Research and Analysis of Human Activity Recognition based on Machine Learning

Yucheng Sheng[1*]

[1]Wilbraham and Monson Academy, Wilbraham, Massachusetts, United States

Abstract. Machine learning has been increasingly important all over the world due to its powerful abilities in numerous areas. It has shown great potential in the scientific community, such as chemistry and medicine, and other fields, like economics. So far, human activity recognition (HAR) is also becoming essential in daily lives, such as robot interactions and health monitoring devices. By combining machine learning, especially the technique of convolutional neural networks (CNNs), and HAR, the efficiency and accuracy of HAR will significantly improve, and create more opportunities for future research and development. There are already several existing areas for deep learning-based HAR, such as self-driving cars and fitness. The purpose of this paper is to organize and review the current progress on deep learning-based HAR techniques, and major applications of it in everyday lives. There are still some challenges in the current techniques, but as more studies and research are conducted in the near future, deep learning-based HAR techniques will be greatly improved and become a critical part in the real world.

1 Introduction

Machine learning, a major component of artificial intelligence, has been rapidly growing in recent decades, drawing people's attention to how machines can think faster than humans and are used in nearly every aspect of the real world. It has already demonstrated its power in numerous areas, including financial prediction, medical diagnostics, and navigational systems. As one of the major branches of machine learning, deep learning and neural networks are also powerful tools in not only these areas but also in various scientific fields [1]. Machine learning, deep learning, and neural networks are three different concepts. Artificial neural networks are a subset of deep learning, which in turn is a subset of machine learning [2]. The scientific community has shown great interest in the application of deep learning in physics, as numerous review articles on various areas of physics have been published since 2017 [1]. Consequently, it is crucial to provide general information about how machine learning holds great potential in mechanics and motion recognition.

Deep learning employs multiple levels of composition to recognize input data and model complex relationships through transformations. It differs from previous machine learning models in the complexity of each layer [3]. Deep learning is usually a simulation of the human

* Corresponding author's email: syucheng356@gmail.com

brain by social media. Although the brain indeed influences the researchers of deep learning, it is worth noting that deep learning is not simply an attempt to imitate the human brain.

Motion recognition means detecting human movements from sensor or visual data. To be more specific, movement means the trajectory and the basic movements of a person, like walking, running, sitting, standing, etc. There is already a considerable amount of research on human activity recognition (HAR) using deep learning, as it is an essential part of the study of human activities. With a deeper understanding of neural networks and significant progress in deep learning, the application of deep learning in HAR is becoming more appealing to many scientists. However, personal privacy, algorithmic bias, and other ethical considerations [4] caused by artificial intelligence are coming into people's sights. The scientific community is seeking resolutions for these concerns and is attempting to enhance the reliability of this application in various domains, such as self-driving vehicles and traffic analysis. Both topics are popular today, and the study of HAR has a significant contribution to the development of transportation. The study is important in the real world, as it enhances AI's ability to detect the movement of other objects or humans and respond promptly.

This study systematically connects and concludes research on human activity recognition, deep learning, mechanics, and data processing and analysis. In addition, it combines methods from all areas to generate an overall progress report on the application of deep learning and neural networks in human activity recognition. The article is organized in the following order, covering deep learning techniques, HAR techniques, and real-world applications.

2 Deep learning techniques

2.1 Artificial Neural Networks (ANNs)

An artificial neural network (ANN) is a mathematical model designed according to the biological structures of a neural network. A simple way to understand ANNs is to imagine numerous input data and output data that are related to some nonlinear relations [5]. An ANN can be used to predict the output accurately and also to identify the relationship between the set of inputs and outputs. Based on Yang's research on ANNs, an ANN has the potential and ability to recognize the complicated relationship between input and output data without a physical model, even though a large scale of input-output data is necessary for the model to learn the pattern between the two data sets [5, 6]. In an ANN model, artificial neurons are connected and receive different signals from other neurons to generate a unique output (Fig. 1).

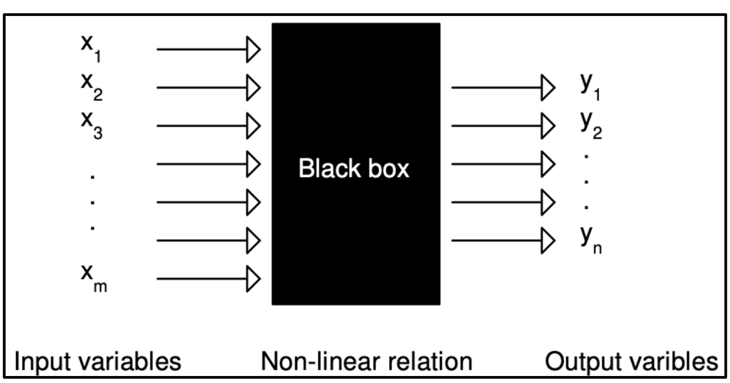

Fig. 1. Artificial neural networks (ANNs) [5]

In the process, the input variable x_i is multiplied by w_i, which is called a weight, and all the products are summed to get the net input, z. The general formula for this process is:

$$z = \sum_{i=1}^{m} w_i x_i + b \tag{1}$$

where m is the number of input variables, and b is the bias, which shifts the sum and the result of the net input.

