Tactile Classification of Object Materials for Virtual Reality based Robot Teleoperation

Bukeikhan Omarali, Francesca Palermo, Kaspar Althoefer, Maurizio Valle, Ildar Farkhatdinov

Tactile Classification of Object Materials for Virtual Reality based Robot Teleoperation

Bukeikhan Omarali[1,2], Francesca Palermo[1,3], Kaspar Althoefer[1], Maurizio Valle[2], Ildar Farkhatdinov[1,4]

***Abstract*— This work presents a method for tactile classification of materials for virtual reality (VR) based robot teleoperation. In our system, a human-operator uses a remotely controlled robot-manipulator with an optical fibre-based tactile and proximity sensor to scan surfaces of objects in a remote environment. Tactile and proximity data and the robot's end-effector state feedback are used for the classification of objects' materials which are then visualized in the VR reconstruction of the remote environment for each object. Machine learning techniques such as random forest, convolutional neural and multimodal convolutional neural networks were used for material classification. The proposed system and methods were tested with five different materials and classification accuracy of 90% and more was achieved. The results of material classification were successfully exploited for visualising the remote scene in the VR interface to provide more information to the human-operator.**

I. Introduction

The last few years have seen rapid adoption of Virtual Reality (VR) technologies in robot teleoperation. The key advantage of VR-based teleoperation interfaces is the three-dimensional (3D) reconstruction of a remotely operated robot and its' environment that improves a human-operator's (further referred to as "operator") sense of presence [1], [2], [3], [4]. However, this also leads to operators mainly relying on visual feedback. An important aspect of teleoperation is providing the operator with reliable information about the objects' materials in a remote environment. For example, the operator may be required to collect all metallic objects in the remote environment which might be difficult if the operator can only rely on the visual reconstruction of the remote environment, which is typically implemented with a point-cloud geometric visualisation based on the data acquired from RGB-D cameras. Often such cameras acquire noisy data that reduce the accuracy of VR reconstruction.

We suggest using tactile sensing and classification to identify materials of objects in the remote environment and visually communicate those materials to the operator in the VR interface. The majority of objects' materials classification techniques are based on machine learning applied to images of objects [5], [6]. A convolutional neural network (CNN) architecture was used to classify materials from patches extracted from photos of objects, followed by localising and segmenting the materials in the original images [7]. In situations with reduced lighting when computer vision methods can fail (for example teleoperation in extreme and hazardous environments [8]), tactile exploration can be used for material recognition [9], [10]. A combination of proximity, tactile and force sensing can provide important information on the explored material such as texture, stiffness and shape [11]. Compliance properties of different objects via supervised classifiers using a hybrid force and proximity finger-shaped sensor was investigated in [12]. A fibre optic-based sensor has been designed and used in [13], [14], [15] to recognise and classify fractures on surfaces implementing a random forest classifier.

In contrast to machine learning algorithms which need engineered features extracted from the data, deep learning models achieve higher performance when trained with high dimensional input, such as audio in speech recognition [16] and figures in image recognition and detection [17]. Various machine learning classifiers (Gaussian, K-nearest

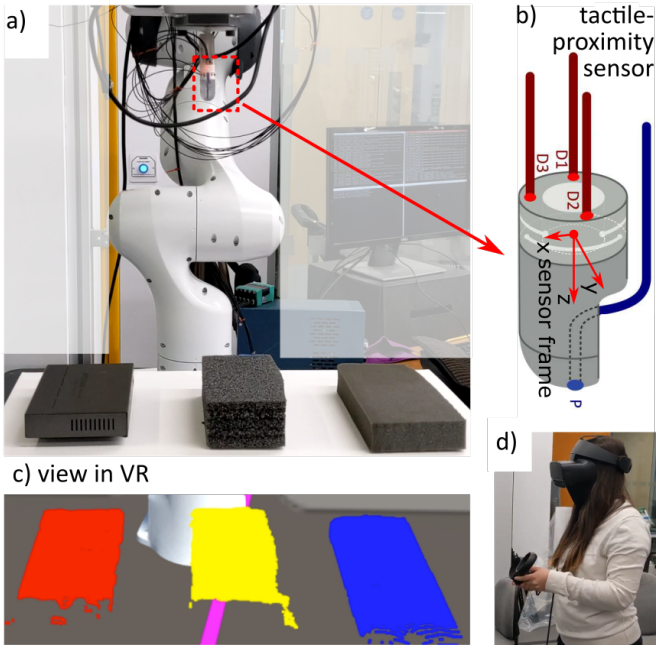

Fig. 1. Overview of the system. **a)** a remotely operated robotic-manipulator with a tactile sensor **(b)** is commanded by a human-operator **(d)** to explore the surfaces of different objects. Material classification results are used to visualise the objects differently in the VR reconstruction **(c).**

This work is supported by the QMUL-Genoa Univ. PhD Program and by the UK EPSRC grants NCNR EP/R02572X/1 and EP/T027746/1.

