Multi-Frame Skeleton-based Fall Detection on RGB Videos

Hamza Ergüder[1], Tuncay Uzun[2]

[1]Yildiz Technical University, Graduate School of Science and Engineering, Department of Electronics and Communication Engineering, Istanbul, Turkey.
[2]Yildiz Technical University, Faculty of Electrical & Electronics, Department of Electronics and Communication Engineering, Istanbul, Turkey.
ORCID Code: [1]0000-0003-4403-9364, [2]0000-0002-4145-0468
* e-mail: [1]hamza.erguder@std.yildiz.edu.tr, [2]uzun@yildiz.edu.tr, mobile phone numbers: [1]+90-536-796-10-00, [2]+90-535-658-70-17

ABSTRACT

Falls are common among the elderly, with one in four people over the age of 65 visiting the emergency department each year due to a fall. Early detection of falls can prevent physical and psychological trauma and potentially save lives. This study presents the development of a video-based algorithm for detecting falls. The UP-Fall dataset, which includes a total of 11 activities, 5 of which are falls, was used and 10-second videos were selected for each action. For high accuracy, state-of-the-art techniques were employed at each step. The YOLO-V7 object recognition algorithm was used to detect the patient in the image. The BoT-SORT tracking algorithm was utilized to follow the same person in the video. The ViT-Pose image pose estimation algorithm was employed to determine the person's joints. From these results, a dataset was prepared to detect the time-dependent changes of the person's bounding box and joints. The dataset consists of 190 (frames) x 18 x 3 (consisting of locations and scores of bounding box, and 17 sets of joint) matrices for each action. The matrices were labeled as "fallen" or "unfallen," indicating whether or not a fall had occurred. An LSTM algorithm was developed, trained and evaluated on this dataset. An accuracy of 92% was obtained in fall detection. In this way, video-based detection was carried out which enables the fall detection on multi-frame. The results demonstrate successful performance in terms of accuracy and the ability to detect falls in multiple frames, compared to other studies in the literature.

Keywords: Computer Vision, Fall Detection, LSTM.