Automated analysis of neck muscle boundaries for cervical dystonia using ultrasound imaging and deep learning

Ian Loram*, *Member, IEEE*, Abdul Siddique, María B. Sánchez, Pete Harding, Monty Silverdale Christopher Kobylecki, and Ryan Cunningham*, *Member, IEEE*

***Abstract*— *Objective:* To provide an automated visualization, pattern analysis and discrimination of neck muscle boundaries comparing cervical dystonia with healthy controls. *Methods:* We recorded transverse cervical ultrasound (US) images and whole-body motion analysis of sixty-one standing participants (35 cervical dystonia, 26 age matched controls). We manually annotated 3,100 images sampling a range of pitch, yaw, and roll head movements (2,000 this dataset, 1,100 previous dataset of 28 healthy participants), and trained, validated and tested a deep residual deconvolutional neural network (DCNN) to classify pixels to 13 categories (five paired neck muscles, skin, ligamentum nuchae, vertebra). For all participants in their normal standing posture, we generated held out DCNN segmented US images, extracted segmentation boundaries, transformed the boundaries to principal components and clustered the dystonia participants into two groups using their component scores. *Results:* For all segments, metrics of segmentation accuracy were Jaccard Index (50±21%) and Hausdorff Distance (5.8±4 mm). For the four deep muscle layers, their boundaries were used to predict central injection sites for each muscle with average precision 94±2%. Using 10-fold cross-validation to select predictive components, linear analysis of component scores identified two significant eigen-patterns discriminating Dystonia from Controls and classifying group membership (Dystonia1, Dystonia2, Control) correctly at 93.4%. Motion analysis showed the groups differed significantly according to head yaw-rotation posture and truncal posture. *Conclusion:* Neck muscle shape alone discriminates dystonia from healthy controls. *Significance:* Using deep learning, US imaging allows online, automated visualization, pattern analysis, diagnostic assessment of cervical dystonia and segmentation of individual muscles for targeted injection.**

***Index Terms*—Deep learning, ultrasound imaging, cervical dystonia, segmentation, muscle boundaries, diagnosis.**

Submitted 30/04/2019. This work was funded by The Dystonia Society grant ("A clinical tool for real-time analysis and visualization of cervical muscles for cervical dystonia"), by Manchester Metropolitan University (MMU) for R Cunningham and by Salford Royal Foundation Trust (SRFT) Clinical Research Fellowship for A Siddique. Loram, Cunningham, Sánchez are with Cognitive Motor Function Research Group, Research Centre for Musculoskeletal Science & Sports Medicine, MMU, M1 5GD. Siddique, Kobylecki, Silverdale are with Dept. of Neurology, SRFT, M6 8HD. Kobylecki, Silverdale are with Division of Neuroscience and Experimental Psychology, University of Manchester, Oxford Road Manchester M13 9PT. Harding is with Elements Technology Platforms Ltd, Sheffield, S1 2NS. *Correspondence: (i.loram@mmu.ac.uk, Ryan.Cunningham@mmu.ac.uk).

I. Introduction

Cervical Dystonia (CD), also called *spasmodic torticollis*, is a painful condition in which the neck muscles contract involuntarily, causing the head to twist, turn, and pull into an abnormal posture. This neurological movement disorder affects an estimated 18,000 adults in the UK [1]. The reported mean duration from symptom onset to diagnosis is 44 months, with consultations sought from a mean of 3.5 different healthcare providers before reaching a diagnosis and receiving effective therapy [2]. For CD, the diagnosis is based on expert clinical assessment since laboratory testing and imaging of the brain or spine is typically unrevealing [2].

Treatment of CD is symptomatic and the established method of choice is injecting the neck muscles with botulinum

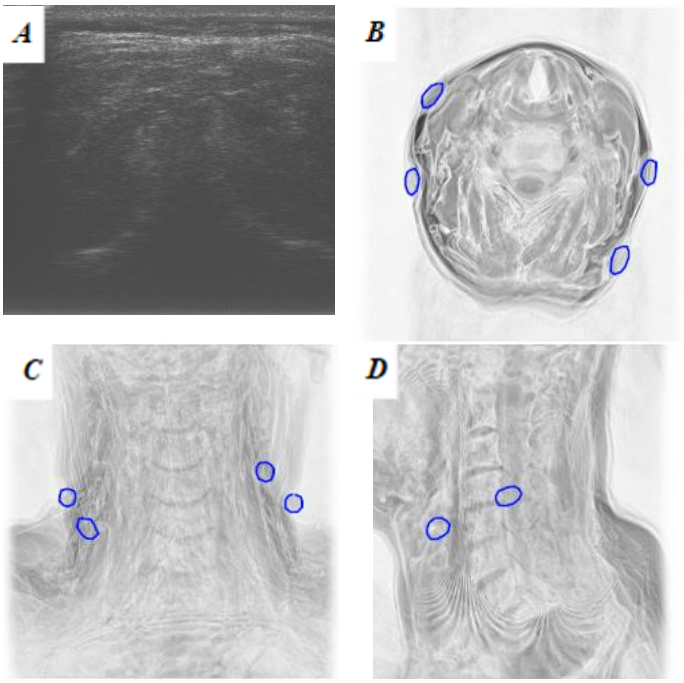

Figure 1. Linked US-MRI dataset
A. Representative axial ultrasound image targeted at vertebral level C4. The probe plane was marked with four cod liver oil capsules. ***B, C, D***: Axial, coronal and sagittal MR images of the same participant showing the ultrasound image plane marked by four cod liver oil capsules (blue circles). Here, the ultrasound image plane intersects the vertebrae between C3 and C4. N.B. Note the challenging nature of the ultrasound image to extract muscle boundaries. For the whole dataset, images were acquired at level 3.8±0.6 (mean±SD).

neurotoxin (BoNT) [3]. Clinical experience shows the main causes for treatment failure are suboptimal neck muscle selection or BoNT dosing, indicating the importance of appropriate targeting of overactive muscles [4], [5]. Furthermore, monitoring the effectiveness of treatment is confounded by use of differing rating scales and assessment methods [6]. There is a clinical need to diagnose CD more promptly, to improve analysis and identification of dystonic muscles, to improve delivery of injection and dose to specific muscles, to provide objective recording of injection sites for retention within medical records and to track longitudinally the effect of injections on individual muscles [7].

A. *Current clinical methods*

The most common method of identifying and injecting the muscles involved in CD is clinical examination and manual needle placement, based on the clinician's knowledge of functional anatomy of the neck muscles, directed by head position, shoulder elevation and assessment of muscle tone and hypertrophy on palpation [8]. Within the clinic, this method may be efficacious for superficial neck muscles but is compromised for deeper muscles - typically the deepest neck muscles (e.g. spinalis cervicis, multifidus) are rarely injected and therefore left undiagnosed and untreated. There is increasing awareness of the clinical relevance of deep cervical muscles in the pathogenesis and potential therapy of CD [9], [10], but the tools to assess and treat these muscles are currently not fully developed. Compared with intramuscular electromyographic (iEMG) mapping of cervical muscle activity, the sensitivity of clinical examination has been reported as 59% and the specificity 75% [8]. The positive predictive value of shoulder elevation and muscle hypertrophy is reportedly only 70% and head position does not provide added value, because individuals with solitary dystonic head postures do not have muscle dystonia following simple patterns [8]. Without iEMG mapping, 41% of dystonic muscles would not be recognized and 25% of inactive muscles would be judged dystonic [8]. However, iEMG is time consuming, requires substantial expertise, is invasive and cannot be performed in individuals on anticoagulants.

Ultrasound (US) offers non-invasive visualisation of muscle structures with easy contralateral comparison, is readily available, and improves the precision of injections [11], [12]. However, use of US requires training, is dependent upon operator expertise, and remains subjective [11].

The objective of this study is to apply deep learning to US imaging to provide an automated, objective visualization, and pattern analysis of neck muscle boundaries and to test whether this analysis has diagnostic value discriminating cervical dystonia from healthy controls. If successful, these methods demonstrate proof of concept for a clinical tool for objective online diagnosis, injection guidance and monitoring, with minimal requirement for operator expertise and minimal burden on clinical time.

