Ensemble of shape functions and support vector machines for the estimation of discrete arm muscle activation from external biceps 3D point clouds

Leandro Abraham[a,b,c,*], Facundo Bromberg[a,c], Raymundo Forradellas[b]

[a] *Laboratorio DHARMa, DeSI, Universidad Tecnológica Nacional, Facultad Regional Mendoza - Rodriguez 273, PC M5502AJE, Mendoza, Argentina - Phone: +54 261 5244566*

[b] *CEAL, Universidad Nacional de Cuyo, Facultad de Ingeniería - Centro Universitario CC 405 , PC M5500AAT, Mendoza, Argentina - Phone: +54 261 4135000 (ext. 2128)*

[c] *Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

Abstract

Background: Muscle activation level is currently being captured using impractical and expensive devices which make their use in telemedicine settings extremely difficult. To address this issue, a prototype is presented of a non-invasive, easy-to-install system for the estimation of a discrete level of muscle activation of the biceps muscle from 3D point clouds captured with RGB-D cameras.

Methods: A methodology is proposed that uses the *ensemble of shape functions* point cloud descriptor for the geometric characterization of 3D point clouds, together with *support vector machines* to learn a classifier that, based on this geometric characterization for some points of view of the biceps, provides a model for the estimation of muscle activation for all neighboring points of view. This results in a classifier that is robust to small perturbations in the point of view of the capturing device, greatly simplifying the installation process for end-users.

Results: In the discrimination of five levels of effort with values up to the maximum voluntary contraction (MVC) of the biceps muscle (3800 g),

*corresponding author

Email addresses: leandro.abraham@frm.utn.edu.ar (Leandro Abraham), fbromberg@frm.utn.edu.ar (Facundo Bromberg), kike@uncu.edu.ar (Raymundo Forradellas)

the best variant of the proposed methodology achieved mean absolute errors of about 9.21 % MVC — an acceptable performance for telemedicine settings where the electric measurement of muscle activation is impractical.

Conclusions: The results prove that the correlations between the external geometry of the arm and biceps muscle activation are strong enough to consider computer vision and supervised learning an alternative with great potential for practical applications in tele-physiotherapy.

Keywords: biceps activation estimation, 3d point clouds, support vector machines, ensemble of shape functions, biomechanics, tele-physiotherapy

1. Introduction

This paper tackles the problem of remotely estimating the level of activation exerted by the biceps muscle when subjected to external forces into a posture of isometric contraction. Our approach aims to solve the problem for typical tele-physiotherapy conditions: indoors, controlled lighting for the span of the study, easy-to-install hardware, and fast calibration of the software. To achieve this, the proposed system autonomously estimates the discrete activation level based on computer vision plus machine learning techniques when given 3D point clouds of the arm captured by a commercial off-the-shelf RGB-D capturing device such as the Microsoft Kinect™. The approach has been designed with easy installation in mind, so emphasis has been given to producing estimations that are robust to imprecise hardware mountings, so robust that the system can be installed even by the patient (whenever the person's medical condition allows) with only a handful of training examples to calibrate to different patients. The approach works under the principle that the activation of a muscle is physically expressed by changes in the geometry of the external (visible) muscle surface. It proposes a supervised learning approach, namely a *support vector machine* (SVM) [1, 2], to learn the mapping between the geometric features of the muscle surface — characterized by the *ensemble of shape functions* point cloud descriptor [3] computed for the point clouds of the muscle — and the activation level for each patient's biceps muscle. A supervised learner provides the benefit of generalizing to unseen situations — in this case a generalization to new measurement viewpoints — caused by potential misplacements of the apparatus during setup at the beginning of each exercise session, or by involuntary movements of the arm during the measurement process.

The present work has been inspired by the recent growth in tele-medicine and tele-physiotherapy [4, 5, 6, 7] and constitutes an attempt to fulfill the need for remote muscle activation sensing [6, 7, 8]. Currently, muscle activation is measured by electromyography (EMG) [9], a solution that although highly accurate, requires specialized, expensive, and intrusive equipment dependent on complex installation procedures by specialized personnel to attach electrodes to the body, thus strongly limiting its practicality in many tele-physiotherapy exercises. Instead, the proposed approach uses unobtrusive off-the-shelf equipment, no more difficult to install and train than some console video-games, at the expense of lower precision in the estimation of activation level, but precise enough for many existing physiotherapy exercises. The literature corroborates the existence of some physiotherapeutic exercises requiring the monitoring of muscle activation at discrete values that allow some rough level of imprecision, making them suitable for this approach.

One such exercise is that known as *hold-relax*, *contract-relax* and *hold-relax with agonist contraction*, a set of important exercises in rehabilitation programs [10]. In all these exercises, monitoring is required to control the maximum intensity of muscle contraction as well as the duration of muscle activation, avoiding efforts longer than the physician-prescribed percentage of maximum voluntary contraction — MVC (the maximum possible contraction that a subject can produce in the muscle by himself with no external forces). At present, in tele-rehabilitation sessions with this type of exercises, monitoring is circumscribed to human measurement of muscle activation (e.g. the physiotherapist or even the patient), as in most cases it is impossible to operate, or even in some cases transport, an EMG for such simple monitoring scenario. As expected, however, this human measurement is not only highly imprecise, but when there is no assistant, it requires the patient to be worried about monitoring the exercise, distracting him from the rehabilitation task. The proposed approach is suitable to address such scenarios, where the system could be calibrated to detect the maximum activation indicated for the patient's treatment (and, possibly, some other intermediate activations for early warning purposes). This calibration could be done in the presence of a trained assistant during the setup session and left to the autonomous system to detect over-activation in future sessions.

A second case under consideration is a common clinical task in rehabilitation consisting in the assessment of muscle health by using the simple, well-known *manual muscle test*. This test consists in the therapist measuring the contraction level of a muscle when subjected to predefined, specific pos-

tures and movements. It is based on a subjective grading system in which the level of contraction is measured grossly by touching and feeling the muscle [11, 12]. In a tele-health scenario, direct touch is impossible, an EMG is not feasible in most cases, and verbal transfer of subjective measurement based on an assistant's touch would be highly inaccurate. The proposed method could provide a way to make the remote execution of this test possible by calibrating the system with the therapist's subjective perception of contraction during a first face-to-face session between therapist and patient. Once calibrated, the system could perform the measurements remotely from the patient's home.

These two cases are just examples of the widespread use of muscle activation in physiotherapy, stemming from the key role it plays in modeling the biomechanics of the musculoskeletal human system, for which it is required, together with joint *kinematics* and the *kinetics* (external forces) exerted on the body [13, 14, 15, 16, 17, 18]. Its importance for medical care should not be understated, as these biomechanical models provide the mechanical structures, laws, and phenomena essential for human balance and movement, allowing the identification of harmful movements, over-exertions, awkward postures, musculoskeletal disorders, and optimal movements, among other states of the human body with a high impact on health. This results in the application of this approach beyond tele-physiotherapy in disciplines like occupational medicine [19], ergonomics [20], and sports [21], among others.

An autonomous system for visual sensing of the activation level of human muscles is an open problem with only a few published contributions (described in some detail in the following section) mostly focused on characterizing the visual aspects of muscle contraction. To the best of the author's knowledge, however, no work has been done to use this information in measuring muscle activation level. Some work has been performed on autonomous visual sensing of other aspects of human biomechanics, such as joint *kinematics* [22], presenting solutions so advanced that there are even existing commercial applications [23, 24], and *kinetics*, which is currently an active line of research with several ongoing efforts to achieve a final solution [25, 26, 27, 28, 29, 30, 31].

The rest of this paper is structured as follows. The methodology for solving the problem is included in Section 2. Section 3 presents the experiments performed for empirically proving the effectiveness of the method. A discussion of results and a comparison of this work with other related work is included in Section 4. Finally, conclusions and future work are outlined in

Section 5.

2. Materials and Methods

The main component of the proposed method is a computational procedure for the estimation of muscular activity based on a *supervised learning* algorithm for mapping the *geometric features* of the muscle's external surface to *muscle activity*. The following section introduces some basic concepts and procedures of these technologies, followed by a detailed explanation of the proposal for using them to solve the problem of remote measurement of biceps activity.

