

Article

Automatic Lower-Limb Length Measurement Network (A3LMNet): A Hybrid Framework for Automated Lower-Limb Length Measurement in Orthopedic Diagnostics

Se-Yeol Rhyou ^{1,2,†} , Yongjin Cho ^{3,*} , Jaechern Yoo ^{1,†}, Sanghoon Hong ², Sunghoon Bae ², Hyunjae Bae ² and Minyung Yu ¹

¹ Department of Electrical and Computer Engineering, College of Information and Communication Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea; fbtpdf@naver.com (S.-Y.R.)

² Cleverus Corp., Seoul 06771, Republic of Korea

³ Department of Orthopedic Surgery, Chosun University College of Medicine, Gwangju 61453, Republic of Korea

* Correspondence: choisidoru@chosun.ac.kr; Tel.: +82-62-220-3147; Fax: +82-62-226-3379

† These authors contributed equally to this work.

Abstract: Limb Length Discrepancy (LLD) is a common condition that can result in gait abnormalities, pain, and an increased risk of early degenerative osteoarthritis in the lower extremities. Epidemiological studies indicate that mild LLD, defined as a discrepancy of 10 mm or less, affects approximately 60–90% of the population. While more severe cases are less frequent, they are associated with secondary conditions such as low back pain, scoliosis, and osteoarthritis of the hip or knee. LLD not only impacts daily activities, but may also lead to long-term complications, making early detection and precise measurement essential. Current LLD measurement methods include physical examination and imaging techniques, with physical exams being simple and non-invasive but prone to operator-dependent errors. To address these limitations and reduce measurement errors, we have developed an AI-based automated lower-limb length measurement system. This method employs semantic segmentation to accurately identify the positions of the femur and tibia and extracts key anatomical landmarks, achieving a margin of error within 4 mm. By automating the measurement process, this system reduces the time and effort required for manual measurements, enabling clinicians to focus more on treatment and improving the overall quality of care.

Keywords: X-ray image; limb length discrepancy; computer-aided diagnosis; deep learning



Academic Editor: Reza K. Amineh

Received: 4 December 2024

Revised: 27 December 2024

Accepted: 30 December 2024

Published: 2 January 2025

Citation: Rhyou, S.-Y.; Cho, Y.; Yoo, J.;

Hong, S.; Bae, S.; Bae, H.; Yu, M.

Automatic Lower-Limb Length

Measurement Network (A3LMNet): A

Hybrid Framework for Automated

Lower-Limb Length Measurement in

Orthopedic Diagnostics. *Electronics*

2025, 14, 160. [https://doi.org/](https://doi.org/10.3390/electronics14010160)

10.3390/electronics14010160

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license

([https://creativecommons.org/](https://creativecommons.org/licenses/by/4.0/)

[licenses/by/4.0/](https://creativecommons.org/licenses/by/4.0/)).

1. Introduction

Limb Length Discrepancy (LLD) is a prevalent condition that can lead to compensatory gait abnormalities, pain, and an increased risk of early-onset degenerative osteoarthritis in the lower extremities [1,2]. Epidemiological studies suggest that mild LLD, defined as a difference of 10 mm or less, affects approximately 60–90% of the general population [3,4]. Although more severe cases are less common, they are clinically significant, as LLD is associated with secondary conditions such as low back pain, scoliosis, and hip or knee osteoarthritis, due to altered biomechanics [5]. This condition not only affects daily activities, but can also lead to long-term complications [6,7]. Therefore, early detection and accurate measurement are crucial for preventing these issues [8,9]. The etiology of LLD includes growth plate injuries during childhood, unilateral bone or soft tissue tumors, congenital conditions, and idiopathic cases where the cause remains unidentified [10,11]. Various

methods have been used to measure limb length, including both physical examination techniques and imaging modalities [12]. Physical exams, such as measuring distances from the umbilicus to the medial malleolus or from the anterior superior iliac spine (ASIS) to the lateral malleolus, offer simple, non-invasive approaches [13]. However, these methods are prone to significant measurement errors and are highly dependent on the skill of operator [14].

Given the limitations of manual methods in accurately assessing LLD, imaging modalities have been developed to address these shortcomings [15,16]. Upright weight-bearing full-length radiography, CT scans, and MRI are among the most commonly used techniques for more accurate quantification of LLD [17,18]. These techniques differ in accuracy, reliability, and the conditions under which they assess limb lengths. Upright weight-bearing radiography is favored for its high reproducibility and relatively low cost, allowing clinicians to assess limb length discrepancies in functional, real-world conditions [19,20]. However, it has a measurement error margin of approximately 4 mm and may slightly magnify images [21]. In contrast, CT scans are more accurate but cannot assess LLD in a weight-bearing position, which is crucial for orthopedic interventions. Despite advancements in imaging technology, these methods remain time-consuming and require skilled interpretation, introducing variability and subjectivity into the process [22,23].

In recent years, artificial intelligence (AI) techniques have emerged as promising tools to enhance the precision and efficiency of medical imaging analysis [24,25]. Traditional imaging methods, though reliable, are prone to observer bias and are labor-intensive [26,27]. AI algorithms, particularly those based on deep learning, have demonstrated significant potential in automating tasks traditionally performed by human operators [28,29]. These algorithms reduce observer variability and accelerate diagnostic workflows [30,31]. AI can process large datasets, recognize subtle patterns, and provide rapid, consistent measurements, all of which are crucial for improving patient outcomes [32].

The objective of this study is to develop an AI module that reduces the time required to measure limb lengths using upright weight-bearing full-length radiography while maintaining high reliability and reproducibility [33]. This module will automatically identify anatomical landmarks and measure limb lengths from radiographs, minimizing human error. By streamlining the diagnostic process through automation, our AI module has the potential to revolutionize LLD assessments in clinical practice. It will be evaluated for its effectiveness and efficiency in clinical settings, ultimately improving patient care through more accurate and timelier LLD assessments. Furthermore, this innovation could lower healthcare costs, reduce radiation exposure, and provide broader access to high-quality diagnostic tools, especially in settings with limited radiological expertise. AI-driven solutions can lead to more personalized, data-driven musculoskeletal care in the future [34–38].

Several studies have explored various approaches for measuring limb length and assessing lower-limb alignment, focusing particularly on traditional imaging modalities and the application of machine learning techniques. One widely used method is upright weight-bearing full-length radiography, which provides functional and reproducible assessments of limb length discrepancies. However, this method is prone to observer variability and requires manual intervention, with potential measurement errors up to 4 mm, as reported in previous studies (Lee et al. [39], Guggenberger et al. [40]). Lee et al. compared upright radiography and supine CT, emphasizing that limb length and alignment assessment is more accurate under weight-bearing conditions. However, this study still relies on manual measurements and does not address potential errors arising during the measurement process.

In recent years, machine learning techniques have increasingly been integrated into medical imaging to reduce observer bias and enhance measurement accuracy. For instance,

Shen et al. [41] introduced a U-Net-based method that automatically segments the femur and tibia with high accuracy. However, this study was limited to specific datasets, and there are challenges in applying it to various pathological conditions or data from different hospitals. Additionally, the method still requires manual intervention for key point extraction, making it not fully automated.

