

Received July 29, 2020, accepted August 31, 2020, date of publication September 4, 2020, date of current version September 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3021943

# Deep Learning Based Systems Developed for Fall Detection: A Review

MD. MILON ISLAM<sup>1</sup>, OMAR TAYAN<sup>2,3</sup>, MD. REPON ISLAM<sup>1</sup>,  
MD. SAIFUL ISLAM<sup>1</sup>, SHEIKH NOORUDDIN<sup>1</sup>,  
MUHAMMAD NOMANI KABIR<sup>4</sup>, AND MD. RABIUL ISLAM<sup>5</sup>

<sup>1</sup>Department of Computer Science and Engineering, Khulna University of Engineering & Technology, Khulna 9203, Bangladesh

<sup>2</sup>Department of Computer Engineering, Taibah University, Medina 41411, Saudi Arabia

<sup>3</sup>IT Research Center for the Holy Quran (NOOR), College of Computer Science and Engineering, Taibah University, Medina 41411, Saudi Arabia

<sup>4</sup>Faculty of Computing, Universiti Malaysia Pahang, Gambang 26300, Malaysia

<sup>5</sup>Department of Electrical and Electronic Engineering, Bangladesh Army University of Engineering and Technology, Natore 6431, Bangladesh

Corresponding author: Omar Tayan (otayan@taibahu.edu.sa)

This work was supported in part by the NOOR Research Center, Taibah University, Madinah, Saudi Arabia, under Grant NRC1-126B.

**ABSTRACT** Accidental falls are a major source of loss of autonomy, deaths, and injuries among the elderly. Accidental falls also have a remarkable impact on the costs of national health systems. Thus, extensive research and development of fall detection and rescue systems are a necessity. Technologies related to fall detection should be reliable and effective to ensure a proper response. This article provides a comprehensive review on state-of-the-art fall detection technologies considering the most powerful deep learning methodologies. We reviewed the most recent and effective deep learning methods for fall detection and categorized them into three categories: Convolutional Neural Network (CNN) based systems, Long Short-Term Memory (LSTM) based systems, and Auto-encoder based systems. Among the reviewed systems, three dimensional (3D) CNN, CNN with 10-fold cross-validation, LSTM with CNN based systems performed the best in terms of accuracy, sensitivity, specificity, etc. The reviewed systems were compared based on their working principles, used deep learning methods, used datasets, performance metrics, etc. This review is aimed at presenting a summary and comparison of existing state-of-the-art deep learning based fall detection systems to facilitate future development in this field.

**INDEX TERMS** Auto-encoder, convolutional neural network, deep learning, fall detection, long short-term memory, recurrent convolutional network, recurrent neural network, review, surveillance system.

## I. INTRODUCTION

Falls are a major cause of serious injuries for the elderly population worldwide. Falls impede their comfortable and independent living. Statistics show that falls are the primary reason behind injury-related death for older aged 80 or more. According to a study performed by the United Nations [1], in 2017, the number of elderly people aged 60 or above were 962 million (13% of the total population of the world). According to the research of World Health Organization (WHO), the number of elderly people may be about 1.2 billion by 2025 in the whole world [2]. This number is expected to more than double by 2050 (2.1 billion) and triple (3.1 billion) by 2100 [1]. Every year, around 2.8 million older people are associated with emergency health issues

The associate editor coordinating the review of this manuscript and approving it for publication was Ah Hwee Tan<sup>id</sup>.

including fall injuries. Adults aged over 65 are more prone to sustaining life-threatening injuries caused by falls. For adults aged over 85, falls were the primary cause behind approximately two-thirds of all reported injury-related deaths [3]. Around 20% falls result in fatal injuries such as hip bone fracture, head trauma, etc. [3]. In USA alone, Nationwide, 29,668 (61.6 per 100,000) U.S. residents aged over 65 years died from injuries related to fall events in 2016 [4].

Generally, most of the falls occur at home due to potential health hazards in the living room [3], [4]. The common hazards include poor lighting, clutter, obstructed ways, slippery floors, pets and unstable furniture [5], [6]. The elderly suffering from neurological diseases such as dementia [7]–[9] and epilepsy [10], [11] are prone to fall and fall related casualties than the average elderly population. The tendency of independent living of the elderly population separate from their family members in western cultures is also a reason

behind fall-related casualties. The fall events themselves are not life threatening, in most cases. But the falls in cluttered environments lead to concussions, hemorrhage, and other severe health-risks that lead to unfortunate deaths. Due to independent living, in absence of fall detection technologies, emergency services cannot respond to the fall events in time. This often leads to severe consequences. Many supported surveillance systems are developed for filling up the necessity of the presence of nurses and support staff at all times.

