University of Sindh Journal of Information and Communication Technology (USJICT)
Volume 6, Issue 2, July 2022
ISSN-E: 2523-1235, ISSN-P: 2521-5582
Website: http://sujo.usindh.edu.pk/index.php/USJICT/

Stacked Autoencoder and Meta-Learning based Heterogeneous Domain Adaptation for Human Activity Recognition

Sajjad Ali Awan[1], Sunder Ali Khowaja[2], Kamran Taj Pathan[3]

[1]Dr. A.H.S Bukhari Postgraduate Center of ICT, University of Sindh, Jamshoro.
[2]Department of Telecommunication Engineering, Faculty of Engineering and Technology, University of Sindh, Jamshoro.
[3]Department of Software Engineering, Faculty of Engineering and Technology, University of Sindh, Jamshoro.
roykhanali@gmail.com, sandar.ali@usindh.edu.pk, Kamran.taj@usindh.edu.pk

Abstract: The field of human activity recognition (HAR) using machine learning approaches has gained a lot of interest in the research community due to its empowerment of automation and autonomous systems in industries and homes with respect to the given context and due to the increasing number of smart wearable devices. However, it is challenging to achieve a considerable accuracy for recognizing actions with diverse set of wearable devices due to their variance in feature spaces, sampling rate, units, sensor modalities and so forth. Furthermore, collecting annotated data has always been a serious issue in the machine learning community. Domain adaptation is a field that helps to cope with the issue by training on the source domain and labeling the samples in the target domain, however, due to the aforementioned variances (heterogeneity) in wearable sensor data, the action recognition accuracy remains on the lower side. Existing studies try to make the target domain feature space compliant with the source domain to improve the results, but it assumes that the system has a prior knowledge of the feature space of the target domain, which does not reflect real-world implication. In this regard, we propose stacked autoencoder and meta-learning based heterogeneous domain adaptation (SAM- HDD) network. The stacked autoencoder part is trained on the source domain feature space to extract the latent representation and train the employed classifiers, accordingly. The classification probabilities from the classifiers are trained with meta-learner to further improve the recognition performance. The data from tar- get domain undergoes the encoding layers of the trained stacked autoencoders to extract the latent representations, followed by the classification of label from the trained classifiers and meta- learner. The results show that the proposed approach is efficient in terms of accuracy score and achieves best results among the existing works, respectively.

Keywords: Human Activity Recognition, Domain Adaptation, Heterogeneous Domain Adaptation, Stacked Autoencoders, Meta-Learning.

I. INTRODUCTION *(HEADING 1)*

In recent times, manufacturing systems have seen strides of advancements, increased competitive performance, and efficient production, through the automation systems. Although many tasks concerning industrial systems have been auto- mated, some processes have limited flexibility as they heavily rely on humans [1]. Furthermore, with the emergence of Indus- try 5.0, human involvement in industrial processes is considered to be inevitable. Therefore, a lot of focus has been drawn to the applications that can help in optimizing the industrial processes, integration of human workers, and their flexibility concerning working conditions. The integration of human workers needs to be made through interfaces powered by software- or intelligence-controlled systems. Many researchers have considered the human activity recognition (HAR) system to be one of the prospects that can provide seamless integration with the industrial processes and manufacturing systems [2]. The HAR addition not only helps in improving the automatic assessment of integration for cyber physical systems (CPS) with humans through quantitative analysis but also helps in gaining insights regarding the bottlenecks in workflows such as monitoring activity duration for a specific activity [3]. This allows an industrial process to be evaluated from anthropocentric perspective [4].

