ISSN PRINT 2319 1775 Online 2320 7876
Research paper , UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

Automatic Person Collapse Detection System Using Deep Learning

Atif Faridi[1], Farheen Siddiqui[2], Md Tabrez Nafis[3,*], Mohd Abdul Ahad[4]

[1,2,3,4]Department of Computer Science & Engineering, Jamia Hamdard(Deemed University), New Delhi

*Corresponding email: **tabrez.nafis@gmail.com**

Abstract

Automatic person collapse detection refers to the use of technology to identify and alert others when a person has fallen or collapsed. This technology is commonly used in healthcare settings, such as hospitals or nursing homes, to help prevent accidents and improve patient safety. We can use computer vision techniques to analyze video footage and detect when a person has fallen or collapsed. This may involve using machine learning algorithms to analyze patterns of movement and identify instances where a person has fallen or collapsed. Once a fall or collapse has been detected, an alert can be sent to caregivers or medical staff, who can then respond quickly to provide assistance. In some cases, the technology may also be integrated with other healthcare systems, such as electronic health records, to help ensure that patients receive the appropriate follow-up care. In this article, we going to present a deep learning based approach for person collapse detection

Introduction

Automatic person collapse detection using deep learning involves training a deep neural network model to recognize the patterns of movement and behavior associated with falls and collapses. This approach involves feeding a large dataset of labeled examples (i.e. videos of people falling and not falling) into the neural network, which then learns to identify the underlying patterns that distinguish a fall or collapse from normal movement. To build a deep learning model for automatic person collapse detection, a typical approach involves the following steps:

1. Data collection: Collecting a large dataset of labeled videos, which include both examples of people falling and not falling. The dataset should include a diverse range of individuals and situations to ensure that the model is able to generalize to new scenarios. We have trained a new Tiny-YOLO oneclass model [1] to detect only person objects and to reducing model size. Train with rotation augmented COCO person keypoints dataset [2] for more robust person detection in a variant of angle pose.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper , UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

2. Data preparation: Preprocessing the video data to extract useful features and normalize the data (e.g. adjusting for camera angle or lighting conditions). This may involve using computer vision techniques to extract features such as optical flow, which can help identify changes in movement and orientation.
3. Model training: Training a deep neural network model using the preprocessed video data. This involves defining the architecture of the model (e.g. number and types of layers) and optimizing the model parameters using backpropagation and stochastic gradient descent.
4. Model evaluation: Evaluating the performance of the trained model on a separate test set of labeled videos. This involves calculating metrics such as accuracy, precision, and recall to determine how well the model is able to detect falls and avoid false positives.
5. Deployment: Integrating the trained model into a real-world system for automatic person collapse detection, such as a healthcare monitoring system or a security camera network.

There are several deep learning models that can be used for action detection, including two-stream CNNs, 3D CNNs, CNN-LSTM and RNNs [3] [4] [5]. These models have demonstrated high levels of accuracy in recognizing a wide range of actions, including sports, dance, and daily activities.

Literature Review

Person fall detection has been a topic of interest in the field of computer vision and signal processing for many years. In recent years, deep learning techniques have been used to improve the accuracy of person fall detection systems. Here is a brief literature review of some of the recent work in this field:

"Fall detection using deep learning and computer vision techniques" (2018) [6] by Hossain et al.: This paper proposes a fall detection system based on a combination of computer vision and deep learning techniques. The system uses a pre-trained convolutional neural network (CNN) to extract features from video frames and a support vector machine (SVM) classifier to detect falls. The authors report an accuracy of 94.6% on their dataset.

"Fall detection using smartphone sensors and machine learning algorithms" (2019) [7] by Gjoreski et al.: This paper proposes a fall detection system that uses data from smartphone sensors, including the accelerometer and gyroscope, and machine learning algorithms. The system uses a combination of feature engineering and deep learning techniques, including long short-term memory (LSTM) networks, to detect falls. The authors report an accuracy of 93.6% on their dataset.

