

Research



Cite this article: Cunningham RJ, Loram ID. 2020 Estimation of absolute states of human skeletal muscle via standard B-mode ultrasound imaging and deep convolutional neural networks. *J. R. Soc. Interface* **17**: 20190715.
<http://dx.doi.org/10.1098/rsif.2019.0715>

Received: 16 October 2019
 Accepted: 6 January 2020

Subject Category:

Life Sciences—Engineering interface

Subject Areas:

biomechanics, biomedical engineering, biometrics

Keywords:

contraction, deep learning, imaging, model, muscle, ultrasound

Author for correspondence:

Ryan J. Cunningham
 e-mail: ryan.cunningham@mmu.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.4814016>.

Estimation of absolute states of human skeletal muscle via standard B-mode ultrasound imaging and deep convolutional neural networks

Ryan J. Cunningham¹ and Ian D. Loram²

¹Department of Computing and Mathematics, and ²Cognitive Motor Function Research Group, Research Centre for Musculoskeletal Science & Sports Medicine, Department of Life Sciences, Manchester Metropolitan University, Manchester, Greater Manchester M1 5GD, UK

RJC, 0000-0001-6883-6515; IDL, 0000-0001-8125-6320

The objective is to test automated *in vivo* estimation of active and passive skeletal muscle states using ultrasonic imaging. Current technology (electromyography, dynamometry, shear wave imaging) provides no general, non-invasive method for online estimation of skeletal muscle states. Ultrasound (US) allows non-invasive imaging of muscle, yet current computational approaches have never achieved simultaneous extraction or generalization of independently varying active and passive states. We use deep learning to investigate the generalizable content of two-dimensional (2D) US muscle images. US data synchronized with electromyography of the calf muscles, with measures of joint moment/angle, were recorded from 32 healthy participants (seven female; ages: 27.5, 19–65). We extracted a region of interest of medial gastrocnemius and soleus using our prior developed accurate segmentation algorithm. From the segmented images, a deep convolutional neural network was trained to predict three absolute, drift-free components of the neurobiomechanical state (activity, joint angle, joint moment) during experimentally designed, simultaneous independent variation of passive (joint angle) and active (electromyography) inputs. For all 32 held-out participants (16-fold cross-validation) the ankle joint angle, electromyography and joint moment were estimated to accuracy $55 \pm 8\%$, $57 \pm 11\%$ and $46 \pm 9\%$, respectively. With 2D US imaging, deep neural networks can encode, in generalizable form, the activity–length–tension state relationship of these muscles. Observation-only, low-power 2D US imaging can provide a new category of technology for non-invasive estimation of neural output, length and tension in skeletal muscle. This proof of principle has value for personalized muscle assessment in pain, injury, neurological conditions, neuropathies, myopathies and ageing.

1. Introduction

There is a current unmet medical demand for personalized *in vivo* skeletal muscle analysis. Muscle-related pain, injury and dysfunction represent an enormous socio-economic cost, including the cost of medical treatment, work absence and long-term decreased ability to perform activities of daily living which exceeds that estimated for heart disease, cancer or diabetes [1,2]. This need arises in conditions of pain/injury (work-related injury, neck–back–leg pain and injury), arthritic conditions, neurological conditions (dystonia, motor neuron disease), myopathies (myositis), neuropathies (nerve injury, spinal cord injury) and changes associated with ageing (motor unit loss) [3,4]. Work-related upper limb and neck musculoskeletal disorders are among the most common occupational disorders around the world [5]. Personalized assessment requires available,

non-invasive, accurate, objective measurement of function and condition for skeletal muscles throughout the body [2–4,6,7].

The mechanical function of muscle is to deliver force. Muscle comprises muscle fibres embedded within a collagenous endomysial network [8]. This dynamic three-dimensional (3D) structure, observable by ultrasound (US) as shape and texture [7], transmits muscle force along the distributed curvilinear path between the origin and insertion of each muscle [3,4]. We hypothesize that the dynamic state of skeletal muscle is encoded by the three-dimensional collagenous structure, and is observable by two-dimensional (2D) US images [3,4].

Intrinsic muscle properties are driven by two main independent inputs: active neural drive and length (origin–insertion distance), which determine the dynamic state of muscle (electronic supplementary material, figure A). The state is termed ‘neurobiomechanical’ because the state vector comprises one neural (activity) and two biomechanical (length, tension) components, defined here as (activity, joint angle, joint moment) [9]. Neural drive causes metabolically active contraction in muscle fibres. This internally generated pattern of tension contracts the internal collagenous structure, which shortens the muscle tissue and stretches the tendon tissue connecting muscle to bone. Joint angle reflects external forces (gravitational, contact, inertial) imposed on muscle. External force stretches the collagenous structure passively from outside and lengthens both the muscle and tendon. Because of the different active versus passive force transmission patterns, we hypothesize that the three components (activity, length, tension) are encoded instantaneously and independently within the collagenous structure and in a form generalizable between individuals.

If correct, this hypothesis provides the basis for a new approach to acquire information from muscle. Particularly significant is the potential to measure neural output from deep muscles. Also significant is the potential to use the intrinsic collagen-encoded activity–length–tension relationship to measure length and tension simultaneously using observation alone.

Current technology provides no general, non-invasive solution for measuring specific muscle states. Magnetic resonance (MR) allows low-frame-rate (less than 10 Hz) imaging of musculoskeletal structures in inactive supine posture or limited movement [10]. Electromyography (EMG), subject to many well-known problems [11], can measure only the neural component excluding the biomechanical (length, tension). Non-invasive, surface EMG is limited to superficial muscles, excluding general access to deep clinically important muscles in the neck, trunk and limbs. Intramuscular EMG can provide invasive measurement. Clinical neurophysiologists typically use needles and avoid deep muscles to prevent thoracic or spinal puncture. Dynamometry provides non-invasive measurement of joint moment and cannot provide muscle-specific measurement.

US allows non-invasive imaging of skeletal muscle to full anatomical depth (5–6 cm in the spine [7], up to 17 cm for the diaphragm [12]). Perturbation methods, such as supersonic shear wave imaging and shear wave elastography (SWE) [13,14], induce a shear wave and measure its propagation through muscle [13]. Using multiple assumptions, the combination of a known stress with an observed strain pattern provides non-invasive estimates of regional stiffness within cross-sectional areas of specific muscles [14,15]. There are limitations to SWE. Transmission power safety regulations limit depth of the adequate shear wave to 3–4 cm [16], which excludes the deepest muscles. SWE does not resolve active from passive force. Correlations with measured force are subjective, requiring calibration to

person-specific maximum voluntary contraction (MVC). SWE has a maximum sampling rate of 1 Hz [14]. Standard frame-rate (25–100 Hz), B-mode imaging is clinically ubiquitous, non-invasive, low cost and portable with minimal exclusion criteria, but does not reveal the tension state of skeletal muscle [16].

In summary, there is an unmet need for a non-invasive estimation of skeletal muscle states. We ask the reader to view electronic supplementary material, video S1, which shows a US recording of the calf muscles undergoing simultaneous, independent change in active and passive input. From any single image, by comparison with any image selected as a baseline, could the reader estimate the instantaneous absolute activity, joint angle and joint moment? This paper demonstrates, by proof of principle, that using deep learning (DL) and standard observation-only 2D US, three components of the dynamic neurobiomechanical state (activity, joint angle, joint moment) can be recovered directly from US images of muscle from people outside the training set.

2. Context of technical contribution

Several authors have highlighted the nonlinear relationship between muscle image features (muscle thickness, length) and a singly varying external input such as EMG or joint angle (e.g. [17–19]). With respect to skeletal muscle, research has focused on computational extraction of predefined, or partially defined, intuitive low-resolution features such as pennation angle, fascicle length, muscle shearing, fascicle curvature, muscle thickness and cross-sectional area [17,20–40]. Common limitations are lack of fully automated segmentation of muscles and features, manual initialization of analysis, confounding effect of non-muscle structures such as blood vessels and cumulative drift arising from feature-tracking methods [18,21,23,29,39–41]. Muscle is a complex 3D time-varying structure in which features leave or enter the image plane; hence, features are inherently impossible to track using pure feature-tracking methods [42,43]. Regulated tracking [23,43] is closely related to feature engineering. Some presupposition about the information content is made and a technique is developed to measure and use that information to regulate spurious tracking points. However, consistent features suitable for regulation are generally lacking.

The intrinsic encoding and estimation of more than one simultaneously varying component (activity, joint angle, joint moment) has never been demonstrated. Currently, no method has demonstrated automated robust (multiple held-out participants) generalization to new participants and no method extracts the complete state in general conditions where the inputs (activity, joint angle) vary independently.