B. *Contribution of this study*

Through the application of deep learning, the ability to extract information from limited quality images (c.f. Fig. 1A) is progressing rapidly. Application to US is under-developed and application to skeletal muscle is rare. This study, builds upon previous work by our group realizing the scientific and clinical value of *in-vivo* skeletal muscle analysis [13]–[18], applying deep learning to skeletal muscle US [19]–[21] and

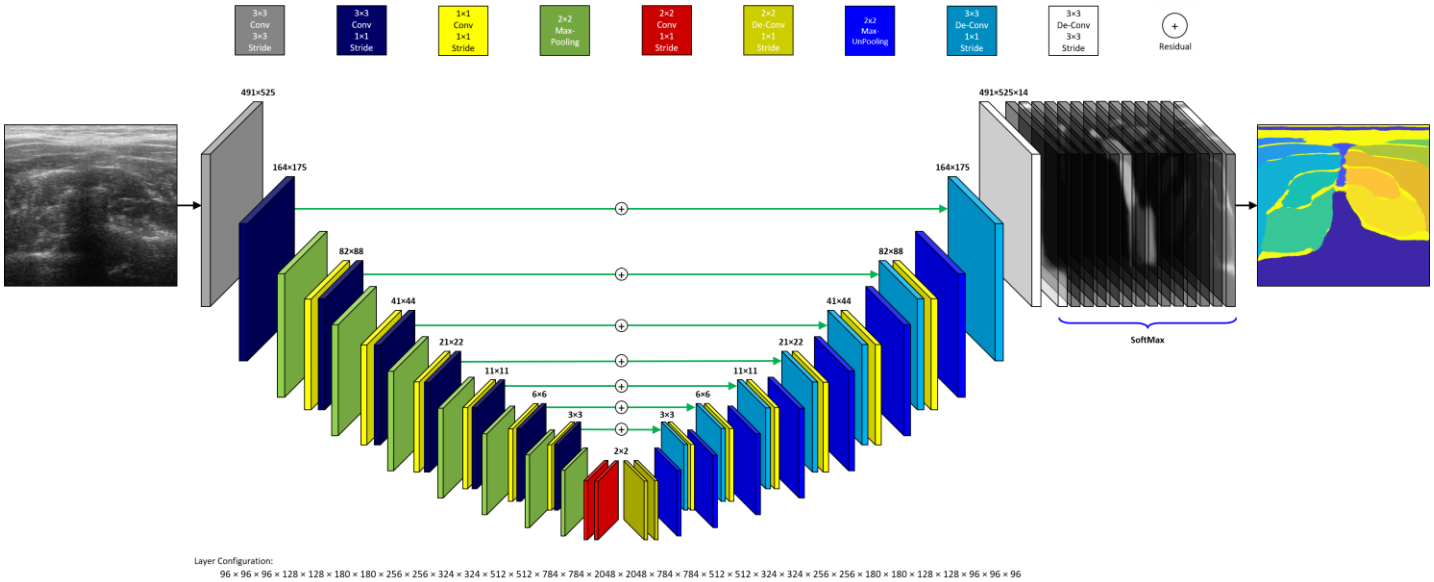

Figure 2. Residual deconvolution network architecture. This diagram shows a residual encode-decoder network with a spatial SoftMax classification layer. On the far left, the raw ultrasound image is input to the encoder network. The first convolutional layer uses a stride of 3 and every convolution layer thereafter uses a stride of 1. The network contains 32 convolutional layers and 14 pooling layers. The pattern of layers is 2 convolutional layers, then a MaxPooling layer. In the decoder part of the network the pattern is the mirror opposite of the encoder network – an UnPooling layer followed by 2 deconvolutional layers. Alternating deconvolutional layers had residual (identity) connections to their reference convolutional layers (shown by green arrows in the diagram). The output layer of the network is a 14 channel SoftMax layer at full image resolution (each pixel has a 14-unit SoftMax function). The argmax of each SoftMax function reveals the predicted classification of each pixel. All convolutional/deconvolutional layers (except output layer) used the Exponential Linear Unit (ELU) nonlinearity. This network has 32,092,992 free parameters and 19,872,665 ELU units.

specifically developing methods for analysis of the neck muscles [22], [23]. Recently we contributed a dataset, a methodology for labelling training images suitable for participants with involuntary head movement, and a benchmark deep learning method for segmenting the neck muscles [23], [24]. Here we apply those methods to the investigation of cervical dystonia.

In this study, we collect a dataset of 61 participants (35 cervical dystonia, 26 age matched controls) in standing posture demonstrating their range of head yaw, pitch and roll rotations. We use our established methods [23] to label a sample of 2,000 US images from this dataset and to train, validate and test a deep deconvolution neural network to segment five paired muscles, vertebra, ligamentum nuchae and skin. Our primary interest is to establish whether from a single axial image, neck muscle shape alone is diagnostic of cervical dystonia. Our secondary interest is whether segmentation is accurate enough to guide injections to deep neck muscles.

We test three hypotheses: (i) segmentation of deep muscles will identify injection points within the designated muscle, (ii) cervical dystonia manifests identifiable, abnormal patterns of muscle shape and (iii) muscle shape alone discriminates significantly, cervical dystonia from healthy controls.

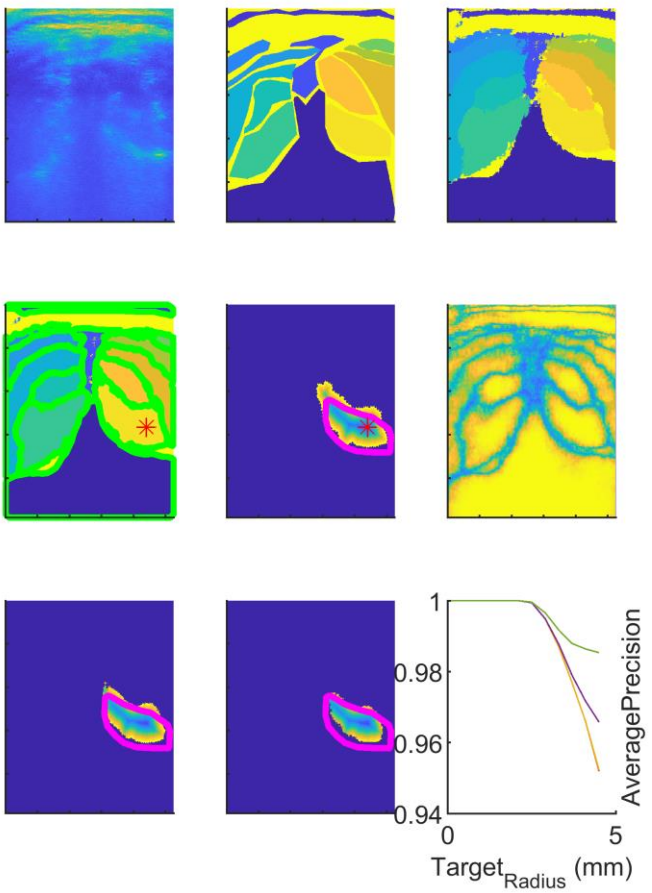

Figure 3. Boundary Extraction and Injection Point location
Figure shows an example of the visualisation and analysis provided by the neural network approach. This participant had cervical dystonia.
Top row: **Left**: ultrasound image. **Middle**: Manually defined labels. **Right**: neural network predicted classification of pixels.
2nd row: **Left:** boundaries (green) estimated around each segment. **Middle**: Injection point (red star) defined as pixel most distant from predicted segment boundary. Colour spectrum blue to yellow shows decreasing distance from predicted boundary and hence increasing target margin around injection point. Boundary of right multifidus label (green). **Right**: confidence of the neural network (yellow = high confidence) in classifying pixels.
3rd row: **Left and middle**: as 2nd row middle, but predicted pixels shown restricted to greater than 60 and greater than 80% respectively. **Right**: Average precision of pixels within target region for right multifidus. Lines show pixels restricted to confidence greater than 0, 20, 40, 60, 80% resp.

The clear visualisation of predicted segment (top right) will be appreciated by clinicians. The visualisation of confidence (2nd row right) gives the user feedback regarding good probe location and orientation.

II. Methods

A. Data collection

Using a probe (7.5 MHz, SonixTouch, Ultrasonix, USA) held transversely to the posterior neck, US images were recorded from 61 adults (35 cervical dystonia, 26 age matched controls, with respective age 61±10, 59±14 years and sex 15, 19 male) while standing and while performing head rotation tasks defining their range of pitch, yaw and head rotation. These experiments, performed in the Cognitive Motor Function laboratory, Dept. of Healthcare Science, Manchester Metropolitan University (MMU), received ethical approval from the NHS Health Research Authority (REC: 15/NW/0016, IRAS:169803) and from MMU Science and Engineering Faculty Ethics Committee. The study was conducted in accordance with the Declaration of Helsinki guidelines. All values are reported as mean±S.D. unless stated otherwise.