In retrospective, the impact of HAR has been explored across variety of domains that include activity recognition for construction workers [5], ergonomic assessment in smart factories [6], and evaluation of assembly task analysis [7]. The afore- mentioned examples help in realizing the importance of HAR in the context of industrial workplaces for tackling unexpected accidents, detection of inefficient tasks, and analysis of the work process, in a timely manner. In order to build a robust and accurate HAR system, we need to have sufficient amount of labeled data, however the process for acquiring such data is time consuming and often expensive. An alternative is to consider existing publicly available datasets, but they cannot be used in a straightforward manner due to the domain differences. To address such difference, domain adaptation methods were proposed that aim to reduce the discrepancies between the source domain (training set) and the target domain (testing set), accordingly [8]- [10]. The domain adaptation concerning HAR infers the class labels for the samples in the target domain by leveraging the class label data from the source domain. For instance, a model trained on the activities of a worker in an assembly line can be used to classify the activities for a worker in an automobile repair shop as both of the domains involve similar body postures and hand movement. Such type of adaptation across different domains is considered to be homogeneous as the labels and sensor modalities were the same. Earlier works mostly focused on the homogeneous domain adaptation while assuming that the feature space for both the source and target domains are same, i.e., number of features, sensor modalities, and so forth. However, with the recent advancements in microelectromechanical systems (MEMS), different types of sensors are used in wearable devices that vary from sampling rate to sensing units, thus, making it difficult for domain adaptation methods to train cross-domain HAR classifiers.

Few works have been carried out on heterogeneous domain adaptation to address the limitations concerning its homogeneous counterpart. In contrast to homogeneous, heterogeneous domain adaptation assumes that both the source and target domains have varying feature spaces. Conventional heterogeneous domain adaptation methods try to align the heterogeneous input space in order to make the feature space homogeneous but in order to attain the task, the method requires additional information such as labeled samples in the target domain and instance correspondence [11]. Although, it helps in overcoming the heterogeneity challenge but in practical cases such information is not available to begin with. Recently some studies have taken a hybrid approach to heterogeneous domain adaptation such as [12] suggesting that the common segment of feature spaces are leveraged to transform them into homogeneous feature spaces (same dimensionality). The approach seems promising concerning HAR as most of the sensing modalities are similar. Earliest attempts made in heterogeneous domain adaptation was only to leverage the common intersection of feature spaces while ignoring/omitting its counterpart. Although it might work in some cases such as distinguishing walking from lying down and sitting but in complex cases it fails to achieve desirable results, for instance, walking vs walking downstairs, running vs jogging, and more. It is because of the ignorance of remaining features that are not common to both the domains. The studies [13] use both the common and different features to retain domain- specific information. The problem with the aforementioned study is they consider the original feature space for processing the common features that potentially might vary in terms of distribution, and they use hand-crafted features for filling out the missing feature space part. Similarly, the study [14] uses converter sub-networks to fill in or align the feature space for both the domains but it also relies on some domain-specific information to perform the task. Therefore, a dire need of heterogeneous domain adaptation method with unsupervised characteristics to transform the heterogeneous feature space. In this paper, we propose stacked auto-encoder and meta- learning based heterogeneous domain adaptation (SAM-HDD) for human activity recognition. Instead of using raw data we directly process the feature space using stacked auto- encoders to make the feature space homogeneous. We then train multiple streams with long-short term memory networks (LSTMs), Random Forests, and Boosting classifier, followed by a Bayesian learning based meta-learner for improving the recognition results. We test the proposed approach on two publicly available datasets to validate its efficacy. The contributions of this work are as follows:

- We propose stacked auto-encoder based architecture for transformation of heterogeneous feature space to homogeneous.
- We propose a Bayesian meta-learner to combine the classification probabilities from LSTM streams.
- We report superior performance in comparison to the existing baselines.

The rest of the paper is structured as follows: Section 2 consolidates related works in the field of homogeneous and heterogeneous domain adaptation. Section 3 presents the details regarding proposed methodology. Section 4 provides quantitative results and Section 5 concludes the work along with future prospects.