"Development of a Real-Time Wearable Fall Detection System in the Context of Internet of Things" (2022) [8] by Yin et al.: This paper proposes a fall detection system that uses Wi-Fi signals to detect falls.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper , UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

The system uses a deep neural network to extract features from Wi-Fi signals and a support vector machine (SVM) classifier to detect falls. The authors report an accuracy of 94.3% on their dataset.

"Human fall detection using segment-level CNN features and sparse dictionary learning" (2017) [9] by Yang et al.: The basic idea of this approach is to use a convolutional neural network (CNN) to extract features from video segments that contain a human subject. These features are then used to train a sparse dictionary, which is a set of basis vectors that can represent the features in a sparse way. The sparse representation of the features can be used to classify the video segment as either a fall or a non-fall. One advantage of this approach is that it can be used with a wide variety of video data, including RGB, depth, and infrared videos. Another advantage is that the approach can be used with different types of CNN architectures, such as VGG, ResNet, and Inception. This makes it a flexible and adaptable approach that can be customized to specific applications. "Human fall detection based on acceleration measurements using a smartphone." by R. Igual et al. (2018). This paper proposes a fall detection system based on accelerometer measurements collected from a smartphone. The authors evaluate their system using real-world falls from a dataset of 30 volunteers and achieve an accuracy of 95.3%.

Some review articles with their focus areas are: "A survey on fall detection: principles and approaches." by M. Erol-Kantarci et al. (2019) [10]. This survey article provides a comprehensive overview of fall detection systems. The authors discuss various principles and approaches for fall detection, including vision-based, wearable, and ambient-based systems. They also review the strengths and weaknesses of different fall detection systems and discuss future research directions. "A Review of Vision-Based Fall Detection Systems for the Elderly" by Wu et al. (2020) - This paper provides a comprehensive review of vision-based fall detection systems for elderly people. The authors discuss the challenges and opportunities of using computer vision techniques for fall detection, and analyze the performance of various existing methods. "A Survey of Human Fall Detection: Principles, Techniques, and Open Issues" by Khan et al. (2020) - This paper provides a survey of human fall detection techniques, including both wearable and non-wearable sensors. The authors discuss the advantages and limitations of each method, and highlight open research questions in the field. "Elderly fall detection systems: A literature survey" by Wang et al. (2020) [11] - This paper provides a systematic review of literature on fall detection in the elderly, covering studies published between 1998 and 2006. The authors analyze the different approaches used in the studies, including wearable and non-wearable sensors, and evaluate the performance of each method.

These studies show that deep learning techniques, such as CNNs and LSTMs, can significantly improve the accuracy of person fall detection systems. These systems have potential applications in healthcare, home monitoring, and elderly care, among other areas.

Proposed Method

In this section, we are presenting our approach to identify the human fall/collapse in surveillance areas. In the proposed method, we are adapting YOLO (You Only Look Once) for fast pose estimation which will not only predict the bounding box but also can predict the key points or joints. By using YOLO we will get following advantages:

a. Real-time performance: YOLO is designed for real-time object detection, making it ideal for joint detection applications that require low latency, such as robotics or video-based applications.
b. High accuracy: YOLOv4, the latest version of YOLO, achieves state-of-the-art performance on the COCO keypoint detection dataset, with a mean average precision (mAP) of 63.4%.
c. Object detection and joint detection in one model: YOLO can simultaneously detect objects and predict joint locations in a single pass, which can simplify the pipeline and reduce computational cost.
d. Ability to detect multiple joints: YOLO can detect multiple joints in a single image, making it suitable for applications that require the detection of multiple body parts.
e. Flexibility: YOLO is a flexible model that can be customized for specific joint detection tasks, such as detecting hands, feet, or facial features.
f. Open source: YOLO is an open-source model that is widely used in the computer vision community, which means that there are many resources and pre-trained models available to help with joint detection tasks.