Similarly, Lee et al. [42] and Zheng et al. [43] developed systems using deep learning to automatically measure limb length discrepancies. However, these studies primarily used non-weight-bearing images, which fail to fully reflect important clinical information. In particular, Lee et al. achieved success in measuring overall limb length discrepancies but struggled with accuracy in specific patient groups and failed to consistently extract anatomical landmarks. Zheng et al. also faced challenges with the generalizability of their models, limiting their applicability to diverse patient populations.

Our study, the Automatic Lower-Limb Length Measurement Network (A3LMNet), follows the same conditions for limb length measurement as previous studies, but leverages A3LMNet, which fully automates the entire process, including femur and tibia segmentation, precise extraction of anatomical landmarks, and limb length measurement. By doing so, we address the issues of manual intervention and data diversity limitations present in previous studies, providing consistent and fast results without the need for expert manual input. Furthermore, our method significantly improves accuracy and time efficiency compared to manual approaches, while maintaining a high level of reliability.

- (a) Segment Femur and Tibia: In this step, the femur and tibia are identified from the X-ray data provided by CSU hospital. Details are described in Section 2.3.1.
- (b) Key point extraction: To measure the lower-limb length, key points must be identified on the femur and tibia. In this process, key points are accurately extracted from the segmented femur and tibia. Details are provided in Section 2.3.2.
- (c) Lower-Limb Length Measurement: The final step involves calculating the Euclidean distance between the extracted key points to obtain the final result. More information can be found in Section 2.3.3.

We demonstrate the effectiveness of a hierarchical approach for femur and tibia segmentation, followed by robust key point extraction and accurate lower-limb length measurement. These techniques yield highly precise results, comparable to those achieved through manual measurements by medical professionals. The hierarchical process progressively refines the segmentation of the femur and tibia, minimizing errors, while the key point extraction method accurately identifies critical landmarks. Finally, our lower-limb length measurement method calculates distances between key points with remarkable accuracy, ensuring reliable inputs for clinical evaluation and decision-making.

2. Materials and Methods

2.1. Dataset Preparation

We utilized 5500 lower-limb raw X-ray images collected from the Department of Orthopedic Surgery at Chosun University (CSU) Hospital, obtained under data use agreements via Institutional Review Board (IRB) approval, with data gathered indiscriminately from individuals of all ages and genders. The size of the images were all different, but about 8000 by 3000. The X-ray equipment used was the digital X-ray machine (GC85A, Samsung Electronics, Suwon 16677, South Korea) and the imaging was consistently performed at a fixed distance of 200 cm. The raw data were provided in DICOM format and subsequently converted to PNG format using the DICOM viewer PACS. The lower-limb length measurements were conducted by experts of CSU hospital. As depicted in Figure 1b, the lower-limb

length was uniformly measured from the most superior point of the femoral head to the midpoint of the most inferior part of the tibia.

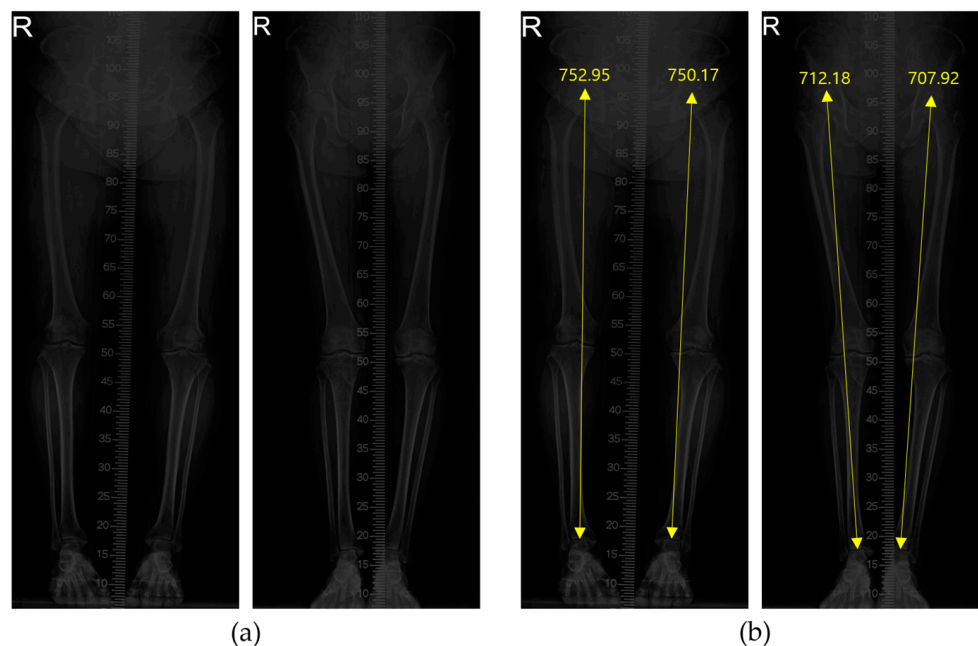


Figure 1. Lower-limb X-ray image collected by Chosun University Hospital. (a) Raw lower-limb X-ray images. (b) Manually measured lower-limb length by experts of CSU hospital.

The lower-limb X-ray dataset was randomly split into train, validation and test image, with a ratio 8:2:1 ratio, respectively, of 4000 train images, 1000 validation images and 500 test images.

2.2. Image Preprocessing

The raw data had an average size of 8000 by 2000 pixels. As the image size was excessively large for our purposes, we resized the images to 1024 by 256 pixels. Additionally, due to the inherently dark nature of raw X-ray images, it was essential to adjust the brightness. We applied histogram equalization to the images. This adjustment made the bone structures appear white and the background dark, enabling clear differentiation. However, when applying histogram equalization to the entire image, an issue arose where the background became black, leading to an overly bright overall image, as shown in Figure 2c. To address this, we implemented cut-off histogram equalization. Cut-off histogram equalization involves assigning a cut-off value and applying histogram equalization only to the range from n to 255, instead of the entire 0 to 255 range. After various experiments, we determined n to be 15 and applied cut-off histogram equalization, yielding the results shown in Figure 2.

2.3. Proposed A3LMNet

The flowchart of proposed A3LMNet is shown in Figure 3.

The below provides a succinct explanation of the steps in the A3LMNet process.

- i. Semantic segmentation (for detail, see Section 2.3.1): first, the preprocessed lower-limb X-ray images are semantically segmented into two classes, the femur and tibia.
- ii. Key point extraction (for detail, see Section 2.3.2): from the semantically segmented regions, we extract two key points necessary for the calculations.
- iii. Lower-limb length calculation (for detail, see Section 2.3.3): to determine the lower-limb length on each side, we calculate the Euclidean distance between the key points on the left and right sides.