2.1. Background knowledge

This sub-section presents the basics of the computer science technologies applied in this research necessary to understand how the method works. Section 2.1.1 explains the functioning of supervised machine learning techniques used to estimate activation levels. Finally, Section 2.1.2 presents the operation of computer vision techniques used to characterize the method's input point clouds.

2.1.1. Supervised learning

Supervised learning is a technology from the artificial intelligence area that consists in learning multivariate, non-linear functions inductively. This work uses classification methods whose functions map the multivariate input to a discrete-valued variable usually referred to as *class*. For instance, this work tackles the measurement of activation in five discrete levels for which the corresponding classes are named with the labels $O0$, $O500$, $O1k$, $O2k$ and $O4k$, respectively.

Supervised learning algorithms construct this function through an inductive process that takes as input a *training set* of possible inputs to the function, that is, some assignment of all the variables in the function domain paired with the value of the function for that input configuration of the variables. Common names given to the value of the function are *label* or *ground-truth* for that input vector. Most commonly, they propose a family of parametrized models and work by searching the space of possible parameter values, and therefore, the space of all models in the family. This results in a model whose output better matches the expected output signaled by the labels. Due to insufficient data or noisy labels, the model learned may

differ from the true underlying classifier. Since in practice the underlying classifier is unknown, the learned model can be evaluated by testing it empirically against some labeled examples left out of the training set. Such set is called *test set* for obvious reasons. The testing proceeds by running the classifier over each (labeled) example in the test set and comparing the output produced by the learned model (hereon called *predicted label*) for the n^{th} example denoted by $\tilde{w}_n$, with the label representing the ground-truth value of the classifier (hereon called *true label*) for the n^{th} example denoted by w_n. All these comparisons, one per test example, are aggregated into different kinds of possible *performance measures* that highlight different qualities of the learned model. The present work considers the *Mean Absolute Error* (MAE) defined by:

$$MAE(testset) = \frac{\sum_{n=1}^{|testset|} |w_n - \tilde{w}_n|}{|testset|} \tag{1}$$

That is, the absolute value mean of the differences between the true and predicted weights. An advantage of using MAE as a performance measure is that its value can be measured with the same physical magnitude as the labels. As explained below, the proposed approach uses weights for labels, so in this case, MAE would be measured in grams. From this definition, better performances are reflected by smaller MAE values.

Interestingly, the supervised learning models work regardless of the strength of non-linearity and dimensionality of the function as well as noise in the training set; for instance, input examples labeled incorrectly, that is, paired with values that do not match the value of the function for that input. Most of the learning algorithms require user given parameters called *hyperparameters* to distinguish them from the internal parameter that defines the configuration of the technique. Users selecting these hyperparameters may result in underperformance. An alternative approach used in this work, known as *tuning*, is to infer them by first splitting the training set into a smaller training set and a *validation set* and iteratively navigating through a representative range of hyperparameter values and to train a model for each value with the training set, to then compute performance measures with the validation set. The hyperparameters selected are those that produce the model with the best average performance of the validation set, over all possible splits.

The proposed work suggests using one of the most effective shallow classifiers: Support Vector Machines (SVM). SVM belongs to the family of *lazy*

classifiers that require the memorization of training examples, but in the case of SVM, it requires only the memorization of a sparse set of examples, called *support vectors* [1, 2]. The algorithm proceeds by constructing a hyperplane that optimally separates the decision boundaries for feature vectors of different classes. These feature vectors may have been projected to spaces of higher dimensionality through a *kernel function*, if they are not linearly separable in their own space. For this work, the *linear* kernel was selected against two alternatives: *radial basis* and *polynomial* kernels [2], discarded through preliminary experiments (results not shown). Intuitively, a good separation that maximizes hyperplane generalization power is achieved by maximizing its margin, defined as the perpendicular distance between decision boundaries. This approach uses the SVM implementation provided in the *caret* package of the R programming language. For the final version of the proposed approach, a value of 0.002 was chosen for hyperparameter C — the strength of the penalty for misclassified examples — obtained from preliminary tuning.

2.1.2. Geometric descriptors

The proposed approach for supervised learning uses visual information of the muscle surface as input. To produce a meaningful model, however, one must transform the raw visual information, that is, the 3D point cloud capture of the muscle surface, into a format with high discriminative power. This is usually referred to as features — functions of the raw input whose outputs are better and more compactly grouped over classes. The best way to prove that the features produce a compact grouping is to run a good supervised learning algorithm. Experimental results later on show the empirical procedure for selecting the best out of three pre-selected 3D geometric feature candidates: *viewpoint feature histogram* (VFH) [32], *clustered viewpoint feature histogram* (CVFH) [33], and *ensemble of shape functions* (ESF) [3]. These are well-known, widely used geometric features that are also conveniently implemented in the *point cloud library* [34] and are briefly described as follows:

Viewpoint Feature Histogram. (VFH) [32]: This descriptor represents the relative orientations of normals and distance between each point and the point cloud centroid. These pairs are encoded with angles between the normals at the considered points and the normal at the centroid. Each of these angles and distances are binned into a histogram.

Clustered Viewpoint Feature Histogram. (CVFH) [33]: This descriptor works by dividing the object in N disjoint smooth regions, each of which is used to compute a VFH histogram. The approach behind CVFH is to use object parts to build the coordinate system while still using the whole view of the object to compute the descriptor. This descriptor adds a *shape distribution component* that encodes information about the relation of each point with the centroid of its corresponding object region summarized for the entire object.

Ensemble of Shape Functions. (ESF) [3]: This descriptor is an ensemble of ten histograms which summarize functions describing characteristic shape and angle properties of the point cloud, resulting in a vector of 640 variables. One group of variables is estimated by sampling point-pairs from the point cloud and building a histogram of the distances between them. Another group of variables is estimated by tracing the lines between random samples of point pairs and summing them in a histogram counting the number of lines on the surface of the point cloud, off the surface of the point cloud, and partially on the surface of the point cloud. The values of another group of variables are computed by encoding into a histogram the angles between the two lines constructed from three random points of the cloud. In like manner, the final group of vector variables was built by encoding the surface area constructed with triplets of points and counting their number on the surface, off the surface, and partially on the surface of the point cloud.

2.2. The proposed approach

This section presents the methodology for the non-invasive estimation of biceps muscle activation level based solely on visual information about the external geometric features of arm deformation when subjected to a discrete set of efforts.

A naïve approach would attempt a manual mapping between the output of one geometric feature and the different activation levels exerted by the muscle. However, not only it is almost impossible to produce a manual mapping because of the large dimensionality of the problem (the ESF feature vector has 640 dimensions), but this mapping is also problematic due to the inherent noise of the signal (e.g., movements of the arm, of the capturing device, illumination, among others). This results in a non-trivial separation surface in feature space between the subspaces corresponding to different activation values. This issue was addressed through an autonomous learning algorithm,

support vector machines, specialized in automatically producing such a mapping, even for complex separating surfaces. To produce good enough mappings, this algorithm, as any other supervised learning algorithm, requires a minimally sufficient number of labeled mapping examples. To achieve this, a careful capturing setup must be designed to produce accurate and simultaneous measurements of both parts of each example. In our scenario this corresponds to capturing the geometrical features of the muscle surface while the activation level is being measured.