The development [44–54] of DL provides a framework for encoding the content of US images in relation to measured data (EMG, angle, moment). DL is a technique for building artificial neural network (ANN) representations of data in a layer-wise fashion, where each layer models increasingly abstract/complex features of the data. DL facilitates modelling of complex features without *a priori* assumptions of the descriptive features. ANNs can learn nonlinear functions to map data (images) to labels (EMG, moment, joint angle). Even without many (or any) labels (which may often be the case with respect to deep muscles) features can be extracted using generative models such as restricted Boltzmann machines [44,55], deep belief networks [56], deep (variational) autoencoders/autoassociators

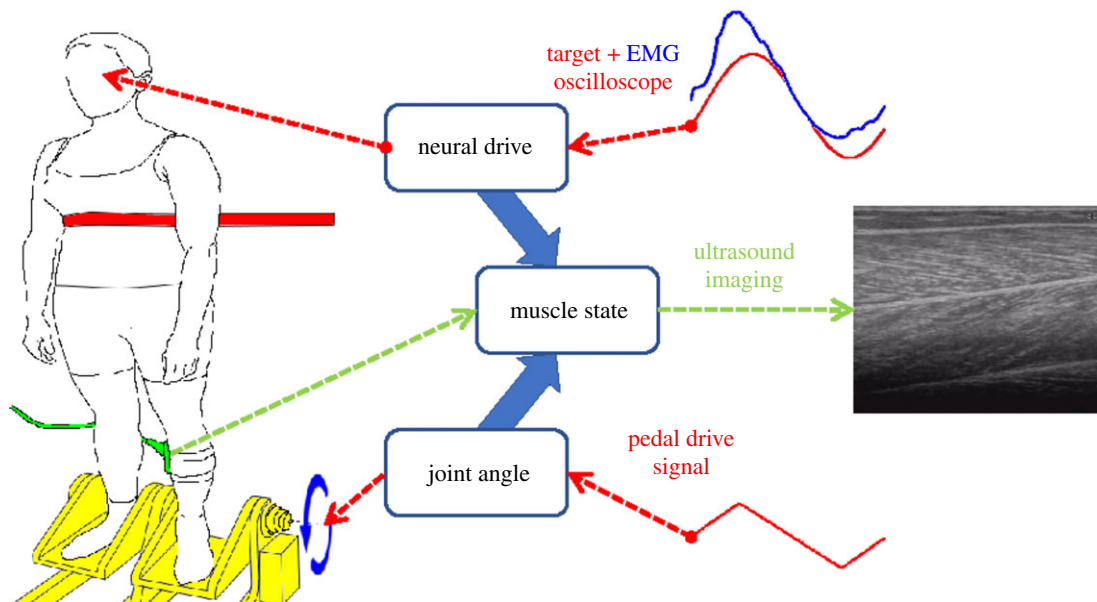


Figure 1. Experimental set-up. The participant stood upright on a foot pedal system (yellow), while strapped (red) at the chest to a backboard and observed an oscilloscope at eye level. A US probe (green) was attached to the left calf to image the gastrocnemius medialis (GM) and soleus (SO) muscles (right: greyscale). A wireless EMG sensor was attached to GM and to SO at standard locations (<http://www.seniam.org/>). By contracting their calf muscles, the participant matched the GM EMG feedback signal (blue) to the target signal (red) presented on the oscilloscope. A pedal signal (red) rotated the pedals and ankle joint angle at the rotational axis (blue arrows).

[46,57–59], or more recently generative adversarial networks (GANs) [60]. These features can be either analysed directly (using statistics or distance metrics) or re-mapped to relatively few labels. If large volumes of labelled data exist, a convolutional neural network (CNN) can be trained directly on the data to predict the labels, which can be continuous or discrete. CNNs work well for understanding the content of static images [48] or speech [61], and more recently deep residual networks (ResNet) have surpassed human-level performance [62] on the ImageNet image recognition competition. CNNs have also demonstrated the ability to track local motion [63], which means that, unlike standard feature tracking, a CNN can measure the dynamic state (two temporally different frames) of local features, while simultaneously having access to the static state (or pose) information. Use of historical states, perhaps with recurrent or long short-term memory networks, can also be valuable. For this investigation of a new hypothesis (that the instantaneous collagen structure encodes states), we avoided temporal ANN models since they cloud the issue, are comparatively difficult and time-consuming to train and would complicate generalization to different US acquisition rates (e.g. ultrafast US > 1000 Hz versus standard 25–100 Hz).

We use DL (CNNs) to map individual frames with a contextual reference frame (prior) to absolute (drift-free) states measured by other means (ankle angle, muscle EMG). This work is not a study of different CNN architectures. This work tests, for the first time, whether or not the three components of the absolute neurobiomechanical state can be estimated using DL. This work is informed by our own prior investigation [64], combined with developments in the field [62,65].

3. Methods

3.1. Experimental design and overview of methods

We test the hypothesis that US images alone contain the information required to model the state of muscle and to resolve

that state into the two independent inputs which created it. We identify the two main independent inputs as muscle activity and joint angle. We select two muscles, gastrocnemius medialis (GM) and soleus (SO), for which both inputs can be manipulated and measured to provide ground truth. We design apparatus and a protocol, which allows us to vary each input independently and simultaneously, to create a dataset of muscle US images populating the space of possible neurobiomechanical states. We investigate a form of supervised learning (CNNs) for their potential to learn, with minimal overfitting, the temporal variation of biomechanical states and we use 16-fold cross-validation providing genuine held-out test results for all 32 of our participants to test our hypothesis.

3.2. Data acquisition

Thirty-two healthy participants (seven female; ages 19–65, mean 27.5) stood upright on a programmable/controllable foot pedal system while strapped at the chest to a backboard (figure 1).

Joint angle refers to plantar-flexion/dorsiflexion of the ankle joint. The participant was restrained, maintaining a straight leg and standing flatfoot on the pedal system. Joint angle is measured by rotation of the pedals from horizontal. The calf muscles deliver force through the Achilles tendon. This force rotates the foot relative to the shin. Joint moment is the rotational effect of the combined muscle forces acting around the joint axis of rotation. For the range of motion studied, ankle moment is approximately proportional to the summed calf muscle force. We measured ankle moment using a calibrated strain gauge mounted on the under-side of the foot pedals.

EMG is the electrical activity arising from the active contraction of the muscle fibres. Surface rather than intramuscular electrodes provide the best global measure of muscle activity [66,67] and so this electrical activity is recorded from the skin surface above the muscle using electrodes (Trigno, Delsys Inc., USA). We recorded electromyographs from the GM muscle belly, EMG from SO at its medial superficial location, ankle joint angle and ankle joint moment, all at 1000 Hz. EMG data were rectified and low-pass filtered to below 10 Hz.

Using a US scanner (Aloka Prosound SSD 4000+; probe 7.5 MHz, width × depth 5.9 × 5.5 cm), we imaged GM and SO

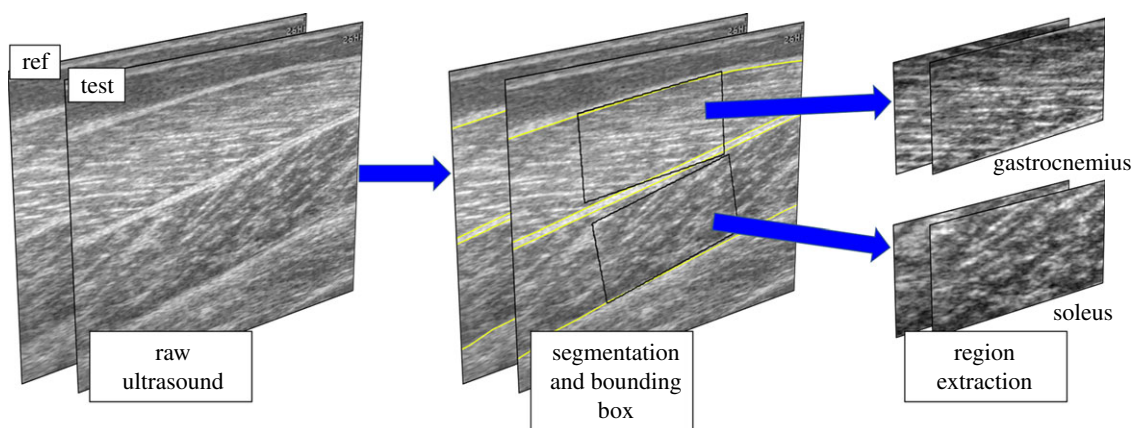


Figure 2. Image segmentation and region extraction pipeline. The fully automatic region extraction process occurs prior to neural network training and testing. Left: a pair of US images is shown, where the ref image represents the *reference* image; the *reference* image accompanies a test image as part of a single input to the neural network. Middle: yellow lines show the automatic segmentation of individual muscles, from which bounding boxes (black rectangles) are positioned on the centroid oriented orthogonal to the main axis of each muscle. Right: a 128×256 -pixel region is extracted from the raw images. For each muscle, the label vector (muscle EMG, joint angle, joint moment) equals the values temporally aligned to the *test* frame, subtracted from the values temporally aligned to the *reference* frame.

simultaneously in their longitudinal fascicle plane. The imaging location, angle and depth were chosen to include within-plane fascicle collagen content of both muscles, but optimized for GM. The probe was strapped to the participant to maintain constant location during movement [18]. US was recorded at 25 Hz using a frame grabber (DT 3120; Data Translation). We used Simulink (Matlab, R2013a; The MathWorks Inc., Natick, MA) to interface with the laboratory equipment (pedal system and EMG). For video synchronization, a hardware trigger was used to initiate the start of each trial.

