



Published in final edited form as:

Comput Methods Programs Biomed. 2021 July ; 206: 106120. doi:10.1016/j.cmpb.2021.106120.

Automated Analysis of Medial Gastrocnemius Muscle-Tendon Junction Displacements in Healthy Young Adults During Isolated Contractions and Walking Using Deep Neural Networks

Rebecca L. Krupenevich, Callum J. Funk, Jason R. Franz

Joint Department of Biomedical Engineering, UNC Chapel Hill & NC State University, Chapel Hill, NC, USA

Abstract

Background and objective: Direct measurement of muscle-tendon junction (MTJ) position is important for understanding dynamic tendon behavior and muscle-tendon interaction in healthy and pathological populations. Traditionally, obtaining MTJ position during functional activities is accomplished by manually tracking the position of the MTJ in cine B-mode ultrasound images – a laborious and time-consuming process. Recent advances in deep learning have facilitated the availability of user-friendly open-source software packages for automated tracking. However, these software packages were originally intended for animal pose estimation and have not been widely tested on ultrasound images. Therefore, the purpose of this paper was to evaluate the efficacy of deep neural networks to accurately track medial gastrocnemius MTJ positions in cine B-mode ultrasound images across tasks spanning controlled loading during isolated contractions to physiological loading during treadmill walking.

Methods: Cine B-mode ultrasound images of the medial gastrocnemius MTJ were collected from 15 subjects (6M/9F, 23 yr, 71.9 kg, 1.8 m) during treadmill walking at 1.25 m/s and during maximal voluntary isometric plantarflexor contractions (MVICs). Five deep neural networks were trained using 480 manually-labeled images, defined as the ground truth, collected during walking, and were then used to predict MTJ position in images from novel subjects 1) during walking (novel-subject), and 2) during MVICs (novel-condition).

Results: We found an average mean absolute error of 1.26 ± 1.30 mm and 2.61 ± 3.31 mm between the ground truth and predicted MTJ positions in the novel-subject and novel-condition evaluations, respectively.

Conclusions: Our results provide support for the use of open-source software for creating deep neural networks to reliably track MTJ positions in B-mode ultrasound images. We believe this

Author for correspondence: Rebecca Krupenevich, rkrup@email.unc.edu, 10304 Mary Ellen Jones Building, Chapel Hill, NC 27599, 919-445-2331.

Declaration of Competing Interest

The authors have no conflicts of interest to disclose

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

approach to MTJ position tracking is an accessible and time-saving solution, with broad applications for many fields, such as rehabilitation or clinical diagnostics.

Keywords

ultrasound; deep learning; plantarflexor; ankle; gait

1. Introduction

Ultrasound imaging is an increasingly popular tool for quantifying *in vivo* muscle-tendon dynamics in humans, with fidelity to do so during functional activities such as walking and running (1, 2). Tracking distinct anatomical landmarks in recorded ultrasound images - namely, the muscle-tendon junction (MTJ) - allows researchers the ability to differentiate muscle from tendon behavior with applications in clinical gait analysis (3), rehabilitation from injury (4, 5), and efficacy of exercise interventions (6). Traditionally, analyzing cine B-mode ultrasound images requires the researcher to manually label landmarks of interest within each frame in a video sequence. This process is time consuming and labor intensive (e.g., a single 2 s video recorded at 70 frames/s would require labeling 140 frames) and is thereby not practical for larger datasets across multiple cohorts with many experimental manipulations or conditions. In recent years, several semi- or fully-automated tracking tools have been developed for the purpose of tracking and quantifying muscle fascicle length and orientation (7-9), but these approaches are generally ineffective at tracking MTJ displacement. Some semi-automated and fully-automated methods for tracking MTJ displacement have been previously developed using optical flow (10, 11) or block-matching approaches (12). However, these methods work best when MTJ displacement between frames is relatively small, and may therefore be unreliable during activities involving large changes in MTJ position over short periods of time, such as walking or maximal voluntary contractions (MVCs). Additionally, these methods often require manual user input to correct tracking errors, which may unintentionally introduce measurement error and/or bias. Thus, there is a need for alternative solutions to accurately and efficiently track the MTJ during a wide range of functional activities.