To capture geometrical features, the proposed approach uses RGB-D devices such as the Microsoft Kinect™ that produce the muscle surface 3D point cloud necessary to compute ESF geometric features. To produce the activation measurement, one would immediately consider an EMG measuring device attached to the biceps. This, however, presents several shortcomings. On the one hand, the electrodes would clutter the image, producing extra noise that hinders the mapping process. But most importantly, in a practical tele-health application, the proposed approach requires that the measurements for training the mapping algorithm be produced for each new patient — even in remote scenarios — where the transport and setup of an EMG would greatly increase the expertise level of the required tele-health assistant. This approach, therefore, proposes an alternative indirect activation measurement based on the fact that some muscle rehabilitation exercises require isometric contractions of the muscle [10], mainly after surgery, as a means of exercising the muscles without forcing them into making a great deal of movement. These are static contractions of the muscle with no movements that stretch or shrink it, with the person staying relaxed so that muscle contraction is the minimum required to hold the desired effort. In such cases, an external force, such as the weight of an object held by the hand, would be transferred almost completely to the muscle, resulting in a direct, monotonous increasing correlation between weight and activation level of the muscle, as proven for instance in [35, 36]. Thus, to produce the mapping example, the 3D point cloud of the biceps muscle was captured while the corresponding hand was holding weights of known value. Those weight values were taken as a measure of the biceps activation level. More details of the capture procedure specific to the case of the biceps muscle are provided in the experimental section.

The procedure just described, although a great improvement over manual or simplistic map- pings, still presents an important shortcoming. In practice, the training and later use of the trained system spans days or even weeks,

making it almost impossible to guarantee a precise alignment between the 3D point of view of the Kinect while producing the training examples (required for the SVM to produce the mapping) and the 3D point of view of the Kinect when it captures biceps 3D point clouds to obtain the activation level later on in the rehabilitation process. This results in a brittle system, since the values produced by all three geometric features are sensitive to changes in point of view, which modifies which parts of the surface are visible and which are occluded. This results in rather large changes in the 3D point clouds. To address this serious shortcoming, the approach proposes expanding the training base of the supervised learning algorithm to a whole new set of points of view. Under this new training scenario, the set of examples mapping to some given weight is now expanded to include 3D captures of the muscle not from one, but from a set of spatially neighboring points of view located within a sphere, with the expectation that the learned model will now be robust to measurements from all points of view within that sphere.

In practice, however, this proposal is quite demanding if not completely unfeasible, as it would not only require that the patient perform an enormous amount of weight lifting, most probably completely counter-indicated by the physician, but also a cumbersome re-positioning of the Kinect to all possible training points of view. Therefore, the approach was extended to consider a simulated generation of training examples which, based only on captures from a single point of view, produces data points for all neighboring points in the sphere through a simulation rather than actual captures. To do this, based on a small set of 3D point clouds obtained by actual captures, 3D point clouds were produced that would be observed from the new point of view by detecting which parts of the arm should not be visible from that new point of view and removing the corresponding 3D points from the original capture. For this operation, the *Hidden Point Removal* (HPR) operator presented in [37] was used. By computing point cloud geometric features for these neighboring points of view and by attaching the corresponding weight of the original captures, a simulated training example is obtained. In principle, new visible areas that now appear in the new point of view should also be added. In this setup the problem was solved by simply ignoring those extra areas. As this only discards useful information, adding it back could only improve the quality of the model, so the errors resulting from the model trained this way should be taken as an upper bound.

The following section presents the experimental setup, where different scenarios and user given choices for properties of the approach were considered,

proving the feasibility of the proposed approach for practical applications.

3. Empirical evaluation

This section describes the experiments conducted to prove the effectiveness of the proposed approach measured in terms of its practicality for solving the problem of tele-physiotherapy, taking as a guiding principle the two practical applications described in the introduction: (i) the *hold-relax*, *contract-relax* and *hold-relax with agonist contraction* exercises that require the monitoring of over-contraction above a therapist-specified maximum value, typically indicated as a percentage of MVC; and (ii) the *manual muscle test* that requires the subjective, touch-based assessment of how much the muscle is contracted, measured at discrete levels calibrated at the physician's interest values. In both cases the measurement is on discrete activation values that range between no contraction and MVC values. It is, therefore, a rough activation estimate sufficient to improve a situation in which EMGs are an unfeasible alternative. To achieve this, five discrete activation values were considered, produced by weights that roughly cover the range from no weight lifted at the hand, only the weight of the forearm itself, to a weight at the biceps muscle equivalent to the typical MVC. According to several studies [38, 39], the MVC of an adult biceps muscle ranges approximately from 200N to 350N. For the biceps to produce such a force through a weight at the hand under isometric contraction (static scenario), one must equate the torques produced by each (see [40] for details). With the hand at roughly 30cm and the biceps muscle at 4cm from the elbow joint, there is a factor of 7.5 between these forces for a static scenario corresponding to isometric contraction. To this, one must add the weight of the forearm itself (approximately 15N) assumed to be applied 15cm from the elbow joint. This adds an extra torque in the same direction as the hand weight, equivalent to 225Ncm. Such combination results in a force (weight) of approximately $(350 \times 4 - 225)/30$ = 39.16N or 3.9kg at the hand to produce 350N of MVC at the biceps. Thus, five discrete levels were considered within this weight range, roughly duplicating at each step: 0g, 505g, 1340g, 1940g, and 3885g each, hereon named with the labels $O0$, $O500$, $O1k$, $O2k$ and $O4k$, respectively. From the above discussion, it can be inferred that a good performance of the method predicting these values can be extrapolated to expect good results in practical cases. However, further testing would be required before implementing this approach in actual treatments.

The performance of the proposed approach, as any other supervised learning algorithm, strongly depends on how well its training set represents the underlying mapping — in this case, the mapping between the geometric features of the muscle surface and its level of activity. To assess this performance, MAE (Eq 1) is reported in grams for a representative range of possible training scenarios that depend on several alternative user choices. On the one hand, the choice of the point cloud descriptor used to generate feature vectors of the training and testing sets — that is, VFH, CVFH, or ESF —. On the other hand, choices relative to the simulation based augmentation of the dataset: the point of view of the captures; the number n of original captures; the $trainRadius$ of the sphere; and the number N of simulated neighboring points of view. Each scenario relies on actual captures of the biceps subjected to efforts of some weight held by the hand. Therefore, the next section explains how the captures were produced, followed by two subsections that prove empirically for which of these alternatives the learned mapping confirms its validity as a practical muscle-activity measuring-process for telerehabilitation. One shows results over the alternative geometric features, and the other over the parameter for producing the augmented datasets.

3.1. Capturing setup

Using a Microsoft Kinect™ RGB-D camera, 3D point clouds were captured from a frontal view of the biceps muscle (see Figure 1) by positioning the Kinect at coordinate (0,0,-0.2), corresponding to 20 cm over the negative Z axis and pointing towards the origin. All captures were performed maintaining uniform illumination and scale conditions in a posture restricted to ensure an isometric contraction of the biceps. The subject was told to maintain the segment of the arm from the shoulder to the elbow in an angle of approximately 45° with respect to the torso, and as the segment of the arm from the elbow to the hand is in horizontal position parallel to the floor plane (X axis), the shoulder-elbow segment is also approximately 45° with respect to the X axis. The whole arm was supported only at the shoulder by the back of a chair on which the subject was seated.

For each of the five weights, four bursts were captured using a tool of the Point Cloud library[34] that captures 3D point clouds with the Kinect at 30 fps. Ten seconds were captured for a total of 300 frames, from which 25 were sampled at random, resulting in four groups of 25 3D point clouds per weight. Three bursts per weight — i.e. $75 \times 5 = 375$ point clouds — were used to construct the captured training set. One burst per weight —

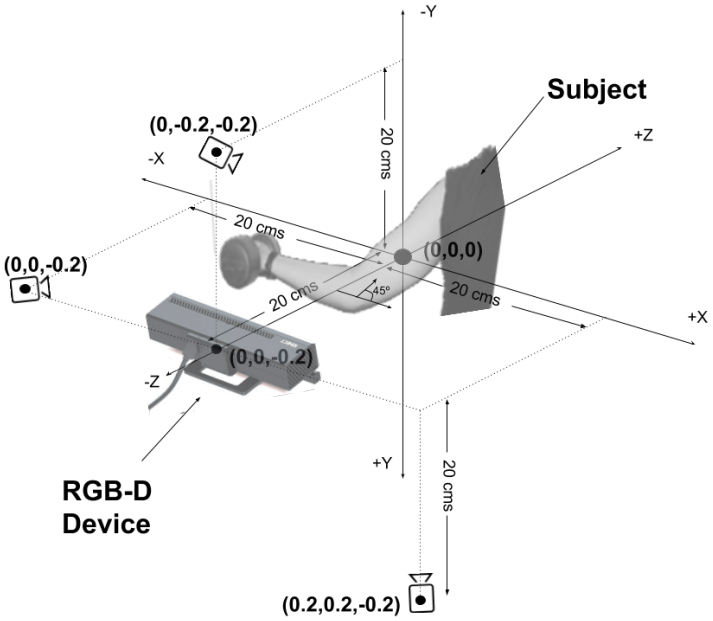

Figure 1: Capture setup

$25 \times 5 = 125$ point clouds — was separated for constructing the captured testing set following the typical $75\% - 25\%$ split.

