International Journal of Recent Engineering Research and Development (IJRERD)
ISSN: 2455-8761
www.ijrerd.com || Volume 09 – Issue 03 || May - Jun 2024 || PP. 181-184

Human Fall Detection Using Yolo V3

Abirami R, Priyadharshini A, Yogeshwari R
UG Scholar, Department of CSE, A.R.J College of Engineering and Technology, Mannargudi, Thiruvarur, Tamilnadu-614 001, India

Kavitha R,
Assistant Professor, Department of CSE, A.R.J College of Engineering and Technology, Mannargudi, Thiruvarur, Tamilnadu-614 001, India

Abstract: The increasing population of elderly individuals has led to a rise in incidents of falls, both at home and outdoors, emphasizing the importance of effective fall detection systems for their health and safety. However, existing methods, particularly those based on video monitoring, often suffer from complexity and redundancy, hampering real-time detection and accuracy. To address these challenges, this paper proposes a novel fall detection method leveraging video data in complex environments to enhance accuracy and speed of detection. The proposed method introduces the YOLO network model as the foundation of the detection algorithm, offering a streamlined approach to identifying fall events. A human fall detection dataset is curated following the format of the Pascal VOC dataset, providing essential training data for the model. Subsequently, the algorithm is fine-tuned and trained on a GPU-based deep learning server to optimize its performance. Evaluation of the proposed YOLO-based model against other detection algorithms demonstrates its superior recognition capabilities. The comparison highlights the efficacy of the developed approach in accurately identifying fall incidents in diverse environments. By combining advanced deep learning techniques with optimized model training, this method offers a promising solution for enhancing fall detection systems and safeguarding the well-being of the elderly population.

Keywords: YOLO network model, the Pascal VOC dataset, GPU-based deep learning server

1. Introduction

The demographic shift towards an aging population has brought attention to the pressing need for effective fall detection systems to ensure the well-being of elderly individuals. Falls among the elderly can lead to serious injuries and complications, impacting their quality of life and imposing significant burdens on healthcare systems. Consequently, there is a growing emphasis on developing advanced technologies to detect and respond to fall events promptly. While various approaches have been explored, many existing methods, particularly those reliant on video monitoring, face challenges related to complexity and accuracy. Current Fall Detection Challenges: Traditional fall detection methods often rely on video monitoring systems, which pose several challenges. These systems tend to be cumbersome, requiring complex setups and extensive computational resources. Moreover, their effectiveness can be compromised by factors such as poor lighting conditions, occlusions, and background clutter in real-world environments. Additionally, existing algorithms may struggle to distinguish between normal activities and fall events accurately, leading to false alarms or missed detections. As a result, there is a pressing need for innovative approaches that can address these limitations and improve the efficiency and accuracy of fall detection systems. Proposed Approach: To address the shortcomings of existing fall detection methods, this paper proposes a novel approach based on leveraging video data in complex environments. Central to this approach is the adoption of the YOLO network model, renowned for its efficiency and effectiveness in object detection tasks. By harnessing the capabilities of deep learning, the proposed method aims to streamline the fall detection process, enabling more accurate and timely identification of fall events. Furthermore, the development of a dedicated human fall detection dataset, modeled after the widely used Pascal VOC dataset, provides essential training data to enhance the model's performance. Through optimization and training on GPU-based deep learning servers, the algorithm is fine-tuned to achieve optimal results in real-world scenarios.

2. Literature Survey

Gang Li, et.al(2020), Learning skeleton information for human action, Human action recognition plays an important role in modern intelligent systems, such as human–computer interaction (HCI), sport analysis, and somatosensory game. Compared with conventional 2-D based human action analysis, using Kinect sensor can obtain depth information of human action, which is significant for human action recognition. In this paper, we propose a joint angle sequence model for recognizing human actions, where depth images are acquired by using

Kinect sensor. We design an improved DTW method to improve the matching accuracy. Comprehensive experiments show the effectiveness and robustness of our proposed method.

