Towards a Wearable Low-Cost Ultrasound Device for Classification of Muscle Activity and Muscle Fatigue

Lukas Brausch
Fraunhofer Institute for Biomedical Engineering
Sankt Ingbert, Germany
Lukas.Brausch@ibmt.fraunhofer.de

Holger Hewener
Fraunhofer Institute for Biomedical Engineering
Sankt Ingbert, Germany
Holger.Hewener@ibmt.fraunhofer.de

Paul Lukowicz
German Research Center for Artificial Intelligence
Kaiserslautern, Germany
Paul.Lukowicz@dfki.de

ABSTRACT

Being able to reliably predict muscle contractions is important for athletes and rehabilitation patients alike. Numerous techniques and surrogates exist for this task. However, they are in general not well suited for everyday use and not able to extract information of muscles located in deeper body layers. To address this shortcoming, we present an approach to classify muscle contractions with raw ultrasound radio-frequency data (A-Scans) collected with a wearable system. It consists of a single element ultrasound transducer connected to custom-built acquisition hardware and an Android app to receive, store and analyze the data. We rely on data from the lower legs of healthy volunteers performing squats as sample exercises and use machine learning methods, ranging from sequence similarity measurement techniques to artificial neural networks, to classify the radio-frequency data. Results of our preliminary experimental setup prove its feasibility to classify muscle contractions based on ultrasound measurements.

KEYWORDS

Wearable Ultrasound System, Machine Learning, Muscle Activity, Muscle Fatigue.

ACM Reference Format:
Lukas Brausch, Holger Hewener, and Paul Lukowicz. 2019. Towards a Wearable Low-Cost Ultrasound Device for Classification of Muscle Activity and Muscle Fatigue. In *Proceedings of the 2019 International Symposium on Wearable Computers (ISWC '19), September 9–13, 2019, London, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3341163.3347749

ISWC '19, September 9–13, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6870-4/19/09.
https://doi.org/10.1145/3341163.3347749

1 INTRODUCTION

The reliable detection of muscle contractions and muscle fatigue remains a challenge with respect to unobtrusive systems suitable for long-term everyday use. Arguably, the most promising non-invasive techniques include *surface electromyography (sEMG)*, *force sensitive resistors*, *inertial measurement units* or *textile capacitive sensors*. However, these techniques do not provide information about muscles deep below the surface tissue [12]. To circumvent this limitation, wearable ultrasound solutions have been proposed [12] to acquire signals stemming from muscles located deep in the human body. In this paper we present a system consisting of a wearable single element ultrasound transducer, custom-designed electronics and an Android app to acquire and process ultrasonic radio-frequency (*RF*) data from eight healthy volunteers (see Table 1). We use digital signal processing (*DSP*) and machine learning (*ML*) approaches to classify muscle contractions based on this data. Our system solely relies on one-dimensional RF data (A-Scans) and **not** on heavily processed B-mode images. The latter are very suitable for the visualization of different muscle groups but the former only provide depth information of the tissue (see Figure 1). Even though A-Scans are much less illustrative due to their one-dimensional nature, they can be acquired without complex or bulky electronics to enable ultrasound imaging, facilitating a low resource wearable solution. This is a fundamental difference to existing systems and a major contribution of the paper.

Table 1: Attributes and available data for each subject.

ID	*Age*	*Gender*	*# A-Scans*	*# Data-Sets*	*# Squats*	*# Measured time (s)*
1	28	m	92,000	7	154	2,193
2	26	m	2,000	2	8	37
3	27	f	18,872	2	35	364
4	24	m	20,000	2	49	463
5	35	m	10,000	1	27	245
6	35	m	10,000	1	21	221
7	29	m	20,000	2	88	338
8	37	m	40,000	4	133	828

2 BACKGROUND AND RELATED WORK

Related work attempting to use only RF data (A-Scans) compared to B-Mode images exists for finger motion classification [8, 12], gesture recognition [7] or wrist angle detection [4, 5] tasks. However, these works either rely on A-Scans derived from B-mode images [8] or are based on signals acquired with multiple single element transducers [7, 12]. Guo et al. present work based on signals stemming from a *single* element transducer only and process these data with classical DSP techniques [4] and ML strategies [5] to predict wrist angles. However, they do not present a wearable solution.

