Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-021-03348-w

ORIGINAL RESEARCH

RGB-D based human action recognition using evolutionary self-adaptive extreme learning machine with knowledge-based control parameters

Preksha Pareek[1] · Ankit Thakkar[1]

Received: 25 September 2020 / Accepted: 6 June 2021


Abstract
Human Action Recognition (HAR) has gained considerable attention due to its various applications such as monitoring activities, robotics, visual surveillance, to name a few. An action recognition task consists of feature extraction, dimensionality reduction, and action classification. The paper proposes an action recognition approach for depth-based input by designing Single Layer Feed forward Network (SLFN) using Self-adaptive Differential Evolution with knowledge-based control parameter-Extreme Learning Machine (SKPDE-ELM). To capture motion cues, we have used Depth Motion Map (DMM) wherein to obtain compact features, Local Binary Pattern (LBP) is applied. Thereafter, for dimensionality reduction, Principal Component Analysis (PCA) is applied to reduce the feature dimensions. For the action classification task, Extreme Learning Machine (ELM) achieves good performance for depth-based input due to its learning speed and good generalization performance. Further, to optimize the performance of ELM classifier, an evolutionary method named SKPDE is used to derive the hidden parameters of ELM classifier. The performance of the proposed approach is compared with the existing approaches Kernel ELM (KELM), L2-Collaborative Representation Classifier (CRC), and Probabilistic CRC (Pro-CRC) using datasets MSRAction3D (with 557 samples), MSRAction3D (with 567 samples), MSRDaily Activity3D, MSRGesture3D, and UTD-MHAD. The proposed approach is also statistically tested using Wilcoxon signed rank-test.

Keywords Human action recognition · Extreme learning machine · Principal component analysis · Self-adaptive differential evolution with knowledge-based control parameters · Depth-based action classification

1 Introduction

Human Action Recognition (HAR) is used to identify action type from video sequence. In a day-to-day life, various types of activities appear such as gestures, actions, and human-object interactions, or group activities (Pareek and Thakkar 2021). In an action type, activity may be performed by single person, for instance, running, having interaction between a person and an object, for example, sitting on sofa or making body gestures such as hand movement. Due to the growing need of automatic identification of human behavior in machine vision, action recognition is an important task that can be applied to different application domains such as automatic driving, Human-Computer Interaction (HCI), surveillance, elderly and youth care, assisted living, security surveillance, intelligent transportation, terrorist activity detection, crowd behavior analysis, virtual reality, gaming, to name a few (Sunny et al. 2015).

HAR focuses on obtaining and analyzing data captured from different monitoring devices such as sensors. Manual analysis of the visual information generated from different monitoring devices is a time-consuming process (Nida et al. 2019). Therefore, in the recent era of computer vision, it is imperative to apply methods to automate the process to understand the visual semantics of human actions (Nida et al. 2019). Situations involving occlusion as well as variations in size or shape of the subject(s) are challenging for action recognition (Sunny et al. 2015). To perform HAR, video sequences can be captured using various sensors such

✉ Preksha Pareek
15EXTPHDE146@nirmauni.ac.in

Ankit Thakkar
ankit.thakkar@nirmauni.ac.in

[1] Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat 382481, India

Published online: 26 June 2021