

CNN-Based Methods for Object Recognition With High-Resolution Tactile Sensors

Juan M. Gandarias¹, Student Member, IEEE, Alfonso J. García-Cerezo, Member, IEEE, and Jesús M. Gómez-de-Gabriel¹, Member, IEEE

Abstract—Novel high-resolution pressure-sensor arrays allow treating pressure readings as standard images. Computer vision algorithms and methods such as convolutional neural networks (CNN) can be used to identify contact objects. In this paper, a high-resolution tactile sensor has been attached to a robotic end-effector to identify contacted objects. Two CNN-based approaches have been employed to classify pressure images. These methods include a transfer learning approach using a pre-trained CNN on an RGB-images dataset and a custom-made CNN (TactNet) trained from scratch with tactile information. The transfer learning approach can be carried out by retraining the classification layers of the network or replacing these layers with an SVM. Overall, 11 configurations based on these methods have been tested: eight transfer learning-based, and three TactNet-based. Moreover, a study of the performance of the methods and a comparative discussion with the current state-of-the-art on tactile object recognition is presented.

Index Terms—Tactile sensors, object recognition, deep learning.

I. INTRODUCTION

SENSE of touch is essential for human beings to perform complex tasks such as object recognition. As for humans, tactile information is also useful for robotics systems [1], [2]. Nowadays research studies in robotics are focusing on making robots more similar to humans. That includes providing the robot tactile perception [3], [4]. However, tactile sensing is not enough, and some level of intelligence is required. This way, tactile perception is still a fundamental problem in field robotics that has not been solved so far [5]. To overcome this, two points have to be addressed: first, obtaining sensing tactile information, and second, having the cognitive capabilities to process that information [6].

On the one hand, multiple tactile sensors (as the one presented in Fig. 1) have been employed in robotics manip-

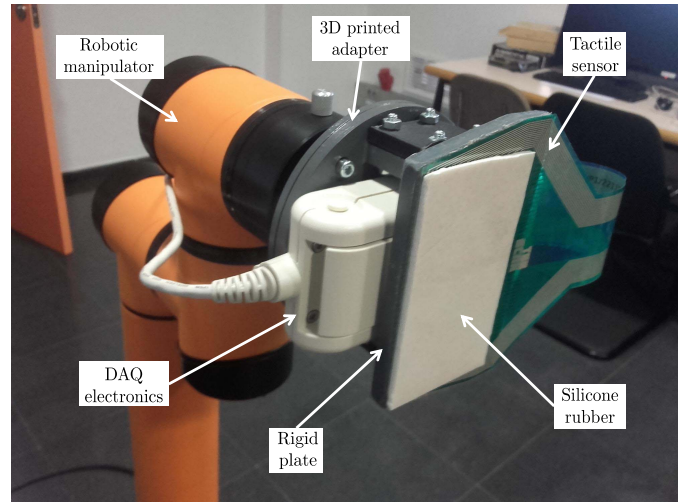


Fig. 1. The tactile end-effector attached on the robotic manipulator AUBO OUR-i5 as a platform to touch objects and collect data.

ulators and grippers [7], [8] to carry out a large variety of applications such as slippage detection [9], [10], tactile object recognition [11], [12] and surface classification [13]–[15]. Tactile sensors estimate the pressure by measuring different physical magnitudes. Thus, the information given depends on the nature of the transducer. For example, typical piezoresistive tactile sensors produce a pressure image where each pixel represents the pressure measured by each tactel [16], while hall-effect sensors are used in combination with magnets and elastic materials [17] to estimate the pressure in larger areas. Other technologies have also been used to estimate contact pressure from other physical magnitudes such as light [18], air pressure [19] or capacitance [20].

On the other hand, novel Artificial Intelligence (AI) techniques can be used for interpreting the information perceived by tactile sensors. Machine learning methods such as Gaussian Processes [21], Bayesian approaches [22], k-mean clustering and Support Vector Machines (SVM) [23] or k-Nearest Neighbour (kNN) [24] have been used for distinguishing contacts. Novel Convolutional Neural Networks (CNNs) are also acquiring excellent results in multiple applications such as visual object recognition [25]. These methods can be used for recognizing objects contacted through tactile sensors [26]–[28].

It is also possible to transfer the knowledge from other domains to reduce the efforts of data collection, building a CNN model, and training the new network from scratch

Manuscript received April 3, 2019; accepted April 10, 2019. Date of publication April 23, 2019; date of current version July 17, 2019. This work was supported in part by the Spanish project DPI2015-65186-R and in part by the European Commission under Grant BES-2016-078237. This is an expanded paper from the IEEE SENSORS 2017 Conference. The associate editor coordinating the review of this paper and approving it for publication was Dr. Marco J. Da Silva. (Corresponding author: Juan M. Gandarias.)

The authors are with the Robotics and Mechatronics Group, Systems Engineering and Automation Department, University of Málaga, Escuela de Ingenierías Industriales, 29071 Málaga, Spain (e-mail: jmgandarias@uma.es).

This paper has supplementary downloadable multimedia material available at <http://ieeexplore.ieee.org> provided by the authors. The Supplementary Material contains code and data used in this work. This code and data is also uploaded in the following repository, where possible errors will be fixed and updated: <https://github.com/TaISLab/CNN-based-Methods-for-Tactile-Object-Recognition>. This material is 2.80 MB in size.

Digital Object Identifier 10.1109/JSEN.2019.2912968

1558-1748 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

TABLE I

MAIN DIFFERENCES BETWEEN RGB AND TACTILE IMAGES. *THE NUMBER OF FEATURES DEPENDS ON EACH IMAGE ITSELF, BUT IN GENERAL TERMS IT COULD BE AFFIRMED THAT TACTILE IMPRINTS HAVE LOWER NUMBER OF FEATURES THAN RGB IMAGES

Factors	RGB	Tactile
Spatial resolution	1.6×10^5	1.4×10^3
Color	Yes	No
Background	Yes	No
Magnitude	Light intensity	Pressure
Number of features*	High	Low
Object in image	Whole object	Depends on the size
Area of detection	Large	Small
Depth of field	Large	Small
Type of perception	Passive	Active
3D visual signs	Shading, texture, focus, movement, perspective	No
Amount of data	1.3×10^6	1.1×10^3

(i.e. transfer learning). This way, some applications can take advantage of the developments made for similar fields (e.g. pressure versus visual image classification), using previous calibration data reduce training time for new devices, and labelling from one classifier can be reused for different input sets [29]. CNN for video image recognition automatically detect features and provide image invariances such as rotation, scale, and brightness [7]. In particular, transfer learning is specially useful when the amount of training data from the domain of interest is limited. In [30], the transfer when the amount of target data is minimal is studied.

The available deep CNNs trained for image recognition can have millions of weights, and the size of the layers can be bigger than the actual input data from the new application domain. The application of these traditional CNNs to tactile object recognition may be not efficient due to the shortage of features in tactile images in comparison with RGB images, and also to the lower resolution of tactile images among other factors. The main differences between RGB and tactile images can be seen in Fig. 2 and are summarized in Table I. Moreover, using these networks for embedded applications, with computational and energy constraints, is not suitable. For embedded applications in domains with a smaller input data size (i.e. lower resolution images) or a lower number of output labels, a custom CNN built from scratch may be done off-line in a different computer) with an eventually limited amount of training data has to be addressed using data augmentation.

