

Automated measurement of foot arch structures on weight-bearing lateral radiographs using a CNN-based deep learning model

Jiaming Xu

The Second Clinical Medical College of Shanxi Medical University

Wei Wang

The Second Clinical Medical College of Shanxi Medical University

Yu Han

The Third Clinical Medical College of Shanxi Medical University

Shipeng Zhu

Shanxi Medical University

Hangtian Dong

Shanxi Medical University

Zongyan Liu

Shanxi Medical University

Shengdong Ren

a1242076107@163.com

The Second Clinical Medical College of Shanxi Medical University

Article

Keywords: Foot Arch Structure, Deep Learning, Convolutional Neural Network, Automated Measurement, X-ray

Posted Date: June 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9872041/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Purpose

The foot arch is crucial for foot stability, and structural abnormalities may cause dysfunction. Manual measurement is limited by accuracy and efficiency. This study develops a rapid, straightforward CNN and deep learning model to measure foot arch data and evaluate its performance.

Design:

Prospective, multicentre, diagnostic accuracy study with blinded comparison between automated model measurements and manual reference standards.

Setting:

Secondary care; data were collected from multiple hospitals in Shanxi Province, China.

Participants:

A total of 897 consecutive volunteers (mean age 37 ± 22 years; 454 females, 443 males; all self-reported Han Chinese ethnicity) undergoing weight-bearing lateral foot radiography were enrolled. Exclusion criteria were: non-standard positioning (e.g., foot rotation, non-weight-bearing); incomplete visualisation of key anatomical structures; presence of severe foot deformities, post-fracture internal fixation, or foot tumours; and poor image quality. After quality review, 1207 images from the 897 participants were included in the final analysis. The independent test set comprised 126 images from 87 participants (46 females, 41 males; mean age 45 ± 19 years).

Primary and secondary outcome measures:

The primary outcomes were the model's measurements of four foot arch angles – medial longitudinal arch (MLA), lateral longitudinal arch (LLA), anterior arch (AA), and posterior arch (PA) – compared to manual gold-standard measurements. Agreement was assessed by mean differences (95% confidence intervals) and area under the ROC curve (AUC) for binary classification (normal vs. abnormal). Secondary outcomes included sensitivity, specificity, precision, F1-score, and overall accuracy derived from confusion matrices.

Methods

Participants were divided into training, validation, and test sets (7:2:1 at the patient level). The four angles were measured manually by two senior doctors as the gold standard. Images were annotated with LabelMe and processed using DeepLabV3 + and YOLOv8. The model was optimised on the validation set and final accuracy assessed on the test set (126 images). Paired t-tests, confusion matrices, and ROC curves compared model measurements to manual annotations.

Results

No significant difference was found between model and manual measurements ($P > 0.05$). Mean differences (95% CI) were: MLA -0.45 ($-1.07, 0.17$); LLA 0.56 ($-0.59, 1.72$); AA 0.52 ($-0.06, 1.09$); PA -0.55 ($-1.12, 0.01$). All AUC values exceeded 0.95, indicating high measurement accuracy.

Conclusion

The CNN-and-deep-learning model accurately and efficiently measures foot arch data, providing a reliable reference for clinical assessment and treatment.

Article summary

Strengths and limitations of this study

1. This was a prospective, multicentre diagnostic accuracy study with blinded comparison between automated model measurements and manual reference standards.
2. Consecutive participants were enrolled from multiple hospitals, and the dataset was split at the patient level into training, validation, and independent test sets in a 7:2:1 ratio.
3. Two senior surgeons performed manual measurements as the gold standard, with a third adjudicating any disagreement exceeding 2° .
4. The model integrated DeepLabV3+ for semantic segmentation of foot bones and YOLOv8 for detection of challenging anatomical landmarks.
5. The study did not differentiate between left and right feet, and separate models for paediatric or severely deformed feet were not developed.

Introduction

Foot arch structures consist of 26 bones, ligaments that link several arthrosis and muscles that keep foot stability[1]. This structure is a stable pedestal support our body weight in quiescent state and can play an important role in buffer and building up energy through running and jumping as a spring[2]. Because of its irreplaceable effect during sports, once the structure is injured, the movement of patients will be affected and serious patients may face the risk of losing moving ability[3]. According to the foot injury cases of National Collegiate Athletic Association (NCAA) collected by researchers in the past ten

years, the foot arch injury is one of the highest surgical rate diseases, while it is also a severe structure damage can end their season or career[4]. Therefore, it is important to assess foot arch structures accurately and timely in order to prevent worsening of symptoms and re-injury due to without a clear diagnosis.

Currently, Transverse Arch and longitudinal arch have been widely used in evaluate foot structures[5]. Common measuring methods are as follows. Routine physical examination: measuring foot arch height (AH) which is the distance of tuberosity of the navicular to the ground[6]. Measuring footprints: getting footprints by ink or pressure sensor and calculating the arch index (AI),the ratio of the area of the central one-third region to the area of the total region excluding the toes[7]. 3D-Scanner:using commercial RGB-D camera[8] to detect 3D characters color data from plantar surface while the data such as AH, AI, arch width(AW)will be measured automatically. However, these methods had the manual errors and coupled with the limited scalability of pressure sensors inherent.

