EPiC Series in Health Sciences

Volume 7, 2024, Pages 1–6

Proceedings of The 24th Annual Meeting of the International Society for Computer Assisted Orthopaedic Surgery

Solutions for Lower Limb Malalignment: A Segmentation-Guided Coordinate Regression Approach for Landmark Detection and Automatic Measurement

Sebastian Amador Sanchez[1, 2], Philippe Vanoverschel[3], Julien Lebleu[4], Andries Pauwels[4], Ward Servaes[4], Wanne Wiersinga[4] and Jef Vandemeulebroucke[1, 2,5]

[1] Vrije Universiteit Brussel (VUB), Department of Electronics and Informatics (ETRO), Brussels, Belgium

[2] imec, Leuven, Belgium

[3] Hip and Knee Unit, Gent, Belgium

[4] moveUP, 1000 Bruseels, Belgium`

[5] Vrije Universiteit Brussel (VUB), Universitair Ziekenhuis Brussel (UZ Brussel), Department of Radiology, Brussels, Belgium

julien@moveup.care,

Abstract

Traditionally, surgeons measure lower limb deformities manually by assessing angles between axes drawn on full lower limb X-rays connecting specific landmarks. This process is considered cumbersome and subject to the surgeon's expertise. Our study aims to alleviate the manual detection of landmarks while enhancing the assessment of lower limb malalignment through an innovative approach that combines coordinate regression and landmark segmentation. While various deep learning solutions exist, our method differs by using landmark segmentation to indicate the possible location of the landmarks; this information is combined with the X-rays to estimate the position of the landmarks via coordinate regression. We named this deep learning architecture segmentation-guided regression.

To address the performance of our proposed approach, we evaluated the detection errors for eight landmarks and measured five malalignment metrics. We also compare our approach against landmark regression and landmark segmentation. While landmark segmentation achieved accurate landmark identification, it faced challenges in malalignment measurement due to incorrectly detected landmarks. On the other hand, regression had no failed detections but exhibited lower landmark detection accuracy. Our

segmentation-guided regression showed a balance, ensuring no incorrect landmark detections, improved landmark accuracy, and precise malalignment quantification.

By encouraging the coordinate regression network to focus on specific areas through segmentation guidance, our method positions landmarks more accurately and effectively measures malalignment. Consequently, our approach provides surgeons with a reliable tool for comprehensive lower limb malalignment assessment, combining the strengths of coordinate regression and landmark segmentation.

1 Introduction

Lower limb deformities are deviations of the physiological axes that translate to lower limb malalignment [1]. When deformities occur in the frontal plane, these are known as varus and valgus deviations [1]. Varus deformity increases the risk of progression of medial osteoarthritis (OA), and valgus has an equivalent effect on lateral OA [2]. To restore the limb's natural alignment and ensure long-term outcomes after placement of the prostheses, lower limb malalignment (LLM) is quantified pre-and post-operatively [3].

Traditionally, malalignment was quantified by manually drawing physiological axes with specific initial and ending landmarks over the lower limb X-rays and measuring the angles formed between these axes, as is shown in Figure 1. Such a process is cumbersome and is subject to the expertise of the physician performing the measurements. Computer-aided diagnosis systems have recently helped physicians outline the axes, but manual positioning still needs to be done [4]. To alleviate this burden, deep learning algorithms have been proposed to quantify the malalignment automatically [5]–[8].

These methods aim to automatically and accurately detect the landmarks required to draw the axes for an LLM test. They employ standard landmark detection techniques: coordinate regression [6], [7] or landmark segmentation [5], [8] to achieve such a goal. Compared to previous works, we propose a novel approach for landmark detection in full lower limb (FLL) X-rays that combines both methodologies: coordinate regression and landmark segmentation. We named our architecture segmentation-guided regression (SGR). The results show that our approach balances accuracy and robustness, making it suitable for clinical applications such as the LLM test.

2 Methods

SGR is a deep learning architecture based on segmentation and regression networks. The segmentation network takes a region of interest (ROI) as input and yields a probability map that informs which pixels of the image correspond to the landmark's coordinate position. Subsequently, the obtained probability map is combined with the ROI and input to a regression network to improve detection accuracy and robustness. Compared to earlier works where independently trained models were employed, SGR was trained end-to-end.

The assessment of SGR consisted of quantifying the detection error of eight different landmarks and measuring five malalignment metrics, as shown in Figure 1. To grasp the improvement of our end-to-end workflow, we contrasted SGR against independent landmark segmentation and coordinate regression. We used mean Euclidean distance as a metric to assess the performance of detecting landmarks. To evaluate the error between the values of angles using the ground truth landmarks and the estimated ones, we employed the mean absolute error (MAE).