This paper is focused on the application of tactile perception techniques for practical autonomous robots, that have limited computational power and energy. In that sense, this work provides comparative results of eleven CNN-based approaches to recognize pressure images in terms of recognition rate

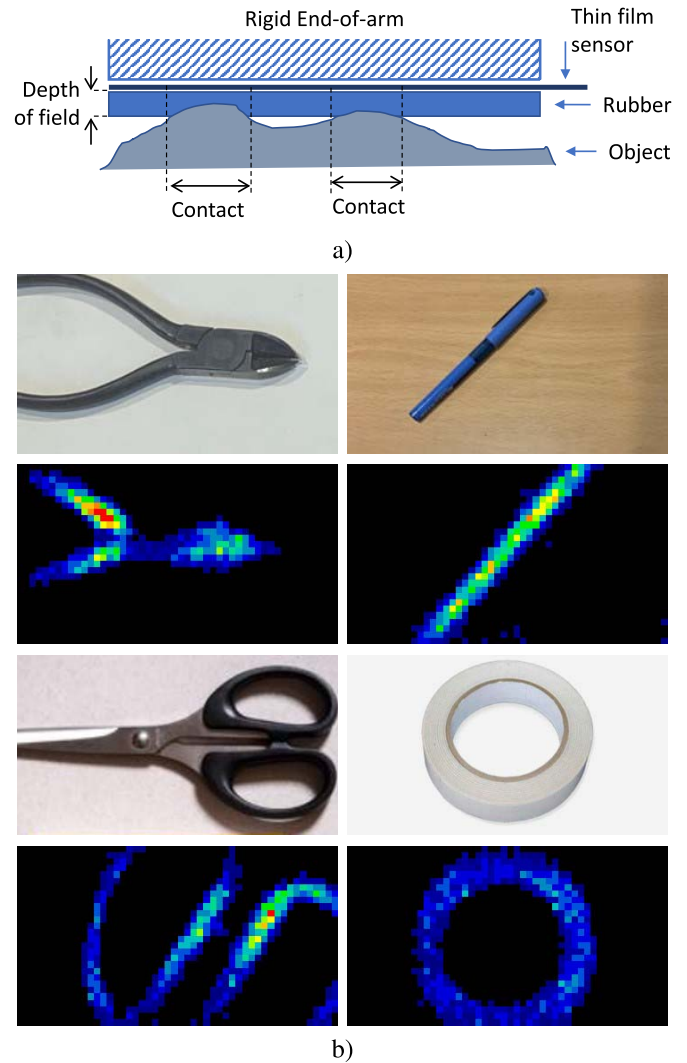


Fig. 2. Differences between visual and pressure images a) Illustration of the depth-of-field in pressure images. b) Examples of video (top) vs pressure (bottom) images for common objects (from left to right and top to bottom: pliers, pen, scissors and tape).

and computational load. Two main approaches are presented: transfer learning from CNNs pre-trained in a dataset of RGB images, and training a custom CNN from scratch using tactile information only. Tactile data is collected by a high-resolution piezoresistive tactile sensor which has been attached to a robotic manipulator as shown in Fig. 1. CNN-based techniques, commonly used in visual object recognition tasks, are used to classify tactile data. In particular, different CNN-based methods are evaluated in a 22-classes experiment. Transfer learning approach is compared with a custom-made CNN in terms of accuracy and classification time. The customized CNN has been named *TactNet* and includes 3 configurations: TactNet-4, TactNet-6 and TactResNet. The results are discussed, and the most beneficial situation for using one method or another is determined based on the experimentation, considering the recognition rate and the classification time. Besides, these approaches are analyzed and compared against the current state-of-the-art. Therefore, this paper can also be used as a guide for further readers whose intend to use CNNs

to classify pressure images. Furthermore, the data and the methods have been made available in the GitHub cloud-based platform, so they can be used with different methods to compare new approaches while providing reproducibility to this work.

II. RELATED WORK

In literature, two main approaches can be followed: surface/material identification and object recognition. Material identification refers to the extraction of information for the superficial properties of an object like roughness, texture, stiffness or thermal conductivity [15], [31]–[33]. Sequential data is important to identify the properties of objects. The interpretation of tactile information as time-series of data is presented in multiple works about material identification [34], [35]. An exploratory motion is carried out with a robotic arm in [36] to obtain dynamic information with a 2D force sensor (normal and shear forces). The control of the actuator is critical to maintain the pressure against the contact material constant and obtain trustworthy information. In that case, a multi-channel neural network is used and a high accuracy is achieved.

On the other hand, object recognition concerns to distinguish objects by its shape [37]–[39]. Apart from the approach, most of these works are based on the use of AI methods to identify tactile information. The use of one method or another depends on the type of the sensed information.

Related works used Neural Networks and Deep Learning techniques to recognize objects [40], [41]. This way, the benefits of using CNNs can be exploited. The translation-invariant property allows the recognition independently of the position of the object in the image [42]. In [23], a CNN-based method is developed to recognize in-contact human hand with artificial skin. The method takes advantage of the translation-invariance of the networks, being able to identify both right and the left hands, indistinctly. The use of Deep Learning with dropout to reduce overfitting is presented in [41]. 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. Luo et al. also propose a novel algorithm which synthesizes both kinesthetic and tactile information, forming a 4D point cloud of the object with the label numbers of the tactile features as an additional dimension to the 3D sensor positions [12].

Other idea consists of applying multi-modal techniques [43]. Fusing haptic and visual data generally presents better results than using a single kind of information [44]. In [45] a multi-modal deep learning method based on a CNN is presented. The network takes the information sensed with an accelerometer and the corresponding image of the surface material as inputs and estimates the contact material.

The information acquired by a pressure sensor array can be represented as a matrix, the same way as an image. Multiple studies treat tactile data as images [7], [23], [24], [46], [47]. Most of them are based on the same two steps: feature extraction from pressure images and obtaining a classifier based on those features [48]. An existing solution uses a variant

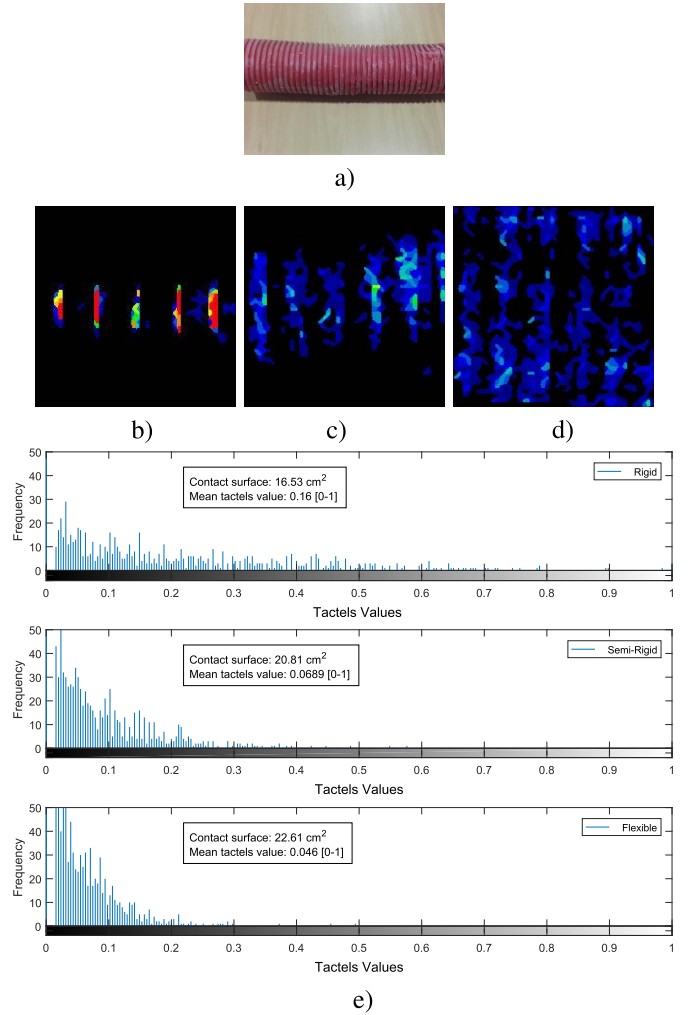


Fig. 3. Pressure images have limited field-of-depth and depend on the compliance of the object and the sensor. a) Sensed object. b) Pressure image from a rigid sensor. c) Pressure image from a semi-rigid sensor. d) Pressure image from a flexible sensor. e) Histograms of the pressure images, including contact surface and mean tactel value for images b) c) and d).

of the *Scale Invariant Feature Transform* (SIFT) descriptor as a feature extractor and a supervised *k-Nearest Neighbour* (kNN) algorithm to get the classifier [24]. However, multiple touches are needed to obtain good results in the classification task.

