Deep n-Shot Transfer Learning for Tactile Material Classification with a Flexible Pressure-Sensitive Skin

Berthold Bäuml   Andreea Tulbure

Abstract— n-shot learning, i.e., learning a classifier from only few or even one training samples per class, is the ultimate goal in minimizing the cost of sample acquisition. This is esp. important for active sensing tasks like tactile material classification. Achieving high classification accuracy from only few samples is typically possible only when pre-knowledge is used. In n-shot transfer learning, knowledge from pre-training on a large knowledge set with many classes and samples per class has to be transferred to support the training for a given task set with only few samples per new class.

In this paper, we show for the first time that deep end-to-end transfer learning is feasible for tactile material classification. Based on the previously presented (TactNet-II) [1], a deep convolutional neural network (CNN) which reaches superhuman tactile classification performance, we adapt state-of-the art deep transfer learning methods.

We evaluate the resulting deep n-shot learning methods with a publicly available tactile material data set with 36 materials [1] in a 6-way n-shot learning task with 30 materials in the knowledge set. In 1-shot learning, our deep transfer learning method reaches 75.5% classification accuracy and in 10-shot more than 90%, outperforming classification without knowledge transfer by more than 40%. This results in an up to 15 time reduction in the number of samples needed to reach a desired accuracy level.

We also provide insights of the inner workings of the derived deep transfer learning methods.

I. INTRODUCTION

Tactile sensing, esp. on hands, is indispensable for robots to perform complex dextrous manipulation tasks or to delicately explore the physical properties of the environment. As the physical interactions between the sensor and surface while sliding over it are involved, usually learning has to be used to interpret the tactile data. Deep learning has shown very robust [2] and even superhuman performance [1] in tactile material classification. But learning typically needs a large number of training samples and in robotics, sample acquisition is costly, even more so for active sensing tasks like tactile sensing.

It is therefore desirable to need only as little training data as possible; in the extreme case, needing only one sample per to be learned class. This is the well known one-shot learning problem [3]. The generalized n-shot learning, where n (small number) samples per class are used, is a topic of intensive research in the machine learning community (see Sec. I-A).

The main idea in n-shot learning methods is to take usage of pre-knowledge to support the learning task at hand.
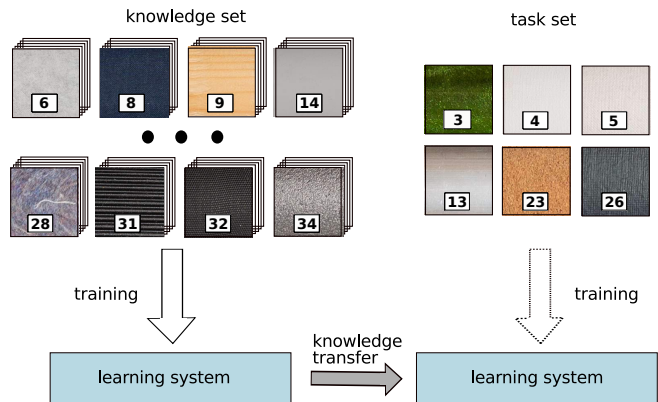
Fig. 1. n-shot transfer learning. From a large number of samples (the knowledge set) a classifier is trained once. The learned knowledge is then extracted and transferred to a new classifier (often of the same architectural type) to support the learning of a new classification task on previously unseen classes with only a few n samples per class. The depicted materials are a subset of the used publicly available data set from [1] of typical materials found in households: (3) jersey, (4) cotton fabric reverse, (5) cotton fabric, (6) linen fabric smooth reverse, (8) linen fabric smooth(9) wood, (12) reflecting fabric, (13) metal, (23) cork, (26) carton, (28) carpet rough, (31) rubber rough horizontal, (32) rubber smooth, (34) foam.

Typically, the pre-knowledge is acquired by training on a large data set once and is then transferred to the new task with previously unseen classes. Fig. 1 depicts such a transfer learning procedure.

In this paper, we investigate and show for the first time the feasibility of deep one-shot transfer learning for a real world tactile material classification task with a flexible pressure-sensitive skin (see Fig. 2 for the experimental setup). Deep one-shot learning means that learning is done end-to-end with no manually pre-constructed feature mapping but directly on the raw high-dimensional sensor signal.

