

Textile Taxonomy and Classification Using Pulling and Twisting

Alberta Longhini, Michael C. Welle, Ioanna Mitsioni and Danica Kragic

Abstract—Identification of textile properties is an important milestone toward advanced robotic manipulation tasks that consider interaction with clothing items such as assisted dressing, laundry folding, automated sewing, textile recycling and reusing. Despite the abundance of work considering this class of deformable objects, many open problems remain. These relate to the choice and modelling of the sensory feedback as well as the control and planning of the interaction and manipulation strategies. Most importantly, there is no structured approach for studying and assessing different approaches that may bridge the gap between the robotics community and textile production industry. To this end, we outline a textile taxonomy considering fiber types and production methods, commonly used in textile industry. We devise datasets according to the taxonomy, and study how robotic actions, such as pulling and twisting of the textile samples, can be used for the classification. We also provide important insights from the perspective of visualization and interpretability of the gathered data.

I. INTRODUCTION

Interaction with deformable objects is an integral part of our everyday life but still a challenge for robotic systems. Work on robotic handling of textile or fabric traces back several decades [1] and, despite the clear need in industrial and domestic applications, many of the problems related to perception, planning and control remain open. From the industrial perspective, textile production and subsequent processes of garment design in fashion industry are largely not automated. Fashion industry is also undergoing an important transformation to address sustainability concerns, given that textile and clothing overproduction has a significant negative impact on the environment.

From the scientific perspective, robotic interaction with deformable materials has gained significant attention recently [2], [3]. Important milestones regarding the modelling, perception, planning, control and simulation of deformable materials have been identified but not yet reached. It may even be so that until robots reach the dexterity, flexibility and sensing that to some extent resembles human capabilities, successful interaction with deformable objects will remain a challenge. In robotics, textile has been used to study manipulation tasks like folding [3]–[5], robot-assisted dressing [6]–[8], garment recognition and classification [9]–[12]. In most of these works, only a subset of textile properties is commonly considered, and textile is merely a tool for testing sensors [13] or control strategies [14].

In our work, we aim to study textile materials and their properties using physical interactions and wrist-mounted force-torque sensing. Similarly to humans, we aim at using

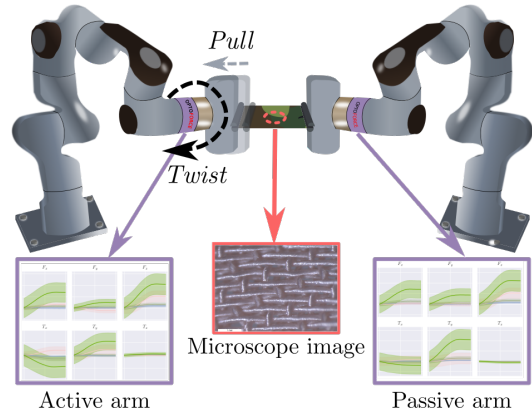


Fig. 1: System setup: Two 7 DoF Franka Emika Panda arms with force-torque sensors on the wrists, twisting and pulling a textile sample. Microscope images are used to define textile classes prior to training.

actions such as pulling and twisting, to learn more about the textile properties, see Fig. 1. The properties are defined using a textile taxonomy that follows the classification used in the textile production industry. Textile properties in general, and thus interaction dynamics, are affected by factors such as fiber material and production method - the fiber may be raw, coated or it may be a blend of several materials. Once used to produce garments or bed-clothing, the properties will change overtime based on washing, wearing, steaming - the textile can become harder or softer, less or more elastic, thinner. The change in properties will also have an impact on the planning and control strategies used to interact with it - how we wash, iron and fold them, how we hold and manipulate garments when dressing somebody, whether we decide to recycle or reuse old garments.

The focus of this paper is to assess how a dual arm robotic system can be used to identify textile production methods through pulling and twisting. We propose to do so by learning a classifier on a dataset of textile samples that are annotated by their construction type, determined by inspecting their microscopic structure. We first outline a textile classification taxonomy related to both fiber type and production method, following notation used in textile industry. We then make a thorough study using a subset of materials and production techniques to assess the validity of our approach. We analyse two manipulation strategies as well as investigate which measurements are most relevant for classification. We conclude by discussing challenges and open problems.

¹The authors are with the Robotics, Perception and Learning Lab, EECS, at KTH Royal Institute of Technology, Stockholm, Sweden albertal@kth.se, mwelle@kth.se, mitsioni@kth.se, dani@kth.se

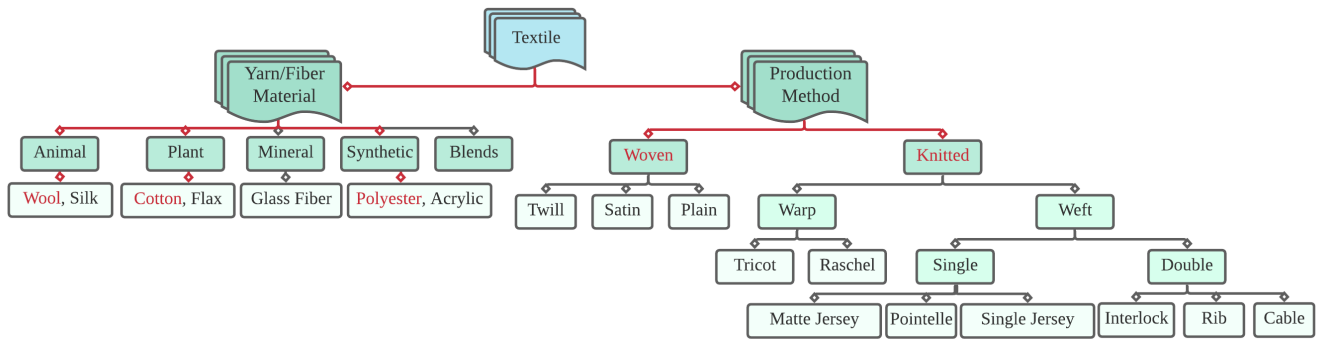


Fig. 2: Textile taxonomy considering yarn/fabric material and production method. Classes considered in this work are highlighted in red.