Next, these 500 point clouds were post-processed manually segmenting out the points not corresponding to the biceps and moving the resulting point cloud, so that the point in the biceps corresponding to the mid-distance between its extremes (elbow and shoulder) was located at the origin of the coordinate system. To accelerate a laborious manual segmentation, the same segmenting planes were chosen for all images of the same burst. In the future, the segmentation process could be automatized by applying algorithms of 3D object detection trained through examples to locate the target detection object (in this case the biceps) in the full point cloud of the scene [41, 42].

In the next section, there follows an experiment for selecting the best feature descriptor among the three considered.

3.2. Selection of geometric descriptor

As already mentioned, the best feature descriptor would be that with the highest discriminative power, only measurable by comparing its performance with some supervised learning algorithm. The natural choice in this work is of course using the linear-SVM. To do this, MAE was evaluated obtained by the

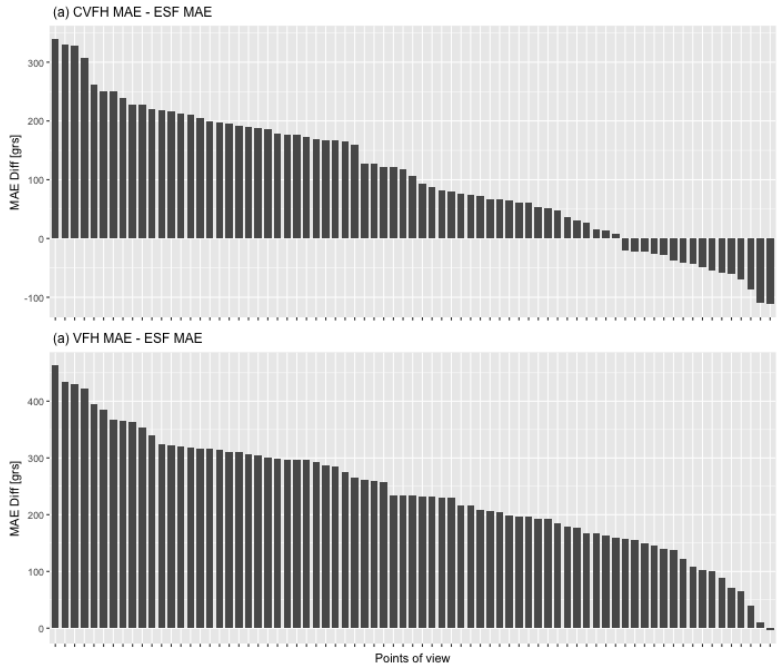

Figure 2: (a) shows the difference of CVFH MAE minus ESF MAE for each of the tested points of view; (b) shows the difference of VFH MAE minus ESF MAE for each of the tested points of view.

linear-SVM over the simulated training and testing sets obtained from the 375 training captures and the 125 testing captures, respectively, over 75 different points of view. Seventy-five points were considered evenly distributed over a 40cm by 40cm by 20cm grid within the prism enclosed by vertices (X=20cm, Y=20cm, Z=−20cm) down to the 3D point (X=−20cm, Y=−20cm, Z=0cm), evenly distributed every 10 cm over the three directions. Three pairs of training and testing examples per point of view were produced, one for each of the three geometric features: VFH, CVFH, and ESF. The experiments were also run with a four-fold cross-validation tuning over hyperparameter C, considering the following as possible values: $C = \{2 \times 10^{-9}, 2 \times 10^{-7}, 2 \times 10^{-5}, 2 \times 10^{-3}, 2 \times 10^{-1}, 2, 2 \times 10^{1}, 2 \times 10^{3}, 2 \times 10^{5}, 2 \times 10^{7}, 2 \times 10^{9}\}$.

Figure 2 shows the difference in the MAE of VFH and CVFH descriptors with ESF for each of the 75 points of view where bars over 0g mean that ESF performed better for the corresponding point of view. Figure 2 (a) shows that ESF performed better thant CVFH in most of the tested points of view. Also Figure 2 (b) shows that ESF obtained a better MAE than VFH in almost all the tested points of view. Based on these results ESF was used as the geometric descriptor of choice for the remaining experiments.

3.3. Selecting optimal data-augmentation parameters

In the following experiments, learning scenarios over four possible *measurement points of view* were considered as depicted in Figure 1, with all of them pointing to the origin of the coordinate system:

- (0,0,-0.2) corresponding to 20cm in front of the biceps,
- (0,-0.2,-0.2) corresponding to 20cm in front and 20cm over the biceps,
- (-0.2,0,-0.2) corresponding to 20cm in front and 20cm to the left of the biceps,
- (0.2,0.2,-0.2) corresponding to 20cm in front, 20cm to the right, and 20cm below the biceps.

To produce the training sets for these learning scenarios, the authors started by $n = \{1, 2, 3\}$ of a total of three captured training bursts and simulated n 3D point clouds for each of these measurement points of view. Next, for each of these possible scenarios, the training set was built by augmenting the captured set through the simulation of new 3D point clouds for points

of view in the spatial surroundings of the original captures. To do this, N points of view were sampled uniformly within a sphere of radius *trainRadius* centered at the *measurement point of view*. Then, for each of the N sampled points of view and each n training burst, two point clouds were sampled from which two simulated point clouds for that point of view were generated, resulting in $2 \times n \times N$ simulated point clouds per measurement point of view. Given all these simulated point clouds, the training set was generated by computing the vector of geometric features for each and appending them the weight lifted for each case as ground-truth label. The following experiments report results for different values of n, N, and $trainRadius$.

The test sets were also generated through a simulation stage, this time using randomly two clouds of the fourth burst reserved for testing purposes. However, in this case 100 points of view were sampled, not from a sphere, but rather a circumference of radius $testRadius$ centered at the measurement point of view whose plane is perpendicular to the line of vision, that is, the line that connects the location of the measurement point of view with the origin of the coordinate system. This choice of testing scenario is one that represents the expected errors in the placement of the Kinect, for which distance is regarded as much easier to guarantee than the angle. As for the training set, from all these simulated point clouds, the test set was generated by computing the vector of geometric features for each and appending them the weight lifted for each case as ground-truth label.

3.3.1. Convergence over n and N

For the number of n bursts and N simulated points of view, both influencing the size of the training set, convergence must be proven, from which one could conclude that enough training data points have been produced. For n a low convergence value is expected, as those bursts must be captured manually for each of the patients, who often have some sort of ailment. In contrast, the N simulated captures are obtained computationally, so there is no major impact on the practicality of the method for high N values. As shown below in detail, for all four measurement points of view, convergence was reached for $n = 3$ bursts and over $N = 1000$ in most cases, with both values within ranges that make the approach practical.

Figure 3 proves empirically that three bursts of 25 captures each are sufficient for convergence when the training set is augmented with $N = 1000$ simulated points of view in a sphere of $trainRadius = 2$cm. The figure shows four bar charts, one per measuring point of view, with each showing

the MAE for the three possible values $n = 1, 2, 3$ (represented in bars of increasing shades of grey) over five testing scenarios for increasing values of *testRadius*: 3mm, 2cm, 4cm, 7cm, and 12cm. As observed in Figure 3, in all cases there is a marked decrease in MAE (quality improvement) between one and two bursts, whereas no major improvement or even decrease in quality (mainly for large values of *testRadius*) occurs between two and three bursts. This convergence is a clear sign that more bursts would show no significant improvement.

