Utilizing Image Processing to Measure Foot Parameters for Insole Production

Pilailuck Panphattarasap[1] (ID) , Pisit Praiwattana[2] (ID) , Jidapa Kraisangka[3] (ID) , Wanida Janvikul[4] (ID) , Boonlom Thavornyutikarn[5] (ID) , Wasana Kosorn[6] (ID) , Nutdanai Nampichai[7] (ID) , Wudhichart Sawangphol[8] (ID) ,

[1]Faculty of Information and Communication Technology, Mahidol University, pilailuck.pan@mahidol.ac.th
[2]Faculty of Information and Communication Technology, Mahidol University, pisit.pra@mahidol.ac.th
[3]Faculty of Information and Communication Technology, Mahidol University, jidapa.kra@mahidol.ac.th
[4]National Metal and Materials Technology Center, National Science and Technology Development Agency, wanidaj@mtec.or.th
[5]National Metal and Materials Technology Center, National Science and Technology Development Agency, boonlomt@mtec.or.th
[6]National Metal and Materials Technology Center, National Science and Technology Development Agency, wasanak@mtec.or.th
[7]National Metal and Materials Technology Center, National Science and Technology Development Agency, nutdanai.nam@mtec.or.th
[8]Faculty of Information and Communication Technology, Mahidol University, wudhichart.saw@mahidol.ac.th

Corresponding author: Wudhichart Sawangphol, wudhichart.saw@mahidol.ac.th

Abstract. To mitigate reliance on the unwieldy and costly device for the insole production, we proposed the image-based algorithm for measuring specific foot parameters to create an insole profile using in the insole production. The images was taken using a camera equipped on the mobile phone. The general idea was to segment the foot sole from the retrieved image and define the salient points of the foot features, and use that to calculate the significant parameters of the insole profile. We ran the experiment over 200 participants with more than 860 of foot images and the best coverage range is within 0.0001 to 0.01 centimeter. The findings of this study demonstrate the viability and effectiveness of employing a single image for the insole production.

1 INTRODUCTION

The foot plays an important role in human mobility, weight support, and even shock absorption. We can also use the appearance of the foot to detect health issues such as arthritis, diabetes, etc. As each person has a different foot structure, the foot can be categorized based on the arch types as flat, normal, and high. The abnormality of a too-flat or too-high arch leads to several common problems including pain, instability, deformities, and issues with balance and gait.

Several studies show that numbers of people across all age groups suffer from foot arch problems that disrupt their everyday life activities [7], [11]. Therefore, to overcome the abnormality and minimize its impact on an individual's health and well-being, one of the solutions is to use the insole, especially by integrating a customizable attribute. The utilization of well-fitted footwear has the potential to contribute to pain alleviation and enhance the alignment of the feet.

As precise measurements of the insole profile are necessary to ensure the proper fit and comfort of the insole, accuracy is indeed crucial for insole personalizing. In recent developments, there has been a growing integration of 3D models in the process of foot reconstruction for insole customization. These 3D models provide a detailed representation of the foot, allowing for precise measurements and customization of insoles. However, the cost of 3D scanners used for creating these models can be relatively high, making them less accessible for widespread use. Additionally, the process often requires the presence of both users and experts to ensure accurate scanning and measurements. Therefore, we looked for solutions that could relieve the cost and difficulty levels of the foot-measuring process.

We, in collaboration with experts in Orthosis and Prosthesis from Thailand National Metal and Materials Technology Center (MTEC), then proposed the use of images due to their relative ease of acquisition. Apart from the mobility issue, the other reasons are based on the possibility of cost reduction and the easiness of acquiring the equipment.

The rest of the paper is organized as follows: Section 2 presents a background study including some background knowledge and an overview of related works. Section 3 describes our proposed approach and methodology. Section 4 presents results and discussion. Finally, Section 5 concludes remarks and together with the possibilities of future work.

2 BACKGROUND STUDY

In this section, we present the important foot parameters used for insole customization, the traditional techniques to measure those parameters along with the recent methods, and the study related to measuring using images.

2.1 The Foot Parameters and Insole

In order to customize the insole for the individual one, the required parameters are as shown in Figure 1. The explanation for each parameter is shown in Table 1. The parameters employed in this study were established by the experts in foot orthosis design, which closely resemble the setup utilized in studies [9] and [17].

The parameters A to I can be obtained by measuring only the foot sole, but for the arch height (parameter **(H)**), we cannot measure the value directly. To do that, the widely employed method is to use the Arch index proposed by [3] to characterize the arch of the foot as shown in [15]; this value is determined by subdividing the foot sole to be the forefoot, midfoot, and hindfoot, denoted as A, B, and C, respectively. The Arch Index (AI) is calculated by finding the ratio of area B to the sum of areas A, B, and C denoted as $AI = B/(A+B+C)$. This index serves as an indicator of the relative proportions and distribution of the arch in relation to the overall foot structure. Based on this criterion in [3], the arch type is determined as follows: arch indices greater than or equal to 0.26 are classified as low, arch indices falling between 0.21 and 0.26 are classified as normal, and
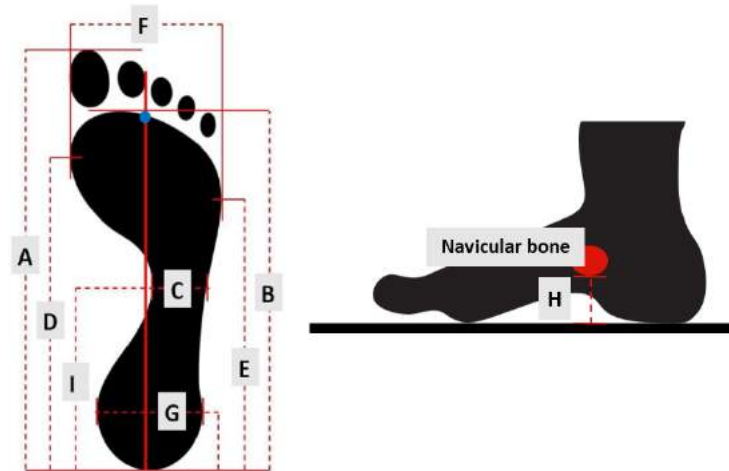
Figure 1: The 8 foot parameters from **(A)** to **(H)** commonly used for the insole customization obtained from the experts.

arch indices less than or equal to 0.21 are classified as high. However, in the earlier stage of our research, we would temporarily omit this value (parameter **(H)**) and focus on the foot sole parameters.