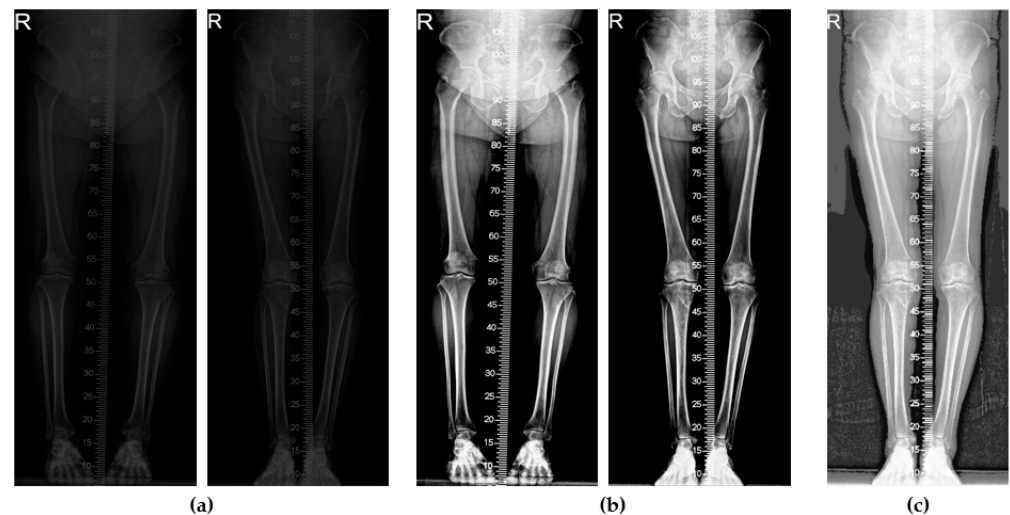


Figure 2. (a) Raw lower-limb X-ray images. (b) The result image of cut-off histogram equalization. (c) The result obtained after applying histogram equalization without cut-off.

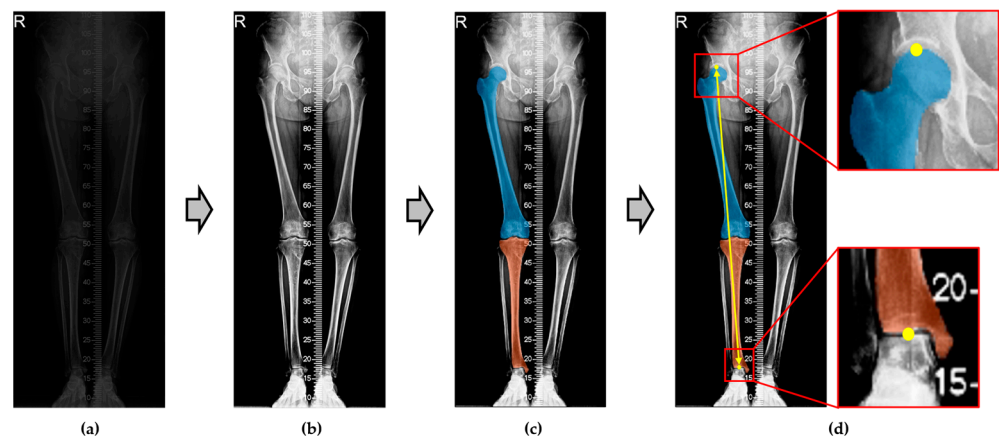


Figure 3. The flowchart of A3LMNet. (a) Raw lower-limb X-ray images. (b) Applying the cut-off histogram equalization. (c) The regions of the femur and tibia delineated by semantic segmentation. (d) Determining lower-limb length using the extracted key points.

The following sections detail the methods required for semantic segmentation, key point extraction, and lower-limb length calculation.

2.3.1. Femur and Tibia Segmentation

To accurately measure lower-limb length, it is essential to identify the precise locations of the femoral head's central uppermost point and the tibia's central lowermost point. We called these “key points”. Consequently, determining the exact positions of the femur and tibia in lower-limb X-ray images is paramount. In this phase, we focused on identifying the regions of the femur and tibia in preprocessed images to accurately locate these key points. To achieve this, we employed one of the deep learning techniques known as semantic segmentation, which has gained widespread use in medical imaging for pinpointing specific anatomical regions. Among the various methods available, we utilized DeepLabV3+ with ResNet 50 as the backbone, leveraging a pretrained network based on the ImageNet dataset, and fine-tuned it for our purposes. For training this network, we prepared a dataset consisting of 4000 training images and 1000 validation images. Each image in the dataset was meticulously annotated by experts from the CSU hospital, labeling the pixels into two classes: femur and tibia, as illustrated in Figure 4.

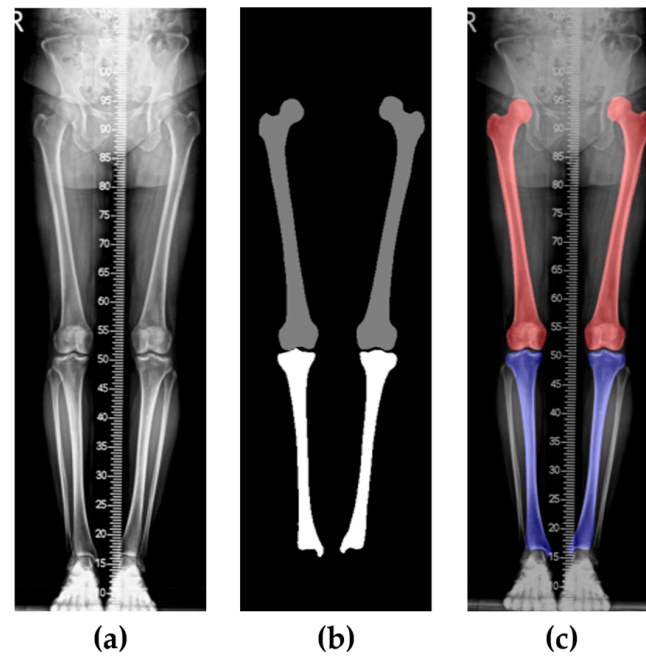


Figure 4. (a) Preprocessed lower-limb X-ray images. (b) An image visualizing pixel labels with the femur in gray and the tibia in white. (c) The overlapped image of (a,b).

Using these data, we trained the semantic segmentation model to accurately identify the regions of the femur and tibia. However, due to the inherent noise in X-ray images, errors can occur. To address this, we implemented Exception Handling after Inference. In a correctly performed segmentation, there should be a total of four regions, with one femur and one tibia on each side. However, due to the nature of semantic segmentation, pixel errors can occur. Therefore, we conducted a post-processing procedure in five steps, to accurately determine the regions of the femur and tibia, as outlined below.

- i. Select the two largest femur and tibia regions:
Let S be the segmented image, where $S(x, y) = 1$ for femur and $S(x, y) = 2$ for tibia.
- ii. Identify connected regions:

$$\left\{ C_{S(x,y),i} \right\}_{i=1}^{n_j}, \quad \text{for } S(x,y) = 1, 2, \quad (1)$$

where $C_{1,i}$ represents connected components of femur and $C_{2,i}$ represents connected components of tibia with $i \in \{1, \dots, n_j\}$ denoting the index of each connected component and $j = 1, 2$ specifying femur ($j = 1$) or tibia ($j = 2$)

- iii. Select the two largest regions:

$$R_{S(x,y),1}, R_{S(x,y),2} = \operatorname{argmax} |C_{S(x,y),i}|, \quad i \in \{1, \dots, n\}, \quad i = 1, 2, \quad (2)$$

where $R_{1,1}$ and $R_{1,2}$ are the selected femur regions and $R_{2,1}$ and $R_{2,2}$ are the selected tibia regions.

