

Improved fascicle length estimates from ultrasound using a U-net-LSTM framework

Letizia Gionfrida, Richard W. Nuckols, Conor J. Walsh, and Robert D. Howe

Abstract— Brightness-mode (B-mode) ultrasound has been used to measure *in vivo* muscle dynamics for assistive devices. Estimation of fascicle length from B-mode images has now transitioned from time-consuming manual processes to automatic methods, but these methods fail to reach pixel-wise accuracy across extended locomotion. In this work, we aim to address this challenge by combining a U-net architecture with proven segmentation abilities with an LSTM component that takes advantage of temporal information to improve validation accuracy in the prediction of fascicle lengths. Using 64,849 ultrasound frames of the medial gastrocnemius, we semi-manually generated ground-truth for training the proposed U-net-LSTM. Compared with a traditional U-net and a CNN-LSTM configuration, the validation accuracy, mean square error (MSE), and mean absolute error (MAE) of the proposed U-net-LSTM show better performance (91.4%, MSE = 0.1 ± 0.03 mm, MAE = 0.2 ± 0.05 mm). The proposed framework could be used for real-time, closed-loop wearable control during real-world locomotion.

I. INTRODUCTION

As the field of wearable robotics moves toward dynamic real-world applications, the ability to continuously observe and respond to users' neuromuscular state may be important for providing individualized and adaptive assistance. B-mode ultrasound has received attention as a low-cost, non-invasive, and mobile modality to estimate muscle force capacity, energy use, and force production [1], [2], [3]. Previously, we used B-mode imaging to estimate muscle force and develop an ankle exosuit muscle-based assistance strategy [1] (Fig. 1). The assistance profiles, derived from B-mode imaging and the individual's biomechanics, reduced energy expenditure across tasks. Although the study demonstrated the benefits of muscle-based control, the B-mode image processing was performed offline and could not be used for real-time control. The investigation of more advanced techniques to analyze B-mode images and extract muscle states in a reliable manner in real-time is still needed.

The muscle states of most interest to be extracted from B-mode image frames are the fascicle lengths and pennation angles. From these, researchers have estimated muscle force production [4]. Measurements of fascicle length changes have been traditionally quantified by manually identifying the fascicle intersection of the upper and lower aponeurosis as

two bright echogenic linear structures [5]. Pennation angles have then been defined as the angle between the fascicle and aponeurosis orientations [5]. However, such manual identification on a frame-by-frame basis is extremely time-consuming, does not leverage the spatial resolution present in ultrasound images, and is prone to intra- and inter-rater variability, referring to the consistency or variability made by the same rater or observer or different raters or observers, respectively, when assessing the same [6].

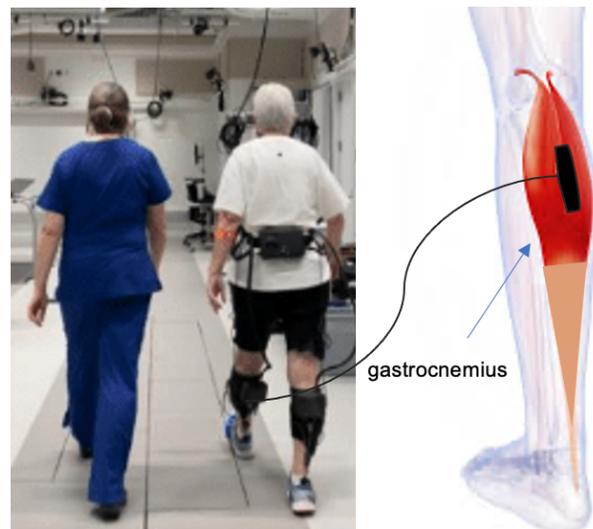


Fig. 1. Representative figure illustrating data collection on older adults, showing the ultrasound placement over the medial gastrocnemius muscles.

Automated image analysis techniques have been developed to address these challenges. Initial work outlined fascicles based on pixel intensity to estimate muscle fiber orientation [7]. However, the technique is affected by the quality of the image frame and often requires image enhancement filters that can generate artifacts [8]. Optical flow, which estimates the motion of individual pixels on the sequence of images, can estimate fascicle length changes across frame sequences [9]. However, these models can generate errors when encountering large movements. UltraTrack is a semi-automated application commonly used to track fascicle lengths and pennation angles that rely on an affine optical flow model [6]. A significant drawback is its

**Research supported by the National Institutes of Health (NIH Grant Number: 5R21AR076686-02).

L. Gionfrida, C. J. Walsh, R. D. Howe are with the Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Science and Engineering Complex, 150 Western Ave, Boston, MA 02134, USA (e-mail: {gionfrida, walsh, howe}@seas.harvard.edu).