It is difficult to make the environment completely fall-proof. Thus, fall detection and rescue services can ensure the safety of the elderly population. It has become a concern to develop intelligent detection and prevention system. Many developed fall detection and monitoring systems have been reviewed in [12]–[23] with various categories and perspectives. The challenges, issues and, advances related to fall detection and fall prevention are discussed in [24]–[27].

Detection is the task of detecting the presence of a particular incident or object in a given context whereas, recognition can be described as figuring out the membership of the instance or incident to a particular class. Hence, fall “detection” can be described as detecting whether fall has occurred from given data, unlike fall “recognition” can be described as recognizing the specific type of fall, such as: forward fall, backward fall, fall from sitting position, fall from standing position, fall from lying position, etc. [11]–[13]. Fall recognition can be very useful in mounting the proper response. For example, fall from standing position might be more fatal than fall from sitting down or lying positions depending on the circumstances [10], [14], [15]. If the proper type of fall is known, effective responses to handle possible complications regarding to that specific type of fall can be mounted. In this literature, we do not explicitly mention whether a reviewed system is a fall detection or a fall recognition system and use the term “fall detection” to represent the all of the general fall detection and activity recognition system.

In recent years, deep learning has been used widely in most of the fields worldwide [28]. In fall recognition, deep learning methods are being used in the last few years effectively [29], [30] than other approaches like threshold based algorithms [31]–[36]. Machine learning approaches are also very common in this field [37]–[44]. Deep learning and machine learning are a subset of artificial intelligence [45], [46]. Machine learning works on structured data for classifying problems [38]. In machine learning based systems, the features that need to be recognized are programmed by the supervisor. On the contrary, deep learning based systems generally perform the necessary pattern recognition task without the employment of any explicit feature extraction methods. Deep learning models consist of multiple layers. Each layer extracts the feature of the given data or performs some transformations on the data. The final layers of the models generally consist of artificial neurons. The data can be in the form of vision based captured image, raw accelerometer data, gyroscope etc. The application of fall detection systems for robust activity monitoring is a widely researched field.

However, no current reviews on the current advances in deep learning based fall detection systems are present. In this literature, we have provided an overall view on the advances in deep learning based fall detection systems based on their working methods, efficiency, performance, and limitations to facilitate further development in this field. The reviewed systems are practical systems that employ deep learning methods for the proper detection and recognition of fall events among Activities of Daily Life (ADL) events. We have presented short summaries of their methodologies that include the details on how these practical systems are implemented in real-life environments. In addition, different methodologies are used for achieving the best accuracy, sensitivity, specificity and response time.

All of the reviewed fall detection systems have some general steps, combined with sensing, data processing, fall event recognition and emergency alert system to rescue the victim. A generalized view of fall detection systems are shown in Figure 1.

Various data acquisition devices such as accelerometers, gyroscopes, RGB cameras, radars are used to record data regarding the fall and Activities in Daily Life (ADL) events. In the data processing stage, data preprocessing, augmentation, feature extraction, etc. are performed for making the data usable in later stages. In the model development stage, the processed data is used for training a deep learning technique. In the detection stage, activity data from sensors are processed and provided to the trained model for proper classification. If a fall event is detected, then emergency services are alerted and rescue service is provided through informing doctors, ambulance, nurses and trained staffs.