- iv. Calculate the centroids of each femur and tibia regions:

$$\bar{x}_{R_{S(x,y),k}}, \bar{y}_{R_{S(x,y),k}} = \frac{1}{|R_{S(x,y),k}|} \sum_{(a,b) \in R_{S(x,y),k}} (a, b), \quad k = 1, 2, \quad (3)$$

where the selected regions $R_{S(x,y),1}$ and $R_{S(x,y),2}$.

v. Classify, based on centroid coordinates:

$$\text{centroid } x - \text{coordinates of } \begin{cases} \text{Femur}_R : \operatorname{argmax}(\bar{x}_{R,1,1}, \bar{x}_{R,1,2}) \\ \text{Femur}_L : \operatorname{argmin}(\bar{x}_{R,1,1}, \bar{x}_{R,1,2}) \\ \text{Tibia}_R : \operatorname{argmax}(\bar{x}_{R,1,1}, \bar{x}_{R,1,2}) \\ \text{Tibia}_L : \operatorname{argmin}(\bar{x}_{R,1,1}, \bar{x}_{R,1,2}) \end{cases} \quad (4)$$

By applying this method, the centroid x-coordinates for the Femur_R , Femur_L , Tibia_R and Tibia_L regions are determined. The corresponding y-coordinates for $R_{S(x,y), k}$, where $k = 1, 2$ are the centroid y-coordinates of each respective region. The $R_{S(x,y), k}$ values of the obtained coordinates define each respective region. It is important to note that only the x-coordinates of the centroids are used for classification because the separation of left (L) and right (R) regions is primarily along the horizontal axis. The x-coordinate alone provides sufficient information to distinguish between left and right sides, while the y-coordinate, representing vertical positioning, does not contribute to this differentiation. This choice simplifies the classification process without compromising its accuracy.

Using the aforementioned method, Femur_R , Femur_L , Tibia_R and Tibia_L are identified, and regions outside these areas are disregarded. This ensures proper functionality, even if errors, as shown in Figure 5, occur.

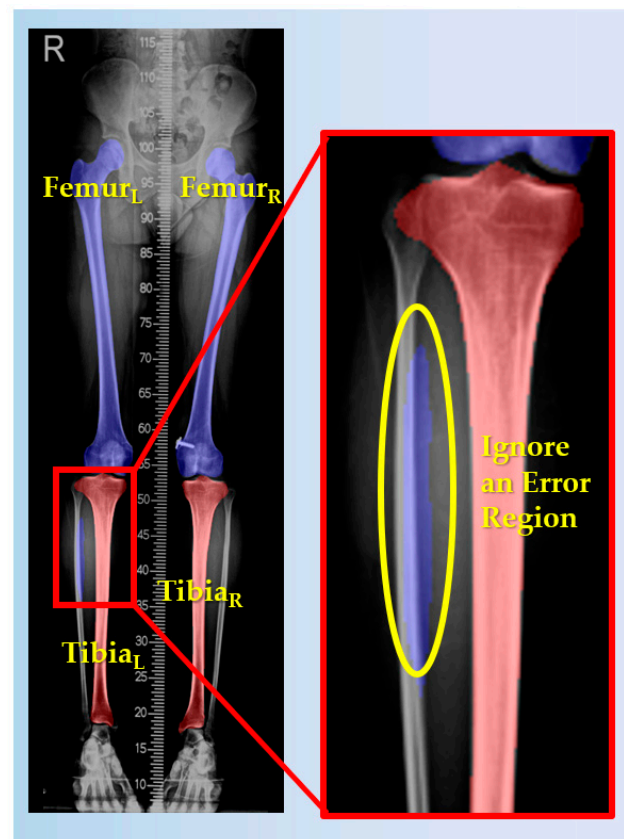


Figure 5. An error caused a femur region to appear next to the tibia, but the error region was ignored through exception handling.

2.3.2. Key Point Extraction for Determining Lower-Limb Length

Upon completing the segmentation, we can extract the key points necessary for measuring lower-limb length. A total of four key points are required: the central uppermost points of the femoral heads on the left and right femurs, and the central lowermost points

of the tibias on the left and right sides. By determining these four points, we can calculate the lower-limb length by measuring the distances between the corresponding points on each side.

- i. Left- and right-femur key points: the key points of the femur can be determined relatively easily. We need to identify the femoral head of the femur, which can be found at the highest y-coordinate point of $Femur_R$ and $Femur_L$.
- ii. Left- and right-tibia key points: identifying the key points of the tibia requires a considerably more complex process. Simply locating the lowest point may yield incorrect coordinates, due to the presence of the medial malleolus. To address this, we developed a specific method to accurately identify the key points, ensuring precise measurement of lower-limb length. This approach guarantees the determination of the optimal points for accurate measurement. Each step of the method is indicated in Figure 6a.
 - A. Find the lowest point of the tibia. (Purple points of Figure 6a).
 - B. Extract the tibia region by up to 10% of its total length above the lowest point. (Orange lines in Figure 6a).
 - C. Find the leftmost and rightmost points in the extracted $Tibia_R$ and $Tibia_L$. (Blue points of Figure 6a).
 - D. Calculate the midpoint of the found points. (Green points of Figure 6).
 - E. Find the lowest point in the tibia region corresponding to the x-coordinate of the midpoint. (Red points of Figure 6a).