The conventional Radiology examination is still a Cost-effective and Convenient way in nearly two decades of professional practice[9,10]. These imaging examinations provide a comprehensive assessment of the foot's arch structure, enabling the detection of subtle injuries that are often missed by physical examination alone[11]. Furthermore, measurements obtained from these images can serve as objective, quantifiable metrics for evaluating the severity of arch structural deterioration. Evaluating their foot arch structures by measuring four different angles MLA angle, LLA angle, AA angle and PA angle. All these four angle can be measured through Weight-bearing Lateral View of the Foot. Compared to conventional measurements, it can evaluate the Transverse Arch and longitudinal arch[12,13], which are the most important structures of foot arch comprehensively. However, Limited by the **human error** and skill of Medical Technician, this method has flaws in both accuracy and efficiency. It is fortunate that the remarkable progress in technology in recent years has enabled artificial intelligence to overcome this shortcoming.

In the past decade, deep learning combined with medical images has led to exciting advances in automated recognition and diagnosis[14]. Artificial intelligence and deep learning have been developed for the diagnosis and classification of multiple fracture recognition, diabetic retinopathy[15,16], optic nerve papilledema[17] and other diseases. this method achieved an excellent accuracy and precision in evaluating the damage of foot arch structure by deep learning. The purpose of this study was to develop an automatic recognition model to automatically measure foot arch data on Weight-bearing Lateral View of the Foot images. The model can accurately and efficiently obtain the required data, providing a reference for clinical assessment and treatment.

Methods

Study Design, Setting and Participants

This was a prospective, multicentre diagnostic accuracy study. Volunteers undergoing weight-bearing lateral foot radiography at multiple hospitals across Shanxi Province, China, between January 2025 and

December 2025 were consecutively enrolled. All data were collected in secondary care settings. All participants were self-reported Han Chinese. The study protocol was approved by the Medical Ethics Committee of Shanxi Medical University (Reference No. 2026YX078) and written informed consent was obtained from all participants (or legal guardians for minors).

Based on the empirical rule for medical imaging AI studies and considering the feasibility of consecutive collection across multiple centres, we targeted a total sample of 897 participants. Through consecutive collection, we obtained 1324 weight-bearing lateral foot radiographs from these 897 patients. After quality review, 117 radiographs were excluded due to the following criteria: (1) non-standard positioning (e.g., foot rotation, non-weight-bearing); (2) incomplete visualisation of all key anatomical structures (calcaneus, talus, first and fifth metatarsal heads); (3) presence of severe foot deformities (e.g., clubfoot, tarsal coalition), post-fracture internal fixation, or foot tumours that could alter normal anatomical landmarks; (4) poor image quality (e.g., overexposure, blurring). In cases where any of the four measured angles could not be reliably annotated due to missing or unrecognisable anatomical landmarks, the entire image was excluded from the analysis. Consequently, 1207 images from 897 patients (mean age, 37 ± 22 years) were included in the final analysis.

Data Partitioning and Outcome Measures

The split was performed at the patient level to avoid data leakage: all images from the same patient were assigned to the same set. Stratification was based on sex and age (three strata: ≤ 18 , 19–64, ≥ 65 years). Participants were divided into training, validation, and test sets in a 7:2:1 ratio, resulting in 823 images (596 patients) for training, 258 images (214 patients) for validation, and 126 images (87 patients) for the independent test set. The training and validation sets were used for model development and optimisation, while the test set was held out for final, independent evaluation.

The primary outcomes were the model's measurements of four foot arch angles – medial longitudinal arch (MLA), lateral longitudinal arch (LLA), anterior arch (AA), and posterior arch (PA) – compared to manual gold-standard measurements. The prespecified clinically relevant thresholds for binary classification were $LLA > 150^\circ$, $MLA > 130^\circ$, $AA < 13^\circ$, and $PA < 15^\circ$. Secondary outcomes included sensitivity, specificity, precision, F1-score, overall accuracy, and area under the ROC curve (AUC). The distribution of patients and images across these sets is detailed in Table 1.

Gold Standard (Manual Measurements)

Two senior orthopaedic surgeons independently performed manual measurements of all four angles using standardised protocols, blinded to the model's predictions. In case of a disagreement exceeding 2° , a third senior surgeon was consulted to reach consensus, and the consensus values served as the ground truth. All manual measurements were performed using a picture archiving and communication system (PACS) workstation with digital protractor tools. No missing data occurred for any of the four measured angles in the final dataset because images with incomplete annotations were excluded a priori (as described above). During model testing, the automated measurements were compared to

these manually obtained ground truth values by a separate researcher who was not involved in model development.