There are several deep learning techniques used for recognition of the fall events. Some systems are developed by using only a single deep learning method and some are developed combining different methods for higher detection rate. In this literature, we have reviewed a total of 37 fall detection systems in total. Out of the 37 papers, a total of 21 systems (56.8% of the total reviewed corpus) used CNN architecture for fall detection. A total of 12 systems (32.4% of the total reviewed corpus) used LSTM architecture for fall detection. A total of 4 systems (10.8% of the total reviewed corpus) used auto-encoder architecture for fall detection. CNN based fall detection systems represent the fall and ADL related data in image form [51], [57]. This can be done either directly by taking images or videos captured by sensors, or by representing other sensor data in image form, etc. CNN architectures are great for finding patterns and shapes within given images [48], [52]. Thus, fall detection systems that employ various CNN based architectures leverage the power of CNN in categorizing or detecting images. LSTMs are an improvement of RNN architecture. Like CNNs, LSTMs can also handle image data. However, their main strength lies in dealing with sequential data, such as time series data [86], [88]. As sensor data regarding fall events are time-specific, LSTMs can be successfully and efficiently used to differentiate between fall and ADL events. In some of

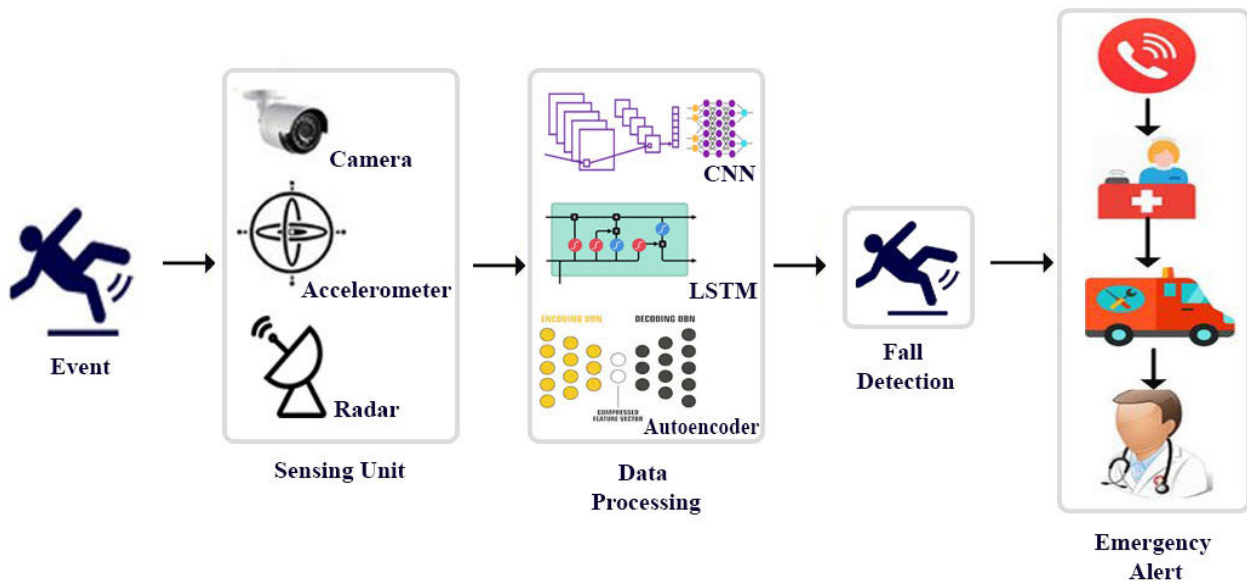


FIGURE 1. Generalized working principle of the fall detection systems.

the reviewed systems, combinations of LSTM and CNN are used for fall detection to overcome the general vision related problems such as image noise, occlusion, incorrect segmentation, perspective, etc. While CNN and LSTM architectures are mostly used for supervised learning, auto-encoders are used for learning efficient data coding in an unsupervised manner [108], [110]. Generally, some techniques are applied on sensor data that transform the sensor data into distinct data for different fall and ADL types. Auto-encoders are then used on the transformed dataset to distinguish between the events. Thus, the reviewed systems were categorized based on the principal method used to infer the events. The categorization focuses on how the different principal methods (CNN, LSTM, and Auto-encoder) handle the event data captured by sensors.

In this review, we have categorized the reviewed systems on the nature of handling the event data used by the deep learning methods. Thus, the reviewed systems can be categorized into the following three major categories:

- (i) Convolutional Neural Network (CNN)
- (ii) Long Short-Term Memory (LSTM) with Recurrent Neural Network (RNN) and Recurrent Convolutional Network (RCN)
- (iii) Auto-encoder

Figure 2 is a general representation of the categorization. The figure also represents the specific architectures that are generally used under every subsection. The use of the different architectures is mostly dependent on the sensor data or the technique used to represent fall and ADL events [48], [49].

