Robust Semi-Supervised Adversarial Subject-Level Transfer Learning for Sensor-Based Human Activity Recognition

Elnaz Soleimani[1,*], Abdelghani Chibani and Ghazaleh Khodabandehlou [*]

***Abstract*— The performance of Human Activity Recognition (HAR) models, particularly deep neural networks, is highly contingent upon the availability of the massive amount of annotated training data which should be sufficiently labeled. Though, data acquisition and manual annotation in the HAR domain are prohibitively expensive due to skilled human resource requirements in both steps. Hence, domain adaptation techniques have been proposed to adapt the knowledge from the existing source of data. More recently, adversarial transfer learning methods have shown very promising results in image classification, yet limited for sensor-based HAR problems, which are still prone to the unfavorable effects of the imbalanced distribution of samples. This paper presents a novel generic and robust approach for semi-supervised domain adaptation in HAR, which capitalizes on the advantages of the adversarial framework to tackle the shortcomings, by leveraging knowledge from annotated samples exclusively from the source subject and unlabeled ones of the target subject. Extensive subject translation experiments are conducted on three large, middle, and small-size datasets with different levels of imbalance to assess the robustness and effectiveness of the proposed model to the scale as well as imbalance in the data. The results demonstrate the effectiveness of our proposed algorithms over state-of-the-art methods, which led in up to 13%, 4%, and 13% improvement of our high-level activities recognition metrics for Opportunity, LISSI, and PAMAP2 datasets, respectively. The LISSI dataset is the most challenging one owing to its less populated and imbalanced distribution. Compared to the SA-GAN adversarial domain adaptation method, the proposed approach enhances the final classification performance with an average of 7.5% for the three datasets, which emphasizes the effectiveness of micro-mini-batch training. The manuscript provides a comprehensive evaluation of model performance, the explanation of the training procedure, the impact of sample population on the classifier performance, and the depiction of elements of the adversarial game.**

I. Introduction

Sensor-based HAR can be formulated as predicting current activity according to a sequence of sensors outputs. Any HAR related dataset has a finite amount of samples that are obtained from a limited number of human subjects. However, considering the requirements of applying HAR in real-world conditions, it is more interesting to evaluate the performance of an HAR model against many human subjects who their behaviors' data have not been included in the training dataset.

The shift between the source and target may root in the learning domain, learning task, or both. Furthermore, the source and the target domains may be dissimilar also in terms of class distributions, typically known as *class imbalance* problem in machine learning, in which, the conditional distributions of feature values are the same in source and target domains, yet the labels may not follow the same distribution in both domains.

Subject level knowledge transfer concentrates on the generalization of the knowledge a machine learning model. The latter is trained from a known subject and should be extended to unknown or unseen subject. Let us consider an HAR system that is supposed to recognise the activities of the inhabitants of smart homes. The inhabitants are considered as new subjects from the perspective of the HAR system. Even if the HAR system can be setup to temporarily collect data and learn inhabitants activities in the same time in a kind of a system initialisation mode, the annotation of the collected samples, by human experts or the inhabitant themselves to apply supervised machine learning will be infeasible. In this case the appropriate approach is a semi-supervised learning, where labeled data is provided in the source domain (subject) while target domains' samples are label excluded. Formally, the objective is to adapt the target domain $D_t = \big(X_t, P(X_t)\big)$ to the source domain $D_s = \big(X_s, P(X_s)\big)$ so as to have enough labeled data to train a HAR model on target domain. It has commonly been assumed that in the class imbalance domain adaptation problems, $P(X_s|Y_s = y_i) = P(X_t|Y_t = y_i)$ is held for all classes i, though the distribution of classes may not be the same in both domains [1]. Taking any pair of subjects in the evaluation sets as the source and target domain, there exists a shift between the class probability distribution of both domains, which means $P(Y_s) \neq P(Y_t)$. The imbalanced class distribution, along with the absence of the labels in the target domain, may pose a significant drawback on the performance attainable by the adaptation process.

Domain adaptation methods can be sorted out into 4 categories based on the type of knowledge transferred: Instance, feature representation, parameter, and relational transfer [2]. The first two categories focus on drawing the samples of both domains closer, by direct transformation or finding a common representation, respectively. Parameter and relational transfer methods transform prior knowledge and parameters and data relationship between domains. Each Transfer Learning approach should address the primary questions of *What, How, and From where to transfer* [3]. The proposed solution is a kind of instance transfer except that it combines the data transformation and classifier training

[*] Laboratoire Images, Signaux et Systemes Intelligents (LISSI), University Paris-Est Creteil, France.
[1] `elnaz.soleimani@u-pec.fr`

procedure. Therefore, it transfers instances and parameters instantaneously. To simplify the problem in terms of sources of transfer, we focused on one-to-one translation for this work. Following section will address the question of *How to transfer*.