Pressure data extracted from tactile arrays can also be considered as a series of images. A flexible high-resolution tactile sensor is used in [27] to classify food textures. A CNN is used with a sequence of pressure images, taken when an experimental device crushes the food imitating the movement of a mouth closing and biting food. Although time series of tactile data is important to record the behaviour of gel-like food, authors found that using only two pressure images produces similar results.

III. METHODOLOGIES

A. Tactile Images

Pressure images have limited field-of-depth and depend on the compliance of the object and the sensor. In Figure 3,

the pressure images obtained with three different sensor compliances, from the same plastic pipe are shown. Also, quantitative measurements have been done and included in the histograms in Fig.3(e), that shows how pressure images obtained from flexible sensors that bend around the objects have higher contact area, but lower dynamic range than the images obtained with a flat rigid sensor [28]. The rigid configuration has higher pressure values than the others and avoid bending the thin-film sensor. This arrangement (as seen in Figure 1) includes a layer of silicon rubber to further protect the film and enhance the depth of field (see Fig. 2) of the sensor, and has been chosen for the rest of the experiments in this paper.

B. Transfer Learning

Transfer learning consists of using a trained neural network for a different purpose for which it was created. Thus, a CNN trained in a large amount of RGB images can be used to classify pressure images. This idea takes advantage of the particular roles played by each parts of a CNN. The first convolutional layers of the network learn to extract features from images, whereas the last layers learn to classify the input data. Hence, this approach consists of using the convolutional layers of a pre-trained CNN and replacing the classification layers with a custom classifier.

The structure of this approach is presented in Fig. 4. The first step consists of training a CNN in a large image dataset. This CNN has two distinguishable parts: the firsts convolutional layers ($\mathcal{C} = [\text{conv}_1, \dots, \text{conv}_k]$) with filters sizes $[f_1 \times f_1], \dots, [f_k \times f_k]$; and the lasts fully-connected layers ($\mathcal{F} = [\text{fc}_1, \dots, \text{fc}_n]$). After the training process, the convolutional part of the network learns to extract features \mathcal{F} from RGB images, while the fully-connected layers learn to classify this features.

Considering this factor, the first part of the trained network ($\hat{\mathcal{C}} = [\text{conv}_1, \dots, \text{conv}_k]$) which has learn to extract features from common RGB images \mathcal{F} , can be exploited to extract features from pressure images \mathcal{F}' . In this step, some trained fully-connected layers ($\hat{\mathcal{F}} = [\hat{\text{fc}}_1, \dots, \hat{\text{fc}}_t]$ with $t < n$) can also be employed.

However, as the structure of the network has been set previously, the size of the input tactile data ($\mathbf{T}(r, c) \forall r = 1, \dots, R, c = 1, \dots, C$) has to match the size of the RGB data used for training the convolutional layers ($\mathbf{I}(r', c', z) \forall r' = 1, \dots, R', c' = 1, \dots, C', z = 1, 2, 3$).

To use the transfer learning approach and take advantage of the pre-learned parameters, the whole structure of the network must be maintained. Hence if the pre-trained network has been trained to classify $\mathbf{I}[R' \times C' \times 3]$ images, this image size has to be kept. Thus it is needed to resize tactile images $\mathbf{T}[R \times C \times 1]$ to $\mathbf{T}''[R' \times C' \times 3]$.

Then the bicubic interpolation method has been applied. This method obtains better quality results than others, although it needs larger amount of calculation [49]. However, the time of this calculation can be disregarded in comparison with the neural network computation time. Each pixel $\mathbf{T}'(r', c')$ of the resized image is obtained from its 16-neighbour pixels

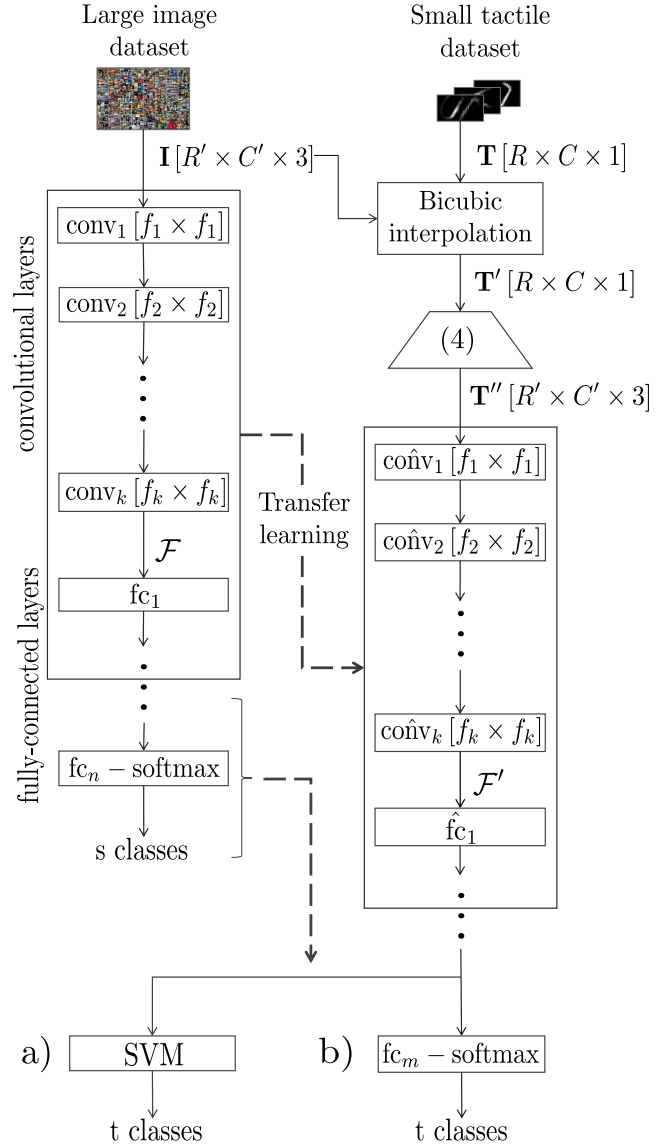


Fig. 4. Structure of the transfer learning approach. The replacement of the classification layers can be carried out using an SVM a) or fine-tuning the classification layers of the network b).

applying the equation (1):

$$\mathbf{T}'(r', c') = \sum_{i=0}^3 \sum_{j=0}^3 [r_i \times c_j] \mathbf{T}(i, j), \quad (1)$$

where $\mathbf{T}(i, j)$ is the 16-neighbour matrix, and r_i and c_j can be obtained by the equations (2) and (3):

$$r_i = \prod_{k=0, k \neq i}^3 \frac{r' - [S_r \times (x + k)]}{[S_r \times (x + i)] - [S_r \times (x + k)]}, \quad (2)$$

$$c_j = \prod_{k=0, k \neq j}^3 \frac{c' - [S_c \times (y + k)]}{[S_c \times (y + j)] - [S_c \times (y + k)]}, \quad (3)$$

x and y are the values of each row and column divided by the scale factor $S_r = R/R'$ and $S_c = C/C'$ respectively. Finally, according to equation 4, the matrix $\mathbf{T}'[R' \times C' \times 1]$ is used in each of the 3 channels of \mathbf{T}'' to match the size of

the original RGB image $\mathbf{I} [R' \times C' \times 3]$, resulting a 3-channel tactile image $\mathbf{T}' [R' \times C' \times 3]$.

$$\mathbf{T}''(r', c', z) = \mathbf{T}'(r', c'), \quad z \in [1, 2, 3], \quad (4)$$

Then, as shown at the bottom of the scheme in Fig. 4, the classification layers of the network can be replaced by two classifiers. On one hand, the substitution of the classification layer by an SVM is contemplated in Fig. 4 a). On the other hand, a simple idea is to train the classification layers again (fine-tuning) on a tactile dataset Fig. 4 b).

C. TactNet

Other idea consists of creating a CNN from scratch and training it from a pressure image dataset, however, a large amount of tactile data is needed. For this work, a CNN for tactile information classification (TactNet) has been created. In this study, this network is configured with 4 and 6 plain layers, following the architecture of the AlexNet [25], and with 6 layers including residual convolutions, following the structure of the ResNet [50].