[1]School of Electronic Engineering and Computer Science, Queen Mary University of London, UK

[2]Cosmic Lab, Department of Electrical, Electronics and Telecommunication Engineering and Naval Architecture, University of Genoa, Italy

[3]Department of Brain Science, Faculty of Medicine, Imperial College of Science, Technology and Medicine, London, UK.

[4]Department of Bioengineering, Imperial College of Science, Technology and Medicine, London, UK.

{b.omarali,i.farkhatdinov}@qmul.ac.uk

neighbours, support vector machine) were compared in [18] against a CNN model to classify explored surfaces via tactile skin sensor attached to an iCub humanoid robot. It was found that the deep learning model for tactile material classification provides more robust performance in comparison to feature learning-based methods. The authors of [19] demonstrated that implementing a multi-modal approach based on CNN with both visual and physical interaction signals achieved more accurate results than vision only as well as superior classification compared to hand-designed features methods. Authors of [20] proposed an algorithm for recognising the object touched via human interactions on an electronic skin. The 3D tactile data generated by the skin were converted into 2D images and used as input for a CNN which outperformed classical tactile data classification algorithms.

We propose a method for tactile classification of materials for VR-based robot teleoperation. In our system, the operator remotely controls a robotic arm with an optical fibre-based tactile and proximity sensor to scan surfaces of objects in a remote environment. Tactile and proximity data as well as the robot's end-effector state feedback are used for the classification of objects' materials. Classification results are then used to visualize objects' materials in VR. Machine learning techniques such as random forest, convolutional neural and multi-modal convolutional neural networks were used for material classification.

To the authors' knowledge, this work is the first attempt to demonstrate the integration of material classification with tactile sensing in VR-based robot teleoperation. Our work proposes a novel material classification technique based on data collected with a fibre optics-based tactile sensor, human-supervisory control and visualisation of objects' materials in VR.

II. System overview

Key components of the system are shown in Fig. 1: Franka Emika's Panda robot (a), a fibre optics tactile sensor (b), an Intel Realsense 435i RGBD camera, Oculus Rift S headset and Oculus Touch handheld controllers (d). Two desktop computers were connected in a wired network: the operator's computer running the VR interface (built with Unity 3D), and the robot's computer running ROS. The robot was controlled by the operator using a VR-based robot teleoperation interface described in our previous work [2]. The operator sets desired motion path by manually placing waypoints in a VR reconstruction of the remote environment. These waypoints are then used to plan the robot's motion, which is previewed and accepted or rejected by the operator. The remote environment was visualized in VR using a point cloud from an Intel Realsense 435i RGB-D camera mounted on the robot's end-effector.

We used an integrated force and proximity finger-shaped sensor described in [12], [21] attached to the robot's end-effector. The sensor consists of 3D printed rigid and soft components that allow the finger to bend during interaction with the environment, as shown in Fig. 1(b). The sensor has three pairs of optical fibre cables (D1, D2, D3) that use reflected lights' intensity to measure the deformation of the finger. The fourth pair of optical fibre cables (P) is used to measure the distance between the tip of the finger and nearby objects. Each pair of the sensor's fibre optic cables was attached to a Keyence FS-N11MN light-to-voltage transducers that communicate with the robot's control computer at 400 Hz.

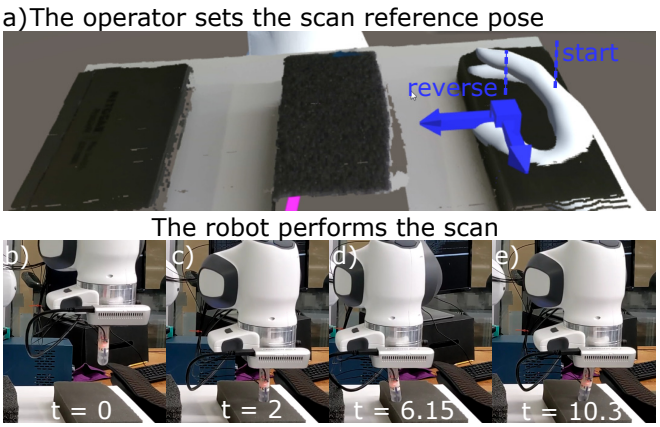

Fig. 2. a) The operator manually sets the scan reference pose: the pose of the 3D blue axis mesh is used as a reference of scan location with the second reference at 56.5mm offset. b) the robot approaches to the scan start position at t = 0; c) the slide begins at t = 2; d) the slide direction reverses at t = 6.15; e) the slide is finished at t = 10.3.

