International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

HUMAN ACTIVITY RECOGNITION

MAHERA FATIMA[1],QAMER FATIMA[2],ABDUL ALEEM[3]
UG SCHOLAR[1&2], ASSISTANT PROFESSOR[3]
DEPARTMENT OF ECE,DECCAN COLLEGE OF ENGINEERING AND TECHNOLOGY.
AFFILIATED TO OSMANIA UNIVERSITY, HYDERABAD

ABSTRACT:

Recently, a significant amount of literature concerning machine learning techniques has focused on automatic recognition of activities performed by people. The main reason for this considerable interest is the increasing availability of devices able to acquire signals which, if properly processed, can provide information about human activities of daily living (ADL). The recognition of human activities is generally performed by machine learning techniques that process signals from wearable sensors and/or cameras appropriately arranged in the environment. Whatever the type of sensor, activities performed by human beings have a strong subjective characteristic that is related to different factors, such as age, gender, weight, height, physical abilities, and lifestyle. Personalization models have been studied to take into account these subjective factors and it has been demonstrated that using these models, the accuracy of machine learning algorithms can be improved. In this work we focus on the recognition of human activities using signals acquired by the accelerometer embedded in a smart phone. The contributions of this research are mainly three. A first contribution is the definition of a clear validation model that takes into account the problem of personalization and which thus makes it possible to objectively evaluate the performances of machine learning algorithms. A second contribution is the evaluation, on three different public datasets, of a personalization model which considers two aspects: the similarity between people related to physical aspects (age, weight, and height) and similarity related to intrinsic characteristics of the signals produced by these people when performing activities. A third and last contribution is the development of a personalization model that considers both the physical and signal similarities. The experiments show that the employment of personalization models improves, on average, the accuracy, thus confirming the soundness of the approach and paving the way for future investigations on this topic

.

INTRODUCTION

Human activity recognition (HAR) is an active research area because of its applications in elderly care, automated homes and surveillance system. Several studies has been done on human activity recognition in the past. Some of the existingwork are either wearable based [1] or non-wearable based [2] [3]. Wearable based HAR system make use of wearable sensors that are attached on the human body. Wearable based HAR system are intrusive in nature. Non-wearable based HAR system do not require any sensors to attach on the human or to carry any device for activity recognition. Non-wearable based approach can be further categorised into sensor based [2] and vision-based HAR systems [3]. Sensor based technology use RF signals from sensors, such as RFID, PIR sensors and Wi-Fi signals to detect human activities. Vision based technology use videos, image frames from depth cameras or IR cameras to classify human activities. Sensor based HAR system are non-intrusive in nature but may not provide high accuracy. Therefore, vision-based human activity recognition system has gained significant interest in the present time. Recognisinghuman activities from the streaming video is challenging. Video-based human activity recognition can be categorized as marker-based and vision-based according to motion features [4]. Marker-based method make use of optic wearable markerbasedmotion capture (MoCap) framework. It can accurately capture complex human motions but this approach has some disadvantages. It require the optical sensors to be attached on the human and also demand the need of multiple camera settings. Whereas, the vision based method make use of RGB or depth image. It does not require the user to carry any devices or to attach any sensors on the human. Therefore, this methodology is getting more consideration nowadays, consequently making the HAR framework simple and easy to be deployed in many applications. Most of the vision-based HAR systems proposed in the literature used traditional machine learning algorithms for activity recognition. However, traditional

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

machine learning methods have been outperformed by deep learning methods in recent time [5]. The most common type of deep learning method is Convolutional Neural Network (CNN). CNN are largely applied in areas related to computer vision. It consists series of convolution layers through which images are passed for processing. In this paper, we use CNN to recognize human activities from Wiezmann Dataset. We first extracted the frames for each activities from the videos. Specifically, we use *transfer learning* to get deep image features and trained machine learning classifiers. We applied 3 different CNN models to classify activities and compared our results with the existing works on the same dataset. In summary, the main contributions of our work are as follows:

1) We applied three different CNN models to classify human recognition activities and we showed the accuracy of 96.95% using VGG-16.

2) We used *transfer learning* to leverage the knowledgegained from large-scale dataset such as ImageNet [6] to the human activity recognition dataset

EXISTING SYSTEM:

The type of information used to build the classifiers divides the approach into three main categories: data-driven, knowledge-driven, and hybrid. Data-driven approaches use data mining and machine learning techniques to learn activity models. Data-driven approaches are able to handle uncertainty and temporal information. The law is that data-driven approaches require large datasets of labelled data to train classifiers. Chen and Nugent provide a recent survey of data-driven approaches. Knowledge-driven approaches use a-priori contextual information to infer the activities performed. The prior knowledge may include for example, the implicit relationships between activities, the related temporal and spatial context, and the entities involved (objects and people) .