These networks are composed by a set of convolutional ($\mathcal{C} = [\text{conv}_1, \dots, \text{conv}_k]$) or residual convolutions ($\mathcal{R} = [\text{res}_1, \dots, \text{res}_k]$) which learn to extract tactile features \mathcal{F}_t from pressure images. A batch normalization layer with $\epsilon = 10^{-4}$ is introduced after each convolution, followed by a rectified linear unit (ReLU) to introduce non-linearities. Some layers are alternated with a max-pooling layer with a stride of 2. Finally, a set of n fully connected layers ($\mathcal{F} = [\text{fc}_n]$) learn to classify the input data.

The details of the three different configurations of TactNet that have been considered in this work are explained below. The networks were implemented in *Matlab R2018b* using the *Neural Network Toolbox*. The implementation details can be found at the GitHub repository¹

1) *TactNet-4*: This is the simplest configuration of TactNet. The network is composed by 3 convolutional layers ($\mathcal{C} = [\text{conv}_1, \text{conv}_2, \text{conv}_3]$) with filters sizes $[5 \times 5]$, 8 $[3 \times 3]$, 16 and $[3 \times 3]$, 32 respectively; and a fully-connected layer ($\mathcal{F} = [\text{fc}_4]$) with 22 neurons followed by a softmax-layer to classify the input tactile data and gives the likelihood of belonging to each class. The structure of this network is shown in Fig. 5 a).

2) *TactNet-6*: In this case, the network is formed by 5 convolutional layers ($\mathcal{C} = [\text{conv}_1, \text{conv}_2, \text{conv}_3, \text{conv}_4, \text{conv}_5]$) with filters sizes $[5 \times 5]$, 8 $[5 \times 5]$, 16 $[3 \times 3]$, 32 $[3 \times 3]$, 64 and $[3 \times 3]$, 128 respectively. After that, 1 fully-connected layer ($\mathcal{F} = [\text{fc}_6]$) with 22 neurons, and a softmax-layer to get the probabilities. The structure of this network is shown in Fig. 5 b).

3) *TactResNet*: This network is formed by 1 convolutional layer ($\mathcal{C} = [\text{conv}_1]$), 4 residual layers ($\mathcal{R} = [\text{res}_2, \text{res}_3, \text{res}_4, \text{res}_5]$), and 1 fully connected layer ($\mathcal{F} = [\text{fc}_6]$) followed by a softmax-layer to get the

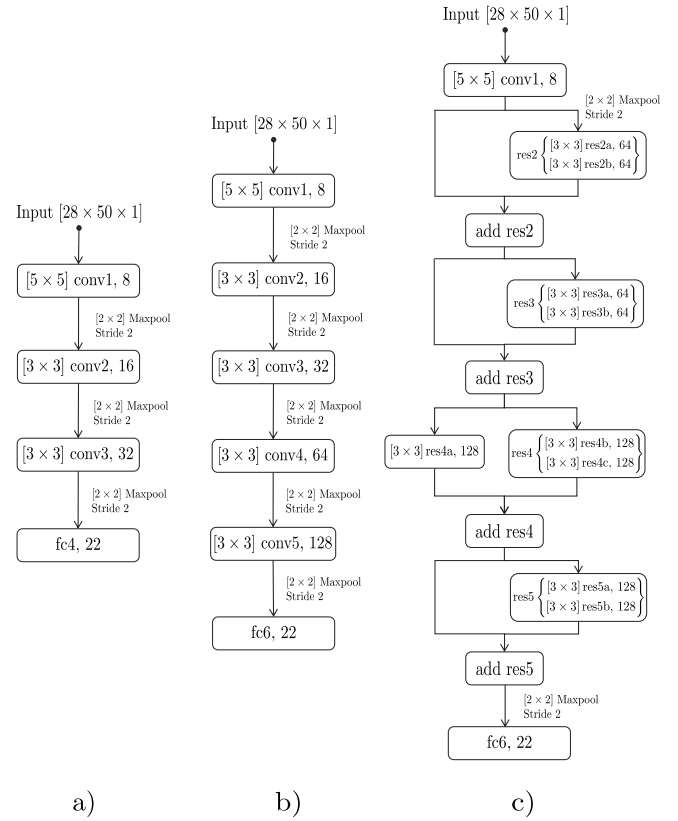


Fig. 5. Structure of the TactNet configurations: TactNet-4 a), TactNet-6 b) and TactResNet c).

probabilities. The filters and specifications of this network are presented in Fig. 5 c).

IV. EXPERIMENTAL SETUP

A. Description of the Experiments

The performance of the proposed methodologies is evaluated based on the recognition rates achieved in an experiment with objects from 22 classes. Tactile information from these objects has been collected to carry out the learning processes.

Tactile data acquisition is a critical task as training data determines the learning process, therefore the tactile sensor is placed in contact with different objects manually. The operator decides when a sample (pressure image) is valid or not based on the visual information provided by the user interface of the *I-Scan* data acquisition software, provided by *Tekscan*. With this software the operator sees the pressure map in real-time.

In Fig. 2, an example of tactile imprints of four objects used for the experiment is presented. The real-time pressure map provided by the data acquisition software is formed by 15 values (colors) image. The color of each pixel depends on the pressure applied in each tactel. Hence, the minimum pressure (black) corresponds to 0 Pa, whereas the maximum pressure (red) corresponds to the maximum pressure admitted by the sensor.

The experimental implementation consists of testing and comparing 11 different classification procedures. Following

¹<https://github.com/TaISLab/CNN-based-Methods-for-Tactile-Object-Recognition>

TABLE II
SUMMARY OF THE CNN-BASED TECHNIQUES USED FOR THE CLASSIFICATION EXPERIMENT (TL = TRANSFER LEARNING, CONV = CONVOLUTIONAL LAYER, RES = RESIDUAL LAYER, SM = SOFTMAX)

Approach	Name	Feat. Extractor	Classifier
TL-SVM (Fig.4a)	SqueezeNet-SVM	SqueezeNet	SVM
	AlexNet-SVM	AlexNet	SVM
	ResNet-SVM	ResNet	SVM
	VGG-SVM	VGG16	SVM
TL-NN (Fig.4b)	SqueezeNet-NN	SqueezeNet	1conv-SM
	AlexNet-NN	AlexNet	2fc-SM
	ResNet-NN	ResNet	1conv-SM
	VGG-NN	VGG16	3fc-SM
TactNet (Fig.5)	TactNet-4	3conv	1fc-SM
	TactNet-6	5conv	1fc-SM
	TactResNet	1conv-4res	1fc-SM

the transfer learning approach, 8 methods have been considered: 4 for the SVM classifier and 4 for re-training the last layers of the network as a classifier (see Fig. 4). The pre-trained networks used for extract features are AlexNet [25], ResNet [50], SqueezeNet [51] and VGG16 [52]. On the other hand, the 3 configurations of TactNet presented in Fig. 5 have been also contemplated. The 11 methods are summarized in Table II. Moreover, the SURF-SVM method described in [11] has been included in the experiment as a comparison point to show the enhancement of the CNN-based methods against a traditional approach that doesn't use deep learning.

To carry out the experiment, two different hardware systems are used. A GPU NVidia GeForce GTX 1050 Ti with 4 GB of RAM and a CPU Intel Core i7-7700HQ CPU @ 2.80 GHz. The methods have been trained using both systems and the classification time has been measured on each one.

B. Sensor Specifications

The tactile end-effector employed for the experiments, shown in Fig. 1, uses a high-resolution tactile sensor that has been attached to the 6 DOF robotic arm *AUBO Our-i5*. An own-designed 3D printed part is used for coupling the sensor to the robotic manipulator. This component has been manufactured using Fused Deposition Modelling (FDM) 3D printing in PLA (PolyLactic Acid) plastic.