II. TEXTILE TAXONOMY

Textile or fabric, is a deformable and flexible material made out of yarns or threads, which are put together by a construction or manufacturing process such as weaving, knitting, crocheting, knotting, tating, felting, bonding or braiding. Most of the everyday clothing items we wear are constructed through weaving or knitting prior to sewing, although in high fashion other processes are used frequently too. Yarns and threads are produced by spinning raw fibers that may have different origin: animal, plant, mineral, synthetic or their blend. We summarize some of these aspects in Fig. 2, showing also some of the further distinctions in terms of differences in the manufacturing process.

Woven fabric is usually produced by using two sets of yarn, while knitted fabric employs a single set. To produce woven fabric, the yarn is interlaced, as opposed to knitting where it is interloped. Due to its construction, woven fabric is often hard and nonelastic, allowing it to hold creases well, and can be only stretched diagonally if the yarn itself is not a blend that includes elastic material. It is commonly used to produce garments such as shirts and jeans. On the other hand, knitted fabric is soft and can be stretched in all directions, making it rather wrinkle-resistant. One example of its frequent use is for t-shirts. One important aspect is that it usually does not stretch equally in all directions - for example, a t-shirt will stretch more horizontally than vertically to more naturally follow body shape. A more complete account of the properties of fabric can be found in [15].

The above is of importance for various robotics applications that consider active interaction with the textile. For assisted dressing applications, it is important for the robotic system to generate relevant control strategies when pulling up pants, helping with the sleeves or pulling down the t-shirt: hard textile may require completely different manipulation strategies and safety considerations than the flexible one. Most of the clothing items will have a content label attached to them and may offer information about the type of yarn used. However, the manufacturing process is never described on the label, neither is the fact that a clothing item may be a combination of woven and knitted parts and a combination

of different type of textiles being put together in the sewing step. For example, both a pair of jeans and a t-shirt may be made with 100% cotton textile that, in the first case, is woven and in the second case, knitted. With the proposed taxonomy and the work in this paper, we report on some initial insights on how some of the textile properties can be examined by using force-torque measurements and actions such as pulling and twisting of the textile samples.

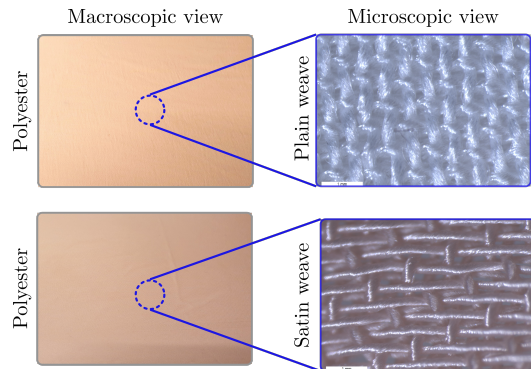


Fig. 3: Textile can look the same given regular camera (left), but very different on the microscopic level (right). The different weaving styles (plain weave and satin weave) with the same material determine the dynamical properties that are important when manipulating fabric.

Even for humans, the manufacturing method may not be visible with the naked eye and the label only provides the yarn material. We use our experience of previously interacting with clothing items to choose appropriate actions when dressing ourselves or others, washing, ironing, repairing or sewing. To shed some light on this, we collected microscope images of our textile samples. While a regular camera image may not provide enough signal resolution in actual robot interaction with the textile, the high-resolution microscope images show different ways of interlacing yarn that has a huge effect on the elasticity of the textile Fig. 3. Similarly, garments with different reflection or texture properties may look rather different under a regular camera but their dynamical properties may be the same if the manufacturing method and yarn type are the same, see Fig. 4.

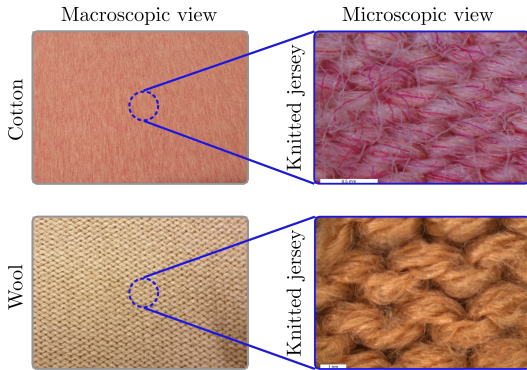


Fig. 4: Textile can look different on the macro level, but very similar on the micro level. Left) two different materials (Wool and Cotton) using a regular camera. The microscope image (right) reveals that they have the same underlying construction - knitted jersey.