The rest of the paper is arranged as follows: Section 2 discusses the literature on the methods and systems developed for fall detection, Discussions and future directions demonstrated in section 3 including the summary and performance analysis. Finally, we conclude this review in section 4.

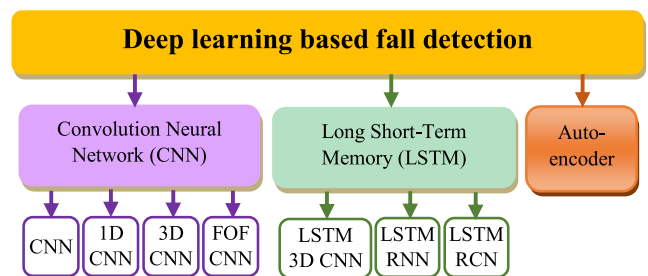


FIGURE 2. Deep learning methods used for fall detection.

## II. LITERATURE ON FALL DETECTION SYSTEMS

In the recent decade, several deep learning methods and systems have been developed for recognizing fall and ADL. Many researchers and organizations have been researching and working hard for developing a highly efficient deep learning method for recognizing fall with nearly zero false alarm in real-world situations. In this review, we tried to exhibit the most vital and most recent researches and methods developed for fall detection with the description considering:

- (i) Introducing the system
- (ii) Working principle of the system
- (iii) Datasets used for training and experiments
- (iv) Critical and performance analysis of the scheme

### A. CONVOLUTIONAL NEURAL NETWORK (CNN) BASED FALL DETECTION SYSTEMS

Most of the reviewed systems in this review use CNN for developing the automatic fall detection systems. CNN takes an image as input, then processes it and classifies the processed image under certain given categories [47]–[52]. CNN technique passes each input image through a series of

convolutional layers with filters, pooling layers, fully connected layers (FC) and softmax function for training and testing purposes. Some of the CNN based fall detection systems are proposed in [53]–[62]. The CNN based fall detection systems that are reviewed in this article are described as follows:

#### 1) CONVOLUTIONAL NEURAL NETWORK (CNN)

Adhikari *et al.* [63] proposed a fall detection system based on video images taken from RGB-Depth camera of Kinect. CNN was used in the system to recognize ADL and fall events. The proposed system used their own data set taken from different indoor environments by recording activities of different people. The dataset contains a total of 21499 images. A (73-27) % split was performed on the dataset for training and testing purposes. The overall accuracy of the proposed system is 74%. The sensitivity of the device is 99% when the user was in lying position. However, the system achieved very poor sensitivity when the user was in crawling, bending and sitting positions. The system also worked in a selected environment and simple deep learning techniques are used whereas fusion based method can provide better results in all environments. The system was developed for a single-person scenario. The use of CNN in video surveillance environment for fall detection is presented by Li *et al.* [64]. To learn human shape deformation feature for determining fall events and ADL, CNN is applied directly to each frame image. URFD dataset was used in this system. 10-fold cross validation with 850 test images per fold was used to evaluate the performance of the system. The performance of this approach achieved sensitivity with 100%, specificity and accuracy with 99.98% in average for 10-fold cross-validation. As the used datasets have almost same background, colors and environment, changes in background and foreground may degrade the performance of the system. However, the performance was not measured for real-life elderly falling event in different viewpoints and environment. Yhdego *et al.* [65] proposed pre-trained kinematics based machine learning approach in the annotated accelerometry datasets. The data of accelerometer is converted into images using continuous wavelet transform. Transfer learning approach is then used to train a deep CNN on the images. The open URFD dataset containing 30 sequences of fall and 40 sequences of normal activities was used. On the dataset, a standard split (80-20) % was performed. The proposed system achieved 96.43% accuracy.