III. Methods

Tactile Data Collection. Supervised teleoperation was used for tactile exploration of five objects made of metal, paper, silicon, hard styrofoam, soft foam. All objects were flat and placed horizontally in the robot's workspace. Fig. 1(a) shows sample objects made of metal, styrofoam and soft foam. An expert operator was asked to set the reference tactile scan pose, which determined the location and the orientation of initial touch between the end-effector mounted tactile sensor and the object as well as the subsequent sliding (scanning motion). Setting the initial conditions for tactile exploration manually allows efficient collection of highly variable data representative of real-world use cases suitable for the classifiers' training and validation.

Objects were scanned using horizontal linear sliding motion that was generated from a reference pose set by the operator. The operator set the reference poses by manually placing an axis 3D mesh object in the VR teleoperation interface, see Fig. 2(a). The reference pose was then used to generate the scan start and reverse positions (both generated at the same height as the reference pose) for the tactile sensor. The slide was always performed along the sensor's y-axis projection on the object's horizontal plane, as it results in more distinct D1 and D2 sensor outputs.

The orientation of the tactile sensor with respect to the scanned object was determined by the orientation of the 3D axis mesh set by the operator in VR. The robot has always approached to and moved away from the scanned object along the sensor's z-axis. Different orientations of the sensor and corresponding approaches have resulted in

different sensor deformation behaviours. The contact force exerted by the robot/finger on the scanned object (and by extent the penetration depth on softer materials) depended on the reference pose height set by the operator.

After the operator set the reference pose, the robot planned and executed a trajectory necessary to approach and get in contact with the scanned object, see Fig. 2(b,c), performed a slide Fig. 2(d), reversed the slide direction, slid back Fig. 2(e) and retracted. The length of the sliding motion was set to 56.5 mm. The motion plans were generated using Moveit and the planned trajectory's scan duration was automatically re-scaled to 8.3 s.

We have recorded raw outputs of the tactile sensor for object classification. We recorded the robot's end-effector's average position error during the scan as the mean of differences between the end-effector's desired and actual positions. The robot's position error can be used as an indicator of objects' softness. For example, the tactile sensor can deform and penetrate into soft foam resulting in a small position error, which is not the case for metal. We also recorded the reference pose orientation. The tactile sensor scan output varied depending on the sensor's orientation with respect to the scanned objects, hence desired orientation was used as a feature in classifiers.

The operator set the reference pose and the robot performed the scan three separate times. Then the operator changed the reference pose by either changing the reference orientation, or reference position on the scanned object's surface, or the reference scan depth. We have recorded 150 samples (scans) per class.

Data pre-processing for classification. A sample was recorded from 2 seconds before the slide start (to include the initial contact between the sensor and the scanned object) to 2 seconds after the slide end (to include the retraction of the sensor from the scanned object). We used spectrograms of raw tactile data for objects' material classification. Spectrograms were generated with 52 time segments and 50 spectral bands (i.e. each channel's spectrogram is a 52×50 matrix). We compared three classifiers: random forest, convolutional neural network (CNN) and multi-modal convolutional neural network (M-CNN). Random forest was made using scikit-learn library, CNNs were made using Keras with Tensorflow. Spectrograms were standardized using the maximum amplitude present in the dataset. The average end-effector position error was standardized using the maximum absolute value present in the dataset. The reference orientation was represented as a unit quaternion. We generated additional 50 synthetic samples per class by randomly copying existing samples and adding corresponding $\pm 5\%$ standard deviation to each of the spectrograms' time segment and frequency band (to each cell of the spectrogram's 52×50 matrix).

The dataset was split into 60%-20%-20% training, validation and test sets, respectively. We distributed samples into sets based on the tactile sensor's orientation with respect to the scanned objects. This was done to allow the classifiers to train with the data collected at different tactile exploration conditions. Fig. 3 demonstrates sensor orientations in radial

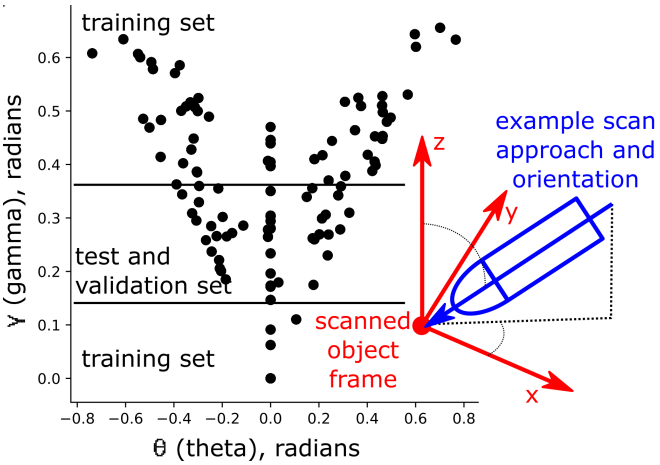

Fig. 3. Distribution of different initial tactile scanning angular orientations of the sensor. γ is the angle between the objects' surface normal and the tactile sensor's z axis. θ is the angle between the projection of the tactile sensor's z-axis to the scanned object's xy-plane and the object's x-axis (object's x and y-axis are collinear with the robot's with an origin and desired scan position).

angles present in the dataset (note that many samples overlap), where γ is the angle between scanned object's z-axis (scanned surface normal) and the tactile sensor's z-axis and θ is the angle from scanned object's x-axis to y-axis (both are collinear to robot's base frame's x and y-axis). The training set contained samples in upper and lower 30% γ angles, validation and testing sets were randomly chosen from the rest. Hence classifiers were tested and validated on tactile exploration conditions that were not present in the training set, ensuring classifiers' robustness.

