Transfer Learning in Assistive Robotics: From Human to Robot Domain

David Ada Adama, *Graduate Student Member, IEEE,* Ahmad Lotfi, *Senior Member, IEEE,*
Robert Ranson and Pedro Trindade

Abstract— **Transfer Learning (TL) aims to learn a problem from a source reference to improve on the performance achieved in a target reference. Recently, this concept has been applied in different domains, especially, when the data in the target is insufficient. TL can be applied across domains or across tasks. However, the challenges related to *what to transfer, how to transfer* and *when to transfer* create limitations in the realisation of this concept in day to day applications. To address the challenges, this paper presents an overview of the concept of TL and how it can be applied in human-robot interaction for assistive robots requiring to learn human tasks in Ambient Assisted Living environments. The differences in feature spaces between a human (source domain) and robot (target domain) makes it difficult for tasks to be directly learned by robots. To address the challenges of this task, we propose a model for learning across feature spaces by mapping the features in the source domain to the target domain features.**

I. INTRODUCTION

Assisted living environments are incorporated with different technological solutions to improve the quality of life and wellbeing. In recent years, there has been a growing interest in the research community on how to develop evolving solutions to aid assisted living. Different techniques have been studied to address the need for technological systems which are intelligent enough to evolve their knowledge to solve task which have not been previously encountered. One of such approach is Transfer Learning (TL), for example, between humans and robots.

TL in computational intelligence is a branch of learning adopted from a concept in psychology concept 'transfer of learning'. TL involves developing computational models which are capable of mimicking humans' ability to learn and reuse knowledge in different but related tasks. For example, the knowledge acquired while learning to eat with a spoon can be applied in learning to use 'chopsticks'. This knowledge is transferred across related tasks. Traditional machine learning techniques work under the assumption that both source and target data are drawn from a similar distribution of information or similar data domains. This assumption holds in situations where the machine learning model is applied in classification of data which occur in both source and target information. However, in situations when source and target data are drawn from different information distribution, the traditional machine learning techniques struggle to correctly identify the target data [1]. This poses a limitation to machine learning techniques being used in such situations [1]. To address the limitation of traditional machine learning techniques, TL models seek to apply knowledge learned from a previous/source information

to a new, but related target information to improve the performance achieved and to reduce time needed in training the model from scratch [2].

Practical implementations of TL aim to transfer as much knowledge from a source task or domain over to the target task or domain. The knowledge transferred varies depending on the application and data from the source available. According to authors in [3], the key challenge in TL is defining the metrics related to *what to transfer*, *how to transfer* and *when to transfer*. This is mainly due to the fact that there are various algorithms that can be applied in TL. In trying to solve this challenge, TL algorithms used so far have focused on three main steps namely.

- Given a target task, select an appropriate source task or sets of tasks from which to transfer knowledge.

- Learn the relationship between a target task and source task, and.

- Transfer knowledge effectively from source task(s) to target task.

These steps have been used by researchers in proposing TL models [3] [4] .

TL has recently attracted interest in recent years due to the potential benefits it offers in artificial intelligence applications including assisted living [1], computer vision [5] and robotics [6]. It has not recorded as much success as the long existing traditional machine learning methods partly due to the challenges which yet remain unresolved in the research community [2]. It has potential to become a fundamental driver for the success of machine learning in the coming years [7] [8].

In relation to assisted living, different applications of TL have been studied. The authors in [1] proposed a model called Fuzzy TL which was applied in an intelligent environment. Data from the source domain was learned by constructing a fuzzy inference system from generated fuzzy rules. The constructed fuzzy inference system is then applied to a new domain referred to as the target domain through stages of adaptation of the generated fuzzy rules with the target data. Results from the model tested on real datasets from two intelligent environments (source and target environments) which were different but related showed the model achieves better performance in the target with transfer of knowledge when compared to performance attained without transfer.

TL has been used in other applications. Of interest to this work is its application in assistive robotics. In [9], the authors proposed a method for improving robot learning manipulation

tasks from data obtained from the robot performing other tasks or from similar robot architectures. The method used in [9] is an attempt to address the challenge of *how to transfer* by considering two steps which include, dimensionality reduction of data obtained from the robot to a low dimensional space and manifold alignment of source and target robot dimensions through a transformation function. The work in [6] also follows a similar approach of finding *how to transfer* between multi-robots. Even though these works achieve impressive performances, the challenges of *what to transfer* and *when to transfer* prove to be difficult in transfer learning applications. Addressing these challenges require consideration of properties related to spatial and temporal occurrences of both source/target domains. In this paper, we consider the case of TL from human to robot domains in trying to address some of the challenges of TL.