Measurement Definitions

In this research we used four angles to evaluate the foot arch structure. The MLA angle is defined by three anatomical landmarks: the most inferior points of the talus, the first metatarsal head, and the calcaneus, measured at the intersection of the lines connecting these points. The LLA angle is defined by the most inferior points of the calcaneocuboid joint, the calcaneus, and the fifth metatarsal head. The AA angle is defined by the most inferior points of the first metatarsal head, the first tarsometatarsal joint, and the calcaneus. The PA angle is defined by the most inferior points of the calcaneus, the calcaneocuboid joint, and the fifth metatarsal head, geometrically formed at the calcaneocuboid joint. All four angles are expressed in degrees ($^{\circ}$). (Fig. 1)

Image Annotation and Model Architecture

All weight-bearing foot X-ray images in the training set were annotated using LabelMe software. The entire foot was segmented into five anatomical regions: the talus region, calcaneus region, first metatarsal bone region, fifth metatarsal bone region, and a special region comprising the navicular, cuboid, and cuneiform bones. The entire foot was segmented into five anatomical regions, as shown in Fig. 2. The annotated images were used to train DeepLabV3+ with a ResNet50 backbone pre-trained on ImageNet, yielding a preliminary model capable of performing initial image segmentation. Subsequently, on the processed images, key anatomical landmarks that are relatively straightforward to identify were directly defined, including the most inferior points of the talus, calcaneus, first metatarsal head, and fifth metatarsal head. The two more challenging landmarks – the most inferior points of the calcaneocuboid joint and the first tarsometatarsal joint – were annotated using LabelMe and used to train the YOLOv8m object detection framework, ultimately producing a model capable of identifying the required target landmarks.

Hyperparameters for DeepLabV3+: input image size 512×512 pixels; ResNet50 backbone; cross-entropy loss; AdamW optimiser with initial learning rate 1×10^{-4} ; batch size 16; early stopping based on validation loss. Hyperparameters for YOLOv8m : CIoU loss; same optimiser settings; image augmentation included random rotation ($\pm 15^{\circ}$), horizontal flipping, and random brightness-contrast adjustment. Training was conducted on a single NVIDIA RTX 4090 GPU. All code was implemented in Python 3.10 using PyTorch 2.0.

Model Calibration on the Validation Set

The entire process of the proposed automated measurement system is illustrated in Fig. 3. The validation set (258 images from 214 participants; 106 males, 108 females; age range 6–74 years; mean age 33.5 ± 21.9 years) was used for model optimisation and parameter calibration. The four foot arch angles predicted by the model on the validation set were subjected to linear regression analysis against

the gold standard values manually annotated by two senior orthopaedic surgeons. Regression equations were fitted using SPSS25®, with the model-predicted values as the independent variable (x) and the manual measurements as the dependent variable (y). The following linear regression equations were obtained: LLA: $y = 0.851x + 22.672$; MLA: $y = 0.789x + 26.726$; AA: $y = 0.764x + 3.345$; PA: $y = 0.976x + 0.677$. These correction parameters were subsequently integrated into the final algorithm to optimise the model's measurement accuracy on the test set.

Performance Evaluation on the Test Set

The final calibrated model was applied to the independent test set (126 images from 87 participants; mean age 45 ± 19 years). Agreement between model measurements and manual gold-standard measurements was assessed using paired t-tests; mean differences and 95% confidence intervals were calculated for each angle. Effect sizes were calculated using Cohen's d. Using manual measurements as the gold standard, angles were classified as "normal" or "abnormal" based on the prespecified clinical thresholds. From the resulting confusion matrices, sensitivity, specificity, precision, F1-score, and overall accuracy were computed. ROC curves were plotted, and AUC was calculated for each foot arch angle..

Statistical Analysis

Statistical analyses were performed using SPSS25®. Normality of continuous data was assessed using the Shapiro–Wilk test. Between-group comparisons were performed using paired t-tests for normally distributed data or Wilcoxon signed-rank tests for non-normally distributed data. A p-value < 0.05 was considered statistically significant. No multiplicity adjustment was applied due to the exploratory nature of the study. Data are presented as mean $\pm$ standard deviation unless otherwise specified.

Reporting Guidelines

This study followed the STARD (Standards for Reporting of Diagnostic Accuracy Studies) and TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) reporting guidelines. Given the artificial intelligence/machine learning nature of our diagnostic accuracy model, we also incorporated the key AI-specific items from the STARD-AI and TRIPOD+AI extensions. The completed checklists are provided in the supplementary materials.

Result

1. Demographic characteristics of the study population

A total of 897 participants (mean age 37 ± 22 years; 454 females, 443 males) were included. The demographic characteristics of the overall dataset and the test set are summarized in Table2.

2. Model Calibration and Linear Regression Analysis

During the model development phase, linear regression analysis was performed using the validation set to calibrate the model's raw outputs. The analysis revealed a strong linear correlation between the model's preliminary predicted angle values and the manual gold standard measurements (Fig.4) .These correction parameters were subsequently integrated into the final algorithm to optimize the model's measurement accuracy on the test set.

3. Agreement Between Model and Manual Measurements

The calibrated deep learning model was applied to the independent test set comprising 126 images. The results showed no statistically significant difference between the angles measured by the model and those measured manually by human raters. The p-values from the paired t-tests were 0.152 for the MLA, 0.338 for the LLA, 0.077 for the AA, and 0.055 for the PA. The mean differences and their 95% confidence intervals for each angle were as follows: MLA: -0.45° (-1.07° , 0.17°); LLA: 0.56° (-0.59° , 1.72°); AA: 0.52° (-0.06° , 1.09°); PA: -0.55° (-1.12° , 0.01°) (Table3). The effect sizes (Cohen's d) for all comparisons were negligible ($d < 0.2$), indicating minimal practical difference despite minor numerical contrasts.