2.2   Traditional Techniques for Foot Measuring

To measure the foot parameters stated in Section 2.1, the basic tools utilized for measuring foot insole size include a foot ruler or tape measure, as well as the Harris mat, which was applied in the clinical use [6]. While these methods are generally considered simple and cost-effective, the measurements obtained may lack precision and be susceptible to errors. Some employ a digital caliper to enhance accuracy. Nonetheless, utilizing this instrument can be time-consuming and necessitates a certain level of skill to ensure effective operation.

Apart from this, there is also the usage of visual tools such as a carbon paper imprint, plaster casting, or foot wet test. Nonetheless, the utilization of these techniques necessitates the preparation of specialized materials and the involvement of a proficient expert to read the measured values, and they exhibit an unsatisfactory level of precision.

2.3   Recent Techniques for Foot Measuring

Instead of manual measurements conducted by human operators, the introduction of foot scanners has emerged as an alternative approach. Foot scanners utilize a combination of cameras and sensors to capture precise measurements of the foot. However, the utilization of foot scanners comes with certain drawbacks, including the associated cost of the device and the requirement for specialized expertise in operating and interpreting the scanner data.

Recently, to increase precision in measurement methodologies, the utilization of the 3D scanner and Computer Aided Design (CAD) are employed. The retrieved 3D foot model will be constructed using either the 3D printer or the Computer Numerical Control (CNC). Based on the work of [8] and [24] that conducted the comparison experiment between the scanner and traditional methods, the result shows that the method using the 3D foot scanning offers a high level of accuracy than other techniques. However, the drawbacks of these methods are associated with their cost, limited accessibility, and expert dependency.

Table 1: The foot parameters notation used

Notation	Description
A	Foot length measured from the heel to the forefoot
B	Foot length measured from the heel to the forefoot excluding the toes
C	Middle foot width
D	The length measured from the heel to the apex of 1st metatarsal head
E	The length measured from the heel to the apex of 5th metatarsal head
F	The width measured from the apex of 1st metatarsal head to the apex of 5th metatarsal head
G	Heel width
H	Arch height
I	The length measured from the heel to the middle foot

2.4 Related Study for Foot Measuring Using Images

To optimize cost-effectiveness and simplify the measurement process, several works have been directed toward image-based approaches. Within the domain of image analysis, various studies have explored the utilization of planar images obtained from foot scanning devices, pressure sensors, or temperature sensors [4]. However, it is important to note that these investigations have primarily focused on applications beyond the direct measurement of foot parameters, such as the detection of diabetic ulcers [2, 16, 12, 14, 22].

To measure foot size, some commercial mobile applications have been developed utilizing some sensors liked the depth sensor or infrared in conjunction with mobile cameras such as Feetmeter[1], FeetSizr[2], and ShoeFitter[3]. These applications are primarily focused on customization for shoes than insole. A commonly employed technique involves capturing a photograph of side of feet with an A4 paper serving as a reference point. This can be achieved by either using the paper as the background or placing it side-by-side with the foot in the image frame. Subsequently, this image serves as the basis for further analysis and processing. In some applications, augmented reality (AR) technology is integrated, leveraging the depth sensor capabilities to enhance the accuracy of distance measurements.

Another aspect to replace the 3D scanner is the use of mobile cameras' photogrammetry techniques, which is to capture multi-view foot photos or videos to generate 3D key points for foot reconstruction [10, 21].

The utilization of scanners equipped with depth or pressure information demonstrates notable accuracy. Nevertheless, these techniques necessitate the use of costly devices or additional sensors. Despite that, we grow interested in simpler alternatives, specifically relying only on foot sole images captured by a mobile camera without the need for additional sensors. To achieve our objective, the initial step involves identifying the foot sole area through the image segmentation techniques as in [5], [12], [14], and [23]. Recent advancements in machine learning and deep learning have shown promising results in effectively classifying foot type [20] and segmenting the foot sole area from images [1, 2]. Once the segmentation process is completed, subsequent steps involve the extraction of relevant foot parameters using image processing algorithms. These parameters are obtained in pixel units, and a conversion process is implemented to translate the measurements back to the standard unit of centimeters.

---

[1] https://www.feetmeter.pro/
[2] https://feetsizr.com/
[3] https://shoefitter.io/

Table 2: The number of participants in each age group.

Age group	Number of participants
20-29	39
30-39	72
40-49	86
50-59	20
more than 59	3

The subsequent sections of this paper will delve further into the details of the proposed approach, including the methodology, experimental setup, and analysis of the obtained results.

3 THE PROPOSED FOOT PARAMETER EXTRACTION

In this section, we present the overall design of our method to segment the foot sole and extract the common foot parameters needed for insole customization, data preparation, and implementation.

3.1 System Design

Firstly, we accept the foot images as input. Then we performed the pre-processing and labeled them in 2 categories: 1. whole foot sole and 2. partial sole area (toe-omitted sole). We then train a deep learning model to segment the labeled areas.

Once finished training, to extract the foot parameters, we segmented the foot image using the deep learning model combine with the color segmentation to get the whole foot sole and partial sole area (toe-omitted sole) from an input image. The segmented products were converted to the binary mask.

In the final stage, we employed image processing techniques to extract the feature points necessary for calculating the essential foot parameters required for insole customization. The general ideas of the foot parameters measuring process is as summarised in Figure 2.

3.2 Participant Statistics

We gathered the data from 220 participants containing the variation of all three foot arch types are 150 normal foot arch images (Figure 3a), 139 flat foot arch images (Figure 3b), and 151 high foot arch images (Figure 3c).

The participants in this study were aged between 22 and 61 years. To provide a comprehensive overview of the age distribution, the participants were separated into five age groups: 20-29, 30-39, 40-49, 50-59, and more than 59 years. The number of participants in each age group is shown in Table 2. The average age of the participants is 38.7. This variation in foot arch type and age was deliberately incorporated to ensure the diversity of our dataset.