II. Related Works

This section consolidates works concerning human activity recognition, transfer learning, and domain adaptation. One of the main characteristics for HAR is human behavior recognition by analyzing the measurements from inertial measurement units (IMUs). HAR has been proven to be effective when analyzing the work processes to improve productivity and implemented in workspaces to avoid occupational safety and health issues [5]-[7]. Human activity monitoring is of utmost importance in operator 4.0-compliant system as it notifies in timely manner if any anomalous activity is detected, or potential threat is sensed. With the emergence of wearable sensors, it’s not easy, affordable, and secure to acquire an individual worker's data including IMUs, heart rate, and others to monitor their health-specific information [15], [16]. Existing studies have extensively used machine learning approaches to recognize human activities from wearable sensors. The study in [17] uses accelerometer data from smartphones to monitor and classify human activities. The study in [18] mainly focuses on the assisted daily living for elderly by recognizing commute-based activities using smartphone embedded sensors. The study in [19] used passive infrared sensors to assess health anomalies in a smart home environment. All of the aforementioned works mainly rely on feature extraction and shallow learning techniques to recognize, assess, and classify human activities. Recently, deep learning techniques including convolutional neural networks (CNN) and LSTMs have been introduced that does not need hand-crafted features to be extracted, rather they extract the features automatically. Many researchers have opted for 1D-CNNs and LSTMs for recognizing and classifying human activities, since then [3]. However, these studies ignore the domain differences between different datasets, suggesting that the method is tailored to serve single environment context, device setup or user. Furthermore, the aforementioned studies require enormous amounts of labelled samples which is hard to get for a new user or different devices. Shallow learning-based approaches can be applied for testing samples from other domains, but it has been proven by existing studies that the yielded performance is not acceptable for a real-world environment.

Recently, transfer learning strategies were introduced to deal with the diversification of environment, device setup, and user differences. These strategies leverage the knowledge from source domains to handle domain adaptation for target domains. The transfer learning approach where labels for target domains are not used are considered to be unsupervised domain adaptation. To mention a few of them, Fernando et al. [8] leveraged the eigenvectors for creating domain-invariant features. Gong et al. [9] uses statistical and geometric properties to characterize domain shift in feature spaces. They termed their method as geodesic flow kernel (GFK) based domain adaptation for HAR. It should be noted that the aforementioned works consider homogeneous feature spaces rather than heterogeneous ones, thus they heavily rely on hand crafted features that can reduce class-wise discrepancies. Khan et al. [20] considered the use of convolutional neural networks that can extract features in an end-to-end manner, but the study still dealt with the homogeneous domains, therefore, it cannot handle diverse feature spaces.

Heterogeneous domain adaptation studies address the limitations of its homogeneous counterpart; however, it is more challenging to implement and attain a good performance. The earlier ideas concerning heterogeneous domain adaptation were to augment the feature spaces of both domains so that they may be compliant in terms of dimensionality. The works [21], [22], mainly focused on the aforementioned idea to realize heterogeneous domain adaptation. Some of the heterogeneous domain adaptations were centered around text-analysis in cross language application. The studies in [11] used some prior information of instance-correspondence to align the feature spaces from the cross-domain datasets. However, the same cannot be performed in HAR as prior instance correspondences are not available. Recently, a study [13] proposed the use of limited labels from the target domain samples as a prior information to align the feature spaces. The study in [13] addressed the heterogeneous domain adaptation task by leveraging the commonalities in domain-specific feature spaces. It makes sense as most of the times, sensor modalities are the same such as accelerometer and gyroscope. The study [14] designed a converter sub-network to fill in the missing part of feature spaces for making both domains compliant. However, both of the approaches require some prior information from the target domain in order to comply with the dimensions of the feature space which is not readily available with new or limited amount of data in real-world environment. In this regard, this work does not consider prior information from the specific domains rather shares auto-encoder weights from one domain to extract features from other domains and leverage the characteristics of both deep and shallow-learning algorithms to develop a meta-learning approach for labeling the target domain samples.

III. Proposed Work

The abstract workflow of our proposed work is shown in Figure 1. The sub-sections will explain the working principles for each of the blocks in a comprehensive manner.

A. *Problem Definition*

We represent the source-specific and target-specific features with X_{ss} and X_{st}, having dimensions $\mathbb{R}^{\omega \times ss}$ and $\mathbb{R}^{\omega \times ts}$, respectively. The notation ω corresponds to the features extracted from a single window length of sensor readings. The proposed study assumes the feature space to not be homogeneous, i.e., the number of features may vary in source and target specific domains, $ss = ts$. The study also assumes that the label information is only incorporated in the features belonging to the source specific domain, i.e., y_{ss}. The aim of this study is to infer the labels from the features extracted from target domain, i.e., y_{ts}.