A. Related Work

Tactile material classification and tactile sensing in general has gained increasing attention in the robotics community in the last years. Reviews of tactile sensors can be found , e.g., in Dahiya et al. [4] and Kappassov et al. [5]. Tulbure and Bäuml [1] give a thorough overview of research in tactile material classification.

To our knowledge, the only work on transfer learning in tactile material classification is [6]. They use the signals of the 2 kHz pressure sensor and the 19 electrodes (50 Hz sample rate) of the BioTac [7] sensors mounted on the fingertips of a Shadow Hand. The hand holds the objects while sliding over the surface with two fingers. From the raw high-dimensional signal, first lower dimensional features

are extracted using manually designed feature descriptors. Their online tactile learning algorithm which is based on a least squares support vector machine (SVM) classifier works in this lower dimensional feature space. For 1-shot learning with a knowledge set of 10 objects and a task set of 12 objects, they reach an excellent accuracy of 97%.

In contrast, the work we present in this paper is based on [1] and uses a flexible pressure-sensitive skin instead of the BioTac sensor. This is an important advantage as the flexible skin can be easily mounted (glued) to an existing robotic structure whereas for the bulky BioTac sensor parts, e.g., the fingertip, have to replaced. We also use the larger data set from [1] with 36 materials which, in addition, is publicly available and allows for comparison to other methods in the future.

However, the most important difference is that we do not use manually designed feature descriptors but use a deep learning based end-to-end learning scheme with automatic feature extraction directly from the 16000 dimensional raw input signal. End-to-end learning is important not only because it makes the time-consuming feature design by an expert obsolete but because the automatic feature extraction gives better performance for the flexible skin sensor as was shown in [2] (different feature descriptors and classification methods including SVMs have been compared to deep learning).

Recently, there has been much interest in n-shot learning in the context of deep learning which usually demands for large training set sizes. Early works [8] use feature extraction from a network trained on the large knowledge set and then use classification in this feature space. A more advanced matching variant is the Siamese network method [9] where twin networks are trained especially for getting a good feature space for later matching via a simple distance function, e.g., the euclidian distance.

Newer work use quite complex (recurrent) neural networks needing episodic training [10] [11] [12]. Those transfer learning networks are not easy to train and usually work only for rather simple base network architectures. In contrast, Bauer et al. [13] show in their recent work that, using a capable feed forward neural network trained on the knowledge set as a basis, simple matching and logistic regression based transfer methods work surprisingly well and outperform all the more complicated methods by a large margin on the usual benchmark tasks.

Therefore, in this work we use the high performing TactNet-II as the basis for the state of the art transfer learning methods from Bauer et al. [13].

II. ROBOT SETUP & n-SHOT LEARNING TASK

A. Setup

1) Robot Setup: Fig. 2 summarizes the experimental setup for recording the tactile data from a flexible tactile skin mounted on a robot's finger tip.

2) Material Data Set: For all the experiments we use the publicly available material data set from [1] ($K = 36$ materials and $m = 100$ samples per class) from everyday
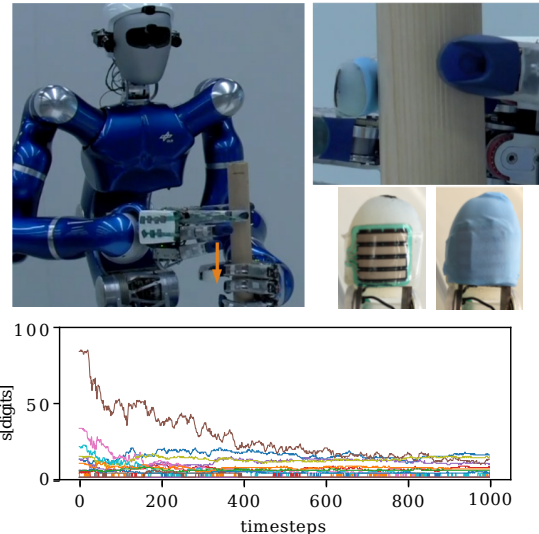

