This is a preprint. This article has been accepted for publication but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ICSENS.2017.8234203, IEEE SENSORS

Human and Object Recognition with a High-Resolution Tactile Sensor

Juan M. Gandarias, Jesús M. Gómez-de-Gabriel, Alfonso García-Cerezo
System Engineering and Automation Department
University of Málaga
Málaga, Spain
jmgandarias@uma.es

***Abstract*—This paper describes the use of two artificial intelligence methods for object recognition via pressure images from a high-resolution tactile sensor. Both methods follow the same procedure of feature extraction and posterior classification based on a supervised *Supported Vector Machine* (SVM). The two approaches differ on how features are extracted: while the first one uses the *Speeded-Up Robust Features* (SURF) descriptor, the other one employs a pre-trained *Deep Convolutional Neural Network* (DCNN). Besides, this work shows its application to object recognition for rescue robotics, by distinguishing between different body parts and inert objects. The performance analysis of the proposed methods is carried out with an experiment with 5-class non-human and 3-class human classification, providing a comparison in terms of accuracy and computational load. Finally, it is discussed how feature-extraction based on SURF can be obtained up to five times faster compared to DCNN. On the other hand, the accuracy achieved using DCNN-based feature extraction can be 11.67% superior to SURF.**

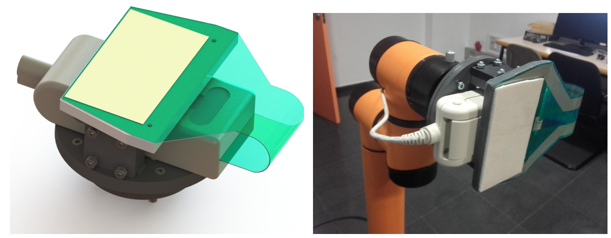
***Kewywords*—Tactile sensors, Object recognition, Rescue robotics, Machine learning**

I. INTRODUCTION

Touch sense is critical for humans. We use the sense of touch to perform complexes tasks such as recognising objects. To perform this task, human beings need two abilities: first, being able to extract information through touch, and second, having cognitive capabilities to process this information [1]. As current trends in robotics are focusing on providing intelligence to robots and making them more similar to humans, tactile sensing in field robotics, is a key problem [2]. The resurgence of artificial intelligence (AI) methods is a great help for interpreting the information acquired. Recent applications propose the use of a tactile sensor to extract information from the object touched and a learning process that use this information to distinguish familiar objects among the collected data [3].

In the literature, a broad variety of works related to object classification can be found, although it is complex to provide a fair comparison between the different approaches, given the differences on the applied hardware [4]. The use of Deep Learning with dropout to reduce overfitting is presented in [5]. This work also describes the benefits of including both kinesthetic and tactile information to object shape recognition, and raises the differences between using planar or curved tactile sensors.

Most of the studies related to object classification are based on the same two steps: feature extraction from pressure images and obtaining a classifier based on those features [6]. An existing solution uses a variant of the *Scale Invariant Feature Transform* (SIFT) descriptor as a feature extractor and a supervised *k-Nearest Neighbour* (kNN) algorithm to get the classifier [7]. A recent work from the same authors proposes a novel algorithm which synthesizes both kinesthetic of 3D positions of the contact and tactile information forming a 4D point cloud of the object [8]. All these works generally present methods for object recognition. However, these methods have not been used in real applications, except those works that use tactile information for grasping objects [9]–[11].



(a) (b)

Fig. 1: The end-tactile-effector (a) and the implementation on the robotic manipulator *AUBO Our-i5* (b).

The use of tactile information for object recognition is essential while searching for victims in first-response and disaster scenarios. In such situations, teleoperation and haptic perception is crucial due to the complexity of the operations [12] and the lack of visual perception in low-light scenarios or in presence of smoke or dust [13].

This paper present two contributions. In one hand, we describe two methods for recognising objects through pressure images. The first method uses the *Speeded-Up Robust Features* (SURF) [14] as a feature extractor, whilst the second method uses the *Deep Convolutional Neural Network (AlexNet)* [15]. Then, both methods include a *Supported Vector Machine* (SVM) [16] to get a classifier. Furthermore we compare both methods results in terms of accuracy and computation time. On the other hand, we propose a real application of tactile sensors and object recognition to the field of rescue robotics. An experiment for testing the performance of the methods is carried out under controller conditions. Fig. 1 presents the used hardware and its accession to the robot. The test demonstrates the capability of the system to recognize not only inert objects, but certain humans parts of the body, with the aim to identify potential victims in disaster scenarios.