Next, the net input z is put into an activation function, or a transfer function, f(x), to generate the actual output of the neuron. It is important to note that all the outputs of the previous layer can influence neurons in the next layer again by another transfer function. There are many transfer functions, like sigmoid, tanh, ReLU, and linear functions. Sigmoid function is a simple example, and the formula containing the net input is:

$$y = \frac{1}{1+e^{-z}} \tag{2}$$

where y is the output of the neuron. This procedure will continue within the hidden layers until the data has been processed and transferred to the output layer, which is the final layer.

The number of layers can significantly impact the model's effectiveness. Generally speaking, more layers mean better performance. ANN is a major component in HAR, providing a reliable method to analyze and predict a person's movement.

2.2 Convolutional Neural Network (CNNs)

A convolutional neural network (CNN) is similar to a traditional ANN, but with only one key difference, which is that CNNs are widely used in image recognition [7]. There are mainly three layers, including convolutional layers, pooling layers, and fully connected layers [8]. The convolutional layer plays an important role, and it is a crucial characteristic of a CNN model (Fig. 2). The image is originally transformed into a two-dimensional array of numbers and processed by a small grid of parameters called a kernel. A CNN shows great potential in image processing compared to an ANN because it is faster to process a set of input data, and it takes up less memory.

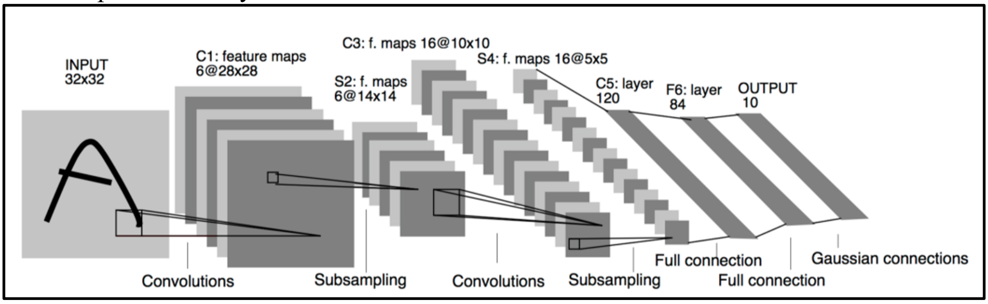

Fig. 2. Different layers and connections of convolutional neural [9]

Since a CNN is widely used in image recognition and identification, it plays an important role in HAR recognition. It is usually combined with cameras and other systems, and can identify human movements in both static and dynamic circumstances.

2.3 Radar

Radar is a useful sensing system that sends out radio waves and receives the signals that reflect from distant objects. According to Li et al., radar systems have been utilized in various applications, including satellite remote sensing and air traffic control, and they have been expanded for HAR tasks [10]. There are mainly three types of radar, including Doppler radar, Frequency-Modulated Continuous-Wave (FMCW) Radar, and Interferometry Radar [10, 11].

Radar is capable of revealing time and space changes throughout the whole process, and has both 2D and 3D maps for future analysis by deep learning.

3 Applications in the real world

3.1 Transportation and automotive systems

3.1.1 Driver activity recognition and monitoring

Accidents due to distraction have been increasing rapidly as the number of vehicles and drivers has continued to grow continuously since the beginning of the twenty-first century. Consequently, introducing a system to monitor drivers' actions during the day and night becomes very important in traffic safety. According to the research and compilation by Kidu et al. [12], a model named the long-term recurrent convolutional network (LRCN) is proposed. This driver behavior recognition system uses both RGB and IR images to minimize the adverse effects of light on data acquisition and analysis. It is possible to apply machine learning in driver action recognition tasks, and it can significantly contribute to the future improvement of traffic safety.

Additionally, by tracking movements with radar systems, driver activity recognition and its accuracy can be significantly improved.

3.1.2 Pedestrian action and recognition

There is always a large number of pedestrian casualties in traffic accidents every year, and therefore, it is beneficial if the driver or a system can know whether a person is going to cross the road. Since analysis of images is also applied in this case, a CNN model is adopted to extract different features of the pedestrian. There are many features that indicate whether a person is going to cross the road, like head orientations and step sizes. For people who are looking at vehicles that are moving toward them, it is less likely for them to cross than people who do not look at the cars and still have large step sizes (Fig. 3). This example shows two key features of pedestrians, and can be used to predict the intentions of different people on the side of the road.