To decide how many N simulated points of view are enough, *trainRadius* was also fixed to 2cm and MAE was reported for an increasing number of simulated training points of view within the corresponding sphere, namely, $N = \{20, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000\}$ for all four measurement points of view. To further improve representativeness, ten training sets per N were generated. For each training set, a mapping was learned and tested for five scenarios of different *testRadius*: 3mm, 2cm, 4cm, 7cm, and 12cm. Figure 4 shows the results of each of the four capture points of view showing each of the five curves, one per *testRadius*. Each figure plots MAE mean and standard deviation for each of the five testing scenarios.

The immediate conclusion is that more data points indeed decrease MAE, at least for small N, together with a convergence over approximately $N = 500$ for which despite some minor oscillations, the MAE mean remains constant. This corresponds to the fact that at approximately $N = 500$, no further improvement in mapping quality can be expected.

3.3.2. Impact of trainRadius on method performance

The experiments conducted so far fixed the *trainRadius* to a value of 2cm. This section shows trends for increasing values of this radius, by evaluating how spatially-expanded models learned from training sets have an impact on measurement with potentially misaligned measuring devices. Thus, a training set was produced for $trainRadius = \{0, 0.5, 1.0, 1.5, 2.0\}$cm. A model for each was learned and tested for five scenarios of different *testRadius*: 3mm, 2cm, 4cm, 7cm, and 12cm. For consistency the same density of points per training sphere was maintained rather than the same total number of points (resulting in a cubic increase in the actual N, as volume grows with $trainRadius^3$). $N = 1000$ data points were simulated for the largest value of *trainRadius*=2cm, and then they were simply sub-sampled for the smaller spheres. There follows the MAE for all four measurements points of view in Figure 5, with each figure containing one curve per *testRadius* ranging over

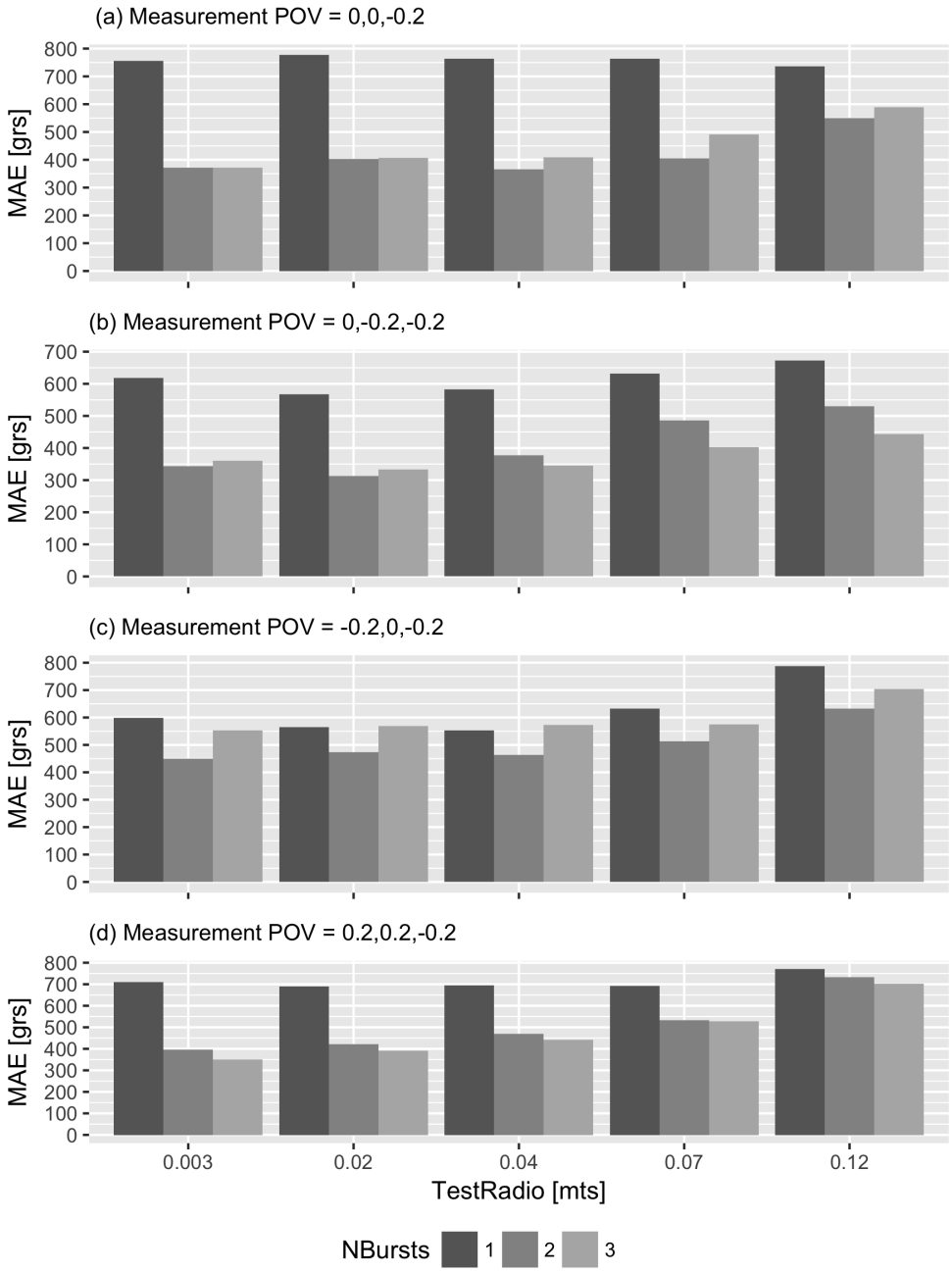

Figure 3: Figures (a), (b), (c), and (d) show the MAE mean for increasing values of n training bursts for $trainRadius = 2$cm and $N = 1000$ over the four measurement points of view.

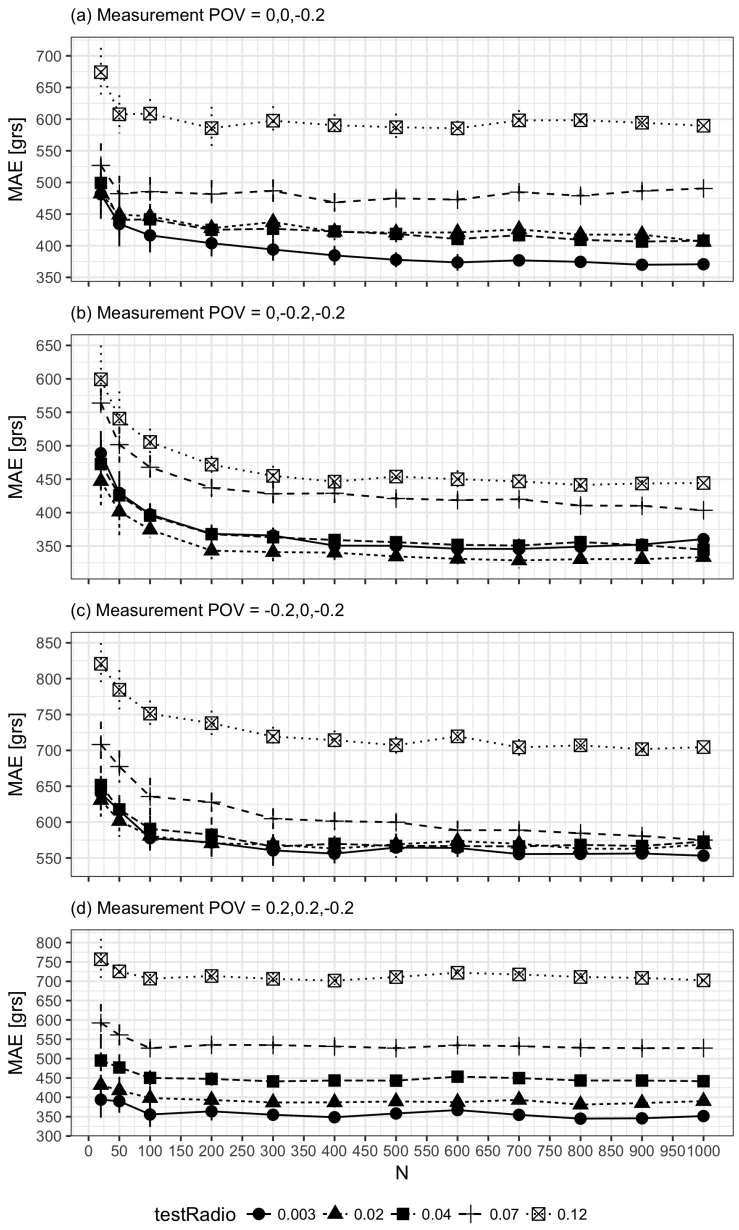

Figure 4: Figures (a), (b), (c) and (d) show the mean and standard deviation of the MAE for a $trainRadius$ = 2cm, increasing values of N over the four measurement points of view, with five curves each corresponding to the five testing scenarios.

all values of $trainRadius$.