3.3 Data Collection

The processes of data gathering were to take photos of the participants' foot soles; the left foot and right foot were taken separately. In each instance, the subjects assumed a sit-to-stand posture with one leg raised, a configuration deemed optimal for obtaining sole area images using solely a mobile phone camera. Figure 4 shows the environment set up for the photo shooting. Note that the current environment had been controlled by having the dark paper set up as the background. While we acknowledge the need for further refinement
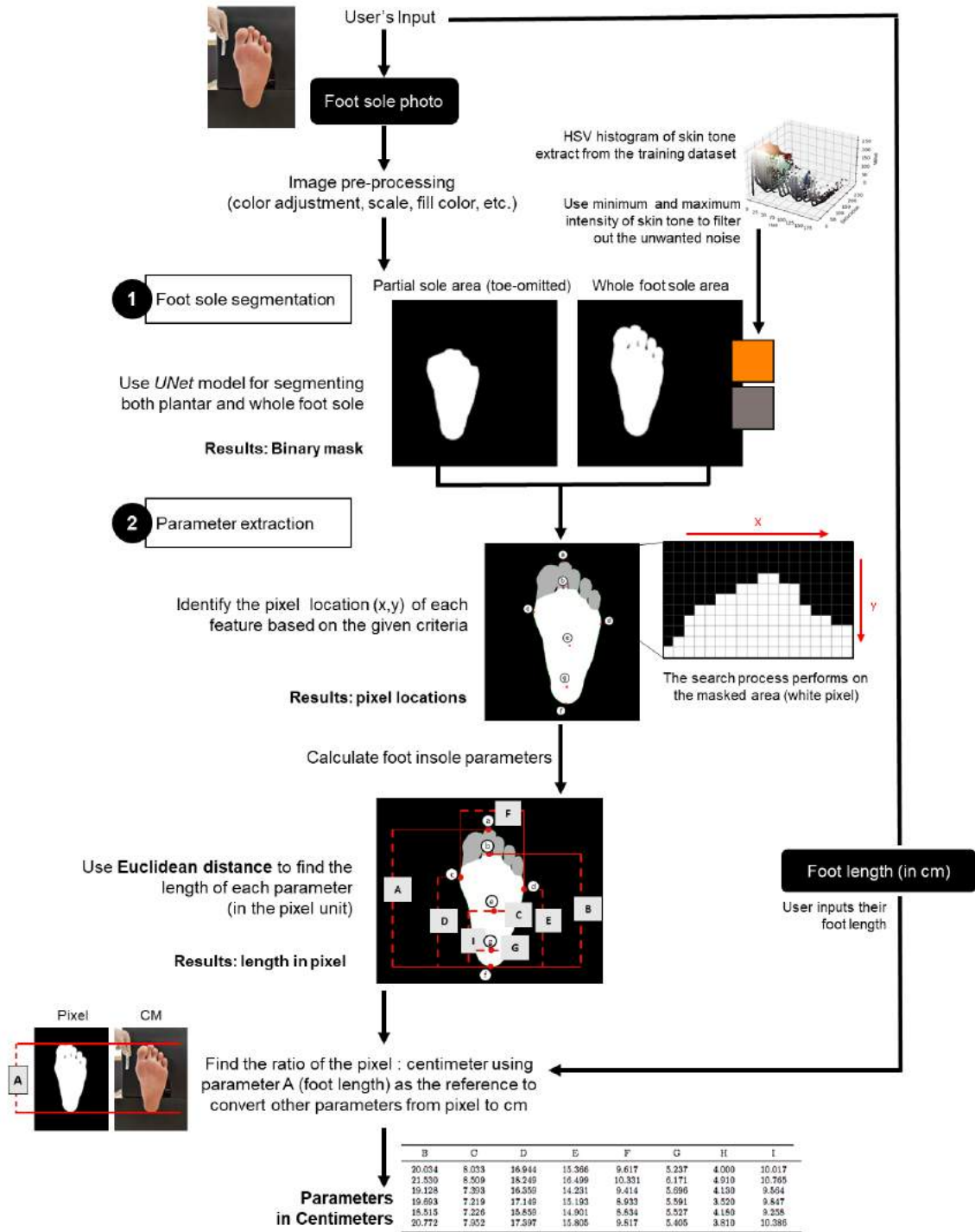
Figure 2: The processes of foot parameter extraction using our method.

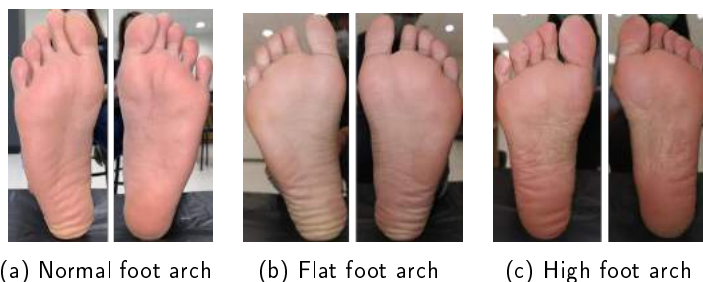(a) Normal foot arch   (b) Flat foot arch   (c) High foot arch

Figure 3: Examples of the left and right feet photos those taken from the participants having (a) Normal foot arch, (b) Flat foot arch and (c) High foot arch for the variation of our dataset.

Figure 4: The set-up of the environment we took the foot sole. The participant was in the sit-to-stand pose with one leg raised. The process involved one laboratory staff member responsible for holding a participant identification tag during image acquisition. This facilitated the subsequent pairing of captured images with corresponding ground truth data.

and validation, this initial investigation serves as a foundational step toward exploring the more practical implications of our methodology. More about this limitation will be discussed further in the conclusion section. We finally obtained 880 images of foot soles from 220 participants.

To obtain the precise ground truth, we had been assisted by the experts in Orthosis and Prosthesis from MTEC who measured the data and recorded all 8 parameters for each participant. The examples of ground truth data we collected from the participants are as shown in Table 3 (left feet) and Table 4 (right feet). To ensure coherence with our training data in sit-to-stand pose, the experts instruct participants to place each foot gently on the foot scanner, avoiding any pressure or weight application.

3.4 Data Preparation

From the obtained images, firstly, we filtered out the defective data such as blurred images. Thus, the number of remaining images is 867, which consists of 259 images of the high arch foot, 321 images of the normal arch foot, and 287 images of the flat arch foot. Next, we pre-processed the remaining images by adjusting the scaling of the dimension down to 1024x1024 pixels so that they could be fed to the deep learning model. As the ratio of the original images is not a 1:1 relationship (Figure 5a), if we performed the scaling without any

Table 3: The table represents the sample of the ground truth of the foot parameters from the left feet of the participants. All are randomly picked up from our training dataset.

No.	Arch Type	Left feet parameters (cm)									
		A	B	C	D	E	F	G	H	I	AI
1	Normal	23.716	20.034	8.033	16.944	15.366	9.617	5.237	4.000	10.017	0.23
2	Normal	25.827	21.530	8.509	18.249	16.499	10.331	6.171	4.910	10.765	0.24
3	High	22.392	19.128	7.393	16.359	14.231	9.414	5.696	4.130	9.564	0.16
4	Normal	23.356	19.693	7.219	17.149	15.193	8.933	5.591	3.520	9.847	0.22
5	High	21.589	18.515	7.226	15.859	14.901	8.834	5.527	4.180	9.258	0.05
6	Flat	23.755	20.772	7.952	17.397	15.805	9.817	5.405	3.810	10.386	0.28

Table 4: The table represents the sample of the ground truth of the foot parameters from the right feet of the participants. All are randomly picked up from our training dataset.