Fig. 2. DLR's *Agile Justin* [14] performing a sweeping motion to identify the material of a tube. For our "Mars Habitat" demonstrator the robot has to discriminate geometrically identical tubes (50 cm × 5 cm.) by touch alone. Agile Justin is equipped with two DLR Hand-II [15]. To the soft finger tip of the index finger of the right hand a flexible tactile skin with a 4 × 4 taxel array from Tekscan [16] (VersaTek® sensor 4256E) is attached providing a spatio-temporal pressure signal at 750 Hz sample rate. The right column depicts a close-up (upper) of the contact situation and how the tactile sensor is taped to the finger tip (lower) with a thin laboratory glove on top to increase the grip. The procedure for exploring a given tube is performed autonomously by the robot: grasp the tube with the left hand to stabilize it; grasp the tube with the thumb and the index finger of the right hand; slide down along the tube at a constant velocity of 3 cm/s with a roughly constant force of 1 N (20% precision) using the hand's joint torque sensors for 2 s. The plot at the bottom shows the resulting raw spatio-temporal signal of the tactile skin. For material classification, the 1.33 s sliding motion, resulting in a $1000 \times 4 \times 4 = 16000$ dimensional sample, is used without any pre-processing.

household objects which has been recorded with the setup described before (see Fig. 1 for a subset of the materials). Tulbure and Bäuml [1] also present human performance data for a classification and differentiation task from a large human study.

B. n-Shot Learning Task

Fig. 1 describes n-shot learning. The knowledge set $\widetilde{\mathcal{D}} = \{\widetilde{x}_i, \widetilde{y}_i\}_{i=1}^{\widetilde{N}}$ has many classes $\widetilde{C}$ and a large number of samples $\widetilde{n}$ per class, i.e., $\widetilde{N} = \widetilde{C}\widetilde{n}$. If the task set $\mathcal{D} = \{x_i, y_i\}_{i=1}^{N}$ consists of C classes with only few samples n per class, i.e., $N = Cn$, the problem is called a C-way n-shot learning problem. The goal in n-shot transfer learning is to reach high accuracy on test samples from the task classes although the task set for training has only few samples. The trick is to transfer knowledge from the large knowledge set (with different classes).

III. n-SHOT TRANSFER LEARNING METHODS

A. Fine-Tuning

As a typical CNN, our TactNet-II architecture (Fig. 3 a) consists of two parts: the automatic hierarchical feature extraction layers and the classification perceptron. The first part extracts important features from the input in a hierarchical fashion from low (elementary) to high semantic

level [17, 18], whereas the second part expands these features into a high dimensional space to fit a classification hyperplane.

The idea is, inspired by Hoffman et al. [8], that the hierarchical features learned are generic when the number of training samples and classes is large enough. So, this features can be learned once from a large knowledge set and then be reused. For a new task, the input samples x are first transformed into the feature space, $u = \phi(x; \varphi)$, with the parameters φ previously learned from the knowledge set. Then, only the classification in the abstract feature space has to be learned from the task set.

A nice property of CNNs is how such an idea directly leads to a concrete implementation: given a C-way n-shot problem, the knowledge set is first used to train a CNN. The trained weights (layers) are then extracted and saved. For training the network on the task set with C new classes, the trained weights are loaded into the network and only a subset of the layers are re-trained, typically a subset of the last layers (classifier perceptron), using n samples per class.

Choosing the subset of retrainable layers results in different variants of n-shot transfer learning CNNs.

OL: All layers except for the output layer (OL) are pre-trained on the knowledge set and then the weights are fixed. Only the OL is then trained on the task set.

FCL-OL: Like in OL, all layers are pre-trained on the knowledge set but this time the OL as well as the fully connected layer (FCL) are retrained on the task set.

NK: As a baseline, also the no-pre-knowledge (NK) method fits into this scheme as an extreme case. In NK, the original TactNet-II is directly trained on the task set, i.e., without any pre-training on the knowledge set. We use the knowledge set only for an individual hyperparameter search for each n of the n-shot learning tasks.

Fig. 3 a gives a graphical summary of these learning variants.

B. Matching

Another method for reusing the learned feature mapping $u = \phi(x; \varphi)$, hence, the knowledge set, is to directly perform a nearest neighbor (NN) classification without any learning on the task set. The idea is, that a simple distance metric in the abstract feature space can cover the class structure better than using the simple metric in the original sample space.