II. Proposed method

Let us consider $X^s = \{(x_s, y_s)^i \mid i = 0 \rightarrow n_s\}$ represents the set of n_s labeled samples from the source domain $D_s = (X_s, P(X_s))$ and $X^t = \{(x_t)^i \mid i = 0 \rightarrow n_t\}$ denotes the set of n_t unlabeled samples from the target domain $D_t = (X_t, P(X_t))$. The proposed adversarial adaptation model consists of a Generator (G), Discriminator (D), and Classifier (C). The generator $G(x, z; \theta_G)$ is a differentiable function represented by a Convolutional Neural Network (CNN) that generates synthetic data, also called fake samples, by using the input and noise vector. The discriminator $D(x; \theta_D)$ is defined as a CNN that outputs a single scalar indicating the probability that x came from the target domain rather than the generator. The classifier $C(x; \theta_C)$ is also a CNN predicts the class of the input. These elements are playing a min-max game together based on the cost function $\mathbb{J}$ which combines the loss functions of adversary and classification tasks as follows [2]:

$$\mathbb{J}(G, C, D, X^s, X^t) = \min_{G,C} \max_{D} \mu \mathbb{J}_D(D, G, X^t, X^s) + \lambda \mathbb{J}_C(C, G, X^s) \quad (1)$$

The impact of the classification $\mathbb{J}_C$ and adversary $\mathbb{J}_D$ task loss on the generation task are reflected and controlled by μ and λ coefficient respectively. Generator supposed to generate artificial data which are similar to the samples of the target domain by descending on the gradient of $\mathbb{J}$. During the training phase, these fake generated samples are getting as similar as possible so that the discriminator will not be able to discriminate them from the original target data. The training samples are commonly feeding into the adversarial frameworks in mini-batch form, to avoid the mode collapse problem in which, the generator learns only to generate fake samples from a few classes (modes) of the data distribution, albeit the samples from the ignored modes appears in the training set [4]. Therefore, the generator collapses into the few modes that discriminator assumes them highly realistic. On the contrary, feeding the samples in mini-batch to the discriminator rather than in isolation, gives a broader horizon to the discriminator and possibly avoids the mode collapse. However, it would not be practical enough in all cases, especially for the highly imbalanced small-size datasets such as LISSI. Mini-batch selection of the dataset would even aggravate the problem in practice since it scales down the sample size. It is highly probable that the less populated classes be left without a representative in some batches, as the sample size shrinks and the proportions of the classes in the sample space cannot be taken for granted. Consequently, the discrimination of that certain classes would be unconcerned in the gradient computation of the batch due to the disappearance of their samples.

To tackle this issue, we applied a *micro-mini batch* strategy of learning. Each mini-batch consists of $\mathbb{C}$ micro-batches whereas $\mathbb{C}$ is the number of classes of activities in the dataset. Micro-batches contain m samples randomly drawn without replacement from each one of $\mathbb{C}$ classes while m can be set by the value of the least populated class population in the dataset. In this way, the proposed adversarial approach does not only prevent the Mode Collapse and its related issues but also enables a kind of parallelism of instance transfer and re-training of the target domain's classifier.

The optimization problem for discriminator can be solved by ascending the gradient of a Mean Squared Error (MSE) loss function $\mathbb{J}_D$:

$$\mathbb{J}_D(D, G, X^t, X^s) = \frac{1}{m\mathbb{C}} \sum_{j=1}^{\mathbb{C}} \sum_{i=1}^{m} \Big[\Big(1 - D\big(x_t^{(i)}\big)\Big)^2_{x_t^{(i)} \in b_t^j} + \Big(1 + D\big(G(x_s^{(i)}, z)\big)\Big)^2_{x_s^{(i)} \in b_s^j} \Big] \quad (2)$$

where b_s^j and b_t^j refer to the j-th micro part of the current mini-batch from the source X^s and target X^t domain samples, respectively. Correspondingly, the classifier attempts to assign a right label to its inputs including source domain data and the synthetic data generated by optimizing the cross-entropy loss function $\mathbb{J}_C$:

$$\mathbb{J}_C(C, G, X^s) = \frac{1}{m\mathbb{C}} \sum_{j=1}^{\mathbb{C}} \sum_{i=1}^{m} \Big[- y_s^{(i)} \log C\big(x_s^{(i)}\big) - y_s^{(i)} \log \Big(C\big(G(x_s^{(i)}, z)\big)\Big) \Big]_{x_s^{(i)}, y_s^{(i)} \in b_s^j} \quad (3)$$

Finally, when the training loss values converge, the training phase can be terminated and the classifier component will be functional independently.