No.	Arch Type	Right feet parameters (cm)									
		A	B	C	D	E	F	G	H	I	AI
1	Normal	23.580	19.795	8.110	17.492	15.211	9.365	5.360	3.960	9.898	0.23
2	Normal	25.954	21.605	8.890	18.979	16.384	10.422	6.401	5.150	10.803	0.23
3	High	22.802	19.092	7.170	16.479	14.825	9.065	5.677	4.550	9.546	0.08
4	Normal	23.734	20.189	7.499	17.074	15.299	9.063	5.709	3.470	10.095	0.23
5	High	21.363	18.347	7.147	15.629	14.684	8.926	5.531	3.940	9.173	0.11
6	Flat	23.148	20.603	7.803	17.394	15.318	9.763	5.553	4.350	10.301	0.27

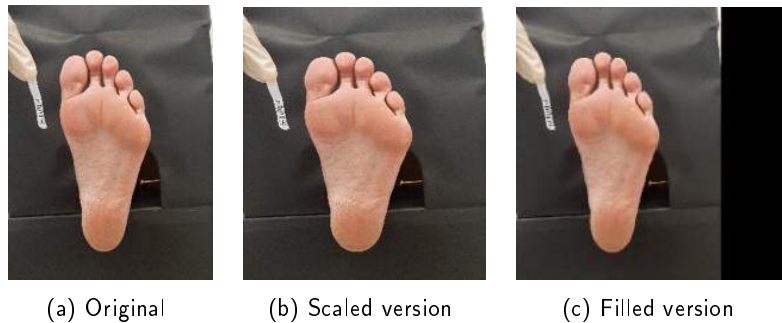(a) Original   (b) Scaled version   (c) Filled version

Figure 5: The comparison between (a) the original, (b) the scale one and (c) the filled with the black pixels in the data pre-processing for converting the original image of dimension down to 1024x1024 so that it fits the training model.
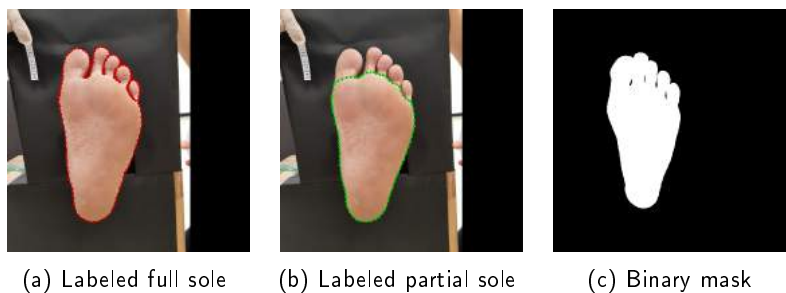

(a) Labeled full sole   (b) Labeled partial sole   (c) Binary mask

Figure 6: The labeled images of (a) the whole sole area and (b) the partial sole area or the partial of the sole cutting of the toes, and (c) the binary mask of (a). The images of (a) and (b) have been trained separately using the *U-Net* model. Note that we performed the (b) part just to retrieve the location of the highest part of the sole area without toes included.

adjusting, the image would be in the wrong scale (Figure 5b). Therefore, we performed the proportional scale instead and filled the blank areas with the black pixels (Figure 5c).

After that, we manually labeled the data using the *LabelMe* application [4], which consists of processes of the label of whole foot sole (Figure 6a) and foot sole omitted toes (Figure 6b). Then, the labeled area will be converted into the binary mask (Figure 6c).

We then used the binary mask and feed that to the deep learning model for the auto-segmentation process.

3.5 Foot Sole Segmentation Processes

The segmentation of the whole foot sole area and the partial sole area (toe-omitted) is achieved through a two-step process that involves the utilization of the deep learning model and color analysis techniques.

Firstly, deep learning techniques are employed to identify and isolate the sole and toe regions based on learned patterns and features. Secondly, color-based analysis is applied to further refine the segmentation by distinguishing the sole region from the surrounding areas based on variations in color properties.

For the deep learning utilization, we exploit the model called *U-Net* [19], which has been widespread for segmentation in the health-science field, and perform the training process. The data used for this process
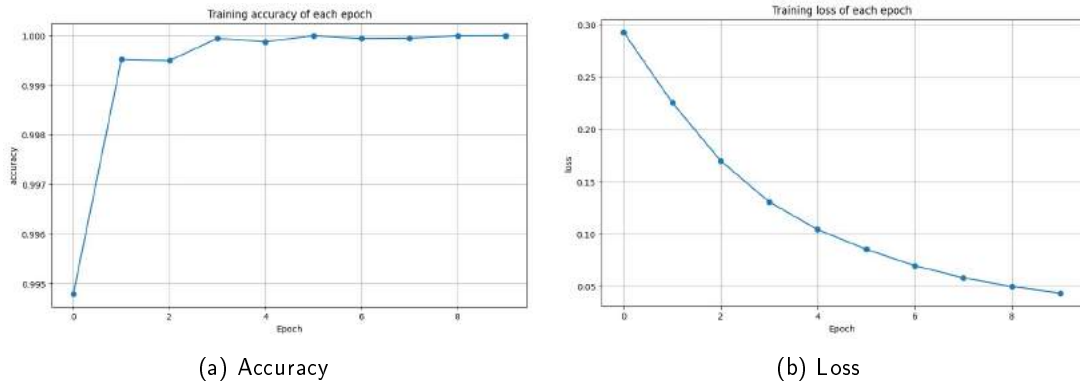
---

[4] http://labelme.csail.mit.edu/Release3.0/

(a) Accuracy (b) Loss

Figure 7: The reported of the training results of whole foot sole segmentation using *U-Net* consisting (a) Accuracy and (b) Loss on 867 images; the learning rate is 0.0001 with 8 epochs.

Table 5: The table represents the performance of the model for the whole foot segmentation.

Dataset	Accuracy	Precision	Recall	F1	Mean IOU
Training	0.995	0.998	0.974	0.986	0.972
Validation	0.996	0.999	0.975	0.987	0.975
Testing	0.994	0.999	0.970	0.983	0.969

was 867 images collected from 220 participants. The training process employs a learning rate of 0.0001 over 8 epochs to enhance model performance and the accuracy and loss during the training process are plotted in Figure 7. We trained the models with 70 percent of samples used for training, 10 percent of samples for validation, and 20 percent of samples for testing. Table 5 and Table 6 represent the performance of the whole foot area.