3.3. Tasks

Three distinct tasks were designed to explore the state–function space of muscle.

3.3.1. Isometric

The pedal system was fixed at a neutral angle (flat feet), and participants observed an analogue oscilloscope. On the oscilloscope, we displayed, side by side, a dot representing the amplitude of their filtered GM EMG signal, and a dot representing the amplitude of a fabricated (target) signal (see §3.2). Participants were asked to contract their calf muscles by pushing down their toes in order to match their EMG with the target signal, while simultaneously keeping their foot in full contact with the static pedals.

3.3.2. Passive

Participants observed an analogue oscilloscope. On the oscilloscope, we displayed, side by side, a dot representing the amplitude of their filtered GM EMG signal and a dot representing the zero amplitude target. Participants were asked to monitor and minimize any EMG activity by relaxing their muscles. The pedal system was driven using a fabricated signal (see §3.4). Participants were asked to allow their ankle to rotate and keep their feet in full contact with the moving pedals.

3.3.3. Combined

The pedal system was fixed at a neutral angle (flat feet) and participants observed an analogue oscilloscope. On the oscilloscope, we displayed, side by side, a dot representing the amplitude of their filtered GM EMG signal and a dot representing the amplitude of a fabricated (target) signal (see §3.4). The pedal system was simultaneously driven using a different fabricated signal (see §3.4). Participants were asked to contract their calf muscles

by pushing down their toes in order to match their EMG with the target signal, while simultaneously keeping their foot in full contact with the static pedals.

Trials were 190 s in length; this consisted of 10 s of neutral standing (i.e. no signals were used to move the pedals or the dot on the screen), followed by 180 s of trial. Data ranges are shown in §4.3.

3.4. Designing the labels

Two signals (active contraction, passive joint rotation) were designed to manipulate the two independent muscle inputs. Both signals were derived from the following bases.

Active contraction (target dot on the screen to guide calf muscle contraction). (1) For the first 10 s signal *a* was used, and every 10 s thereafter we alternated between signals *a* and *b*. (2) After 30 s signal *c* was used, and every 30 s thereafter either signal *a* or *b* was used depending on the first rule.

Passive joint rotation (pedal angle). (1) For the first 20 s signal *a* was used, and every 20 s thereafter we alternated between signals *a* and *b*. (2) After 60 s signal *c* was used, and every 60 s thereafter either signal *a* or *b* was used depending on the first rule. The signals were designed to produce transient correlations, de-correlations and anti-correlations to maximize exploration of the muscle input state space. The correlation of the two signals was $r = 0.33$, $p = 0$ (Pearson) and $r = 0.34$, $p = 0$ (Spearman). Pedal rotations ranged from 9.32° dorsiflexion to 13.79° plantar-flexion with respect to the neutral angle. Pedal velocities were distributed from approximately -10 to $+10^\circ \text{ s}^{-1}$ (electronic supplementary material, figure B).

$$\left. \begin{aligned} a &= \sin\left(0.4t\pi - \frac{\pi}{2}\right), \\ b &= \sin\left(0.5t\pi - \frac{\pi}{2}\right) \end{aligned} \right\} \quad (3.1)$$

and

$$c = \sin\left(\sin\left(t\frac{\pi}{30} - \frac{\pi}{2}\right)30\pi - \frac{\pi}{2}\right).$$

3.5. Segmentation and region extraction

To map muscle-specific EMG to a muscle-specific image, we extracted regions of superficial (GM) and deep (SO) muscle tissue (figure 2). Analysis of restricted, standardized regions enabled us to maximize the spatial resolution while reducing the computational dimensions and complexity. First, an expert (R.J.C.) annotated the internal boundaries of the medial GM and

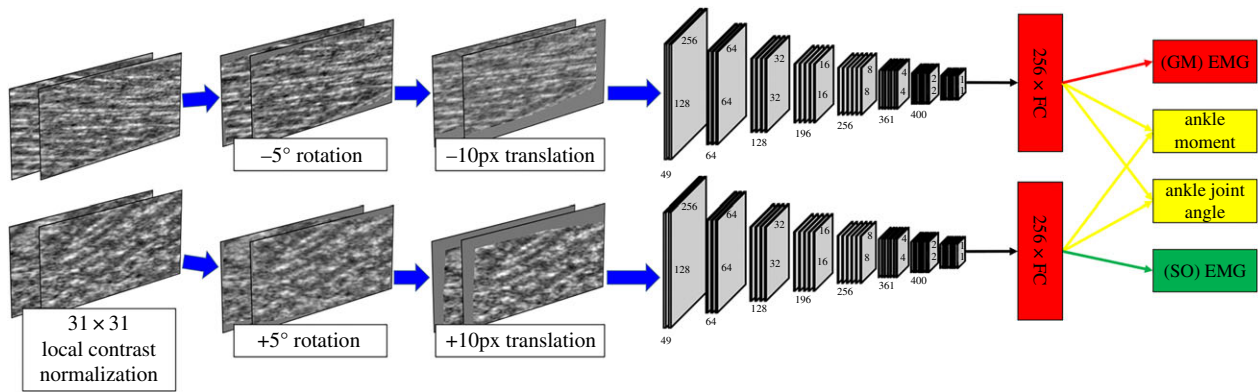


Figure 3. Data augmentation and CNN pipeline. Online data augmentation. First, we normalize contrast by removing mean and variance, sampling from patches of ± 15 pixels. Then we sample a random rotation parameter ($\pm 5^\circ$) per muscle and rotate each pair of muscle images independently. Finally, we sample a random translation parameter (± 10 pixels) per muscle and translate each pair of muscle images independently. CNN pipeline: two CNNs with shared parameters and a dense layer (FC) encode each muscle separately. In the final layer, four linear units connect to the FC layers of each CNN: the ankle moment and ankle joint angle units are fully connected to the units of both CNNs, while the GM EMG unit is only connected to the CNN which receives input from the GM image. Also, the SO EMG unit is only connected to the CNN which receives input from the SO image. (Online version in colour.)

SO muscles in 500 randomly selected images, of which 100 were selected randomly for testing. After interpolating the annotations to a standard 80-point vector, a principal component model was constructed from the remaining 400 images. An active shape model [68], constructed from just 10 principal components ($> 99\%$), was used to guide a heuristic search with a large profile search range (± 30 pixels about the contour). No initial segmentation was used; the increased profile range was an ample aid to match distant contours. The search was conducted at full resolution ± 10 pixels about each contour point. For more details, see [7]. The entire dataset ($> 400\,000$ images) was segmented.

To standardize the image input, we extracted a rectilinear region ($x \times y = 256 \times 128$ pixels $\approx 29.4 \times 14.6$ mm; figure 2) about the centroid of each muscle, orthogonal to the main axis of the muscle. The main axis was calculated as the linear least-square fit to mean segmentation over the whole trial sequence. This region captures the muscle tissue rather than the tendon, which connects muscle to bone.

3.6. Neural network architecture

Our primary concern was to choose an architecture for which the model was large enough to minimize adequately the training error. The second main concern was to maximize generalization and minimize computation. Using our previous experience [64], our strategy was to train a large model, with state-of-the-art regularization (dropout) in multiple layers, while evaluating performance on held-out validation data. Unlike our original study [64], we additionally address the deep muscle (SO), and this decision inspired a CNN architecture wherein a CNN model is created and applied per muscle, and the weights were shared between models as a regularizer (figure 3). We are not the first to use this type of architecture [69], though the application of it is novel.

The learning objectives are different for each of the muscle-specific CNN models. The muscle-specific EMG should be predicted from the relevant muscle to ensure that the estimation of muscle activity came from the target muscle. However, the ankle angle and ankle moment should be estimated from the states of both muscles combined. To meet this objective, information/gradient flow was gated using a binary mask for the relevant learning objectives (figures 2 and 3). Data augmentation was implemented online at the input to each CNN (figure 3).

3.7. Online data augmentation

To aid convergence as well as generalization, local contrast normalization (LCN) was applied to each image via a graphics

processing unit (GPU) with a local field of 31 square pixels (figure 3). During training, to help prevent overfitting and account for (intra/inter)-participant variation in muscle region extraction, linear transformations (rotation and translation) were applied to the reference image and the target image per muscle. The same transformation was applied between reference and target images, but not between muscles (i.e. each muscle had its own transformation). Rotations and translations were randomly sampled from a uniform distribution between -5° and 5° and -16 and 16 horizontal/vertical pixels, respectively. Transformations were carried out on a GPU using linear interpolation with no extrapolation, where zeros filled the extra pixels (figure 3).

3.8. Training and cross validation

The input to the neural network was a pair of images, a reference image (frame 1) and a target image (any frame within a trial), and the labels. The labels for each input were the test frame-aligned EMG, ankle angle and ankle moment minus the reference frame-aligned EMG, ankle angle and ankle moment. Labels were normalized to unit standard deviation. The learning objective was to predict the difference in the independent states of the muscles between two images. To reduce bias in the input channel of the network corresponding to the reference image, we doubled the training set by swapping the reference and target frames. We randomly sampled reference and target frames to increase further the size of the training set. The final training set contained over 1 million pairs of images.