Deep learning approaches to image processing have been successful in medical image analysis (13) and may be a promising technique for automated MTJ tracking. Recent advancements in deep neural networks (DNNs), such as the use of convolutional layers, have improved performance across a wide range of feature detection tasks (14, 15). DNN-based architectures consist of multiple hidden layers (i.e. deep layers) that 'learn' common features of the training images, such as shading or shapes, and then use that learned knowledge to recognize those features in future images(16). The addition of deep convolutional and deconvolutional layers allows DNNs to detect learned features anywhere in the image, which makes this approach particularly well-suited for analysis of medical images (e.g., ultrasound). Leitner et al. (2020) recently demonstrated that deep learning with a convolutional neural network can successfully track the MTJ during isolated maximal contractions and passive rotation (17). However, to our knowledge, these methods have not been evaluated on ultrasound images recorded during more functional activities with physiological loading such as walking. Further, it is unknown the extent to which a DNN

that was trained on images from one movement type can be generalized to tracking a novel movement.

Widespread adoption of deep-learning methods for image processing in the biomechanics community has been limited by the availability of accessible solutions for researchers without extensive experience in deep learning. User-friendly open-source software packages, such as DeepLabCut (18, 19), have removed this barrier to deep learning methods but their efficacy has not yet been demonstrated for tracking MTJ displacements in cine B-mode ultrasound images. Therefore, the purpose of this paper was to evaluate the efficacy of deep neural networks constructed with DeepLabCut (18, 19) at accurately tracking medial gastrocnemius MTJ positions in cine B-mode ultrasound images across tasks spanning controlled loading during isolated contractions to physiological loading during treadmill walking. We aim to benchmark the performance of these analytical techniques against conventional manual tracking approaches. Pretrained models presented in this paper are publicly available at <https://github.com/rlkrup/MTJtrack>

2. Methods

2.1 Subjects

15 healthy young adult subjects (6M/9F, 23 yr, 71.9 kg, 1.8 m) participated in this study. Prior to participation, subjects were screened and excluded if they reported injury or fracture to the lower-extremity within the previous six months, neurological disorders affecting the lower-extremity, or were currently taking medications that cause dizziness. All subjects provided written informed consent according to the University of North Carolina Biomedical Sciences Institutional Review Board.

2.2 Data Collection

Ultrasound images of the gastrocnemius MTJ used in this study were collected as part of a larger experiment that took place over the course of two sessions. For both sessions, prior to data collection, subjects pre-conditioned their triceps surae by walking for six minutes at 1.25 m/s on an instrumented treadmill (Bertec, Columbus, Ohio) (20). During the first session, cine B-mode ultrasound video was recorded while subjects performed a maximal voluntary isometric plantarflexor contraction (MVIC) (Biodex Medical Systems, Shirley, NY). During the second session, a 10-s ultrasound video was recorded while subjects walked for two minutes at 1.25 m/s. Ultrasound images were recorded using A 10 MHz, 60 mm linear array ultrasound transducer (LV7.5/60/128Z-2, Teleded Echo Blaster 128, Lithuania) operating at 38-68 frames/s. Three subjects were collected at 38 frames/s due to an unanticipated change to the settings file, the remaining 12 subjects were collected with frame rates between 60-76 frames/s – the later range of frame rates reflects differences in window sizes (e.g 80% vs 100% window) between subjects.

2.3 Data Processing

From the isolated contraction data, we trimmed the video so that it begins with the subject in a relaxed state and ends when the subject reaches peak torque (corresponding to greatest MTJ displacement). From the walking data, two consecutive full strides of walking were

identified and isolated using GRF data to determine the timing of gait events. One investigator manually labeled the gastrocnemius MTJ in each frame of each video, the results of which we present as our ground truth measurement. The position of the MTJ was identified as the most distal insertion of the muscle into the free tendon (Fig. 1A). To account for between-subject differences in stance time and framerates, 80 frames were randomly chosen from each subject (1200 frames total) for use in training and evaluation of the networks. The full subject set of walking videos (n=15) was then randomly split into five different combinations of training set (n=12) and test set (n=3), such that each subject was represented in the test set in one of the five combinations (Fig. 2). The MVIC videos were set aside for use in a later testing stage (novel condition testing), as described below, and were not used to train the networks.

2.4 Network

We used DeepLabCut (Version 2.1.6.4) (18) to construct and train our networks. Specifically, each network used 960 labeled images taken from 12 subjects. Networks were based on MobileNetV2-1.0, pretrained on ImageNet (21) with default parameters for 18,000 iterations with a batch size of 20. We selected our frame number parameter by systematically varying the size of the training set to train 45 distinct networks (five each of 80, 60, 40, 20, 10, 5, 2, and 1 frame(s) per subject included in the training sets).