For 1-NN, given the task set $\mathcal{D} = \{x_i, y_i\}_{i=1}^{N}$ and a distance metric $d(u_1, u_2)$ in the feature space, a sample x is classified as y^* according to

$$i^* = \arg\min_i d\left(\phi(x; \varphi), \phi(x_i; \varphi)\right), \quad y^* = y_{i^*}^*.$$

We use the feature mapping of TactNet-II without the output layer and the simple euclidian metric $d(u_1, u_2) = \|u_1 - u_2\|_2$.

Fig. 3 b depicts this *NN* method as a network structure, where two identical shortened TactNet-IIs are fed in parallel with x and the x_i.

An advanced variant of this matching idea is the *Siamese* network [9]. In this method, also the training on the knowledge set for getting the feature mapping is done in the
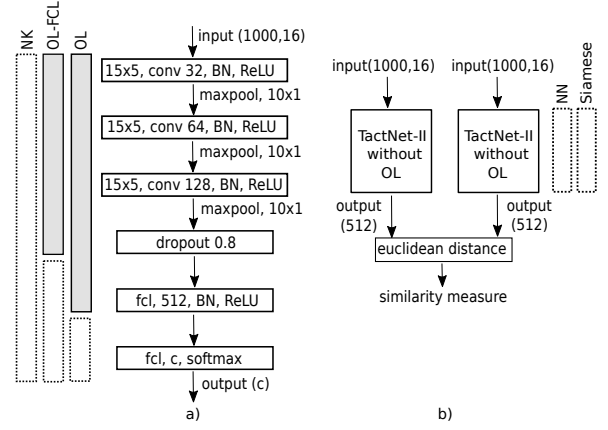

Fig. 3. Deep transfer learning network architectures. a) The TactNet-II CNN [1]. In transfer learning, first the full TactNet-II is trained on the knowledge set. For the n-shot learning only, parts of the network are retrained. The boxes on the left indicate for the fine-tuning methods OL and OL-FCL which layers are kept fixed (gray) and which are re-trained (empty). As baseline, the NK method without knowledge transfer retrains the full network. b) Matching network architecture. During testing, a pair of a test sample and a sample from the task set are mapped in parallel to the feature space by two trained identical TactNet-IIs. The Euclidean distance is then computed in this feature space. For the NN method, a standard pre-trained TactNet-II is used whereas the Siamese method trains the networks explicitly for discrimination between two given classes.

matching setup with the two identical twin networks (hence, the name Siamese). The idea is, that this way the mapping learns better features for the matching than when trained for a classification task.

For training on the knowledge set, the network is fed with pairs of training samples $(\tilde{x}_i, \tilde{y}_i), (\tilde{x}_j, \tilde{y}_j)$ and it has to predict if they belong to the same ($\tilde{y}_i = \tilde{y}_j$) or different ($\tilde{y}_i \neq \tilde{y}_j$) classes. This is a binary classification problem with label $t = 0$ (same) or $t = 1$ (different) and the contrastive loss function is used. For a single training pair the loss is

$$E(\varphi) = (1 - t)d^2 + t \max(0, \epsilon - d)^2,$$

with $d = d\left(\phi(\tilde{x}_i; \varphi), \phi(\tilde{x}_j; \varphi)\right), \ \epsilon > 0.$

For training, $N_{\text{Siam}}/2$ pairs with same and $N_{\text{Siam}}/2$ pairs with different class labels are randomly drawn from the knowledge set.

At test time on the task set, the same scheme as in NN is used. Fig. 3 b shows the network structure.

C. Concept Learning

In all above methods, the transfer of knowledge is performed solely via the learned mapping into the feature space. Bauer et al. [13] describe a simple probabilistic method that additionally transfers knowledge from the learned weights of the output layer in a OL fine-tuning setting like in Sec. III-A.

Given the weights W of the C class neurons in the output layer, which have to be trained from the transformed task set samples $u = \phi(x; \varphi)$, with $x \in \mathcal{D}$ and $\widetilde{W}$ the weights of the $\widetilde{C}$ class neurons that have been trained on the knowledge set $\widetilde{\mathcal{D}}$. The full Bayesian approach in Bauer et al. [13] using the assumptions of a large number of classes $\widetilde{C}$ and samples per class $\tilde{n}$ in the knowledge set and a small number of samples per class n in the task set, leads then to the following maximum a posteriori (MAP) point estimate for

the predictive probability distribution for a new sample of a task class

$$p(y|x, \mathcal{D}, \widetilde{\mathcal{D}}) \approx p(y|x, W_{\mathrm{MAP}}),$$

with $W_{\mathrm{MAP}} = \arg\max\limits_{W} p(\mathcal{D}|W)p(W|\widetilde{W})$.