II. METHODS

In this paper, the use of two methods to classify pressure images is presented. Both methods implement the classification structure in two steps: extract features and train a classifier.

The first method is based on SURF descriptor. Its workflow is illustrated in Fig. 2. This method needs to include an intermediate step between the features extraction and the supervised learning. As in [17], a *k-means* unsupervised algorithm is implemented to cluster features into a dictionary, generating a framework of *Bag of Words* (BoW). Then, a supervised *SVM* is trained to generate a classifier.

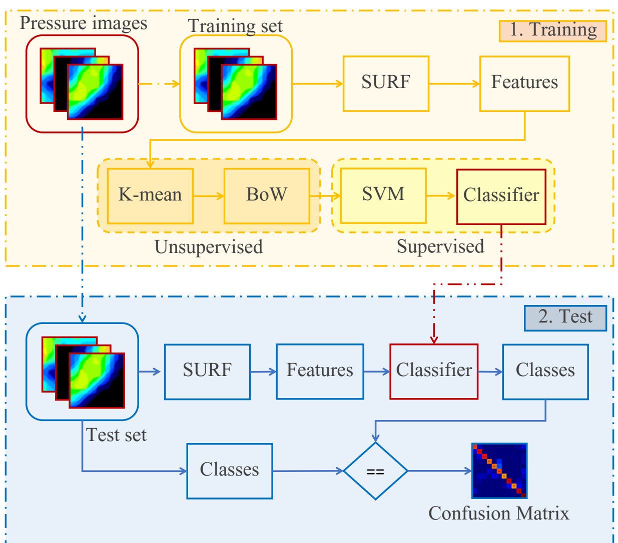


Fig. 2: Method 1: Features extraction with *SURF*, clustering with *k-means* and *BoW*, and classification with *SVM*.

The second method implements a DCNN to extract features. This procedure, previously presented in [18], consists of using a pre-trained network to classify conventional images taking with a camera. Activations of the last layer before the classification are used to describe features. After that, a supervised *SVM* is trained to get a classifier. The implementation scheme can be seen in figure 3.

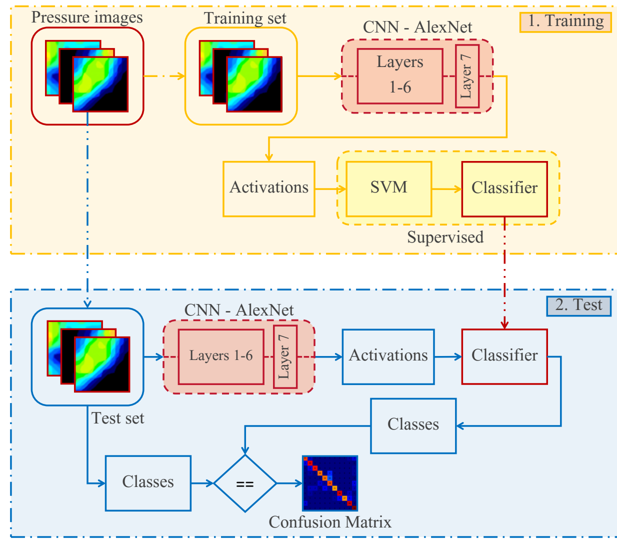


Fig. 3: Method 2: Features extraction with *AlexNet* and classification with *SVM*.

III. EXPERIMENTS AND RESULTS

A. Experimental Setup

A high-resolution tactile sensor has been attached to the 6 DOF robotic arm *AUBO Our-i5*. The *Tekscan* pressure mapping sensor 6077 is conformed by 1400 resistive *sensels* of pressure distributed on 28 rows and 50 columns with a size of 53.3 mm x 95.3 mm. The sensor is covered by a silicone rubber as a contact interface, protecting the device while conducting external forces.

