An Audit of Open-Source Software for Ultrasound-Based Muscle Fascicle Tracking

Nicole Banka [1], Kayce Ochoa [2], John F. Drazan[1]
[1]Fairfield University, Fairfield, CT, [2]Central High School, Bridgeport, CT
Jdrazan@fairfield.edu

Disclosures: Nicole Banka (N), Kayce Ochoa (N), John Drazan (N)

INTRODUCTION: Ultrasound is a non-ionizing imaging modality for the visualization of muscle contractions during static and dynamic movement, making it a valuable tool for studying in vivo muscle function. It allows researchers to quantify musculoskeletal architecture and capture interactions between muscles and their surrounding structures[1]. Its portability, relatively low cost, and strong agreement with MRI for cross-sectional area measurements further support its use in movement and rehabilitation research[2]. By collecting real-time sequential images, ultrasound can reveal valuable information regarding tissue level deformations of muscle architecture such as fascicle length and pennation angle. Although manual tracking of these features can produce highly accurate data, the process is time-consuming and therefore limits its practicality in research studies, particularly with larger sample sizes or varied movement types. Automated computational methods offer high throughput alternatives, but their reliability without supervision remains a concern. Different tracking algorithms may also be prone to certain types of error depending on their computational design, including sensitivity to speckle noise, drift error, and limitations arising from training on specific datasets. Such constraints can reduce performance when analyzing lower-quality ultrasounds or muscle images outside of the training set. The objective of this study is to evaluate which recently published, open-source automated feature-tracking algorithms yield the most accurate results for fascicle length and pennation angle, when compared to manually validated data.

METHODS: A literature search of Google Scholar was conducted to identify open-source fascicle-tracking algorithms published between 2019 and 2025. Eligible algorithms were required to be publicly available, trained on datasets that included the medial or lateral gastrocnemius muscle, and provided automated or semi-automated tracking of fascicle length and pennation angle. Four tracking algorithms following these criteria were identified, along with one additional tracking algorithm from 2016 that was included due to its widespread use. Twenty-five open-source ultrasound videos of the medial gastrocnemius muscle were downloaded from an online repository, showing isometric and isokinetic maximal effort contractions at 30, 120, 210, and 500 degrees per second. Issues in documentation and execution were noted to assess usability. When necessary, codes were modified for compatibility with current software versions. Videos were manually scaled in each program whenever the option was available. Output data on fascicle length and pennation angle from each algorithm was collected and compared to manually validated measurements. Root mean square error (RMSE) was calculated for each video and then averaged to quantify the error present between the predicted and validated manual values for each program.

RESULTS: The average RMSE values (in mm) for each program were: 2.13 ± 1.06 for [3], 4.53 ± 3.35 for [4], 5.91 ± 3.47 for [5], 17.90 ± 5.22 for [6], and 581.75 ± 206.87 for [7]. Three trackers required code modifications, and one tracker needed the US videos to be extensively edited to detect features.

DISCUSSION: The average RMSE values indicated that most algorithms did not achieve acceptable accuracy, largely due to inconsistencies between the high-resolution training data used to develop the trackers and the lower-quality ultrasound clips analyzed here. This discrepancy limited the ability of the algorithms to reliably detect muscle features, resulting in frequent tracking errors. Fascicle length errors within a few millimeters are generally acceptable, but RMSE values greater than 5 mm indicated a noticeable level of discrepancy that would limit reliability for meaningful analysis. The worst-performing algorithm likely suffered from a frame-scaling issue as it did not allow for manual scaling. In addition, three types of issues were consistently encountered: coding modifications, video quality, and software add-ons. These challenges were particularly evident in the context of an undergraduate and high school research pursuit, where limited coding experience made troubleshooting time-intensive and restrictive. Collectively, the findings highlight the need for automated feature-tracking algorithms that can produce reliable data across image qualities, provide clearer documentation, and allow for more flexible user control. Future evaluations using more diverse ultrasound datasets would also be valuable, as they could better demonstrate algorithm reliability under conditions that more closely match training data. Such improvements are critical for increasing both accuracy and accessibility in orthopedic research.

SIGNIFICANCE/CLINICAL RELEVANCE: The ability to quickly and reliably collect data on fascicle length and pennation angle from ultrasound videos using open-source tracking algorithms offers researchers an important tool for studying muscle deformation across a variety of movements, however, the accessibility and applicability of these opensource algorithms were unknown. Our work highlights the need to create clearer documentation and analysis pipelines to allow for widespread use of these tools across the discipline.

REFERENCES: [1] Cronin, N. J. *et al. Gait Posture* 37, 305–312 (2013). [2] Crofts, G. *et al. Gait Posture* 39, 35–39 (2014). [3] Drazan, J. F. *et al. PeerJ* 7, e7120 (2019). [4] Ritsche, P. *et al. Ultrasound Med. Biol.* 50, 258–267 (2024). [5] Van Der Zee, T. J. *et al. PeerJ Comput. Sci.* 11, e2636 (2025). [6] Farris, D. J. *et al. Comput. Methods Programs Biomed.* 128, 111–118 (2016). [7] Van Der Zee, T. J. *et al. PLOS ONE* 17, e0265752 (2022)
ACKNOWLEDGEMENTS: This work was supported in part by the Sapere Aude Fund.

IMAGES AND TABLES:

Trackers	Average RMSE relative to manual measurement(mm)	Time to track 80 frames	Advantages	Weaknesses/Issues
[3]	2.13 ± 1.06	27 s	-Strongest RMSE value -Can validate data w/ manual input	-Program glitches and validation fails with videos >300 frames
[4]	4.53 ± 3.35	4 min, 45 s	-Trainable on new data sets -Fully automated tracking	-Frequent gaps in tracking -Doesn't allow for manual corrections
[5]	5.91 ± 3.47	32 s	-Allows for manual corrections -Still tracks features after they leave frame	-Poor tracking on low contrast videos
[6]	17.90 ± 5.22	3 s	-Allows for manual corrections -Insensitive to speckle noise	-Poor tracking and prone to drift error -Training data not made publicly available
[7]	581.75 ± 206.87	2 min, 32 s	-Not prone to drift error	-Poor tracking and sensitive to speckle noise -Doesn't allow for manual corrections