This is the same maximization expression as in standard MAP estimation but with a prior $p(W|\widetilde{W})$ for the weights W that depends on the weights $\widetilde{W}$ trained on the knowledge set.

With $W = (w_1, \ldots, w_C)$ are the weights of the C task class neurons and $\widetilde{W} = (\widetilde{w}_1, \ldots, \widetilde{w}_{\widetilde{C}})$ of the $\widetilde{C}$ knowledge class neurons and the probabilistic model assumption that the prior is a univariate Gaussian distribution, the prior can be written as

$$p(W|\widetilde{W}) = \prod_{c}^{C} \mathcal{N}(w_c|\mu_{\widetilde{w}}, \sigma_{\widetilde{w}}^2),$$

$$\mu_{\widetilde{w}} = \frac{1}{\widetilde{C}}\sum_{c}^{\widetilde{C}} \widetilde{w}_c, \quad \sigma_{\widetilde{w}}^2 = \frac{1}{\widetilde{C}}\sum_{c}^{\widetilde{C}} \frac{1}{F}(\widetilde{w}_c - \mu_{\widetilde{w}})^{\mathrm{T}}(\widetilde{w}_c - \mu_{\widetilde{w}}).$$

$\mu_{\widetilde{w}}$ and $\sigma_{\widetilde{w}}^2$ are the empirical mean and variance of the trained weights of the knowledge set class neurons and F is the dimensionality of the feature space (i.e., $F = \dim(u)$). Important to note is that the prior $p(W|\widetilde{W})$ is not centered at 0 but shifted by $\mu_{\widetilde{w}}$. In summary, the concept learning method of Bauer et al. [13] biases the weights of the task class neurons to be close to the mean of the knowledge class neuron weights and the strength of this bias is inversely proportional to the variance of the knowledge class neuron weights.

In the error function, the prior results in a L_2 regularization term but with the calculated optimal strength and the shift

$$E(W) = -\log p(\mathcal{D}|W) + s\sum_{c}^{C}(w_c - m)^{\mathrm{T}}(w_c - m),$$

$$s = 1/(2\sigma_{\widetilde{w}}^2), \quad m = \mu_{\widetilde{w}}.$$

We call the method using the full regularization term *OL-mean*. When only the calculated optimal regularization strength s but not the shift term ($m = 0$) is used, we call it *OL-opt*. The method OL from Sec III-A is the case where the strength s is set to the default value as used in the TactNet-II training.

IV. EVALUATION & DISCUSSION

A. Cross-Validation

To evaluate the n-shot learning performance of a classifier with the material data set from Sec. II-A with $K = 36$ classes, we use a double cross-validation scheme. On the class level, we split the $K = 36$ materials randomly into $C = 6$ task set classes and the $\widetilde{C} = 30$ knowledge set classes, i.e., we perform 6-way n-shot learning. If not mentioned otherwise, all results are reported for averaging over $S = 100$ random splits (but the same random splits are used for all learning methods).
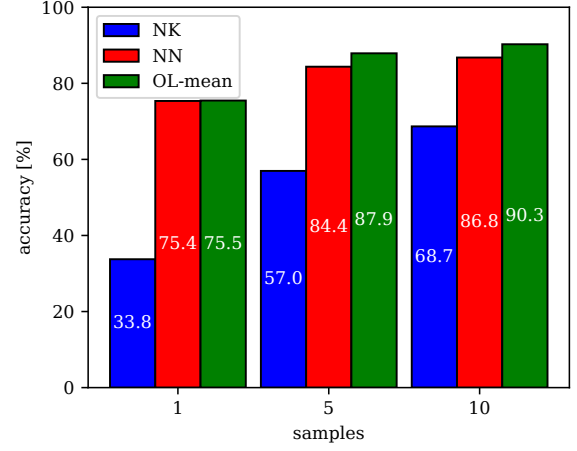

Fig. 4. Performance comparison of the best transfer learning methods and the NK method without pre-knowledge.