2.4 Evaluation

After training, each network was used to predict the position of the MTJ in the videos from each network's respective test subjects (Fig. 2). Each prediction is accompanied by a confidence score (0-1) which represents the probability that the MTJ is visible in the frame (18). To reduce noise from outlier data points, we applied a modified median filter to predictions with low confidence scores (<0.98). The default window size was seven frames, such that a low confidence prediction at frame i is replaced with the median of the set of predictions from frames $i-3$ to $i+3$. This window size was reduced at the beginning and end of videos to maintain symmetry around the target frame. For example, at frame $i=2$ and $i=3$, the window size is reduced to three (frames 1-3) and five (frames 1-5) frames respectively. We report both unfiltered and filtered data in the results.

Novel Subject Evaluation: The trained networks were used to track MTJ position during walking for the three respective test subjects corresponding to each network (Fig. 1B).

Novel Condition Evaluation: To further characterize generalizability, the networks were also used to track MTJ position during MVICs for the three respective test subjects corresponding to each network (Fig 1B).. One subject was excluded from this analysis due to data loss (n=14). Novel condition images were not used in any of the training sets.

Network performance was quantified using root mean square error (RMSE), mean absolute error (MAE), and the absolute value of MTJ excursion discrepancies. RMSE was calculated for the Euclidean distance between the predicted and ground truth MTJ positions. MAE was calculated as the Euclidean distance between the predicted and ground truth MTJ positions. RMSE and MAE were calculated for each individual subject. We also report overall RMSE, which represents the RMSE of all subjects, and average MAE, calculated as the group

average MAE. Finally, we report the percentage of valid or invalid predictions. A valid prediction was defined as being within a 5 mm radius of the ground truth (17).

2.5 Speed Benchmarking

We evaluated GPU inference speed on Nvidia Tesla V100-SXM2 16 GB and Nvidia GeForce GTX1080 with a batch size of 64. CPU inference speed was evaluated on a 2.5 GHz Intel Xeon (E5-2680 v3) processor with batch sizes of 64 and 8. Average inference speed (frames/s) is reported for the novel-subject evaluation.

3. Results

3.1 Training parameters

After a steep initial drop at ~3,000 training iterations, RMSE fell below 5 mm and remained stable as the number of iterations increased. RMSE was minimized at ~18,000 training iterations. (Fig. 3A). For training set size (i.e. total number of frames in training set), we found that RMSE and MAE decreased with a greater number of frames and were both relatively stable from 500 to 1000 frames. A training set size of 480 total frames (40 frames per subject) was sufficient to achieve an overall RMSE of less than 3 mm and an overall MAE less than 2 mm (Fig. 3B).

3.2 Novel-subject evaluation: medial gastrocnemius MTJ displacements during walking

Overall unfiltered RMSE in the novel-subject evaluation was 2.72 mm, with an average unfiltered MAE of 1.26 ± 1.30 mm (Table 1). The modified median filter reduced overall RMSE by 0.35 mm (2.37 mm) and MAE by 0.09 mm (1.17 ± 1.23 mm). 94 and 95% of unfiltered and filtered predicted novel-subject MTJ positions were classified as valid (i.e., 5 mm of ground truth) (Table 2). MAE for three subjects in the novel-subject evaluation was greater than two standard deviations from the mean (Fig. 3A). Removing these subjects resulted in a smaller overall RMSE (Unfiltered: 1.36 mm; Filtered: 1.11 mm) and a smaller average MAE (Unfiltered: 0.70 ± 0.49 mm; Filtered: 0.65 ± 0.35 mm).

3.3 Novel-condition Evaluation: medial gastrocnemius MTJ displacements during MVICs

Overall unfiltered RMSE in the novel-condition evaluation was 6.23 mm, with an average unfiltered MAE of 2.61 ± 3.31 mm (Table 1). The modified median filter reduced overall novel-condition RMSE by 1.18 mm (5.05 mm), and average MAE by 0.43 mm (2.18 ± 2.64 mm) (Table 1). 88 and 90% of unfiltered and filtered predicted MTJ positions were classified as valid (Table 2). MAE for three subjects was greater than two standard deviations from the mean (Fig. 3B). Removing these subjects resulted in smaller overall RMSE (Unfiltered: 2.48 mm; Filtered: 2.10 mm) and smaller average MAE (Unfiltered: 1.05 ± 0.83 mm; Filtered: 0.92 ± 0.60 mm).

3.4 Inference Speed

Inference speed was fastest on Nvidia GeForce GTX1080 (25.55 ± 1.39 frames/s), followed by Nvidia Tesla V100-SXM2 (9.10 ± 1.10 frames/s). CPU inference speed was considerably

slower than GPU inference speed and was slowest with a smaller batch size (64 batch: 0.90 ± 0.04 frames/s, 8 batch: 0.59 ± 0.01 frames/s).