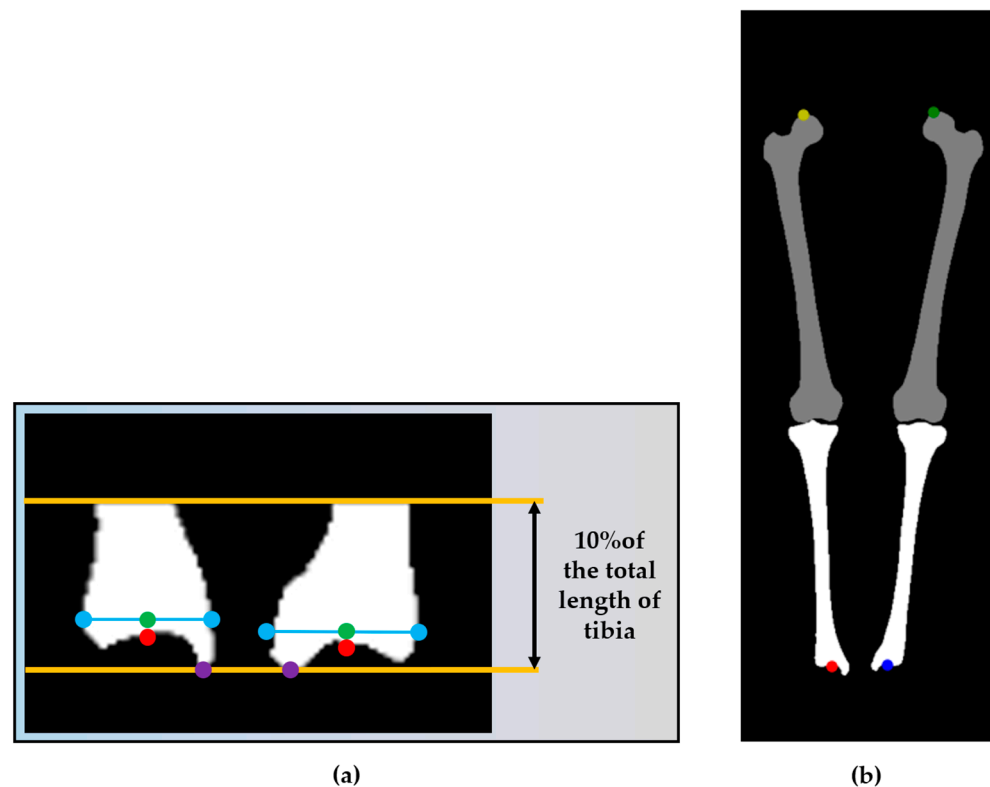


Figure 6. Key point extraction process and results. (a) Each point represents the outcome of the five steps, A to E, of the methods mentioned above. (b) Key points extracted for each femur and tibia.

In Figure 6b, the yellow dot represents the key point of the left femur and is designated as KP_{Femur_L} . Similarly, the green point is KP_{Femur_R} , the red point is KP_{Tibia_L} and the blue point is KP_{Tibia_R} .

2.3.3. Calculating the Euclidian Length Between Two Key Points

The final step involves measuring the lower-limb length by calculating the distances between KP_{Femur_L} and KP_{Tibia_L} , as well as between KP_{Femur_R} and KP_{Tibia_R} . This can be achieved by determining the Euclidean distance. However, this method measures the distance in pixels rather than the actual physical distance. To convert the pixel distance into the actual physical distance, the measured pixel value is divided by a ground-truth value measured by an expert, resulting in the scaling factor s . This scaling factor s is then used to convert pixel measurements to the real length of the lower limb, as indicated by the following equation.

$$\begin{aligned} Lower\ Limb\ Length_{Left} &= \frac{\sqrt{(x_{KP_{Tibia_L}} - x_{KP_{Femur_L}})^2 + (y_{KP_{Tibia_L}} - y_{KP_{Femur_L}})^2}}{s} \\ Lower\ Limb\ Length_{Right} &= \frac{\sqrt{(x_{KP_{Tibia_R}} - x_{KP_{Femur_R}})^2 + (y_{KP_{Tibia_R}} - y_{KP_{Femur_R}})^2}}{s}, \end{aligned} \quad (5)$$

where $x_{KP_{Femur_L}}$ and $y_{KP_{Femur_L}}$ are the x and y coordinate of KP_{Femur_L} and the others follow a same pattern.

3. Results

Our proposed “Automated Lower-Limb Length Measuring Network”, A3LMNet, was implemented using Python on a computer with a GeForce RTX 3090, GPU 24 GB. Lower-limb X-ray images were collected from the CSU hospital to verify the performance of the proposed model, where experts annotated the images and measured length. As demonstrated in the following experiments, our proposed algorithm accurately identified key points in lower-limb X-ray images, allowing for the precise measurement of lower-limb length. This advancement reduces the burden on experts who would otherwise need to measure lengths manually, thereby freeing up valuable time for discussing more effective treatment methods or conducting more in-depth research.

3.1. Performance of Femur and Tibia Segmentation

Given a preprocessed lower-limb X-ray image, semantic segmentation is a crucial step in A3LMNet for detecting key points on the femur and tibia. The image is segmented into two distinct regions: femur and tibia, and this segmentation is used to identify key points necessary for accurately measuring the length of the lower limb. For semantic segmentation, a pretrained model, DeepLab V3+ with a ResNet-50 backbone, was trained using transfer learning with the lower-limb X-ray images. In this study, cross-entropy loss was utilized to adjust model weights during neural network training. The quality of semantic segmentation was assessed using metrics including mean accuracy, mean Intersection over Union (IoU), and Boundary F-1 Score (BF1 Score). The performance outcomes for semantic segmentation, specifically pertaining to liver detection, are summarized in Table 1. Additionally, since deviations are not typically provided for in these metrics, we performed K-fold validation to calculate the deviation for Mean IoU separately, and included these results in Table 1.

Table 1. Performance of semantic segmentation for femur and tibia.

Class	Mean Accuracy (σ)	Mean IOU (σ)	BF1 Score
Femur	0.958 (0.018)	0.982 (0.011)	0.970
Tibia	0.963 (0.015)	0.984 (0.009)	0.973

The Mean IoU is the most commonly used evaluation metric for semantic segmentation tasks, enabling the evaluation of predicted pixel accuracy against the ground truth. Similarly, the BF1 score measures the similarity between the boundaries of segmented images and those of the ground truth. As demonstrated in Table 1, all three metrics show exceptionally high values, indicating that our preprocessing methods have had a significantly positive impact on the X-ray images. Despite the inherent noise present in medical images, our preprocessing steps have yielded excellent results. This has been crucial in accurately identifying the key points required for precise measurements.

3.2. Performance of the Key Point Extraction

Using semantic segmentation, we segmented the femur and tibia regions. From these segmented regions, we extracted the key points necessary for measuring lower-limb length. The key point on the femur was easily identified as the uppermost point of the femoral head. However, due to the presence of the medial malleolus, extracting the key point on the tibia required a more complex, five-step process. We evaluated the accuracy of the key points extracted through this process. The accuracy assessment involved calculating the pixel value differences between the coordinates of the identified key points, specifically KP_{Femur_L} , KP_{Femur_R} , KP_{Tibia_L} and KP_{Tibia_R} , and their corresponding ground-truth coordinates, as determined by the following formula.

$$\begin{aligned} \text{Deviation of } KP_{Femur_L} &= \sqrt{(x_{GTFemur_L} - x_{KPFemur_L})^2 + (y_{GTFemur_L} - y_{KPFemur_L})^2} \\ \text{Deviation of } KP_{Tibia_L} &= \sqrt{(x_{GTTibia_L} - x_{KPTibia_L})^2 + (y_{GTTibia_L} - y_{KPTibia_L})^2}, \end{aligned} \quad (6)$$

where $\text{Deviation of } KP_{Femur_L}$ and KP_{Tibia_L} follow a same pattern.

The deviation of key points was calculated to assess the accuracy of the identified key points. Due to the varying sizes of the original datasets, making precise comparisons was challenging. Therefore, we resized all images to 8000 by 2000 pixels for consistent testing. Table 2 presents the deviation results for key points across five test datasets. The key point deviation, with a maximum of 25.18 pixels, represents only about 0.3% of the diagonal length of the 8000 by 2000 pixel resized images provided by CSU. This minimal error indicates that the identified key points are highly accurate, supporting the reliability and clinical applicability of the lower-limb length measurements discussed in the following section.

Table 2. The ground-truth, measured-length, and deviation values for the lower-limb length measured by A3LMNet with 10 out of 1000 test data.