For each split, the knowledge training (and potentially hyperparameter training) is performed using all $\widetilde{n} = m = 100$ samples per class of the knowledge classes. For the n-shot training, the task set contains only n samples for each task class. But we use cross-validation on the sample level where for $R = 10$ runs, randomly n new samples out of the $m = 100$ samples for each of the C classes in the task set are chosen and all other $m - n$ samples per class are used for testing. The accuracy is averaged of all runs.

TABLE I

N-SHOT TRANSFER PERFORMANCE

method	$n = 1$	$n = 5$	$n = 10$
NK	33.8	57.0	68.7
Siamese	66.7	67	75.6
NN	75.4	84.4	86.8
OL-FCL	67.0	83.8	88.8
OL	75.0	87.5	89.9
OL-opt	75.1	87.6	90.2
OL-mean	**75.5**	**87.9**	**90.3**

B. Transfer Learning Performance

Table I lists the accuracy for all methods from Sec. III and Fig. IV-A compares the most important results.

1) Method Comparison: Using no pre-knowledge (NK method) performs poorly when using 1-shot learning and even for 10-shot learning the accuracy is with $< 70\%$ unusably low. But the n-shot transfer learning methods work very well and reach accuracies of $> 75\%$ even for 1-shot learning and $> 90\%$ for 10-shot learning. Compared to using no pre-knowledge, the accuracy gain is as high as 40%.

For 1-shot learning, the fine-tuning as well as the matching based NN methods perform comparably well but for large n the accuracy of the NN method is almost 4% lower than the best OL methods. This shows that learning of a shallow classifier in the feature space, although on only few examples, still covers better the class structure of the tactile data.

2) Matching Methods: Other than expected, the Siamese network performs $> 10\%$ worse than the NN method although it explicitly trains a feature mapping for pairwise
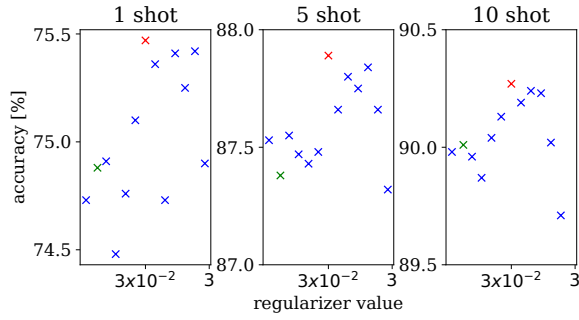
Fig. 5. Dependence of the accuracy on the regularization strength for OL-mean. The red cross marks the calculated strength, the green cross marks the standard strength of TactNet-II as it is found by hyperparameter search for training on the knowledge set. The calculated strength is always optimal proving that the probabilistic model assumption of concept learning are valid for TactNet-II.

matching via a distance function. We suspect that this is due to the inefficiency of the needed training with pairs of samples which would lead to $(\widetilde{C}\widetilde{n})^2 = 9 \cdot 10^6$ of pairs to present all information of the knowledge set to the network. But for performance reasons we can only provide 27000 pairs per epoch (a usual ratio in training Siamese networks).

3) Fine Tuning Methods: In the fine-tuning methods, the OL-FCL method performs worst, as expected, due to the high number of weights that have to be trained from only few task set samples. For 1-shot learning, the full concept learning method OL-mean with the calculated optimal regularization strength and shift term performs 0.5% better than OL with no shift term and the default regularization strength of TactNet-II (found by hyperparameter search on the knowledge set).

Fig. IV-B.3 shows for OL-mean that the calculated regularization strength is actually optimal. This proves that the assumptions of the probabilistic concept learning model are valid in case of the tactile material classification task and the TactNet-II network architecture.

C. Analysis

1) Easy/Hard Material Combinations: For some combinations of materials, the classification is harder than for others. Therefore, Table II lists the individual accuracies for the hardest and the easiest combinations out of the $S = 100$ material splits we evaluated. In 1-shot learning, the difference in the accuracies is $> 40\%$ for OL-mean but even for the hardest combination, transfer learning performs 30% better than when using no pre-knowledge.