L. Gionfrida is with the Department of Informatics, King's College London, Strand, London, WC2R 2LS, United Kingdom (e-mail: letizia.gionfrida@kcl.ac.uk).

R. W. Nuckols is with the Department of Systems Design Engineering University of Waterloo, University Ave W Waterloo, ON N2L 3G1, Canada (e-mail: muckols@uwaterloo.ca).

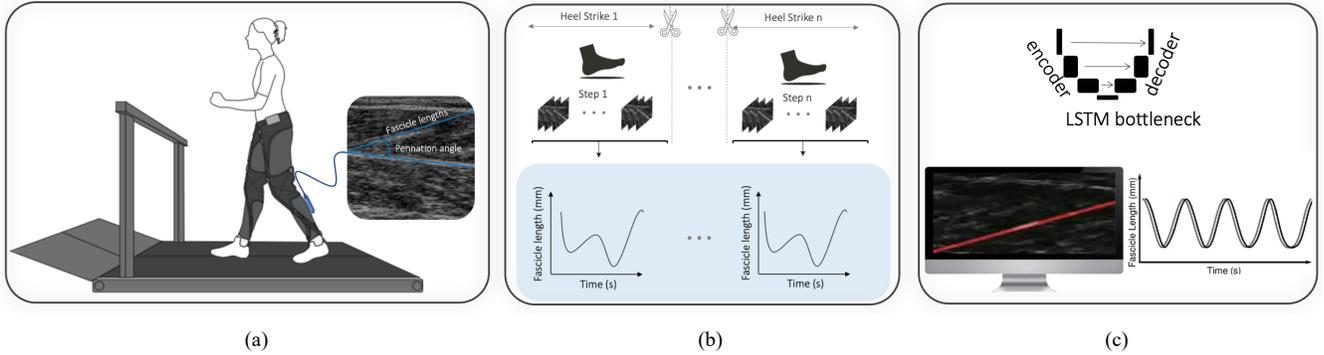


Fig. 2. The proposed framework where an ultrasound probe is placed over the medial gastrocnemius muscle as users walk at varying speeds (1.00, 1.25, 1.5, and 1.75 m/s) and tasks (level ground and incline) (a), the data is labelled based on pre-recorded heel strikes (b), and the U-net-LSTM architecture extracts the fascicles lengths compared against the labelled sequence (c).

tendency to accumulate frame-by-frame shifts [10]. It also relies on users' initial manual annotation, after which the optic flow model is applied, raising inter- and intra-operator variability that translates into inconsistent and unreliable fascicle tracking over long sequences [11].

More recently, machine learning has presented the opportunity to automate the estimation of muscle architecture during dynamic motion tasks (e.g., ankle motion constrained, level-ground, and sloped walking at different speeds) [12], [13]. A preliminary attempt at using support vector regression [14] has the drawback of acting as a “black box,” not allowing researchers to visually confirm the fascicle being tracked, and presenting only moderate accuracy (average correlation between predicted values and the ground-truth, $r = 0.65$) [12].

Efforts to determine muscle architecture from B-mode and capture the relationship between spatial and temporal structure using a more generalizable convolutional neural network (CNN) are presented in [15]. Several studies have also illustrated that a hybrid CNN-LSTM architecture has better performance in terms of visually learning time-sequenced information than independent CNN and LSTM models [16], [17]. However, CNN-LSTM may not effectively capture the spatial features of the muscle fascicles in B-mode, especially when the images are of mid-to-low resolution [18]. Therefore, given the sequential nature of the continuous muscle recording in ultrasound imaging, a deep learning model that can read and visualize sequential data while interpreting long-term temporal dependencies would be preferred to increase robustness.

U-net models have proven to be successful and accurate (98% validation accuracy over a fascicle classification task) for image segmentation of muscle fascicles, allowing researchers to validate the performances of the system by visually identifying fascicles and angles [19], [20]. However, the U-net architecture is designed for image segmentation, but it does not consider long-term temporal dependencies, which is an important aspect of B-mode images [13]. U-net architectures are sensitive to variability, and their performance degrades dramatically when the training set comes from one participant and the validation set comes from a different participant [21].