Subsequently, we applied the same technique to develop a segmentation model for the partial sole area (as in Figure 6b). It is important to note that this step was undertaken solely to eliminate the toes, thereby enabling us to pinpoint the location of the highest part of the partial sole area with the toes omitted. The performance is comparable to that of the whole area model.

Despite the satisfactory performance demonstrated by the deep learning technique, we addressed the noise pixel challenge as illustrated in Figure 8a. This issue arose during the data gathering process when laboratory personnel holding participant tag labels inadvertently moved too close to participants, causing noise from hand proximity to the foot. Even though this noise may arise from the data gathering process, we perceive it as an unavoidable circumstance, as in real-world scenarios, noise from the environment cannot be entirely

Table 6: The table represents the performance of the model for the foot sole segmentation.

Dataset	Accuracy	Precision	Recall	F1	Mean IOU
Training	0.998	0.993	0.993	0.993	0.986
Validation	0.998	0.992	0.994	0.993	0.987
Testing	0.997	0.993	0.987	0.988	0.980

(a) Binary mask   (b) Histogram of the skin tone range and minimum and maximum values
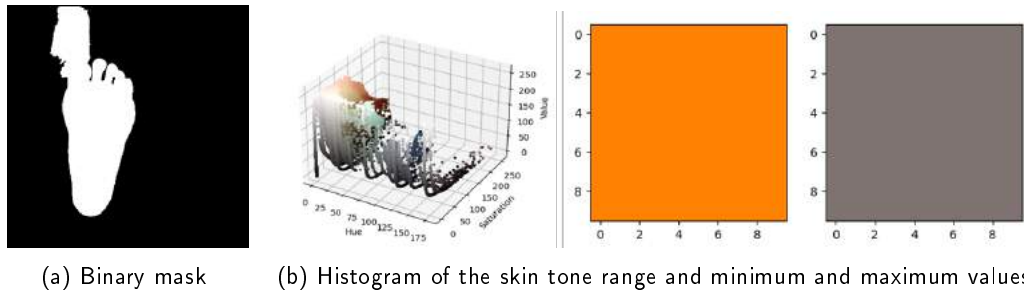
Figure 8: The example of (a) the binary mask that contains the unwanted part and (b) the HSV histogram used to extract minimum and maximum intensity of the participant skin tone range.

eliminated.

Therefore, we have opted to introduce a skin-tone segmentation process to effectively filter out such noise. As shown in Figure 8b, we initially build the HSV color histogram based on the skin color of the training data, and we employ this histogram to filter participants' skin color based on the minimum and maximum intensity thresholds corresponding to the participants' skin tone range and utilize these thresholds to filter out unwanted components from the environment.

After obtaining the binary masks from the deep learning technique and the color filter, the next step is to merge these results to obtain the final binary mask representing the location of the entire foot sole.

3.6 Parameter Extraction

To obtain 8 foot parameters of insole profiles as stated in Section 2.1, we apply an image processing technique to pin 7 feature points as shown in Figure 9. For instance, determining parameter **(A)**, which represents the foot length from heel to forefoot, involves identifying the pixel locations corresponding to both the heel and forefoot, denoted as feature points **(a)** and **(f)**, respectively. Therefore, to acquire each feature point from **(a)** to **(f)**, we conduct a search across the binary mask of size I_{width} x I_{height} to identify the most appropriate pixel location (x_i, y_j) where i is the pixel location in $[0, I_{width}]$ and j is the pixel location in $[0, I_{height}]$ according to the criteria as followed:

(a) or the highest points of the forefoot: We obtain this by searching within the toes' area and choosing the pixel location (x_a, y_a) where y_a is the smallest y value among others.

(b) or the highest points of the foot excluding the toes: We obtain this by searching within the segmented partial sole area (toe-omitted) and choosing the pixel location (x_b, y_b) where y_b is the smallest y value among others.

(c) or the point of the apex of 1st metatarsal head: This value depends on the side of the foot. If it is the left side, the pixel location (x_c, y_c) is where x_c is the maximum x within the right area of the image which is $[(I_{width}/2) + 1, I_{width}]$ range. Otherwise, the pixel location (x_c, y_c) is where x_c is the minimum x within the left area of the image which is $[0, (I_{width}/2)]$ range.

(d) or the point of the apex of 5th metatarsal head: This value depends on the side of the foot. If it is the left side which is $[(I_{width}/2) + 1, I_{width}]$ range, the pixel location (x_d, y_d) is where x_d is the minimum x within the right area of the image. Otherwise, the pixel location (x_d, y_d) is where x_d is the maximum x within the left area of the image which is $[0, (I_{width}/2)]$ range.
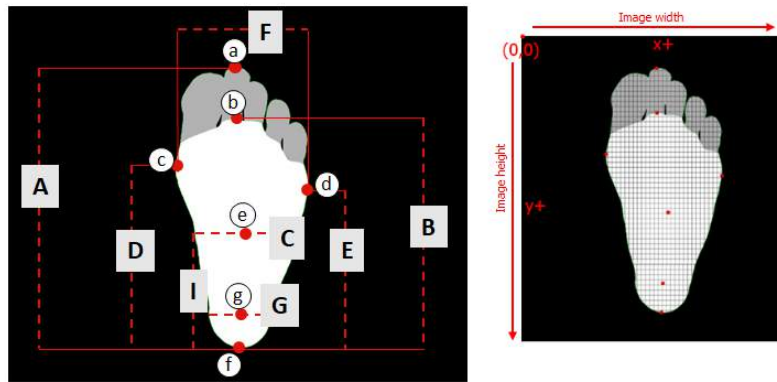
Figure 9: To obtain the insole profile parameters **(A)** to **(I)** (excluded **(H)**) described in Section 2.1, the 7 feature points **(a)** to **(g)** are extracted from the binary mask of the foot sole and partial sole area. The approach involves localizing the pixel coordinates (x, y) that adhere to the specified criteria.

(e) or the middle point of the foot: We obtain this by identifying the middle point (x_e, y_e) of the line from point **(a)** to point **(f)**.

(f) or the lowest point of the foot (the heel): We obtain this by searching within the hindfoot's area. To identify the hindfoot's area, we divide the foot into 3 parts: forefoot, midfoot, and hindfoot. To do that, Given the line from **(a)** to **(f)**, we cut that line into 3 and searched over the third area. Then, we choose the pixel location (x_f, y_f) where y_f is the largest y value among others within the hindfoot's area.

(g) or the point indicates the center location of the hindfoot: We obtain this by searching within the hindfoot's area and calculating the same way as in point **(f)**. Then, for each masked pixel row in the hindfoot's area, we identify the pixel location (x_g, y_g) by calculating the middle point of the widest row.