B. *Stacked Autoencoder*

This work uses stacked autoencoder as a feature space transformation technique rather than a conventional feature extractor. The underlying theory behind stacked autoencoders is similar to that of principle component analysis (PCA), i.e., representing the original representation in a compressed form. The key difference between PCA and stacked autoencoder is the usage of linear and non-linear functions for representing the complex data, respectively. Existing studies have extensively used stacked autoencoders for feature extraction, transformation, detection, and classification tasks [23], [24]. Some studies also prefer to use the sparse autoencoders to improve the performance as the sparsity constraint enforces the activation of limited neurons. This strategy results in a kind of natural feature selection that most of the times improves the recognition performance [23]. However, in this study, the stacked autoencoders are not used for recognition, rather they are used for feature space transformation for creating a homogeneity between source and target domain. Therefore, we assume that the sparse stacked autoencoders could work in the capacity of discriminating representation that could extract a common subspace from source and target domain feature space, respectively. We adopt the sparse stacked autoencoder network architecture from the study [23]. The architecture of stacked autoencoder is such that the number of neurons is the same at input and output layers but gets significantly lower in the hidden layer (bottleneck). The function that the networks seek to learn for extracting discriminating representations is shown in equation 1.

$$\widehat{\boldsymbol{X}} = \boldsymbol{f}_{\boldsymbol{w},\boldsymbol{b}}(\boldsymbol{x}) \tag{1}$$

where the notations w and b refer to the weights and biases associated with the input x and reconstructed output $\hat{x}$ and f

refers to the activation function, accordingly. The network then computes the average activation over the training set for the k^{th} hidden layer neuron as shown in equation 2.

$$\boldsymbol{\alpha_k = \frac{1}{k}\sum_{k=1}^{k} act(x_n)} \qquad (2)$$

The notation K refers to the total number of examples provided to the input layer and act_k refers to the *kth* neuron activation for n^{th} input value. In order to make the representation sparse, a sparsity constraint (fixed value) is introduced such that $\alpha_k = \overline{\alpha}$. The value for the sparsity parameter has been selected as 0.05 in existing studies, accordingly. The sparsity term can be incorporated in the objective function as shown in equation 3.

$$\boldsymbol{\rho_{sparse}(w, b) = \rho(w, b) + \gamma \sum_{k=1}^{f_3} KL}\ (\alpha_k || \overline{\alpha}\ ) \qquad (3)$$

In the above equation the cost function is denoted by ρ, the network is trained by minimizing the cost function with reference to the corresponding weights and biases. The term γ corresponds to the sparsity coefficient that controls this parameter. The notation f_3 represents the number of neurons in the bottleneck layer and the KL represents the Kullback-Leibler divergence between the Bernoulli random variable $\boldsymbol{\overline{\alpha}}$ and the mean of the average activation $\boldsymbol{\alpha}$, respectively. The KL divergence for the aforementioned parameters is defined in equation 4.

$$\boldsymbol{KL}\ (\alpha_k || \overline{\alpha}\ ) = \boldsymbol{(1 - \overline{\alpha}\ )\ log \frac{1-\overline{\alpha}}{1-\alpha_k}} + \overline{\alpha}\ log \frac{\overline{\alpha}}{\alpha_k} \qquad (4)$$

The KL divergence can also be referred to as relative entropy between the mean of a random variable and average of an activation function. The corresponding weights and biases are then updated using stochastic gradient descent (SGD) function, accordingly. Once, the sparse stacked autoencoders are trained on the source domain, the weights and biases for the encoder part are shared with the sparse stacked autoencoders for the target domain to transform heterogeneous feature space into a homogeneous one. The representation obtained from the bottleneck layer of the source domain would then be trained using different classifiers, accordingly. The details for the employed classifiers are provided in the subsequent subsection.

C. Classification Methods

As shown in figure 1, we use three different types of classification methods to train the transformed features extracted from stacked autoencoder's bottleneck layer. The three classification methods include eXtreme Gradient boosting (XGBoost), random forests (RF), and long-short term memory networks (LSTM) [13], [23], [24]. All three classifiers have different characteristics, such that two of them belong to the family of shallow learning methods while the third one uses deep learning architecture. Amongst the two shallow learning methods the XGBoost is categorized as boosting while the RF is categorized as bagging tree method, respectively. The boosting strategy mainly trains in a recursive manner suggesting that among many sub classifiers the weak classifiers are train recursively to transform into the stronger ones. On the other hand, bagging trees use multiple decision trees with random pruning and selective features. A maximum voting strategy is carried out amongst the bagging trees to select the classification label. Recently, deep learning strategies have gained a lot of interest in research community due to their end-to-end classification and promising recognition results. However, in this case, an end-to-end learning is not adopted due to the heterogeneity of the input feature space between two domains. The LSTMs make use of gates such as forget, cell, input, and output gate to update the learned values which then can be used for further classification. These three classifiers are trained simultaneously independent of each other so that a decision-level fusion module can be applied, accordingly.