A private dataset from the Hip and Knee Unit consisting of 919 FLL X-rays, including pre- and post-operative images, was employed for training and testing the networks. The radiologist annotated

the X-rays using the software V7 Darwin*, an online platform that allows the annotation of medical images. Labelling consisted of placing bounding boxes surrounding the desired joint to detect an ROI. Next, landmarks were annotated on each leg side, following the definitions of Figure 1. Due to commercial restrictions, supporting data and annotations are unavailable for sharing.

All the models were developed using Python v3 8.6. The deep learning library employed was Pytorch v1.7.1 and its sub-library Torchvision v0.8.2. The neural networks were trained using an Nvidia Tesla P100 GPU with 16GB of memory.

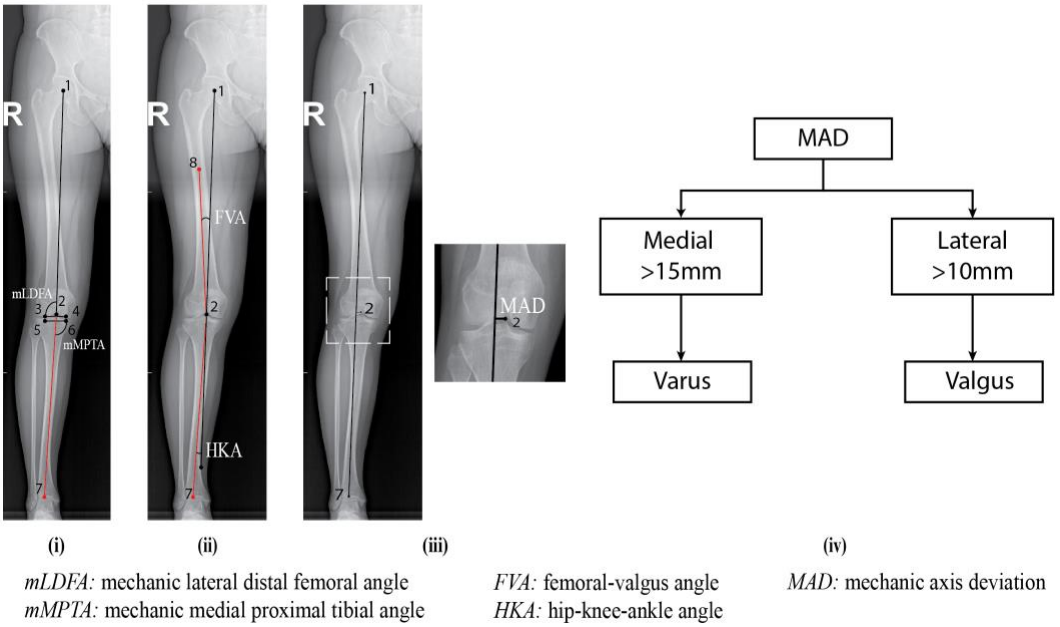

Figure 1: Axes (mechanical in black and anatomical in red) and angles measured in a malalignment test. **(i)** *mLDFA*: Angle measured between the axis that runs from the head of the femur centre to the centre of the knee (1→2) and the line that runs tangent to the femoral condyles (3→4). *mMPTA*: Angle formed between the axis that connects the centre of the knee and the centre of the ankle (2→7) and the tangent line that joins the tibial plateaus (5→6). **(ii)** *FVA*: Angle between the axis that connects the head of the femur centre and the centre of the knee (1→2) and the axis that runs from the middle of the diaphysis to the centre of the knee (8→2). HKA: Angle between the extended axis that joins the head of the femur centre and the centre of the knee (1→2), and the axis formed by connecting the centre of the knee and the ankle (2→7). **(iii)** *MAD*: Distance from the centre of the knee (2) to the axis that connects the head of the femur centre and the centre of the ankle (1→7). **(iv)** Evaluation of the deformity based on the MAD and the position of the axis (1→7) with respect to the centre of the knee. Varus and valgus deformities are based on the distance between the centre of the knee and the axis that connects the head of the femur to the centre of the ankle. Definitions taken from [1].

3 Results

Table I compares the three methods using Euclidean distance errors between ground truth and estimated landmarks. Landmark segmentation accurately detects hip and knee landmarks, while SGR performs better at the ankle joint. Overall, SGR outperforms pure regression with lower detection errors.

* https://www.v7labs.com/

With a 4 mm distance threshold for successful detection, SGR detects 83.83% of the landmarks below this boundary, landmark segmentation 86.57%, and pure regression achieves 79.28%.

To assess landmark detection's impact on lower limb malalignment measurements, an analysis checked if all the essential landmarks for drawing physiological axes on both legs were detected in X-rays. A successful analysis required detecting the complete set of landmarks; if a landmark was missing or an extra one was detected in any ROI, the analysis was deemed unsuccessful. Figure 2 contrasts the three methods, based on the previously described evaluation, with regard to the mean absolute error in each malalignment metric.