Classification with Random forest. Random Forest [22] is a machine learning algorithm used for classification and regression built on an ensemble of multiple learning trees. Insensitive to over-fitting, it can produce reasonable predictions with little tuning and provides an effective way of handling missing data. Random Forest has been vastly used in remote sensing [23], [15], [13]. We implemented a Random Forest classifier with 1000 estimators. The number of estimators was determined by a grid search. The classifier was given concatenated spectrograms, average end-effector position error and reference orientation as inputs.

Classification with Convolutional Neural Network. CNNs are commonly used in image recognition due to their ability to learn cross-correlations between multiple channels (RGB in case of images) and shift invariancy. They can also be used to learn patterns between multiple sensor signals as demonstrated in [24], [25]. We suggest a similar approach to classify objects' materials - using tactile sensor's multiple sensing channels. The CNN classifier only takes spectrograms as an input and its' model architecture is shown in Fig. 4. There was a dropout with 30% probability between fully connected layers. The output of the last convolutional layer was batch normalized. The model was trained with an early stopping triggered by no improvement on validation loss. We used the Adam optimizer with an exponentially

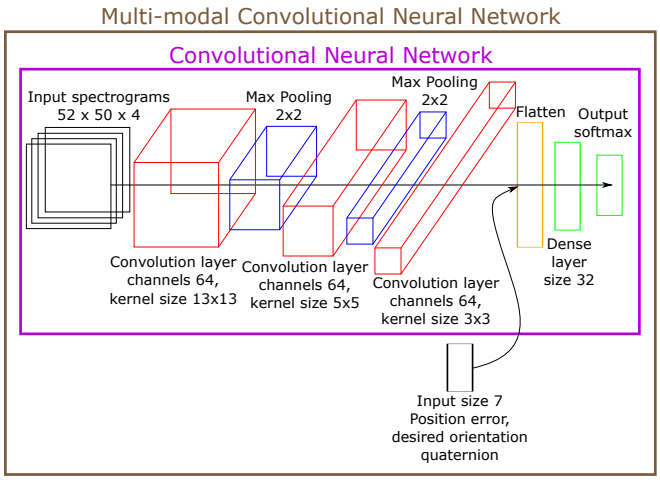

Fig. 4. Convolutional neural network (CNN) and multi-modal CNN classifiers. Classifiers are similar except for the additional input layer that contains the robot's end-effector position error and scan orientation quaternion.

decaying learning rate. The model's and training hyper-parameters were determined using grid search.

Classification with Multi-modal Convolutional neural network. We added an extra input to the CNN described above that included the robot's end-effector's average position error and the reference orientation. The model architecture is detailed in Fig. 4. The extra inputs were concatenated with the flattened output of the last convolution layer. The multi-modal CNN retained the 30% probability dropout between fully connected layers and batch normalization after the last convolutional layer. The M-CNN was trained with settings similar to the CNN.

Visualization of object's materials in VR. We propose two visualization methods that communicate scanned objects' predicted classes to the operator in the VR-based robot teleoperation interface. Both methods rely on naive Octomap-based segmentation introduced in our previous work [2], which allows the operator to segment pointclouds into separate objects. Octomap [26] is a probabilistic mapping method that splits the space (for example the remote environment) into occupied, vacant and unknown cubes (voxels) using pointcloud as an input. Our naive segmentation detects when such occupied cube has been pointed and clicked by the operator and segments all points (of the pointcloud) contained within it as well as within other connected cubes. All segmented points are then stored locally as static segmented pointclouds.

1) Objects' predicted classes as color-coded Octomaps: As the operator sets the reference pose using the 3D axis mesh, the axis mesh checks for collisions with the Octomap using Unity's collider system. These collisions do not have any physical meaning and are simply used to detect which part of the Octomap the scan will be performed on. Once a collision is detected the corresponding Octomap cube and connected cubes are segmented and copied from the "live" Octomap and stored locally. Once the classification is finished the predicted class is used to colour the segmented Octomap according to the color-code given to the operator in the VR interface.

