

عنوان مقاله:

A Hybrid Deep Learning Architecture Using 3D CNNs and GRUs for Human Action Recognition

محل انتشار:

ماهنامه بین المللی مهندسی، دوره 33، شماره 5 (سال: 1399)

تعداد صفحات اصل مقاله: 7

نویسندگان:

M. Savadi Hosseini - *Human-Computer Interaction lab., Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran*

F. Ghaderi - *Human-Computer Interaction lab., Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran*

خلاصه مقاله:

Video contents have variations in temporal and spatial dimensions, and recognizing human actions requires considering the changes in both directions. To this end, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) and their combinations have been used to tackle the video dynamics. However, a hybrid architecture usually results in a more complex model and hence a greater number of parameters to be optimized. In this study, we propose to use a stack of gated recurrent unit (GRU) layers on top of a two-stream inflated convolutional neural network. Raw frames and optical flow of the video are processed in the first and second streams, respectively. We first segment the video frames in order to be able to track the video contents in more details and by using 3D CNNs extract spatial-temporal features, called local features. We then import the sequence of local features to the GRU network, and use a weighted averaging operator to aggregate the outcome of the two processing flows, called global features. The evaluations confirm acceptable results for the two HMDB51 and UCF101 datasets. The proposed method resulted in a 1.6% improvement in the classification accuracy of the HMDB51 challenging dataset compared to the best reported results.

کلمات کلیدی:

Inflated convolutional neural networks, recurrent gate unit, action recognition, two-stream architecture

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1021750>