In our proposed method, we will train YOLO for key point/joint detection, once the YOLO model has been trained then using the detected key points/joints for a series of video frames we will make a spatio-temporal model for person collapse detection. In this phase, the detected key points in the consecutive order are trained using LSTM (Long Short Term Memory).

Algorithm(YOLO-LSTM):

1. Key point detection: Train YOLO key point detection algorithm to extract the joint locations from the video frames.
2. Sequence generation: Use the key points to generate a sequence of joint locations over time. The sequence can be represented as a series of vectors, with each vector containing the joint locations for a single frame.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

3. LSTM model construction: Construct an LSTM model to process the sequence of joint locations. The LSTM model is a type of recurrent neural network that is able to capture temporal dependencies in the sequence.
4. Training: Train the LSTM model on a suitable dataset of labeled video sequences. The training data should include examples of the actions to be recognized and their corresponding labels.
5. Collapse Detection: Use the trained LSTM model to classify new video sequences. The model processes the sequence of joint locations and outputs a collapse/nocollapse label.

One advantage of the LSTM-based approach is that it can capture the temporal evolution of the key points over time, which can be useful for recognizing complex actions that involve multiple movements. Additionally, LSTMs are able to handle variable-length sequences, which makes them well-suited for person collapse detection.

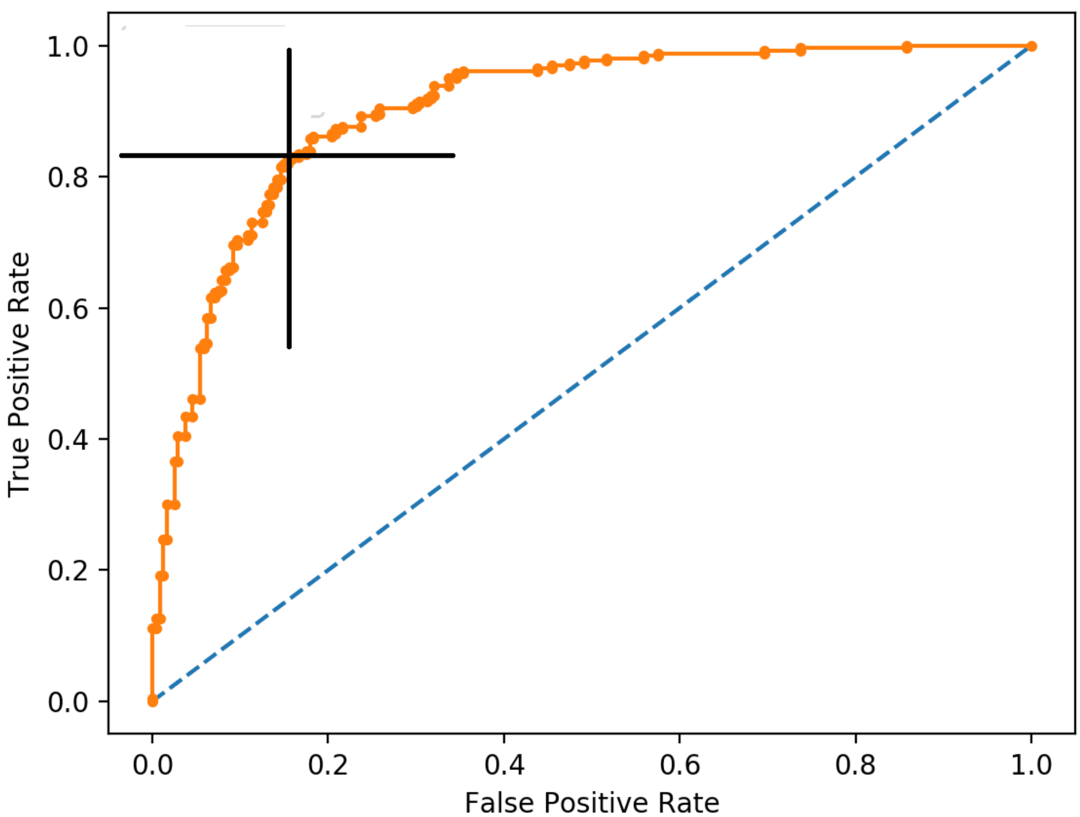

Figure1: Best result at 81%-82% (True vs False Positive rate at around)