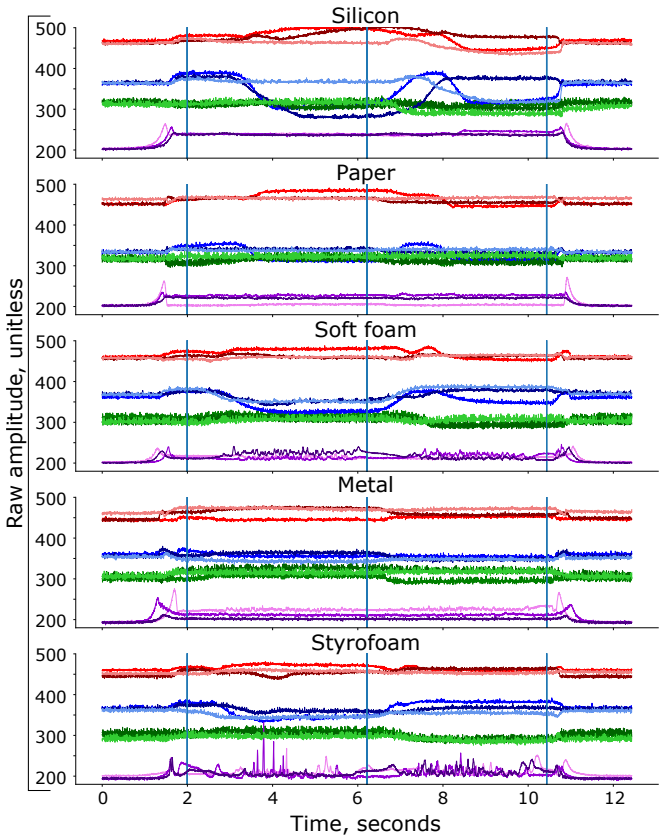

Fig. 5. The raw output of the tactile sensor for different materials. Reds: D1, blues: D2, greens: D3, purples: P. Leftmost vertical blue line represents the start of the sliding motion; the middle vertical blue line represents the reverse in slide direction; the rightmost vertical blue line represents the end of the slide.

2) Objects' predicted classes as color-coded pointclouds: This visualization method is similar to the one above except instead of segmenting and storing Octomaps, it segments and stores corresponding pointclouds. Once the classification is finished the predicted class is used to color the segmented pointcloud according to the labels' color-codes.

IV. Results

Tactile data. Fig. 5 presents three samples of raw sensor output for each material recorded at different sensor orientations and heights. There is a noticeable variance between different samples of each material. The leftmost vertical lines indicate the start of the slide, the middle one indicates the slide direction reversal, and the last vertical line indicates the stop of the slide.

For all materials the proximity sensor (P) spikes up on the sensor approach and retraction (before the slide begins and after it finishes respectively). The proximity sensor reading depends on the distance to the object. The proximity sensor only detects in a limited range and once the object is too close (or the sensor touches the object) the proximity sensor value drops to the baseline. The amplitude of the spike is

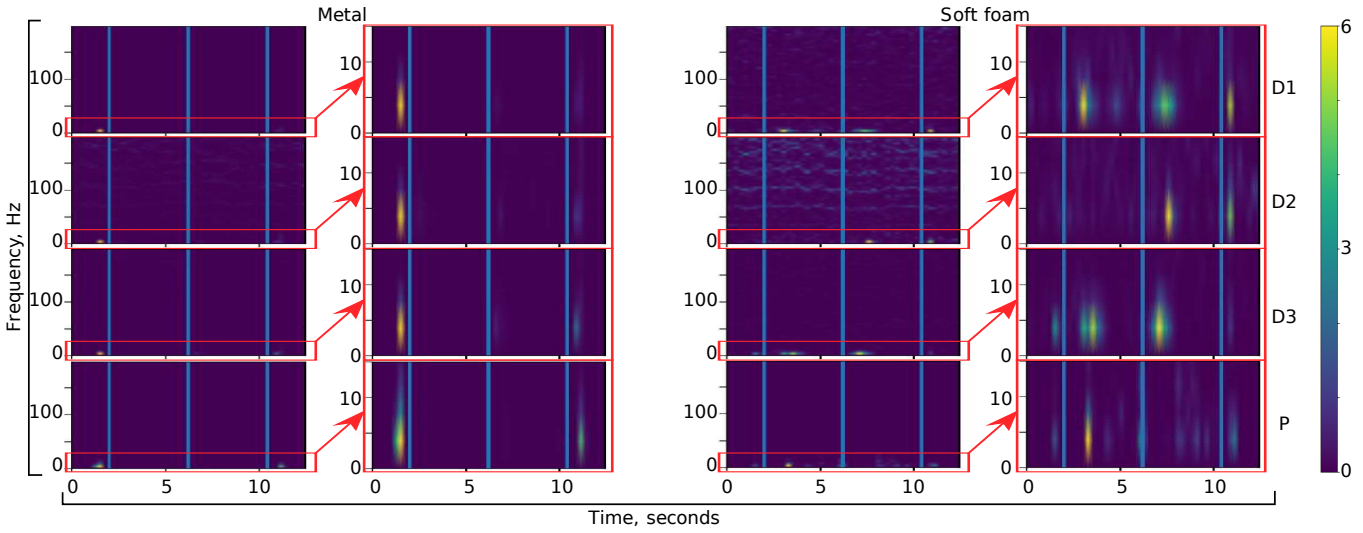

Fig. 6. Sample spectrograms (not normalized) for metal and soft foam. There are large spikes in 0-15Hz frequency range for both materials (arrows indicate zoom-in regions) and smaller spikes in high frequency for softer/rougher soft foam.