In summary, each iteration of the training procedure consists of 3 steps for the mini-batch update of D, C, and G, respectively. Reordering the steps may affect convergence flow. The components of the model together struggle to close in the distribution of target domain samples on those of source domain where the labels are available. Having samples with approximately the same distribution, source domain labels are compatible to be exploited in supervised training of the classifier. As the distribution of generated samples $P(G(X^s, z))$ getting closer to $P(X^t)$, C's performance improves since the source labels more deeply cohere with target inputs.

III. Evaluation

We evaluate the proposed model by conducting extensive experiments on three datasets of different size (Opportunity,

TABLE I

COMPARISON OF THE THE GENERIC PROPOSED APPROACH PERFORMANCE AND GFK [5], STL[6] AND SA-GAN[2] MODEL, IN TERMS OF WEIGHTED F1 MEASURE ON **OPPORTUNITY** DATASET. THE MOST DOMINANT PERFORMANCE IN EACH TRANSFORMATION EXPERIMENT MARKED IN BOLD.

Source Subject	Target Subject	Distance	No Transfer	STL	GFK	SA-GAN	Proposed Model	Supervised
1	2	46.69	0.45	0.65	0.59	0.73	**0.74**	0.75
	3	45.10	0.27	0.37	0.43	0.45	**0.58**	0.71
	4	77.15	0.40	0.47	0.55	0.49	**0.57**	0.59
2	1	40.47	0.48	0.52	**0.62**	0.56	0.56	0.65
	3	34.38	0.44	0.46	0.51	**0.52**	0.40	0.71
	4	72.80	0.29	**0.46**	0.40	0.39	0.42	0.59
3	1	38.38	0.23	0.40	0.45	0.42	**0.52**	0.65
	2	37.54	0.21	0.54	0.53	**0.61**	0.52	0.75
	4	73.69	0.31	0.37	0.44	0.44	**0.50**	0.59
4	1	73.53	0.26	0.38	0.51	0.51	**0.52**	0.65
	2	70.80	0.29	0.54	0.45	0.55	**0.68**	0.75
	3	69.44	0.24	0.48	0.37	0.49	**0.53**	0.71

TABLE II

COMPARISON OF THE THE GENERIC PROPOSED APPROACH PERFORMANCE AND GFK [5], STL [6] MODEL, IN TERMS OF WEIGHTED F1 MEASURE ON **PAMAP2** DATASET. THE MOST DOMINANT PERFORMANCE IN EACH TRANSFORMATION EXPERIMENT MARKED IN BOLD.

Source Subject	Target Subject	Distance	No Transfer	STL	GFK	SA-GAN	Proposed Model	Supervised
1	5	91.82	0.37	0.62	0.72	0.69	**0.77**	0.98
	6	91.58	0.32	0.56	0.64	0.66	**0.70**	0.97
	8	107.07	0.04	0.57	0.49	0.65	**0.72**	0.92
5	1	91.82	0.32	0.76	0.66	0.71	**0.76**	0.99
	6	42.13	0.64	0.83	0.75	0.83	**0.83**	0.97
	8	56.01	0.26	0.52	0.69	0.66	**0.73**	0.92
6	1	91.58	0.16	0.67	0.56	0.61	**0.78**	0.99
	5	42.13	0.41	0.74	0.75	0.79	**0.83**	0.98
	8	56.76	0.17	**0.86**	0.58	0.63	0.82	0.92
8	1	107.07	0.10	0.54	0.58	0.68	**0.76**	0.99
	5	56.01	0.25	0.55	0.41	0.60	**0.73**	0.98
	6	56.76	0.27	0.58	0.61	**0.73**	0.71	0.97

TABLE III

RESULTS OF APPLYING SUBJECT TO SUBJECT TRANSFER LEARNING ON **LISSI** DATASET.

	S1→S2	S6→S2	S1→S3	S6→S3	S1→S4	S6→S4	S1→S5	S6→S5
STL	0.40	0.39	0.30	0.45	0.44	0.47	0.53	0.60
GFK	0.51	0.35	0.44	0.50	0.25	0.55	0.70	0.49
SA-GAN	0.56	0.54	0.66	0.70	0.54	0.67	0.76	0.78
Proposed model	0.58	0.56	0.70	0.65	0.52	0.64	0.76	0.70
Supervised	0.71	0.71	0.85	0.85	0.82	0.82	0.88	0.88

PAMAP2 and LISSI datasets) to assess its functionality and robustness. We opted for W-F1 measure as the evaluation metric since it gives better insight comparing to accuracy, precision, and recall deliberating imbalance distribution of classes in the dataset. Set of subject translation experiments have been held to practically demonstrate the necessity of domain adaptation and its effectiveness as well. For each dataset samples of each subject are considered independently as the source or target domain. For each target domain (subject) a classifier is adapted, examined and compared with two other adaptation methods, namely, STL[6] and GFK[5] as well as adaptation performance upper bound. Table I to IV reports the obtained results.

