Cleaning tasks knowledge transfer between heterogeneous robots: a deep learning approach

Jaeseok Kim[1], Nino Cauli[2], Pedro Vicente[2], Bruno Damas[2,3], Alexandre Bernardino[2], José Santos-Victor[2] and Filippo Cavallo[1]

***Abstract*—This is an extended abstract of a paper with the same name published in the Journal of Intelligent & Robotic Systems, Springer, on August 2019 [1]. In this paper, a robot is taught to perform two different cleaning tasks over a table, using a learning from demonstration paradigm. Robustness to robot posture and illumination changes is achieved using data augmentation techniques and camera images transformation. This robustness allows the transfer of knowledge regarding execution of cleaning tasks between heterogeneous robots operating in different environmental settings. To demonstrate the viability of the proposed approach, a CNN network trained in Lisbon to perform cleaning tasks, using the iCub robot, is successfully employed by the DoRo robot in Peccioli, Italy.**

***Index Terms*—Learning from demonstration, Transfer learning, Data augmentation, Convolutional neural networks, Task parametrized Gaussian mixture models**

I. Introduction

In order to adapt to unknown environment and acquire new skills, cleaning robots should be able to learn from past and new experience. In Learning from Demonstration (LfD) algorithms robot cleaning skills are derived from observations of human demonstrations and generalized to new environments. Dynamic Movement Primitives (DMPs) and Gaussian Mixture Models (GMM) are typically used to encode action in LfD.

Calinon *et al.* [2] proposed the Task-Parameterized Gaussian mixture model (TP-GMM), a technique to generalize trajectories from demonstrated ones using task parameters (frames). While several TP-GMM systems have been successfully used to generate robotic cleaning motions, none of them is able to autonomously learn the task parameters from raw images.
One powerful solution to extract information from raw pixel data and learn important features on the images are Convolutional neural networks (CNNs). Pervez *et al.* [3] proposed to use a CNN to learn the parameters of a TP-DMP directly from camera images, calling the system Deep-DMP (D-DMP). D-DMP was used to swipe different objects from a table.

In a recent work of our [4], we used a similar approach to learn the parameters of a TP-GMM to control a robot performing sweeping and wiping movements while cleaning a table. In this paper, we extend the works presented in [4] and [5] using a CNN/TP-GMM system, trained on a dataset collected on the iCub robot in Lisbon-Portugal, to control the Domestic Robot (DoRo) in Peccioli-Italy while cleaning a table. The main contributions of this paper are:

1) **Demonstration of successful transferring of knowledge from a robot to another**
2) **Finding an optimal number of demonstrations needed to learn a cleaning motion**
3) **Proving the importance of domain randomization in our scenario**

II. Proposed approach

The goal of this paper is to transfer the knowledge acquired by the iCub robot in Lisbon, during a kinesthetic demonstration of a cleaning task, to the DoRo robot in Peccioli. Two different cleaning movements are taught to the iCub in order to clean a table: a sweeping motion to remove lentils from the table and a wiping motion to clean marker scribbles. In order to generalize to different robot camera positions and table heights, camera images are transformed to a canonical virtual image plane, similarly to what has been done in [5]. The canonical virtual camera is placed at a fix distance from the table, right on top of it, generating a bird-view image. Specific sizes and positions of objects placed on the table correspond to particular sizes and positions in the virtual image plane.

From the virtual images, the robot estimates the correct cleaning hand trajectories using the same architecture introduced in [4]: a CNN estimate the initial, intermediate and final positions of the desired trajectory used to create the parameters of the TP-GMM and GMR algorithm is used to estimate the desired trajectory from the TP-GMM. We analytically calculate the reference frames orientations from the reference frame positions predicted by the CNN. To collect the dataset for training we placed the iCub robot in front of a white table of size 50x50 cm. For each demonstration some dirt was placed on the table (lentils clusters or marker scribbles). A human guided the iCub right hand cleaning as much as possible of the dirt spot with a specific motion for each dirt type. This 659 demonstrations dataset was then augmented resulting in a bigger dataset.

III. Results

The evaluation of the cleaning task is defined according to the different type of dirt presented on the environment. For the

[1]BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy. j.kim @sssup.it, filippo.cavallo @santannapisa.it

[2]Institute for Systems and Robotics, Instituto Superior Tecnico, Universidade de Lisboa, Portugal {ncauli,pvicente,bdamas,alex, jasv} @isr.tecnico.ulisboa.pt

[3]CINAV — Centro de Investigação Naval, Almada, Portugal

This work was partially supported by Fundação para a Ciência e a Tecnologia (project UID/EEA/50009/2013 and Grant PD/BD/135115/2017) and the RBCog-Lab research infrastructure. We acknowledge the support of NVIDIA Corporation with the donation of the GPU used for this research.

	Marker		Lentils	
	Area cleaned	Standard Deviation	Distance reduced	Standard Deviation
Cauli et al. (iCub) [5]	80%	15%	45%	2%
Our results (DoRo)	75%	20%	50%	10%

TABLE I: Comparison between the previous results of [5] on the iCub robot and our results on the DoRo robot. The test scenario is the same on both. The two systems have a different network architecture and a different data augmentation strategy.