PROPOSED SYSTEM:

This method exploits the similarity between users to weight training data and thus to improve the recognition accuracy. Unfortunately, results achieved by these researchers are not reproducible because the dataset used for the experimentation is not publicly available and moreover, the authors mainly focused on the automatic annotation of inertial signals and not classification of activities of subjects. The approach proposed by deserves further investigation and thus it has been the starting point of the research we performed and whose results are presented in this paper.

Advantages:

- ➢ A subject-dependent strategy takes advantage of this redundancy and specializes very well the classifier especially when the training set is made of only data from the subject under test.
- ➢ Hybrid approaches combine data- and knowledge driven approaches to take the advantages from each of them.

Disadvantages:

- ➢ One of the most relevant difficulty to face with new situations is due to the population diversity problem, that is, the natural differences between users' activity patterns, which implies that different executions of the same activity are different.
- ➢ This problem, activity classification models should be able to generalize as much as possible with respect to the final user and the real execution context.

LITERATURE SURVEY

IN "NONINVASIVE SENSOR BASED AUTOMATED SMOKING ACTIVITY DETECTION"AUTHORS:B. BHANDARI, J. LU, X. ZHENG, S. RAJASEGARAR, AND C. KARMAKAR Although smoking prevalence is declining in many countries, smoking related health problems still leads the preventable causes of death in the world. Several smoking intervention mechanisms have been introduced to help smoking cessation. However, these methods are inefficient since they lack in providing real time personalized intervention messages to the smoking addicted users. To address this challenge, the first step is to build an automated smoking behavior detection system. In this study, we propose an accelerometer sensor based non-invasive and automated framework for smoking behavior detection. We built a prototype device to collect data from several participants performing smoking and other five confounding activities. We used three different classifiers to compare activity detection performance using the extracted features from accelerometer data. Our evaluation demonstrates that the proposed approach is able to classify smoking activity among the confounding activities with high accuracy. The proposed system shows the potential for developing a real time automated smoking activity detection and intervention framework.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

IN "COMPRESSIVE REPRESENTATION FOR DEVICE-FREE ACTIVITY RECOGNITION WITH PASSIVE RFID SIGNAL STRENGTH"AUTHORS: L. YAO, Q. Z. SHENG, X. LI, T. GU, M. TAN, X. WANG, S. WANG,AND W. RUAN

Understanding and recognizing human activities is a fundamental research topic for a wide range of important applications such as fall detection and remote health monitoring and intervention. Despite active research in human activity recognition over the past years, existing approaches based on computer vision or wearable sensor technologies present several significant issues such as privacy (e.g., using video camera to monitor the elderly at home) and practicality (e.g., not possible for an older person with dementia to remember wearing devices). In this paper, we present a low-cost, unobtrusive, and robust system that supports independent living of older people. The system interprets what a person is doing by deciphering signal fluctuations using radio-frequency identification (RFID) technology and machine learning algorithms. To deal with noisy, streaming, and unstable RFID signals, we develop a compressive sensing, dictionary-based approach that can learn a set of compact and informative dictionaries of activities using an unsupervised subspace decomposition. In particular, we devise a number of approaches to explore the properties of sparse coefficients of the learned dictionaries for fully utilizing the embodied discriminative information on the activity recognition task. Our approach achieves efficient and robust activity recognition via a more compact and robust representation of activities. Extensive experiments conducted in a real-life residential environment demonstrate that our proposed system offers a good overall performance and shows the promising practical potential to underpin the applications for the independent living of the elderly.

IN "SPARSE COMPOSITION OF BODY POSES AND ATOMIC ACTIONS FOR HUMAN ACTIVITY RECOGNITION IN RGB-D VIDEOS"AUTHORS: I. LILLO, J. C. NIEBLES, AND A. SOTO This paper presents an approach to recognize human activities using body poses estimated from RGB-D data. We focus on recognizing complex activities composed of sequential or simultaneous atomic actions characterized by body motions of a single actor. We tackle this problem by introducing a hierarchical compositional model that operates at three levels of abstraction. At the lowest level, geometric and motion descriptors are used to learn a dictionary of body poses. At the intermediate level, sparse compositions of these body poses are used to obtain meaningful representations for atomic human actions. Finally, at the highest level, spatial and temporal compositions of these atomic actions are used to represent complex human activities. Our results show the benefits of using a hierarchical model that exploits the sharing and composition of body poses into atomic actions, and atomic actions into activities. A quantitative evaluation using two benchmark datasets illustrates the advantages of our model to perform action and activity recognition.