D. Decision-Level Fusion

The workflow for decision-level fusion using Bayesian meta-learner in the proposed work is shown in Figure 2. The classification probabilities from all three classifiers will be forwarded to the decision-level module that combines the aforementioned probabilities in a directed acyclic graph (DAG) manner to yield the final output label. We adopt the Bayesian meta-learner from the study [24], [25]. The computation for the fusion method is shown in equation 5.

$$\boldsymbol{\sigma(P_i | feat_{SAE}) = \frac{\sigma(P_i)\sigma(feat_{SAE}|P_i)}{\sum_{i=1}^{n}\sigma(P_i)\sigma(feat_{SAE}|P_i}} \qquad (5)$$

where σ refers to the classification probabilities obtained using the three classifiers, $\sigma(P_i|feat_{SAE})$ refers to the probability when a particular classification label is likely to occur with the provision of stacked autoencoder feature vector. Based on the classification probabilities, a Bayes classification rules as proposed in [24] will be used to yield the final output label, accordingly.

IV. Experimental Setup

This section mainly defines the experimental setup employed in this study. We employed two datasets, i.e., daily and sports activities (DSA) [27] and Opportunity (OPP) [28]. The reason for considering these datasets is that they both contain motion sensor data and have 4 common action categories, namely, lying, sitting, walking, and standing. In this study, we apply our cross-domain adaptation on two body parts, i.e., left and right wrist, which mimics the use of smartwatches, thus, makes it compliant with real world adaptation. In the experiments, we use one of the aforementioned datasets as the source domain and other as the target domain, respectively. The experiments are also conducted considering the intra-dataset, i.e., left - right

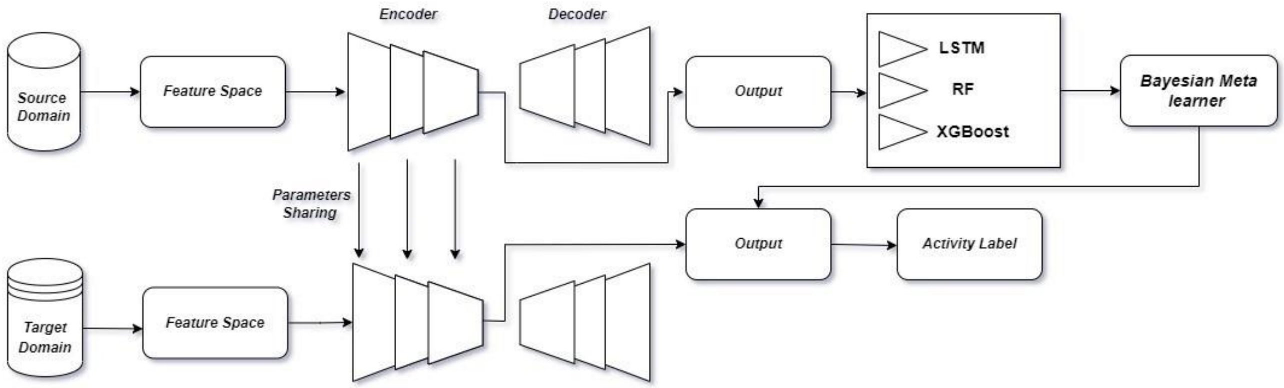

Figure 1 Proposed Workflow for SAM-HDD Network

wrists, and cross-dataset, i.e., DSA - OPP. We implemented our method using keras, tensorflow, and sklearn in python. For the LSTM we use a single layer with 256 hidden units. The learning rate was selected to be 0.001, with ADAM optimizer and default parameters. We used five layers for the stacked autoencoder with filter size of 32, 16, and 8 for the encoding layers. For the RF and XGBoost we use the depth of 10 and 20, learning rate of 0.1 and 0.01, respectively. As the validation set is non-existent due to the consideration that labels are unavailable in the target domain, all the aforementioned parameter values are chosen, empirically.