Note that in our implementation, we use a pixel coordinate system where the origin $(0, 0)$ resides at the top-left corner of the image; the x-axis extends horizontally to the right, while the y-axis progresses vertically downward. If the search results yield two locations, we opt for the first one. However, if more than two pixels are returned, we select the middle point among them.

After acquiring the aforementioned feature points, the subsequent step involved computing the length using the Euclidean distance between two feature points, with each foot parameter's value determined as follows:

(A) obtained by calculating the length from point **(a)** to point **(f)**.

(B) obtained by calculating the length from point **(b)** to point **(f)**.

(C) obtained by calculating the length of the row containing **(e)**.

(D) obtained by calculating the length from point **(c)** to **(f)**.

(E) obtained by calculating the length from point **(d)** to **(f)**.

(F) obtained by calculating the length from point **(c)** to **(d)**.

(G) obtained by calculating the length of the row containing **(g)**.

(a) Example A

(b) Example B

Figure 10: The examples of the parameter extraction using our method. The red dots represent the location of the predicted parameters.

(I) obtained by calculating the length from point e to point **(f)**.

Based on the outcomes of this experiment, the results will be expressed as the length of each parameter in terms of pixels. Therefore, as all the parameters we retrieved in this process are in pixel units, we calibrated them to be in centimeters unit for practical use. Specifically, the foot height (equally to parameter **(A)**) was acquired from the user in centimeters as derived in Equation 1.

$$R_{pc} = \frac{L_{af}}{L_A} \tag{1}$$

where R_{pc} is the ratio of pixels per centimeter L_{af} is parameter **(A)**, which is the length from point **(a)** to point **(f)**, and L_A is the foot height in centimeters retrieved from the user (in this study, we used the ground truth). This value served as the reference ratio to convert the measured length of the foot parameter from **(B)** to **(I)** in pixels to centimeters. Thus, once we obtain the R_{pc} value, we apply it as a multiplier to our results to convert from pixels to centimeters. We have made this selection under the assumption that in real-world scenarios, patients can accurately measure the length of their own feet. In addition, this method also shares similarities with the utilization of A4 paper as a reference point in various research projects and commercial foot measurement applications discussed before.

The outcomes are shown in Figure 10; red dots represent the location of key features we predicted.

Table 7: The table representing the difference error between the prediction and the ground truth from Figure 10a and 10b. Note that each alphabet labeling is referred to the foot parameters corresponding to Section 2.1. The unit is centimeters.

	Type	B	C	D	E	F	G	I
	Ground Truth	23.3616	9.0321	19.8686	17.5204	10.5878	70.418	11.6808
Figure 10a	**Prediction**	22.5688	9.8408	18.177	15.2085	11.9147	7.1976	11.3047
	Difference	0.7928	0.8087	1.6916	2.3119	1.3269	63.2204	0.3761
	Ground Truth	19.559	7.108	17.128	16.361	8.75	5.593	9.78
Figure 10b	**Prediction**	20.1957	6.9005	17.5945	15.2462	8.7069	5.5999	10.1159
	Difference	0.6367	0.2075	0.4665	1.1148	0.0431	0.0069	0.3359

Table 8: The table represents the maximum, the minimum and the average difference error between the prediction and the ground truth. Note that each alphabet labelling is referred to the Section 2.1.

Parameters	Maximum difference errors (cm)	Minimum difference errors (cm)	Average difference errors (cm)
B	5.5225	0.0039	0.5326
C	2.4355	0.0001	0.3803
D	8.0874	0.0023	0.9787
E	5.6405	0.0122	1.1337
F	2.8150	0.0003	0.5723
G	4.4015	0.0002	0.2248
I	3.1866	0.0005	0.2720

4 RESULTS AND DISCUSSION

Our methodology, involving the capture of sole images for measurement, is distinct from alternative technologies, each characterized by disparate mediums of operation. It is important to note that direct comparisons with high-precision 3D scanners may not be directly applicable, given our emphasis on mobility, cost reduction, and a diminished reliance on either machinery or human expertise. Consequently, we reported our performance metrics independently.

To evaluate the performance, we tested over 800 images and compared our prediction with the ground truth measured by the expert. Table 7 shows the difference between the ground truth and the prediction of the 7 foot parameters (from **(B)** to **(G)**, except **(H)**), and Table 8 presents the difference error statistics of the results. Our optimal outcomes were noted when the parameters demonstrated an error range within 0.0001 to 0.01 centimeters.

To gain more understanding, from the box plot of the errors as shown in Figure 11a, we can see that the maximum errors are the outliers. Furthermore, Figure 11b shows that only the 75th percentile of the parameters **(B)**, **(C)**, **(F)**, **(G)** and **(I)** contain the average of the errors less than 0.8 centimeters, which highlights the efficiency of our approach.

However, we can see that the parameters **(D)** and **(E)**, i.e., the length from the heel to the apex of 1st and 5th respectively, are the features contain the highest values of difference of more than one centimeters.
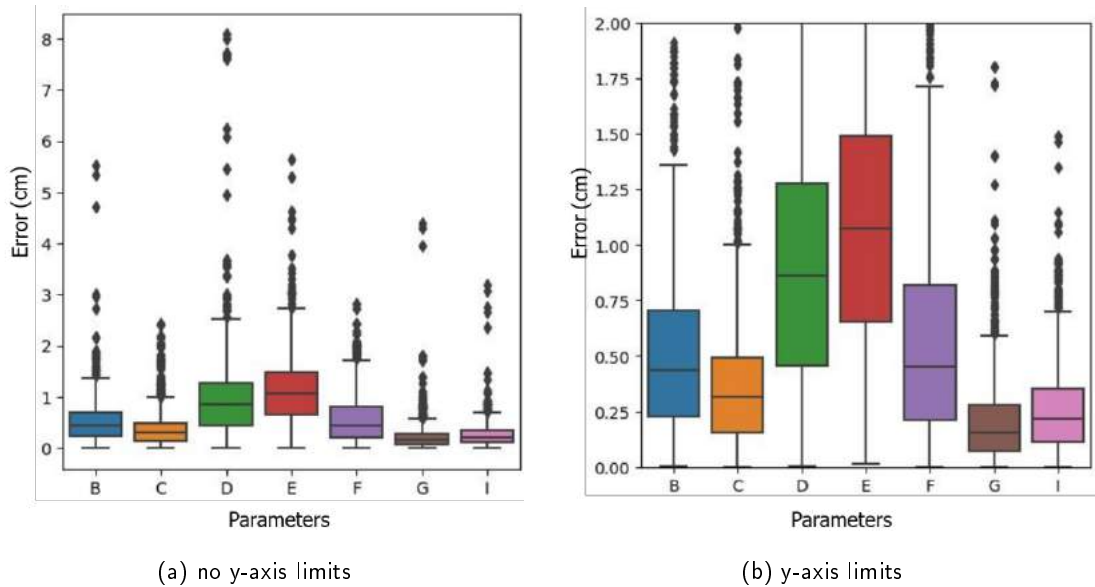
(a) no y-axis limits   (b) y-axis limits

Figure 11: The box plot representing the average error of each parameter depicted as (a) no y-axis limits and (b) y-axis limits at 2.00.