In this article, we propose a deep U-shaped network-LSTM (U-net-LSTM) framework for the time-sequenced prediction of muscle fascicles. The main contributions of the presented work are:

A deep U-net-LSTM framework that is able to incorporate long-term temporal dependencies of B-mode images into the segmentation process. This allows for more accurate segmentation of the fascicle being tracked.

A framework that could improve generalizability over input images from different participants and tasks compared to standard U-nets, that learns and adapts to the temporal characteristics of the muscle fascicles over time.

In the next sections, the work is organized as follows: Section II describes the methodology, including the dataset used in the study and the proposed U-net-LSTM framework. Section III illustrates the valuation results in terms of training and validation accuracy, mean square error (MSE), and mean absolute error (MAE) of the proposed deep U-Net-LSTM framework, comparing its performance against that of other deep learning frameworks. Section IV discusses the strengths and limitations of the proposed methodology. Finally, Section V concludes with an extension of the approach and future directions.

II. METHODS

Image frames are labeled using a combination of manual fascicle detection to generate binary masks and UltraTrack [6] applied on a small sub-section of the recording to avoid spatial drift [22]. Next, the U-net-LSTM network is trained and tested to estimate the fascicle lengths (Fig. 2).

A. Dataset and Labelling

The dataset [1] contained 64,849 image frames from nine healthy participants walking on a treadmill, as illustrated in Fig. 2 (a), on level ground at walking speeds of 1.00, 1.25, 1.5, and 1.75 m/s, and at 5.71° (10% incline) at 1.25 m/s, using an instrumented treadmill (Bertec, Columbus, OH, USA; 1200 Hz) that measured ground reaction forces. Participants wore a MicroUs Telemed ultrasound transducer that captured B-mode images of the medial gastrocnemius muscle with a 75 mm probe and a 5 MHz center frequency.

The study is approved by the Harvard Longwood Campus Institutional Review Board, protocol number IRB14-3608s. All methods are carried out in accordance with the approved study protocol and written informed consent was obtained prior to the start of the study for all participants.

The labeling approach aimed to generate ground-truth masks of the fascicle lengths to train the U-net-LSTM framework, defined by selecting pixels that belonged to muscle fascicles. A combination of manual labeling and semi-automated labeling (with UltraTrack) is utilized. The division into manual labeling and semi-automated labeling for 60,549 frames is done to minimize the spatial drift caused by the affine optical flow introduced by UltraTrack over longer sequences. Based on pre-recorded heel strikes, each sequence of a participant walking (on average 1441 ultrasound frames) is divided into nine smaller sub-sequences. We manually segmented the first image frames of the nine sequences for each participant and task individually, to mimic the gold-standard manual labeling process. Then, we continued labeling in a more sustainable way using UltraTrack. The masks are then reconstructed based on the manually identified fascicles and the output fascicle length values from UltraTrack to generate the masks.

To ensure that the labeling is appropriate, all the fascicle lengths partitioned by the pre-recorded heel strikes are plotted following the approach proposed by Lai et al. [23] to confirm correct labeling, as illustrated in Fig. 2 (b).

B. U-net-LSTM framework

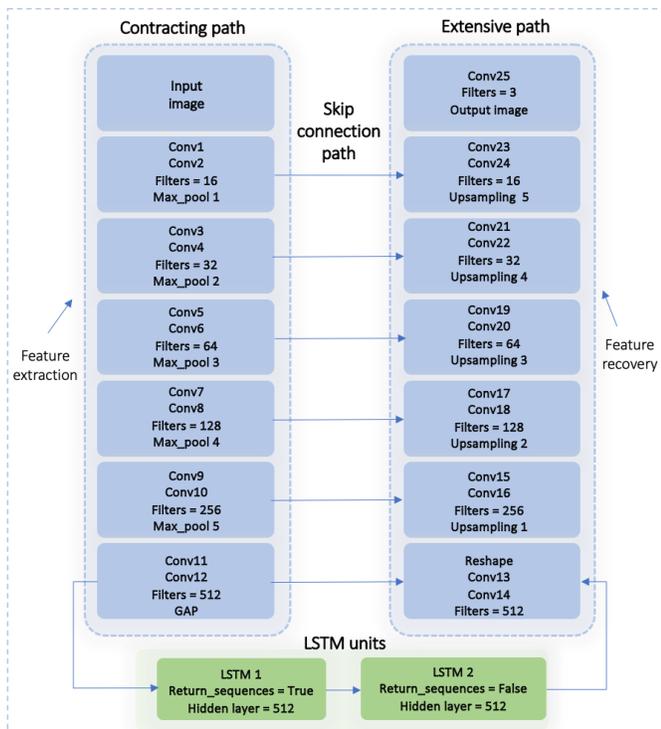


Fig. 3. The proposed deep U-Net-LSTM framework.