Linked US-MRI dataset: Participants stood upright, observing a monitor at 1m distance, just below eye level. Three or more or more axial US images of the posterior neck targeted at level C4 were recorded, each with renewed probe placement (Fig. 1A). Four cod liver oil capsules were taped (using Transpore medical tape) to the neck, two either side of the neck approximately in the image plane of the probe. The probe was removed, leaving the capsules in place, and an MRI scan (0.3T open MRI scanner, G-Scan, Esaote, Italy) was obtained with participants lying supine on the scanning bed and their neck positioned central within a cervical imaging coil. Axial scans (Spin T1-weighted HF, matrix 512×512) were performed in a range from the upper jaw line to the clavicle, orthogonal to the spine, in 19 equidistant sections (Fig. 1B).

Linked US-Vicon dataset: Forty seven retroreflective markers were attached to the body to allow motion analysis of eighteen body segments (head, neck, thorax, pelvis, thighs, shanks, feet, clavicles, upper arms, forearms, hand) [14]. Participants stood in the middle of the calibrated volume and were instructed to perform pitch (flexion/extension), yaw (right/left) and roll (right side/left side) head rotations, turning their heads as far as possible in both directions and each trial was repeated starting in the opposite direction. Body motion was recorded by a 9 camera Vicon MX motion capture system. For each trial, the US probe was held to the posterior neck targeted at level C4, to allow free movement of the head and image of 5 bilateral layers of muscles. Images were saved digitally at 10 Hz with start time synchronized to the Vicon

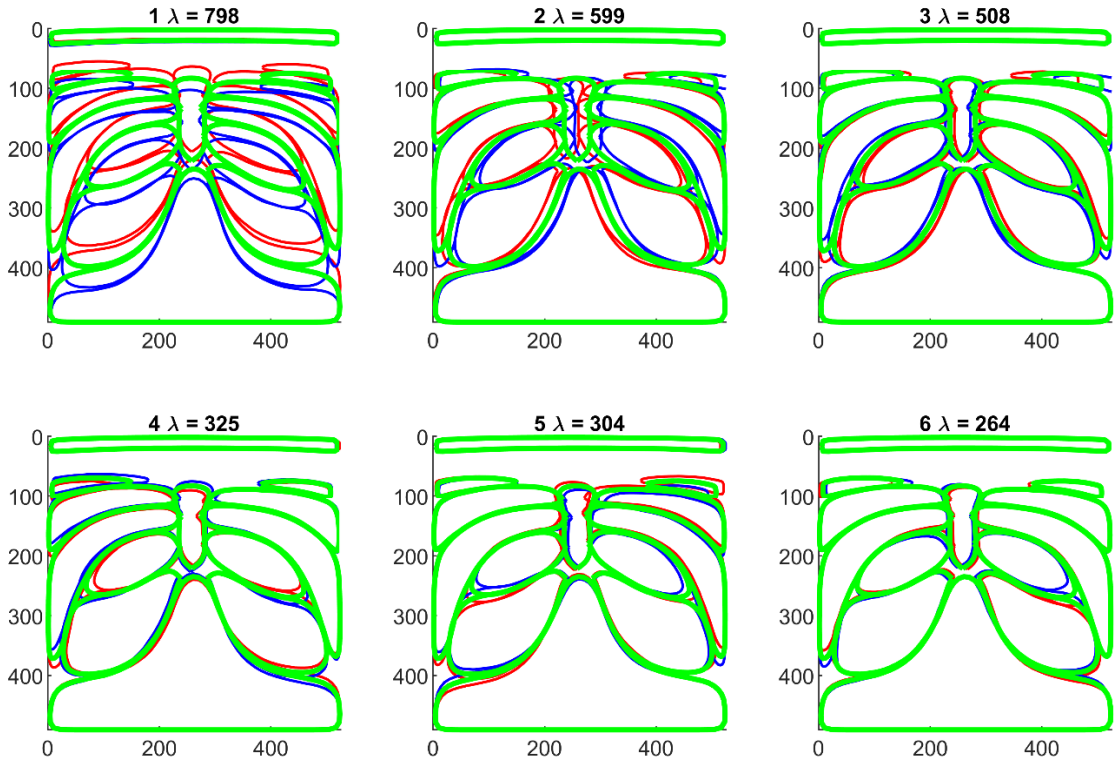

Figure 4 Principal Component shape patterns
One segmentation is described by 2,600 points (13 Segments x 100 horizontal + 100 vertical coordinates). Our test set of 434 segmented images are recomposed into principal components. Figure shows first six eigenvectors.
λ: Eigen value; Green: mean shape; Blue/Red: mean ± one eigenvalue of eigenvector. Individual segmentations are described by their component scores. Image left shows anatomical left.

recording.

B. Image Labelling

US images (192 total, ~3 per participant) from the linked US-MRI dataset were labelled manually by annotating the boundaries around ten muscles, vertebra, ligamentum nuchae and skin. As described previously [22], [23], MRI images showing the same cod liver oil capsule marked plane were annotated and registered to the US images to guide annotation of the US images.

Two thousand US images (~30 per participant) from the linked US-Vicon data, sampling uniformly the range of pitch, yaw and roll head rotations were labelled manually for the same 13 segments as described previously [23].

C. Machine learning

Two thousand examples from the linked US-Vicon dataset, were combined with 1,100 labelled images from our previous dataset representing healthy human head motion [24] to provide a training dataset. Each US image and corresponding label was flipped about the vertical line of symmetry, to artificially expand the training data set to 6,200 examples. Applying methods developed and software written within this group [23] an encode-decoder neural-network with a spatial SoftMax classification layer was trained (c.f. Fig. 2 for details). A class-weighted cross-entropy cost function was used to minimize the training error between labels and network prediction. Network training consisted of online learning, interrupted every quarter pass (550 learning iterations) through the training set, to record cross entropy test results from the 2 test sets. If the cross-entropy loss for either test set was lower than any previous recorded loss for that test set, the network was saved to long term storage. When neither test set recorded a lower loss for 25 consecutive test iterations, training was terminated.

Image-labels from the linked US-MRI dataset were combined with a previous dataset of 25 linked US-MRI image-labels [22] and were augmented by reflection to provide a validation/test dataset of 434 examples, 86 participants. Participants were divided into two groups: for each neural network, one group was assigned to select the training iteration with lowest cost function (validation) and the selected network was tested by the other group, and *vice versa* for both networks. This process yielded held-out test results for all images in the linked US-MRI dataset.

D. Boundary extraction and injection point analysis

The linked US-MRI dataset of 434 images was used for boundary extraction and injection point analysis (Fig. 3). To the predicted segmentation (Fig. 3) we applied an 8 x 8 pixel median filter, filled holes, smoothed the boundaries and extracted boundaries using MATLAB functions (*medfilt2, imfill, imclose* and *bwtraceboundary* respectively. All boundaries were extracted clockwise, starting from a key point, defined as the most medial pixel for muscles, and

Metric	Vertebra	Multifidus	Spinalis. Cervicis	Spinalis Capitis	Splenius Capitis	Trapezius	Lig. Nuchae	Skin	All Segments
JI (%)	78±11	53±13	52±18	52±13	50±17	40±21	34±17	44±19	**50±21**
HD (mm)	5.5±2.4	6.5±2.2	5.2±2.0	6.1±2.4	6.3±3.6	6.1±4.5	7±3.7	2.8±2.9	**5.8±4.0**
MHD (mm)	1.4±0.8	2.4±1.0	1.9±0.9	1.7±0.7	1.7±1	2.2±2.6	2.4±1.7	0.83±0.6	**1.9±1.8**

Table 1. Boundary accuracy between predicted and manually annotated boundaries. Jaccard Index (JI) show percentage intersection over union. Hausdorff Distance (HD) shows the greatest distance, and Modified Hausdorff Distance (MHD) shows the mean distance between predicted and manually annotated boundaries. This table reports mean ± S.D. for all participants from the expanded test dataset (N=86).

interpolated to 100 evenly spaced points (Fig. 3). Accuracy of boundaries was assessed using Jaccard Index (JI), Hausdorff Distance (HD) and modified Hausdorff Distance (MHD). While HD shows the worst (largest) distance between ground-truth and predicted boundaries, MHD shows the average distance between ground-truth and predicted boundaries [25].