4. Classification Performance

To assess the deep learning model's performance in foot arch angle measurement, we took the clinical manual measurements as the gold standard. The results were classified into normal or abnormal categories according to established thresholds (LLA Angle $>150^{\circ}$, MLA Angle $>130^{\circ}$, AA Angle $<13^{\circ}$, PA Angle $<15^{\circ}$). Table4 presents the confusion matrices for all four foot arch angles. The precision, sensitivity, specificity, and F1 score values for each angle are shown in Table5. The AUCs for the four angles were 0.952, 0.980, 0.966 and 0.951. Fig.5 shows the ROC curves and AUC values.

Discussion

The foot arch plays an important role in maintain foot function. Suffered foot arch injury will lead to pain symptoms immediately and result in serious damage of foot biological structures[18]. Currently, imaging examination such as X-ray, had great superiority in displaying the damage of foot structures directly[19]. On the other hand, Artificial intelligence such as CNN and DL can resolve the dilemma caused to manual measurements[20, 21]. A key strength of this study lies in the integration of medical imaging with artificial intelligence, yielding a rapid and accurate tool for the assessment of foot arch structure.

Previously, there were two main methods that researchers evaluate the foot arch structures. One of them was obtaining patients'AI of foot. This kind of measurement can be done by both manual work and machine[22]. The manual measurement was getting the footprints and calculating the AI artificially. Using machines such as optical imaging, camera and scanner, the AI would be acquired automatically by analyzing the data from equipment mentioned before. Nowadays, these methods are still widely used. The other one was measuring the height of the tuberosity of the navicular or measuring the complex foot data contained height, length and volume by 3d-scanner directly[23, 24].This model addresses both the

inaccuracy of manual measurements and the limited accessibility of advanced instruments, such as 3D scanners.

Our research demonstrated the difference between the data measured by model and data measured by manual measurement is not significant. The P-value of MLA Angle, LLA Angle, AA, and PA are 0.152, 0.338, 0.077 and 0.055 respectively. All of them are bigger than 0.05, and the differences were not statistically significant. Therefore, the data measured by two methods are remarkably consistent. In order to estimate magnitude of difference more comprehensively, we further calculate the effect size (Cohen's d). The effect sizes for all parameters were negligible. This fact is particularly similar to the result of statistical test, which means though there are minor contrast in mean, but the statistical and practical significance is negligible.

Based on the confusion matrix, the model demonstrates excellent and robust overall performance. The accuracy rate exceeded 90% for all angles, indicating that its predictions are consistent with the actual conditions in the vast majority of cases. The model shows great performance in sensitivity, specificity and precision.

The above metrics collectively reveal that the model is not only effective at identifying the majority of positive cases but also highly reliable at excluding negative instances, thereby accurately filtering out potential interference.

Furthermore, the model's F1-Score reached 0.920, 0.930, 0.907 and 0.889. This composite metric further confirms that strikes a good balance between precision and recall, underscoring the high consistency and reliability of its predictive outcomes.

The ROC analysis demonstrates that our model possesses excellent discriminatory power, as evidenced by high AUC values (0.952–0.980). The AUC, which is equivalent to the probability that the model will correctly rank a randomly chosen positive instance higher than a negative one, provides a threshold-independent measure of diagnostic accuracy[25]. Unlike conventional accuracy metrics that depend on a single cut-off value, the ROC curve integrates performance across all possible classification thresholds, making it particularly suitable when clinical thresholds may be subject to variation[26]. In our study, all four AUC values exceeded 0.95, indicating that the model has a > 95% probability of correctly distinguishing abnormal from normal foot arch angles—a level widely considered as excellent diagnostic performance[27]. These properties justify the use of ROC and AUC as the primary metrics for evaluating the model's binary classification accuracy.

Our study has several limitations. First, our study included children and adolescents, whose foot arch morphology may differ from that of adults[28, 29]. This could adversely affect image segmentation and keypoint detection.. Second, for plain radiographs of extremely deformed feet, the angles measured by our model remain subject to a degree of uncertainty, which can be attributed to the limited number of abnormal samples in our study. Third, our model did not distinguish between left and right feet during training and evaluation. Although the four measured angles are defined symmetrically, treating left and

right foot images as equivalent may have slightly affected measurement precision in specific subgroups. Fourth, while we employed paired t-tests, confusion matrices, and ROC curves to evaluate agreement and classification performance, we did not employ additional statistical methods such as Bland-Altman plots, intraclass correlation coefficient (ICC), mean absolute error (MAE), or root mean square error (RMSE) for a more comprehensive evaluation. Inclusion of these metrics in future studies will enable more comprehensive and transparent performance evaluation.

To further enhance model accuracy, future work will include stratifying abnormal samples into distinct subgroups (e.g., children, adolescents, and severe deformities) and training separate models for each, as well as incorporating foot-side labels to develop side-specific models. These approaches aim to optimize measurement for atypical arch morphologies and address potential lateralization effects.

Conclusion

This study presents an AI-based measurement that enables the rapid and accurate acquisition of key foot arch parameters, thereby facilitating the assessment of functional impairment and activity limitations. The method provides a reliable basis for subsequent clinical evaluation. Compared to traditional measurement techniques, it offers distinct advantages in both efficiency and objectivity.