The deep U-net-LSTM framework is designed to predict time-sequenced fascicle lengths from B-mode ultrasound image frames. It combines a U-shaped network [22], two LSTM units [24], and a skip connection (Fig. 3). The U-shaped

framework is made of a contracting (or encoding) path and an expansive (or decoding) path. The contracting path is utilized to extract complex features and follows the architecture of a convolutional network. It is made of two 3×3 unpadded convolutions, each followed by a rectified linear unit (ReLU) activation function and a 2×2 max pooling operation that enables downsampling; the number of feature channels doubles after the max pooling operation.

The input layer of the encoding path accepts time-sequenced images as inputs. To reshape the output of the convolutional network layers into a format that can be fed into the LSTM layers, from a three-dimensional tensor to a two-dimensional matrix, a global average pooling (GAP) layer is used to reduce the dimensions of input images and prevent overfitting in the framework.

The LSTM section is placed between the contracting and the extensive paths and is used to learn temporal information. This is composed of two LSTM layers connected in series, with 512 hidden layers each. In the first LSTM layer, the `return_sequences` parameter is set to `True`, meaning that the layer can output a sequence of hidden states that represent the temporal features of the input data. These hidden states are then fed into the second LSTM layer to further capture the temporal dependencies in the data. The `return_sequences` parameter of the second LSTM layer is then set to `False`, meaning that only the last predicted vectors of the LSTM layer are outputs. After the two LSTM layers are processed in series, a reshape function is used to transform the shape of the output of the LSTM layers and prepare it for further processing in downstream layers.

The skip connection section sends more high-level semantics to the expansive path. For instance, as the conv10 layer and the conv16 layer are in symmetry with the same size feature map, the time series features output by the conv10 layer are sent to the conv16 layer. After the channels concatenate, the shape will change.

Finally, the expansive path recovers the vectors of the same size as the input by upsampling. The expansive path consists of a 2×2 upsampling layer, and two 3×3 convolutions followed by a ReLU activation function.

C. Training and Validation

The training and validation procedures of the proposed framework for fascicle length prediction are summarized as follows:

1. Each B-mode ultrasound image frame is imported and resized to 512×512 pixels from the original resolution of 581×681 pixels to build a database for training and validation (Fig. 4). Once resized, the frames are labeled, based on pre-recorded heel-strikes, as described above, to generate the binary masks to train the proposed U-net-LSTM (Fig. 4). The imported image frames are split between training and validation with an 80:20 ratio.
2. The training involves forward-propagating image frames through the U-net-LSTM, calculating the loss between the predicted fascicle lengths and the labeled fascicle lengths, and back-propagating to update the model parameters. This study utilizes the Adam optimizer [25] to adjust the weights. To prevent overfitting in the training process,

data augmentation techniques, including spatial cropping and rotation, are used. The total number of epochs is set to 120, and the initial learning rate is 0.001. The environments are a Google NVIDIA T4 Tensor Core GPU [26], Python (Python Software Foundation, version 3.7), and PyTorch [27].

3. The output images from the U-net-LSTM are then extracted (Fig. 4). Then two endpoints of each fascicle in the segmented images are identified and the distance between them is measured to obtain the fascicle length. This process is repeated for each segmented image in the output from the U-net-LSTM, and the calculated fascicle lengths are combined to obtain the fascicle lengths estimate for the full sequence over time (Fig. 4).
4. After the training to generate the output images, the U-net-LSTM framework's performances are evaluated. The validation accuracy, as a common metric to assess the correctly classified pixels in the segmentation map, is considered. To evaluate the framework's performance in mm, the mean absolute error (MAE) and mean squared error (MSE), which are more sensitive to individual pixel values, are also considered. The calculation formulas are:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\vartheta_i - \hat{\vartheta}_i)^2 \quad (1)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\vartheta_i - \hat{\vartheta}_i| \quad (2)$$

where n represents the number of samples (total number of pixels over the sequence), i represents the ranges of samples pixels across time, ϑ_i the ground-truth masks, and $\hat{\vartheta}_i$ the predicted result.