For each segment, the central predicted injection point was defined as the pixel of maximum distance (d_{max}) from any boundary point (Fig. 3D). We iteratively increased the margin around this injection point by distance t = 0,1,2… d_{max} mm. The pixels enclosed by this boundary at distance $d_{max} - t$ from the predicted segment boundary provided a series of target injection regions (Fig. 3). By comparison with corresponding pixels in the manually labeled image, we compute average precision, for varying target region. This analysis was iterated using pixels only of predicted confidence (SoftMax scores) greater than 0, 0.2, 0.4, 0.6 and 0.8 respectively (Figs. 3, 6).

E. Principal component model of muscle shapes

We constructed a principal component model representing all shapes from the linked MRI-US dataset (434 cases, 86 participants). For one image, the pattern of 13 segments is described by 2,600 numbers (100 horizontal, and 100 vertical coordinates per segment). These 2,600 variables, 434 cases, were reduced to 100 principal components intended to capture all potentially important variation. Each component represents a pattern of variation from the mean shape (Fig. 4).

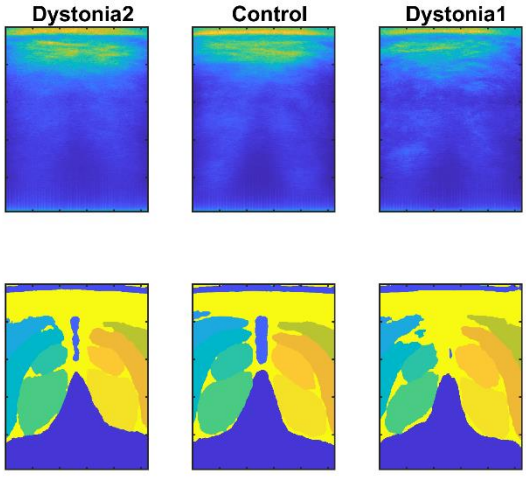

Figure 5 Clustering Dystonia into two groups
Dystonia participants were clusterred into two groups (Dystonia1, Dystonia2) using principal component scores and k-means algorithm. Group sizes were, Dystonia2 (n-22), Age matched Control (n=26) and Dystonia1 (n=13). For each group we show: **Top Row**. Median ultrasound images, **Bottom Row**. Median pixel classification image. Left side of all images represents the left anatomical side of the participant.

F. Analysis of muscle shape in cervical dystonia v controls

Analysis of muscle shape between conditions was restricted to the 35 cervical dystonia and 26 age matched controls (n=61) and excluded reflected images. For each participant, we computed their mean shape description (100 scores) to give 61 independent cases for statistical analysis.

Individual cases of dystonia deviate from normality in muscle shape in potentially positive and negative senses. The average, dystonic shape pattern may be inseparable from the average age matched shape pattern. To provide a simplest possible separation of signed descriptors, the 100-element shape descriptions of dystonia participants were partitioned into two clusters (Dystonia1, Dystonia2) using k-means with a correlation distance metric (Fig. 5).

To identify components distinguishing group membership (Dystonia1, Dystonia2, Control) we computed a univariate ANOVA for each principal component (n=61), (Fig. 7). Then, using MATLAB functions *sequentialfs* (10-fold cross validation, 50 monte-Carlo repetitions, forward entry starting with significant univariate components), and *classify* (‘diaglinear’, naive Bayes), we selected the combination of principle components which predicts group membership.

To reduce the selected components to the two linear discriminant functions, maximizing separation of the groups, we performed one-way Multivariate Analysis of Variance (n=61), using MATLAB function *manova1* (Figs. 8, 9);

G. Association of US muscle shape with whole body posture

For each participant, we computed their median multi-segment posture (51 angular components from 17 joints) from all their trials in the *Linked US-Vicon dataset*. To identify postural joint angles discriminating group membership (Dystonia1, Dystonia2, Control), we computed a univariate ANOVA (n = 61) for each joint rotation (Fig. 10).

III. RESULTS

We report three main findings: (i) accuracy of segmentation and of injection points within neck muscles, (ii) analysis of patterns of muscle shape associated with cervical dystonia and (iii) discrimination of cervical dystonia from muscle shape alone.

A. Accuracy of segment boundaries and injection points

For the test dataset (linked US-MRI, n=434, 86

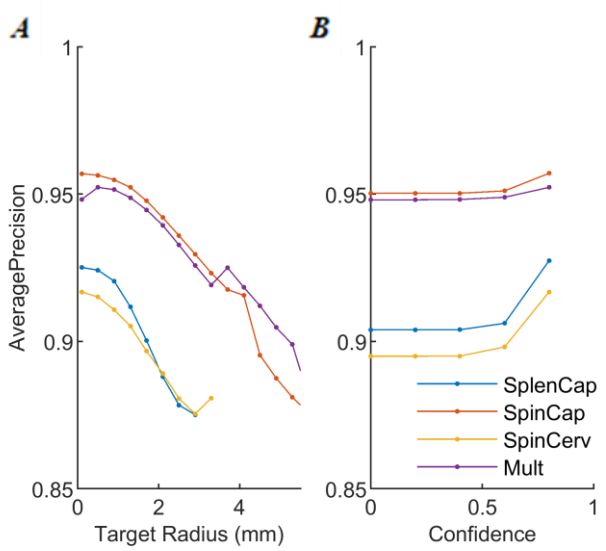

Figure 6 Precision of Injection points
A: Average Precision for all pixels of confidence more than 80% in the target region for a variety of sizes of target region. ***B***: Average Precision for varying levels of confidence. Precision is number accurate as a percentage of pixels predicted to be within muscles. Average precision is precision averaged with respect recall, sorted in descending confidence.

participants), segments overall were predicted to accuracy 50±21% JI, 5.8±4 mm HD and 1.9±1.8 mm MHD and these values were typical for muscles deep to the surface (Multifidus, Spinalis Cervicis, Spinalis Capitis, Splenius Capitis), (Table 1). A clinically meaningful assessment of accuracy is provided by the question "would an injection into the predicted muscle segment inject the correct muscle?" As the size of target region around the central predicted injection point increased by several millimeters, average precision decreased slightly, but remained higher than 90% (Fig. 6). By using SoftMax confidence at 80% or more to select injection points (Fig. 3), accuracy indicated by average precision is improved (Fig. 6). Setting minimum prediction confidence to 80%, accuracy of injection points for the deep muscles (multifidus, spinalis cervicis, spinalis capitis, splenius capitis) is indicated by average precision 93.8±2% (Table 2). These results support our hypothesis (i) that segmentation of deep muscles will identify injection points within the designated muscle.

B. Principal component analysis of neck muscle shape

The principal component model of 100 eigen-vectors encapsulates 98.6% of the variance from the test dataset (linked US-MRI, n=434, 86 participants). The first component shows variation in depth of the neck muscles and vertebra typical of different size necks. The second component shows lateral variation in midline boundaries and components three onwards show asymmetric variation in muscle shapes and positions (Fig. 4). These asymmetric components have potential to describe cervical dystonia.

The differences in shape between dystonia and age-match controls are subtle rather than striking (Fig. 5). Using their component scores, dystonia participants were clustered into two groups resulting Dystonia1 (n=13) and Dystonia2 (n=22). By comparison with Controls (n=26), the median classification of image central pixels show more background, which indicates inconsistent variable locations of these muscles. However, Dystonia 2 indicates asymmetric compression manifested on the right neck, smaller spinalis cervicis (left and right), larger right multifidus, smaller and deeper right splenius capitis and left deviation of the midline relative to the vertebra (Fig. 5).

Univariate ANOVA analysis revealed seven components differ significantly between groups (Dystonia1, Control, Dystonia2), (Fig. 7). Component 6 (most significant), showing superficial displacement-enlargement in right multifidus (c.f. Fig. 4) is positive for Dystonia2 and negative for Dystonia1. Component 4 (second most significant), showing rightwards displacement of the midline is negative for Dystonia2 and positive for Control/Dystonia1. Fuller analysis requires a combination of components.

C. Discrimination between dystonia and age matched controls using neck muscle shape alone

Using 10-fold cross-validation, a linear naïve Bayes classifier selected 22 components to predict group membership (Dystonia1, Dystonia2, Control). These components classified group membership correctly at 93.4% for leave-one-case-out.