IN "IMAGENET: A LARGE-SCALE HIERARCHICAL IMAGE DATABASE"AUTHORS :J. DENG, W. DONG, R. SOCHER, L. LI, KAI LI, AND LI FEI-FEI

The explosion of image data on the Internet has the potential to foster more sophisticated and robust models and algorithms to index, retrieve, organize and interact with images and multimedia data. But exactly how such data can be harnessed and organized remains a critical problem. We introduce here a new database called "ImageNet", a large-scale ontology of images built upon the backbone of the WordNet structure. ImageNet aims to populate the majority of the 80,000 synsets of WordNet with an average of 500-1000 clean and full resolution images. This will result in tens of millions of annotated images organized by the semantic hierarchy of WordNet. This paper offers a detailed analysis of ImageNet in its current state: 12 subtrees with 5247 synsets and 3.2 million images in total. We show that ImageNet is much larger in scale and diversity and much more accurate than the current image datasets. Constructing such a large-scale database is a challenging task. We describe the data collection scheme with Amazon Mechanical Turk. Lastly, we illustrate the usefulness of ImageNet through three simple applications in object recognition, image classification and automatic object clustering. We hope that the scale, accuracy, diversity and hierarchical structure of ImageNet can offer unparalleled opportunities to researchers in the computer vision community and beyond.

IN "A VISION-BASED TRANSFER LEARNING APPROACH FOR RECOGNIZING BEHAVIORAL SYMPTOMS IN PEOPLE WITH DEMENTIA"AUTHORS:Z. WHARTON, E. THOMAS, B. DEBNATH, AND A. BEHERA With an aging population that continues to grow, dementia is a major global health concern. It is a syndrome in which there is a deterioration in memory, thinking, be-havior and the ability to perform activities of daily living. Depression and aggressive behavior are the most upsetting and challenging symptoms of dementia. Automatic recognition of these behaviors would not only be useful to alert family members and caregivers, but also helpful in planning and managing daily activities of people with dementia (PwD). In this work, we propose a vision-based approach that unifies transfer learning and deep convolutional

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

neural network (CNN) for the effective recognition of behavioral symptoms. We also compare the performance of state-of-the-art CNN features with the hand-crafted HOG-feature, as well as their combination using a basic linear SVM. The proposed method is evaluated on a newly created dataset, which is based on the dementia storyline in ITVs Emmerdale episodes. The Alzheimer's Society has described it as a "realistic portrayal" 1 of the condition to raise awareness of the issues surrounding dementia.

MODULES:

Human activity recognition:

Human activity recognition (HAR) is a field of research that aims at defining and experimenting new techniques able to automatically recognize human activities exploiting signals recorded by wearable and/or environmental devices. In the majority of cases, environmental devices require an installation in the home environment and devices such as cameras are perceived as intrusive devices, especially by elderly people. For these reasons, in recent years the focus has shifted to the use of wearable devices. Among them, special attention is currently being paid to smart phones, smartwatches, and fitness devices.

Knowledge driven approach:

Knowledge-driven approaches use a-priori contextual information to infer the activities performed. The prior knowledge may include for example, the implicit relationships between activities, the related temporal and spatial context, and the entities involved (objects and people) . Knowledge-driven approaches are semantically clear and easy to get started. However, they suffer in handling uncertainty and temporal information. Approaches may be further classified in logic-, ontology-, and mining-based. Examples of knowledge-driven approaches are those from and from more approaches are discussed.

Personalization in har :

Although research on activity recognition techniques from wearable devices is very active, the resulting systems are limited in their ability to generalize to new users and/or new environments, and require considerable effort and customization to achieve good performance in a real-context. One of the most relevant difficulty to face with new situations is due to the population diversity problem, that is, the natural differences between users' activity patterns, which implies that different executions of the same activity are different.

Personalization methods:

In the previous section we discussed literature methods that exploit information about the user under test to improve the accuracy of recognition algorithms. More performing models include personalization prospective both in term of physical characteristics and in term of combination of different classifiers. In this work we propose personalization models based on similarity between users in term of physical attributes and/or signals patterns. Personalization models are used to weight users training data of the classifier that in our case is the AdaBoost classifier. We demonstrate that a classifier trained on data personalized in this way is more powerful, in terms of recognition accuracy, with respect to a classifier trained without personalization.