also determined by the object's surface roughness, color and reflectivity. In the case of rough and porous materials (soft foam, medium styrofoam), the proximity sensor generates a noisy output even during the touch.

There was a delay between the beginning of the slide and deformation sensors responses. The delay was the time necessary for the sensing finger to deform, (the sensor deformation can be seen in Fig 2d,e). Similarly, responses of the deformation sensors were delayed after slide direction reversal and sensor retraction.

Spectral analysis. Example spectrograms are shown in Fig. 6. Spectrograms were distinct per material, which was noticeable visually. Spectrograms differed most in the 0-15Hz frequency band, where there were large amplitude spikes during the sensor's initial contact with an object, during deformation and sensor retraction. There was also a noticeable high-frequency signal present in rougher/softer materials as well, (see D2 for soft foam).

End-effector position error. The average end-effector position errors occurred along the scanned surface's normal as the robot failed to push in as deep as the operator intended to. Table I presents means and standard deviations of the robot's end-effector's average position error along the scanned surface's normal per material. As expected the harder materials had larger errors: the sensor deformed and pushed in deep into the soft foam but could not do the same with metal. However, T-test showed no statistical significance.

Classification metrics. Accuracy, precision, recall, and f1-score classification metrics are used to validate and compare the used classification models. The results for the analysis with the implemented classifiers are presented in Table II. Random Forest, which is robust to outliers and requires little parameter optimisation, achieves the best results. The MCNN, with the extra inputs of the robot's end-effector's average position error and reference orientation, achieves

TABLE I

END-EFFECTOR AVERAGE POSITION ERROR PER SAMPLE

	Mean (mm)	Standard deviation (mm)
Soft foam	-0.166	0.179
Styrofoam	-0.124	0.1477
Paperbook	-0.264	0.415
Silicon	-0.0817	0.1484
Metal	-0.232	0.220

TABLE II

COMPARISON OF RESULTS FOR THE EXPERIMENTS

Model	**Accuracy**	**Precision**	**Recall**	**F1**
Random Forest	94.5	95.0	94.5	94.4
CNN	84.5	85.1	84.5	84.1
M-CNN	92.0	92.3	92.0	92.1

results comparable to the Random Forest classifier while the baseline CNN generates the worst outcome. Thus, even with the small size of the dataset of the network, the extra inputs make the M-CNN model more robust than the CNN and able to generalise better. Confusion matrices for CNN, M-CNN, and Random forest classifiers are shown in Fig. 7. One of the most challenging materials to classify for all three models is the paperstack, which is frequently confused with the silicon class. This may be due to a similar pattern in the frequency domain.

Predicted class visualization in VR. The Fig. 8 shows the operator's VR view of classification results as colored Octomaps and colored pointclouds as well as the GUI description of the color-coding. The operator can toggle between visualization modes and "live" pointcloud using a button press on a handheld controller. In the current implementation, both methods use static segmented clones of pointcloud and Octomap respectively, which means that updates to "live" pointclouds or Octomap of objects (for example if objects move or RGB-D camera changes position