As it can be observed, for all four measurement points of view and all $testRadius$ in each figure, the curves show a clear tendency to reduce MAE with increasing $trainRadius$, proving empirically that the spatial expansion of the training set indeed produces a better mapping, as shown by lower values of MAE. Moreover, only the figure corresponding to (-0.2, 0,-0.2) shows convergence, proving through all the other cases that there may be further improvements for even larger spatial expansions of the training set.

4. Discussion

The experimental results presented in Section 3 prove the existence of correlations between the muscular effort and 3D images of the biceps amenable for extraction by autonomous systems that are both non-intrusive (no electrode or cumbersome apparatus required) and simple enough to be easily installed by non-medical technicians to be operated later on by the user. Moreover, the results of our experiments show that this visual measurement approach is capable of finding discriminative patterns between discrete levels of effort, ranging from zero to roughly MVC, with errors low enough to compete with an EMG.

To further this argument, Table 1 shows the values of the above results for $trainRadius$=2cm and $n = 3$ training bursts, the values for these parameters that showed the best results in all cases, all measuring points of view, and all testing scenarios. For each $testRadius$ (column TeR in Table 1), the table shows the performance measurement values at $N^{\star}$ — the value of N with the best MAE. Although the problem tackled in this work is one of classification, in order to be able to somehow compare the results with previous similar works the coefficient of determination R^2 was also calculated by transforming the classes to their values in grams as for the MAE and reported it in this table. The table shows the best results for the point of view of (0,-0.2,-0.2)(Table 1 (b)) corresponding to an upper-front view of the biceps. We can see for this case that measurements of up to 4cm away from the measurement point of view would incur in a MAE smaller than 350g. On the contrary, measurements up to 12cm away from the measurement point of view would incur in a MAE smaller than 450g.

Although no equal works that estimate activation level exclusively from 3D point clouds were found, the errors obtained by the proposed method are

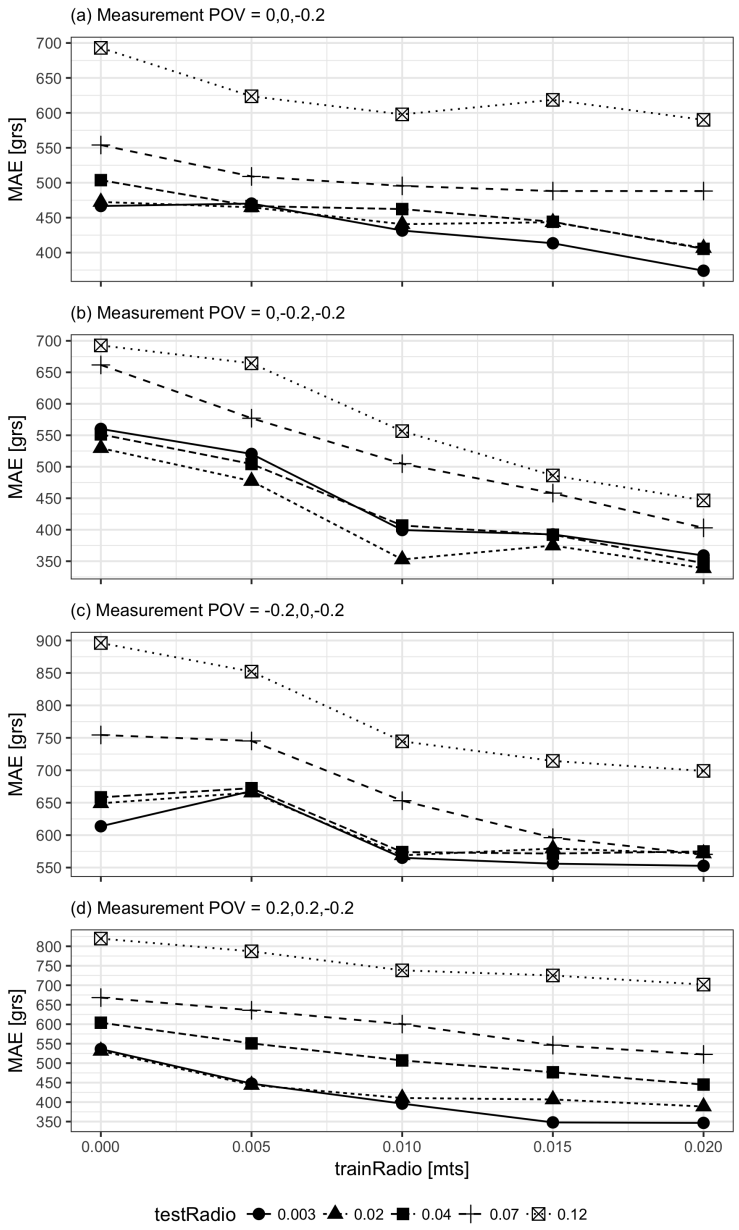

Figure 5: Figures (a), (b), (c), and (d) show the MAE mean for increasing values of *trainRadius* over the four measurement points of view, with five curves each corresponding to the five testing scenarios.

equivalent to those reported for EMG force measurements [43, 44]. Considering that in the problem under study, the MVC corresponds to the highest level of activation measured (class $O4k$) – a weight of 3885g — the errors obtained by this method are between 9% (350g) and 11% (450g) of the MVC. Interestingly, these errors are equivalent to those reported for EMG force measurements [43, 44] that range between 9% and 10% of MVC values .

Similar to this work, there is a group of works that tackle the problem of predicting muscle activity, but from kinematic and kinetic information. Some of the most recent works of this group are [45, 46, 47, 48]. The work in [45] compares a linear logistic regression model with artificial neural networks and other prediction models in the task of predicting the activity of six muscles in the right lower extremity. As input to the models, they provide joint angles and external forces. The best results show a determination coefficient of 0.42 between the predicted and real activity. In [46] artificial neural networks are tested to predict the EMG activity of 12 arm muscles. Inputs to the models include hand position, hand orientation, and thumb grip force. The best results present a determination coefficient of 0.6, and this measurement specifically for the biceps is 0.5. In [47] regression equations are built to predict the 10th percentile, the median, and the 90th percentile of muscle activity around the shoulder joint, given an arm posture and net shoulder moments. The determination coefficient of all regression equations ranges between 0.228 and 0.818. Finally in [48] a multi-dimensional wavelet neural network is proposed to predict human lower extremity muscle activity based on ground reaction forces and joint angles. The best results show a determination coefficient above 0.9.

For the reviewed works, kinematic information is highly informative mainly because they analyze movement. However, it could be the case that for the same posture or movement, the level of contraction may be different, and using only kinematics to infer activation may fail. For this reason, the mentioned works add external force information (kinetics) as input to their models. Although direct force measurement is helpful, in a tele-physiotherapy setting like the one considered in this work, adding force sensors could be an extra and impractical requirement to fulfill. To avoid the need for these sensors, the proposed approach measures the effect of external forces on activation in an indirect and non-invasive way by considering the arm's surface deformation. Additionally, since static postures are being analyzed including kinematic information, in this approach the latter is not necessary as it is a constant value. However, if movement were to be analyzed, kinetic

information must definitely be included.