To train our models, we minimized the mean absolute error (MAE) between the model output and the normalized labels (EMG, joint moment, joint angle) using adaptive moment estimation (ADAM: [70]) with $\alpha = 0.999$ and $\beta = 0.9$ and a learning rate of 5×10^{-5} . To prevent saturating units [64], exponential linear units were used in all layers except the output layer, which was linear. Prior to training, all biases were initialized to 0, and all weights were initialized using

$$\text{var}(w) = \frac{2}{fan_in},$$

where w is the normal distributed weight vector of a single unit/node and fan_in is the size of the input vector to that unit.

The train, test and validation errors were measured periodically during training to allow selection of optimal models using test and validation errors. We used 16-fold cross-validation: for each fold unique combinations of validation and test data were used to assess performance of the model within the fold. Within

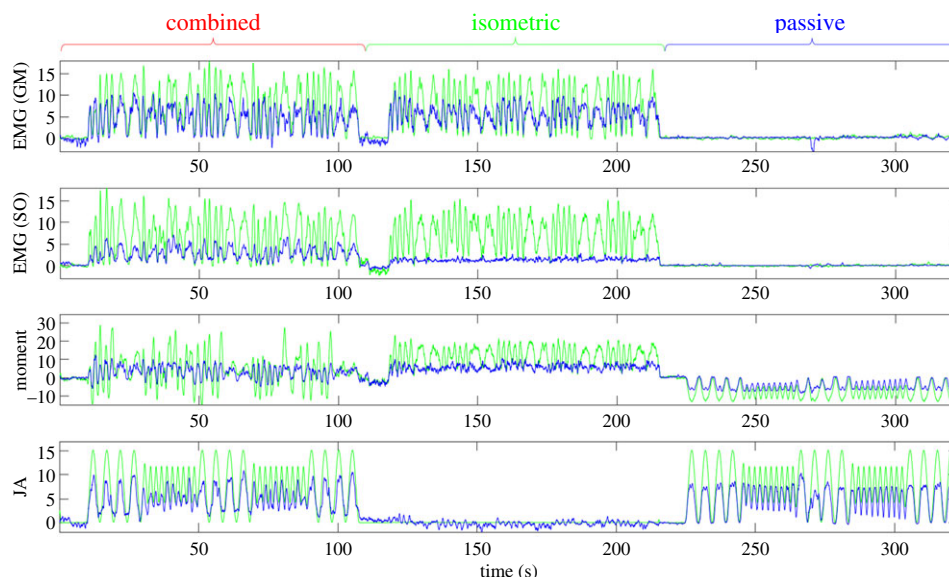


Figure 4. Representative example of neural network performance. Each panel compares neural network output (blue) with labels (green) for a single participant over all three trial conditions (*combined*: left, *isometric*: middle, *passive*: right). Units are mV, Nm and degrees for EMG, moment and joint angle (JA), respectively, relative to the reference frame for each section (*combined*, *isometric*, *passive*).

each fold, a test set of one held-out participant (approx. 12 500 samples) and a validation set of one held-out participant (approx. 12 500 samples) was created (where none of the test/validation participants were used in training). For each fold, the validation set was used to choose the optimal test model and the test set was used to select the optimal validation model. This process yielded 16 unique neural networks with genuine held-out results for all 32 participants. To regularize our models, we used a dropout scheme similar to [65], where dropout was applied to every layer with larger dropout rates in layers closer to the output of the model. That dropout strategy circumvented the need to try variations of dropout, requiring repeated training and model evaluation as in [64]. As additional regularization and detection of convergence, early stopping was used where the model with the lowest test/validation error was taken after both test and validation errors did not decrease for more than eight error evaluations.

Errors are reported using MAE of all samples (i of n) for each signal. To report accuracy of estimate Y in the context of time-varying signal y , we use

$$\text{Accuracy} = 100 - \text{SMAPE},$$

where SMAPE (symmetric mean absolute percentage error) is

$$\text{SMAPE} = \frac{100\%}{n} \sum_{i=1}^n \frac{|Y_i - y_i|}{(|Y_i| + |y_i|)/2}.$$

3.9. Software and tools

All ANN and segmentation software was developed from first principles by R.J.C. using C/C++ and CUDA-C (NVIDIA Corporation, California). No libraries other than std CUDA libraries (runtime version 8.0 `cuda.h`, `cuda_runtime.h`, `curand.h`, `curand_kernel.h`, `cuda_occupancy.h`, and `device_functions.h`), the C++ 11 std library and OpenMP were used.

4. Results

4.1. Segmentation and region extraction

Our requirement for segmentation is accuracy within the dataset rather than generalization. Cropping well within the muscle

boundary and averaging over the sequence requires millimetre rather than sub-millimetre accuracy. For the 100 randomly selected test images, segmentation agreed with the manual annotations to 0.3 mm^2 (approx. 99% intersection over union) and segmented at approximately 10 images per second.

4.2. Representative neural network output

The CNN estimates the neurobiomechanical state of a muscle (EMG, moment, angle) for each frame independently of all other frames (except the reference). The extended representative sequence from one participant shown in figure 4 illustrates there is no drift in the estimated state components.

Increased activity shortens muscle, stretches the tendon and increases tension, whereas positive joint rotation (plantar-flexion) shortens the muscle passively and decreases muscle (and tendon) tension. Activity and positive joint rotation both reduce strain in (shorten) the collagen structure but have opposite effects on tension in the same collagen structure. A key question is whether the CNN can resolve these independent active and passive changes in muscle.

Consider figure 4; here we describe, qualitatively, the accuracy of the estimate in terms of the temporal pattern of states between frames, the local timing of the estimated versus actual pattern, the scale and the bias of the local pattern. Figure 4 shows that the CNN captures the pattern of GM EMG during isometric and combined conditions and correctly shows no activity during passive joint rotation. The CNN estimates change in GM activity independently from joint rotation in isometric, passive and combined conditions. However, the scale is too small, giving an estimated GM EMG of approximately 50% of the true signal. For the deep muscle SO, the CNN distinguishes activity (*combined*, *isometric*) from the passive inactive condition as elevated bias: the CNN also captures correctly the pattern of EMG activity during combined conditions; however, the scale of the estimate is substantially too low. The CNN correctly estimates isometric activity as increasing moment, positive joint rotation as decreasing moment, and alternating sign of joint moment during combined conditions. However, again, the scale of the estimate is

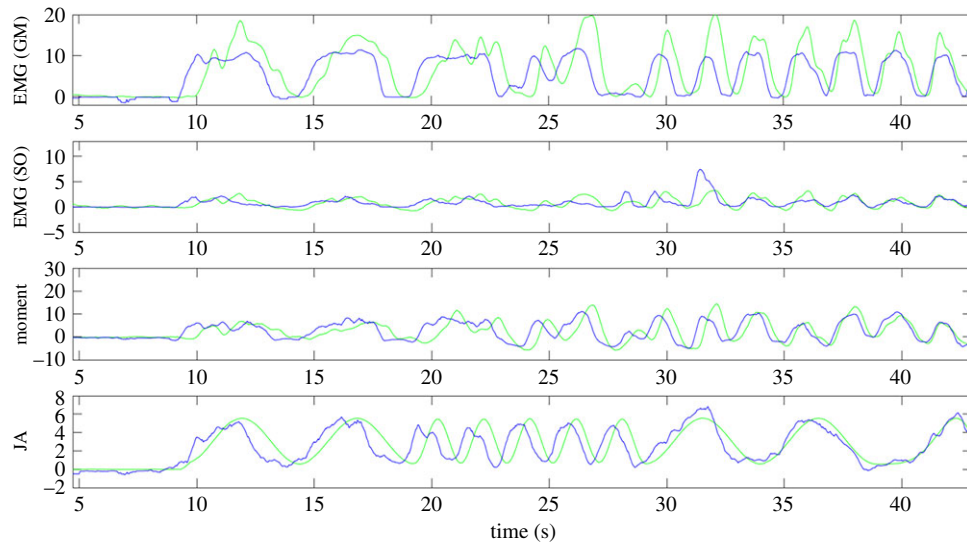


Figure 5. Zoom portion of a representative participant. The zoom shows approximately 40 seconds of contiguous data from a single participant during a trial constructing combined, independent modulations of ankle joint angle and GM/SO EMG. The figure illustrates how the neural network has separated and modelled the four independent signals in a representative case. The participant differs from that shown in figure 4. Units are mV, Nm and degrees for EMG, moment and joint angle (JA), respectively. (Online version in colour.)