Declarations

Ethics statement

The Medical Ethics Committee of Shanxi Medical University approved the study protocol (Reference Number: 2026YX078). All methods were carried out in accordance with relevant guidelines and regulations. Prior to enrollment, we obtained written informed consent from all participants or their legal guardians (if under 18 years of age) for the use of their clinical data for research and publication.

Acknowledgments

The authors thank all participants for their cooperation, and the staff at the collaborating hospitals for their assistance in data collection..

Declaration of Interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

JX: Writing – review & editing, Writing – original draft. WW: Writing – review & editing, Writing – original draft. YH: Writing – original draft, Writing – review & editing. SZ: Writing – review & editing, Writing –

original draft. HD: Writing – review & editing, Writing – original draft. ZL Writing – original draft, Writing – review & editing. SR: Writing – review & editing, Writing – original draft.

Data Availability Statement

The raw weight-bearing lateral foot radiographs cannot be shared publicly due to patient privacy and institutional ethics approval (Medical Ethics Committee of Shanxi Medical University, Reference No. 2026YX078). The anonymised measurement data (four foot arch angles for each image) that support the findings of this study are fully presented within the paper in the Results section and Tables 3–5. All relevant derived data are already included in the manuscript.

Additional information

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This study was supported by The Research Project of the Health Commission of Shanxi Province (No. 2025YD032).

References

1. Davis, D. J., Challis, J. H. Characterizing the mechanical function of the foot's arch across steady-state gait modes. *J. Biomech.* **151**, 111529 (2023).
2. Welte, L., Holowka, N. B., Kelly, L. A., et al. Mobility of the human foot's medial arch helps enable upright bipedal locomotion. *Front. Bioeng. Biotechnol.* **11**, 1155439 (2023).
3. Renjewski, D., Lipfert, S., Günther, M. Foot function enabled by human walking dynamics. *Phys. Rev. E* **106**, 064405 (2022).
4. Chan, J. J. Epidemiology of Severe Foot Injuries in US Collegiate Athletes. *Orthop. J. Sports Med.* **9**, 23259671211001131 (2021).
5. Draghi, F., Gitto, S., Bortolotto, C., et al. Imaging of plantar fascia disorders: findings on plain radiography, ultrasound and magnetic resonance imaging. *Insights Imaging* **8**, 69–78 (2017).
6. Finestone AS, Petrov K, Agar G, Honig A, et al. Pattern of outsole shoe heel wear in infantry recruits. *J. Foot Ankle Res.* **5**, 27 (2012).
7. Menz, H. B., Fotoohabadi, M. R., Wee, E., et al. Visual categorisation of the arch index: a simplified measure of foot posture in older people. *J Foot Ankle Res* **3**, 5(1):10 (2012).
8. Chun, S., Kong, S., Mun, K.-R., et al. A foot-arch parameter measurement system using a RGB-D camera. *Sensors* **17**, 1796 (2017).
9. Tárnoki DL, Karlinger K, Ridge CA, et al. Lung imaging methods: indications, strengths and limitations. *Breathe Sheff. Engl.* **20**, 230127 (2024).

10. Miyamoto T, Otake Y, Nakao S, *et al.* 4D-foot analysis on effect of arch support on ankle, subtalar, and talonavicular joint kinematics. *J. Orthop. Sci. Off. J. Jpn. Orthop. Assoc.***28**, 1337–1344 (2023).
11. Ning X, Ru T, Zhu J, *et al.* Deep learning-assisted segmentation of X-ray images for rapid and accurate assessment of foot arch morphology and plantar soft tissue thickness. *Sci. Rep.***14**, 20028 (2024).
12. Puszczalowska-Lizis, E., Krawczyk, K., Omorczyk, J. Effect of longitudinal and transverse foot arch on the position of the hallux and fifth toe in preschool children in the light of regression analysis. *Int. J. Environ. Res. Public. Health***19**, 1669 (2022).
13. Wildenauer, E., Muller, W. The effects of the musculus fibularis longus on the transverse and lateral longitudinal arch of the foot. *Arch. Orthop. Unfallchir.***46**, 406–410 (1954).
14. Jiang, Y., Yang, M., Wang, S., *et al.* Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Commun. Lond. Engl.***40**, 154–166 (2020).
15. Miere A, Capuano V, Kessler A, *et al.* Deep learning-based classification of retinal atrophy using fundus autofluorescence imaging. *Comput. Biol. Med.***130**, 104198 (2021).
16. Chen Y, Sun Q, Li Z, Zhong Y, *et al.* Development and validation of a deep learning model using convolutional neural networks to identify femoral internal fixation device in radiographs. *Skeletal Radiol.***52**, 1577–1583 (2023).
17. Milea D, Najjar RP, Zhubo J, *et al.* Artificial intelligence to detect papilledema from ocular fundus photographs. *N. Engl. J. Med.***382**, 1687–1695 (2020).
18. Liu J, Deng M, Wang W, *et al.* A foot structure study of new arch flexibility grading system based on three-dimensional arch volume. *Chin. J. Traumatol.***26**, 329–333 (2023).
19. Teh, J., Suppiah, R., Sharp, R., *et al.* Imaging in the assessment and management of overuse injuries in the foot and ankle. *Semin. Musculoskelet. Radiol.***15**, 101–114 (2011).
20. Tran KA, Kondrashova O, Bradley A, *et al.* Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med.***13**, 152 (2021).
21. Yoon, T., Kang, D. Bimodal CNN for cardiovascular disease classification by co-training ECG grayscale images and scalograms. *Sci. Rep.***13**, 2937 (2023).
22. Zheng, T., Yu, Z., Wang, J. *et al.* A new automatic foot arch index measurement method based on a flexible membrane pressure sensor. *Sensors***20**, 2892 (2020).
23. Rogati, G., Leardini, A., Ortolani, M., *et al.* Semi-automatic measurements of foot morphological parameters from 3D plantar foot scans. *J. Foot Ankle Res.***14**, 18 (2021).
24. Liu Q, Zhao C, Yang X, *et al.* Biomechanics of transverse axis of medial longitudinal arch of children's foot based on 3D scanning. *Front. Pediatr.***11**, 1197439 (2023).
25. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med.* **4**, 627–35 (2013).
26. Obuchowski NA. Receiver operating characteristic curves and their use in radiology. *Radiology.* **229**, 3-8 (2003).

