

Multimodal Deep Learning for Human Activity Recognition

Authors : Ons Slimene, Aroua Taamallah, Maha Khemaja

Abstract : In recent years, human activity recognition (HAR) has been a key area of research due to its diverse applications. It has garnered increasing attention in the field of computer vision. HAR plays an important role in people's daily lives as it has the ability to learn advanced knowledge about human activities from data. In HAR, activities are usually represented by exploiting different types of sensors, such as embedded sensors or visual sensors. However, these sensors have limitations, such as local obstacles, image-related obstacles, sensor unreliability, and consumer concerns. Recently, several deep learning-based approaches have been proposed for HAR and these approaches are classified into two categories based on the type of data used: vision-based approaches and sensor-based approaches. This research paper highlights the importance of multimodal data fusion from skeleton data obtained from videos and data generated by embedded sensors using deep neural networks for achieving HAR. We propose a deep multimodal fusion network based on a twostream architecture. These two streams use the Convolutional Neural Network combined with the Bidirectional LSTM (CNN BILSTM) to process skeleton data and data generated by embedded sensors and the fusion at the feature level is considered. The proposed model was evaluated on a public OPPORTUNITY++ dataset and produced an accuracy of 96.77%.

Keywords : human activity recognition, action recognition, sensors, vision, human-centric sensing, deep learning, context-awareness

Conference Title : ICGPCC 2023 : International Conference on Green, Pervasive and Cloud Computing

Conference Location : Venice, Italy

Conference Dates : November 13-14, 2023