For visualization, discriminant function analysis reduced these 22 components to two eigenfunctions (DF1, DF2) both of which were significant ($p=6.9\times10^{-8}$, p=0.039, Fig. 8). With a separation of 20 units (mahalanobis distance), Dystonia2 is substantially different from Dystonia1 and Control, which

Segment	Target Radius (mm)	Distance From Edge (mm)	N Pos ($\times10^6$)	N Neg ($\times10^6$)	Acc (%)	TP Rate (%)	FN Rate (%)	FP Rate (%)	TN Rate (%)	AP (%)
Multifidus	0.3	5.57	4	55.2	93.4	1.6	98.4	0.0108	100	96.2
Spinalis Cervicis	0.5	3.39	2.26	56.9	96.4	5.6	94.4	0.024	100	94.7
Spinalis Capitis	0.1	3.37	2.39	56.8	96	0.233	99.8	0.00152	100	93.6
Splenius Capitis	0.3	2.53	1.58	57.6	97.4	3.1	96.9	0.0171	100	90.8
Trapezius	1.1	0.0599	0.53	58.7	99.4	51.8	48.2	0.129	99.9	89.1
Deep Muscles	**0.3 ±0.2**	**3.71 ±1**	**2.56±1**	**56.7 ±1**	**95.8 ±1.7**	**2.63 ±2**	**97.4 ±2**	**0.013 ±0.01**	**100 ±0.01**	**93.8 ±2**

Table 2. Muscle injection accuracy. Shows classification accuracy for pixels of confidence greater than 80% within target region. Values are mean for all participants in test set (N=86). Deep Muscles shows mean ± s.d. for multifidus to splenius. Injection point is the pixel furthest from all boundary points. Target Radius: decrease in distance from edge to target region increase target area. Distance from Edge: distance of target boundary from segment boundary. N Pos, N Neg: Number of pixels in image. Acc: Percentage of pixels classified correctly. TP Rate, FP Rate, FN Rate, TN Rate: True positive, false positive, false negative, and true negative rate respectively. AP: Average precision (precision averaged with respect to recall sorted in decreasing confidence).

differ from each other by 8 units (Fig. 8A). Dystonia2 is separated cleanly from the other two groups by DF1 alone, while two cases of Dystonia1 are mixed with Control (Fig. 8B).

The two discriminant functions (DF1, DF2) represent patterns of altered neck muscle shape (Fig. 9).

DF1 (positive for Dystonia2, negative for Dystonia1, Fig. 9A), represents depth-wise compression on the right muscles, reduced area of right and left Spinalis Cervicis, and enlargement of the right multifidus (Fig. 9B). Correlation of all principal components with DF1 (the structure matrix) shows DF1 is correlated particularly with components 6 (r=0.68, p=1.6x10-9) and 4 (r=-0.56, p=2.8x10-6), the two most significant (c.f. Fig. 4). Correlation of muscle areas (units: percentage area of all segments) with DF1 at ($p<0.01$) shows reduced area of right Spinalis Cervicis (r=-0.64, $p=3x10^{-8}$), left Spinalis Cervicis (r=-0.52, $p=2x10^{-5}$) and right Splenius Capitis (r=-0.43, $p=5x10^{-4}$), and increased area of left Spinalis Capitis (r=0.48, $p=8x10^{-5}$) and right Multifidus (r=0.36, p=0.004).

DF2 scores are positive in both Dystonia1 and Dystonia2 (Fig. 9C) and represent leftwards displacement of the midline (Fig. 9B). DF2 is associated most significantly with principal component 15 (r=-0.51, $p=2x10^{-5}$), (3rd most significant c.f. Fig. 4) and with increased area of right Splenius Capitis (r=0.36, p=0.004).

These results above, confirm our hypotheses (ii) that cervical dystonia manifests identifiable, abnormal patterns of muscle shape and (iii) that automated analysis of muscle shape alone allows significant discrimination of cervical dystonia from age matched controls.

D. *Correspondence between US and whole-body posture*

Using the Dystonia groups defined using US image analysis alone we examined the correspondence with whole body standing posture. The median standing postures of these groups (Dystonia2, Control, Dystonia1) show differences in whole body truncal alignment and head turn. Referenced to the head, Dystonia 2 shows a neck displaced to their right and a trunk displaced oppositely to their left (Fig. 9D). However, Dystonia1 shows also a neck displaced to their right while their trunk is displaced to their right as well (Fig. 9D). Dystonia1 is not a mirror opposite to Dystonia2.

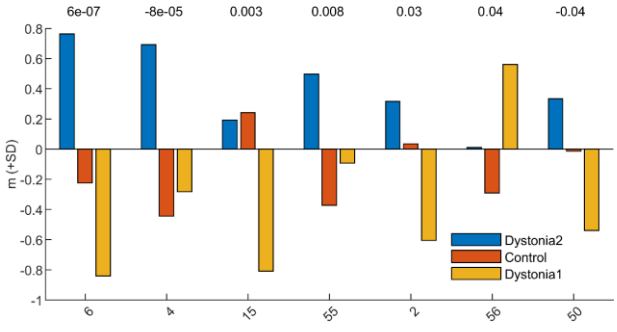

Figure 7. Univariate analysis of Ultrasound Principal Components to discriminate Dystonia from Controls
Each principal component of US boundaries was tested for significant difference between groups (Dystonia2, Control, Dystonia1) using one-way ANOVA (n=61). We show components sorted in decreasing order of significance to $p<0.05$.
Vertical axis indicates mean of group centre in units of whole sample standard deviation. To ease viewing Dystonia2 is always shown as positive and a negative value is indicated as a negative prefix to the p-value which is displayed on the top line.

Univariate ANOVA of joint angles shows the most significant difference between groups lies in head yaw (AOAngle_z) with Dystonia1 showing right yaw 7.6º±6º and Dystonia2 showing left yaw 1.2º±4º (Fig. 10). Significant differences in frontal rotation at the ankle joint (LAnkleAngle_y), hip extension (RHipAngle_x, LHipAngle_x) and frontal thoracopelvic rotation (ThoracopelvicAngle_y) all indicate changes in truncal alignment. Using the same procedure as above to select predictive joint angles resulted in correct leave-one-out classification of 60.0%. Following reduction to two discriminant functions, the first discriminant function was significant ($p = 3.8x10^{-5}$) correlated most highly with head yaw (AOAngle_z, r=0.53, $p=1x10^{-5}$). The second was not significant (p=0.09), but correlated most highly with frontal ankle angle (LAnkleAngle_y, r =0.49, $6x10^{-5}$)

This whole-body motion analysis, provides validation that diagnosis of dystonia condition from just three US images, has functional correlates in the standing posture.

E. *Comparison of diagnosis by posture with diagnosis by US*

Diagnosis by US reported above was superior to diagnosis by standing posture. Using the same procedures, we provide a comparative analysis using whole body motion data. We converted 51 components of joint rotation (17 joints x 3 degrees of rotation) into 51 principle components, we clustered Dystonia participants into two groups, we selected principle components of joint rotation to predict condition (Dystonia2, Control, Dystonia1), and we reduced selected predictors to two discriminant functions. Using all 51 joint rotations, the leave-one-case-out classification of group (Dystonia2, Control, Dystonia1) was 68.9% (Table 3 SM). By reducing joints used for clustering to truncal joints (atlanto-occiptial, neck-thorax and thoraco-pelvis angles) leave-one-case-out classification of group was improved to 78.7% (Table 3 SM).

Clustering dystonia participants using truncal joint angle, produced two discriminant functions most similar in separation and body posture to that obtained by US muscle shape (c.f. p-values, Mahalanobis separations and Structure matrices DF1, DF2, Table 3 SM).

IV. Discussion

A. *Contribution of this study: the main results*

This study reports results of the first application of deep learning to the segmentation, analysis and visualization of axial neck US images to participants with cervical dystonia.

