CSET 2024

The 2nd International Conference on Sustainability and Emerging Technologies

DONG NAI TECHNOLOGY UNIVERSITY

https://dntu.edu.vn/

206 Nguyen Khuyen Str.,Quarter 5, Trang Dai Ward, Bien Hoa City, Dong Nai Province

CSET 2024

The 2nd International Conference on Sustainability and Emerging Technologies

(CSET-24-034

EFFICIENT MULTI-PERSON ACTION RECOGNITION USING YOLOV7-POSE AND DEEP LEARNING MODELS

Trinh Dinh Thang[1*], Hamka Mudin Parah[2], Nguyen Khanh An[3], Nguyễn Đức Mạnh[1]

1 Faculty of Information Technology, Dong Nai Technology University, Bien Hoa City, Vietnam

2 Faculty of Engineering, Padang State University, Kota Padang, Padang City, Indonesia

3 Binh Thuan College, Phan Thiet City, Vietnam

*Corresponding author: trinhdinhthang@dntu.edu.vn

ABSTRACT

In this paper, we propose a highly efficient methodology for multi-person action recognition. Our approach combines the rapid pose estimation capabilities of YOLOv7-Pose— an advanced model that overcomes traditional speed limitations associated with pose estimation such as OpenPose and PoseNet— with deep learning models like Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Spatial Temporal-Graph Convolution Network (ST-GCN) to classify actions; it identifies and analyzes the actions of multiple individuals within a single scene simultaneously. We experiment, analyze, and compare the performance of these models in terms of accuracy, testing time, and computational complexity. Our results underscore that combining YOLOv7-Pose with ST-GCN attains superior accuracy; meanwhile, employing YOLOv7-Pose alongside LSTM demonstrates the quickest testing time. This underscores how effectively our proposed methodology balances precision and efficiency in multi-person action recognition tasks.

Keywords: *Deep learning, LSTM, Multi-person action recognition, ST-GCN, YOLOv7-Pose…*

CSET 2024

The 2nd International Conference on Sustainability and Emerging Technologies

THANK YOU TO ALL PARTICIPANTS

DONG NAI TECHNOLOGY UNIVERSITY

DNTU