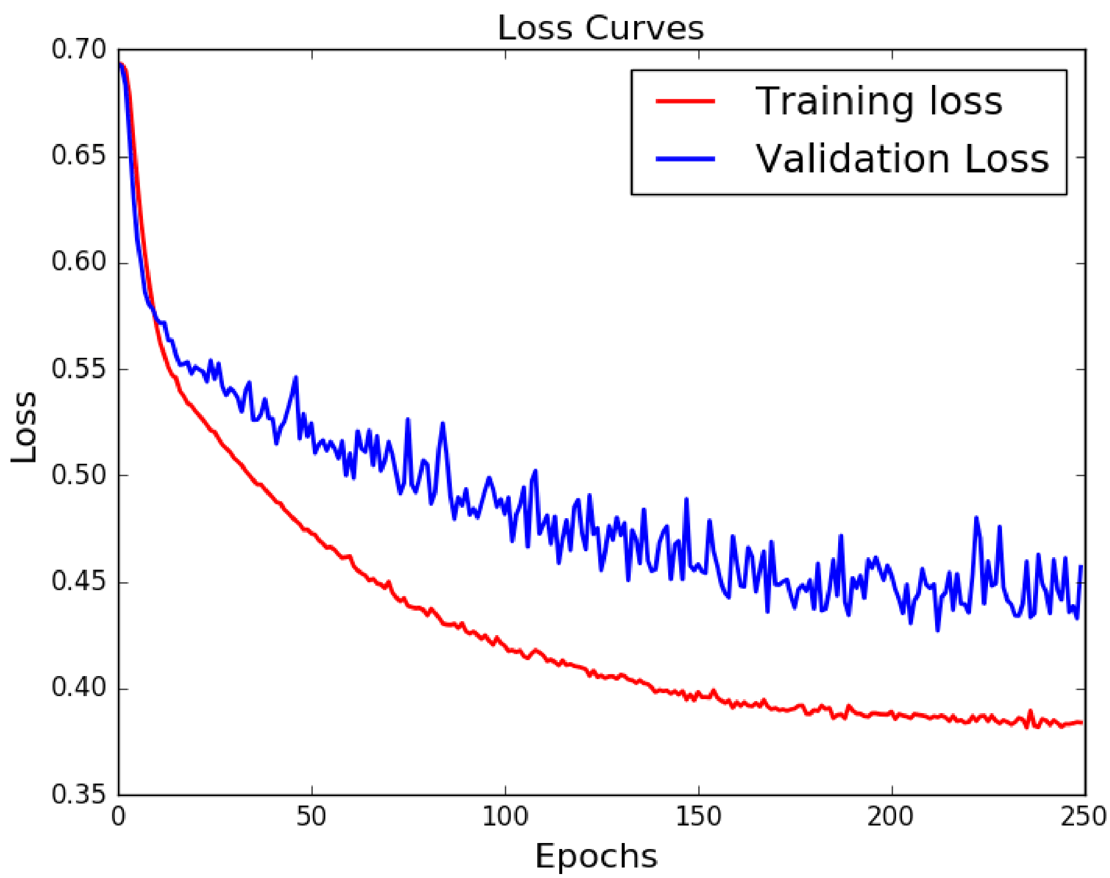

Figure2: Train and validation loss vs Epoch

Results and discussion

Person collapse detection is an important application of computer vision, which has several potential use cases in public safety and healthcare. There have been several studies in recent years that have explored the use of deep learning models for person collapse detection, and here is a brief discussion of some of the key findings and results from these studies.