27. Çorbacioğlu ŞK, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. *Turk J Emerg Med.* **23**, 195-198 (2023).
28. Chang, C.-H., Yang, W.-T., Wu, C.-P., et al. Would foot arch development in children characterize a body maturation process? A prospective longitudinal study. *Biomed. J.* **45**, 828–837 (2022).
29. Escalona-Marfil C, Prats-Puig A, Ortas-Deunosajut X, et al. Children's foot parameters and basic anthropometry - do arch height and midfoot width change? *Eur. J. Pediatr.* **182**, 777–784 (2023).

Tables

Tables 1 to 5 are available in the supplementary files section

Figures



Figure 1

Schematic illustration of the MLA, LLA, AA, and PA angles. Each angle is formed by the intersection of lines connecting three key anatomical landmarks.

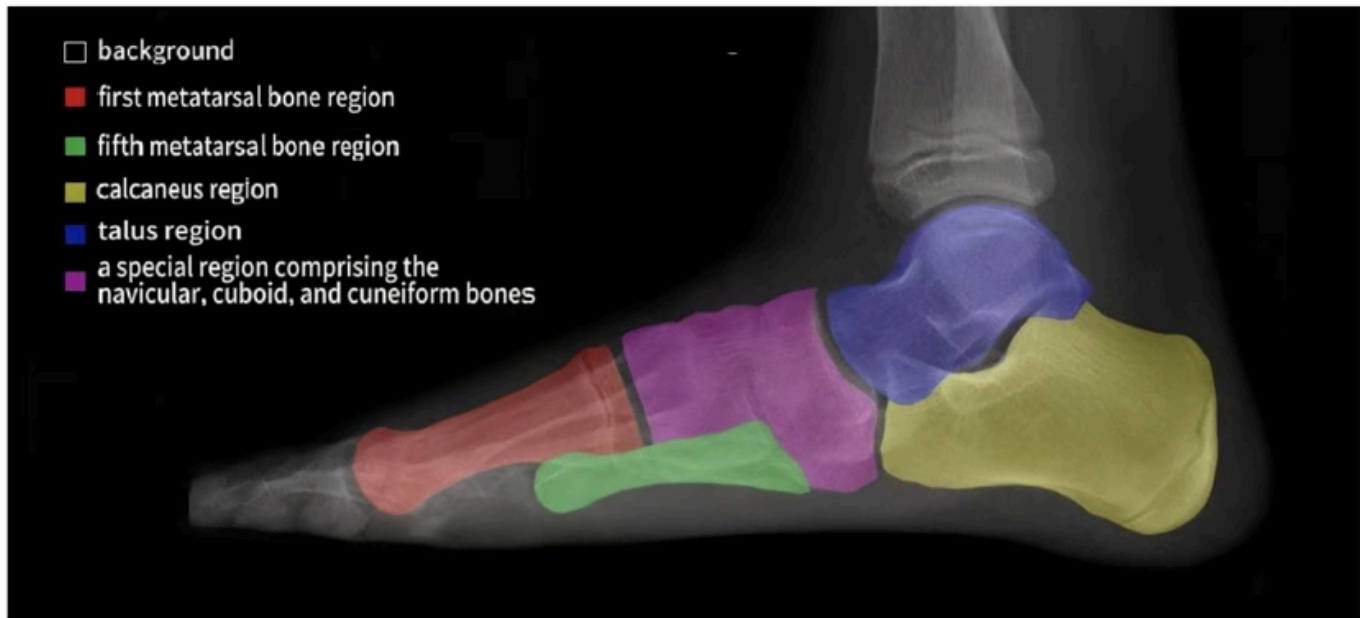


Figure 2

The entire foot was segmented into five distinct regions: the talus region, calcaneus region, first metatarsal bone region, fifth metatarsal bone region, and a special region comprising the navicular, cuboid, and cuneiform bones.