3.5 MTJ Excursion

Average MTJ excursion was not different between manual and automatically labeled data for walking in the X direction (Unfiltered: 95% CI [-0.590 9.64], Filtered: 95% CI [-0.89 5.69]) and Y direction (Unfiltered: 95% CI [-1.58 6.19], Filtered: 95% CI [0.05 0.50]) or for MVC data in the X direction (Unfiltered: 95% CI [-10.20 0.83], Filtered: 95% CI [-12.44 -0.31]) or Y direction (Unfiltered: 95% CI [-9.00 2.60], Filtered: 95% CI [-15.04 0.80]).

4. Discussion

The purpose of this paper was to evaluate the efficacy of deep neural networks constructed with open-source software (18, 19) to accurately track medial gastrocnemius MTJ positions in cine B-mode ultrasound images collected during muscle actions spanning isolated contractions to walking. MTJ position data are important for understanding the interaction between the Achilles tendon and gastrocnemius muscle and for guiding clinical decision making. As one example, ultrasound imaging has been previously used to characterize muscle-tendon disruption during walking in children with cerebral palsy – and plays a key role in informing surgical interventions (22). Traditional analysis of muscle-tendon kinematics using *in vivo* ultrasound imaging requires a time-consuming and labor-intensive process of manual labeling from one frame to the next. Here, we show that deep neural networks, trained using a small subject pool typical of biomechanics research studies, are effective and efficient for automated tracking of MTJ positions across a diverse range of contraction types. Accordingly, based on our cumulative findings, we feel confident advocating for the use of these deep neural networks as a suitable alternative to manual tracking of MTJ position from *in vivo* ultrasound images.

Our novel-subject evaluation during walking (Unfiltered RMSE: 2.72 mm, Filtered RMSE: 2.37 mm) performed better than the optical flow method presented by Cenni et al. (2019), who reported an RMSE of 6.3 mm for automated tracking and 4.7 mm for automated tracking with manual adjustments using images recorded from healthy adults during treadmill walking. Additionally, our networks outperformed those introduced recently by Leitner et al. (2020), who used deep learning methods to train and track MTJ positions during isolated contractions (17). Leitner et al. (2020) report a novel-subject MAE of 2.55 mm, with 88% of frames classified as valid using a 5 mm tolerance radius, compared to our novel-subject MAE of 1.26 mm with 94% of frames classified as valid. We note three subjects with outlier MAE values in the novel-subject evaluation (Fig.3A). These outliers may be the result of our intentionally small subject pool to represent sample sizes common in biomechanics studies; our networks were not trained on an expansive catalog of MTJ characteristics, resulting in larger errors when presented with an uncharacteristic MTJ feature, such as narrow space between the deep and superficial aponeuroses (i.e. small medial gastrocnemius muscle width), or aponeuroses that don't fully converge (see Supplemental Figure 1).

Overall RMSE was higher (i.e., worse performance) in the novel-condition evaluation than in the novel-subject evaluation. The relatively worse performance in the novel-condition evaluation may be due to differences in tissue deformation (i.e. muscle shape change) in maximal (MVICs) vs submaximal (walking) contractions. Thus, since our networks were only trained using images collected during walking, they may have been less capable of identifying MTJ image characteristics during maximal contractions. We purposefully withheld MVC images from the training set in order to assess the generalizability of the training set in a dynamic condition other than walking. We found that even when presented with novel-condition images, our networks performed with similar RMSE compared to semi-automated tracking techniques (10), and similar RMSE and MAE compared to fully automated deep-learning techniques (17)

Filtering the data with a modified median filter resulted in moderate reductions in overall RMSE and MAE for both the novel-subject and novel-condition evaluations. The filter was only applied to individual frames in which the confidence score of the predicted MTJ position was less than was <0.98 . Subjects with MAE values greater than two standard deviations from the mean had a higher proportion of predictions with confidence scores <0.98 (Novel-subject evaluation: 28%, Novel-condition evaluation: 54%) compared to the remaining subject set (Novel-subject evaluation: 3%, Novel-condition evaluation: 13%). As such, when these subjects were removed, the difference between unfiltered and filtered values was relatively small in both the novel-subject and novel-condition evaluations. A modified median filter may be an ideal method for maintaining data integrity, while also smoothing erroneous data points.

To complement network performance metrics, we also now calculate the absolute difference between MTJ excursion using manually-labeled and automatically-labeled MTJ position data. Our excursion values are similar to literature values for both walking (10) and isolated contractions (23). Absolute longitudinal difference between manually and automatically tracked MTJ excursion (Unfiltered: 4.79 ± 9.10 mm, Filtered: 2.68 ± 5.80 mm) was comparable to Cenni et al. (2019) who reported a 4.9 ± 5.6 mm difference in MTJ excursion between manually and automatically labeled images.