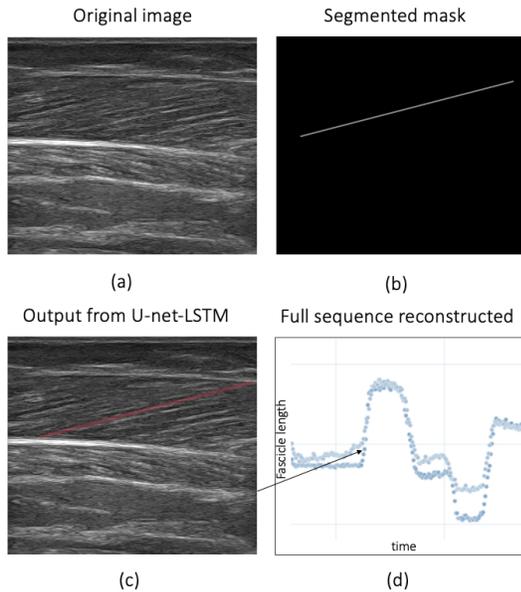


Fig. 4. An example of the original B-mode ultrasound image frame (a); the segmented mask (ground-truth) (b); the output image from the U-net-LSTM framework with the red line visualizing the fascicle predicted (c); representation of the two reconstructed fascicle length sequences over time based on individual output images (the arrow that goes from the output of the U-net-LSTM (c) to one dot value of the scatter plot shows that to each output frame from the U-net-LSTM corresponds a dot of the scatter plot) (d). This example shows a single fascicle for clarity; typical frames include dozens of labeled fascicles.

D. Comparison with other frameworks

An ablation study is conducted, where individual components of the proposed U-net-LSTM framework are removed to evaluate their contribution to the overall performance. The ablation study focused on evaluating the influence of two components, the skip connection components of the U-net-LSTM, as in traditional U-net [20], and the LSTM units of the U-net-LSTM, as in a conventional CNN-LSTM structure [15]. To do so, the performance of the U-net-LSTM are contrasted with those of a traditional U-net by removing the series of the two LSTMs and with those of a conventional CNN-LSTM by replacing the convolutional layers to enable sequences of images as input instead of individual images.

III. RESULTS

The proposed deep U-net-LSTM framework trained on 120 epochs for eight participants walking on a treadmill, as in Fig. 2 (a), on level ground at walking speeds of 1.00, 1.25, 1.5, and 1.75 m/s, and at 5.71° (10% incline) at 1.25 m/s presents training and validation accuracy of 98.2% and 91.4%, respectively (Fig. 5). The average MSE and MAE across participants and tasks with their standard deviations (SD) for the U-net-LSTM are $0.1 \pm 0.03\text{mm}$ and $0.2 \pm 0.05\text{mm}$, respectively. The average detection time, as to the time it takes for the U-net-LSTM framework to extract fascicle information from a full sequence of participants' walking tasks, is 0.35 seconds.

The results of the ablation process that removed components of the U-net-LSTM to evaluate the individual components' overall contribution to the proposed framework's performance are shown in Table I. Overall, accuracy, MSE, and MAE are all worse if any of the components are omitted from the framework. The MSE and MAE across participants for each walking condition for the ablation study are presented in Table II. Again, the removal of any component from the framework results in a steep decline in performance for all walking conditions.

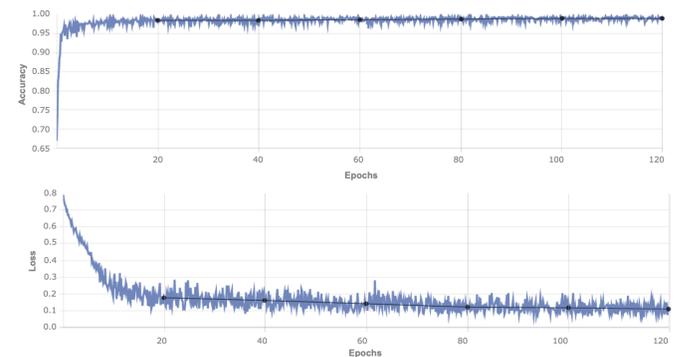


Fig. 5. The proposed U-net-LSTM framework training accuracy and loss curves for 120 epochs.