In this paper, we therefore set out to investigate how force-torque measurements together with actions such as pulling and twisting may be used to classify the textiles according to the proposed taxonomy. We chose pulling and twisting since these are also the two most common actions humans use for inspecting textile properties.

III. RELATED WORK

In robotics, textile materials have been considered from perception, learning, planning and control perspectives. Most notable applications consider folding, assisted dressing or material classification [2], [16].

Despite the broad interest in the computer vision community, most works concentrate on building clothing item taxonomies [9] rather than identifying material properties. Problems such as garment motion prediction [17], classification [10], dressing 3D simulated humans [18], have also been addressed. It has also been shown that wrinkle detection may be helpful for classification [19], [20]. However, with vision alone it may be difficult to estimate the physical attributes of textiles [21] although the results in [22] indicate that vibrations captured in video can be correlated to the stiffness and density of fabrics.

In robotics, identifying textile properties is important, but there is no common taxonomy that allows for comparison and benchmarking of the proposed approaches. Recent work in [23] proposes a taxonomy of 184 materials including leather, fur and plant fiber but there is no focus on textile in particular or the production method. Haptic feedback has often been used to label various types of materials [13], [24], [25]. The authors in [26] study compliance and texture to classify 32 materials including textile. Non-contact techniques have been used in [27] to distinguish among five material categories, one of which was textile. Thus, none of these works focuses specifically on textile material, or considers fiber and production method in particular. When considering textile classification, it has been studied from the fiber material perspective [7], [28] or properties such as thickness, softness and durability [29]. Material texture

identification has been addressed in [30]–[32], without considering the difference between fiber material and textile production method. Given these, we believe that our initial study and outlined taxonomy provides examples of how textile classification can be studied in a more structured manner.

IV. DATA COLLECTION AND DATASET DESIGN

For this initial study, we relied on 40 textile samples, 10 each of polyester and wool, and 20 cotton samples. From the samples, we cut out pieces 40×17 cm in size. Polyester samples are woven and wool samples are knitted. Out of the 20 cotton samples, 10 are woven and 10 are knitted. We cut the pieces so that the yarn direction is along the axis of pulling, with the more elastic direction of stretching being orthogonal to the axis of pulling as can be seen in Fig. 5. Two Franka Emika Panda arms are equipped with wrist-mounted Optoforce 6-axis Force-Torque (FT) sensors and flat 3D-printed grippers.

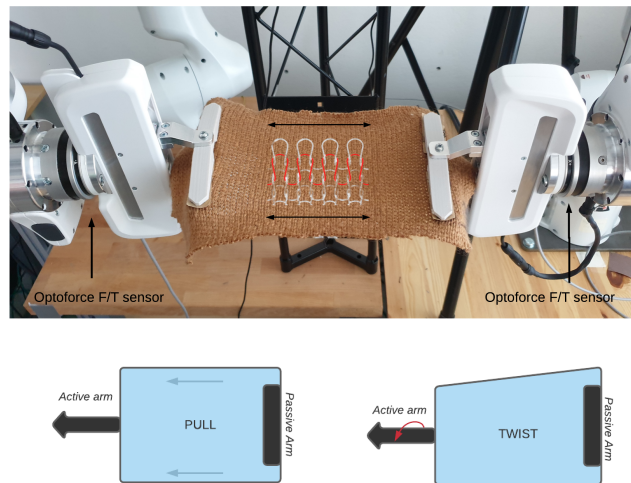


Fig. 5: (Top) Flat 3D-printed grippers holding wool sample. The yarn direction coincides with the pulling direction. FT sensors are mounted on the wrist of the manipulators. (Bottom) Schematic example of *pull* and *twist*.

For data collection, we aimed to capture the samples' properties by means of two exploratory procedures [33], *pulling* and *twisting*, and investigate if they are consistently classifiable. To further analyse different data collection strategies, we decided to let just one arm move (also called *active arm*) while the other one is kept still (*passive arm*), see Fig. 5. A precise definition of these two manipulation actions is the following:

- **Pull** The active arm exerts force on the sample by steadily moving 2 cm away from the static passive arm, maintaining a motion direction parallel to the grasping plane.
- **Twist** The active arm's end-effector rotates 90 degrees, while the arm is pulling to ensure the sample is stretched adequately to capture its reaction to torsion.

Each textile sample was held with a grasping force of 20N by both robot arms and pulled and twisted 20 times. As each sequence of pulling and twisting may result in a slight offset of the contact point, we re-positioned the sample in the hand to the original starting points after every five pulls/twists. Force and torque signals were recorded for a duration of 2s from each sensor at a frequency of 1kHz. Thus, for each textile sample, we have 20 examples, one for each arm, with $2 \times 1000 \times 6$ raw FT measurements.

A. Dataset design

We first sub-sampled the raw measurements for each FT dimension. We performed average downsampling to 150 values as a trade-off between noise reduction and information loss. Given these, we build 6 datasets:

- $\mathcal{D}_{active}^{twist}$, $\mathcal{D}_{active}^{pull}$, $\mathcal{D}_{passive}^{twist}$, $\mathcal{D}_{passive}^{pull}$: 4 datasets corresponding to the two actions for each arm individually.
- \mathcal{D}^{twist} and \mathcal{D}^{pull} : 2 datasets corresponding to the two actions and the integrated measurements from the two arms.