We note a few limitations to this technique. First, ground truth positions are manually labeled and are therefore prone to human error. However, manual labeling is the current gold standard for identifying MTJ position in ultrasound images. In addition, the networks we describe here are only appropriate for tracking videos in which the MTJ is visible in all frames. Future studies should explore options to manage frames in which the MTJ is not visible, such as specifically training networks to infer position when the target is out of view (24). We also note that our networks were trained using images from healthy young adults, which may not generalize well to pathological populations due to differences in image features (e.g. echogenicity (25)). Investigating the clinical application of these networks is warranted. Our networks were based on MobileNetV2-1.0 (21), which is a shallower network (compared to ResNet network architectures) and may not be suitable for large images without downscaling. However, this network allows for faster training and analysis and can be run without a high-end GPU, which makes it ideal for researchers with limited computing resources. The fast processing speed of MobileNet also makes it ideal for future

studies that aim to track in real-time. Finally, except for increasing the batch size, we used default DeepLabCut settings in order for these methods to be easily replicated. However, fine-tuning the settings may result in faster or more accurate performance.

There are many potential future directions for automated ultrasound image processing in the biomechanics and rehabilitation fields. For example, automated tracking with a trained network does not require the level of skill and experience of manual MTJ tracking. Thus, publicly available trained networks may encourage the adoption of tracking techniques in other fields, such as clinical diagnostics. Additionally, further development of these networks may facilitate real-time MTJ tracking, which could be a useful tool in rehabilitation and sports performance settings. Direct measurement of MTJ position from cine B-mode ultrasound images is important for reliably estimating tendon behavior and muscle-tendon interaction during functional activities. Although it is possible to indirectly measure tendon behavior by subtracting muscle fascicle length from muscle-tendon unit length, this indirect measurement frequently yields implausible outcomes (26). Thus, tools that can quickly and accurately estimate MTJ position have the potential to not only improve data processing speed, but also accelerate scientific discovery. Our results provide support for the use of open-source software for creating deep neural networks to reliably track MTJ positions in B-mode ultrasound images. In addition to the models presented in this paper, we also provide a single model trained on all available data presented in this paper, available here: <https://github.com/rlkrup/MTJtrack>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by grants from the National Institutes of Health (F32AG067675 to RLK and R01AG058615 to JRF). We also thank the University of North Carolina at Chapel Hill and the Research Computing group for providing computational resources and support.

References

1. Zelik KE, Adamczyk PG. A unified perspective on ankle push-off in human walking. *J Exp Biol.* 2016;219(23):3676–83. [PubMed: 27903626]
2. Cronin NJ, Lichtwark G. The use of ultrasound to study muscle–tendon function in human posture and locomotion. *Gait Posture.* 2013;37(3):305–12. [PubMed: 22910172]
3. Zhao H, Ren Y, Wu Y-N, Liu SQ, Zhang L-Q. Ultrasonic evaluations of Achilles tendon mechanical properties poststroke. *J Apply Physiol.* 2009;106(3):843–9.
4. Silbernagel KG, Steele R, Manal K. Deficits in heel-rise height and Achilles tendon elongation occur in patients recovering from an Achilles tendon rupture. *Am J Sports Med.* 2012;40(7):1564–71. [PubMed: 22593092]
5. Rees J, Lichtwark G, Wolman R, Wilson A. The mechanism for efficacy of eccentric loading in Achilles tendon injury; an in vivo study in humans. *Rheumatology.* 2008;47(10):1493–7. [PubMed: 18647799]
6. Maeda N, Urabe Y, Tsutsumi S, Sakai S, Fujishita H, Kobayashi T, et al. The acute effects of static and cyclic stretching on muscle stiffness and hardness of medial gastrocnemius muscle. *J Sports Sci Med.* 2017;16(4):514. [PubMed: 29238251]