The data acquisition system is constituted by the tactile sensor model 6077, the *Evolution Handle* and the *I-Scan* software, which are provided by *Tekscan* (South Boston, MA, USA). The high-resolution tactile-array has a total of 1400 pressure sensels (also known as tactels or taxels). Each tactel has a size of 53.3×95.3 mm, conforming a set of resistive pressure sensors with density 27.6 tactels/cm² distributed in a matrix of 28 rows by 50 columns. The main features of the sensor are detailed in Table III.

The tactile sensor is covered by a silicone rubber as a contact interface which protects the sensor and conducts external forces. A silicon rubber pad (sourced by *RS* with

TABLE III
SPECIFICATIONS OF THE HIGH-RESOLUTION TACTILE SENSOR MODEL 6077 FROM *Tekscan*

Parameter	Value
Max. pressure	34 KPa
Number of tactels	1400
Tactels density	27 tactels/cm ²
Temperature range	-40 °C to +60 °C
Matrix height	53.3 mm
Matrix width	95.3 mm
Thickness	0.102 mm

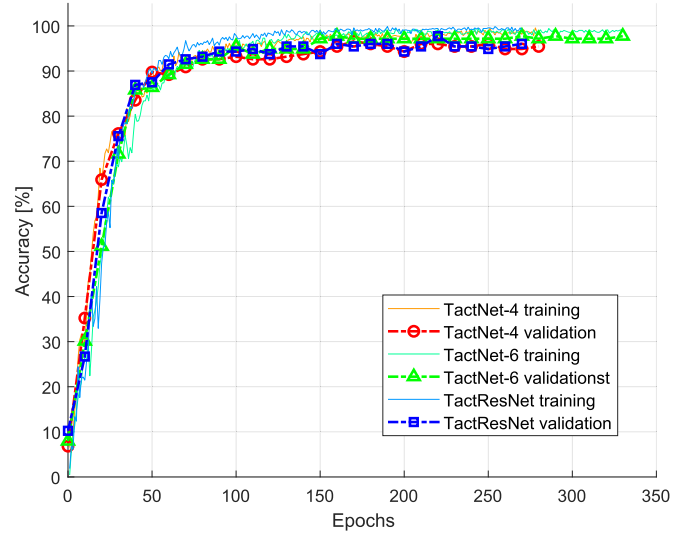


Fig. 6. Learning process of the 3 TactNet configurations.

code 733 – 6713) provides soft contact between the sensor and the objects.

C. Dataset

For training the models, the use of two datasets is considered. On one hand, CNNs used in transfer learning models are pre-trained on more than a million images and can classify images into 1000 object categories. The ImageNet dataset is formed by common RGB images and is used in the ImageNet Large-Scale Visual Recognition Challenge [53].

On the other hand, a dataset formed by tactile images has been collected. This dataset is used to train the TactNet models and the classifiers of the transfer learning methods. A total of 1100 pressure images have been used to feed each method. These images are divided into 22 classes labelled as: *Adhesive, allen key, arm, ball, bottle, box, branch, cable, cable pipe, caliper, can, finger, hand, highlighter pen, key, pen, pliers, rock, rubber, scissors, sticky tape and tube*.

Learning processes require three subsets of tactile data: training set, validation set and test set. The training set is composed by 704 images, 32 images for each label, whilst the validation and test set are composed by 176 and 220 images respectively.

TactNet models are trained using data augmentation techniques. Exploiting the translation, rotation and scale invariant

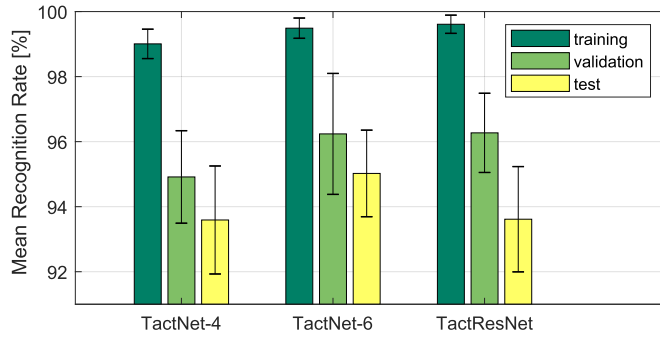


Fig. 7. Accuracy achieved after the 20 samples of learning process by each TactNet configurations with the training, validation and test sets. The error bars represent the one standard deviation.

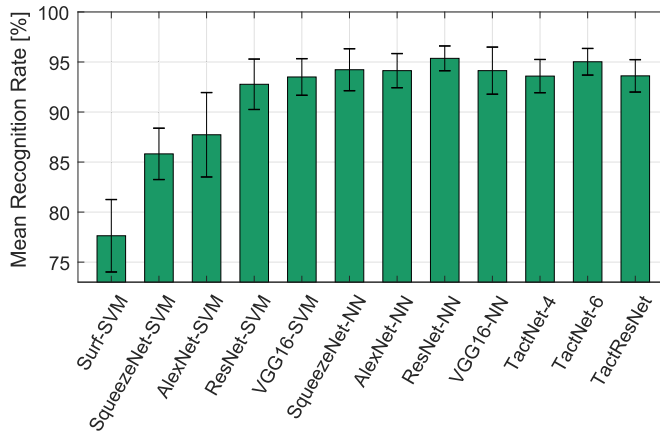


Fig. 8. Comparison of the recognition rate achieves by each method. The error bars represent the one standard deviation.

traits introduced by data augmentation, the models are able to recognize touched objects independently of their location in the image, orientation or contact pressure [54]. Note that ImageNet dataset already includes data augmentation techniques, therefore these invariant qualities are also presented in transfer learning models.

In the case of the learning process of the different configurations of TactNet, the data augmentation techniques considered have been reflections, rotations and translations in X and Y axis. Hence, the amount of tactile data available for training the networks is 4224, while the validation and test set are 1056 and 1320 respectively.

V. RESULTS

In this section, the results achieved by the applied methods according to the experiment statement are shown and explained.

Fig. 6 shows the learning processes of each TactNet configuration from scratch. In this graph, it is shown that the accuracy achieved by each configuration is close to 100%.

As the images used for the training process affect the performance of the network, 20 training sets have been randomly collected from the dataset to train each method, then 20 validation and 20 test sets were formed, also randomly, with the rest of the data not used for training. The accuracy achieved

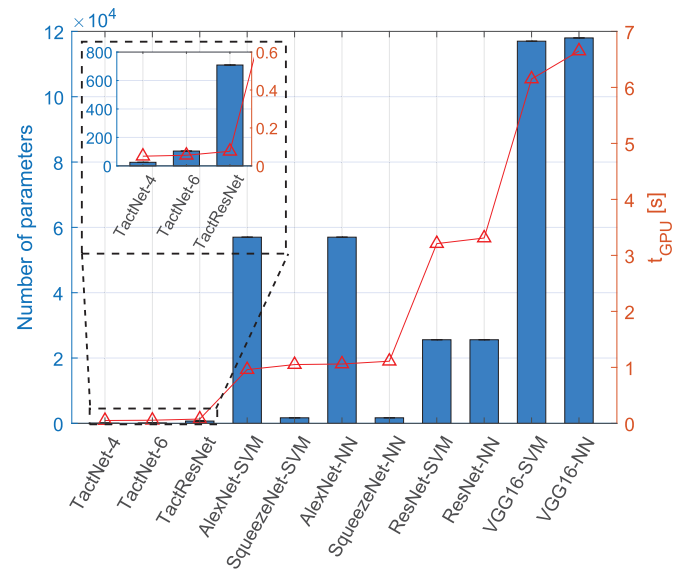


Fig. 9. Comparison of the number of parameter of each method, and the computation time for each one. The computation time considered in this graph corresponds to the classification time of the methods running on a NVidia GeForce GTX 1050Ti GPU.