Table 1. Summary of neural network test results. This summarizes the results presented in electronic supplementary material, table A. We present the mean, standard deviation, median, minimum and maximum values for symmetric mean absolute accuracy ($100 - \text{SMAPE}$) and mean absolute error (MAE), calculated over all 32 participants, for each of the four labels: EMG (GM/SO), ankle joint moment and ankle joint angle (JA). Samples are images per participant: the sum reports all images tested from the complete dataset of 32 participants.

metric	samples	100 – SMAPE				MAE			
		EMG				EMG			
		GM (%)	SO (%)	moment (%)	JA (%)	GM (mV)	SO (mV)	moment (Nm)	JA (°)
mean	12 594	56.90	45.91	46.97	54.57	3.06	2.61	6.11	2.50
standard deviation	2613	11.02	11.06	8.87	7.94	1.02	1.57	3.04	1.28
median	13 944	59.37	45.33	47.60	53.82	2.88	2.18	5.07	2.19
min	8058	35.61	24.26	27.49	43.33	0.59	0.55	2.58	1.06
max	14 971	72.45	69.06	61.54	67.30	5.84	7.56	15.97	6.17
sum	403 023	—	—	—	—	—	—	—	—

too low. The CNN captures the pattern of joint rotation and absence of rotation, but at a scale that is too low.

Figure 5 illustrates the ability of the CNN to extract simultaneous, independent changes in components of the neurobiomechanical state in both the superficial and the deep muscles. The CNN captures correctly the distinct, independent pattern and scale of EMG in GM and SO muscles. The CNN captures correctly the distinct, independent patterns and scales of joint rotation, activity and joint moment. While pattern and scale are correct in all quantities, there is a temporal error, decreasing through time, between US image-derived estimates and the synchronized electrically recorded signals. This temporal error reduces the accuracy reported for these CNN estimates.

To summarize, the CNN separates the independent signals. During passive conditions the neural network was robust at predicting little to no active EMG, but a good proportion of passive motion. During isometric conditions, the CNN

predicted little to no passive motion, but a good proportion of active EMG. The reported accuracy is adversely affected mostly by errors in amplitude of prediction and also by some time-varying temporal misalignment.

4.3. Summary of neural network performance

For each muscle, and using just a single frame referenced to a common baseline frame, the neural network estimated meaningful values of all three signals (EMG, joint moment, joint angle). The MAE values for GM EMG, SO EMG, joint moment and joint angle were 3.1 ± 1.0 mV, 2.6 ± 1.6 mV, 6.1 ± 3.0 Nm and $2.5 \pm 1.3^\circ$, respectively (table 1). In context, these errors represent 0.58, 0.56, 0.54 and 0.71 of the standard deviation of each signal, and 0.0056, 0.029, 0.056, 0.11 of the functional range of each signal where the ranges were 50 mV, 90 mV, 109 Nm and 23° , respectively (table 2). These results summarize $12\,600 \pm 2600$ (mean \pm s.d.) samples

Table 2. Label statistics. Table shows the distribution of labels recorded over all 32 participants. Negative joint angle represents dorsiflexion (decrease in angle between foot and shin referenced from flatfoot, i.e. 90°). Positive angle represents plantar-flexion (increase in angle between foot and shin referenced from 90°).

	EMG			
	GM (mV)	SO (mV)	moment (Nm)	JA (°)
mean	5.62	4.67	16.92	1.24
standard deviation	5.32	4.70	11.30	3.51
median	2.74	2.87	15.23	0.66
min	0.0026	0.0021	-14.77	-9.32
max	50.21	89.58	94.35	13.79

tested per participant and 403 023 samples in total, tested using cross-validation.

Accuracy, as a percentage of the time-varying signal, was $56.9 \pm 11\%$, $45.9 \pm 11\%$, $47.0 \pm 8.9\%$ and $54.6 \pm 7.9\%$, respectively (table 1). Performance varied across all 32 participants, with coefficients of variation in accuracy of 19%, 24%, 19% and 15%, respectively (table 1). There were no individuals for whom estimation of any signal failed (i.e. accuracy < 0 or equivalently error larger than signal). Results for all 32 individuals are shown in electronic supplementary material, table A.

Some very high individual accuracies were recorded for GM EMG with three participants at over 70% and 13 at over 60%. There were only four participants with accuracy less than 40%. While accuracy was lowest for SO EMG, accuracy for several participants approached 70%. Estimation of ankle angle was the most stable, with a high average accuracy of 55% and the lowest standard deviation at 7.9%.

5. Discussion

5.1. The main finding

For two muscles, this investigation tests the hypothesis that three components (activity, length, tension) of the dynamic muscle state are encoded instantaneously within the 3D collagenous structure and are observable in generalizable form by 2D US images. We used novel data collection to generate 403 023 images from 32 participants containing independent and combined modulation of passive joint rotation and active neural input to the muscle state. We used deep CNNs to test whether the complex nonlinear muscle state is encoded in 2D US images. Our results from 32 genuinely held-out participants reveal that, to approximately 50% accuracy (table 1), the absolute values of EMG of each of the superficial and deep muscles, the ankle joint angle and the ankle joint moment are each encoded objectively, in generalizable form, in 2D US images of the GM and SO calf muscles.

5.2. Technical discussion

This article demonstrates the successful application of CNNs to predict continuous state variables rather than classify objects. This is US medical imaging analysis of physiological function rather than anatomical structure. A survey of the literature

indicates many papers applying CNNs to classification but few applying CNNs to regression. There are only a few published investigations of the architectures, hyperparameters and results of the application of CNNs to regression and of the application of CNNs to modelling of complex systems such as skeletal muscle. Our use of DL to test a scientific hypothesis is novel. If the CNN can encode the neurobiomechanical state from US images of muscle, it demonstrates that the state is encoded objectively within that muscle tissue.

This application is challenging and hence the neural network's ability to predict the absolute labels is surprisingly robust. The prediction of active and passive states from single images removes the possibility of signal drift. The CNN resolves active and passive states even when they vary independently and incoherently (figure 5, EMG (GM) and JA at 20–45 s). The encoding of this information in the US image is not trivial or intuitive. To illustrate the achievement, observe the motion of the muscle in electronic supplementary material, video S1, and note that from a *single frame* and a *reference frame* the active and passive states of each muscle are estimated to approximately 50% accuracy.

The necessity of the reference frame is to reduce bias and promote generalization between participants. We anticipate a reference frame would not be necessary to generalize to new motions within the same participant. Within a participant, generalization is relevant to prosthetics, where a system is trained on person-specific actions to control a prosthetic limb [32,33]. Our existing evidence suggests that a within-participant system would be more accurate than a general system [71].

Several sources of error limit the accuracy of this generalized system to approximately 50% (table 1).

- (i) Temporal misalignment between US images and measured signals occurs and is variable (figure 5, cf. 8–10 s and 40–45 s). EMG and joint moment signals contain more of their power at higher frequencies and are thus more sensitive to temporal misalignment. Substantial temporal misalignment is a known phenomenon related to image timing in clinical US interfaces [72]. While clinical US interfaces such as the one we used give no control over image timing, accurate image timing is possible using low-level US systems [72]. For this study, misalignment means the reported accuracy is less than it would be if temporal alignment were always correct (figure 5).
- (ii) Un-encoded variation between participants: contrary to our hypothesis, some variation between participants is intrinsically not encoded within the US images. Variation between participants in limb strength, muscle mass, muscle cross-sectional area, depth of fat layer, electrical electrode–skin–muscle impedance, electrode placement and location of the foot on the footplate are not available to the CNN. These variables alter the mapping between the US image and measured signals (EMG, joint angle, joint moment). Our results provide the first benchmark of the generalizable content of US images. Injecting some prior knowledge into the system such as anthropometric data may improve accuracy. Within a participant, generalization would avoid these issues and may show higher accuracy.
- (iii) Imperfect ground truth: the measured signals have limited accuracy. Imperfect placement of EMG

electrodes, partial sampling of the whole muscle volume, crosstalk from adjacent muscles, electrical noise and interference, foot placement, slight knee flexion, heel raising and toe curling limit accuracy of EMG, joint angle and joint moment signals.

- (iv) US probe placement: US acquires a single plane from a 3D muscle structure. The extent to which generalization depends upon the specific plane imaged is uncertain. More data are required to test the effect of variation in probe placement.

Given the uncertainties listed above, it remains remarkable that, from a single 2D US image referred to a baseline image, absolute values of each component of the neurobiomechanical state can be estimated on new participants to a benchmark accuracy of approximately 50%.

Our empirical demonstration provides the first proof of principle that the estimation of specific muscle states in deep muscle is possible in general conditions of combined/isolated active and passive changes. The deep SO gave EMG accuracy comparable within 10% to the superficial GM. While we placed the probe to acquire GM and SO within a single image, one plane is not optimal for both muscles. We chose to optimize for GM. Thus, our benchmark for SO represents a lower limit to what is possible for this muscle.

5.3. Application to other muscles

In this investigation, we chose muscles for their suitability to test our hypothesis that muscle states are encoded in their collagen structure and are observable by US. This investigation required the muscles to be observable by US and required the main inputs of EMG and passive joint rotation to be measurable. This investigation also required it to be possible to control the two inputs experimentally to produce a dataset covering the space of single and simultaneous combined, independent variation of neural activity and passive joint rotation. The calf muscles are a relatively well-understood muscle group with access to control and measure these labels.