7. Drazan JF, Hullfish TJ, Baxter JR. An automatic fascicle tracking algorithm quantifying gastrocnemius architecture during maximal effort contractions. *PeerJ*. 2019;7:e7120. [PubMed: 31304054]
8. Seynnes OR, Cronin NJ. Simple Muscle Architecture Analysis (SMA): An ImageJ macro tool to automate measurements in B-mode ultrasound scans. *Plos one*. 2020;15(2):e0229034. [PubMed: 32049973]
9. Farris DJ, Lichtwark GA. UltraTrack: Software for semi-automated tracking of muscle fascicles in sequences of B-mode ultrasound images. *Comput Meth Prog Bio*. 2016;128:111–8. doi: 10.1016/j.cmpb.2016.02.016.
10. Cenni F, Bar-On L, Monari D, Schless SH, Kalkman BM, Aertbeliën E, et al. Semi-automatic methods for tracking the medial gastrocnemius muscle–tendon junction using ultrasound: a validation study. *Exp Physiol*. 2020;105(1):120–31. [PubMed: 31677311]
11. Zhou G-Q, Zhang Y, Wang R-L, Zhou P, Zheng Y-P, Tarassova O, et al. Automatic myotendinous junction tracking in ultrasound images with phase-based segmentation. *BioMed Res Int*. 2018;2018.
12. Lee SS, Lewis GS, Piazza SJ. An algorithm for automated analysis of ultrasound images to measure tendon excursion in vivo. *J Appl Biomech*. 2008;24(1):75–82. [PubMed: 18309186]
13. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: a review. *J Med Syst*. 2018;42(11):226. [PubMed: 30298337]
14. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.
15. Krizhevsky A, Sutskever I, Hinton GE, editors. *Imagenet classification with deep convolutional neural networks*. *Adv Neural Inform Process Syst*; 2012.
16. He K, Zhang X, Ren S, Sun J, editors. *Deep residual learning for image recognition*. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016.
17. Leitner C, Jarolim R, Konrad A, Kruse A, Tilp M, Schröttner J, et al. Automatic Tracking of the Muscle Tendon Junction in Healthy and Impaired Subjects using Deep Learning. arXiv preprint arXiv:200502071. 2020.
18. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat neuroscience*. 2018;21(9):1281–9. [PubMed: 30127430]
19. Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nat protocols*. 2019;14(7):2152–76. [PubMed: 31227823]
20. Hawkins D, Lum C, Gaydos D, Dunning R. Dynamic creep and pre-conditioning of the Achilles tendon in-vivo. *J Biomech*. 2009;42(16):2813–7. [PubMed: 19762028]
21. Mathis A, Yüsekçöğül M, Rogers B, Bethge M, Mathis MW. Pretraining boosts out-of-domain robustness for pose estimation. arXiv preprint arXiv:190911229. 2019.
22. Barber L, Carty C, Modenese L, Walsh J, Boyd R, Lichtwark G. Medial gastrocnemius and soleus muscle-tendon unit, fascicle, and tendon interaction during walking in children with cerebral palsy. *Dev Med Child Neurol* 2017;59(8):843–51. [PubMed: 28369824]
23. Arampatzis A, Stafilidis S, DeMonte G, Karamanidis K, Morey-Klapsing G, Brüggemann G. Strain and elongation of the human gastrocnemius tendon and aponeurosis during maximal plantarflexion effort. *J Biomech*. 2005;38(4):833–41. [PubMed: 15713305]
24. Zaghi-Lara R, Gea MÁ, Camí J, Martínez LM, Gomez-Marin A. Playing magic tricks to deep neural networks untangles human deception. arXiv preprint arXiv:190807446. 2019.
25. Pitcher CA, Elliott CM, Panizzolo FA, Valentine JP, Stannage K, Reid SL. Ultrasound characterization of medial gastrocnemius tissue composition in children with spastic cerebral palsy. *Muscle Nerve*. 2015;52(3):397–403. [PubMed: 25556656]
26. Zelik KE, Franz JR. It's positive to be negative: Achilles tendon work loops during human locomotion. *Plos one*. 2017;12(7):e0179976. [PubMed: 28671955]

Highlights

- Muscle-tendon junction (MTJ) position is important for understanding dynamic tendon behavior
- Obtaining MTJ position in cine B-mode ultrasound images is labor intensive
- Deep neural networks created with open-source software can reliably track MTJ position
- This approach is accessible and efficient, with applications in rehabilitation or clinical diagnostics

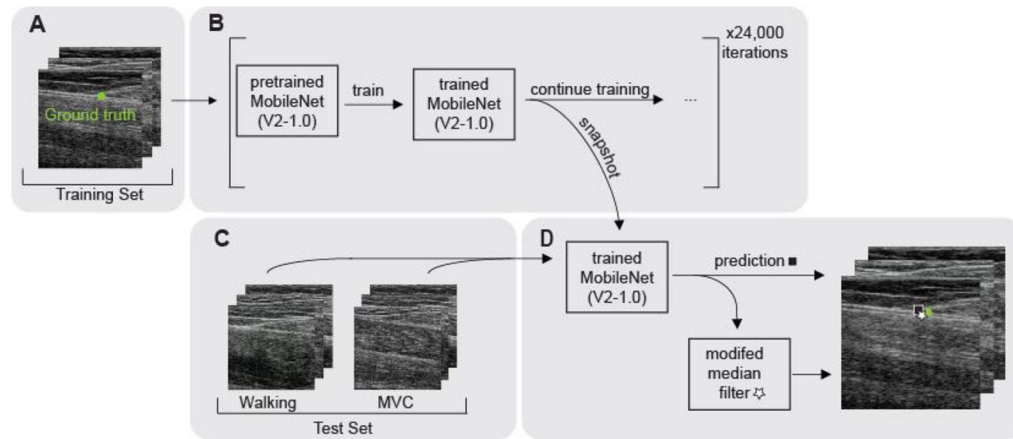


Figure 1.