CONCLUSION AND FUTURE WORK

Recently, a significant amount of literature concerning machine learning techniques has focused on automatic human activity recognition (HAR) by using accelerometer recorded by smartphones. Real-world HAR systems may achieve not satisfying recognition accuracy in real world applications because HAR techniques may struggle to generalize to new users and/or new environments. Several factors may affect the accuracy of activity recognition methods: i) position of the device; ii) differences between different brands of sensors; iii) human characteristics. While factors related to the position and the characteristics of the devices have been largely investigated, few works have explored the effects of human characteristics on recognition accuracy. In this paper we have experimented several personalization methods on three public datasets in order to make the results reproducible and thus allowing future research on this topic. The personalization methods experimented are based on the concept of similarity between users. This means that users may have similar physical characteristics or have similar accelerometer signals and that, such a similarity can be employed to weight training data in a way that data belonging to more similar subjects to the subject under test count more than data of less similar subjects.

REFERENCES

[1] B. Bhandari, J. Lu, X. Zheng, S. Rajasegarar, and C. Karmakar, “Noninvasive sensor based automated smoking activity detection,” in Pro-ceedings of Annual International Conference of the IEEE

Engineering in Medicine and Biology Society (EMBC). IEEE, 2017, pp. 845–848.

[2] L. Yao, Q. Z. Sheng, X. Li, T. Gu, M. Tan, X. Wang, S. Wang, and W. Ruan, "Compressive representation for device-free activity recognition with passive rfid signal strength," IEEE Transactions on Mobile Computing, vol. 17, no. 2, pp. 293–306, 2018.

[3] I. Lillo, J. C. Niebles, and A. Soto, "Sparse composition of body poses and atomic actions for human activity recognition in rgb-d videos," Image and Vision Computing, vol. 59, pp. 63–75, 2017.

[4] W. Zhu, C. Lan, J. Xing, W. Zeng, Y. Li, L. Shen, and X. Xie, "Co-occurrence feature learning for skeleton based action recognition using regularized deep lstm networks," in Thirtieth AAAI Conference on Artificial Intelligence, 2016.

[5] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

[6] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 2009, pp. 248–255.

[7] A. Jalal, N. Sarif, J. T. Kim, and T.-S. Kim, "Human activity recognition via recognized body parts of human depth silhouettes for residents monitoring services at smart home," Indoor and built environment, vol. 22, no. 1, pp. 271–279, 2013.

[8] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in Advances in neural information processing systems, 2014, pp. 568–576.

[9] G. Gkioxari, R. Girshick, and J. Malik, "Contextual action recognition with r* cnn," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1080–1088.

[10] L. Wang, Y. Xiong, Z. Wang, and Y. Qiao, "Towards good practices for very deep two-stream convnets," arXiv preprint arXiv:1507.02159, 2015.

[11] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Deep end2end voxel2voxel prediction," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2016, pp. 17–24.

[12] P. Wang, W. Li, J. Wan, P. Ogunbona, and X. Liu, "Cooperative trainingof deep aggregation networks for rgb-d action recognition," in Thirty-Second AAAI Conference on Artificial Intelligence, 2018.

[13] S. Ji, W. Xu, M. Yang, and K. Yu, "3d convolutional neural networks for human action recognition," IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 1, pp. 221–231, 2013.

[14] P. Khaire, P. Kumar, and J. Imran, "Combining cnn streams of rgb-d and skeletal data for human activity recognition," Pattern Recognition Letters, vol. 115, pp. 107–116, 2018.

[15] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks," in Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2014, pp. 1725–1732.

[16] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3d convolutional networks," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 4489–4497.

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[18] Z. Wharton, E. Thomas, B. Debnath, and A. Behera, "A vision-based transfer learning approach for recognizing behavioral symptoms in people with dementia," in 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2018, pp. 1–6.

[19] J. Cai, X. Tang, and R. Zhong, "Silhouettes based human action recognition by procrustes analysis and fisher vector encoding," in International Conference on Image and Video

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Processing, and Artificial Intelligence, vol. 10836. International Society for Optics and Photonics, 2018, p. 1083612.

[20] S. S. Kumar and M. John, "Human activity recognition using optical flow based feature set," in Proceedings of IEEE International Carnahan conference on security technology (ICCST). IEEE, 2016, pp. 1–5.

[21] W. Feng, H. Tian, and Y. Xiao, "Research on temporal structure for action recognition," in Chinese Conference on Biometric Recognition. Springer, 2017, pp. 625–632.

[22] P. Y. Han, K. E. Yee, and O. S. Yin, "Localized temporal representation in human action recognition," in Proceedings of International Confer-ence on Network, Communication and Computing. ACM, 2018, pp.261–266.