Figure 2 shows that landmark segmentation achieves malalignment errors equal to or below SGR and pure regression. However, landmark segmentation analyses only 86% of the images satisfactorily. In contrast, SGR and pure regression successfully analyse the complete set of X-rays at the cost of less accuracy. SGR has the lowest errors, surpassing coordinate regression across all five malalignment metrics.

Moreover, SGR correctly classifies all test X-rays as healthy or pathological based on the definition for identifying varus and valgus deformities, see Figure 1. When labelled as pathological, the deformity side and type (varus or valgus) are accurately classified, with precision, recall, and accuracy equal to 1.0. Since SGR can robustly and accurately detect all the required landmarks to measure lower limb malalignment, we compared it to what is currently reported in the literature. Table II compares SGR to prior literature, highlighting its quality compared to the current state-of-the-art. SGR outperforms Nguyen et al. [6] across all metrics, with lower HKA errors than other studies.

Method	*HF*	*R-FL*	*L-FL*	*R-TL*	*L-TL*	*CK*	*R-AL*	*L-AL*
Regression	2.03 (1.48)	3.39 (2.88)	2.95 (2.26)	3.71 (2.87)	3.40 (2.50)	1.81 (1.79)	2.59 (2.02)	2.83 (2.43)
Landmark Segmentation	**1.21 (1.06)**	**2.53 (2.26)**	**2.40 (1.56)**	**2.41 (2.88)**	**2.38 (2.17)**	**0.88 (2.61)**	2.52 (4.08)	2.79 (3.72)
SGR	1.55 (1.08)	3.03 (2.71)	2.49 (1.87)	3.00 (2.42)	3.30 (2.43)	1.34 (1.43)	**2.42 (1.93)**	**2.54 (2.24)**

Table I. Mean Euclidean (standard deviation) distance errors [mm] on each landmark for each addressed approach. The best metrics are represented using bold-faced numbers; the second-best metrics are represented via underlined numbers. **Notes***:* *HF*: head of the femur, *R-FL:* right femur landmark, *L-FL:* left femur landmark, *R-TL:* right tibia landmark, *L-TL:* left tibial landmark, *CK:* centre of the knee, *R-AL:* right ankle landmark, *L-AL:* left ankle landmark.

Model	*MAD[1] [mm]*	*mLDFA[1] [°]*	*mMPTA[1] [°]*	*FVA[1] [°]*	*HKA[1] [°]*	*HKA[2] [°]*
Ours	1.37 ± 1.63	0.73 ± 0.69	1.09 ± 1.02	0.19 ± 0.19	0.34 ± 0.43	-0.06 ± 0.55
Pei et al. [5]	-	-	-	-	-	-0.49 ± 0.75
*Tack et al. [7]**	-	-	-	-	-	0.09 ± 0.73
	-	-	-	-	-	0.18 ± 0.67
*Nguyen et al. [6]***	-	0.90 ± 0.80	1.15 ± 0.94	0.64 ± 0.50	0.67 ± 0.42	-
	-	1.14 ± 0.89	1.03 ± 0.67	1.30 ± 0.57	0.54 ± 0.49	-0.40 ± 0.68

Table II. Comparison of the segmentation-guided regression to what is reported in the literature. Malalignment metrics are represented in the form of error value ± standard deviation. **Notes:** [1] mean absolute error, [2] mean error, [*] two datasets were employed, [**] top corresponds to the right leg and bottom to the left leg measurements.