A) MTJ position was manually labeled (ground truth) for each frame of B-mode ultrasound video collected during walking. Five training sets were created, consisting of 80 frames from 12 subjects. B) Networks were based on MobileNetV2-1.0, pretrained on ImageNet with default parameters, and then trained on our training sets for 24,000 iterations. During training, the networks were saved intermittently (i.e. a snapshot was taken) to allow for evaluating the effect of training time on performance. C) Each training set was associated with a test set consisting of three subjects that were not included in the respective test set (i.e., novel subjects). D) The trained networks were used to predict the location of the MTJ in each frame of the videos from the novel subjects in the test set. A modified median filter was applied to predictions with low confidence scores (<0.98) to reduce noise from outlier data points.



Figure 2.

Illustration of training/test set shuffles. Each box represents a subject; each row represents one network. White boxes represent subjects that were used for training the network and black boxes represent subjects that were used for testing the respective network. Test subjects were not used to train the respective network. The full subject set of walking videos (n=15) was randomly split into five different combinations of training set (n=12) and test set (n=3).

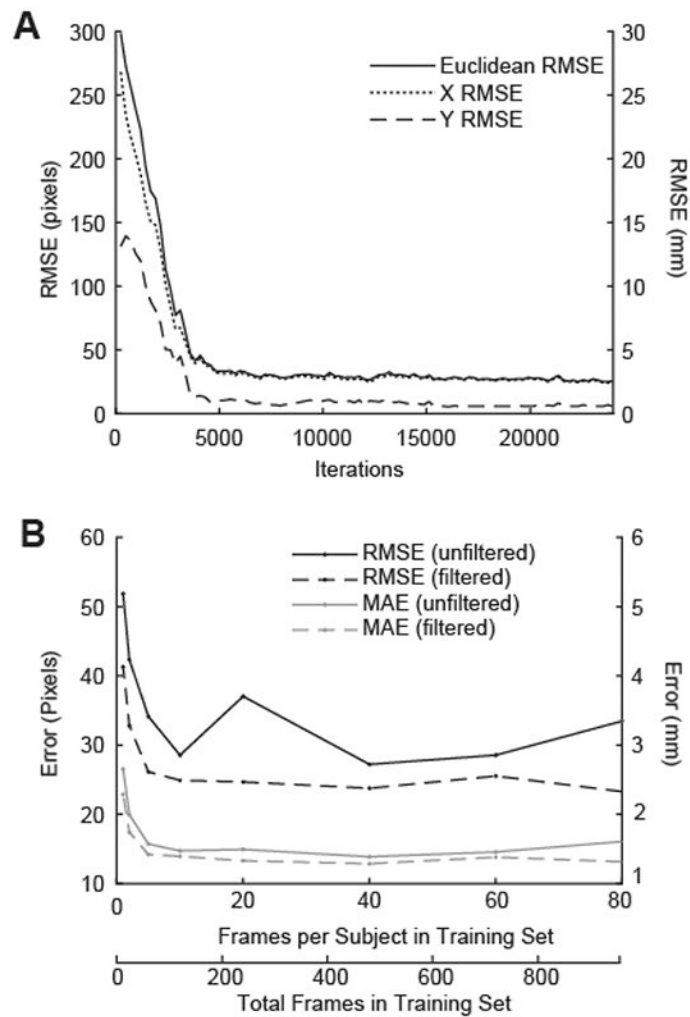


Figure 3.

A) Root mean square error (RMSE) between ground truth and predicted MTJ position was evaluated every 240 iterations during training. RMSE showed a steep initial drop and subsequent plateau around 4,000 iterations, and was minimized at 18,000 iterations. B) RMSE and MAE decreased with an increasing number of frames, but was relatively stable from 500 to 1000 frames. 480 labeled frames (40 frames per subject) were sufficient to achieve an overall RMSE of less than 3 mm and an overall MAE less than 2 mm. RMSE is reported in pixels (left axis) and mm (right axis).