Test Data Number	Ground Truth (mm)		Measured Length (mm)		Deviation (mm)	
	Left	Right	Left	Right	Left	Right
1	793.2	787.2	792.94	787.43	0.26	0.23
2	828.6	827.5	830.56	826.74	1.96	0.76
3	790.6	784.5	790.09	783.66	0.51	0.84
4	889.6	886.0	892.12	883.52	2.52	2.48
5	815.9	817.2	817.05	816.43	1.15	0.77
6	852.4	860.2	849.73	858.08	2.67	2.12
7	743.9	725.9	743.59	724.90	0.31	1.00
8	756.9	781.7	755.65	783.79	1.25	2.09
9	766.9	763.6	766.93	766.54	0.03	2.94
10	691.5	694.7	693.27	694.83	1.77	0.13
Mean	-	-	-	-	1.57	1.45

As seen in the table above, the deviation of the key points does not exceed 30, indicating a high level of accuracy in key point extraction. This high accuracy suggests that the measurement of lower-limb length in the next phase will be highly reliable. Consequently, this method can be trusted and utilized by clinical experts.

3.3. Performance of Measuring the Lower-Limb Length

Using the highly reliable key points identified earlier, we measured the lower-limb length, which is our final objective. Table 3 presents the ground-truth distances, measured lengths, and deviations for the 10 test datasets.

Table 3. Key points of the Femur and Tibia: ground truth, extracted values, and errors.

Test Data Number	Ground Truth (x, y)		Extracted Key Points (x, y)		Deviation	
	<i>Femur_L</i> <i>Tibia_L</i>	<i>Femur_R</i> <i>Tibia_R</i>	<i>Femur_L</i> <i>Tibia_L</i>	<i>Femur_R</i> <i>Tibia_R</i>	<i>Femur_L</i> <i>Tibia_L</i>	<i>Femur_R</i> <i>Tibia_R</i>
1	(516, 6901)	(1460, 6854)	(523, 6899)	(1445, 6860)	7.28	16.15
	(841, 798)	(1207, 809)	(859, 813)	(1219, 805)	23.43	12.65
2	(619, 6952)	(1584, 6921)	(601, 6954)	(1570, 6922)	18.11	14.04
	(863, 917)	(1213, 917)	(875, 914)	(1195, 922)	12.37	18.68
3	(556, 6982)	(1474, 6922)	(531, 6985)	(1468, 6930)	25.18	10.00
	(519, 940)	(1307, 940)	(539, 953)	(1289, 953)	23.85	22.20
4	(546, 6659)	(1532, 6613)	(531, 6665)	(1523, 6625)	16.16	15.00
	(788, 1125)	(1181, 1125)	(797, 1125)	(1172, 1148)	9.00	24.70
5	(470, 7048)	(1488, 7110)	(453, 7047)	(1460, 7102)	17.03	29.12
	(920, 975)	(1266, 1005)	(930, 969)	(1258, 1000)	11.66	9.43

Since the values we obtained are in pixels, there is a discrepancy between the pixel measurements and the actual physical distances. To address this issue, we determined a scaling factor, denoted as s , which was found to be most appropriately set to 74.9 based on extensive data analysis. Therefore, we divided the measured pixel values by 74.9, resulting in the following adjusted results, seen in Table 2. The accuracy of the lower-limb length measurements was evaluated by comparing our results with ground-truth measurements provided by clinical experts. The average deviation from the ground truth was within the targeted 3 mm threshold, which is well within the acceptable range for clinical applications. This minor deviation indicates that our method provides precise measurements, crucial for accurate diagnosis and treatment planning.

4. Discussion and Conclusions

This study aimed to improve the accuracy and efficiency of LLD measurement using upright radiography, through the application of AI. The results showed that the AI module provided faster and more reliable measurements compared to manual methods, significantly reducing inter-observer variability. This is a crucial achievement in enhancing the consistency of LLD evaluations in clinical settings, easing the burden on clinicians while ensuring precise diagnoses. To achieve these results, we developed A3LMNet. In the first stage, semantic segmentation of the femur and tibia achieved a mean accuracy of 0.958, 0.963 and Mean IOU of 0.963, 0.984 respectively. The second stage, key point extraction, showed an error rate of up to 25.18 pixels, which corresponds to only 0.3% of the entire image, a remarkably small figure. Lastly, in the most important stage, lower-limb length measurement, the test showed a maximum error of 2.9 mm, much lower than the 4 mm error margin typical in manual measurements. Additionally, this method was not only

more accurate, but also approximately 90% faster than traditional methods. It takes approximately 6 s from loading the file to generating the results for the X-ray image. Compared to other studies, such as Moon et al. [18], our approach demonstrated a processing speed improvement of approximately 6.38 s (6 s vs. 12.38 s) and a reduction in average error margins for lower-limb length measurement. Specifically, our method achieved average errors of 1.57 mm for the left side and 1.45 mm for the right side, compared to 2.3 mm reported in Moon et al. These results highlight the superior efficiency and accuracy of our method in LLD measurement. However, it is important to note that the method proposed by Moon et al. is commendable for its ability to analyze not only limb length, but also additional parameters such as angles, providing a more comprehensive assessment. We respect their approach for its broader applicability and detailed evaluation framework.

Returning to our methodology, this error margin is clinically acceptable, suggesting that AI could be used as a practical diagnostic tool in radiographic analysis. It has the potential to save time for many experts, and to significantly contribute to improving the quality of medical care. The ability of the AI module to automatically identify anatomical landmarks while maintaining reproducibility underscores the importance of automation technology in the field of orthopedics. The improvement in accuracy and speed through the AI module could be particularly useful in busy clinical settings, enabling healthcare professionals to make faster and more reliable diagnoses. Especially in cases where repeated measurements are necessary for LLD diagnosis, automated tools can reduce observer fatigue and errors, ensuring consistency throughout the diagnostic process.

In conclusion, this study demonstrated that AI-based technology can significantly enhance the efficiency and accuracy of LLD measurements. The results of this study hold clinical significance in terms of reducing healthcare costs and minimizing radiation exposure, while also showing potential for use as a diagnostic tool in environments with limited radiological expertise. Further clinical research will be necessary to expand the role of AI and increase its applicability in musculoskeletal assessments. It is crucial to assess whether the performance of the AI module is consistent across different machines and clinical environments, not just limited to specific hospitals or devices. Additional studies are needed to train and test the AI model on a variety of radiographic machines and hospital settings, to ensure broader applicability.

However, this study has some limitations. We used X-ray images from CSU Hospital, and if X-rays from other hospitals have different distances between the source and the subject, accurate distance measurement may not be possible. To address this issue, we are working on an algorithm that can incorporate information about the X-ray distance as part of our future work. Additionally, this study focused solely on radiographic evaluation, and future research should aim to integrate functional assessments for a more comprehensive evaluation of LLD. Future advancements in deep learning techniques and the use of larger datasets for training AI models will allow for more reliable measurements, even in complex anatomical structures. Furthermore, combining other types of medical imaging, such as 3D CT and MRI, could enable AI to contribute not only to LLD evaluation, but also to treatment planning.