We consider 4 baselines for this study, i.e., support vector machine (SVM), domain-specific feature transfer (DSFT) [13], heterogeneous deep convolutional neural network (HDCNN) [20], and subspace alignment (SUB) [8]. We use the common feature space as HDCNN, SUB, and SVM can only handle homogeneous feature spaces.

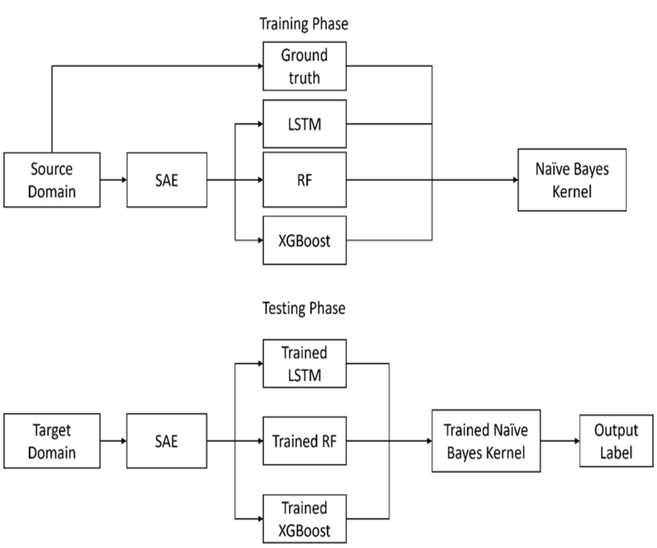

Figure 2 Workflow for decision-level fusion using Bayesian Meta-learner.

V. Result and Discussion

In this section, we report the accuracy score on the target domain as reported in Table 1. The accuracy score represents the ratio of correctly classified samples over the total number of samples. The first segment of results shows the intra-dataset cross variance on OPP dataset, i.e., right to left, and left to right wrist, the second segment shows the intra-dataset cross variance on DSA dataset in a similar manner, and the third segment illustrates the cross-data variance on OPP and DSA dataset, respectively. We also vary the sensor modality accordingly, i.e., from gyroscope to magnetometer. Our results are compared with the existing approaches such as DSFT, HDCNN, SUB, and SVM, accordingly. The results clearly indicates that the proposed approach outperforms the existing ones in all the domain adaptation tasks concerning the two datasets. Furthermore, it is also to be highlighted that the cross-dataset domain adaptation is the hardest task among all evaluated in Table 1, which is indicated by the lower accuracy scores. We believe that it is due to the heterogeneity in sensor modalities such as sampling rate, reading sensitivity, or value range. The said heterogeneity can be handled with SAM- HDD network if the data from heterogeneous devices can be used to train the stacked autoencoders for the source domain, respectively. Overall, we assume that the quantitative results suggest that the use of stacked autoencoder and decision- level fusion for the cross-domain adaptation is progressive and contributes to the improved recognition performance, accordingly.

VI. Conclusion

This work presents a stacked autoencoder and meta-learning based heterogeneous domain adaptation (SAM-HDD) network for human action recognition from wearable sensors. The SAM-HDD learns to transform the feature space into latent representation via stacked autoencoders using the source domain data. The encoder shares parameters to extract the latent representations from target domain data, accordingly.

Table 1 Accuracy Score for the Proposed SAM-HDD and its comparison with existing methods