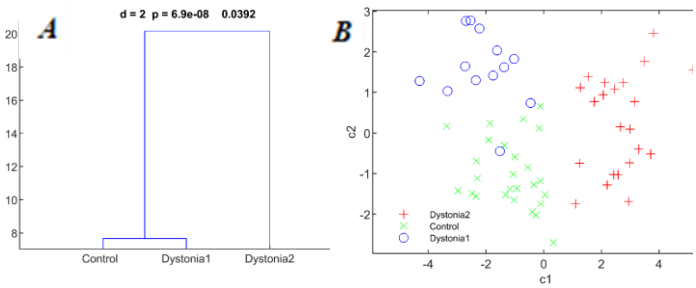

Figure 8. Diagnosis of Cervical Dystonia using principal components of neck muscle shape

Using principal components of US segment boundaries from all participants (35 dystonia, 26 age matched controls), we have assessed the extent to which shape alone can distinguish participants with dystonia from those without dystonia. A subset of 22 principal components were selected using a naïve Bayes classifier with 10-fold cross validation. These principal components were reduced to two discriminant eigen functions (DF1, DF2), i.e. combinations of components which maximise separation between the three groups (Dystonia2, Dystonia1, Control). Using Wilk's lambda, both eigenfunctions were significant (p=0.000000069, 0.039). N.B. "Muscle" shape refers to all 13 segments (muscle, vertebra, ligament, skin).

A. Mahalanobis distance (mean linkage) between group centres

B. Participants (n=61), plotted according to their canonical discriminant function scores (c1, c2 for DF1, DF2 respectively).

Dystonia2 is separated cleanly from both Dystonia1 and Controls by DF1. Dystonia1 is separated partially from Controls by DF2.

From a sample of 35 participants with cervical dystonia, and 26 age matched controls, we extracted neck muscle boundaries, reformulated into principal components and, using component scores, clustered dystonia participants into two main groups (Dystonia2, Dystonia1). The most salient findings are: -

(i) Cervical dystonia can be discriminated from age matched healthy controls, using an axial US image of the neck muscles (Fig. 8). Leave-one-out classification of group membership (Dystonia2, Dystonia1, Control) using naïve Bayes classification was correct at 93.4% (Table 3)

(ii) Cervical dystonia is associated with visible patterns of neck muscle shape (Fig. 9). This sample showed two significant eigen-patterns. The first (DF1), associated most strongly with component 6 (Figs. 4, 6), shows significant changes in the deep muscles (spinalis cervicis, spinalis capitis, multifidus) (Fig. 9B), provides clean signed separation of Dystonia2 from Dystonia1/Controls (Fig. 8B, Fig. 9A) and is associated with right head rotation (Fig. 10). The second (DF2), associated most strongly with component 15 (Figs. 4, 6), shows changes in the superficial muscles (splenius) and midline (Fig. 9B), separates Dystonia1/2 from controls and is most positive for Dystonia1 (Figs. 8B, 9C) and is associated with altered truncal (ankle, hip, pelvis) alignment (Fig. 10).

(iii) Segmentation is accurate enough to guide injections to specific muscles (Table 2).

B. Rationale for methods of analysis

The development of methods for annotating images, training neural networks and evaluating deep learning architectures to segment muscle boundaries in US images of the posterior neck is discussed more fully in our preceding submission [23]. Here, we apply the deep learning methods developed in our lab (Fig. 2) which provide immediate online segmentation, we acquire data from participants with cervical dystonia, extract muscle boundaries, and analyze the muscle shape associated with cervical dystonia. The need for visualization led us to extract linear components of shape (Fig. 4), which can be combined to provide an understandable analysis and visualization of neck muscle patterns in cervical dystonia (Figs. 4,5, 9). Our results demonstrate that US images of the neck muscles contain the information necessary to visualize, understand and potentially diagnose cervical dystonia.

C. Scientific and clinical value of results

Cervical dystonia is a neurological disorder of sensorimotor integration characterized by abnormal postures of the head and neck. Abnormal involuntary activation (dystonia) of neck muscles is a primary symptom, but is a direct cause of pain, abnormal whole-body posture, and constraints on movement. Neck muscles traverse the primary link between the head (which is the source of visual-vestibular head referenced sensory frames, location of sensory integration and motor planning) and the mass distribution of the body (trunk, upper and lower limbs). Abnormal action of neck muscles causes local changes in head and shoulder position and to maintain vertically aligned balance, these local changes require compensatory changes in whole body posture of the trunk and limbs. Neck muscles provide sophistical proprioceptive sensation and have a primary role in integration of head-referenced with ground referenced coordinate frames which is also subject to interference by abnormal neck muscle activity. Altered body posture and sensory feedback is a consequence of abnormal neck muscle action. Thus, analysis of the neck muscles provides direct insight into cervical dystonia.

US imaging analysis can quantify dystonic muscle attributes (Fig. 9). US does not require participants with movement disorders to remain still and avoids limitations of MRI imaging. US is relatively low cost and available in clinics. The confidence measure provided by this neural network analysis (Fig. 3) gives inexperienced operators feedback to improve the quality of their US probe location and image. This analysis within clinic could facilitate communication between patient and clinician and would inform patients about their neck muscles and their specific dystonia. The objective recording of images and analysis provides a potential tool for guiding and recording the location of injections, for monitoring change and improvement with treatment, and thus is expected to improve the patient experience. In addition, our findings reinforce the potential critical role in CD of deep neck muscles which have previously not been amenable to assessment or treatment.

D. Relationship to previous work

The application of machine learning and specifically deep learning to analysis of ultrasound images of muscle is rare [19], [21], [26]. While under developed, the domain of muscle diagnosis is valuable since unlike visual observation, manual