Figure 1 shows an illustration of TL between a human and an assistive robot in which the robot learns to perform a similar task by extracting relevant properties from the activity source (a human). In this work, we aim to follow a similar approach to TL in the context of transfer of human activity between a human and a robot by: 1) identifying requirements for TL in applications using human/robot as source/target domains respectively, 2) propose a method to address the differences between both domains through a remapping of feature spaces. From the review of related works [1] [9], it is evident that once an optimal mapping between source and target domains is known, *what/when to transfer* would be achievable.

II. METHODOLOGY FOR TRANSFER LEARNING BY FEATURE SPACE REMAPPING

From the introduction to TL presented in Section I, consider a source domain D_S with a feature space F_S and a target domain D_T with a feature space F_T such that $D_S \neq D_T$. TL aims to learn a task in D_S and the knowledge acquired is used in solving a different but related task in D_T. An overview of the method proposed to address the challenges of TL discussed in this work for TL by a remapping of feature spaces between source and target domains is presented in Figure 2. Information from both domains is required as inputs from which the feature spaces are constructed. For a model applied in a domain to be effectively transferred to a different domain, the features related to both domains need to be studied. The proposed approach assumes transfer is achieved when an effective mapping of F_S is obtained in D_T.

In the case where $F_S = F_T$ there can be a direct mapping from source to target to achieve transfer. This case is a much simpler case of TL where the challenges of *what/when to transfer* can be addressed with less computational effort. However, for example in applications such as human-robot interaction where a robot is required to learn an action from a human, this can be activities like *pick and place objects* in assisted living environments. The human and robot are considered as the source and target domains respectively. The differences in both feature spaces makes it not feasible for a direct mapping of features across the robot/human domains. This work assumes the robot domain needed for transfer of knowledge differs in feature space from that of a human, that is, $F_S \neq F_T$ and therefore for TL across such domains we

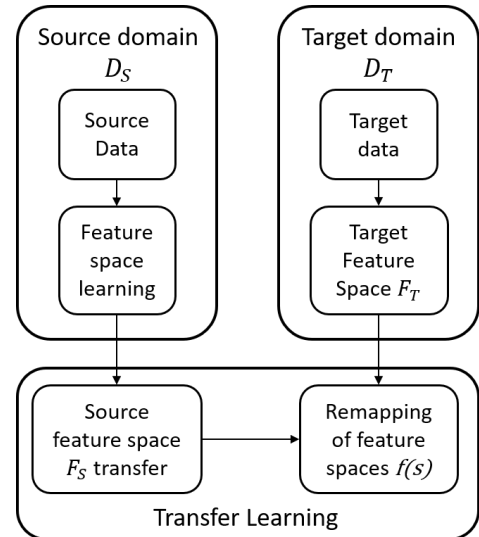*Figure 1. An illustration of Transfer Learning with an Assistive Robot.*

Figure 2. Transfer learning overview by a remapping of features in both source and target domains

---

Algorithm 1. TL by feature space remapping from source to target domains.

Input: Source domain feature space F_S and Target domain feature space F_T.
Output: Mapping function $f(s)$ from F_S to F_T.
1. Check and remove all duplicate features in F_S and F_T.
2. For every observation in the source domain D_S^i, a weight W_S^i is estimated for each feature – for $i > 0$.
3. Similarly, weights are constructed for the target features and represented by a matrix W_T.
4. For identical features in F_S and F_T, return corresponding weights W_S and W_T.
5. For the non-identical features in F_T, find correlation between weights W_S and W_T.
6. $f(s)$ is obtained by running a similarity function on weights W_S and W_T obtained, and a transformation of learned model to the target domain.

---

propose a remapping of feature space from source to target domains.

The proposed method for a remapping of feature spaces is summarised in Algorithm 1. The method requires both source and target domain feature spaces as inputs and the output obtained is a mapping function $f(s)$ which is a transformation of source features into relevant target features. Duplicate features within the feature spaces are discarded and weights W_S and W_T are assigned to features through a measure of feature importance in both domains. Identical features are extracted in a matrix while a method of correlation is applied to the weights of non-identical features to deduce a relationship between the features. Once this stage is completed, a mapping function is defined which is used in the transformation from F_S to F_T. It is worth noting that the proposed TL by feature remapping method is generalizable to different applications. This is possible if the feature spaces for transfer of knowledge are identified and not specific to an application or information distribution.

III. APPLICATION OF METHODOLOGY FOR TRANSFER LEARNING OF HUMAN ACTIVITIES FROM HUMAN TO ROBOT DOMAINS

The application of the method discussed in the previous section is applied in TL of human activities in an assisted living environment. A human performing an activity is assumed to be the source domain D_S with feature space F_S and an assistive robot needed in learning to perform a similar activity is assumed to be the target domain D_T with feature space F_T. The goal is to be able to learn an activity from a human and transfer the knowledge acquired to an assistive robot which would be capable of learning similar activity within an assisted living environment such as the example presented in the illustration in Figure 1.