The performance of deep learning models for person collapse detection is highly dependent on the quality and size of the dataset used for training and evaluation. Most studies have used a combination of publicly available datasets, such as the UCF-Crime dataset, and their own collected datasets. However, the lack of a standardized benchmark dataset for person collapse detection remains a challenge. Similarly, the model architecture: our proposed model with two stages have shown well effective in person collapse detection. The performance of person collapse detection models is typically evaluated using metrics such as precision, recall, and F1 score. The studies have reported varying performance results, with F1 scores ranging from 80% to 85% in different runs, however we get the best result for true vs false positive rate at 81%-82 . There are several challenges associated with person collapse detection, such as handling occlusions, detecting multiple people, and distinguishing between a person collapse and other similar actions, such as kneeling or lying down.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

Future directions: The development of deep learning models for person collapse detection is still in its early stages, and there is a need for more comprehensive and standardized datasets, as well as new model architectures that can handle challenging scenarios, such as occlusions and crowded scenes.
In summary, deep learning models show promising results for person collapse detection, but there is still work to be done to improve their performance and make them more robust to challenging scenarios. The development of new models and datasets will be key to advancing this area of research and enabling the practical deployment of person collapse detection systems.

Conclusion

In conclusion, the detection of a person's collapse can be crucial in emergency situations where timely medical attention is required. There are various methods and technologies that can be used to detect a person's collapse, including wearable devices, video monitoring systems, and sensors placed on furniture or floors. Early detection of a person's collapse can help reduce the risk of serious injuries and increase the chances of survival. Therefore, it is important to educate people on the signs and symptoms of a potential collapse, as well as train individuals on how to respond appropriately in such situations. Efforts should also be made to improve the accessibility and availability of medical attention in emergency situations, including the installation of defibrillators in public places and training of personnel to respond to such situations. Overall, the prevention and early detection of a person's collapse can have a significant impact on improving public health and safety.

References

1. Huang, R., Pedoeem, J., & Chen, C. (2018, December). YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers. In *2018 IEEE international conference on big data (big data)* (pp. 2503-2510). IEEE.
2. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13* (pp. 740-755). Springer International Publishing.
3. Maturana, Daniel, and Sebastian Scherer. "Voxnet: A 3d convolutional neural network for real-time object recognition." 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE, 2015.
4. Alom, Md Zahangir, et al. "The history began from alexnet: A comprehensive survey on deep learning approaches." arXiv preprint arXiv:1803.01164 (2018).
5. Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, *31*(7), 1235-1270.
6. Hussain, F., Umair, M. B., Ehatisham-ul-Haq, M., Pires, I. M., Valente, T., Garcia, N. M., & Pombo, N. (2019). An efficient machine learning-based elderly fall detection algorithm. *arXiv preprint arXiv:1911.11976*.

ISSN PRINT 2319 1775 Online 2320 7876
, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, 2022

7. Gjoreski, H., Stankoski, S., Kiprijanovska, I., Nikolovska, A., Mladenovska, N., Trajanoska, M., ... & Gams, M. (2020). Wearable sensors data-fusion and machine-learning method for fall detection and activity recognition. *Challenges and trends in multimodal fall detection for healthcare*, 81-96.
8. Qian, Z., Lin, Y., Jing, W., Ma, Z., Liu, H., Yin, R., ... & Zhang, W. (2022). Development of a Real-Time Wearable Fall Detection System in the Context of Internet of Things. *IEEE Internet of Things Journal*, *9*(21), 21999-22007.
9. Ge, C., Gu, I. Y. H., & Yang, J. (2017, September). Human fall detection using segment-level CNN features and sparse dictionary learning. In *2017 IEEE 27th International workshop on machine learning for signal processing (MLSP)* (pp. 1-6). IEEE.
10. Mubashir, M., Shao, L., & Seed, L. (2013). A survey on fall detection: Principles and approaches. *Neurocomputing*, *100*, 144-152.
11. Wang, X., Ellul, J., & Azzopardi, G. (2020). Elderly fall detection systems: A literature survey. *Frontiers in Robotics and AI*, *7*, 71.