In the subsequent analysis, we found that the percentage of calculated variables **(B)**, **(C)**, **(F)**, **(G)**, and **(I)** with an error less than 0.8 centimeters exceeds 70%, particularly with variables **(C)**, **(G)**, and **(I)** have a percentage of error less than 0.8 centimeters exceeding 90%. However, the percentage of error for both parameters **(D)** and **(E)**, with values less than 0.8 centimeters, is relatively lower, at 46.13% and 33.27%, respectively.

In order to deepen our comprehension, we retrieved test data for analysis, revealing that one of the primary factors contributing to the outlier is attributable to the tilt of the foot as shown in Figure 12. This resulted in a misalignment of the foot, consequently influencing the precision of measurements for all parameters. Given that our methodology relies solely on images, even a minor tilt introduces the possibility of errors. This is attributed to the heightened sensitivity required for the precise determination of pixel locations corresponding to feature points. Specifically, as the average highest point of the forefoot (referred to as parameter **(A)** in Figure 1) is 23.89 centimeters, an incorrect selection of one pixel could result in errors for ± 0.22 centimeters. This finding also corresponds to our range of errors.

Another factor is due to capturing images in the sit-to-stand pose. As indicated in studies [13], [18], the absence of human weight pressure may lead to a slight displacement of foot parameters. This assertion is further validated by expert opinion, suggesting that adopting the sit-to-stand pose introduces potential stress on leg muscles and a reduction in weight-bearing, thereby impacting pressure dynamics compared to using a pressure scanner. The foot sole is ideally assessed under the load of the entire body weight. However, those studies also indicated that the parameter most significantly impacted is the navicular bone position, which consequently influences the arch height (parameter **(H)** in Figure 1); other parameters on the foot sole are less affected.

While we acknowledge the challenges posed by the sit-to-stand posture, it remains more practical for our research objectives focused on mobility and accessibility. Importantly, the sit-to-stand pose is easier for patients to capture their photo, aligning with our primary research goals. In addition, as per discussions with the experts, the insole production entails a degree of flexibility, allowing the acceptable deviation threshold for
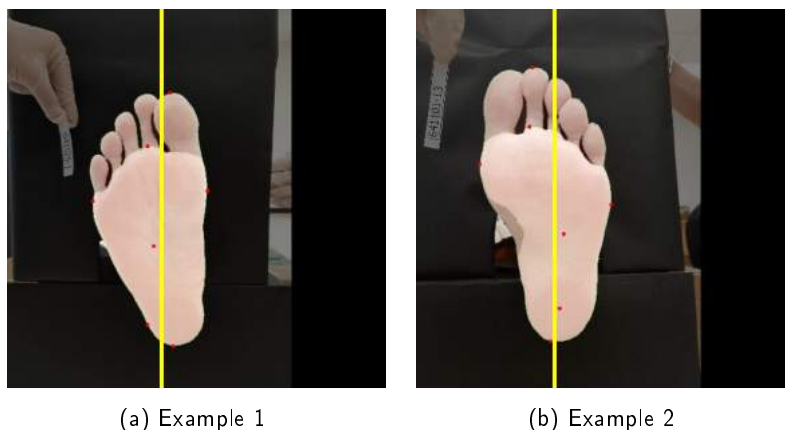
(a) Example 1   (b) Example 2

Figure 12: Examples of the feet that contain the high tilt. We can see the red dot at the heel position of each figure does not align with the vertical axis represented by the yellow line.

faulty lines to be within 1 cm. This number also correlates with the process of purchasing an insole in the real world, where users are typically advised to leave approximately 0.5-1 cm of space from their foot length. Notably, our approach demonstrates a high degree of compliance with these criteria.

Even though the current approach demonstrates promising results, achieving maximum precision necessitates further refinement of the insole profile. To this end, we propose a strategic enhancement of our methodology through the calibration of key points under both pressure and non-pressure conditions, with the aim of attaining better accuracy.

5 CONCLUSIONS

In our experimental findings, our work serves as a proof of concept for utilizing images for measurement purposes. Optimal outcomes were observed when the parameters exhibited an error range within 0.0001 to 0.01 centimeters, and an average range of approximately 0.001 to 1 centimeters, which shows the promising potential of the employed technique. In addition, our approach of utilizing a single image for the foot measurement offers a more compact approach compared to using multi-viewed or scanned images; it reduces the data complexity and storage requirements, making it more efficient and practical for everyday use.

However, there are some limitations associated with using a single image for foot measurement. One of the major concerns is foot alignment when capturing the images. Our future plans involve refining the data collection process by introducing intermediate mechanisms designed to filter out instances of incorrect foot alignment.

Another is by addressing the sit-to-stand issue, given its prevalence as a common pose for photographing individuals' soles, our focus will be on parameter tuning to mitigate the absence of weight pressure.

Furthermore, it is imperative to note that in this experimental setup, the evaluation of arch height (parameter **(H)** in Figure 1), a metric crucial for insole customization, has not been incorporated as of yet. We planned to integrate the use of the sided foot image into our project later to complete obtaining all parameters with the sit-to-stand parameters calibration.

In conclusion, our research underscores the viability of employing images captured by commonly owned mobile phones, coupled with advanced image processing and machine learning techniques, to facilitate the customization of insoles. This approach offers the potential to diminish reliance on costly devices and sensors, while also minimizing the necessity for expert intervention, thereby rendering the process more accessible at a

household level.

ACKNOWLEDGEMENTS

Pilailuck Panphattarasap, http://orcid.org/0000-0003-0210-8237
Pisit Praiwattana, http://orcid.org/0000-0002-6710-7801
Jidapa Kraisangka, http://orcid.org/0000-0002-4281-3603
Wanida Janvikul, http://orcid.org/0000-0002-2560-0075
Boonlom Thavornyutikarn, http://orcid.org/0000-0002-9814-0065
Wasana Kosorn, http://orcid.org/0000-0001-9339-0594
Nutdanai Nampichai, http://orcid.org/0000-0003-0926-2261
Wudhichart Sawangphol, http://orcid.org/0000-0001-7872-2482

REFERENCES

[1] Arteaga-Marrero, N.; Hernandez, A.; Villa, E.; Gonzalez-Perez, S.; Luque, C.; Ruiz-Alzola, J.: Segmentation approaches for diabetic foot disorders. Sensors, 21(3), 934, 2021. http://doi.org/10.3390/s21030934.