Preksha Pareek, et,al (2021), A survey on video-based Human Action Recognition, Human Action Recognition (HAR) involves human activity monitoring task in different areas of medical, education, entertainment, visual surveillance, video retrieval, as well as abnormal activity identification, to name a few. Due to an increase in the usage of cameras, automated systems are in demand for the classification of such activities using computationally intelligent techniques such as Machine Learning (ML) and Deep Learning (DL). In this survey, we have discussed various ML and DL techniques for HAR for the years 2011–2019. The paper discusses the characteristics of public datasets used for HAR. It also presents a survey of various action recognition techniques along with the HAR applications namely, content-based video summarization, human–computer interaction, education, healthcare, video surveillance, abnormal activity detection, sports, and entertainment. The advantages and disadvantages of action representation, dimensionality reduction, and action analysis methods are also provided. The paper discusses challenges and future directions for HAR.

Farhat Afza, et.al(2021) A framework of human action recognition In this article, we implement an action recognition technique based on features fusion and best feature selection. In the proposed method, HSI color transformation is performed in the first step to improve the contrast of video frames and then extract their motion features by optical flow algorithm. The frames fusion approach extracts the moving regions that find out by optical flow. After that, extract shape and texture features fused by a new parallel approach name length control features. A new Weighted Entropy-Variances approach is applied to a combined vector and selects the best of them for classification. Finally, features are passed in M-SVM for final features classification into relevant human actions. The experimental process is conducted in four famous action datasets- Weizmann, KTH, UCF Sports, and UCF.

3. Proposed System

The proposed fall detection system leverages the YOLO network model to enhance the accuracy and efficiency of detecting falls in complex environments, specifically tailored to address the needs of the elderly population. The system's foundation lies in the construction of a human fall detection dataset adhering to the Pascal VOC format, facilitating the extraction of relevant features crucial for identifying fall events. This dataset undergoes preprocessing and segmentation to optimize feature extraction, ensuring that the YOLO model receives high-quality training data. Through the utilization of a GPU deep learning server, the YOLO network model is meticulously optimized and trained on the dataset, honing its ability to accurately detect human falls

4. System Architecture

A system architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.

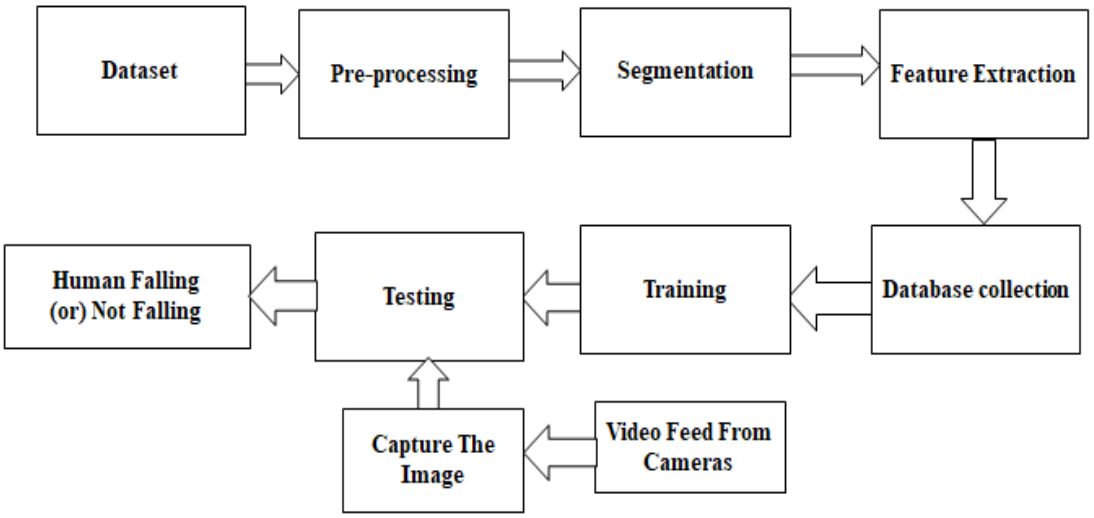

Figure: Block diagram of System Architecture

5. Implementation

DATA COLLECTION: There are two class of images has been collected in the module. One class is human fall detected and another one is not human fall. Each and very class has 150 + images.