	80% (527 original samples)			**20%** (132 original samples)		
	O	**T**	**P**	**O**	**T**	**P**
Effective Training Samples	527	5797	11067	132	1452	2772
Train Loss [$\times 10^{-3}$]	**2.70**	2.96	3.04	**1.29**	**2.09**	**2.81**
Val Loss [$\times 10^{-3}$]	**2.65**	1.65	1.94	**27.54**	**2.76**	**3.15**

TABLE II: Loss after training the Network for 30000 iterations. The dataset is composed with 80% and 20% of the initial dataset (527 and 132 samples, respectively) to train the Network using 3 types of data. (**O**: Original; **T**: Translation and illumination and **O**; **P**: Perlin noise and **T**.

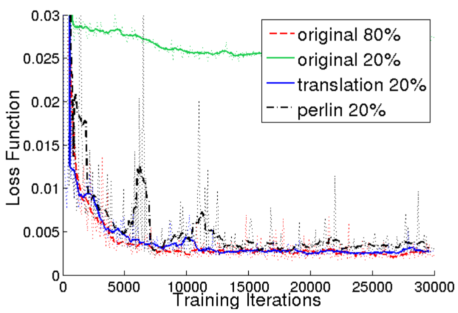

Fig. 1: Validation loss during training on 4 different dataset combinations. (**O80%**, **O20%**, **T20%**, and **P20%**)

case of marker scribbles, the percentage of dirty area after each repetition was calculated. For the lentils case, the performance is the reduction in distance of the weighted centroid of the dirt region from the bottom right corner of the table (expressed in percentage from the initial position).

Network tests: To access the performance of the Network according to the data present in the training set, we run the Network several times with different types of data augmentation and with different amounts of initial kinesthetic teaching examples. We have created 12 (different) training sets combining four (4) percentages of the original dataset with three (3) data types (**O**, **T** and **P**). The performance of the Network on the 80% and 20% validation sets taking into consideration the augmentation strategy performed can be seen in Table II. Figure 1 shows the validation loss during the training of 4 different dataset combinations. The Perlin noise will be essential when generalizing to another background (on the DoRo robot). After this evaluation, we conclude that **T20%** and **P20%** are suitable to test on the real robot and are a good trade-off between number of kinesthetic teaching and accuracy achieved. It is worth to notice that only 132 original examples are needed to learn 2 specific cleaning movements. Therefore, we assume achievable an open-ended learning scenario, where new kinesthetic demonstrations are taught to the robot online.

Robot experiments: The proposed architecture was tested on a real scenario using the DoRo robot to determine the transferring capabilities of the cleaning system to a different robotic platform. The robot should try to clean the dirty table (with marker scribbles or cluster of lentils) using a budget of five (5) repetitions. The DoRo robot performed 15 cleaning experiments on marker scribbles setting the table at 3 different heights. We reduced the dirt in 75% of its initial area with a standard deviation of 20%. In the lentils case, the table was set at the same 3 different heights and the robot performed 15 different experiments. The percentage of the initial distance from the bottom right corner of the table (the target point when cleaning this type of dirt) was reduced in 50% with a standard deviation of 10%. Table I shows the comparison between the results obtained on the DoRo and the results of [5] obtained on the iCub. The results on the DoRo are close to the results obtained on the iCub, showing how a system trained on one robot can be used to control a second one.

IV. Conclusions and Future work

We presented a framework for learning how to perform a given cleaning task from human kinesthetic demonstrations, directly from raw camera images, and later transferring the knowledge gathered in this process to a different robot. The use of a virtual camera and data augmentation strategies reduced the need for a large training set (only 20% of the recorded data was needed). A future direction could be the extension of our model to handle online open-ended learning scenarios. This can be done in two ways: 1) updating the training set with new kinesthetic demonstrations when needed and retrain the model on it; 2) using RL [6] to learn proper cleaning trajectories online based on a proper reward function.

References

[1] J. Kim, N. Cauli, P. Vicente, B. Damas, A. Bernardino, J. Santos-Victor, and F. Cavallo, “Cleaning tasks knowledge transfer between heterogeneous robots: a deep learning approach,” *Journal of Intelligent & Robotic Systems*, Aug 2019.

[2] S. Calinon, T. Alizadeh, and D. G. Caldwell, “On improving the extrapolation capability of task-parameterized movement models,” in *IEEE IROS*, 2013.

[3] A. Pervez, Y. Mao, and D. Lee, “Learning deep movement primitives using convolutional neural networks,” in *IEEE Humanoids*, 2017.

[4] J. Kim, N. Cauli, P. Vicente, B. Damas, F. Cavallo, and J. Santos-Victor, “icub, clean the table! a robot learning from demonstration approach using deep neural networks,” in *IEEE ICARSC*, 2018.

[5] N. Cauli, P. Vicente, J. Kim, B. Damas, A. Bernardino, F. Cavallo, and J. Santos-Victor, “Autonomous table-cleaning from kinesthetic demonstrations using deep learning,” in *IEEE ICDL-EpiRob*, 2018.

[6] C. Devin, P. Abbeel, T. Darrell, and S. Levine, “Deep object-centric representations for generalizable robot learning,” in *IEEE ICRA*.