Although these works estimate the continuous level of muscle activation, since the problem posed is one of classification, in order to be able to somehow compare the results of this work with similar previous works, the coefficient of determination R^2 was calculated by transforming classes to their values in grams as for MAE. For the best measurement point of view (0,-0.2,-0.2) (Table 1 (b)), R^2 values for each $testRadius$ are: $R^2 = 0.67$ for $testRadius =$ 0.3cms; $R^2 = 0.7$ for $testRadius =$ 2cms; $R^2 = 0.68$ for $testRadius =$ 4cms; $R^2 = 0.64$ for $testRadius =$ 7cms and $R^2 = 0.58$ for $testRadius =$ 12cms. The presented works report determination coefficients between 0.228 and 0.9. Here it is worth mentioning that exclusively for the biceps, the value obtained by [46] was 0.5. On the contrary, the reported determination coefficient for the proposed work is between 0.58 and 0.7. With the exception of some cases where these works report correlations above 0.7, the results of this method are comparable with those reported by the reviewed works. These results are encouraging considering that, although discrete levels for static postures are estimated, the task faced is more challenging. This consideration is based on the fact that kinematic together with kinetic information define almost completely the value of muscle activation, and is therefore, much more informative than geometric information only.

With regard to point cloud description, another group of works have been dedicated to solving the problem of characterizing skin deformation caused by underlying muscle contraction both from 2D [49, 50] and 3D images [51]. The problem solved by these methods differs from that solved by the proposed method in that they do not attempt to estimate muscle activation level, but rather seek to build a representation of deformation due to muscle contraction, without prediction. From these works, [51] is the closest to this approach since they also prove the correlation of geometrical deformation of the arm's surface with underlying muscle activity, although they do not pretend to validate the effectiveness of their proposed method to quantify and predict muscle activation. The method proposed in this work to characterize arm deformation could be used in the future as another way to characterize the approach input point clouds.

From this analysis, it can be concluded that the work proposed in the present study constitutes a positive result toward a practical measurement of biceps muscle activation apparatus in the context of tele-rehabilitation.

TeR	$N^\star$	**MAE**	R^2
0.3	900	$369.98_{(7.31)}$	$0.63_{(0.01)}$
2	1000	$406.20_{(0)}$	$0.6_{(0)}$
4	900	$406.62_{(4.57)}$	$0.6_{(0)}$
7	400	$468.55_{(14.51)}$	$0.52_{(0.02)}$
12	600	$585.53_{(12.10)}$	$0.39_{(0.01)}$

(a) Measurement POV = 0,0,-0.2

TeR	$N^\star$	**MAE**	R^2
0.3	700	$345.73_{(15.42)}$	$0.67_{(0.01)}$
2	700	$328.55_{(11.73)}$	$0.7_{(0.01)}$
4	1000	$344.60_{(0)}$	$0.68_{(0)}$
7	1000	$403.50_{(0)}$	$0.64_{(0)}$
12	800	$441.43_{(9.34)}$	$0.58_{(0.01)}$

(b) Measurement POV = 0,-0.2,-0.2

TeR	$N^\star$	**MAE**	R^2
0.3	1000	$553.20_{(0)}$	$0.43_{(0)}$
2	900	$562.80_{(3.5)}$	$0.42_{(0)}$
4	300	$565.78_{(17.5)}$	$0.45_{(0.02)}$
7	1000	$574.95_{(0.6)}$	$0.43_{(0)}$
12	900	$701.76_{(8.21)}$	$0.25_{(0.01)}$

(c) Measurement POV = -0.2,0,-0.2

TeR	$N^\star$	**MAE**	R^2
0.3	800	$344.98_{(7.71)}$	$0.6_{(0.01)}$
2	800	$381.12_{(12.34)}$	$0.56_{(0.01)}$
4	300	$441.20_{(12.99)}$	$0.5_{(0.02)}$
7	900	$527.02_{(8.32)}$	$0.41_{(0.01)}$
12	400	$701.68_{(11)}$	$0.13_{(0.03)}$

(d) Measurement POV = 0.2,0.2,-0.2

Table 1: MAE and R^2 of the approach for $n = 3$ train bursts and $trainRadius = 2cm$ for the best performing N for each $testRadius$ (column TeR, reported in cms). Tables (a), (b), (c) and (d) for each of the measurement points of view.

5. Conclusions and future work

This work presents the first steps towards the estimation of the arm muscle activation level from biceps 3D point clouds using Computer Vision and Machine Learning. Its main objective is the early exploration of these technologies for the measurement of muscle activation level remotely from external arm images in order to satisfy the requirements of telemedicine settings. The solution to the specific problem tackled in this work for measuring discrete levels of biceps activation in isometric contraction could be put into practice with small improvements in the monitoring of intensity and time of muscle contraction in biceps isometric contraction tele-rehabilitation exercises and as a way to automate the execution of the well-known *manual muscle test* for measuring biceps strength. The proposed approach consists in a framework that considers the generation of a feature vector for the point clouds of the muscle area using the ensemble of shape functions 3D geometric descriptors for a given 3D image captured at some recommended point of view and its spatial neighborhood. It also uses the training of a model for autonomously estimating the level of effort of new examples using the support vector machine supervised machine-learning technique. These early results are the first steps in this line of research that aims to develop a new technol-

ogy for remote muscle activity sensing as a tool to improve musculoskeletal system biomechanical sensing. One important finding of this work is that spatially augmenting the amount and variability of training data helps to increase estimation accuracy, or in other words, helps to reduce the number of actual captures required for training a model with enough accuracy. Although this is a solid first step toward achieving a practical application in the future, the efficiency obtained is limited to settings of isometric contraction and to the biceps muscle.

Future work will entail evaluating the methodology by applying deep learned features instead of the hand crafted ones used in this work, or even replacing the whole pipeline with a deep learning approach that integrates the whole cloud characterization and model learning process into one deep neural network. Alternatively, the proposed framework could be tested using [51] as a feature to characterize the point clouds.

Acknowledgment We thank the *Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)* for the funding of the author Leandro Abraham through a doctoral fellowship. This work was also supported by the *Fondo para la investigación científica y tecnológica (FONCYT)* of the Argentinian government [grant number PICT-2014-0567].

Conflict of interest statement

The authors of this work disclose no financial or personal relationships with other people or organizations that could influence or bias the present work.

[1] V. N. Vapnik, V. Vapnik, Statistical learning theory, volume 1, Wiley New York, 1998.

[2] C. M. Bishop, et al., Pattern recognition and machine learning, volume 1, springer New York, 2006.

[3] W. Wohlkinger, M. Vincze, Ensemble of shape functions for 3d object classification, in: Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on, IEEE, pp. 2987–2992.

[4] M. M. Vollenbroek-Hutten, H. J. Hermens, R. Kadefors, L. Danneels, L. J. Nieuwenhuis, M. Hasenbring, Telemedicine services: from idea to implementation (2010).

[5] C. Schönauer, T. Pintaric, H. Kaufmann, S. Jansen-Kosterink, M. Vollenbroek-Hutten, Chronic pain rehabilitation with a serious game using multimodal input, in: Virtual Rehabilitation (ICVR), 2011 International Conference on, IEEE, pp. 1–8.

[6] R. G. Bults, D. F. Knoppel, I. A. Widya, L. Schaake, H. J. Hermens, The myofeedback-based teletreatment system and its evaluation, Journal of telemedicine and telecare 16 (2010) 308–315.

[7] M. Rogante, S. Silvestri, M. Grigioni, M. Zampolini, Electromyographic audio biofeedback for telerehabilitation in hospital, Journal of telemedicine and telecare 16 (2010) 204–206.

[8] L. L. Andersen, C. H. Andersen, O. S. Mortensen, O. M. Poulsen, I. B. T. Bjørnlund, M. K. Zebis, Muscle activation and perceived loading during rehabilitation exercises: comparison of dumbbells and elastic resistance, Physical therapy 90 (2010) 538–549.

[9] C. J. De Luca, The use of surface electromyography in biomechanics, Journal of applied biomechanics 13 (1997) 135–163.