In summary, the datasets were labelled to represent the samples of the taxonomy in Fig. 2 highlighted in red. Therefore, by considering the "Production Method" branch we obtained 2 classes: woven and knitted, while on the "Yarn/Fiber Material" branch, we obtained 3 classes: wool, polyester and cotton.

V. DATA VISUALIZATION AND DATASET INSIGHTS

We first inspect the generated datasets to assess to what extent the collected data, labelled according to the proposed taxonomy, are representative for classification. To this end, we employ t-SNE [34] and project the datasets into a two-dimensional space. More specifically, we want to answer the following questions:

- Can the generated data and employed actions show a clear distinction between woven and knitted textiles?
- Is there a difference between pulling and twisting, in terms of how informative they are, for the woven vs knitted classification?
- What is the effect of fiber type on the classification performance and can we distinguish not only the production method, but also the fiber type given our datasets?

A. Insight 1: Production Method

First, we visually inspect if the datasets are representative of the production method. Fig.6 shows the distribution of the data projected in 2D when only cotton is considered, where the left side of the figures shows how the data are separated by the production method, woven (also called Cotton-Twill) and knitted (also called Cotton-Jersey). The top row corresponds to the active arm and the bottom row to the passive arm during a pulling trial. From the figure we observe that by considering the actual production method we obtain clearly structured groups of samples that would have been otherwise masked by categorizing them as the same material.



Fig. 6: Effect of splitting construction method on Cotton samples. t-SNE plot of measurements obtained from the active (top row) and passive (bottom row) arms during pulling Cotton samples.

B. Insight 2: Pulling vs Twisting

Second, we assess whether there is an advantage in using both pulling and twisting. As a first step, Fig. 7 shows that measurements for the two actions exhibit different behaviors. For example, the force measurements during pulling exhibit more variety than they do for twisting, indicating that they may carry more information for the different classes.

To further examine how important the different measurements are and how they can affect classification, we train a simple SVM [35] classifier to predict the sample's class, while trained on individual signals (F_x , F_y , F_z , T_x , T_y , T_z). The classifier has a linear kernel and it is implemented with Scikit-learn [36]. Table I shows the test set accuracy based on the individual signals for the *pull* (left) and the *twist* (right) strategies of the *active* arm, when learning on a material-based dataset and a construction-based ones.

Signal	Pull		Twist	
	Material	Construction	Material	Construction
F_x	30%	82%	45%	57%
F_y	32%	63%	38%	52%
F_z	70%	100%	53%	73%
T_x	42%	75%	38%	63%
T_y	44%	83%	38%	63%
T_z	40%	53%	45%	59%
All	80%	100%	70%	87%

TABLE I: SVM test set performance based on the individual signals for a dataset with 3 classes for material distinction and 2 classes for the production methods.

From the SVM results, we observe that for the material-based classification, the accuracy scores for the two actions

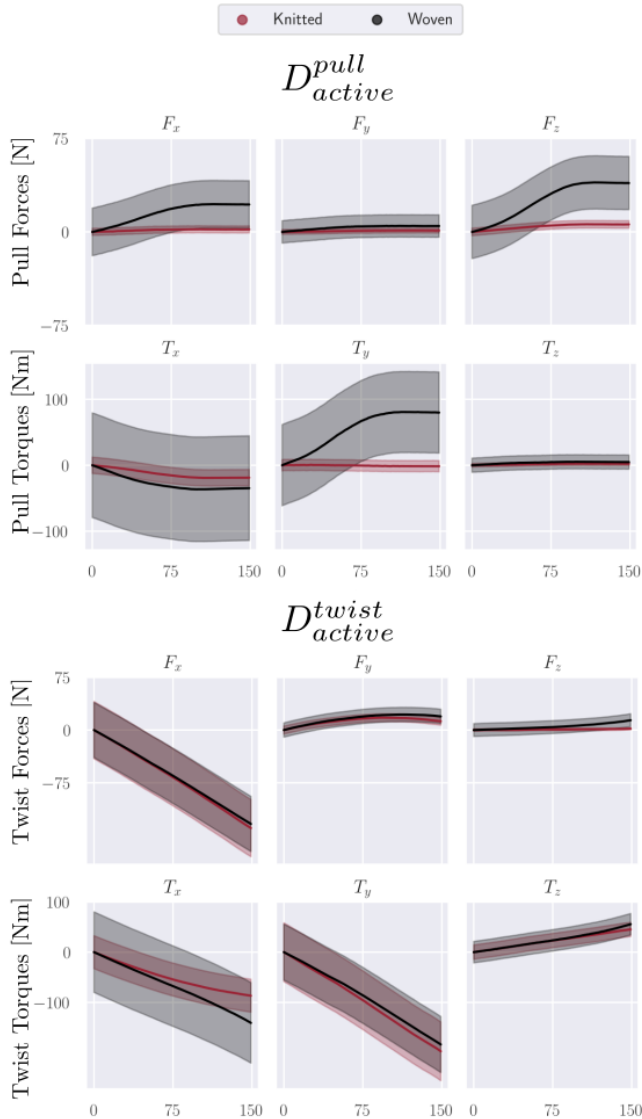


Fig. 7: Pull vs twist measurements on the active arm.

are comparable and rather low. However, when considering the proposed labels, the accuracy increases and in almost all signal cases, pulling outperforms twisting.