2) Uncertainty: Table II also shows the standard deviation σ_R over the $R = 10$ cross-validation runs for the two extreme material combinations and the average $\bar{\sigma}_R = 1/R \sum_{r=1}^{R} \sigma_R^{(r)}$. The σ_R represent the uncertainty due to different sets of n randomly chosen samples out of the m samples for each material in a given task set. As expected, the uncertainty is highest for 1-shot learning ($\approx 5\%$). But with increasing n, $\bar{\sigma}_R$ rapidly shrinks to about 1% for OL-mean but stays at $\approx 3\%$ for NK. This shows that the knowledge transfer not only increases the accuracy but also decreases the uncertainty.

3) Feature Mapping: All the transfer learning methods from Sec. III depend on the assumption that TactNet-II learns a good feature mapping $u = \phi(x; \varphi)$ when trained
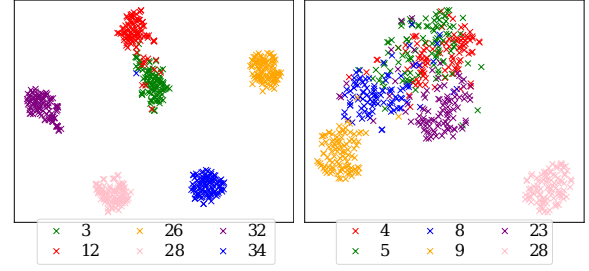

Fig. 6. t-SNE of all $m = 100$ samples per material for the easy and hard combination. Images of the materials are depicted in Fig. 1. *Left*: Easy combination with (3) jersey, (12) reflecting fabric, (26) carton, (28) carpet rough, (32) rubber smooth, (34) foam. *Right*: Hard combination with (4) cotton fabric reverse, (5) cotton fabric, (8) linen fabric smooth, (9) wood, (23) cork.

on the knowledge set so that classification in the feature space is easier than directly in the input space. To check this assumption, we performed t-SNE [19] on the mapped input samples of the two extreme material combinations. Fig. IV-C.3 shows that for the easy combination the classes are clearly separated and even for the hard combination only two classes get heavily mixed.

TABLE II
EASY/HARD MATERIAL COMBINATIONS

task	method	$n = 1$		$n = 5$		$n = 10$	
		acc	σ_R	acc	σ_R	acc	σ_R
hard	OL-mean	56.0	5.5	75.1	2.9	80.8	2.1
	NK	26.3	4.3	43.8	3.5	53.4	5.2
easy	OL-mean	90.4	4.7	96.8	0.6	97.3	0.4
	NK	43.7	4.3	77.0	3.6	89.8	2.6
mean	OL-mean	75.5	5.5	87.9	1.8	90.3	1.3
	NK	33.8	5.3	57.0	5.1	68.7	3.4

D. Sample Efficiency

In robotics, the usefulness of transfer learning depends on the reduction in the number of samples needed to reach a desired accuracy level. For this, Fig. IV-D compares the accuracies of transfer learning (OL-mean) and learning without pre-knowledge (NK) for $n = 1, \dots, 80$. For 1-shot learning, NK reaches the same level only for 15 times more samples. In case, an accuracy of $> 90\%$ is desired, 10-shot transfer learning still needs about 5 times less samples than NK.

V. CONCLUSIONS

In this paper, we have shown for the first time that deep end-to-end learning for tactile material classification is feasible. By adapting state of the art deep transfer learning methods to our previously reported TactNet-II [1] with superhuman performance, we reached classification accuracy in this real world setting of 75.5% for 6-way 1-shot learning and over 90% for 10-shot learning. This leads to a 15 resp. 5 time reduction in the number of samples needed compared to learning without knowledge transfer.

This high classification accuracy also shows that the knowledge set with 30 materials is large enough to cover relevant knowledge for all 100 tested task set material combinations.
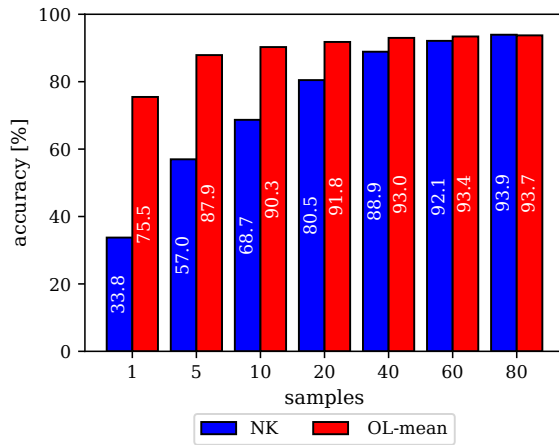
Fig. 7. Performance comparison of n-shot transfer learning OL-mean and learning without prior knowledge NK.