[2] Bouallal, D.; Douzi, H.; Harba, R.: Diabetic foot thermal image segmentation using double encoder-resunet (de-resunet). Journal of Medical Engineering & Technology, 46(5), 378–392, 2022. http://doi.org/10.1080/03091902.2022.2077997.

[3] Cavanagh, P.R.; Rodgers, M.M.: The arch index: a useful measure from footprints. Journal of biomechanics, 20(5), 547–551, 1987. http://doi.org/10.1016/0021-9290(87)90255-7.

[4] Chae, J.; Kang, Y.J.; Noh, Y.: A deep-learning approach for foot-type classification using heterogeneous pressure data. Sensors, 20(16), 4481, 2020. http://doi.org/10.3390/s20164481.

[5] Chun, S.; Kong, S.; Mun, K.R.; Kim, J.: A foot-arch parameter measurement system using a rgb-d camera. Sensors, 17(8), 1796, 2017. http://doi.org/10.3390/s17081796.

[6] Cisneros, L.L.; Fonseca, T.H.; Abreu, V.C.: Inter-and intra-examiner reliability of footprint pattern analysis obtained from diabetics using the harris mat. Brazilian Journal of Physical Therapy, 14, 200–205, 2010. http://doi.org/10.1590/S1413-35552010000300011.

[7] Dabholkar, T.; Agarwal, A.: Quality of life in adult population with flat feet. Int. J. Health Sci. Res, 10(8), 2020.

[8] Farhan, M.; Wang, J.Z.; Bray, P.; Burns, J.; Cheng, T.L.: Comparison of 3d scanning versus traditional methods of capturing foot and ankle morphology for the fabrication of orthoses: a systematic review. Journal of Foot and Ankle Research, 14, 1–11, 2021. http://doi.org/10.1186/s13047-020-00442-8.

[9] Jurca, A.; Žabkar, J.; Džeroski, S.: Analysis of 1.2 million foot scans from north america, europe and asia. Scientific reports, 9(1), 19155, 2019.

[10] Kok, F.; Charles, J.; Cipolla, R.: Footnet: An efficient convolutional network for multiview 3d foot reconstruction. In H. Ishikawa; C.L. Liu; T. Pajdla; J. Shi, eds., Computer Vision – ACCV 2020, 36–51. Springer International Publishing, Cham, 2021. ISBN 978-3-030-69544-6. http://doi.org/10.1007/978-3-030-69544-6_3.

[11] López López, D.; Bouza Prego, M.; Requeijo Constenla, A.; Saleta Canosa, J.L.; Casasnovas, A.B.; Tajes, F.A.: The impact of foot arch height on quality of life in 6-12 year olds. Colombia Médica, 45(4), 168–172, 2014.

[12] Maldonado, H.; Bayareh, R.; Torres, I.; Vera, A.; Gutiérrez, J.; Leija, L.: Automatic detection of risk zones in diabetic foot soles by processing thermographic images taken in an uncontrolled environment. Infrared Physics & Technology, 105, 103187, 2020. http://doi.org/10.1016/j.infrared.2020.103187.

[13] McPoil, T.G.; Cornwall, M.W.; Medoff, L.; Vicenzino, B.; Forsberg, K.; Hilz, D.: Arch height change during sit-to-stand: an alternative for the navicular drop test. Journal of foot and ankle research, 1, 1–11, 2008.

[14] Meneses-Claudio, B.; Alvarado-Díaz, W.; Flores-Medina, F.; Vargas-Cuentas, N.I.; Roman-Gonzalez, A.: Detection of suspicious of diabetic feet using thermal image. International Journal of Advanced Computer Science and Applications, 10(6), 2019.

[15] Menz, H.B.; Fotoohabadi, M.R.; Wee, E.; Spink, M.J.: Visual categorisation of the arch index: a simplified measure of foot posture in older people. Journal of foot and ankle research, 5(1), 1–7, 2012. http://doi.org/10.1186/1757-1146-5-10.

[16] Najafi, B.; Reeves, N.D.; Armstrong, D.G.: Leveraging smart technologies to improve the management of diabetic foot ulcers and extend ulcer-free days in remission. Diabetes/metabolism research and reviews, 36, e3239, 2020. http://doi.org/10.1002/dmrr.3239.

[17] Ozdemir, M.; Cascini, G.; Verlinden, J.C.: A mass personalization framework for knitted footwear. In 9th International Conference on Mass Customization and Personalization–Community of Europe (MCP–CE 2020), 175–183, 2020.

[18] Pohl, M.B.; Farr, L.: A comparison of foot arch measurement reliability using both digital photography and calliper methods. Journal of foot and ankle research, 3, 1–6, 2010.

[19] Ronneberger, O.; Fischer, P.; Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention (MICCAI), vol. 9351 of LNCS, 234–241. Springer International Publishing, Cham, 2015. http://doi.org/10.1007/978-3-319-24574-4_28.

[20] Sawangphol, W.; Panphattarasap, P.; Praiwattana, P.; Kraisangka, J.; Noraset, T.; Prommin, D.: Foot Arch Classification via ML-based Image Classification. Computer-Aided Design and Applications, 20(4), 600–613, 2023. http://doi.org/10.14733/cadaps.2023.600-613.

[21] Shilov, L.; Shanshin, S.; Romanov, A.; Fedotova, A.; Kurtukova, A.; Kostyuchenko, E.; Sidorov, I.: Reconstruction of a 3d human foot shape model based on a video stream using photogrammetry and deep neural networks. Future Internet, 13(12), 315, 2021. http://doi.org/10.3390/fi13120315.

[22] Tulloch, J.; Zamani, R.; Akrami, M.: Machine learning in the prevention, diagnosis and management of diabetic foot ulcers: A systematic review. IEEE Access, 8, 198977–199000, 2020. http://doi.org/10.1109/ACCESS.2020.3035327.

[23] Wang, H.; Liu, F.; Fan, R.: A research on foot size measurement algorithm based on image. In Journal of Physics: Conference Series, vol. 1903, 012004. IOP Publishing, 2021. http://doi.org/10.1088/1742-6596/1903/1/012004.

[24] Witana, C.P.; Xiong, S.; Zhao, J.; Goonetilleke, R.S.: Foot measurements from three-dimensional scans: A comparison and evaluation of different methods. International Journal of Industrial Ergonomics, 36(9), 789–807, 2006. http://doi.org/10.1016/j.ergon.2006.06.004.