Obtaining sufficient data from a robot to train a model for performing activities is a daunting task with a lot of complexities. However, sufficient data for a model to learn an activity can be obtained as humans perform activities and transferred to a robot. This would also enable assistive robots learn human activities by observing while a human performs activities. As shown in Figure 3, human activity data is obtained from visual cues as activities are performed. This could be information such 3D skeleton joint information, RGB or depth frames. The position and orientation features of joints of the human body are extracted. In addition, features of temporal occurrence, velocity, space and motion energy are formulated from the visual information of the activity performed. These features from D_S are used in a learning model for identifying the task performed within the activity.

For a robot to be able to learn to replicate a similar activity, it needs to understand the feature space of the activity source and how it can be transformed into its own space. The TL model requires the robot feature space F_T as input as well. This feature space can be in the form of joint positions and orientations, forward or inverse kinematics of the robot being used. The process of finding the mapping function $f(s)$ for the transformation of F_S to F_T following the algorithm described in Algorithm 1 is followed. The features learned from this step are used by the assistive robot in performing the learned activity.
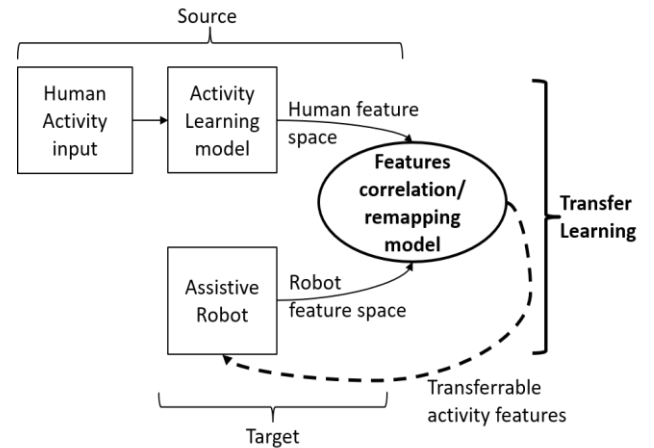

Figure 3. Human activity TL from human to robot domains by features remapping

IV. DISCUSSION AND CONCLUSION

In this paper, the concept of TL and its applications in an assisted living environment is discussed with a proposed application in assistive robotics - which is increasingly being incorporated in assisted living environments and explored by the software industry to provide meaningful services to the end-users. Although we propose a model for TL by learning the relationship between source and target features which is the basis of the knowledge being transferred, the key challenges of TL mentioned in Section I leave a gap in the ability to adequately model this concept. More applications of TL in technology could be a potential driver of the next revolution in artificial intelligence as seen from the benefits discussed earlier in this work which should not be underestimated. More research needs to be done on improving TL models to generalize across various applications. For future work, experiments will be carried out to validate the concept of TL and how it can successfully be applied to assistive robots. We also plan to investigate algorithms such as deep neural networks and domain adaptation methods in addressing the challenge of *what to transfer*.

V. REFERENCES

[1] J. Shell and S. Coupland, "Fuzzy Transfer Learning: Methodology and application," *Information Sciences,* vol. 293, pp. 59-79, 2015.

[2] D. J. Cook and K. D. Feuz, "Transfer Learning across Feature-Rich Heterogeneous Feature Spaces via Feature-Space Remapping (FSR)," *ACM Transactions on Intelligent Systems and Technology,* vol. 6, no. 1, p. 27, 2015.

[3] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering,* vol. 22, no. 10, pp. 1345-1359, 2010.

[4] D. Cook, K. D. Feuz and N. C. Krishnan, "Transfer learning for activity recognition: a survey," *Knowledge Information Systems,* vol. 36, no. 3, pp. 537-556, 2013.

[5] MathWorks, "Transfer Learning Using AlexNet," MathWorks Inc, 2018. [Online]. Available: https://www.mathworks.com/help/deeplearning/examples/transfer-learning-using-alexnet.html. [Accessed 30 12 2018].

[6] M. K. Helwa and A. P. Schoellig, "Multi-robot transfer learning: A dynamical system perspective," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vancouver, BC, 2017.

[7] DTAI - Ku Leuven, "Deep Transfer: Generalizing Across Domains," EU FP7 - People, 2012 - 2015. [Online]. Available: https://dtai.cs.kuleuven.be/projects/deeptransfer. [Accessed 30 12 2018].

[8] European Commission, "Transfer Learning within and between Brains," FP7-IDEAS-ERC, 2014 - 2019. [Online]. Available: https://cordis.europa.eu/project/rcn/191311/factsheet/en. [Accessed 30 12 2018].

[9] B. Bócsi, L. Csató and J. Peters, "Alignment-based transfer learning for robot models," in *International Joint Conference on Neural Networks (IJCNN)*, Dallas, TX, 2013.