Unlike related works, a single touch is used to classify new images. However, considering that the area of the sensor may be smaller than the surface of the objects, the most representative portion of the objects has to be touched. A total of 400 pressure images have been used to feed each method. These images are divided into eight classes labelled as: *Finger*, *Hand*, *Arm*, *Pen*, *Scissors*, *Pliers*, *Sticky Tape*, and *Allen Key*. The same set of images has been employed to evaluate both methods. The training set is composed by 160 images, 20 images for each label, whilst the test set is composed by 240, 30 images for each label.

B. Results

Figures 4 and 5 show the confusion matrix obtained by applying the methods. The classification accuracy achieved with the SURF-based is 80% whilst with the DCNN-based is 91.67%, that is, DCNN-based presents an improvement of 11.67% of accuracy with respect to the SURF-based. However, the computation time to classify a new image is around 0.01s in SURF-based in contrast with the 0.7s of the DCNN-based. A summary of the results can be seen in table I.

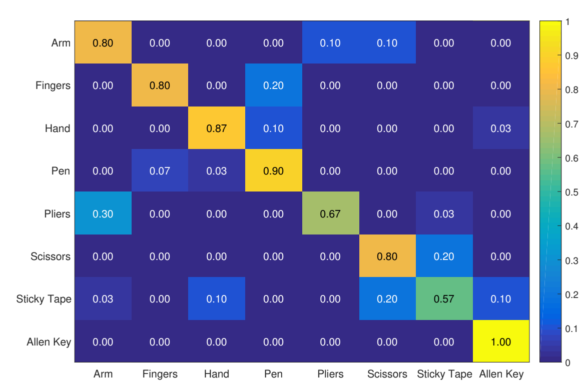


Fig. 4: Confusion matrix of the SURF-based method.

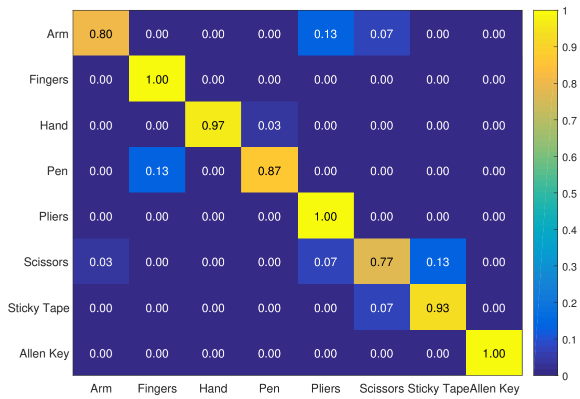


Fig. 5: Confusion matrix of the DCNN-based method.

TABLE I: Summary of results

Method	Accuracy (%)	Improvement (%)	Time (s)
SURF + SVM	80	-	0.01
CNN + SVM	91.67	11.67	0.7

IV. CONCLUSIONS AND FUTURE WORK

Two methods for object recognition using pressure images obtained by a high-resolution tactile sensor have been described. These methods have the same structure based on a features extractor from the pressure images, and a supervised learning algorithm to get a classifier. To extract features, the first method implements a *SURF* descriptor, whilst the second uses a pre-trained DCNN. Then, a *SVM* for each method was trained to get a classifier with the aim to label each image in a pre-determined class. An object recognition application to the field of rescue robotics was also presented. This application consists on classify, using only one touch, inert objects along with humans parts of the body. An experiment with 5-class object and 3-class human parts of the body classification has been carried out to compare both methods in terms of accuracy and computation time. The results show that DCNN-based method has achieved an 11.67% improvement with respect to the method 1, however, the computational time was 0.7s in the method 2 opposite to 0.01s in the method 1. Although the computation time provided shall not be a conclusive evidence, it sheds light on the computational load ratio between methods. Thus, it is demonstrated that the presented methods are valid to classify and distinguish humans parts of the body of inert objects. In future work, we aim to compare our results with existing solutions, and to take advantages of using combined tactile and kinesthetic information. Also, other sensors will be integrated to extract additional information of the state of the victim.

ACKNOWLEDGMENT

This work was partially supported by the Spanish project DPI2015-65186-R and the European Commission under grant agreement BES-2016-078237.

REFERENCES

[1] Ravinder S. Dahiya, Giorgio Metta, Maurizio Valle, and Giulio Sandini. Tactile sensing-from humans to humanoids. *IEEE Transactions on Robotics*, 26(1):1–20, 2 2010.