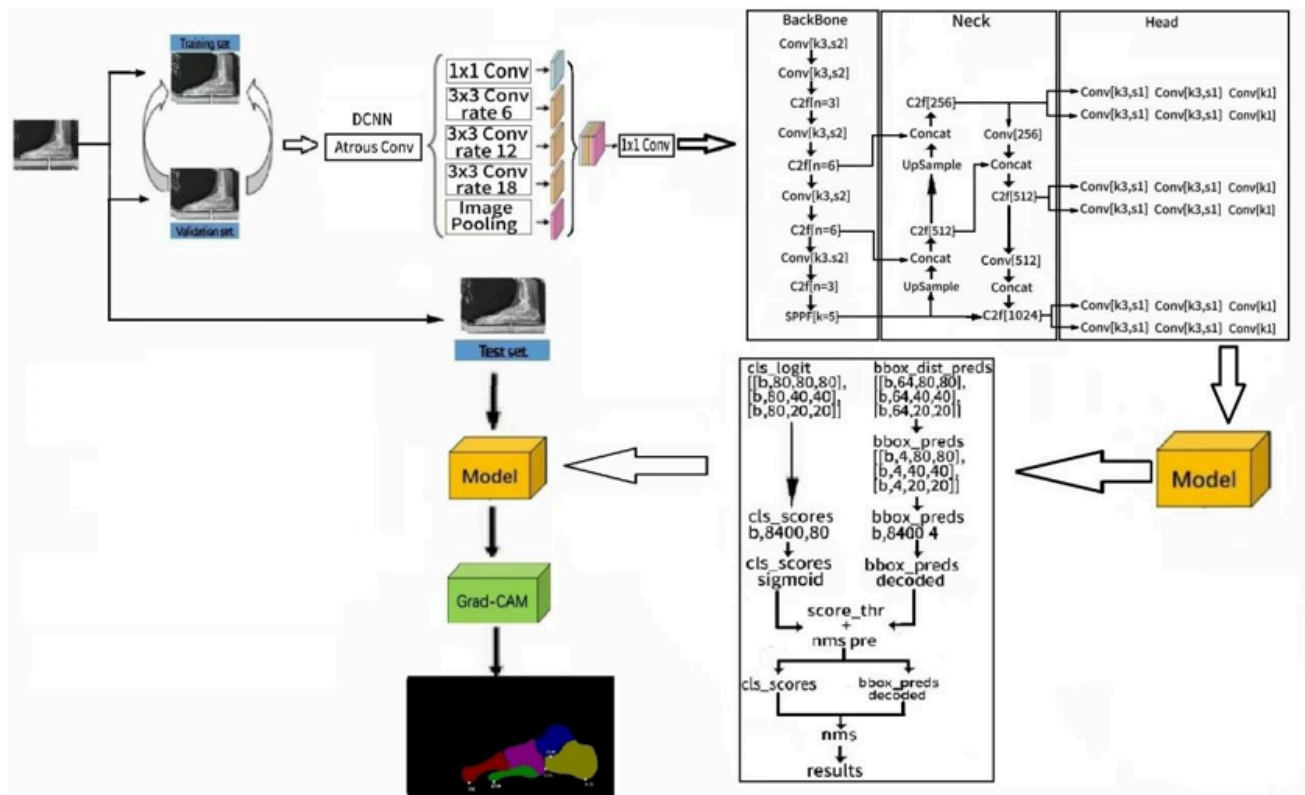


Figure 3

Overview of the automated foot arch measurement system

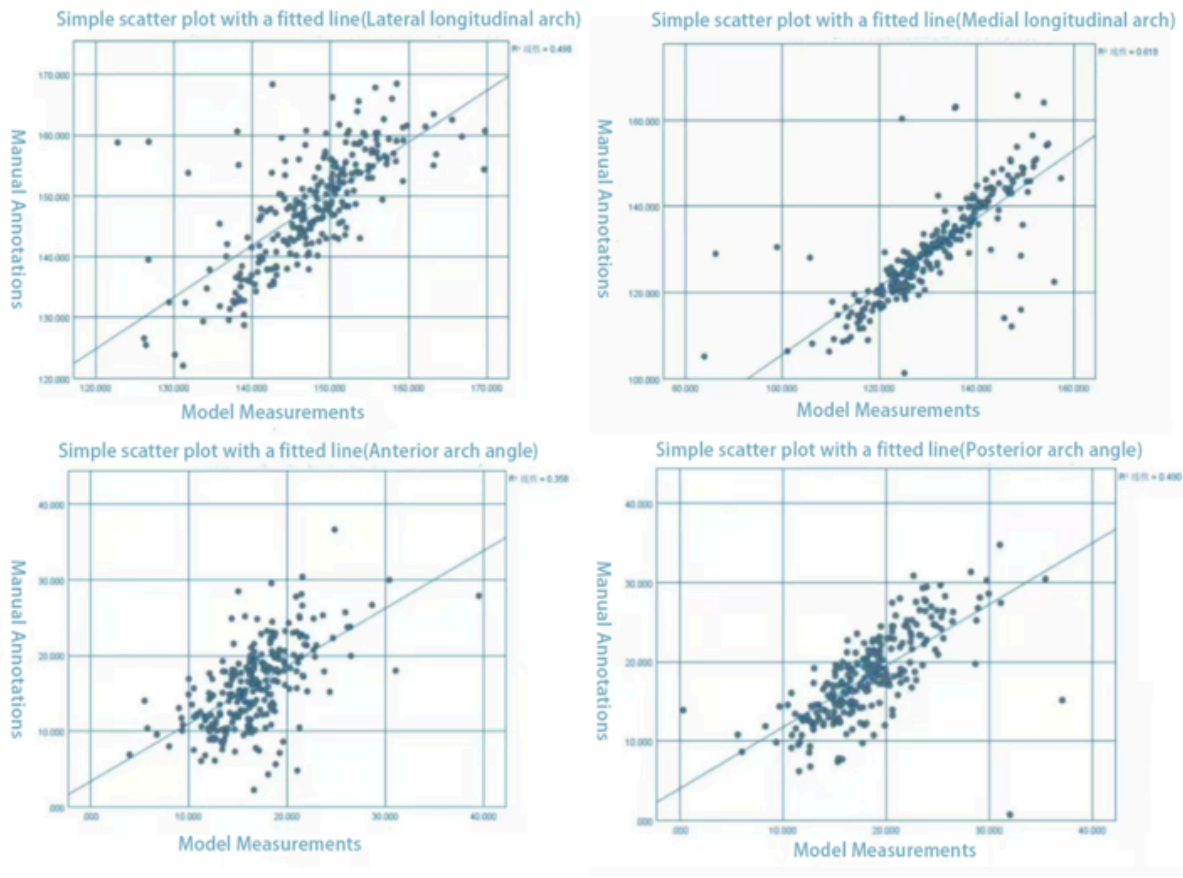


Figure 4

Linear regression equations were fitted for each of the four foot arch angles as follows:

LLA: $y = 0.851x + 22.672$; MLA: $y = 0.789x + 26.726$; AA: $y = 0.764x + 3.345$; PA: $y = 0.976x + 0.677$