These results reinforce that following construction-based taxonomy is advantageous as well as the notion that the sensory feedback varies a lot depending on how textile is manipulated. It is therefore of fundamental importance to understand how to choose the proper exploration strategy. Moreover, as mentioned in Section II, the way in which textile threads are interlocked leads to different elastic properties. Knitted textiles for example, can be stretched lengthwise or along the horizontal direction. Woven textiles instead, are usually not stretchable, apart from a bias direction that for example, for denim is the diagonal one. All these concepts play an important role in classification and highly increase the complexity of the task.

C. Insight 3: Fiber Material vs Production Method

Lastly, we assess to what extent the fiber type can be identified in addition to the production method. An example of this can be seen in Fig. 8 for dataset $\mathcal{D}_{passive}^{twist}$ that depicts the difference between samples categorized using just their production method and sample categorized using both their material and production method. More specifically, the figure portrays the differences between a simple categorization of knitted/woven and a further distinction of the Cotton class, which is split into Cotton-Twill (woven) and Cotton-Jersey (knitted). We can see that Cotton-Twill visually belongs to a separate cluster as observed in section V-A. It can be also noticed that some of the Cotton-Twill samples are closer to Polyester as to Wool, while Cotton-Jersey is closer to Wool than to Polyester.

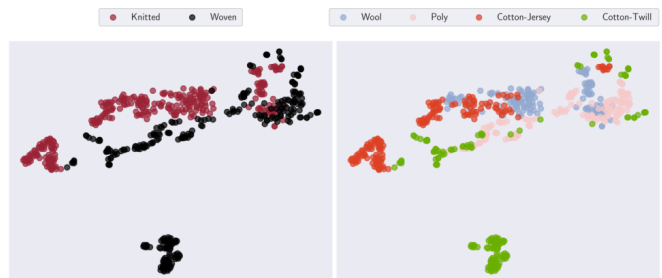


Fig. 8: Twist on the passive arm: visualization with respect to (left) production method, (right) fiber material.

We can also observe the effect of the proposed taxonomy on the individual signal level, by considering the dataset $\mathcal{D}_{active}^{pull}$ for example. Fig. 9 depicts how the split of Cotton by construction method highlights the difference between the mean force used for Cotton-Twill and Cotton-Jersey at the end of the pulling action, further showing the necessity of splitting the Cotton class. Moreover, besides the detectable distinction among Polyester, Wool, Cotton-Twill and Cotton-Jersey signals, woven materials keep being the ones with higher mean force while knitted ones are in general less tension-resistant, reflecting the behaviour of the construction methods.

VI. CLASSIFICATION AND INTERPRETABILITY

The next step is to assess the classification performance using a more complex architecture, like a CNN model. The input of the network are vectors of the six concatenated FT measurements, the network consists of four 2D convolutional layers and its activations are rectified linear units (ReLU). Furthermore, we adopt rectangular kernels of size 5×1 which convolve across measurements of the same signal as done in [37]. The output sizes of the convolutional layers are respectively 24, 12, 8 and 4.

The features learned from the last convolutional layer are flattened and fed to a fully connected layer with 48 hidden neurons. The outputs of this block are the predicted class probabilities. We also consider the case of the joined measurements for the *active* and *passive* arm, using the same

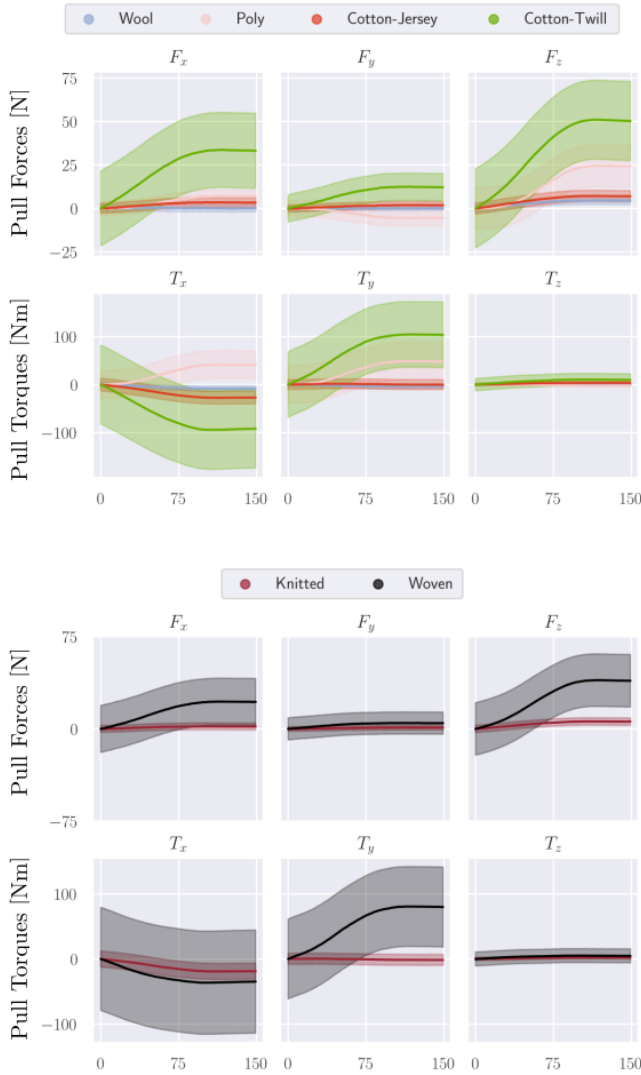


Fig. 9: Mean and standard deviation of the measurements sensed by the active arm while pulling.

architecture but adjusting the size of the fully-connected layer to accommodate the 150×12 signal.