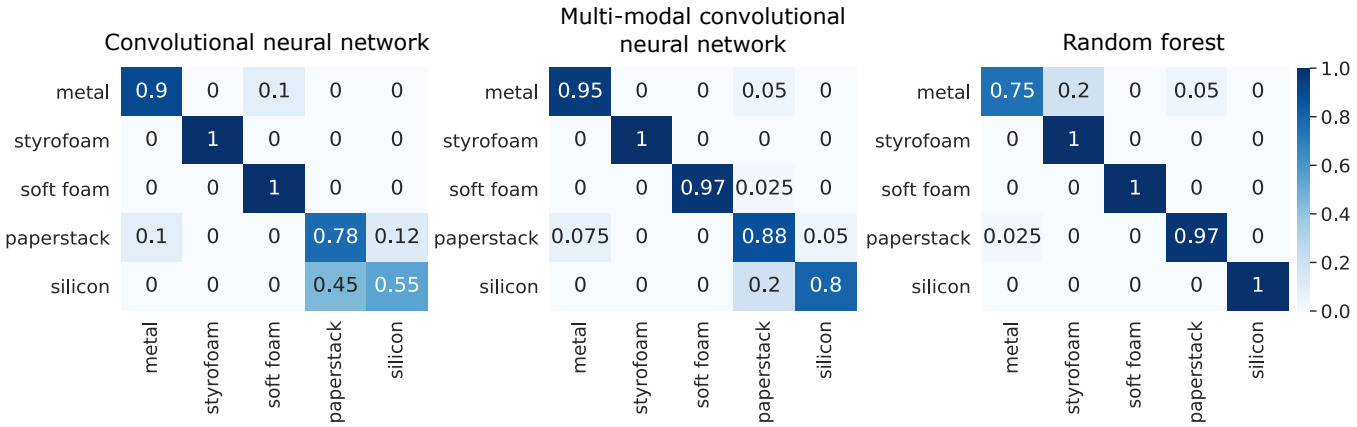

Fig. 7. Confusion matrices of Random Forest, Convolutional Neural Network, Multi-modal Convolutional Neural Network classifiers.

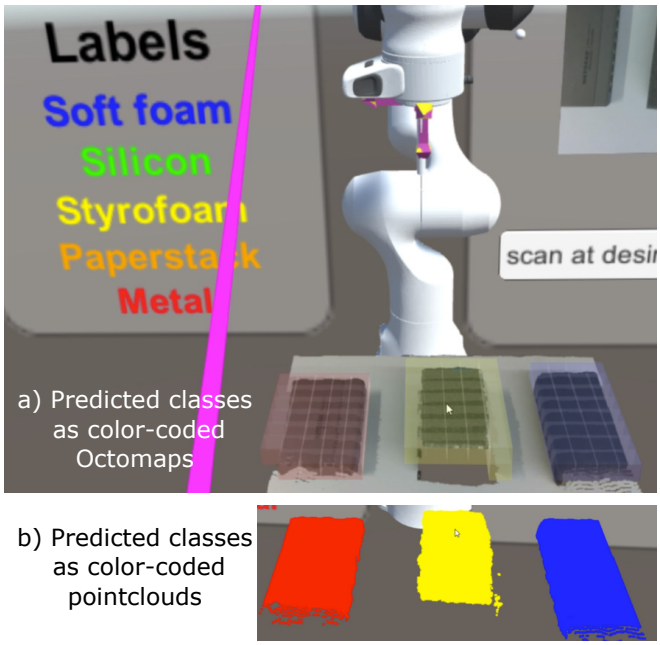

Fig. 8. Classification results visualized in Virtual Reality interface: a) classification results as color-coded Octomaps overlaid on corresponding objects; b) classification results as static color-coded segmented pointcloud clones.

with respect to objects) would not be reflected in classified clones. This presents a technical challenge that requires object tracking as well as continuous segmentation.

V. Discussion and Conclusion

Summary. We presented a method for tactile classification of materials for VR-based robot teleoperation. We have collected a dataset using a teleoperated robot. The dataset contains five objects with a wide range of different sensor-to-object relative orientations and interaction forces. We have compared three classifiers using spectrograms of raw tactile data and robot's end-effector feedback: random forest, convolutional neural network and multi-modal convolutional neural network. We also introduced two visualisation methods to display the classification results to a human-operator in the VR using color-coded pointclouds and corresponding bounding Octomaps.

The results demonstrated that the random forest classifier had the highest accuracy at 94.5% followed by the multi-modal convolutional neural network with 92%. The results of material classification were successfully employed for visualising the remote scene in the VR interface to provide more information to a human-operator.

Limitations. The proposed method currently works only with flat surfaces. In the case of non-flat surfaces, a similar approach can be used where the operator sets multiple waypoints that can be used to generate a non-flat trajectory.

The dataset was collected using constant scan length and duration. It is possible that the random forest classifier trained on this dataset may not be able to generalize well to scans performed at higher or lower speeds. The dataset was collected using teleoperation to be representative of in-field operations. This however prevented us from collecting a larger amount of samples that would improve classifier training. Although the dataset contains samples collected at various positions, orientations and depths, it was collected by a single expert operator. It is a small chance that classifier could be overfitting towards this operator's specific mannerisms of setting the desired scan position.

Future work The future work would benefit from larger sample sizes collected by multiple expert operators as well as from an extended object set. The latter would also allow moving from hard-coded material based labels to more general roughness/hardness estimation.

Due to the imprecise nature of pointclouds the operator may put the reference pose too deep into the object potentially damaging the sensor or not deep enough, resulting in poor classification. In future work, we will make better use of the proximity sensor and haptic feedback to place the sensor more accurately at the object's surface.

Proposed methods require a full scan to be completed before the object's class is determined. A potentially interesting follow-up topic is real-time objects' materials classification using Long Short-term Memory networks.