A representative fascicle length estimate from the various configurations for a randomly selected participant and walking task in the validation set is illustrated in Fig. 6. Here, the fascicle lengths of each segmented image frame in output from the U-net-LSTM framework and the U-net are measured using

the distance between the two endpoints and computed for the full B-mode image sequence.

TABLE I. ABLATION COMPARISON BETWEEN THE PROPOSED U-NET-LSTM FRAMEWORK, A CNN-LSTM, AND A U-NET.

Frameworks	Dataset Accuracy		MSE \pm SD (mm)	MAE \pm SD (mm)
	Training (%)	Validation (%)		
Traditional CNN-LSTM (+LSTM)	85	73	5 \pm 0.67	3 \pm 0.62
Traditional U-net (+Skip connection component)	91	59	4.2 \pm 0.4	1.9 \pm 0.71
Proposed U-net-LSTM (+LSTM + skip connection component)	98.2	91.4	0.1\pm0.03	0.2\pm0.05

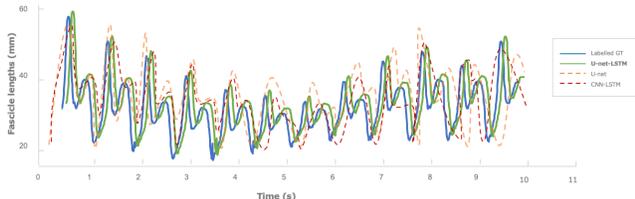


Fig. 6. A comparison of the fascicles in output from the proposed U-net-LSTM framework (in green), with a traditional U-net and CNN-LSTM of fascicle lengths from randomly selected participant 2 when walking on level ground at 1.25 m/s.

TABLE II. MEAN ABSOLUTE ERRORS AND MEAN SQUARED ERRORS (IN MILLIMETERS) FOR DIFFERENT DEEP LEARNING FRAMEWORKS ACROSS WALKING TASKS.

Frameworks		Walking velocities (m/s)				Incline 1.25m/s
		1.00	1.25	1.50	1.75	
Traditional CNN-LSTM (+LSTM)	MSE	4.9	4.2	6.2	6.1	5.3
	MAE	2.8	3.2	3.1	3.6	2.4
Traditional U-net (+Skip connection component)	MSE	2.9	5.1	4.7	6.5	3.1
	MAE	1.3	0.5	1.8	2.1	2
Proposed U-net-LSTM (+LSTM + skip connection component)	MSE	0.4	0.8	0.05	0.7	0.1
	MAE	0.15	0.1	0.05	0.2	0.3

IV. DISCUSSION

This article proposes a deep U-net-LSTM framework that aims to incorporate temporal information to improve estimates of fascicle lengths. The framework is made of a deep U-shaped network, two LSTM units in series, and a skip connection component. The LSTM units are used to model the temporal

dynamics of the B-mode ultrasound video recordings by selectively remembering and forgetting fascicle length information from previous time steps. In comparison with a traditional U-net [20] and a CNN-LSTM [15], the proposed framework showed greatly improved performance in terms of validation accuracy, MSE, and MAE.

Our experiments illustrated that it is possible to detect and visualize fascicle architectures across two new participants and tasks (only used during the validation phase) when analyzing B-mode ultrasound frames. The predicted results for the validation set from the U-net-LSTM are very close to the ground-truth results, indicating that the proposed framework may have good stability to predict the results of additional future timesteps for participants on which it is not trained. Compared with the traditional U-net [20] and CNN-LSTM [15], the U-net-LSTM shows that the validation accuracy improves from 59% and 73% to 91.4%, respectively, and the MSEs and MAEs are reduced by one order of magnitude with the proposed approach.

The results show that across frameworks, the performance generally degrades with an increase in walking speed, while incline walking does not affect the performance of the deep learning frameworks. This is expected as the difference between frames increases as walking speeds increase. The higher validation accuracy and lower MSE and MAE values suggest that the proposed U-net-LSTM framework could generalize better to unseen data compared to the traditional U-net and CNN-LSTM, suggesting that the U-net-LSTM can be more effective at predicting and modeling temporal sequential changes in fascicle lengths over time.

The ablation process shows that when the LSTM component is added, the mean MSE and MAE on the validation datasets are reduced, which verifies the effectiveness of the LSTM part in the proposed framework. When the skip connection part is added, the mean MSE and MAE are also reduced, which verifies the effectiveness of the skip connection part. These results suggest that together, the skip connection component and LSTM units positively impact the validation accuracy of the proposed U-net-LSTM. When both parts are added, the proposed framework achieves the best performance in terms of the selected evaluation criteria based on the validation dataset.