Our analysis shows that the probabilistic concept learning method [13] which was previously only used in a vision benchmark example is applicable to the real world material classification task. By using the TactNet-II, it outperforms all other tested deep n-shot transfer learning methods for all n.

In future work, we will further examine methods for sample efficient training in tactile material classification, namely semi-supervised learning and guided active learning.

ACKNOWLEDGMENTS

REFERENCES

[1] A. Tulbure and B. Bäuml, "Superhuman performance in tactile material classification and differentiation with a flexible pressure-sensitive skin," in *Proc. IEEE/RAS International Conference on Humanoid Robots*, 2018 (submitted). [Online]. Available: https://drive.google.com/file/d/1ybR2Bk5UN_J1R-vrcyARC67qKrBYQk7J/view?usp=sharing

[2] S. Baishya and B. Bäuml, "Robust material classification with a tactile skin using deep learning," in *Proc. IEEE International Conference on Intelligent Robots and Systems*, 2016.

[3] L. Fei-Fei, R. Fergus, and P. Perona, "One-shot learning of object categories," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 594–611, 2006.

[4] R. S. Dahiya *et al.*, "Directions toward effective utilization of tactile skin: A review," *IEEE Sensors Journal*, vol. 13, no. 11, 2013.

[5] Z. Kappassov, J. A. C. Ramon, and V. Perdereau, "Tactile sensing in dexterous robot hands - review," *Robotics and Autonomous Systems*, vol. 74, no. Part A, pp. 195–220, 2015.

[6] M. Kaboli, R. Walker, and G. Cheng, "Re-using prior tactile experience by robotic hands to discriminate in-hand objects via texture properties," in *Proc. IEEE International Conference on Robotics and Automation*, 2016.

[7] N. Wettels, V. Santos, R. Johansson, and G. Loeb, "Biomimetic tactile sensor array," *Advanced Robotics*, vol. 22, no. 8, pp. 829–849, 2008.

[8] J. Hoffman *et al.*, "One-shot adaptation of supervised deep convolutional models," in *Proc. International Conference in Learning and Representation*, 2014.

[9] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," in *Proc. ICML Deep Learning workshop*, 2015.

[10] O. Vinyals *et al.*, "Matching networks for one shot learning," in *Advances in Neural Information Processing Systems 29*, D. D. Lee *et al.*, Eds. Curran Associates, Inc., 2016, pp. 3630–3638. [Online]. Available: http://papers.nips.cc/paper/6385-matching-networks-for-one-shot-learning.pdf

[11] J. Snell, K. Swersky, and R. S. Zemel, "Prototypical networks for few-shot learning," *CoRR*, vol. abs/1703.05175, 2017. [Online]. Available: http://arxiv.org/abs/1703.05175

[12] S. Ravi and H. Larochelle, "Optimization as a model for few-shot learning," in *Proc. International Conference on Learning Representations*, 2017.

[13] M. Bauer *et al.*, "Discriminative k-shot learning using probabilistic models," in *NIPS workshop on Bayesian Deep Learning*, 2017.

[14] B. Bäuml *et al.*, "Agile Justin: An upgraded member of DLR's family of lightweight and torque controlled humanoids," in *Proc. IEEE International Conference on Robotics and Automation*, 2014.

[15] J. Butterfaß, M. Grebenstein, H. Liu, and G. Hirzinger, "DLR-Hand II: Next generation of a dextrous robot hand," in *Proc. IEEE International Conference on Robotics and Automation*, 2001, pp. 109–114.

[16] Tekscan. [Online]. Available: https://www.tekscan.com

[17] Y. Xu *et al.*, "Deep learning of feature representation with multiple instance learning for medical image analysis," in *Proc. IEEE Int. Conf on Acoustics, Speech and Signal Processing*, 2014.

[18] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. NIPS*, 2012.

[19] L. van der Maaten and G. Hinton, "Visualizing high-dimensional data using t-sne," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.