Dataset	Target	Source	x_{ts}	x_{ss}	DSFT	HDCNN	SUB	SVM	SAM-HDD
OPP	Right	Left	Gyr	Mag	58.36%	65.42%	62.27%	61.36%	65.89%
OPP	Right	Left	Mag	Gyr	55.48%	65.11%	62.05%	61.17%	65.54%
OPP	Left	Right	Gyr	Mag	58.22%	66.63%	64.38%	53.85%	67.04%
OPP	Left	Right	Mag	Gyr	55.27%	66.34%	64.09%	53.56%	66.79%
DSA	Right	Left	Gyr	Mag	75.84%	62.47%	59.76%	60.42%	79.21%
DSA	Right	Left	Mag	Gyr	78.94%	63.23%	60.48%	60.93%	81.10%
DSA	Left	Right	Gyr	Mag	68.11%	67.34%	57.38%	56.87%	72.15%
DSA	Left	Right	Mag	Gyr	66.48%	67.61%	57.66%	56.21%	71.45%
Cross	OPP Left	DSA Left	Gyr	Mag	26.34%	39.76%	34.88%	21.05%	52.32%
Cross	DSA Left	OPP Left	Gyr	Mag	45.41%	41.68%	31.32%	28.77%	48.89%

Three different classification methods are employed namely, XGBoost, RF, and LSTM to train the network on latent representations followed by the training of a meta-learner that combines the classification probabilities of the three classifiers. The trained classifier and meta-learner is then used to classify action labels from the latent representations of the target domain dataset. We performed extensive evaluation on intra dataset domain variance and cross-dataset domain variance for the efficacy of the proposed approach and compared it with existing studies. The experimental results indicate that the proposed approach outperforms the existing ones in terms of accuracy score. It should be noted that unlike existing methods, the proposed approach does not rely on extracting common feature spaces or fill the feature spaces with the common sub- space as performed in existing studies. The proposed method has been evaluated on only two datasets and homogeneous labels which we consider as the limitation of this work. We intend to extend this work by considering more human activity recognition datasets for evaluating cross-domain adaptation and to use personalized models for coping with heterogeneous label spaces, accordingly.

ACKNOWLEDGMENT

We are thankful to the authors of OPP and DSA activity datasets for making them publicly available.

REFERENCES

[1] Kadir, B.A., Broberg, O. and Souza da Conceic¸a˜o, C., 2018. Designing human-robot collaborations in industry 4.0: explorative case studies. In DS 92: Proceedings of the DESIGN 2018 15th International Design Conference (pp. 601-610).

[2] Pacaux-Lemoine, M.P., Trentesaux, D., Rey, G.Z. and Millot, P., 2017. Designing intelligent manufacturing systems through Human-Machine Cooperation principles: A human-centered approach. Computers and Industrial Engineering, 111, pp.581-595.

[3] Khowaja, S.A., Yahya, B.N. and Lee, S.L., 2020. CAPHAR: context-aware personalized human activity recognition using associative learning in smart environments. Human-centric Computing and Information Sciences, 10(1), pp.1-35.

[4] Rauch, E., Linder, C. and Dallasega, P., 2020. Anthropocentric perspective of production before and within Industry 4.0. Computers and Industrial Engineering, 139, p.105644.

[5] Akhavian, R. and Behzadan, A.H., 2016. Smartphone-based construction workers' activity recognition and classification. Automation in Construction, 71, pp.198-209.

[6] Nath, N.D., Akhavian, R. and Behzadan, A.H., 2017. Ergonomic analysis of construction worker's body postures using wearable mobile sensors. Applied ergonomics, 62, pp.107-117.

[7] Malaise´, A., Maurice, P., Colas, F., Charpillet, F. and Ivaldi, S., 2018, March. Activity recognition with multiple wearable sensors for industrial applications. In ACHI 2018-Eleventh International Conference on Advances in Computer-Human Interactions.

[8] Fernando, B., Habrard, A., Sebban, M. and Tuytelaars, T., 2014. Subspace alignment for domain adaptation. arXiv preprint arXiv:1409.5241.

[9] Gong, B., Shi, Y., Sha, F. and Grauman, K., 2012, June. Geodesic flow kernel for unsupervised domain adaptation. In 2012 IEEE conference on computer vision and pattern recognition (pp. 2066-2073). IEEE.

[10] Zhang, J., Li, W., Ogunbona, P. (2017). Joint geometrical and statistical alignment for visual domain adaptation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1859–1867).

[11] Li, H., Pan, S. J., Wan, R., Kot, A. C. (2019). Heterogeneous transfer learning via deep matrix completion with adversarial kernel embedding. In Proc. 33rd AAAI Conf. Artif. Intell.(AAAI).