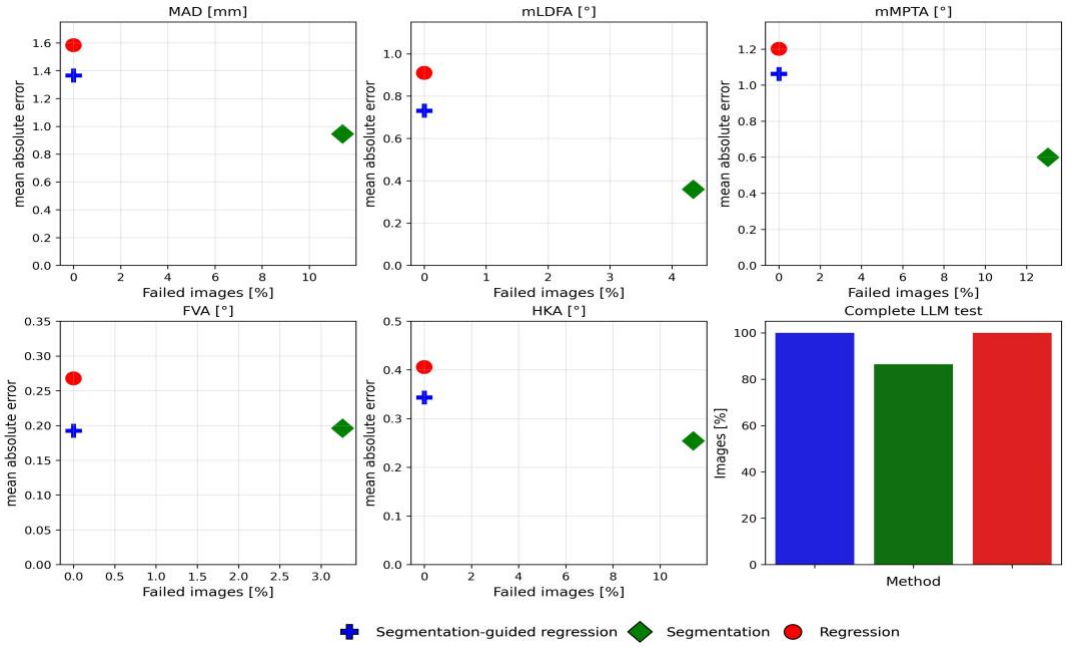

Figure 2: Robustness and accuracy evaluation of the three methods. On the y-axis, MAE between the values obtained using the estimated landmark position and the ground truth coordinates. On the x-axis, percentage of images where landmark detection failed. In addition, a bar plot indicating the percentage of correctly analysed images is given. Landmark segmentation failed to analyse 14% of the images, while SGR and coordinate regression analysed the complete 100% of them.

4 Discussion

We introduced segmentation-guided regression (SGR), a novel landmark detection method, and evaluated its performance in automatically measuring malalignment on full lower limb X-rays. Compared to landmark segmentation, SGR yields less precise detections and metrics. However, due to its design incorporating a coordinate regression network, our proposed method consistently detects the exact number of landmarks, ensuring reliable malalignment measurement. When comparing SGR to pure regression, our technique demonstrates more accurate metrics. This is attributed to our methodology incorporating a probability map before regression, providing information about potential landmark locations to the network, and improving final landmark estimations.

In contrast to other methods, our approach demonstrates similar or superior results in lower limb malalignment. SGR yields a balance between accuracy and robustness that is more valuable for clinical applications. Therefore, we view our approach as innovative, potentially enhancing surgical decision-making and optimising postoperative outcomes by providing surgeons with a comprehensive automated tool for assessing lower limb malalignment.

5 References

[1] N. M. Luís and R. Varatojo, "Radiological assessment of lower limb alignment," *EFORT Open Rev.*, vol. 6, no. 6, pp. 487–494, 2021, doi: 10.1302/2058-5241.6.210015.

[2] B. Zampogna *et al.*, "Assessing Lower Limb Alignment: Comparison of Standard Knee Xray vs Long Leg View," *Iowa Orthop. J.*, vol. 35, pp. 49–54, 2015.

[3] J. J. Cherian, B. H. Kapadia, S. Banerjee, J. J. Jauregui, K. Issa, and M. A. Mont, "Mechanical, anatomical, and kinematic axis in TKA: Concepts and practical applications," *Curr. Rev. Musculoskelet. Med.*, vol. 7, no. 2, pp. 89–95, 2014, doi: 10.1007/s12178-014-9218-y.

[4] N. Awang, R. Sulaiman, A. Shapi'I, A. H. A. Rashid, M. F. M. Amran, and S. Osman, "A comparative study of computer aided system preoperative planning for high tibial osteotomy," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9429, pp. 189–198, 2015, doi: 10.1007/978-3-319-25939-0_17.

[5] Y. Pei *et al.*, "Automated measurement of hip–knee–ankle angle on the unilateral lower limb X-rays using deep learning," *Phys. Eng. Sci. Med.*, vol. 44, no. 1, pp. 53–62, 2021, doi: 10.1007/s13246-020-00951-7.

[6] T. P. Nguyen, D. S. Chae, S. J. Park, K. Y. Kang, W. S. Lee, and J. Yoon, "Intelligent analysis of coronal alignment in lower limbs based on radiographic image with convolutional neural network," *Comput. Biol. Med.*, vol. 120, no. March, p. 103732, 2020, doi: 10.1016/j.compbiomed.2020.103732.

[7] A. Tack, B. Preim, and S. Zachow, "Fully automated Assessment of Knee Alignment from Full-Leg X-Rays employing a "YOLOv4 And Resnet Landmark regression Algorithm" (YARLA): Data from the Osteoarthritis Initiative," *Comput. Methods Programs Biomed.*, vol. 205, p. 106080, 2021, doi: 10.1016/j.cmpb.2021.106080.

[8] S. Simon *et al.*, "Fully automated deep learning for knee alignment assessment in lower extremity radiographs: a cross-sectional diagnostic study," *Skeletal Radiol.*, vol. 51, no. 6, pp. 1249–1259, 2022, doi: 10.1007/s00256-021-03948-9.