Another application of CNN is applied by Yu *et al.* [66], where background subtraction method is applied to extract the human body silhouette. In the proposed system, CNN is applied to pre-processed extracted silhouettes that corresponds to human movement like standing, sitting, bending and lying down. Codebook background subtraction method is a common approach for distinguishing a moving object from the background which is used here. A custom posture dataset containing 3216 postures (804 stands, 769 sits, 833 bends, and 810 lies) was used for testing the system. The proposed system provides better performance (96.88% accuracy) than

traditional machine learning based system. Fall detection method based on deep learning and image processing in cloud environment is introduced by Shen *et al.* [67]. The frequency images taken through high resolution camera are passed to the Deep-Cut neural network model to detect the key points of human body. The detected key points are transferred into the deep neural network. Various morphological tests of the body and different types of the fall show the effectiveness of the system. A custom dataset was created by recording 44 videos, out of which, 42 videos were used for training purposes, and 2 videos were used for testing purposes. The system obtained an accuracy of 98.05%. Santos *et al.* [68] proposed a CNN-based deep learning technique which consists of 3 convolutional layers, 2 max-pooling and 3 fully-connected layers for fall detection in an Internet of Things (IoT) and fog computing environment. The performance of the system is evaluated using three open data sets and against extant research. A standard (80-20) % split was performed on the dataset. The obtained accuracy of the developed system is 93.8%.

A multi-sensor fusion based fall detection system employing deep learning has been proposed by Zhou *et al.* [69]. For obtaining the signals of human motion a continuous wave radar is adopted and short time Fourier transform (STFT) is used to acquiring the time frequency (TF) micro-motion features. An optical camera is used to capture the image sequence of human actions. Two CNNs named as Alex-Net and SSD (single shot multi-box detector) Net are used to classify the TF features. Finally, the joint decision is used for detecting the fall events. The training dataset of TF images consist of 300 falls and 600 non-falls and the fine tuning dataset for the SSD-Net in the optical camera contain 400 images. For testing, 1300 radar signals (time duration 1s-2s) and corresponding image sequences (325 falling, 325 walking, 325 squatting, 325 standing up) were prepared. The system achieved performance with accuracy of 99.85%. He *et al.* [70] developed a low power FD-CNN (Fall Detection Convolutional Neural Network) based fall sensing technology. The system developed an interrupt-driven sensor board integrated with MPU6050 and low power ZigBee for sampling and caching angular velocity and 3 axial acceleration. The collected data was then mapped into 3-channel RGB bitmap images. This data was combined with transformed SisFall and MobiFall datasets and used for training the FD-CNN. A (90-10) % split was performed on the combined dataset to obtain the training and testing data. The performance of proposed technology achieved the average accuracy of 98.61%, while its average specificity and sensitivity are 99.80% and 98.62%, respectively, while using a 10-fold cross-validation procedure. The performance may be improved if edge computing technologies with a low-power logistic network is used. A radar based automatic fall detection system has been developed by Sadreazami *et al.* [71]. A healthcare environment in a room had been set up and radar data had been collected in that setting. A time-series data had been derived. A deep convolutional network was

trained using the derived time series data. The proposed system provides significantly lower CPU time and higher classification metrics. The accuracy, precision, sensitivity, and specificity obtained by the proposed system are 92.72%, 94.21%, 93.44% and, 91.67%, respectively, while using a 5-fold cross-validation procedure.

A wide generalized method combining CNN along with an efficient change detection algorithm is proposed in [72] to tackle the limitations of generalization and the problem of resource optimization. The developed wide generalization method operated on the raw accelerometer data preserving with low resource demands. The developed system used SmartFall dataset where data are collected from experiments made with seven different persons wearing Microsoft Band 2 on his/her wrist. The CUSUM (cumulative sum) algorithm was used for the better performance of the system. The dataset was collected by logging data from 7 different persons. The dataset contains a total of 51192 samples. A standard (66-44) % split was performed on the dataset. The accuracy and balanced accuracy (BACC) of the proposed system is 96.79% and 95.35% on used dataset. The model can be more efficiently trained and tested with real-life datasets for better performance in the actual field of usage. Lu *et al.* [73] proposed a CNN based fall detection system. The developed system works in 3 steps, firstly, the system sent the optical flow images to network for avoiding appearance-based features. Secondly, 3D CNN was trained on different datasets for the acquisition of general features. Finally, for overcoming the problems of small size fall datasets the proposed system applied transfer learning. The experimentation was done with the developed system using three datasets URFD, Multicam and FDD. On the datasets, a standard (80-20) % split was done. The system achieved the average sensitivity and accuracy of 94% for both. The proposed framework performed best with the URFD dataset with accuracy of 99% and sensitivity of 100%. The system may only work well in terms of single-person fall detection, whereas multi-person fall detection is also important. Wang *et al.* [74] proposed a system for fall detection as well as to recognize ADL. The proposed system is combined with tri-axial accelerometer, gyroscope sensor in the smart insole. CNN is used for improving the accuracy level with less computational complexity. Dataset for the developed system is collected by volunteers wearing smart insole. The main point of the system is it can apply CNN directly in the raw sensor data collected by the tri-axial accelerometer and gyroscope. A custom dataset containing 800 falls (200 each for lying in bed, bowing, lying down, continuous walking and jogging) was collected. The average accuracy of the proposed system is approximately 98.61%. In addition, 97.58% and 99.58% are achieved for sensitivity and specificity respectively.