[10] C. Kisner, L. A. Colby, J. Borstad, Therapeutic exercise: foundations and techniques, Fa Davis, 2017.

[11] M. L. Palmer, M. E. Epler, M. F. Epler, Fundamentals of musculoskeletal assessment techniques, Lippincott Williams & Wilkins, 1998.

[12] H. Hislop, D. Avers, M. Brown, Daniels and Worthingham's Muscle Testing-E-Book: Techniques of Manual Examination and Performance Testing, Elsevier Health Sciences, 2013.

[13] Q. Shao, T. S. Buchanan, Electromyography as a tool to estimate muscle forces, Standard handbook of biomedical engineering & design (2004).

[14] K. T. Manal, T. S. Buchanan, Biomechanics of human movement, Standard handbook of biomedical engineering & design (2004) 26.

[15] J. Cholewicki, S. M. McGill, Emg assisted optimization: a hybrid approach for estimating muscle forces in an indeterminate biomechanical model, Journal of biomechanics 27 (1994) 1287–1289.

[16] J. Cholewicki, S. M. McGill, R. W. Norman, Comparison of muscle forces and joint load from an optimization and emg assisted lumbar spine model: towards development of a hybrid approach, Journal of biomechanics 28 (1995) 321–331.

[17] L. Vigouroux, F. Quaine, A. Labarre-Vila, D. Amarantini, F. Moutet, Using emg data to constrain optimization procedure improves finger tendon tension estimations during static fingertip force production, Journal of biomechanics 40 (2007) 2846–2856.

[18] D. Amarantini, G. Rao, E. Berton, A two-step emg-and-optimization process to estimate muscle force during dynamic movement, Journal of Biomechanics 43 (2010) 1827–1830.

[19] D. B. Chaffin, G. Andersson, B. J. Martin, et al., Occupational biomechanics, Wiley New York, 1984.

[20] S. Kumar, Biomechanics in ergonomics, CRC Press, 1999.

[21] P. McGinnis, Biomechanics of sport and exercise, Human Kinetics, 2013.

[22] T. Dutta, Evaluation of the kinect sensor for 3-d kinematic measurement in the workplace, Applied ergonomics 43 (2012) 645–649.

[23] BioVirtua, Biovirtua — humanizing telehealth, 2017.

[24] Respondwell, Respondwell rehabilitation, 2017.

[25] M. A. Brubaker, L. Sigal, D. J. Fleet, Estimating contact dynamics, in: 2009 IEEE 12th International Conference on Computer Vision, IEEE, pp. 2389–2396.

[26] T. Pham, N. Kyriazis, A. A. Argyros, A. Kheddar, Hand-object contact force estimation from markerless visual tracking, 2016. Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence.

[27] T.-H. Pham, A. Kheddar, A. Qammaz, A. A. Argyros, Towards force sensing from vision: Observing hand-object interactions to infer manipulation forces, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2810–2819.

[28] N. Chen, S. Urban, J. Bayer, P. Van Der Smagt, Measuring fingertip forces from camera images for random finger poses, in: Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on, IEEE, pp. 1216–1221.

[29] N. Chen, S. Urban, C. Osendorfer, J. Bayer, P. Van Der Smagt, Estimating finger grip force from an image of the hand using convolutional neural networks and gaussian processes, in: Robotics and Automation (ICRA), 2014 IEEE International Conference on, IEEE, pp. 3137–3142.

[30] T. R. Grieve, J. M. Hollerbach, S. A. Mascaro, Optimizing fingernail imaging calibration for 3d force magnitude prediction, IEEE transactions on haptics 9 (2016) 69–79.

[31] T. R. Grieve, C. E. Doyle, J. M. Hollerbach, S. A. Mascaro, Calibration of fingernail imaging for multidigit force measurement, in: Haptics Symposium (HAPTICS), 2014 IEEE, IEEE, pp. 623–627.

[32] R. B. Rusu, G. Bradski, R. Thibaux, J. Hsu, Fast 3d recognition and pose using the viewpoint feature histogram, in: Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, IEEE, pp. 2155–2162.

[33] A. Aldoma, M. Vincze, N. Blodow, D. Gossow, S. Gedikli, R. B. Rusu, G. Bradski, Cad-model recognition and 6dof pose estimation using 3d cues, in: Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, IEEE, pp. 585–592.

[34] R. B. Rusu, S. Cousins, 3D is here: Point Cloud Library (PCL), in: IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China.

[35] J. H. Lawrence, C. De Luca, Myoelectric signal versus force relationship in different human muscles, Journal of Applied Physiology 54 (1983) 1653–1659.

[36] P. Sbriccoli, I. Bazzucchi, A. e. a. Rosponi, M. Bernardi, G. De Vito, F. Felici, Amplitude and spectral characteristics of biceps brachii semg depend upon speed of isometric force generation, Journal of Electromyography and Kinesiology 13 (2003) 139–147.

[37] S. Katz, A. Tal, R. Basri, Direct visibility of point sets, in: ACM Transactions on Graphics (TOG), volume 26, ACM, p. 24.

[38] S. J. De Serres, R. M. Enoka, Older adults can maximally activate the biceps brachii muscle by voluntary command, Journal of Applied Physiology 84 (1998) 284–291.

[39] S. Baudry, T. Rudroff, L. A. Pierpoint, R. M. Enoka, Load type influences motor unit recruitment in biceps brachii during a sustained contraction, Journal of neurophysiology 102 (2009) 1725–1735.

[40] J. R. Cameron, J. G. Skofronick, R. M. Grant, Physics of the Body, 2nd Edition, Medical Physics Publishing Corporation, 1999.

[41] G. Pang, U. Neumann, 3d point cloud object detection with multi-view convolutional neural network, in: Pattern Recognition (ICPR), 2016 23rd International Conference on, IEEE, pp. 585–590.

[42] A. Velizhev, R. Shapovalov, K. Schindler, Implicit shape models for object detection in 3d point clouds, in: International Society of Photogrammetry and Remote Sensing Congress, volume 2.

[43] P. Spielholz, B. Silverstein, M. Morgan, H. Checkoway, J. Kaufman, Comparison of self-report, video observation and direct measurement methods for upper extremity musculoskeletal disorder physical risk factors, Ergonomics 44 (2001) 588–613.

[44] J. Potvin, S. Brown, Less is more: high pass filtering, to remove up to 99% of the surface emg signal power, improves emg-based biceps brachii muscle force estimates, Journal of Electromyography and Kinesiology 14 (2004) 389–399.

[45] M. Sekiya, T. Tsuji, Inverse estimation of multiple muscle activations based on linear logistic regression, in: Rehabilitation Robotics (ICORR), 2017 International Conference on, IEEE, pp. 935–940.

[46] R. Tibold, A. J. Fuglevand, Prediction of muscle activity during loaded movements of the upper limb., Journal of neuroengineering and rehabilitation 12 (2015) 6–6.

[47] X. Xu, R. W. McGorry, J.-H. Lin, A regression model predicting isometric shoulder muscle activities from arm postures and shoulder joint moments, Journal of Electromyography and Kinesiology 24 (2014) 419–429.

[48] M. Mosafavizadeh, L. Wang, Q. Lian, Y. Liu, J. He, D. Li, Z. Jin, Muscle activity prediction using wavelet neural network, in: Wavelet Analysis and Pattern Recognition (ICWAPR), 2013 International Conference on, IEEE, pp. 241–246.

[49] M. Carli, M. Goffredo, M. Schmid, A. Neri, Study of muscular deformation based on surface slope estimation, in: Electronic Imaging, International Society for Optics and Photonics, pp. 60640U–60640U.

[50] M. Goffredo, M. Carli, S. Conforto, D. Bibbo, A. Neri, T. D'Alessio, Evaluation of skin and muscular deformations in a non-rigid motion analysis, in: Medical Imaging, International Society for Optics and Photonics, pp. 535–541.

[51] M. Goffredo, P. Piras, V. Varano, S. Gabriele, C. D'Anna, S. Conforto, Shape analysis of bicipital contraction by means of rgb-d sensor, parallel transport and trajectory analysis, in: XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016, Springer, pp. 634–639.