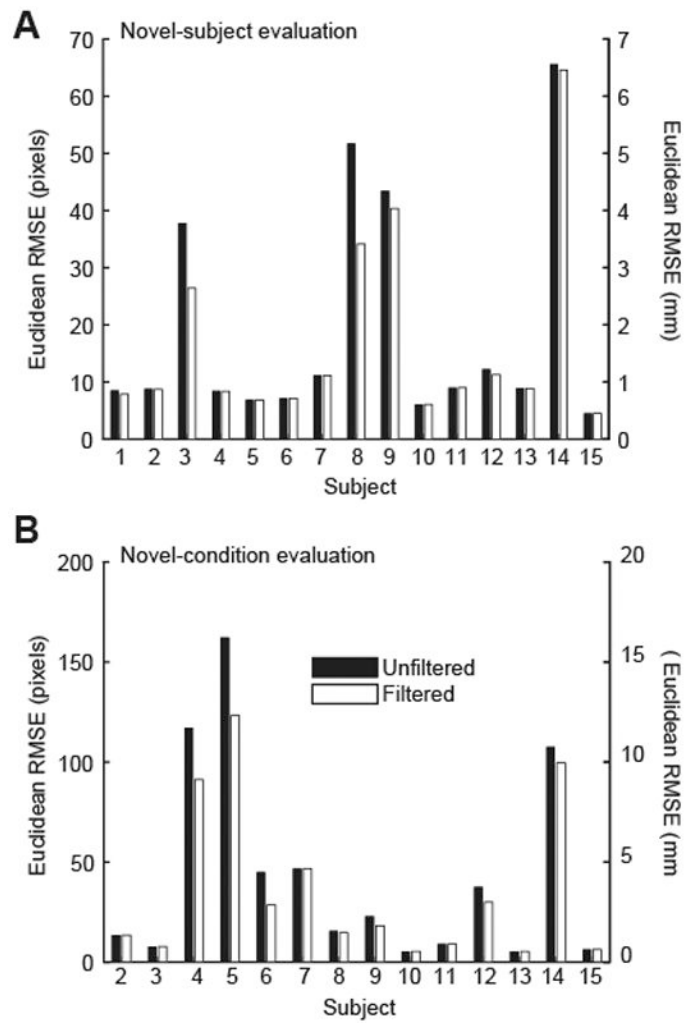


Figure 4. Unfiltered (black) and filtered (white) root mean square error (RMSE) for A) each individual subject in the novel subject evaluation (walking), and B) each individual subject in the novel condition evaluation (MVIC). RMSE is reported in pixels (left axis) and mm (right axis).

Table 1.

Model performance metrics (mm)

	RMSE	MAE (mean±SD)
Novel-subject (walking)		
Unfiltered	2.72	1.26±1.30
Filtered	2.37	1.17±1.23
Novel-condition (MVIC)		
Unfiltered	6.24	2.61±3.31
Filtered	5.05	2.18±2.64

RMSE: root mean square error, MAE: mean absolute error

Table 2.

Subject-specific and average prediction validity (%)

	Novel-subject (walking)		Novel-condition (MVC)	
	Unfiltered	Filtered	Unfiltered	Filtered
Sub001	100	100	-	-
Sub002	100	100	100	100
Sub003	90	96	100	100
Sub004	100	100	68	80
Sub005	100	100	48	55
Sub006	100	100	93	93
Sub007	100	100	98	98
Sub008	88	91	100	100
Sub009	71	73	95	98
Sub010	100	100	100	100
Sub011	100	100	100	100
Sub012	100	100	83	93
Sub013	100	100	100	100
Sub014	61	63	48	50
Sub015	100	100	100	100
Mean	94	95	88	90

Values represent the proportion of predictions that were within 5 mm of the ground truth.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

MTJ Excursion (mm)

	Novel-subject (walking)		Novel-condition (MVCs)	
	<i>Unfiltered</i>	<i>Filtered</i>	<i>Unfiltered</i>	<i>Filtered</i>
MTJ Excursion Manual				
<i>longitudinal</i>	19.20±5.37	-	26.97±15.57	-
<i>transverse</i>	1.26±0.42	-	10.01±14.87	-
MTJ Excursion Automatic				
<i>longitudinal</i>	23.72±10.29	21.59±8.13	22.28±12.47	20.59±10.51
<i>transverse</i>	3.57±7.08	1.54±0.46	6.81±11.57	2.89±2.62
Absolute Discrepancy				
<i>longitudinal</i>	4.79±9.10	2.68±5.80	4.69±9.55	6.38±10.51
<i>transverse</i>	2.36±6.99	0.38±0.31	3.20±10.05	7.12±13.72

Group mean ± standard deviation. Absolute discrepancy is the absolute difference between manual and automatic MTJ excursion.