Zhang and Zhu [75] used cellphones for detecting real-time human activities. Smartphones are a staple of our daily life. The proposed system is based on deep CNN that works on raw 3-axis accelerometer data streams. A challenging dataset UniMiB SHAR (University of Milano Bicocca

Smartphone-based Human Activity Recognition) was used. Three 1D images that are generated from multiple views of the human activities are used for the training stage. The images are applied to CNN model for finding the hidden patterns. The dataset containing a total of 11,770 activity data of 17 different types was collected through patients with a wide range of ages. The reviewed system achieved a detection rate of 91.5%, while using 5-fold cross-validation. Multi-stream model concept is used by Cameiro *et al.* [76] for fall detection. The system proposes a multi-stream model that takes high-level handcrafted features as input. Optical flow, RGB and human estimated pose are provided as handcrafted feature generators. Then, modified CNNs like CNN for optical flow to extract vectors, CNN for RGB video data and CNN for human estimated pose are passed through dense VGG-16 layered classifiers to classify the multiple situations as well as fall detection. URFD and FDD datasets were used for training the models. The developed system obtained an accuracy of 98.77%, using 5-fold shuffled cross-validation.

Casilari *et al.* [77] proposed a deep CNN based fall detection system. The developed system detects fall events by recognizing the pattern from tri-axial transportable accelerometer. The proposed system employs a large set of data up to 14 publicly available datasets including MobiAct, SisFall, MobiFall, UniMiB SHAR, UP-Fall, etc. The system mainly analyzed those datasets characteristics, available features for fall detection and measured performance metrics using the proposed architecture all over the 14 datasets. The proposed method achieved a notable accuracy, sensitivity and specificity of 99.22%, 98.64% and 99.63% respectively on SisFall dataset. For improving real-life performance and decreasing the limitation of extrapolation capability, more efficient CNN network or another efficient architecture (e.g. LSTM) would be deployed. Espinosa *et al.* [78] developed a CNN based fall detection system utilizing UP-Fall detection multi-modal dataset. The system worked only with vision dataset and used multiple cameras for fall detection. Input images were analyzed in a fixed time slot and features were extracted by implementing an optical flow method by obtaining information from two consecutive images relative motion. The proposed system was tested on the public datasets and performed well with detection accuracy of 95.64%, sensitivity of 97.95%, specificity of 83.08% on Multicam dataset and accuracy of 82.26%, sensitivity of 71.67%, and specificity of 77.48% on UP-Fall dataset. The performance can be improved for real-time application along with increasing privacy.

## 2) ONE DIMENSIONAL (1D) CONVOLUTIONAL NEURAL NETWORK (CNN)

Cho and Yoon [79] applied SVD (Singular Value Decomposition) on accelerometer data for 1D CNN based fall detection. The proposed system worked with UniMiB SHAR, SisFall, and UMAFall public ADL and falls public datasets containing signals for triaxial acceleration for both ADLs and falls. The experimental triaxial acceleration dataset was collected

from a single accelerometer. The performance of the system using the UniMiB SHAR dataset has obtained the accuracy of 75.65% by combining the features SMV (Signal Magnitude Vector) and SVD. The performance using the other two datasets has an accuracy of 63.08% for SisFall and 64.69% for UMAFall datasets respectively by combining the raw features and SVD, based on Leave One Subject Out (LOSO) cross-validation procedure. Fusion based method and advanced CNN approaches can be used for higher performance rate.