IMAGE LABELING: Label IMG is open source tool which is used to label the image. Each and very images are labeled by using this tool.

YOLO TRAINING: In this module, darknet frame work and darknet50 has been utilized in order to train the model. The overall architecture of the yolo algorithm is presented figure 3.2.3. Number of epochs is 4000. After complete the training process the weights file will generated and stored.

FRAME CONVERSION: Captured video is converted into the frames in the module.

HUMAN FALL DETECTION: In this module, the webcam capture is given to the input of the deployed model. The trained model process 45 frames per second. If the model detect the human fall, which will sent the massage and show in terminal.

6. Results

The human fall detection system can incorporate the use of different sensors, such as a thermal camera installation to monitor the activities of older adults in lightly lit environments and at night or a fisheye lens installation to expand the detection range. In addition, different shooting angles and complex occlusion situations can be further evaluated.

Figure 2: Fall Detect

7. Conclusion

In the project, the human fall detection system with mail alert has been implemented. Transfer learning method based training process reduces the training time of the proposed method. The loss function of the proposed method is 0.245.Growing demand on services oriented to elderly makes justified the development of improved system to help elderly live longer in their home increasing their quality of life. The product presented here represents an important step beyond the actual state of the art in services to elderly. Indeed, the service offers complete activity monitoring, automatic human fall detection and user localization on a small autonomous mobile module both for indoor and outdoor use. The system, composed by a mobile module worn by the user and a call centre to analyze and save the information, has been developed as easy to use and reliable, and final user requirements have been taken into account on every stage of the development. The first tests and validations, both realized in laboratory conditions as well as with final users (elderly in gerontology centre) show that the system meets well the requirements, is reliable (above 90 % of human fall detection) and well accepted by final users.

8. References

[1]. N. Spolaôr, et al., A systematic review on content-based video retrieval, Eng. Appl. Artif. Intell. 90 (2020) 103557.
[2]. A. Keshavarzian, S. Sharifian, S. Seyedin, Modified deep residual network architecture deployed on serverless framework of IoT platform based on human activity recognition application, Future Gener. Comput. Syst. 101 (2019) 14–28.
[3]. A.D. Antar, M. Ahmed, M.A.R. Ahad, Challenges in sensor-based human activity recognition and a comparative analysis of benchmark datasets: A review, in: 2019 Joint 8th International Conference on

Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition, IcIVPR, IEEE, 2019.

[4]. K.A. da Costa, et al., Internet of things: A survey on machine learning-based intrusion detection approaches, Comput. Netw. 151 (2019) 147–157.

[5]. J.K. Aggarwal, M.S. Ryoo, Human activity analysis: A review, ACM Comput. Surv. 43 (3) (2011) 1–43.

[6]. S. Pirbhulal, et al., Mobility enabled security for optimizing IoT based intelligent applications, IEEE Netw. 34 (2) (2020) 72–77.

[7]. B. Ali, et al., A volunteer supported fog computing environment for delay-sensitive IoT applications, IEEE Internet Things J. (2020).

[8]. S. Zhao, et al., Pooling the convolutional layers in deep convnets for video action recognition, IEEE Trans. Circuits Syst. Video Technol. 28 (8) (2017) 1839–1849.

[9]. R. Girdhar, et al., Actionvlad: Learning spatio-temporal aggregation for action classification. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.

[10]. R. Hou, C. Chen, M. Shah, An end-to-end 3d convolutional neural network for action detection and segmentation in videos, 2017, arXiv preprint arXiv:1712.01111.