Fig. 3. Pictures of different pedestrian actions. (Left: Pedestrian looking at the oncoming car. Right: Pedestrian walking across the street) [13]

With the help of machine learning, the prediction of pedestrians can be achieved. This system is crucial in the development of self-driving cars and other AI robots, but the accuracy of the system still requires future improvements.

3.2 Fitness, sports, and physical activity tracking

3.2.1 Fitness and physical activities recognition and adjustments

In recent years, wearable devices and sensors have become more and more popular in our daily lives. Many of them rely on human-computer interactions, and therefore, learning how people act has become crucial for computers. According to the review article written by Zhang et al. on the topic of sensor-based HAR techniques, the application of machine learning in human physical activities recognition can not only help with movement correction, but also with some entertainment equipment like virtual reality (VR) [14]. Moreover, CNNs can strengthen the ability to collect movement features of the mobile device and make the result more accurate.

3.2.2 Sports

Machine learning can also recognize and collect different features from complex HARs, like different sports. Wearable devices, like smart watches, are suitable for detection and analysis. In the process, the movement data is cut into fragments over a period, and the features are extracted through a CNN model. In the study by Zhuang et al., triaxial acceleration and angular velocity of different activities are collected, and the multi-dimensional data set is converted into a two-dimensional image, which can then be processed by a CNN [15].

4 Conclusion

The rapid development of machine learning techniques, especially CNNs, has brought us many opportunities in almost every area. With machine learning, human activity recognition (HAR) becomes more flexible and reliable, and its popularity reaches an all-time high. HAR, based on machine learning, has already shown great potential in medical care, transportation, fitness, and physical activity tracking.

This review has examined some basic machine learning and HAR techniques, and their applications in the real world. Recent advances demonstrate that machine learning models, including CNNs, recurrent convolutional networks, and CNN-LSTM models, significantly improve the recognition accuracy in many areas. In addition, the improvement of diverse sensing devices, such as wearable sensors, Wi-Fi signals, and radar systems, provides a strong foundation for further study of deep learning in HARs. As the techniques are gradually attracting more attention in numerous areas, the future of HAR based on machine learning is bright and clear.

However, there are still several unresolved challenges. Generalization across users, real-world environments, privacy protection, and sensor displacements is still a major limitation. Furthermore, the efficiency of calculations and energy consumption still needs a deeper investigation. After these challenges are solved, the integration between wearable devices and machine learning will expand to more new areas, and more useful functions will emerge as more research towards HAR is done. Future research directions are expected to focus on the integration of different sensors, like combining sound and images to create multi-dimensional results. Other directions are the study of privacy-preserving systems, unsupervised learning techniques, and a lighter machine learning model. By advancing these aspects, machine learning-based HAR techniques will become more adaptable, efficient, and widely adopted in real-world scenarios in the near future.

References

1. L. Herrmann and S. Kollmannsberger. Deep learning in computational mechanics: a review. Computational Mechanics 74(2): 281–331. (2024)
2. K. Sharifani, and M. Amini. Machine learning and deep learning: A review of methods and applications. World Information Technology and Engineering Journal 10(7): 3897–3904. (2023)
3. S. J. D. Prince. Understanding deep learning. MIT Press. (2023)
4. M. A. Hossen and P. E. Abas. Machine learning for human activity recognition: State-of-the-art techniques and emerging trends. Journal of Imaging 11(3): 91. (2025)
5. J. Zupan. Introduction to artificial neural network (ANN) methods: What they are and how to use them. Acta Chimica Slovenica 41(3): 327–352. (1994)
6. K. Yang. Artificial neural networks (ANNs): A new paradigm for thermal science and engineering. ASME Journal of Heat Transfer 130(9): 093001. (2008)
7. Z. Li, et al. A survey of convolutional neural networks: analysis, applications, and prospects. IEEE Transactions on Neural Networks and Learning Systems 33(12): 6999–7019. (2021)
8. R. Yamashita, et al. Convolutional neural networks: An overview and application in radiology. Insights into Imaging 9(4): 611–629. (2018)
9. J. Koushik. Understanding convolutional neural networks. arXiv preprint arXiv:1605.09081. (2016)
10. X. Li, Y. He, and X. Jing. A survey of deep learning-based human activity recognition in radar. Remote Sensing 11(9): 1068. (2019)
11. R. El Hail, et al. Radar-Based Human Activity Recognition: A Study on Cross-Environment Robustness. Electronics 14(5): 875. (2025)
12. T. Kidu et al. An intelligent real-time driver activity recognition system using spatio-temporal features. Applied Sciences 14(17): 7985. (2024)
13. D. Varytimidis, et al. Action and intention recognition of pedestrians in urban traffic. Proceedings of the 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS): 153–160. IEEE. (2018)
14. S. Zhang, et al. Deep learning in human activity recognition with wearable sensors: A review on advances. Sensors 22(4): 1476. (2022)
15. Z. Zhuang and Y. Xue. Sport-related human activity detection and recognition using a smartwatch. Sensors 19(22): 5001. (2019)