However, the significance of this investigation lies in the potential to measure neural output more generally from deep muscles. The significance lies also in the potential to measure tension generally from individual muscles rather than joint moments. Here the CNN was trained to predict joint moment, which is measurable. However, the prediction was derived from an image of muscle, which means that information, in essence muscle tension, is encoded within the muscle.

Following our test of principle, the practical question becomes how to acquire the labels and training data to train a system on more general muscle groups? In principle, though with greater practical difficulty, it is possible to record EMG with needles/wires from deep muscles inaccessible via surface electrodes. This could give training labels for muscle in a complex system like the neck, back or forearm. It is also possible to use unsupervised learning (like Bayesian GAN [73]) on a large data collection of US only, and reserve supervised learning for a smaller data collection of US with EMG labels. Another possibility is to measure dynamometric or kinematics signals (e.g. head torque), and predict those signals directly from the image, in such a way that the network learns a spatial localization mapping from the labels to the image (like class activation mapping [74]). Combined with an accurate segmentation [7,75], an activity map of generated

head force could provide an estimated muscle-specific contribution to gross head rotational force.

5.4. Scientific and clinical significance

For two muscles (GM, SO), we present the first generalized prediction of independent components of the neurobiomechanical state of skeletal muscle (activity, joint rotation, joint moment) directly from standard frame-rate (25 Hz) 2D, B-mode US under general conditions of independently varying inputs (figures 4 and 5). This result reports a scientific discovery. Previously, it was unknown whether muscle tissue encodes, simultaneously, the activity, origin–insertion length and tension state of the muscle. Skeletal muscle tissue is relatively generic between muscles. The US images of GM and SO are similar to many or most muscles viewed longitudinally in the plane of their fascicles. Generally, muscles connect to bone via a series tendon; however, specific architecture differs between muscles. GM and SO are pennate (fibres are at an angle to the force-generating axis), whereas some muscles are parallel (fibres are parallel to the force-generating axis, e.g. biceps). Thus, the specific encoding of activity, origin–insertion length and tension state within the collagen structure will differ between muscles. We predict that the discovery of state encoding in the collagen structure will generalize to other muscles. We predict that specific systems should be trained on specific muscles, although some generalization between muscles may be possible.

This scientific discovery has technological significance. *Muscle activity* is an amplified version of neural output or motor command delivered by peripheral nerves from the spinal cord to the muscle [76]. Non-invasive measurement of activity in deep muscles is currently impossible. In science, we need activity, particularly from deep muscles in the neck, back, lower and upper limbs, to understand how control of the muscular system is organized. Muscular control is hierarchical and synergistic in nature, and currently that science is immature simply because we cannot measure activity easily in all the important deep muscles [6]. In medicine, healthy control of muscles breaks down for many possible reasons. In myopathies or injury, the muscle is inflamed or diseased and delivers inadequate output from neural drive. In neuropathies, neural drive to the muscle is inadequate, for example from breaks in the peripheral nerves or spinal cord injury, or from demyelination of upper (central) or lower (peripheral) motor neurons. In neurological conditions control of muscles is disordered, e.g. dystonia, Parkinson's disease or cerebellar ataxia, resulting in abnormal patterns and timing of activity. Measurement of activity can discriminate myopathies from neuropathies from neurological conditions and localize the impairment. In the clinic, staff rely routinely on manual palpation and very rarely on needle EMG since that is an expert skill in very short supply. Our US approach, developed to its potential, could provide easy discrimination between these conditions and assessment of abnormal muscle activity. In rehabilitation, prostheses are controlled where possible using available activity signals from muscles. While it is possible to control prostheses directly from brain interfaces, peripheral neural output provides better quality signals since these are already pre-processed motor signals, and muscle activity is simply an amplified version of peripheral neural output [76]. There is current interest in using wearable US to drive prosthetic devices [32,33].

5.4.1. Muscle tension

Measurement of individual muscle force requires a strain gauge to be inserted surgically in series with the individual muscle, typically in the tendon joining that muscle to bone [77]. For surgery, orthopaedics, rehabilitation and biomechanics the force of individual muscles is needed to determine the contribution that individual muscles make to joint moments and joint stability as to whether the balance between muscles is correct and whether surgical correction, physiotherapy or altered training is required. In this study, we have validated the CNN estimation of the force state using a joint moment. However, the estimate was derived from muscle tissue, and thus our results demonstrate the principle that force can be estimated directly from muscle tissue. This study provides objective evidence to justify surgical implantation of strain gauges to provide muscle-specific force labels.

5.4.2. Muscle stiffness

Currently, observational US can only measure muscle strain, and because stress is unknown US cannot measure mechanical properties such as force or stiffness. The results published here reveal that analysis of the full biomechanical state (length and tension) is possible using observational imaging. Muscle force and stiffness are a consequence of intrinsic muscle properties operating on the inputs (neural command and joint rotation). The proof of concept demonstrated in this paper is that multi-layered neural networks with DL methods (convolutions, pooling, dropout, etc.) can model directly from US images the intrinsic muscle properties and the independent inputs, which together determine the mechanical output. This result is possible because the collagenous structure of skeletal muscle is observable, and also because muscle activity and passive joint rotation create different patterns of strain within the structure [9]. Force generated internally by activity within individual motor units has a different strain pattern from the force transmitted externally into the muscle between origin and insertion. Bypassing human preconception, ANNs can learn those dynamic nonlinear patterns and provide spatio-temporal representations of the muscle state for our scientific and diagnostic benefit.

These results imply that perturbation methods (e.g. SWE) may not be required to measure the biomechanical state. In practice, further development will be required to translate this proof of principle into technology applicable to all muscles of medical interest. Standard US machines are more available and cheaper than shear wave imaging machines and they input less acoustic power to the patient. For dynamic structures as complex as skeletal muscle, data-driven modelling of muscle properties using DL should be more accurate than using generic

stress–strain relationships and assumptions of material properties to interpret shear wave velocity maps.

6. Conclusion

Currently, there is an unmet need for technology to provide non-invasive assessment of skeletal muscle state in general conditions. Limitations in current technology (EMG, dynamometry, SWE) mean that many important muscles (e.g. deep muscles in the neck, back, thorax/abdomen and limbs) are inaccessible to full diagnostic analysis. This paper demonstrates an approach which can contribute new assessment of the muscle system.

We have presented a novel experiment for the generation of hundreds of thousands of accurately labelled muscle US images for modelling functional muscle states using US. We have demonstrated that skeletal muscle encodes three components of the neurobiomechanical state within its tissue structure, observable by US. We have presented the first generalized prediction of muscle-specific EMG, joint angle and joint moment from standard frame-rate B-mode 2D US images. Existing methods rely on simple measures in isolated cases (isometric only, or passive only) which do not generalize. We have demonstrated the efficacy of CNNs to this domain, which encourages the application of DL to skeletal muscle US. This approach has potential applications for clinical assessment, monitoring of treatment, biofeedback for behavioural therapy and interfacing with prosthetics in a large range of conditions of substantial socioeconomic impact as stated in the Introduction.

Ethics. These experiments were approved by the Research Ethics Committee of the Faculty of Science and Engineering, Manchester Metropolitan University (MMU), Manchester, UK. Participants gave (written) informed consent to these experiments, which conformed to the standards set by the latest revision of the Declaration of Helsinki. Experiments were performed at the Cognitive Motor Function Laboratory, in the School of Healthcare Science at MMU.

Data accessibility. This paper is accompanied by electronic supplementary material media (videos) and documents. The documents ‘Supplementary Material’ and ‘AllParticipants’ give descriptions of all supplementary material and present full results for all participants, respectively.

Authors’ contributions. This study was planned and the manuscript was prepared by R.J.C. and I.D.L. Data collection was conceived by R.J.C. and I.D.L. and conducted by R.J.C. DL code/analysis was written by R.J.C. Further analyses were performed by R.J.C. and I.D.L. R.J.C. and I.D.L. interpreted analysis and results. R.J.C. and I.D.L. gave final approval for publication.

Competing interests. We declare that we have no competing interests

Funding. During this project, R.J.C. was funded by the School of Healthcare Science, Manchester Metropolitan University.

Acknowledgements. We thank all the participants who gave their time generously. With appreciation, we thank Des Richards for his work building the bespoke foot pedal apparatus for our experiment.