3 MATERIALS AND METHODS

Our system consists of a custom-designed electronics board (with the dimensions of approx. 90 mm × 30 mm × 13 mm), sending the RF data acquired with a single element ultrasound transducer to an Android app via a wireless connection. In our experimental setup we asked participants to place the transducer *anywhere* above the *Gastrocnemius* calf muscle, located on the back of the lower leg, and perform squats as long as possible, which resulted in 21 datasets. The amount of datasets and squats for each subject differs and depends on the level of exhaustion experienced. We obtained 1 hour, 18 minutes and 9 seconds of data containing 212,872 A-Scans and 515 squats (see Table 1). The data acquisition without emphasis on the best possible ultrasound transducer position might have resulted in worse signals but we feel this approach adequately mimics real-life scenarios. The volunteers annotated each A-Scan either as *contracted* or *non-contracted* by manually pressing a button during each squat. This annotated data serves as ground truth for our binary classification algorithms making predictions based on the amplitude values of each A-Scan. We make this data publicly accessible [2] and encourage further investigations by others.

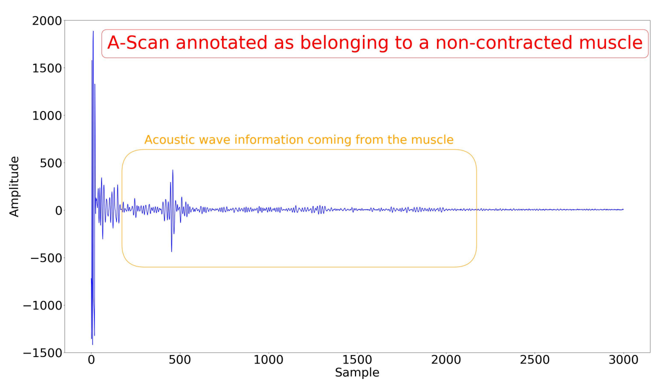

Figure 1: A-Scan with 3,000 samples and measurement depth of 46.2 mm. The highlighted box contains information from the muscle.

Figure 1 shows a sample A-Scan acquired in our experimental setup, which has been annotated as belonging to a non-contracted muscle state. It consists of 3,000 samples and has a tissue penetration depth of 46.2 mm. We also performed post-processing incorporating assumptions about the expected behavior of real muscle activity to smooth out misclassified predictions. Assumptions include that muscle contractions are continuous and cannot occur consecutively within a very short period of time, which excludes any physically impossible classifications. For analysis, we deployed a threshold DSP algorithm comparing the *maximum*, *minimum*, *mean*, *variance* and *standard deviation* values of each A-Scan. The ML methods include custom versions of a *multilayer perceptron (MLP)* [3, 11], a *fully convolutional network (FCN)* [11], a *1-D convolutional network (ConvNet1D)* [3], a *radial basis function network (RBF)* [6, 10] and a *residual neural network (ResNet)* [11]. We also applied a *k-nearest neighbor (k-NN)* algorithm using the *dynamic time warping (DTW)* distance [1]. Table 2 and Table 3 provide details about the deployed feedforward and convolutional neural networks, respectively.

Table 2: Parameters of feedforward neural networks

<table>
<tr><th>Parameter</th><th>MLP</th><th>RBF</th></tr>
<tr><td># Hidden layers</td><td>3</td><td>1</td></tr>
<tr><td># Neurons</td><td>(500|500|500)</td><td>1000</td></tr>
<tr><td>Dropout values</td><td>(0.2|0.2|0.3)</td><td>0.2</td></tr>
<tr><td>Activation function ∀ hidden layers</td><td>relu</td><td>Gaussian similarity function</td></tr>
<tr><td>Activation function for output layer</td><td>sigmoid</td><td>sigmoid</td></tr>
<tr><td>Optimizer</td><td>Adam</td><td>Adam</td></tr>
<tr><td>Learning rate</td><td>0.0001</td><td>0.0001</td></tr>
<tr><td>Cost function</td><td>cross-entropy</td><td>mean squared logarithmic</td></tr>
<tr><td>Early stopping metric</td><td>loss of validation set</td><td>loss of validation set</td></tr>
<tr><td>Early stopping patience</td><td>15</td><td>20</td></tr>
<tr><td>Batch size</td><td>10</td><td>10</td></tr>
<tr><td>Validation set</td><td>last 10%</td><td>last 10%</td></tr>
<tr><td>Batch size</td><td>10</td><td>10</td></tr>
<tr><td># Training epochs</td><td>5000</td><td>1000</td></tr>
</table>