Author Contributions: Conceptualization, S.-Y.R. and Y.C.; data curation, S.-Y.R., Y.C. and J.Y.; formal analysis, S.-Y.R., Y.C. and J.Y.; funding acquisition, Y.C.; investigation, Y.C.; methodology, S.-Y.R., Y.C. and J.Y.; project administration, Y.C. and J.Y.; resources, Y.C. and J.Y.; software, S.-Y.R., J.Y., S.H., S.B., H.B. and M.Y.; supervision, Y.C. and J.Y.; validation, S.-Y.R., Y.C. and J.Y.; visualization, S.-Y.R., Y.C. and J.Y.; writing—original draft, S.-Y.R.; writing—review and editing, Y.C. and J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and ICT (MSIT, Korea) Balanced National Development Account, grant number ITAH0603230110010001000100100.

Institutional Review Board Statement: The local ethics committee granted ethical approval for this retrospective study and the ethics board waived written and informed consent due to the study's retrospective nature. The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of CHOSUN UNIVERSITY HOSPITAL (CHOSUN-2023-11-013 and 28 November 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data available on request due to restrictions.

Conflicts of Interest: Authors Se-Yeol Rhyou, Sanghoon Hong, Sunghoon Bae, and Hyunjae Bae were employed by the company Cleverus corp. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

1. Issei, M.; Mizuho, O.; Makoto, T. Effect of Leg Length Discrepancy on Dynamic Gait Stability. *Prog. Rehabil. Med.* **2023**, *8*, 20230013.
2. Sam, K.; Eli, C. Relationship and Significance of Gait Deviations Associated with Limb Length Discrepancy: A Systematic Review. *Gait Posture* **2017**, *57*, 115–123.
3. Martin, A.; Patrick, F.; Axel, K. Leg Length Discrepancy: A Systematic Review on the Validity and Reliability of Clinical Assessments. *PLoS ONE* **2021**, *16*, e0261457.
4. Guichet, J.; Spivak, F.; Trouilloud, P.; Grammont, P. Lower Limb-Length Discrepancy: An Epidemiologic Study. *Clin. Orthop. Relat. Res.* **1991**, *272*, 235–241. [[CrossRef](#)]
5. Sheha, E.; Steinhaus, M.; Kim, H.J.; Cunningham, M.; Fragomen, A.T.; Rozbruch, S.R. Leg-Length Discrepancy, Functional Scoliosis, and Low Back Pain. *JBJS Rev.* **2018**, *6*, e6. [[CrossRef](#)] [[PubMed](#)]
6. Mekki, K.; Davis, T.; Sakalian, P.; Pino, A.; Corces, A.; Roche, M. Leg Length Discrepancy Before Total Knee Arthroplasty Is Associated with Increased Complications and Earlier Time to Revision. *Arthroplasty* **2024**, *6*, 5. [[CrossRef](#)] [[PubMed](#)]
7. Starobrat, G.; Danielewicz, A.; Szponder, T.; Wojciak, M.; Sowa, I.; Róžańska-Boczula, M.; Latański, M. The Influence of Temporary Epiphysiodesis of the Proximal End of the Tibia on the Shape of the Knee Joint in Children Treated for Leg Length Discrepancy. *Clin. Med.* **2024**, *13*, 1458. [[CrossRef](#)] [[PubMed](#)]
8. Nazmy, H.; Solitro, G.; Domb, B.; Amirouche, F. Comparative Study of Alternative Methods for Measuring Leg Length Discrepancy after Robot-Assisted Total Hip Arthroplasty. *Bioengineering* **2024**, *11*, 853. [[CrossRef](#)] [[PubMed](#)]
9. Larson, N.; Nguyen, C.; Do, B.; Kaul, A.; Larson, A.; Wang, S.; Wang, E.; Bultman, E.; Stevens, K.; Pai, J.; et al. Artificial Intelligence System for Automatic Quantitative Analysis and Radiology Reporting of Leg Length Radiographs. *J. Digit. Imaging* **2022**, *35*, 1494–1505. [[CrossRef](#)]
10. Shailam, R.; Jaramilo, D.; Kan, J.H. Growth Arrest and Leg-Length Discrepancy. *Pediatr. Radiol.* **2013**, *43*, 155–165. [[CrossRef](#)] [[PubMed](#)]
11. Zakrzewski, A.; Jain, V. Etiology of Lower Limb Deformity. In *Pediatric Lower Limb Deformities*; Springer International Publishing: Cham, Switzerland, 2024; pp. 3–17.
12. Sabharwal, S.; Kumar, A. Methods for Assessing Leg Length Discrepancy. *Clin. Orthop. Relat. Res.* **2008**, *466*, 2910–2922. [[CrossRef](#)] [[PubMed](#)]
13. Khalifa, A. Leg Length Discrepancy: Assessment and Secondary Effects. *Orthop. Rheumatol.* **2017**, *6*, 1. [[CrossRef](#)]
14. Eichler, J. Methodological Errors in Documenting Leg Length and Leg Length Discrepancies. In *Leg Length Discrepancy: The Injured Knee*; Springer: Berlin/Heidelberg, Germany, 1977; pp. 29–39.
15. Birkenmaier, C.; Levrard, L.; Melcher, C.; Wegener, B.; Rieke, J.; Holzapfel, B.; Baur-Melnyk, A.; Mehrens, D. Distances and Angles in Standing Long-Leg Radiographs: Comparing Conventional Radiography, Digital Radiography, and EOS. *Skelet. Radiol.* **2024**, *53*, 1517–1528. [[CrossRef](#)] [[PubMed](#)]
16. Christopher, H.W.; Gerety, E.L. Leg Length Measurement: The Discrepancy and Beyond. *EPOS ECR 2019*. 2019. Available online: <https://epos.myesr.org/poster/esr/ecr2019/C-1654> (accessed on 1 December 2024).
17. Liodakis, E.; Kenawey, M.; Doxastaki, I.; Krettek, C.; Haasper, C.; Hankemeier, S. Upright MRI Measurement of Mechanical Axis and Frontal Plane Alignment as a New Technique: A Comparative Study with Weight Bearing Full Length Radiographs. *Skelet. Radiol.* **2011**, *40*, 885–889. [[CrossRef](#)] [[PubMed](#)]
18. Moon, K.R.; Lee, B.D.; Lee, M.S. A Deep Learning Approach for Fully Automated Measurements of Lower Extremity Alignment in Radiographic Images. *Sci. Rep.* **2023**, *13*, 14692. [[CrossRef](#)] [[PubMed](#)]