IV. CONCLUSIONS AND FUTURE WORK

This study set out to propose a generic adversarial framework for knowledge transfer in the domain of Human Activity Recognition. The proposed semi-supervised model has been evaluated against three datasets with different challenges to assess its robustness to the scale and imbalance of the data. The findings of our research are quite convincing, and thus the following conclusions can be drawn:

TABLE IV
THE MAIN CLASSIFICATION METRICS OF A SAMPLE CLASSIFICATION TASK ON OPPORTUNITY, LISSI, AND PAMAP2 DATASETS.

Opportunity Dataset				LISSI Dataset				PAMAP2 Dataset			
Subject 4 → Subject 2				Subject 1 → Subject 5				Subject 5 → Subject 6			
class	precision	recall	support	class	precision	recall	support	class	precision	recall	support
Null	**0.98**	0.56	292	Kneeling	0.63	0.83	53	Ironing	0.87	0.99	137
Relaxing	0.81	**0.84**	190	Lying	0.78	0.78	79	Lying	0.89	**0.99**	85
Coffee time	0.34	0.33	225	Relaxing	**0.99**	0.87	142	Sitting	0.86	0.99	84
Early morning	0.61	0.84	372	Sitting	0.74	0.47	49	Standing	0.88	0.65	89
Clean up	0.34	0.57	224	Sit to Stand	0.80	**0.88**	100	Walking	0.87	0.86	93
Sandwich time	0.94	0.71	679	Standing	0.87	0.71	76	Running	0.82	0.94	82
				Dance walk	0.69	0.51	57	Cycling	**0.93**	0.88	74
				Warm up	0.38	0.73	41	Ascending	0.59	0.82	49
								Descending	0.80	0.49	41
								Cleaning	0.70	0.43	76
Accuracy	0.67		1982	Accuracy	0.76		597	Accuracy	0.83		810
W-Avg	0.74	0.67	1982	W-Avg	0.79	0.76	597	W-Avg	0.84	0.83	810

Compared to the SA-GAN adversarial domain adaptation method, our proposed model enhances the final classification performance with an average of 7.5% for the three datasets, which reinforce the effectiveness of micro-mini-batch training approach. The proposed model provides striking results on the PAMAP2 benchmark medium-size multi-class dataset. It improved the adversarial domain adaptation performance by applying a micro-mini-batch learning technique on Opportunity large-scale yet highly imbalanced dataset. Interestingly, the proposed model revealed competitive results compared to other states of the art models on the LISSI dataset, which is very challenging in terms of both the number of samples and balance of classes. Our comprehensive assessment was carried out over high-dimensional data of highly abstract activities in all three datasets. Besides, the proposed approach is not HAR-exclusive and it can be potentially utilized to solve other domain adaptation problems. The results support the effectiveness of the proposed model to address the imbalanced learning challenges. Further studies with more focus on the lack of samples problem will be undertaken. In future investigations, it might be possible to use multiple sources of knowledge or a combination of transferred models from different source domains, and the source/model selection policies. In addition, the integration of the ontology-based reasoning to the present approach could be a mean of improvement of the classification results obtained by Machine Learning.

REFERENCES

[1] J. Jiang, "A literature survey on domain adaptation of statistical classifiers," *URL: http://sifaka. cs. uiuc. edu/jiang4/domainadaptation/survey*, vol. 3, pp. 1–12, 2008.

[2] E. Soleimani and E. Nazerfard, "Cross-subject transfer learning in human activity recognition systems using generative adversarial networks," *CoRR*, vol. abs/1903.12489, 2019.

[3] D. Feng, M. Kaboli, and G. Cheng, "Active prior tactile knowledge transfer for learning tactual properties of new objects," *Sensors*, vol. 18, no. 2, p. 634, 2018.

[4] A. Srivastava, L. Valkov, C. Russell, M. U. Gutmann, and C. Sutton, "Veegan: Reducing mode collapse in gans using implicit variational learning," in *Advances in Neural Information Processing Systems*, pp. 3308–3318, 2017.

[5] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2066–2073, IEEE, 2012.

[6] J. Wang, Y. Chen, L. Hu, X. Peng, and S. Y. Philip, "Stratified transfer learning for cross-domain activity recognition," in *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 1–10, IEEE, 2018.