Input Dataset	Materials	Construction
$\mathcal{D}_{active}^{pull}$	87.5%	100%
$\mathcal{D}_{passive}^{pull}$	85.8%	96.7%
$\mathcal{D}_{active}^{twist}$	76.7%	95.0%
$\mathcal{D}_{passive}^{twist}$	78.3%	89.2%
\mathcal{D}^{pull}	95.0%	100%
\mathcal{D}^{twist}	79.0%	91.7%

TABLE II: Test accuracy with a CNN model for all the different datasets when following the material-based labelling and the proposed, construction-based one.

We partitioned each dataset into 90/10 train/test splits. Table II summarizes the classification results. Firstly, we observe that the construction-based labeling outperforms the material-based one for any dataset or action. More specifically, using our taxonomy both actions provide enough information for accurate predictions. However, twisting is consistently less accurate than pulling for all labeling and datasets considered.

These observations are validated in the case of the joined datasets with \mathcal{D}^{pull} achieving excellent performance in distinguishing the construction method of the textile, leading to the conclusion that pulling is a better option for classification.

A. Interpretability and measurement assessment

To further examine the effect of the different measurements for classification, we interpret the results from the CNN model through GradCAM [38]. GradCAM is an interpretability technique that produces visual explanations in the form of heatmaps that portray which parts of the input contribute the most to the predicted label. We follow the same methodology as in [39] to produce and inspect the contribution of each feature in samples from datasets \mathcal{D}^{pull} and $\mathcal{D}_{active}^{pull}$ of Table II.

An example of the heatmaps can be seen in Fig. 10 for two correctly classified samples of woven cotton and knitted cotton from the dataset \mathcal{D}^{pull} . Every row corresponds to a different measurement channel and its color is defined by how important it is for the prediction. The importance is scaled between 0 and 1 and follows the colormap on the right of the images. Fig. 10 shows that for both cotton instances, the network is focusing on the same features between the passive and the active arm. Concretely, for woven cotton, the forces on both axes Z are the most important features, followed by the torques T_x^p, T_x^a and some parts of T_y^p, T_y^a . However, for knitted cotton, the network utilized all the force measurements for both arms and the torque on axis Z for the active arm. These results indicate that even when utilizing the material based-labeling, the CNN network focuses on different patterns when classifying samples of the same material but different construction methods.

Finally, we inspect two classification results from dataset $\mathcal{D}_{active}^{pull}$ when it is labelled according to the proposed taxonomy. The left heatmap corresponds to a correct classification of a woven sample and the right heatmap on the correct classification of a knitted one. The important features agree with the intuition gained from Table I as the decisions are heavily based on the ones highlighted also by the SVM, namely forces F_z and torques T_x, T_y .

We note that the visualization shows only where the neurons of the network are most active for single examples. It is possible for a network to construct multiple patterns to classify the same class, making generalisation difficult. We can however, observe that certain measurements are more important for classification than others, which is a valuable insight when designing future active exploration strategies.

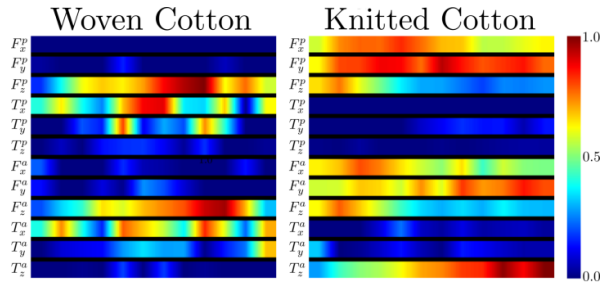


Fig. 10: Heatmaps of feature importance for the classification using dataset \mathcal{D}^{pull} . The intensity of each row (with red being the most important) denotes what the network focuses on to classify the sample.

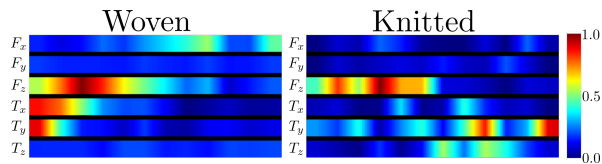


Fig. 11: Heatmaps for the dataset $\mathcal{D}_{active}^{pull}$: activations for a correctly classified Woven sample (left) and a correctly classified Knitted sample (right).

VII. DISCUSSION AND CONCLUSION

In this work, we outlined a textile taxonomy and showed our initial results on textile sample classification using pulling and twisting actions. The focus of the study was to assess to what extent a taxonomy used in textile industry is a viable model to structure robotic interaction and provide a basis for a whole new area of structured studies of this class of deformable materials.

We provided insights on how a combination of different actions and FT measurements vary with respect to textile production method and fiber material. Pulling and twisting, as inspired by the human interaction with textile, are viable choices of actions and they provide relevant information for classification. One interesting question that arises is what other actions can be potentially employed and to what extent dexterous hand/finger motion could be exploited in addition to pulling and twisting. Other works in literature demonstrated the use of specialized fingers and sensors for this purpose and it is yet to be seen to what extent we can consider such solutions to become commercial.