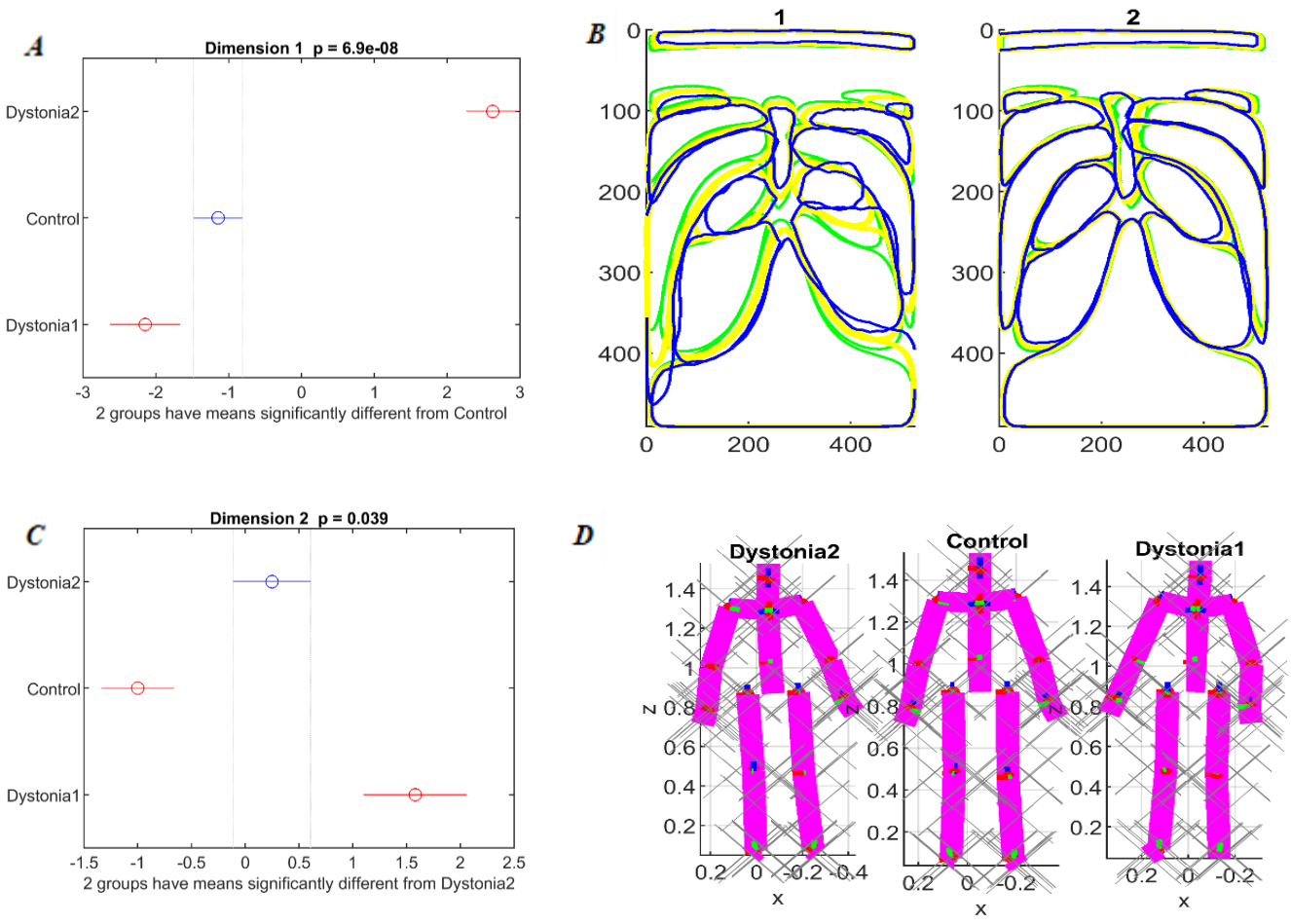

Figure 9 Analysis of Dystonia using neck muscle shape
"Muscle" shape alone can discriminate two subtypes of dystonia from age matched controls.
Panel A: Mean ± 95% confidance intervals of cannonical scores assigned to each individual from the first eigen pattern (*c1*). This pattern discriminates significantly (p=0.000000069) each of the three groups (Dystonia2, Control, Dystonia1).
Panel B: The two discriminant "eigen patterns" distinguishing cervical dystonia1 and dystonia2 from normal controls.
Green is sample mean principal component. Yellow and Blue show respectively +0.5 and +1 standard deviation of the eigen function. Image left shows anatomical left.
For DF1, dystonia2 is positive, dystonia1 is negative and age matched controls are close to zero. DF1 shows deep squashing on the right, and deep twisting with reduced size of both spinalis cervicis, enlargement of right multifidus.
For DF2 controls are negative, with Dystonia1 positive and Dysonia2 close to zero. DF2 shows leftward superficial deviation of the midline.
Panel C: Mean ± 95% confidance intervals of cannonical scores (c2) from the second discriminant function. This pattern also separates significantly (p=0.039) each of the three groups.
Panel D: Median posture (Dystonia2, Control, Dystonia1) facing the reader. Shows joint angles, median from all trials for each participant, averaged across all participants in group. This head reference presentation shows the kinematic chain reconstructed from the head segment which is presented vertical and forward looking for each group. Dystonia2 show neck deviated to their left, trunk deviated to their right. Dystonia1 shows neck deviated to their right, trunk also deviated to their right. Red, green and blue lines show extension, frontal, and axial axes of rotation respectively for each segment.

palpation or surface electromyography, ultrasound can see muscles deep within the body. Segmentation is the foundation of muscle specific analysis and recent methods providing segmentation of the neck muscles include computer vision [22] and deep learning approaches [23]. This study applies the most recent iteration in deep learning methods developed by this group for this application. Following [23], this study uses direct manual annotation of US images to provide training labels. This approach allow us to develop training datasets for participants with a movement disorder who cannot remain still in an MRI machine [23]. Using metrics of JI and HD (Table 1), the accuracy of segmentation achieved is consistent with existing benchmarks [22], [23]. The metric MHD (Table 1) shows boundaries are typically accurate 1.9±1.8 mm, and for deep muscles this accuracy allows for injection at average precision more than 90% for target sizes of several millimeter (SM Fig. A) and a margin from the muscle boundary of 3.7±1 mm (Table 2). These findings have important clinical implications, as freehand injections of botulinum toxin have been shown to have potentially suboptimal accuracy [27]. Manual annotations provide only an approximation to the true muscle boundaries. With training, neural networks learn image features that correlate consistently with the labels. In principle, machine learning discards random error in human labels, and tends to the on-average correct answer within and between labelers.

Prior to this study, it was an open question whether information contained within images of the neck muscles was

of any value for diagnosis and understanding of cervical dystonia. This study affirms the US information content with respect to objective clinical labels (control, dystonia) and with respect to objective motion analysis of whole-body posture. When dividing the 35 dystonia participants into two groups, these segmentations provided leave-one-out correct classification of group membership (Dystonia2, Dystonia1, control) at 93.4% (Table 3, Fig. 8). This classification using US neck muscle shape was more accurate than classification using body posture (Table 3). Furthermore, analysis of dystonia participants using US neck muscle shape identified distinct, patterns of head and truncal posture (Figs. 9D, 10).

This study demonstrates proof of concept of the feasibility of US imaging analysis of the neck muscles for understanding and diagnosing cervical dystonia. This proof of concept motivates further development of US technology. If deployed widely in clinics, there is potential to collect large quantities of data from the estimated 18,000 adults with this condition in the UK [1]. Combined with further exploration of neural network methods, there is potential for this tool to be become very robust, and for a new domain (ultrasound muscle analysis) to be established. Evaluation of the effect of therapeutic interventions e.g. BoNT on the patterns of change on US would also be critical to determine the utility of this tool to monitor changes in dystonia severity, and to evaluate its utility as a potential biomarker.

E. Limitations

The current work contains several limitations. First, we do not present a full classifier. We have a relatively small number of cases, which may not encompass the full and expanding spectrum of neck movements seen in cervical dystonia [11]. Further validation in a larger and independent clinical cohort would be desirable. A clinical classifier would most logically be embedded within the neural network architecture. Second this analysis is limited to an axial image at level C4. A fuller range of probe locations/orientations and muscle images would be desirable. Third, this work predicts and interprets muscle shape, excluding prediction of texture and muscle activity. Further work will exploit the ultrasound information content revealing muscle function [19], [20] as well as geometry.

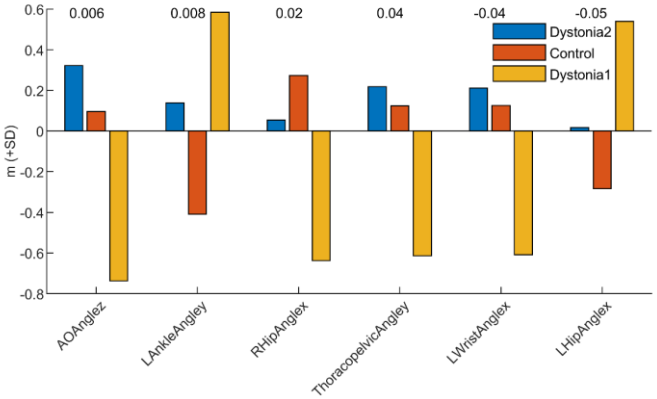

Figure 10 Univariate analysis of postural joint rotations to discriminate Dystonia from Controls
Each participant was represented by a median posture comprising 51 components of joint rotation (3 degrees of rotation x 17 joints).
Each joint rotation component was tested for significant difference between groups (Dystonia2, Control, Dystonia1) using one-way ANOVA (n=61). We show components sorted in decreasing order of significance to $p<0.05$. Vertical axis indicates mean of group centre in units of whole sample standard deviation. To ease viewing Dystonia2 is always shown as positive and a negative value is indicated as a negative prefix to the p-value which is displayed on the top line.

V. Conclusion

This study provides the first application of deep learning to US imaging of the neck muscles in cervical dystonia and provides an automated objective visualization and pattern analysis of neck muscle boundaries. These results demonstrate that muscle boundaries extracted from a single axial image of the neck muscles have the information content to discriminate cervical dystonia from healthy controls and to visualize and understand the dystonic pattern of neck muscles. This proof of concept demonstrates potential for a clinical tool to provide objective online diagnosis of cervical dystonia, guidance and objective logging of injection sites, and objective monitoring of the effect of treatment with minimal requirement for operator expertise and minimal burden on clinical time. This work supports a case for further evaluation of an automated US-based tool in a larger longitudinal dataset.

Acknowledgment

With appreciation we acknowledge the contribution of all our participants who willingly gave their time, the support of Dr Gregory May and Brian Bate to assist data collection and the encouragement and support of Joanne Day to develop this work to assist those with cervical dystonia.

References

[1] G. Defazio, J. Jankovic, J. L. Giel, and S. Papapetropoulos, "Descriptive epidemiology of cervical dystonia," *Tremor Other Hyperkinet. Mov. (N. Y).*, 2013.

[2] E. Tiderington *et al.*, "How long does it take to diagnose cervical dystonia?," *J. Neurol. Sci.*, vol. 335, no. 1–2, pp. 72–74, 2013.

[3] A. Albanese *et al.*, "Practical guidance for CD management involving treatment of botulinum toxin: a consensus statement," *J. Neurol.*, vol. 262, no. 10, pp. 2201–2213, 2015.