Table 3: Parameters of convolutional neural networks

<table>
<tr><th>Parameter</th><th>ConvNet1D</th><th>FCN</th><th>ResNet</th></tr>
<tr><td># Layers</td><td>4</td><td>3</td><td>9</td></tr>
<tr><td>Layer type</td><td>1-D convolutional</td><td>2-D convolutional</td><td>2-D convolutional</td></tr>
<tr><td># Filters</td><td>(64|64|128|128)</td><td>(128|256|128)</td><td>(64|64|64|128|128|128|128|128|128)</td></tr>
<tr><td>Kernel sizes</td><td>(3|3|3|3)</td><td>$\left(\begin{pmatrix}8\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}5\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}3\\ \times\\ 1\end{pmatrix}\right)$</td><td>$\left(\begin{pmatrix}8\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}5\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}3\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}8\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}5\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}3\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}8\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}5\\ \times\\ 1\end{pmatrix} \mid \begin{pmatrix}3\\ \times\\ 1\end{pmatrix}\right)$</td></tr>
<tr><td>Normalizations ∀ convolutional layers</td><td>-</td><td>batch</td><td>batch</td></tr>
<tr><td>Pooling types</td><td>(-|max|-|global average 1-D)</td><td>(-|-|global average 2-D)</td><td>(-|-|-|-|-|-|-|-|global average 2-D)</td></tr>
<tr><td>Pooling sizes</td><td>(-|3|-|-)</td><td>(-|-|-)</td><td>(-|-|-|-|-|-|-|-|-)</td></tr>
<tr><td>Dropout values</td><td>(-|-|-|0.5)</td><td>(-|-|-)</td><td>(-|-|-|-|-|-|-|-|-)</td></tr>
<tr><td>Activation function ∀ convolutional layers</td><td>relu</td><td>relu</td><td>relu</td></tr>
<tr><td>Activation function for output layer</td><td>sigmoid</td><td>softmax</td><td>softmax</td></tr>
<tr><td>Optimizer</td><td>rmsprop</td><td>Adam</td><td>Adam</td></tr>
<tr><td>Learning rate</td><td>0.001</td><td>0.0001</td><td>0.0001</td></tr>
<tr><td>Cost function</td><td>cross-entropy</td><td>cross-entropy</td><td>cross-entropy</td></tr>
<tr><td>Early stopping metric</td><td>loss of validation set</td><td>loss of training set</td><td>loss of validation set</td></tr>
<tr><td>Early stopping patience</td><td>40</td><td>10</td><td>15</td></tr>
<tr><td>Batch size</td><td>10</td><td>10</td><td>100</td></tr>
<tr><td>Validation set</td><td>last 10%</td><td>last 10%</td><td>last 10%</td></tr>
<tr><td># Training epochs</td><td>1000</td><td>1200</td><td>1500</td></tr>
</table>

4 EVALUATION AND RESULTS

We applied dimensionality reduction techniques such as *Linear discriminant analysis*, *(Kernel) principal component analysis* and *t-Distributed Stochastic Neighbor Embedding* to attain low-dimensional visualizations of the data. Based on those visualizations, we concluded that the A-Scans are much more likely to group together with respect to the dataset they belong to instead of their respective annotations. Thus, we focus on evaluating *different* datasets (abbreviated as d_1 to d_8) stemming from the *same* person and *same* transducer position and present the corresponding results in Table 4. We state the amount of A-Scans in each dataset to emphasize the performance differences if more data is available and present the best performing DTW window size for each dataset. We only show F_1 scores for the ML methods before (in brackets) and after post-processing as the DSP techniques resulted in an average F_1 score below 70%. F_1 scores $\geq$ 80% are shown in green, F_1 scores between 70% and 80% are shown in yellow and F_1 scores < 70% are shown in red.

Table 4: Rounded F_1 scores (in %) for different methods before (in brackets) and after applying post-processing.