TABLE IV
PERFORMANCE COMPARISON BETWEEN METHODS USED
FOR TACTILE OBJECT RECOGNITION

Model	Parameters	Accuracy[%]	t _{GPU} [s]	t _{CPU} [s]
SURF-SVM	-	77.64	0.025	0.029
SqueezeNet-SVM	1.7M	85.82	1.05	7.934
AlexNet-SVM	57 M	87.73	0.96	4.141
ResNet-SVM	25.6M	92.77	3.21	23.848
VGG16-SVM	117 M	93.50	6.15	73.355
SqueezeNet-NN	1.7M	94.23	1.11	8.515
AlexNet-NN	57 M	94.14	1.06	4.982
ResNet-NN	25.6M	95.36	3.31	27.039
VGG16-NN	118 M	94.14	6.65	64.488
TactNet-4	25 k	93.59	0.051	0.094
TactNet-6	104 k	95.02	0.056	0.103
TactResNet	790 k	93.61	0.077	0.465

by each TactNet configurations over the 20 training samples is presented in Fig. 7. Each bar represents the mean recognition rate achieved by each configuration with the training, validation and test sets. The error bars represent the one standard deviation over the 20 samples. As expected, the recognition rate achieved with the training rate is the highest while the one achieved with the test set is the lowest. It is also shown that the mean recognition rate is over 93% for all the methods.

Fig. 8 presents a comparison graph of the mean recognition rate achieved by the transfer learning approaches and TactNet configurations. The method SURF-SVM has also been evaluated to be compared against the CNN-based approaches. As in the case of the TactNet, the training process of the transfer learning methods and SURF-SVM has been carried out with 20 samples of randomized data.

TABLE V
STATE-OF-THE-ART ON TACTILE OBJECT RECOGNITION. DUE TO THE NATURE OF THIS WORK, AND THE VARIETY OF THE APPROACHES, ONLY THOSE STUDIES WHICH TREAT TACTILE DATA AS COMMON IMAGES ARE CONSIDERED. THE LAST THREE CONFORM THE METHODS PRESENTED IN THIS PAPER

Year	Method	Tactile sensor	N° of classes	Accuracy [%]	N° of touches	Type of objects
2009	Bag of Features [47]	2 fingers tactile array	21	84.6	10	Household + industrial
2012	PCA + SOM + BoK + ANN [56]	Schunk dexterous hand	10	78.88	4	Household
2014	Stacked DAEs [41]	Twendy-one hand	20	88.00	4	Household
2015	Tactile-SIFT [24]	$[6 \times 14]$ tactile array	18	91.33	15	Household
2015	Cascade of 2 NN [57]	$[8 \times 16]$ tactile array	25	96.00	1	Symbols (Letters)
2016	Zernike moment + iClap [12]	$[6 \times 14]$ tactile array	20	85.36	20	Household
2016	kernel sparse coding [35]	Barret hand	10	> 90.00	1	Household
2017	SURF + k-mean + SVM [11]	$[28 \times 50]$ Tactile array	8	80.00	1	Household + body parts
2017	CNN + SVM [11]	$[28 \times 50]$ Tactile array	8	91.67	1	Household + body parts
2017	SURF + k-mean + SVM [23]	Array with 768 tactels	2	98.15	1	Body parts
2017	CNN [23]	Array with 768 tactels	2	98.33	1	Body parts
2017	SVM [58]	NTU five-finger hand	2	96.67	1	Cylindrical, spherical or cubes
2018	ROTCovPCE-mv [26]	Multiple hardware	5 - 10	81.66 - 99.03	Sequential data	Household
2018	LDS-FCM [59]	Multiple hardware	5 - 10	91 - 100	1	Household
2018	Transfer Learning (VGG16-SVM)	$[28 \times 50]$ Tactile array	22	93.50	1	Household + body parts
2018	Transfer Learning (ResNet-NN)	$[28 \times 50]$ Tactile array	22	95.36	1	Household + body parts
2018	TactNet-6	$[28 \times 50]$ Tactile array	22	95.02	1	Household + body parts

Another aspect that has to be considered is the classification time. This is highly dependent on the number of parameters of a CNN and the hardware, and according to [55] the number of parameters of a network is proportional to the number of operations, the size and the inference time. Therefore, the higher number of parameters a network has, the more memory and processing time is required. Fig. 9 shows the number of parameters of the CNN used in this work and the computation time for each one running on the NVidia GPU.

On the other hand, the spatial resolution of the sensor plays an important role on the performance of the task. Therefore, a classification experiment with the best methods of each approach: VGG16-SVM for transfer learning with SVM (see Fig.4a), ResNet-NN for transfer learning with fine-tuning (see Fig.4b) and TactNet6 for TactNet (see Fig. 5b), has been carried out using approximately $\times 1/16$, $\times 1/8$, $\times 1/4$, $\times 1/2$ and $\times 1$ resolution of the tactile sensor. The results are presented in Fig. 10. The experiment is done using only one touch to classify the objects, and the resolution of the sensor is decreased by software using the bicubic interpolation (see section IIIB). The structure of TactNet-6 is changed to match the dimensions of the input data for each resolution, and some hyperparameters are changed to ensure good training process. This changes are not necessary in case of transfer learning approaches as the input data is resized again to match the original input size of the network.

VI. ANALYSIS AND DISCUSSION

A summary of the results is presented in Table IV. This shows that both transfer learning and TactNet obtain high accuracy rates on tactile object recognition, and that using a CNN-based approach can raise the mean recognition accuracy up to 17.72% against SURF-SVM in the best case (ResNet-NN). In fact, the worst CNN method in terms of accuracy

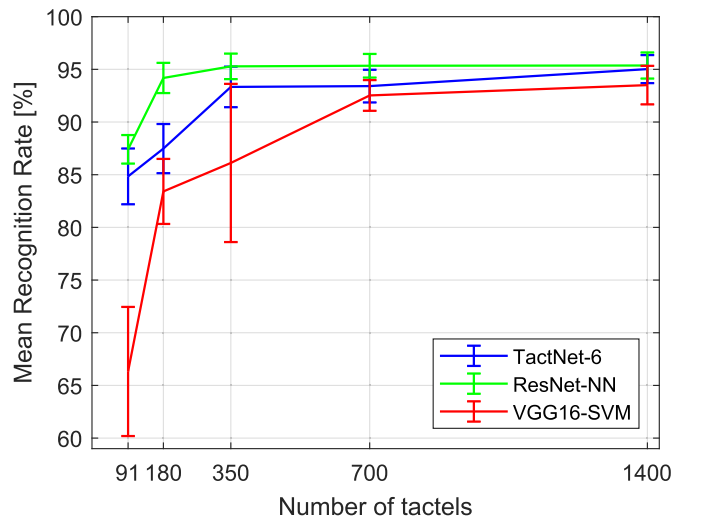


Fig. 10. Comparison of the performance of the best methods of each CNN-based approach (see Table II and Table IV) for different sensor's resolutions using one touch only. The tactels' distribution for each resolution is: 91 – $[7 \times 13]$, 180 – $[10 \times 18]$, 350 – $[14 \times 25]$, 700 – $[20 \times 35]$ and 1400 – $[28 \times 50]$.

(SqueezeNet-SVM) has an improvement of 8.08% against SURF-SVM. The recognition rate in the case of transfer learning depends on the CNN and the classifier, however if a large CNN is used and the classifier consists of fine-tuning the fully-connected layers of the network, the mean recognition rate and the classification time increases. Nevertheless, the classification times of the TactNet configurations are significantly shorter than in transfer learning methods. These results show the better performance of CNN-based method against traditional approaches, especially when using TactNet configurations, which present similar recognition rates to the best transfer learning methods, but need substantially shorter classification times. Moreover, a further reduction in

computing time can be obtained for the TactNet approach by using lower resolutions.

As different researchers are using different hardware, data and methods, to compare the proposed method with existing related works, a table with the qualitative/quantitative features (see Table V) of relevant methods that use tactile information as images, has been included. Different methodologies have been employed, mostly applying computer vision techniques for extracting features, followed by a classifier, as well as the use of Artificial Neural Networks (ANN). It can be seen that a large variety of tactile sensors, number of classes, number of touches and type of objects can be used, and all these aspects have an influence on the accuracy independent of the followed approach. In general, the larger number of classes, the smaller accuracy, and the larger number of touches, the higher accuracy.