[12] Wu, F., and Huang, Y. (2016). Sentiment domain adaptation with multiple sources. In Proceedings of the 54th annual meeting of the association for computational linguistics (Volume 1: Long Papers) (pp.301–310).

[13] Wei, P., Ke, Y. and Goh, C.K., 2018. A general domain specific feature transfer framework for hybrid domain adaptation. IEEE Transactions on Knowledge and Data Engineering, 31(8), pp.1440-1451.

[14] Prabono, A.G., Yahya, B.N. and Lee, S.L., 2021. Hybrid domain adaptation with deep network architecture for end-to-end cross-domain human activity recognition. Computers Industrial Engineering, 151, p.106953.

[15] Khowaja, S.A., Prabono, A.G., Setiawan, F., Yahya, B.N. and Lee, S.L.,2021. Toward soft real-time stress detection using wrist-worn devices for human workspaces. Soft Computing, 25(4), pp.2793-2820.

[16] Khowaja, S.A., Prabono, A.G., Setiawan, F., Yahya, B.N. and Lee, S.L., 2018. Contextual activity based Healthcare Internet of Things, Services, and People (HIoTSP): An architectural framework for healthcare monitoring using wearable sensors. Computer Networks, 145, pp.190-206.

[17] Wannenburg, J. and Malekian, R., 2016. Physical activity recognition from smartphone accelerometer data for user context awareness sensing. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 47(12), pp.3142-3149.

[18] Jahn, A., Kroll, D., Lau, S.L. and David, K., 2017, November. An activ- ity history based approach for recognizing the mode of transportation. In 2017 IEEE conference on open systems (ICOS) (pp. 21-25). IEEE.

[19] Aramendi, A.A., Weakley, A., Goenaga, A.A., Schmitter-Edgecombe, M. and Cook, D.J., 2018. Automatic assessment of functional health decline in older adults based on smart home data. Journal of biomedical informatics, 81, pp.119-130.

[20] Khan, M.A.A.H., Roy, N. and Misra, A., 2018, March. Scaling human activity recognition via deep learning-based domain adaptation. In 2018 IEEE international conference on pervasive computing and communications (PerCom) (pp. 1-9). IEEE.

[21] Dai, W., Chen, Y., Xue, G.R., Yang, Q. and Yu, Y., 2008. Translated learning: Transfer learning across different feature spaces. Advances in neural information processing systems, 21.

[22] Prettenhofer, P., Stein, B. (2010). Cross-language text classification using structural correspondence learning. In Proceedings of the 48th annual meeting of the association for computational linguistics (pp.1118–1127).

[23] Khowaja, S.A. and Khuwaja, P., 2021. Q-learning and LSTM based deep active learning strategy for malware defense in industrial IoT applications. Multimedia Tools and Applications, 80(10), pp.14637-14663.

[24] P. Khuwaja, S. A. Khowaja and K. Dev, "Adversarial Learning Networks for FinTech Applications Using Heterogeneous Data Sources," in IEEE Internet of Things Journal, vol. 10, no. 3, pp. 2194-2201, 1 Feb.1, 2023, doi: 10.1109/JIOT.2021.3100742.

[25] Khowaja, S.A., Khuwaja, P. and Ismaili, I.A., 2019. A framework for retinal vessel segmentation from fundus images using hybrid feature set and hierarchical classification. Signal, Image and Video Processing, 13(2), pp.379-387

[26] Khowaja, S.A., Yahya, B.N. and Lee, S.L., 2017. Hierarchical classification method based on selective learning of slacked hierarchy for activity recognition systems. Expert Systems With Applications, 88, pp.165-177.

[27] Barshan, B. and Yu¨ksek, M.C., 2014. Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units. The Computer Journal, 57(11), pp.1649-1667.

[28] Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S.T., Tro¨ster, G., Milla´n, J.D.R. and Roggen, D., 2013. The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition. Pattern Recognition Letters, 34(15), pp.2033-2042.

[29] S. A. Khowaja, P. Khuwaja, K. Dev and M. A. Jarwar, "PROMPT: Process Mining and Paravector Tensor-Based Physical Health Monitoring Framework," in IEEE Sensors Journal, vol. 23, no. 2, pp. 989-996, 15 Jan.15, 2023, doi: 10.1109/JSEN.2022.3195613.