[2] Alessandro Roncone, Matej Hoffmann, Ugo Pattacini, Luciano Fadiga, and Giorgio Metta. Peripersonal space and margin of safety around the body: Learning visuo-tactile associations in a humanoid robot with artificial skin. *PLoS ONE*, 11(10):e0163713, 10 2016.

[3] Ana Maria Cretu, Thiago Eustaquio Alves De Oliveira, Vinicius Prado Da Fonseca, Bilal Tawbe, Emil M. Petriu, and Voicu Z. Groza. Computational intelligence and mechatronics solutions for robotic tactile object recognition. In *WISP 2015 - IEEE International Symposium on Intelligent Signal Processing, Proceedings*, pages 1–6. IEEE, 5 2015.

[4] Stefan Escaida Navarro, Nicolas Gorges, Heinz Wörn, Julian Schill, Tamim Asfour, and Rudiger Dillmann. Haptic object recognition for multi-fingered robot hands. In *Haptics Symposium 2012, HAPTICS 2012 - Proceedings*, pages 497–502. IEEE, 3 2012.

[5] Alexander Schmitz, Yusuke Bansho, Kuniaki Noda, Hiroyasu Iwata, Tetsuya Ogata, and Shigeki Sugano. Tactile object recognition using deep learning and dropout. *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 1044–1050, 11 2014.

[6] Shan Luo, Wenxuan Mou, Min Li, Kaspar Althoefer, and Hongbin Liu. Rotation and Translation Invariant Object Recognition with a Tactile Sensor. In *IEEE Sensors Conference*, volume 2014-Decem, pages 1030–1033. IEEE, 11 2014.

[7] Shan Luo, Wenxuan Mou, Kaspar Althoefer, and Hongbin Liu. Novel Tactile-SIFT Descriptor for Object Shape Recognition. *IEEE Sensors Journal*, 15(9):5001–5009, 9 2015.

[8] Shan Luo, Wenxuan Mou, Kaspar Althoefer, and Hongbin Liu. Iterative Closest Labeled Point for Tactile Object Shape Recognition. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 10 2016.

[9] Miao Li, Yasemin Bekiroglu, Danica Kragic, and Aude Billard. Learning of grasp adaptation through experience and tactile sensing. In *IEEE International Conference on Intelligent Robots and Systems*, pages 3339–3346. IEEE, 9 2014.

[10] Joaquim Filipe, Institute for Systems, Control Technologies of Information, Communication, Universite de Haute-Alsace, IEEE Control Systems Society, IEEE Robotics and Automation Society, and International Federation of Automatic Control. A Tactile-based Grasping Strategy for Deformable Objects Manipulation and Deformability Estimation. 2015.

[11] Yang Gao, Lisa Anne Hendricks, Katherine J. Kuchenbecker, and Trevor Darrell. Deep learning for tactile understanding from visual and haptic data. In *Proceedings - IEEE International Conference on Robotics and Automation*, volume 2016-June, pages 536–543. IEEE, 5 2016.

[12] Yugang Liu and Goldie Nejat. Robotic Urban Search and Rescue: A Survey from the Control Perspective. *Journal of Intelligent & Robotic Systems*, 72(2):147–165, 11 2013.

[13] Anuradha Ranasinghe, Nantachai Sornkarn, Prokar Dasgupta, Kaspar Althoefer, Jacques Penders, and Thrishantha Nanayakkara. Salient Feature of Haptic-Based Guidance of People in Low Visibility Environments Using Hard Reins. *IEEE Transactions on Cybernetics*, 46(2):568–579, 2 2016.

[14] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359, 2008.

[15] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks, 2012.

[16] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 9 1995.

[17] Alexander Schneider, Jurgen Sturm, Cyrill Stachniss, Marco Reisert, Hans Burkhardt, and Wolfram Burgard. Object identification with tactile sensors using bag-of-features. *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 243–248, 10 2009.

[18] Yufei Wang and Garrison W. Cottrell. Bikers are like tobacco shops, formal dressers are like suits: Recognizing urban tribes with caffe. In *Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015*, pages 876–883. IEEE, 1 2015.