Although the CNN-LSTM framework presents good training and validation accuracy, comparable with that of the U-net-LSTM, it also presents higher errors. This may be related to the fact that a CNN-LSTM framework is usually designed to capture temporal dependencies, while U-Net-LSTM models consider both temporal dependencies and image segmentation features. The proposed U-net-LSTM leverages the strength of a traditional U-net to conduct segmentation on an image and then passes the segmented image through an LSTM network to capture the temporal dependencies between frames in a sequence. This highlighted the strength of the proposed framework.

The proposed approach also facilitates interpretability, due to visualization of the segmentation results. This allows users to immediately confirm the quality of the output segmentations by simply examining ultrasound images with superimposed segmentation results, as illustrated in Fig. 4(c). This can help

with debugging and validation, which could be particularly important for subjects with musculoskeletal pathologies [28] or for potential real-time control applications.

We acknowledge that partial manual fascicle detection used to train the U-net-LSTM, combined with UltraTrack, can lead to errors in the labeled dataset. Instead of labor-intensive manual labeling, which would take approximately eight hours per participant per task, the proposed labeling approach takes approximately 20 minutes for one participant across all tasks, and for this reason, it was preferred to generate the labeled ground-truth. Also, the validation of the proposed framework on different populations can be a preferred strategy to validate the proposed method and will be performed in future investigations. The average detection time, which refers to the time it takes for the proposed U-net-LSTM framework to extract fascicle lengths from a participant's task, is 0.35s, which suggests that it is possible to achieve real-time prediction through optimization of the approach so that the proposed U-net-LSTM framework can be evaluated in real-world human-in-the-loop assistance deployments in future investigations.

V. CONCLUSION

This paper presents a framework for autonomous fascicle length detection from B-mode ultrasound. It combines a U-net architecture for good segmentation performance with an LSTM component that considers temporal information to improve accuracy. The proposed framework provides accurate pixel-wise fascicle identification. Preliminary execution time estimates suggest that this framework may be useful for muscle-driven closed-loop control and visualization. In future work, we aim to test the framework for the development of exoskeleton assistance profiles tailored to the participant's individual biomechanics and adjusted to the specific task.