A one dimensional CNN based method named as MyNet1D-D is proposed in [80] by Tsai and Hsu in which the traditional skeleton feature extraction algorithm is used for transforming the depth image information into the skeleton information and extracting seven highlighted feature points from the skeleton joints. The system proposed a robust deep one-dimensional CNN architecture with a small number of parameters and calculations which is suitable for implementation of embedded system. The proposed system performs better on the NTU RGB+D popular benchmarked dataset. The proposed system was implemented on the NVIDIA Jetson Tx2 integrated machine for achieving real-time prediction of falls. RGB camera may be used instead of a depth Kinect sensor to reduce the cost of the system as it focuses on the implementation of embedded system.

### 3) THREE DIMENSIONAL (3D) CONVOLUTIONAL NEURAL NETWORK (CNN)

A 3D CNN for detecting fall event was developed in [81] by capturing both spatial and temporal information available in frames taken from video. The Kinect depth camera is used for taking videos and collecting the information. Adam optimizer algorithm (a gradient based algorithm) was used to train the model. The proposed system used the SDUFall dataset containing 6 actions that were performed by 20 young people. The average accuracy of 97.58% was achieved by the proposed system using the SDUFall dataset. The performance of the proposed system is measured using a simulated dataset whereas the real-life data of fall events of elderly people is more important and may degrade the achieved performance. The pre-impact fall detection approach based on 3D CNN using RGB images is introduced by Li *et al.* [82]. Pre-training data is served to the 3D CNN to train the model and distinguish the normal movement and fall based on Spatio-temporal patterns. Secondly, new samples and pre-training data are combined to provide fine-tuning data for fine-tuning the model. The dataset of the proposed system is collected by 10 trainers who provide 25 trials of normal movement and 25 trials of falling. A custom dataset was created containing 455 trials of collected data from 10 trainees. The developed system achieved 100% accuracy within 0.5 seconds of falling based on 225 testing trials from 5 trainees. Though the reviewed system obtained reasonable accuracy, further studies can be performed on using other features highlighting the real-time scenarios.

Hwang *et al.* [83] proposed an approach that applies 3D-CNN to analyze the continuous motion data acquired

from depth cameras. To avoid overfitting, data augmentation methods are applied. The system considered a total of 5 random trials, for every 240 and 24 videos for training and evaluation purposes are extracted getting from TST fall detection dataset. The evaluation result is significantly improved and achieved the classification accuracy of 92.4 ~ 96.9%. Kasturi *et al.* [84] introduced a system consisting 3D-CNN which used video information from Kinect camera. Several frames from the video forming a stacked cube is passed as the input to 3D-CNN. Publicly available UR fall detection dataset is used for testing the proposed system and it achieved 100% accuracy on both training and validation phase. A standard (80-20) % split was performed on the dataset. Datasets consist of falls while walking, sudden falls and falls from chairs and other activities like walking, sitting, bending scenarios are considered.

### 4) FEEDBACK OPTICAL FLOW CONVOLUTIONAL NEURAL NETWORK (FOF CNN)

Hsieh and Jeng [85] developed an IoT based home intelligent system for fall detection that worked on home environment. The proposed system used Feedback Optical Flow Convolutional Neural Network. The system used Feature Feedback Mechanism Scheme (FFMS) and 3D-CNN for feedback on the safe movement and obtained higher accuracy with less computation. The KTH dataset consisting six type of actions and 25 subjects was used. The average accuracy of the proposed system was achieved  $92.65 \pm 2.3\%$ .

The summary of the reviewed CNN based systems is shown in Table 1, organized with the following criteria: the category of the systems, devices used for acquisition of raw data, deep learning methods and other techniques used for the systems, whether a system only detects fall or ADLs, used datasets, etc. Table 1 presents an overview of the CNN based reviewed systems at a glance.

Most of the reviewed systems detect fall events. However, some of them also detect ADLs. Different types of datasets from different sources have been used in those systems for training and experiment. But almost all datasets are generated from the simulation of fall and ADLs of different subjects. For sensing the fall event, some of the systems used RGB depth camera which is in the vision device category [63], [64], [73], etc. Some of the systems used accelerometer, UWB radar, etc. which is in the sensor device category. Various datasets have been used in the reviewed systems. Some of them used publicly available dataset such as KTH dataset [85], Uni-MiB SHAR [75], [77], [79] SisFall [70], [77], [79] UMA fall [79], UP-Fall [78], Multicam [78], URFD [64], [65], [73] and [76]. Some of them used simulated datasets generated by trainers.

