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. ACM, New York, NY, USA, 2565–2574. https://doi.org/10.1145/1978942.1979317

[41] Eryun Liu. 2017. Infant footprint recognition. In *Proceedings of the IEEE International Conference on Computer Vision*. 1653–1660.

[42] G Lowe. 2004. Sift-the scale invariant feature transform. *Int. J* 2, 91-110 (2004), 2.

[43] Scott M Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. In *Advances in neural information processing systems*. 4765–4774.

[44] Audrey Macdonald, Nina Petrova, Suhail Ainarkar, John Allen, Peter Plassmann, Aaron Whittam, John Bevans, Francis Ring, Ben Kluwe, Rob Simpson, et al. 2016. Thermal symmetry of healthy feet: a precursor to a thermal study of diabetic feet prior to skin breakdown. *Physiological measurement* 38, 1 (2016), 33.

[45] Jyoti Malik, Dhiraj Girdhar, Ratna Dahiya, and G Sainarayanan. 2014. Reference threshold calculation for biometric authentication. *IJ Image, Graphics and Signal Processing* 2 (2014), 46–53.

[46] Wojciech Michal Matkowski, Tingting Chai, and Adams Wai Kin Kong. 2019. Palmprint recognition in uncontrolled and uncooperative environment. *IEEE Transactions on Information Forensics and Security* 15 (2019), 1601–1615.

[47] Yuxin Meng, Duncan S Wong, Roman Schlegel, et al. 2012. Touch gestures based biometric authentication scheme for touchscreen mobile phones. In *International conference on information security and cryptology*. Springer, 331–350.

[48] A Muthukumar and A Kavipriya. 2019. A biometric system based on Gabor feature extraction with SVM classifier for Finger-Knuckle-Print. *Pattern Recognition Letters* 125 (2019), 150–156.

[49] Kapil Kumar Nagwanshi and Sipi Dubey. 2012. Biometric authentication using human footprint. *International Journal of Applied Information Systems* 3, 7 (2012), 1–6.

[50] Kazuki Nakajima, Yoshiki Mizukami, Kanya Tanaka, and Toshiyo Tamura. 2000. Footprint-based personal recognition. *IEEE Transactions on Biomedical Engineering* 47, 11 (2000), 1534–1537.

[51] SM Ngwira. 2015. Survey of biometric authentication for e-assessment. In *2015 International Conference on Computing, Communication and Security (ICCCS)*. IEEE, 1–4.

[52] KC Parsons. 1992. Contact between human skin & hot surfaces equivalent contact temperature. In *Proc. ICEE*.

[53] Ge Peng, Gang Zhou, David T Nguyen, Xin Qi, Qing Yang, and Shuangquan Wang. 2016. Continuous authentication with touch behavioral biometrics and voice on wearable glasses. *IEEE transactions on human-machine systems* 47, 3 (2016), 404–416.

[54] NL Petrova, A Whittam, A MacDonald, S Ainarkar, AN Donaldson, J Bevans, J Allen, P Plassmann, B Kluwe, F Ring, et al. 2018. Reliability of a novel thermal imaging system for temperature assessment of healthy feet. *Journal of foot and*