19. Rodríguez-Blanco, M.; Sánchez, G.L.; Calvo-Lobo, J.M.; Gómez, E.A.; Morales, P.V.M. Radiographic Assessment of Lower-Limb Discrepancy. *J. Am. Podiatr. Med. Assoc.* **2017**, *107*, 393–398.
20. Chua, C.; Tan, S.; Lim, A.; Hui, J. EOS Low-Dose Radiography: A Reliable and Accurate Upright Assessment of Lower-Limb Lengths. *Arch. Orthop. Trauma Surg.* **2022**, *142*, 735–745. [[CrossRef](#)]
21. Park, K.R.; Lee, J.H.; Kim, D.S.; Ryu, H.; Kim, J.H.; Yon, C.J.; Lee, S.W. The Comparison of Lower Extremity Length and Angle between Computed Radiography-Based Teleoroentgenogram and EOS[®] Imaging System. *Diagnostics* **2022**, *12*, 1052. [[CrossRef](#)] [[PubMed](#)]
22. Bhati, D.; Neha, F.; Amiruzzaman, M. A Survey on Explainable Artificial Intelligence (XAI) Techniques for Visualizing Deep Learning Models in Medical Imaging. *J. Imaging* **2024**, *10*, 239. [[CrossRef](#)] [[PubMed](#)]
23. Li, X.; Zhang, L.; Yang, J.; Teng, F. Role of Artificial Intelligence in Medical Image Analysis: A Review of Current Trends and Future Directions. *J. Med. Biol. Eng.* **2024**, *44*, 231–243. [[CrossRef](#)]
24. Mall, P.; Singh, P.; Srivastav, S.; Narayan, V.; Paprzycki, M.; Jaworska, T.; Ganzha, M. A Comprehensive Review of Deep Neural Networks for Medical Image Processing: Recent Developments and Future Opportunities. *Healthc. Anal.* **2023**, *4*, 100216. [[CrossRef](#)]
25. Pinto-Coelho, L. How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications. *Bioengineering* **2023**, *10*, 1435. [[CrossRef](#)] [[PubMed](#)]
26. Li, M.; Jiang, Y.; Zhang, Y.; Zhu, H. Medical Image Analysis Using Deep Learning Algorithms. *Front. Public Health* **2023**, *11*, 1273253. [[CrossRef](#)] [[PubMed](#)]
27. Salle, G.D.; Fanni, S.C.; Aghakhanyan, G.; Neri, E. Current Applications of AI in Medical Imaging. In *Introduction to Artificial Intelligence. Imaging Informatics for Healthcare Professionals*; Klontzas, M.E., Fanni, S.C., Neri, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2023; pp. 151–165.
28. Younas, F.; Usman, M.; Yan, W.Q. A Deep Ensemble Learning Method for Colorectal Polyp Classification with Optimized Network Parameters. *Appl. Intell.* **2023**, *53*, 2410–2433. [[CrossRef](#)]
29. Szilágyi, L.; Kovács, L. Special Issue: Artificial Intelligence Technology in Medical Image Analysis. *Appl. Sci.* **2024**, *14*, 2180. [[CrossRef](#)]
30. Flory, M.N.; Napel, S.; Tsai, E.B. Artificial Intelligence in Medical Imaging: Opportunities and Challenges. *Semin. Ultrasound CT MRI* **2024**, *45*, 152–160. [[CrossRef](#)]
31. Kübler, J.; Brendel, J.; Küstner, T.; Walterspiel, J.; Hagen, F.; Paul, J.; Nikolaou, K.; Gassenmaier, S.; Tsiflikas, I.; Burgstahler, C.; et al. Artificial Intelligence-Enhanced Detection of Subclinical Coronary Artery Disease in Athletes: Diagnostic Performance and Limitations. *Int. J. Cardiovasc. Imaging* **2024**, *40*, 2503–2511. [[CrossRef](#)] [[PubMed](#)]
32. Gala, D.; Behl, H.; Shah, M.; Makaryus, A. The Role of Artificial Intelligence in Improving Patient Outcomes and Future of Healthcare Delivery in Cardiology: A Narrative Review of the Literature. *Healthcare* **2024**, *12*, 481. [[CrossRef](#)] [[PubMed](#)]
33. Erne, F.; Grover, P.; Dreischarf, M.; Reumann, M.K.; Saul, D.; Histing, T.; Nüssler, A.K.; Springer, F.; Scholl, C. Automated Artificial Intelligence-Based Assessment of Lower Limb Alignment Validated on Weight-Bearing Pre- and Postoperative Full-Leg Radiographs. *Diagnostics* **2022**, *12*, 2679. [[CrossRef](#)]
34. Sun, T.; Wang, J.; Suo, M.; Liu, X.; Huang, H.; Zhang, J.; Zhang, W.; Li, Z. The Digital Twin: A Potential Solution for the Personalized Diagnosis and Treatment of Musculoskeletal System Diseases. *Bioengineering* **2023**, *10*, 627. [[CrossRef](#)]
35. Ali, A.; Omid, A.; Hamid, K.; Nathalie, B.; Massimo, S.; Filippo, M.; Rajendra, A. Interpretation of Artificial Intelligence Models in Healthcare. *J. Ultrasound Med.* **2024**, *43*, 1789–1818.
36. Tang, D.; Chen, J.; Ren, L.; Wang, X.; Li, D.; Zhang, H. Reviewing CAM-Based Deep Explainable Methods in Healthcare. *Appl. Sci.* **2024**, *14*, 4124. [[CrossRef](#)]
37. Obuchowicz, R.; Strzelecki, M.; Piórkowski, A. Clinical Applications of Artificial Intelligence in Medical Imaging and Image Processing—A Review. *Cancers* **2024**, *16*, 1870. [[CrossRef](#)] [[PubMed](#)]
38. Raju, V.; Sakshi, D.; Abhisek, V. Artificial Intelligence (AI): A Potential Game Changer in Regenerative Orthopedics—A Scoping Review. *Indian J. Orthopaedics* **2024**, *58*, 1362–1374.
39. Lee, S.J.; Lee, H.J.; Kim, J.I.; Oh, K.J. Measurement of the Weight-Bearing Standing Coronal and Sagittal Axial Alignment of Lower Extremity in Young Korean Adults. *J. Korean Orthop. Assoc.* **2011**, *46*, 191–199. [[CrossRef](#)]
40. Guggenberger, R.; Pfirrmann, C.; Koch, P.; Buck, R. Assessment of Lower Limb Length and Alignment by Biplanar Linear Radiography—Comparison with Supine CT and Upright Full-Length Radiography. *Am. J. Roentgenol.* **2014**, *202*, W161–W167. [[CrossRef](#)] [[PubMed](#)]
41. Shen, W.; Xiong, W.; Zhang, H.; Sun, Z.; Ma, J.; Ma, X.; Zhang, S.; Guo, S.; Wang, Y. Automatic Segmentation of the Femur and Tibia Bones from X-ray Images Based on Pure Dilated Residual U-Net. *Inverse Probl. Imaging* **2021**, *15*, 1333–1346. [[CrossRef](#)]

42. Lee, C.S.; Lee, M.S.; Byon, S.S.; Kim, S.H.; Lee, B.I.; Lee, B.D. Computer-Aided Automatic Measurement of Leg Length on Full Leg Radiographs. *Skelet. Radiol.* **2021**, *51*, 1007–1016. [[CrossRef](#)]
43. Zheng, Q.; Shellikeri, S.; Huang, H.; Hwang, M.; Sze, R.W. Deep Learning Measurement of Leg Length Discrepancy in Children Based on Radiographs. *Radiology* **2020**, *296*, 152–158. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.