References

- Gaskin DJ, Richard P. 2012 The economic costs of pain in the United States. *J. Pain* **13**, 715–724. (doi:10.1016/j.jpain.2012.03.009)
- Falla D, Hodges PW. 2017 Individualized exercise interventions for spinal pain. *Exerc. Sport Sci. Rev.* **45**, 105–115. (doi:10.1249/JES.000000000000103)
- Maurits NM, Bollen AE, Windhausen A, De Jager AEJ, Van Der Hoeven JH. 2003 Muscle ultrasound analysis: normal values and differentiation between myopathies and neuropathies. *Ultrasound Med. Biol.* **29**, 215–225. (doi:10.1016/S0301-5629(02)00758-5)
- Pillen S, van Alfen N. 2011 Skeletal muscle ultrasound. *Neurol. Res.* **33**, 1016–1024. (doi:10.1179/1743132811Y.0000000010)
- Hoe VCW, Urquhart DM, Kelsall HL, Sim MR. 2012 Ergonomic design and training for preventing work-related musculoskeletal disorders of the upper limb

- and neck in adults. *Cochrane Database Syst. Rev.* **8**, CD008570.
6. Loram ID, Bate B, Harding P, Cunningham R, Loram A. 2017 Proactive selective inhibition targeted at the neck muscles: this proximal constraint facilitates learning and regulates global control. *IEEE Trans. Neural Syst. Rehabil. Eng.* **4320**, 357–369. (doi:10.1109/TNSRE.2016.2641024)
 7. Cunningham RJ, Harding PJ, Loram ID. 2017 Real-time ultrasound segmentation, analysis and visualisation of deep cervical muscle structure. *IEEE Trans. Med. Imaging* **36**, 653–665. (doi:10.1109/TMI.2016.2623819)
 8. Gillies AR, Lieber RL. 2011 Structure and function of the skeletal muscle extracellular matrix. *Muscle Nerve* **44**, 318–331. (doi:10.1002/mus.22094)
 9. Davis J, Kaufman KR, Lieber RL. 2003 Correlation between active and passive isometric force and intramuscular pressure in the isolated rabbit tibialis anterior muscle. *J. Biomech.* **36**, 505–512. (doi:10.1016/S0021-9290(02)00430-X)
 10. Deligianni X, Pansini M, Garcia M, Hirschmann A, Schmidt-Trucksäss A, Bieri O, Santini F. 2017 Synchronous MRI of muscle motion induced by electrical stimulation. *Magn. Reson. Med.* **77**, 664–672. (doi:10.1002/mrm.26154)
 11. Türker K. 1993 Electromyography: some methodological problems and issues. *Phys. Ther.* **73**, 698–710. (doi:10.1093/ptj/73.10.698)
 12. Sarwal A, Walker FO, Cartwright MS. 2013 Neuromuscular ultrasound for evaluation of the diaphragm. *Muscle and Nerve* **47**, 319–329. (doi:10.1002/mus.23671)
 13. Song P, Zhao H, Manduca A, Urban MW, Greenleaf JF, Che S. 2012 Comb-push ultrasound shear elastography (CUSE): a novel method for two-dimensional shear elasticity imaging of soft tissues. *IEEE Trans. Med. Imaging* **31**, 1821–1832.
 14. Hug F, Tucker K, Gennison JL, Tanter M, Nordez A. 2015 Elastography for muscle biomechanics: toward the estimation of individual muscle force. *Exerc. Sport Sci. Rev.* **43**, 125–133. (doi:10.1249/JES.000000000000049)
 15. Dieterich AV, Andrade RJ, Le Sant G, Falla D, Petzke F, Hug F, Nordez A. 2016 Shear wave elastography reveals different degrees of passive and active stiffness of the neck extensor muscles. *Eur. J. Appl. Physiol.* **117**, 171–178. (doi:10.1007/s00421-016-3509-5)
 16. Ryu J, Jeong WK. 2017 Current status of musculoskeletal application of shear wave elastography. *Ultrasonography* **36**, 1–13. (doi:10.14366/usg.13047)
 17. Hodges PW, Pengel LHM, Herbert RD, Gandevia SC. 2003 Measurement of muscle contraction with ultrasound imaging. *Muscle Nerve* **27**, 682–692. (doi:10.1002/mus.10375)
 18. Loram IIDD, Maganaris CN, Lakin M. 2006 Use of ultrasound to make noninvasive *in vivo* measurement of continuous changes in human muscle contractile length. *J. Appl. Physiol.* **100**, 1311–1323. (doi:10.1152/jappphysiol.01229.2005)
 19. Shi J, Zheng YP, Chen X, Huang QH. 2007 Assessment of muscle fatigue using sonomyography: muscle thickness change detected from ultrasound images. *Med. Eng. Phys.* **29**, 472–479. (doi:10.1016/j.medengphy.2006.07.004)
 20. Maganaris CN, Baltzopoulos V. 1999 Predictability of *in vivo* changes in pennation angle of human tibialis anterior muscle from rest to maximum isometric dorsiflexion. *Eur. J. Appl. Physiol. Occup. Physiol.* **79**, 294–297. (doi:10.1007/s004210050510)
 21. Darby J, Hodson-Tole EF, Costen N, Loram ID. 2012 Automated regional analysis of B-mode ultrasound images of skeletal muscle movement. *J. Appl. Physiol.* **112**, 313–327. (doi:10.1152/jappphysiol.00701.2011)
 22. Li JZ, Zhou Y, Yang YF, Chen B, Wang L, Zheng YP. 2012 Estimation of longitudinal muscle motion using a primal-dual algorithm. In *Proc. 5th 2012 Biomedical Engineering Int. Conf., Ubon Ratchathani, Thailand, 5–7 December 2012*. New York, NY: IEEE.
 23. Darby J, Li B, Costen N, Loram I, Hodson-Tole E. 2013 Estimating skeletal muscle fascicle curvature from B-mode ultrasound image sequences. *IEEE Trans. Biomed. Eng.* **60**, 1935–1945. (doi:10.1109/TBME.2013.2245328)
 24. Kwah LK, Pinto RZ, Diong J, Herbert RD. 2013 Reliability and validity of ultrasound measurements of muscle fascicle length and pennation in humans: a systematic review. *J. Appl. Physiol.* **114**, 761–769. (doi:10.1152/jappphysiol.01430.2011)
 25. Rana M, Hamarneh G, Wakeling JM. 2013 3D fascicle orientations in triceps surae. *J. Appl. Physiol.* **115**, 116–125. (doi:10.1152/jappphysiol.01090.2012)
 26. Herbert R, Heroux M, Diong J, Bilston L, Gandevia S, Lichtwark G. 2014 Changes in the length and three dimensional orientation of muscle fascicles and aponeuroses with passive length changes in human gastrocnemius muscles. *J. Physiol.* **2**, 1–34. (doi:10.1113/jphysiol.2014.279166)
 27. Wakeling JM, Randhawa A. 2014 Transverse strains in muscle fascicles during voluntary contraction: a 2D frequency decomposition of B-mode ultrasound images. *Int. J. Biomed. Imaging* **2014**, 1–9. (doi:10.1155/2014/352910)
 28. Cunningham RJ, Harding PJ, Loram ID. 2017 Deep residual networks for quantification of muscle fiber orientation and curvature from ultrasound images. In *Annu. Conf. on Medical Image Understanding and Analysis, Edinburgh, UK, 11–13 July 2017*, pp. 63–73. Cham, Switzerland: Springer.
 29. Rana M, Wakeling JM. 2011 In-vivo determination of 3D muscle architecture of human muscle using free hand ultrasound. *J. Biomech.* **44**, 2129–2135. (doi:10.1016/j.jbiomech.2011.05.026)
 30. Cunningham R, Sánchez M, May G, Loram I. 2018 Estimating full regional skeletal muscle fibre orientation from B-mode ultrasound images using convolutional, residual, and deconvolutional neural networks. *J. Imaging* **4**, 29. (doi:10.3390/jimaging4020029)
 31. Ikegawa S, Funato K, Tsunoda N, Kanehisa H, Fukunaga T, Kawakami Y. 2008 Muscle force per cross-sectional area is inversely related with pennation angle in strength trained athletes. *J. Strength Cond. Res.* **22**, 128–131. (doi:10.1519/JSC.0b013e31815f2fd3)
 32. Castellini C, Passig G, Zarka E. 2012 Using ultrasound images of the forearm to predict finger positions. *IEEE Trans. Neural Syst. Rehabil. Eng.* **20**, 788–797. (doi:10.1109/TNSRE.2012.2207916)
 33. Sierra Gonzalez D, Castellini C. 2013 A realistic implementation of ultrasound imaging as a human-machine interface for upper-limb amputees. *Front. Neurobot.* **7**, 17. (doi:10.3389/fnbot.2013.00017)
 34. Rutherford OM, Jones DA. 1992 Measurement of fibre pennation using ultrasound in the human quadriceps *in vivo*. *Eur. J. Appl. Physiol. Occup. Physiol.* **65**, 433–437. (doi:10.1007/BF00243510)
 35. Häkkinen K, Keskinen KL. 1989 Muscle cross-sectional area and voluntary force production characteristics in elite strength- and endurance-trained athletes and sprinters. *Eur. J. Appl. Physiol. Occup. Physiol.* **59**, 215–220. (doi:10.1007/BF02386190)
 36. Herbert RD, Moseley AM, Butler JE, Gandevia SC. 2002 Change in length of relaxed muscle fascicles and tendons with knee and ankle movement in humans. *J. Physiol.* **539**, 637–645. (doi:10.1113/jphysiol.2001.012756)
 37. Hoang PD, Gorman RB, Todd G, Gandevia SC, Herbert RD. 2005 A new method for measuring passive length-tension properties of human gastrocnemius muscle *in vivo*. *J. Biomech.* **38**, 1333–1341. (doi:10.1016/j.jbiomech.2004.05.046)
 38. Zhou Y, Zheng YP. 2008 Estimation of muscle fiber orientation in ultrasound images using revolving hough transform (RVHT). *Ultrasound Med. Biol.* **34**, 1474–1481. (doi:10.1016/j.ultrasmedbio.2008.02.009)
 39. Rana M, Hamarneh G, Wakeling JM. 2009 Automated tracking of muscle fascicle orientation in B-mode ultrasound images. *J. Biomech.* **42**, 2068–2073. (doi:10.1016/j.jbiomech.2009.06.003)
 40. Cronin NJ, Carty CP, Barrett RS, Lichtwark G. 2011 Automatic tracking of medial gastrocnemius fascicle length during human locomotion. *J. Appl. Physiol.* **111**, 1491–1496. (doi:10.1152/jappphysiol.00530.2011)
 41. Namburete ALL, Rana M, Wakeling JM. 2011 Computational methods for quantifying *in vivo* muscle fascicle curvature from ultrasound images. *J. Biomech.* **44**, 2538–2543. (doi:10.1016/j.jbiomech.2011.07.017)
 42. Yeung F, Levinson SF, Parker KJ. 1998 Multilevel and motion model-based ultrasonic speckle tracking algorithms. *Ultrasound Med. Biol.* **24**, 427–441. (doi:10.1016/S0301-5629(97)00281-0)
 43. Yeung F, Levinson SF, Fu D, Parker KJ. 1998 Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh. *IEEE Trans. Med. Imaging* **17**, 945–956. (doi:10.1109/42.746627)
 44. Mayraz G, Hinton GE. 2002 Recognizing handwritten digits using hierarchical products of experts. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 189–197. (doi:10.1109/34.982899)
 45. LeCun Y, Bengio Y, Hinton G. 2015 Deep learning. *Nature* **521**, 436–444. (doi:10.1038/nature14539)