ankle research 11, 1 (2018), 1–6.
[55] Sarah Prange, Ahmed Shams, Robin Piening, Yomna Abdelrahman, and Florian Alt. 2021. Priview–exploring visualisations to support users' privacy awareness. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.
[56] Shantanu Rane. 2014. Standardization of biometric template protection. *IEEE MultiMedia* 21, 4 (2014), 94–99.
[57] Stephan Richter, Christian Holz, and Patrick Baudisch. 2012. Bootstrapper: recognizing tabletop users by their shoes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1249–1252.
[58] Imad Rida, Noor Almaadeed, and Somaya Almaadeed. 2019. Robust gait recognition: a comprehensive survey. *IET Biometrics* 8, 1 (2019), 14–28.
[59] Brian D Ripley. 2007. *Pattern recognition and neural networks*. Cambridge university press.
[60] Jody L Riskowski, Alyssa B Dufour, Thomas J Hagedorn, Howard J Hillstrom, Virginia A Casey, and Marian T Hannan. 2013. Associations of foot posture and function to lower extremity pain: results from a population-based foot study. *Arthritis care & research* 65, 11 (2013), 1804–1812.
[61] Arun A Ross and Rohin Govindarajan. 2005. Feature level fusion of hand and face biometrics. In *Biometric technology for human identification II*, Vol. 5779. International Society for Optics and Photonics, 196–204.
[62] Alireza Sahami Shirazi, Yomna Abdelrahman, Niels Henze, Stefan Schneegass, Mohammadreza Khalilbeigi, and Albrecht Schmidt. 2014. Exploiting Thermal Reflection for Interactive Systems. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. ACM, New York, NY, USA, 3483–3492. https://doi.org/10.1145/2556288.2557208
[63] Baljit Singh Saini, Navdeep Kaur, and Kamaljit Singh Bhatia. 2016. Keystroke dynamics for mobile phones: A survey. *Indian Journal of Science and Technology* 9, 6 (2016), 1–8.
[64] Bassam Sayed, Issa Traoré, Isaac Woungang, and Mohammad S Obaidat. 2013. Biometric authentication using mouse gesture dynamics. *IEEE systems journal* 7, 2 (2013), 262–274.
[65] Christoph Schröder, Sahar Mahdie Klim Al Zaidawi, Martin H.U. Prinzler, Sebastian Maneth, and Gabriel Zachmann. 2020. Robustness of Eye Movement Biometrics Against Varying Stimuli and Varying Trajectory Length. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3313831.3376534
[66] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 815–823.
[67] Fabius Steinberger, Marcus Foth, and Florian Alt. 2014. Vote With Your Feet: Local Community Polling on Urban Screens. In *Proceedings of The International Symposium on Pervasive Displays* (Copenhagen, Denmark) *(PerDis '14)*. Association for Computing Machinery, New York, NY, USA, 44–49. https://doi.org/10.1145/2611009.2611015
[68] Hong-Ren Su, Kuang-Yu Chen, Wei Jing Wong, and Shang-Hong Lai. 2017. A deep learning approach towards pore extraction for high-resolution fingerprint recognition. In *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2057–2061.
[69] Lichao Sun, Yuqi Wang, Bokai Cao, S Yu Philip, Witawas Srisa-An, and Alex D Leow. 2017. Sequential keystroke behavioral biometrics for mobile user identification via multi-view deep learning. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 228–240.
[70] Kalaivani Sundararajan and Damon L. Woodard. 2018. Deep Learning for Biometrics: A Survey. *ACM Comput. Surv.* 51, 3, Article 65 (May 2018), 34 pages. https://doi.org/10.1145/3190618
[71] Andreas Uhl and Peter Wild. 2008. Footprint-based biometric verification. *Journal of Electronic Imaging* 17, 1 (2008), 011016.
[72] Eduardo Velloso, Dominik Schmidt, Jason Alexander, Hans Gellersen, and Andreas Bulling. 2015. The feet in human–computer interaction: A survey of foot-based interaction. *ACM Computing Surveys (CSUR)* 48, 2 (2015), 1–35.
[73] Ruben Vera-Rodriguez, Julian Fierrez, John SD Mason, and Javier Ortega-Garcia. 2013. A novel approach of gait recognition through fusion with footstep information. In *2013 International Conference on Biometrics (ICB)*. IEEE, 1–6.
[74] Changsheng Wan, Li Wang, and Vir V Phoha. 2018. A survey on gait recognition. *ACM Computing Surveys (CSUR)* 51, 5 (2018), 1–35.
[75] Mark Weiser. 1993. Hot topics-ubiquitous computing. *Computer* 26, 10 (1993), 71–72.
[76] Jacob O. Wobbrock and Julie A. Kientz. 2016. Research Contributions in Human-Computer Interaction. *Interactions* 23, 3 (apr 2016), 38–44. https://doi.org/10.1145/2907069
[77] Ziyi Xu, Quchao Wang, Duo Li, Menghan Hu, Nan Yao, and Guangtao Zhai. 2020. Estimating departure time using thermal camera and heat traces tracking technique. *Sensors* 20, 3 (2020), 782.
[78] Takeshi Yamakawa, Kazuhiko Taniguchi, Kazunari Asari, Syoji Kobashi, and Yutaka Hata. 2008. Biometric personal identification based on gait pattern using both feet pressure change. In *2008 World Automation Congress*. IEEE, 1–6.
[79] Chao Yan, Bailing Zhang, and Frans Coenen. 2015. Multi-attributes gait identification by convolutional neural networks. In *2015 8th International Congress on Image and Signal Processing (CISP)*. IEEE, 642–647.
[80] Soweon Yoon and Anil K Jain. 2015. Longitudinal study of fingerprint recognition. *Proceedings of the National Academy of Sciences* 112, 28 (2015), 8555–8560.
[81] Nan Zheng, Aaron Paloski, and Haining Wang. 2011. An efficient user verification system via mouse movements. In *Proceedings of the 18th ACM conference on Computer and communications security*. 139–150.
[82] Dexing Zhong, Xuefeng Du, and Kuncai Zhong. 2019. Decade progress of palmprint recognition: A brief survey. *Neurocomputing* 328 (2019), 16–28.
[83] Yu Zhong and Yunbin Deng. 2015. A survey on keystroke dynamics biometrics: approaches, advances, and evaluations. In *Recent Advances in User Authentication Using Keystroke Dynamics Biometrics*. Number 1. Science Gate Publishing, 1–22.