[4] G. Reichel, "Cervical dystonia: A new phenomenological classification for botulinum toxin therapy," *Basal Ganglia*, vol. 1, no. 1, pp. 5–12, 2011.

[5] H. A. Jinnah, E. Goodmann, A. R. Rosen, M. Evatt, A. Freeman, and S. Factor, "Botulinum toxin treatment failures in cervical dystonia: causes, management, and outcomes," *J. Neurol.*, vol. 263, no. 6, pp. 1188–1194, Jun. 2016.

[6] W. H. Jost, H. Hefter, A. Stenner, and G. Reichel, "Rating scales for cervical dystonia: a critical evaluation of tools for outcome assessment of botulinum toxin therapy," *J. Neural Transm.*, vol. 120, no. 3, pp. 487–496, 2013.

[7] M. F. Contarino, M. Smit, J. van den Dool, J. Volkmann, and M. A.

J. Tijssen, “Unmet Needs in the Management of Cervical Dystonia.,” *Front. Neurol.*, vol. 7, p. 165, 2016.

[8] J. A. Van Gerpen, J. Y. Matsumoto, J. E. Ahlskog, D. M. Maraganore, and P. G. McManis, “Utility of an EMG mapping study in treating cervical dystonia,” *Muscle and Nerve*, 2000.

[9] A. Schramm, D. Huber, C. Mobius, A. Munchau, Z. Kohl, and T. Baumer, “Involvement of obliquus capitis inferior muscle in dystonic head tremor.,” *Parkinsonism Relat. Disord.*, vol. 44, pp. 119–123, Nov. 2017.

[10] L. Tatu and W. H. Jost, “Anatomy and cervical dystonia : ‘Dysfunction follows form’.,” *J. Neural Transm.*, vol. 124, no. 2, pp. 237–243, Feb. 2017.

[11] A. Schramm, T. Bäumer, U. Fietzek, S. Heitmann, U. Walter, and W. H. Jost, “Relevance of sonography for botulinum toxin treatment of cervical dystonia: an expert statement,” *Journal of Neural Transmission*. 2015.

[12] A. I. Grigoriu, M. Dinomais, O. Rémy-Néris, and S. Brochard, “Impact of Injection-Guiding Techniques on the Effectiveness of Botulinum Toxin for the Treatment of Focal Spasticity and Dystonia: A Systematic Review,” *Archives of Physical Medicine and Rehabilitation*, vol. 96, no. 11. p. 2067–2078.e1, 2015.

[13] 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,” *J. Appl. Physiol.*, vol. 100, no. 4, pp. 1311–1323, 2005.

[14] I. D. Loram, B. Bate, P. Harding, R. Cunningham, and A. Loram, “Proactive Selective Inhibition Targeted at the Neck Muscles: This Proximal Constraint Facilitates Learning and Regulates Global Control,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 4, 2017.

[15] K. Bibbings, P. J. Harding, I. D. Loram, N. Combes, and E. F. Hodson-Tole, “Foreground Detection Analysis of Ultrasound Image Sequences Identifies Markers of Motor Neurone Disease across Diagnostically Relevant Skeletal Muscles,” *Ultrasound Med. Biol.*, Mar. 2019.

[16] I. D. Loram, C. N. Maganaris, and M. Lakie, “The passive, human calf muscles in relation to standing: The non-linear decrease from short range to long range stiffness,” *J. Physiol.*, vol. 584, no. 2, pp. 661–675, 2007.

[17] P. J. Harding, I. D. Loram, N. Combes, E. F. Tole, E. F. Hodson-tole, and E. F. Tole, “Ultrasound Based Detection of Fasciculations in Healthy and Diseased Muscles,” *IEEE Trans. Biomed. Eng.*, vol. Accepted f, no., p. , 2015.

[18] I. D. Loram, C. N. Maganaris, and M. Lakie, “Active, non-spring-like muscle movements in human postural sway: How might paradoxical changes in muscle length be produced?,” *J. Physiol.*, vol. 564, no. 1, pp. 281–293, 2005.

[19] R. Cunningham, M. Sánchez, G. May, and I. Loram, “Estimating Full Regional Skeletal Muscle Fibre Orientation from B-Mode Ultrasound Images Using Convolutional, Residual, and Deconvolutional Neural Networks,” *J. Imaging*, vol. 4, no. 2, p. 29, Jan. 2018.

[20] R. J. Cunningham, P. J. Harding, and I. D. Loram, “The application of deep convolutional neural networks to ultrasound for modelling of dynamic states within human skeletal muscle,” bioRxiv preprint https://doi.org/10.1101/157479, 2017.

[21] R. J. Cunningham, P. Harding, and I. D. Loram, “Deep residual networks for quantification of muscle fiber orientation and curvature from ultrasound images,” in *Communications in Computer and Information Science*, vol. 723, Springer, Cham, 2017, pp. 63–73.

[22] R. J. Cunningham, P. J. Harding, and I. D. Loram, “Real-Time Ultrasound Segmentation, Analysis and Visualisation of Deep Cervical Muscle Structure,” *IEEE Trans. Med. Imaging*, vol. 36, no. 2, pp. 653–665, Feb. 2017.

[23] R. J. Cunningham, M. B. Sánchez, and I. D. Loram, “Ultrasound Segmentation of Cervical Muscle during head motion: a Dataset and a Benchmark using Deconvolutional Neural Networks.” engrXiv, Feb-2019.

[24] I. D. Loram, R. J. Cunningham, and M. B. Sánchez, “Ultrasound/MRI Neck Muscle Image Semantic Segmentation Dataset.” IEEE Dataport, 2019.

[25] M. P. Dubuisson and A. K. Jain, “A modified Hausdorff distance for object matching,” in *Proceedings of 12th International Conference on Pattern Recognition*, 1994, vol. 1, pp. 566–568 vol.1.

[26] P. Burlina, S. Billings, N. Joshi, and J. Albayda, “Automated diagnosis of myositis from muscle ultrasound: Exploring the use of machine learning and deep learning methods,” *PLoS One*, vol. 12, no. 8, p. e0184059, 2017.

[27] A. Schnitzler, N. Roche, P. Denormandie, C. Lautridou, B. Parratte, and F. Genet, “Manual needle placement: accuracy of botulinum toxin A injections.,” *Muscle Nerve*, vol. 46, no. 4, pp. 531–534, Oct. 2012.

Source	**N Available for clustering**	**N Selected**	**LOO (%)**	**Eigenfunction p values**	**Main Separation** D2 v D1/Control (mahalanobis)	**Lesser Separation D1 v Control (mahalanobis)**	**DF1**	**DF2**
Ultrasound	100	22	93.4	6.9e-8, 04	20	8	AOz,	Lankley
Trunk	9	9	78.7	3.6e-8, 02	13.6	6	AOz, NeckThoraxy, x, Lclavicley	Lankley
Trunk + shoulders	18	18	67.2	001	7.6	4.3	Thoracopelvicx, Lclaviclex, z, Rhipx, Lhipx, Lclaviclez, Rwritsy	Lankley, Rankley
Whole body	51	14	68.9	4.10E-06	11.2	2.5	Lwristz	Lankley, Rclaviclez,x

Supplementary Material Table 3. Comparison of Ultrasound and Posture to discriminate Dystonia from Controls.
Shows results of classification and descriptive discrimination using Ultrasound neck muscle boundaries and various combinations of postural joint rotation components to cluster and select predictive components. **Source**: Variables used to cluster Dystonia participants into two groups, and then examined to explain the difference between dystonia and controls. Ultrasound (neck muscle boundaries), Trunk (Atlantooccipital, Neck-thorax, thoraco-pelvic), Trunk + Shoulders (as per trunk, clavicular-thoracic), Whole body (51 components from 17 joints). **N Available** for clustering: Number of components of joint rotation or neck muscle boundary principal components used to cluster dystonia participants into two groups. **N Selected**: Number of components selected to classify group membership (Dystonia2, Dystonia1, Control). **LOO**: Percentage of leave-one-out correct classification of group membership (Dystonia2, Dystonia1, Control).
Eigenfunction p values: Significance of two discriminant eigenfunction created from selected components. **Main separation**: Dystona2 v Dyston1-Control (Mahalanobis distance) i.e. generalised units of covariance. **Lesser separation**: Dystonia1 v Control (Mahalanobis distance); **DF1**: Variables most highly correlated with the first discriminant function. **DF2**: Variables most highly correlated with the second discriminant function.