Combining multiple actions, as well as passive and active interaction, is also an interesting aspect to be explored. We may start with pulling/stretching and based on the first step classification, subsequent routines may be performed more suitably for identifying categories of interest, such as for example fiber material, elasticity, whether the textile is wet or dry, etc. Here, reinforcement learning may be used to learn actions that maximize the utility of the sensor readings for discriminating various textile properties.

We also performed an initial study using visual feedback under pulling and twisting. However, for the considered

categories, regular cameras do not provide enough resolution to bring sufficient information on the production method or the fibre type. One could potentially rely on the reflectance properties of textile materials, but most of the works in this area that stem from the computer vision community, are not applicable in uncontrolled settings that would occur in real-life applications. An idea supported by our experimental/empirical observations was the fact that creases and wrinkles on the textile fabric may be a useful feature to exploit for certain applications. When pulling or stretching the fabric in many different directions, creases and wrinkles will vary dependent on the properties of the textile: dense and hard textile creases differently from soft and thin textile. In such cases, integrating vision and FT may be useful. Careful consideration on what visual features are used needs to be taken into account. For example, using flow-based methods [40] or specified wrinkle detectors [41] dealing with various texture properties may be considered.

An additional important aspect to be considered is the ability to assess how textile properties change over time. Certain textiles are made to be more durable, fibers are blended, their use and handling in terms of washing, ironing, folding, will affect how clothing items deteriorate over time. In other words, the information of the fiber content usually available on the label sewn on the clothing item, may be helpful but it is not fully relevant. For example, a T-shirt made out of cotton, may be more elastic and thicker when new, and rather thin and almost non-elastic after many washings. Thus, its handling in terms of washing and ironing will be different, as well as one may decide to keep or reuse a newer one, and recycle a well-used one.

The outlined taxonomy, visualization, CNN classification and measurement interpretability are important tools that can provide more insight into the difficulty of the considered problem. The taxonomy provides a structured approach to study textile materials and has not been previously considered in the area of robotics. We also need an approach that brings the robotics community closer to textile production industry and this is one way of achieving that. We provided several examples of how the generated textile material classes are a viable approach and how these can be studied together with actions such as pulling and twisting.

Initial classification results using deep neural networks show a good potential and we will build on these with a more extensive database of samples, actions and multimodal sensory feedback. More specifically, we will study a richer set of pulling actions, with samples of different sizes also considering standardized textile for the purpose of repeatability, reproducibility and replicability. We believe that this study is an important step toward a more robust and versatile textile handling and manipulation for applications such as various household tasks, assisted dressing and recycling.

VIII. ACKNOWLEDGEMENTS

This work has been supported by the European Research Council, Swedish Research Council and Knut and Alice Wallenberg Foundation.