Table 4: Precision and Accuracy scores of Random forest model for footwear-dependent floor-dependent conditions.

	Visual		Thermal		Visual & Thermal	
Condition	Precision	Accuracy	Precision	Accuracy	Precision	Accuracy
Linoleum-Personal Shoes	24.7 ± 6.1%	25.4 ± 4.7%	57.8 ± 7.8%	69.8 ± 5.8%	85.1 ± 5.5%	90.9 ± 0.3%
Linoleum-Dedicated Shoes	23.9 ± 8.4%	24.2 ± 4.5%	60.3 ± 1.4%	69.8 ± 1.1%	68.5 ± 8.1%	75.5 ± 0.6%
Linoleum-Socks Only	42.3 ± 2.1%	55.5 ± 2.2%	42.3 ± 0.2%	55.5 ± 1.3%	63.4 ± 2.1%	68.2 ± 1.4%
Laminate-Personal Shoes	29. ± 4.4%	27.3 ± 2.5%	46.6 ± 0.8%	55.5 ± 0.9%	55.2 ± 0.1%	62.3 ± 1.0%
Laminate-Dedicated Shoes	21.0 ± 4.0%	24.2 ± 2.0%	43.1 ± 1.6%	53.9 ± 1.5%	51.2 ± 0.1%	58.3 ± 0.8%
Laminate-Socks Only	22.3 ± 2.1%	20.6 ± 0.2%	41.8 ± 1.2%	52.3 ± 0.1%	49.6 ± 0.7%	55.9 ± 0.7%
Carpet-Personal Shoes	23.0 ± 1.6%	25.7 ± 5.1%	74.8 ± 1.7%	97.9 ± 0.4%	74.5 ± 1.6%	78.1 ± 1.1%
Carpet-Dedicated Shoes	25.0 ± 7.0%	27.3 ± 7.1%	57.2 ± 1.8%	68.2 ± 0.1%	72.2 ± 0.6%	74.2 ± 7.1%
Carpet-Socks Only	25.0 ± 2.0%	73.1 ± 1.1%	37.3 ± 4.2%	47.6 ± 6.7%	57.6 ± 0.4%	60.3 ± 4.8%

Table 5: Precision and Accuracy scores of Random forest model for footwear-dependent floor-independent classifier, footwear-independent floor-dependent classifier and footwear-in dependent floor-independent classifier. Similar to Table 3, the reported results are averaged by the dependent condition but split by the independent condition on which it was evaluated.

		Visual		Thermal		Visual & Thermal	
Classifier	Testing Condition (Training Conditions)	Precision	Accuracy	Precision	Accuracy	Precision	Accuracy
Footwear Dependent-Floor Independent	Laminate(Linoleum & Carpet)	21.8 ± 1.6%	23.1 ± 2.3%	55.7 ± 4.1%	53.9 ± 5.9%	59.6 ± 7.1%	58.4 ± 0.7%
	Linoleum (Laminate & Carpet)	24.3 ± 3.5%	27.2 ± 2.9%	68.1 ± 4.6%	66.6 ± 4.6%	73.9 ± 5.0%	72.1 ± 0.4%
	Carpet (Laminate & Linoleum)	18.7 ± 1.0%	20.6 ± 1.1%	61.5 ± 4.9%	59.7 ± 4.5%	61.1 ± 4.0%	60.0 ± 4.0%
Footwear Independent-Floor Dependent	Personal Shoes (Same Shoes & Socks)	25.5 ± 5.5%	25.6 ± 5.1%	58.3 ± 1.5%	67.1 ± 1.5%	71.3 ± 1.2%	69.9 ± 1.1%
	Same Shoes (Personal Shoes & Socks)	25.5 ± 0.7%	23.3 ± 2.1%	56.9 ± 1.4%	53.3 ± 1.5%	63.3 ± 1.6%	69.9 ± 7.2%
	Socks Only (Personal Shoes & Same Shoes)	25.5 ± 3.1%	37.4 ± 2.4%	40.8 ± 4.1%	51.1 ± 2.1%	56.3 ± 0.4%	63.3 ± 1.1%
Footwear Independent-Floor Independent		17.6 ± 3.6	17.4 ± 2.3%	53.3 ± 0.7%	48.6 ± 7.1%	56.7 ± 2.5%	51.9 ± 2.1%