REFERENCES

- [1] R. Nuckols, S. Lee, K. Swaminathan, D. Orzel, R. Howe, and C. Walsh, 'Individualization of exosuit assistance based on measured muscle dynamics during versatile walking', *Science Robotics*, vol. 6, no. 60, p. eabj1362, 2021.
- [2] I. D. Loram, C. N. Maganaris, and M. Lakie, 'Use of ultrasound to make noninvasive in vivo measurement of continuous changes in human muscle contractile length', *Journal of applied physiology*, vol. 100, no. 4, pp. 1311–1323, 2006.
- [3] T. J. Dick, A. S. Arnold, and J. M. Wakeling, 'Quantifying Achilles tendon force in vivo from ultrasound images', *Journal of biomechanics*, vol. 49, no. 14, pp. 3200–3207, 2016.
- [4] S. Bohm, F. Mersmann, A. Santuz, and A. Arampatzis, 'The force-length-velocity potential of the human soleus muscle is related to the energetic cost of running', *Proceedings of the Royal Society B*, vol. 286, no. 1917, p. 20192560, 2019.
- [5] S. Kurokawa, T. Fukunaga, and S. Fukashiro, 'Behavior of fascicles and tendinous structures of human gastrocnemius during vertical jumping', *Journal of Applied Physiology*, 2001.
- [6] D. J. Farris and G. A. Lichtwark, 'UltraTrack: Software for semi-automated tracking of muscle fascicles in sequences of B-mode ultrasound images', *Computer methods and programs in biomedicine*, vol. 128, pp. 111–118, 2016.
- [7] Y. Zhou and Y.-P. Zheng, 'Estimation of muscle fiber orientation in ultrasound images using revolving hough transform (RVHT)', *Ultrasound in medicine & biology*, vol. 34, no. 9, pp. 1474–1481, 2008.
- [8] H. Zhao and L.-Q. Zhang, 'Automatic tracking of muscle fascicles in ultrasound images using localized radon transform', *IEEE transactions on biomedical engineering*, vol. 58, no. 7, pp. 2094–2101, 2011.
- [9] B. Van Hooren, P. Teratsias, and E. F. Hodson-Tole, 'Ultrasound imaging to assess skeletal muscle architecture during movements: a systematic review of methods, reliability, and challenges', *Journal of Applied Physiology*, vol. 128, no. 4, pp. 978–999, 2020.
- [10] T. J. van der Zee and A. D. Kuo, 'TimTrack: A drift-free algorithm for estimating geometric muscle features from ultrasound images', *PLoS One*, vol. 17, no. 3, p. e0265752, Mar. 2022, doi: 10.1371/journal.pone.0265752.
- [11] J. F. Drazan, T. J. Hullfish, and J. R. Baxter, 'An automatic fascicle tracking algorithm quantifying gastrocnemius architecture during maximal effort contractions', *PeerJ*, vol. 7, p. e7120, Jul. 2019, doi: 10.7717/peerj.7120.
- [12] L. G. Rosa, J. S. Zia, O. T. Inan, and G. S. Sawicki, 'Machine learning to extract muscle fascicle length changes from dynamic ultrasound images in real-time', *PLoS one*, vol. 16, no. 5, p. e0246611, 2021.
- [13] G.-Q. Zhou, P. Chan, and Y.-P. Zheng, 'Automatic measurement of pennation angle and fascicle length of gastrocnemius muscles using real-time ultrasound imaging', *Ultrasonics*, vol. 57, pp. 72–83, Mar. 2015, doi: 10.1016/j.ultras.2014.10.020.
- [14] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, 'Support vector regression machines', *Advances in neural information processing systems*, vol. 9, 1996.
- [15] J. Donahue *et al.*, 'Long-term recurrent convolutional networks for visual recognition and description', presented at the Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 2625–2634.
- [16] L. Gionfrida, W. M. Rusli, A. E. Kedgley, and A. A. Bharath, 'A 3DCNN-LSTM Multi-Class Temporal Segmentation for Hand Gesture Recognition', 2022.
- [17] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx, 'A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data', *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 62–77, 2020.
- [18] R. J. Cunningham and I. D. Loram, 'Estimation of absolute states of human skeletal muscle via standard B-mode ultrasound imaging and deep convolutional neural networks', *Journal of the Royal Society Interface*, vol. 17, no. 162, p. 20190715, 2020.
- [19] W. Zheng, L. Zhou, Q. Chai, J. Xu, and S. Liu, 'Fully Automatic Analysis of Muscle B-Mode Ultrasound Images Based on the Deep Residual Shrinkage U-Net', *Electronics*, vol. 11, no. 7, Art. no. 7, Jan. 2022, doi: 10.3390/electronics11071093.
- [20] O. Ronneberger, P. Fischer, and T. Brox, 'U-net: Convolutional networks for biomedical image segmentation', presented at the International Conference on Medical image computing and computer-assisted intervention, 2015, pp. 234–241.
- [21] T. Muramatsu, T. Muraoka, Y. Kawakami, A. Shibayama, and T. Fukunaga, 'In vivo determination of fascicle curvature in contracting human skeletal muscles', *Journal of Applied Physiology*, vol. 92, no. 1, pp. 129–134, 2002.
- [22] N. J. Cronin, T. Finni, and O. Seynnes, 'Fully automated analysis of muscle architecture from B-mode ultrasound images with deep learning', *arXiv preprint arXiv:2009.04790*, 2020.
- [23] A. Lai, G. A. Lichtwark, A. G. Schache, Y.-C. Lin, N. A. Brown, and M. G. Pandy, 'In vivo behavior of the human soleus muscle with increasing walking and running speeds', *Journal of applied physiology*, vol. 118, no. 10, pp. 1266–1275, 2015.
- [24] S. Hochreiter and J. Schmidhuber, 'Long short-term memory', *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [25] D. P. Kingma and J. Ba, 'Adam: A method for stochastic optimization', *arXiv preprint arXiv:1412.6980*, 2014.
- [26] E. Bisong, *Building machine learning and deep learning models on Google cloud platform*. Springer, 2019.
- [27] A. Paszke *et al.*, 'Pytorch: An imperative style, high-performance deep learning library', *Advances in neural information processing systems*, vol. 32, 2019.
- [28] L. Gionfrida, R. W. Nuckols, C. J. Walsh, and R. D. Howe, 'Age-Related Reliability of B-Mode Analysis for Tailored Exosuit Assistance', *Sensors*, vol. 23, no. 3, p. 1670, 2023.