(x represents the model's preliminary prediction, and y represents the calibrated prediction/target manual measurement).

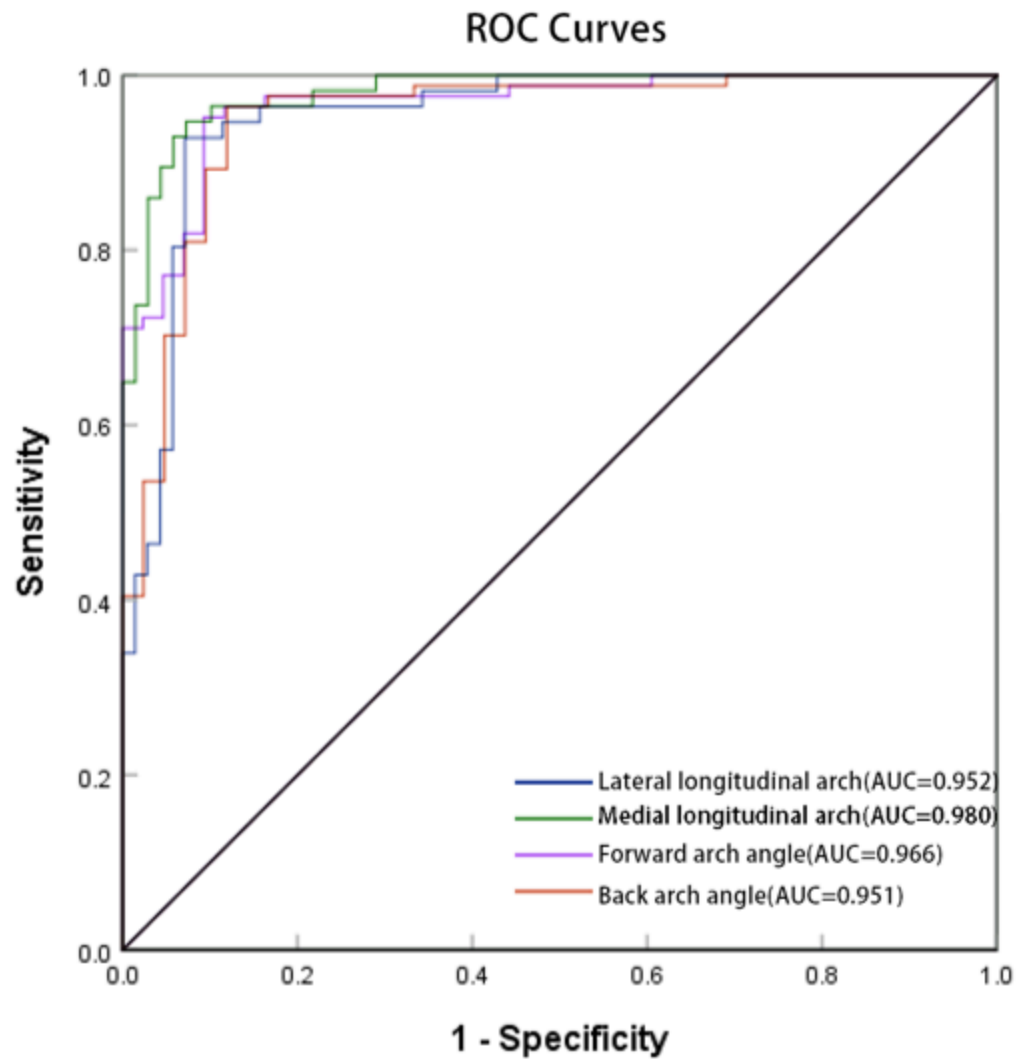


Figure 5

ROC curves for the four foot arch angles.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Tables.docx](#)