With the above considerations, the CNN-based methods presented in this paper offers advantages with respect to the state-of-the-art. In fact, this is the second research work with the largest number of objects in comparison with the state-of-the-art. Only [57] has more classes, with 25 diverse symbols (letters), so that they may be easier to classify. Besides, the methods presented in this paper obtain the highest accuracy of all of those works that have more than 2 classes, with the exception of [57]. Our proposals perform better than other methods with one order of magnitude less in the number of classes and using only one touch.

VII. CONCLUSION

CNN-based methods for tactile object recognition based on high-resolution pressure images, to classify objects along with human-body parts (22 classes) using only one touch, have been presented, evaluated and compared. Tactile images have some similarities to video images, but have strong differences in depth-of-field, resolution, apart from other qualitative considerations that have been discussed in this work. A custom-built CNN with a layer size that matches the size of the tactile image, and has been trained from scratch, in opposition to the use of transfer learning methods, that employ pre-trained CNNs in large video image dataset. In this approach, the classification layers of a CNN are replaced by an SVM or fine-tuned and re-trained in a set of up-sized tactile images with good recognition rates but a high computational cost. Comparative experiments with 11 deep-learning methods (8 transfer learning-based and 3 TactNet-based), have been carried out using the same tactile image dataset on the same hardware to get accuracy and evaluation times, showing the importance of using input layers with just the size of the sensor. Similar recognition rates with a huge reduction of classification time are achieved by TactNet. Full training of from-scratch networks is a worthwhile one-time operation. Other recent methods (such as SURF-SVM) applied to the recognition of tactile images have been qualitative and quantitatively compared with the presented approach without increasing the accuracy over the deep learning methods. A further reduction in the size of the tactile images for recognition with the presented approaches has been experimentally evaluated with different resolutions showing that for our dataset, a

significant reduction in the number of tactels of the pressure image can be done while keeping a high recognition rate. This is important because a reduction in complexity and cost of the tactile sensor and computational power can be done for this kind of images. Finally, tactile sensing is active, therefore, it requires physical interaction between the sensing surface and the environment. In robotics, the use of series of tactile information based on force/displacement will be considered. Considering sequences of tactile images, dynamic information of the contact can be used as input to the CNN to classify and identify physical properties of objects. In this case the performance and computational requirements of the 3D CNNs will be studied optimized to provide real-time execution.

REFERENCES

- [1] S. Luo, J. Bimbo, R. Dahiya, and H. Liu, "Robotic tactile perception of object properties: A review," *Mechatronics*, vol. 48, pp. 54–67, Dec. 2017.
- [2] A. Trujillo-León, W. Bachta, and F. Vidal-Verdú, "Tactile sensor-based steering as a substitute of the attendant joystick in powered wheelchairs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 7, pp. 1381–1390, Jul. 2018.
- [3] C. Bartolozzi, L. Natale, F. Nori, and G. Metta, "Robots with a sense of touch," *Nature Mater.*, vol. 15, no. 9, pp. 921–925, Aug. 2016.
- [4] L. Jamone, L. Natale, G. Metta, and G. Sandini, "Highly sensitive soft tactile sensors for an anthropomorphic robotic hand," *IEEE Sensors J.*, vol. 15, no. 8, pp. 4226–4233, Aug. 2015.
- [5] A. Roncone, M. Hoffmann, U. Pattacini, L. Fadiga, and G. Metta, "Peripersonal space and margin of safety around the body: Learning visuo-tactile associations in a humanoid robot with artificial skin," *PLoS ONE*, vol. 11, no. 10, Oct. 2016, Art. no. e0163713.
- [6] R. S. Dahiya, G. Metta, M. Valle, and G. Sandini, "Tactile sensing-from humans to humanoids," *IEEE Trans. Robot.*, vol. 26, no. 1, pp. 1–20, Jan. 2010.
- [7] J. M. Gandarias, J. M. Gómez-de-Gabriel, and A. J. García-Cerezo, "Enhancing perception with tactile object recognition in adaptive grippers for human-robot interaction," *Sensors*, vol. 18, no. 3, p. 692, Feb. 2018.
- [8] S. Chitta, J. Sturm, M. Piccoli, and W. Burgard, "Tactile sensing for mobile manipulation," *IEEE Trans. Robot.*, vol. 27, no. 3, pp. 558–568, Jun. 2011.
- [9] J. W. James, N. Pestell, and N. F. Lepora, "Slip detection with a biomimetic tactile sensor," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3340–3346, Oct. 2018.
- [10] R. A. Romeo, C. M. Oddo, M. C. Carrozza, E. Guglielmelli, and L. Zollo, "Slippage detection with piezoresistive tactile sensors," *Sensors*, vol. 17, no. 8, p. 1844, Aug. 2017.
- [11] J. M. Gandarias, J. M. Gómez-de-Gabriel, and A. J. García-Cerezo, "Human and object recognition with a high-resolution tactile sensor," in *Proc. IEEE Sensors Conf.*, Oct./Nov. 2017, pp. 1–3.
- [12] S. Luo, W. Mou, K. Althoefer, and H. Liu, "Iterative closest labeled point for tactile object shape recognition," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 3137–3142.
- [13] Q. Yuan and J. Wang, "Design and experiment of the NAO humanoid robot's plantar tactile sensor for surface classification," in *Proc. 4th Int. Conf. Inf. Sci. Control Eng. (ICISCE)*, 2017, pp. 931–935.
- [14] J. Hoelscher, J. Peters, and T. Hermans, "Evaluation of tactile feature extraction for interactive object recognition," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, Nov. 2015, pp. 310–317.
- [15] R. Li and E. H. Adelson, "Sensing and recognizing surface textures using a gelsight sensor," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2013, pp. 1241–1247.
- [16] F. Vidal-Verdú, Ó. Oballe-Peinado, J. A. Sánchez-Durán, J. Castellanos-Ramos, and R. Navas-González, "Three realizations and comparison of hardware for piezoresistive tactile sensors," *Sensors*, vol. 11, no. 3, pp. 3249–3266, 2011.
- [17] D. S. Chaturanga, Z. Wang, Y. Noh, T. Nanayakkara, and S. Hirai, "Magnetic and mechanical modeling of a soft three-axis force sensor," *IEEE Sensors J.*, vol. 16, no. 13, pp. 5298–5307, Jul. 2016.
- [18] B. Ward-Cherrier *et al.*, "The TacTip family: Soft optical tactile sensors with 3D-printed biomimetic morphologies," *Soft Robot.*, vol. 5, no. 2, pp. 216–227, Apr. 2018.

