Chapter 8
An Investigation Into the Use of Deep Learning to Recognize Human Activity

Tuhin Kumar Bera
https://orcid.org/0009-0005-5428-1295
Haldia Institute of Technology, India

Pinaki Pratim Acharjya
https://orcid.org/0000-0002-0305-2661
Haldia Institute of Technology, India

ABSTRACT

Human activity recognition (HAR) is a study area concerned with the voluntary detection of routine human activity using sensor-based time-series data. However, sensor-based HAR systems have higher misclassification rates for complex actions like running, jumping, wrestling, and swinging because of sensor reading errors. The HAR system's overall performance is decreased by these sensor errors, which produce the worst classification outcomes imaginable. For complex tasks, better accuracy can be attained using vision-based HAR systems. This technique builds a deep convolutional neural network, and then uses it to extract features from the input sequence in order to gather data. Then, the temporal connections between the images will be ascertained using LSTM. This model's accuracy was suggestively higher than that of other cutting-edge deep neural network models after it was effectively validated on the UCF50 dataset. The implementation of the models to maximize their efficacy has been covered in this chapter.

DOI: 10.4018/978-1-6684-7702-1.ch008

INTRODUCTION

Human Activity Recognition (HAR) is an exciting and rapidly growing field that has gained significant attention in recent years due to its potential to impact a wide range of applications, including healthcare, sports, entertainment, and many others. The goal of HAR is to automatically recognize and classify human activities, such as walking, running, jumping, and other physical activities, from data collected from wearable sensors (Pham, C., Nguyen, T. H., Nguyen, N. H., Nguyen, D. T., & Le, T. L., 2020), cameras like Kinect (Phyo, C. N., Zin, T. T., & Tin, P., 2019), CCTV (Du, Y., Lim, Y., & Tan, Y., 2019), smartphones (Qi, J., Wang, Z., Lin, X., & Li, C., 2018) or other devices.

Human activity recognition is a crucial aspect of the development of wearable de- vices, health monitoring systems, smart homes, security systems and human-computer interaction. The widespread use of wearable sensors, such as smartwatches, fitness trackers, security cameras has created an increasing demand for reliable and accurate activity recognition systems. These systems are designed to provide valuable insights into a person's daily activity patterns, physical activity levels, and overall health and well-being (Attal, F., Mohammed, S., Dedabrishvili, M., Chamroukhi, F., Oukhellou, L., &Amirat, Y., 2015). The development of human activity recognition systems is a challenging task, as thedata collected from wearable sensors can be highly variable and complex. For ex- ample, different people may exhibit different patterns of movement when performing the same activity, and the same person's movements may vary over time, depending on factors such as their age, physical condition, and environment. Additionally, the data collected from wearable sensors may contain noise, artifacts, and other sources of interference, which can make it difficult to accurately recognize and classify human activities (Mukhopadhyay, S. C., 2015).

On the other hand, vision-based HAR systems provide more detailed and accurate information about human activity, as they capture a person's movements in 3D space (Beddiar, D. R., Nini, B., Sabokrou, M., & Hadid, H., 2020). These systems can also incorporate additional contextual information, such as the environment and objects in the scene, to provide a more complete picture of the activity being performed (Du, Y., Lim, Y., & Tan, Y., 2019). However, vision-based systems can be more complex to implement and require the use of specialized hardware, such as video cameras or depth sensors. To overcome these challenges, human activity recognition systems typically rely ondeep learning algorithms, such as neural networks, decision trees, and support vector machines. These algorithms are trained using large datasets of labelled activity data, and they learn to recognize patterns in the data that correspond to different activities (Gu, F., Chung, M.-H., Chignell, M., Valaee, S., Zhou, B.,