46. Kingma DP, Welling M. 2013 Auto-encoding variational Bayes. (<https://arxiv.org/abs/1312.6114>)
47. LeCun Y, Boser BE, Denke JS, Henderson D, Howard RE, Hubbard WE, Jackel LD. 1990 Handwritten digit recognition with a back-propagation network. In *Advances in neural information processing systems 2* (ed. DS Touretzky), pp. 396–404. Cambridge, MA: MIT Press.
48. Krizhevsky A, Sutskever I, Hinton GE. 2012 Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems 25* (eds F Pereira, JC Burges, L Bottou, KQ Weinberger), pp. 1097–1105. Cambridge, MA: MIT Press.
49. He K, Zhang X, Ren S, Sun J. 2015 Deep residual learning for image recognition. *Arxiv.Org* **7**, 171–180.
50. Nair V, Hinton GE. 2010 Rectified linear units improve restricted Boltzmann machines. In *Proc. 27th Int. Conf. Mach. Learn., Haifa, Israel, 21–24 June 2010*, pp. 807–814. International Machine Learning Society.
51. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. 2014 Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958.
52. Noh H, Hong S, Han B. 2015 Learning deconvolution network for semantic segmentation. In *Proc. of the IEEE Int. Conf. on Computer Vision, Santiago, Chile, 11–18 December 2015*, pp. 1520–1528. Piscataway, NJ: IEEE Computer Society.
53. Hinton GE. 2002 Training products of experts by minimizing contrastive divergence. *Neural Comput.* **14**, 1771–1800. (doi:10.1162/089976602760128018)
54. Higgins I, Matthey L, Glorot X, Pal A, Uria B, Blundell C, Mohamed S, Lerchner A. 2016 Early visual concept learning with unsupervised deep learning. (<https://arxiv.org/abs/1606.05579>)
55. Smolensky P. 1986 Information processing in dynamical systems: foundations of harmony theory. *Parallel Distrib. Process. Explor. Microstruct. Cogn.* **1**, 194–281.
56. Hinton GE, Dayan P, Frey BJ, Neal RM. 1995 The ‘wake-sleep’ algorithm for unsupervised neural networks. *Science* **268**, 1158–1161. (doi:10.1126/science.7761831)
57. Plaut DC, Hinton GE. 1987 Learning sets of filters using back-propagation. *Comput. Speech Lang.* **2**, 35–61. (doi:10.1016/0885-2308(87)90026-X)
58. Masci J, Meier U, Cireşan D, Schmidhuber J. 2011 Stacked convolutional auto-encoders for hierarchical feature extraction. In *International Conference on Artificial Neural Networks*, pp. 52–59. Berlin, Germany: Springer.
59. Hinton GE, Salakhutdinov RR. 2006 Reducing the dimensionality of data with neural networks. *Science* **313**, 504–507. (doi:10.1126/science.1127647)
60. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. 2014 Generative adversarial nets. In *Advances in neural information processing systems 27* (eds Z Ghahramani, M Welling, C Cortes, ND Lawrence, KQ Weinberger), pp. 2672–2680. Cambridge, MA: MIT Press.
61. Hinton G *et al.* 2012 Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Process. Mag.* **29**, 82–97. (doi:10.1109/MSP.2012.2205597)
62. He K, Zhang X, Ren S, Sun J. 2015 Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In *Proc. of the IEEE Int. Conf. on Computer Vision, Santiago, Chile, 11–18 December 2015*, pp. 1026–1034. Piscataway, NJ: IEEE Computer Society.
63. Dosovitskiy A, Fischer P, Ilg E, Hausser P, Hazirbas C, Golkov V, Van Der Smagt P, Cremers D, Brox T. 2015 FlowNet: learning optical flow with convolutional networks. In *Proc. of the IEEE Int. Conf. on Computer Vision, Santiago, Chile, 11–18 December 2015*, pp. 2758–2766. Piscataway, NJ: IEEE Computer Society.
64. Cunningham RJ, Harding PJ, Loram ID. 2017 The application of deep convolutional neural networks to ultrasound for modelling of dynamic states within human skeletal muscle. (<https://arxiv.org/abs/1706.09450>)
65. Clevert D-A, Unterthiner T, Hochreiter S. 2015 Fast and accurate deep network learning by exponential linear units (ELUs). (<https://arxiv.org/abs/1511.07289>)
66. Vieira TM, Loram ID, Muceli S, Merletti R, Farina D. 2012 Recruitment of motor units in the medial gastrocnemius muscle during human quiet standing: is recruitment intermittent? What triggers recruitment? *J. Neurophysiol.* **107**, 666–676. (doi:10.1152/jn.00659.2011)
67. Vieira TMMM, Loram ID, Muceli S, Merletti R, Farina D. 2011 Postural activation of the human medial gastrocnemius muscle: are the muscle units spatially localised? *J. Physiol.* **589**, 431–443. (doi:10.1113/jphysiol.2010.201806)
68. Cootes TF, Taylor CJ, Cooper DH, Graham J. 1995 Active shape models-their training and application. *Computer Vision and Image Understanding* **61**, 38–59. (doi:10.1006/cviu.1995.1004)
69. Bromley J, Guyon I, Lecun Y, Säckinger E, Shah R. 1993 Signature verification using a ‘Siamese’ time delay neural network. In *Advances in neural information processing systems 6* (eds JD Cowan, G Tesauro, J Aspector), pp. 737–744. Cambridge, MA: MIT Press.
70. Kingma DP, Ba J. 2014 Adam: a method for stochastic optimization. (<https://arxiv.org/abs/1412.6980>)
71. Cunningham R. 2015 *The application of B-mode ultrasonography for analysis of human skeletal muscle*. Manchester, UK: Manchester Metropolitan University.
72. Miguez D, Hodson-Tole EFF, Loram I, Harding PJJ. 2017 A technical note on variable inter-frame interval as a cause of non-physiological experimental artefacts in ultrasound. *R. Soc. open sci.* **4**, 170245. (doi:10.1098/rsos.170245)
73. Saatchi Y, Wilson AG. 2017 Bayesian GAN. In *Advances in neural information processing systems 30* (eds I Guyon, UV Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan, R Garnett), pp. 3622–3621. Cambridge, MA: MIT Press.
74. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. 2016 Learning deep features for discriminative localization. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, Las Vegas, NV, 27–30 June 2016*, pp. 2921–2929. New York, NY: IEEE.
75. Cunningham R, Sánchez MB, Loram ID. 2019 Ultrasound segmentation of cervical muscle during head motion: a dataset and a benchmark using deconvolutional neural networks. (doi:10.31224/osf.io/fsa3c)
76. Jiang N, Falla D, d’Avella A, Graitmann B, Farina D. 2010 Myoelectric control in neurorehabilitation. *Crit. Rev. Biomed. Eng.* **38**, 381–391. (doi:10.1615/CritRevBiomedEng.v38.i4.30)
77. Ravary B, Pourcelot P, Bortolussi C, Konieczka S, Crevier-Denoix N. 2004 Strain and force transducers used in human and veterinary tendon and ligament biomechanical studies. *Clin. Biomech.* **19**, 433–447. (doi:10.1016/j.clinbiomech.2004.01.008)