- [19] D. Gong, R. He, J. Yu, and G. Zuo, "A pneumatic tactile sensor for co-operative robots," *Sensors*, vol. 17, no. 11, p. 2592, 2017.
- [20] P. Maiolino, M. Maggiali, G. Cannata, G. Metta, and L. Natale, "A flexible and robust large scale capacitive tactile system for robots," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3910–3917, Oct. 2013.
- [21] Z. Yi *et al.*, "Active tactile object exploration with Gaussian processes," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 4925–4930.
- [22] T. Corradi, P. Hall, and P. Irvani, "Bayesian tactile object recognition: Learning and recognising objects using a new inexpensive tactile sensor," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 3909–3914.
- [23] A. Albin, S. Denei, and G. Cannata, "Human hand recognition from robotic skin measurements in human-robot physical interactions," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 4348–4353.
- [24] S. Luo, W. Mou, K. Althoefer, and H. Liu, "Novel tactile-SIFT descriptor for object shape recognition," *IEEE Sensors J.*, vol. 15, no. 9, pp. 5001–5009, Sep. 2015.
- [25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [26] L. Cao, F. Sun, X. Liu, W. Huang, R. Kotagiri, and H. Li, "End-to-end ConvNet for tactile recognition using residual orthogonal tiling and pyramid convolution ensemble," *Cognit. Comput.*, vol. 10, pp. 718–736, Jun. 2018.
- [27] A. Shibata, A. Ikegami, M. Nakauma, and M. Higashimori, "Convolutional neural network based estimation of gel-like food texture by a robotic sensing system," *Robotics*, vol. 6, no. 4, p. 37, Dec. 2017.
- [28] J. M. Gandarias, J. M. Gómez-de-Gabriel, and A. J. García-Cerezo, "Tactile sensing and machine learning for human and object recognition in disaster scenarios," in *Proc. 3rd Iberian Robot. Conf.* Cham, Switzerland: Springer, 2017, pp. 165–175.
- [29] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- [30] L. Mihalkova and R. J. Mooney, "Transfer learning from minimal target data by mapping across relational domains," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, 2009, pp. 1163–1168.
- [31] D. Feng, M. Kaboli, and G. Cheng, "Active prior tactile knowledge transfer for learning actual properties of new objects," *Sensors*, vol. 18, no. 2, p. 634, Feb. 2018.
- [32] M. Kaboli and G. Cheng, "Robust tactile descriptors for discriminating objects from textural properties via artificial robotic skin," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 985–1003, Aug. 2018.
- [33] S. S. Baishya and B. Bäuml, "Robust material classification with a tactile skin using deep learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 8–15.
- [34] M. Madry, L. Bo, D. Kragic, and D. Fox, "ST-HMP: Unsupervised spatio-temporal feature learning for tactile data," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May/Jun. 2014, pp. 2262–2269.
- [35] H. Liu, D. Guo, and F. Sun, "Object recognition using tactile measurements: Kernel sparse coding methods," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 3, pp. 656–665, Mar. 2016.
- [36] M. Kerzel, M. Ali, H. G. Ng, and S. Wermter, "Haptic material classification with a multi-channel neural network," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2017, pp. 439–446.
- [37] N. Jamali and C. Sammut, "Majority voting: Material classification by tactile sensing using surface texture," *IEEE Trans. Robot.*, vol. 27, no. 3, pp. 508–521, Jun. 2011.
- [38] H. Liu, X. Song, T. Nanayakkara, L. D. Seneviratne, and K. Althoefer, "A computationally fast algorithm for local contact shape and pose classification using a tactile array sensor," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2012, pp. 1410–1415.
- [39] U. Martinez-Hernandez, T. J. Dodd, and T. J. Prescott, "Feeling the shape: Active exploration behaviors for object recognition with a robotic hand," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 12, pp. 2339–2348, Dec. 2018.
- [40] A. Khasnobish *et al.*, "Object-shape recognition from tactile images using a feed-forward neural network," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2012, pp. 1–8.
- [41] A. Schmitz, Y. Bansho, K. Noda, H. Iwata, T. Ogata, and S. Sugano, "Tactile object recognition using deep learning and dropout," in *Proc. IEEE-RAS Int. Conf. Humanoid Robots*, Nov. 2014, pp. 1044–1050.
- [42] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, Jan. 1997.
- [43] P. Falco, S. Lu, A. Cirillo, C. Natale, S. Pirozzi, and D. Lee, "Cross-modal visuo-tactile object recognition using robotic active exploration," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May/Jun. 2017, pp. 5273–5280.
- [44] Y. Gao, L. A. Hendricks, K. J. Kuchenbecker, and T. Darrell, "Deep learning for tactile understanding from visual and haptic data," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2016, pp. 536–543.
- [45] H. Zheng, L. Fang, M. Ji, M. Strese, Y. Özer, and E. Steinbach, "Deep learning for surface material classification using haptic and visual information," *IEEE Trans. Multimedia*, vol. 18, no. 12, pp. 2407–2416, Dec. 2016.
- [46] S. Luo, X. Liu, K. Althoefer, and H. Liu, "Tactile object recognition with semi-supervised learning," in *Proc. Int. Conf. Intell. Robot. Appl.*, 2015, pp. 15–26.
- [47] A. Schneider, J. Sturm, C. Stachniss, M. Reiser, H. Burkhardt, and W. Burgard, "Object identification with tactile sensors using bag-of-features," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2009, pp. 243–248.
- [48] S. Luo, W. Mou, M. Li, K. Althoefer, and H. Liu, "Rotation and translation invariant object recognition with a tactile sensor," in *Proc. IEEE Sensors Conf.*, Nov. 2014, pp. 1030–1033.
- [49] M. B. Hisham, S. N. Yaakob, R. A. A. Raof, A. B. A. Nazren, and N. M. Wafi, "An analysis of performance for commonly used interpolation method," *Adv. Sci. Lett.*, vol. 23, no. 6, pp. 5147–5150, 2017.
- [50] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [51] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer. (2016). "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size." [Online]. Available: <https://arxiv.org/abs/1602.07360#>
- [52] K. Simonyan and A. Zisserman. (2014). "Very deep convolutional networks for large-scale image recognition." [Online]. Available: <https://arxiv.org/abs/1409.1556>
- [53] O. Russakovsky *et al.*, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.
- [54] E. Kauderer-Abrams. (2017). "Quantifying translation-invariance in convolutional neural networks." [Online]. Available: <https://arxiv.org/abs/1801.01450>
- [55] A. Canziani, A. Paszke, and E. Culurciello. (2016). "An analysis of deep neural network models for practical applications." [Online]. Available: <https://arxiv.org/abs/1605.07678>
- [56] S. E. Navarro, N. Gorges, H. Wörn, J. Schill, T. Asfour, and R. Dillmann, "Haptic object recognition for multi-fingered robot hands," in *Proc. Haptics Symp.*, 2012, pp. 497–502.
- [57] A.-M. Cretu, T. E. A. de Oliveira, V. P. da Fonseca, B. Tawbe, E. M. Petriu, and V. Z. Groza, "Computational intelligence and mechatronics solutions for robotic tactile object recognition," in *Proc. IEEE Int. Symp. Intell. Signal Process. (WISP)*, May 2015, pp. 1–6.
- [58] W.-Y. Lee, M.-B. Huang, and H.-P. Huang, "Learning robot tactile sensing of object for shape recognition using multi-fingered robot hands," in *Proc. 26th IEEE Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Aug./Sep. 2017, pp. 1311–1316.
- [59] C. Liu, W. Huang, F. Sun, M. Luo, and C. Tan, "LDS-FCM: A linear dynamical system based fuzzy C-means method for tactile recognition," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 1, pp. 72–83, Jan. 2019.



Juan M. Gandarias received the B.S. and M.S. degrees in mechatronics engineering from the University of Málaga, in 2015 and 2017, respectively, where he is currently pursuing the Ph.D. degree in mechatronics engineering with the Robotics and Mechatronics Laboratory, System Engineering and Automation Department.

He has participated as a researcher in projects about rescue robotics and educational robotics. His research interests include physical human-robot interaction, tactile sensing, haptic perception, and education.



Alfonso J. García-Cerezo received the Industrial Electrical Engineer and the Doctoral Engineer degrees from the Escuela Técnica Superior de Ingenieros Industriales of Vigo in 1983 and 1987, respectively.

From 1983 to 1988, he was an Associate Professor with the Department of Electrical Engineering, Computers and Systems, University of Santiago de Compostela, where he was an Assistant Professor, from 1988 to 1991. Since 1992, he has been a Professor of System Engineering and Automation

with the University of Malaga, Spain. From 1993 to 2004, he was the Director of the Escuela Técnica Superior de Ingenieros Industriales de Malaga. He is currently the Head of the Department, and the responsible of the Instituto de Automática Avanzada y Robótica de Andalucía, Málaga. He has authored or coauthored about 150 journal articles, conference papers, book chapters, and technical reports. His current research interests include mobile robots and autonomous vehicles, surgical robotics, mechatronics, and intelligent control.



Jesús M. Gómez-de-Gabriel received the Engineering and Ph.D. degrees in computer science from the University of Málaga, Málaga, Spain, in 1990 and 1999, respectively.

He is currently an Associate Professor with the Department of Systems Engineering and Automation, University of Málaga. He has led a project on tele-surgery and has participated in medical robotics projects and other projects related to mobile robots and factory automation. His current research interests include physical human-robot

interaction (pHRI), Internet-of-Things, education in engineering, and mobile robotics.