REFERENCES

[1] D. Whitney, E. Rosen, D. Ullman, E. Phillips, and S. Tellex, "ROS Reality: A Virtual Reality Framework Using Consumer-Grade Hardware for ROS-Enabled Robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1–9.

[2] B. Omarali, B. Denoun, K. Althoefer, L. Jamone, M. Valle, and I. Farkhatdinov, "Virtual reality based telerobotics framework with depth cameras," in *29th IEEE International Conference on Robot and Human Interactive Communication*, 2020, pp. 1217–1222.

[3] S. Kohn, A. Blank, D. Puljiz, L. Zenkel, O. Bieber, B. Hein, and J. Franke, "Towards a Real-Time Environment Reconstruction for VR-Based Teleoperation Through Model Segmentation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2018.

[4] B. Omarali, K. Althoefer, F. Mastrogiovanni, M. Valle, and I. Farkhatdinov, "Workspace scaling and rate mode control for virtual reality based robot teleoperation," in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2021, pp. 607–612.

[5] M. Cimpoi, S. Maji, and A. Vedaldi, "Deep filter banks for texture recognition and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3828–3836.

[6] J. Xue, H. Zhang, K. Dana, and K. Nishino, "Differential angular imaging for material recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 764–773.

[7] S. Bell, P. Upchurch, N. Snavely, and K. Bala, "Material recognition in the wild with the materials in context database," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3479–3487.

[8] I. Vitanov, I. Farkhatdinov, B. Denoun, F. Palermo, A. Otaran, J. Brown, B. Omarali, T. Abrar, M. Hansard, C. Oh *et al.*, "A suite of robotic solutions for nuclear waste decommissioning," *Robotics*, vol. 10, no. 4, p. 112, 2021.

[9] F. De Boissieu, C. Godin, B. Guilhamat, D. David, C. Serviere, and D. Baudois, "Tactile texture recognition with a 3-axial force mems integrated artificial finger." in *Robotics: Science and Systems*. Seattle, WA, 2009, pp. 49–56.

[10] A. Drimus, M. B. Petersen, and A. Bilberg, "Object texture recognition by dynamic tactile sensing using active exploration," in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2012, pp. 277–283.

[11] T. P. Tomo, S. Somlor, A. Schmitz, L. Jamone, W. Huang, H. Kristanto, and S. Sugano, "Design and characterization of a three-axis hall effect-based soft skin sensor," *Sensors*, vol. 16, no. 4, p. 491, 2016.

[12] J. Konstantinova, G. Cotugno, A. Stilli, Y. Noh, and K. Althoefer, "Object classification using hybrid fiber optical force/proximity sensor," in *2017 IEEE SENSORS*. Glasgow, UK: IEEE, 2017, pp. 1–3.

[13] F. Palermo, J. Konstantinova, K. Althoefer, S. Poslad, and I. Farkhatdinov, "Implementing tactile and proximity sensing for crack detection," in *IEEE International Conference on Robotics and Automation*, 2020.

[14] F. Palermo, L. Rincon-Ardila, C. Oh, K. Althoefer, S. Poslad, G. Venture, and I. Farkhatdinov, "Multi-modal robotic visual-tactile localisation and detection of surface cracks," in *2021 IEEE 17th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2021, pp. 1806–1811.

[15] F. Palermo, J. Konstantinova, K. Althoefer, S. Poslad, and I. Farkhatdinov, "Automatic fracture characterization using tactile and proximity optical sensing," *Frontiers in Robotics and AI*, vol. 7, 2020.

[16] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath *et al.*, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal processing magazine*, vol. 29, no. 6, pp. 82–97, 2012.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[18] S. S. Baishya and B. Bäuml, "Robust material classification with a tactile skin using deep learning," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 8–15.

[19] 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)*. IEEE, 2016, pp. 536–543.

[20] M. Alameh, A. Ibrahim, M. Valle, and G. Moser, "Dcnn for tactile sensory data classification based on transfer learning," in *2019 15th Conference on Ph.D Research in Microelectronics and Electronics (PRIME)*, 2019, pp. 237–240.

[21] J. Konstantinova, A. Stilli, A. Faragasso, and K. Althoefer, "Fingertip proximity sensor with realtime visual-based calibration," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Daejeon, South Korea: IEEE, 2016, pp. 170–175.

[22] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[23] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31, 2016.

[24] S. Raghu, N. Sriraam, Y. Temel, S. V. Rao, and P. L. Kubben, "Eeg based multi-class seizure type classification using convolutional neural network and transfer learning," *Neural Networks*, vol. 124, pp. 202–212, 2020.

[25] Y.-H. Kwon, S.-B. Shin, and S.-D. Kim, "Electroencephalography based fusion two-dimensional (2d)-convolution neural networks (cnn) model for emotion recognition system," *Sensors*, vol. 18, no. 5, 2018.

[26] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, pp. 189–206, 2013.