REFERENCES

- [1] K. Paraschidis, N. Fahantidis, V.-Petridis, Z. Doulgeri, L. Petrou, and G. Hasapis, "A robotic system for handling textile and non rigid flat materials," *Computers in Industry*, vol. 26, no. 3, pp. 303–313, 1995, computer Integrated Manufacturing and Industrial Automation.
- [2] J. Sanchez, J.-A. Corrales, B.-C. Bouzgarrou, and Y. Mezouar, "Robotic manipulation and sensing of deformable objects in domestic and industrial applications: a survey," *The International Journal of Robotics Research*, vol. 37, no. 7, pp. 688–716, 2018.
- [3] A. Doumanoglou, J. Stria, G. Peleka, I. Mariolis, V. Petrik, A. Kargakos, L. Wagner, V. Hlaváč, T. Kim, and S. Malassiotis, "Folding clothes autonomously: A complete pipeline," *IEEE Transactions on Robotics*, vol. 32, no. 7, pp. 688–716, 2018.
- [4] Y. Li, Y. Yue, D. Xu, E. Grinspun, and P. K. Allen, "Folding deformable objects using predictive simulation and trajectory optimization," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 6000–6006.
- [5] M. Lippi, P. Poklukar, M. C. Welle, A. Varava, H. Yin, A. Marino, and D. Kragic, "Latent space roadmap for visual action planning of deformable and rigid object manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020.
- [6] A. Kapusta, W. Yu, T. Bhattacharjee, C. K. Liu, G. Turk, and C. C. Kemp, "Data-driven haptic perception for robot-assisted dressing," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2016, pp. 451–458.
- [7] G. Chance, A. Jevtić, P. Caleb-Solly, and S. Dogramadzi, "A quantitative analysis of dressing dynamics for robotic dressing assistance," *Frontiers in Robotics and AI*, vol. 4, pp. 1–14, 05 2017.
- [8] I. Garcia-Camacho, M. Lippi, M. C. Welle, H. Yin, R. Antonova, A. Varava, J. Borras, C. Torras, A. Marino, G. Alenya *et al.*, "Benchmarking bimanual cloth manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1111–1118, 2020.
- [9] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "Deepfashion: Powering robust clothes recognition and retrieval with rich annotations," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1096–1104.
- [10] J. Huang, R. S. Feris, Q. Chen, and S. Yan, "Cross-domain image retrieval with a dual attribute-aware ranking network," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1062–1070.
- [11] Y. Ge, R. Zhang, X. Wang, X. Tang, and P. Luo, "Deepfashion2: A versatile benchmark for detection, pose estimation, segmentation and re-identification of clothing images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5337–5345.
- [12] T. Ziegler, J. Butepage, M. C. Welle, A. Varava, T. Novkovic, and D. Kragic, "Fashion landmark detection and category classification for robotics," in *2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. IEEE, 2020, pp. 81–88.
- [13] E. Kerr, T. McGinnity, and S. Coleman, "Material recognition using tactile sensing," *Expert Systems with Applications*, vol. 94, pp. 94–111, 2018.
- [14] V. Petrik and V. Kyrki, "Feedback-based fabric strip folding," *arXiv preprint arXiv:1904.01298*, 2019.
- [15] S. Grishanov, "2 - structure and properties of textile materials," in *Handbook of Textile and Industrial Dyeing*, ser. Woodhead Publishing Series in Textiles, M. Clark, Ed. Woodhead Publishing, 2011, vol. 1, pp. 28–63.
- [16] T. C. Jiménez P., "Perception of cloth in assistive robotic manipulation tasks," vol. 19, pp. 409–431, 2020.
- [17] C. Patel, Z. Liao, and G. Pons-Moll, "Tailornet: Predicting clothing in 3d as a function of human pose, shape and garment style," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [18] B. L. Bhatnagar, G. Tiwari, C. Theobalt, and G. Pons-Moll, "Multi-garment net: Learning to dress 3d people from images," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019.
- [19] L. Sun, S. Rogers, G. Aragon-Camarasa, and J. P. Siebert, "Recognising the clothing categories from free-configuration using gaussian-process-based interactive perception," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 2464–2470.
- [20] J. Silvestre-Blanes, J. Berenguer-Sebastiá, R. Pérez-Lloréns, I. Miralles, and J. Moreno, "Garment smoothness appearance evaluation through computer vision," *Textile Research Journal*, vol. 82, no. 3, pp. 299–309, 2012.
- [21] S. Luo, J. Bimbo, R. Dahiya, and H. Liu, "Robotic tactile perception of object properties: A review," *ArXiv*, vol. abs/1711.03810, 2017.
- [22] A. Davis, K. L. Bouman, J. G. Chen, M. Rubinstein, F. Durand, and W. T. Freeman, "Visual vibrometry: Estimating material properties from small motion in video," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.
- [23] M. Strese, L. Brudermueller, J. Kirsch, and E. Steinbach, "Haptic material analysis and classification inspired by human exploratory procedures," *IEEE Transactions on Haptics*, vol. 13, pp. 404–424, 2020.
- [24] A. J. Fishel and E. G. Loeb, "Bayesian exploration for intelligent identification of textures," *Front. Neurobot.*, 2012.
- [25] J. M. Romano and K. J. Kuchenbecker, "Methods for robotic tool-mediated haptic surface recognition," in *2014 IEEE Haptics Symposium (HAPTICS)*, 2014, pp. 49–56.
- [26] M. Kerzel, M. Ali, H. G. Ng, and S. Wermter, "Haptic material classification with a multi-channel neural network," in *2017 International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 439–446.
- [27] Z. Erickson, E. Xing, B. Srirangam, S. Chernova, and C. Kemp, "Multimodal material classification for robots using spectroscopy and high resolution texture imaging," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2020.
- [28] A. A. Khan, M. Khosravi, S. Denei, P. Maiolino, W. Kasprzak, F. Mastrogiovanni, and G. Cannata, "A tactile-based fabric learning and classification architecture," in *2016 IEEE International Conference on Information and Automation for Sustainability (ICIA/S)*, 2016, pp. 1–6.
- [29] W. Yuan, Y. Mo, S. Wang, and E. Adelson, "Active clothing material perception using tactile sensing and deep learning," 05 2018, pp. 1–8.
- [30] J. Sinapov, V. Sukhoy, R. Sahai, and A. Stoytchev, "Vibrotactile recognition and categorization of surfaces by a humanoid robot," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 488–497, 2011.
- [31] F. V. Drigalski, M. Gall, S.-G. Cho, M. Ding, J. Takamatsu, T. Ogasawara, and T. Asfour, "Textile identification using fingertip motion and 3d force sensors in an open-source gripper," *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 424–429, 2017.
- [32] S. Luo, W. Yuan, E. Adelson, A. Cohn, and R. Fuentes, "Vitac: Feature sharing between vision and tactile sensing for cloth texture recognition," *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2722–2727, 2018.
- [33] S. J. Lederman and R. L. Klatzky, "Haptic classification of common objects: Knowledge-driven exploration," *Cognitive Psychology*, vol. 22, no. 4, pp. 421–459, 1990.
- [34] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.
- [35] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [36] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [37] Y. Gao, L. A. Hendricks, K. J. Kuchenbecker, and T. Darrell, "Deep learning for tactile understanding from visual and haptic data," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 536–543.
- [38] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," *International Journal of Computer Vision*, vol. 128, no. 2, p. 336–359, Oct 2019.
- [39] I. Mitsioni, J. Mänttari, Y. Karayiannidis, J. Folkesson, and D. Kragic, "Interpretability in contact-rich manipulation via kinodynamic images," 2021.
- [40] V. Scholz and M. A. Magnor, "Cloth motion from optical flow," in *VMV*, vol. 4, 2004, pp. 117–124.
- [41] M. Takamatsu, E. Iwata, T. Furukawa, and M. Hasegawa, "A study on garment wrinkle detection through a 3d camera scanning with normal map and hough transform," in *International Workshop on Advanced Image Technology (IWAIT) 2019*, vol. 11049. International Society for Optics and Photonics, 2019, p. 110492U.