Train/Test dataset (# A-Scans)	*MLP*	*ConvNet1D*	*RBF*	*FCN*	*ResNet*	*DTW (w=var.)*
d_1/d_2 (1,000)	79 (79)	72 (69)	81 (81)	85 (85)	79 (67)	91 (90) (w=2)
d_2/d_1 (1,000)	77 (76)	73 (70)	69 (67)	77 (72)	57 (57)	75 (73) (w=2)
d_3/d_4 (3,000)	85 (83)	61 (59)	80 (78)	61 (57)	86 (79)	89 (83) (w=5)
d_4/d_3 (3,000)	73 (71)	77 (74)	85 (82)	75 (71)	69 (66)	88 (82) (w=2)
d_5/d_6 (10,000)	86 (85)	89 (86)	85 (84)	84 (82)	87 (80)	90 (86) (w=3)
d_6/d_5 (10,000)	90 (88)	87 (85)	89 (87)	88 (88)	89 (81)	88 (85) (w=7)
d_7/d_8 (10,000)	94 (93)	65 (65)	80 (79)	90 (89)	94 (90)	97 (96) (w=4)
d_8/d_7 (10,000)	86 (85)	97 (83)	76 (76)	87 (85)	93 (89)	95 (93) (w=5)
Average F_1 score	84 (82)	77 (74)	81 (79)	81 (79)	82 (76)	89 (86)

5 CONCLUSION AND FUTURE WORK

With an average F_1 score of up to 89 % after post-processing, the results prove the validity of our novel approach for wearable muscle contraction classifications by exploiting raw ultrasound RF data and ML methods. For comparison, it has been shown that a comparatively complex approach using a Discrete wavelet transform on sEMG signals to classify muscle contractions "results in a classification accuracy of 88.90 %" [9]. The performance of our methods improves for almost all datasets containing more A-Scans, which strongly hints at further improvements if more data is acquired. We reckon a truly robust solution would require data from a wide range of participants by taking all possible factors that might have on an effect on the signals, such as age, gender, ethnicity, fitness level, etc, into account. In the future, we expect to include additional muscle fatigue classification algorithms to enable a wearable system feasible for many gym and rehabilitation scenarios. Pending further investigations, first evaluation results show promising results with F_1 scores > 80% for selected datasets.

REFERENCES

[1] Donald J Berndt and James Clifford. 1994. Using dynamic time warping to find patterns in time series.. In *KDD workshop*, Vol. 10. Seattle, WA, 359–370.

[2] Lukas Brausch, Holger Hewener, and Paul Lukowicz. 2019. Muscle Contraction A-Scan data annotated by volunteers. Data uploaded to OpenML.org, https://www.openml.org/d/41971.

[3] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. 2019. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery* (2019), 1–47.

[4] Jing-Yi Guo, Yong-Ping Zheng, Qing-Hua Huang, and Xin Chen. 2008. Dynamic monitoring of forearm muscles using one-dimensional sonomyography system. *Journal of rehabilitation research and development* (2008).

[5] Jing-Yi Guo, Yong-Ping Zheng, Hong-Bo Xie, and Terry K Koo. 2013. Towards the application of one-dimensional sonomyography for powered upper-limb prosthetic control using machine learning models. *Prosthetics and orthotics international* 37, 1 (2013), 43–49.

[6] Jing-Yi Guo, Yong-Ping Zheng, Hong-Bo Xie, and Terry K Koo. 2013. Towards the application of one-dimensional sonomyography for powered upper-limb prosthetic control using machine learning models. *Prosthetics and Orthotics International* 37, 1 (2013), 43–49. PMID: 22683737.

[7] Nalinda Hettiarachchi, Zhaojie Ju, and Honghai Liu. 2015. A New Wearable Ultrasound Muscle Activity Sensing System for Dexterous Prosthetic Control. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*. 1415–1420.

[8] Youjia Huang, Xingchen Yang, Yuefeng Li, Dalin Zhou, Keshi He, and Honghai Liu. 2018. Ultrasound-based sensing models for finger motion classification. *IEEE journal of biomedical and health informatics* 22, 5 (2018), 1395–1405.

[9] Tanu Sharma and Karan Veer. 2016. EMG classification using wavelet functions to determine muscle contraction. *Journal of medical engineering & technology* 40, 3 (2016), 99–105.

[10] Petra Vidnerova. 2019. RBF-Keras: an RBF Layer for Keras Library. https://github.com/PetraVidnerova/rbf_keras.

[11] Zhiguang Wang, Weizhong Yan, and Tim Oates. 2017. Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*. IEEE, 1578–1585.

[12] Xingchen Yang, Xueli Sun, Dalin Zhou, Yuefeng Li, and Honghai Liu. 2018. Towards Wearable A-Mode Ultrasound Sensing for Real-Time Finger Motion Recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26, 6 (June 2018), 1199–1208.