## B. LONG SHORT-TERM MEMORY (LSTM) BASED FALL DETECTION SYSTEMS

In sequence modeling, LSTM is one of the popular and most used recurrent structures. For controlling information flow in the recurrent computations, it uses gates [86], [87].

**TABLE 1.** Summary of the systems developed for fall detection using CNN.

Authors	Year	Device Category	Capturing Device	Used Method	Detection	Used Dataset
Adhikari <i>et al.</i> [63]	2017	Vision	RGB-Depth camera of Kinect	CNN	ADL, Fall	Own dataset
Li <i>et al.</i> [64]	2018	Vision	Microsoft Kinect cameras	CNN with 10-fold cross-validation	ADL, Fall	UR fall detection dataset (URFD)
Yhdego <i>et al.</i> [65]	2019	Sensor	Annotated accelerometer camera	CNN	ADL, Fall	UR fall detection dataset (URFD)
Yu <i>et al.</i> [66]	2017	Vision	High resolution camera	CNN	ADL, Fall	N/A
Shen <i>et al.</i> [67]	2019	Vision	High resolution camera	DeepCut Neural Network	Fall	Own dataset with 44 videos
Santos <i>et al.</i> [68]	2019	Sensor	Accelerometer	CNN consisted with 3 convolutional layers and 2 max-pool	ADL, Fall	Three open data sources
Zhou <i>et al.</i> [69]	2018	Multi-sensor	Continuous wave radar, optical camera	Three CNN based on Alex-Net and SSD-Net	Fall	Own dataset
He <i>et al.</i> [70]	2019	Sensor	3axial accelerometer, gyroscope	Fall Detection CNN (FD-CNN)	Fall	SisFall and MobiFall dataset
Sadrezami <i>et al.</i> [71]	2019	Sensor	UWB radar	CNN	Fall	Xethru X4M03 development kit
Tasoulis <i>et al.</i> [72]	2019	Sensor	Accelerometer	CNN and cumulative sum (CUSUM) algorithms	Fall	SmartFall
Lu <i>et al.</i> [73]	2019	Vision	RGB camera	CNN	Fall	URFD, Multi Cam and FDD
Wang <i>et al.</i> [74]	2019	Sensor	3axial accelerometer, gyroscope	CNN	ADL, Fall	By volunteer wearing insole
Zhang and Zhu [75]	2019	Sensor	Raw 3 axis accelerometer	CNN	ADL, Fall	UniMiB SHAR
Cameiro <i>et al.</i> [76]	2019	Vision	RGB camera	CNN for optical flow	Fall	URFD and FDD dataset
Casilari <i>et al.</i> [77]	2020	Sensor	tri-axial transportable accelerometer	CNN	Fall	14 public datasets including MobiAct, SisFall, MobiFall, UniMiB SHAR, UP-Fall, etc.
Espinosa <i>et al.</i> [78]	2019	Vision	RGB multi-cameras	CNN	Fall	UP-Fall, Multicam
Cho and Yoon [79]	2019	Sensor	Accelerometer	Singular Value Decomposition (SVD), 1D CNN, Signal Magnitude Vector (SMV)	ADL, Fall	UniMiB SHAR, SisFall and UMAFall
Tsai and Hsu [80]	2019	Vision	Depth camera	1D CNN called MyNet1D-D	Fall	NTU RGB+D dataset
Rahnemoonfar and Alkittawi [81]	2019	Vision	Kinect depth camera	3D CNN, Adam Optimizer algorithm	Fall	SDUFall dataset
Li <i>et al.</i> [82]	2019	Vision	RGB images	3D CNN	ADL, Fall	Collected by 10 trainers, 50 trials for normal and falling movement
Hwang <i>et al.</i> [83]	2017	Vision	Depth cameras	3D-CNN	Fall	TST fall dataset
Kasturi <i>et al.</i> [84]	2019	Vision	Kinect camera	3D-CNN	Fall, ADL	UR fall detection dataset (URFD)
Hsieh and Jeng [85]	2017	Vision	RGB cameras	FOF CNN and 3DCNN	ADL, Fall	KTH dataset

LSTM networks can hold long term memories very well. The memory may or may not be retained by the network depending on the data. The gating mechanisms of LSTM preserves the long-term dependencies in the network. The network can store or release memory using the gating mechanism [88]–[90].

Some of the systems have used aforementioned features of LSTM in developing fall detection systems [91]–[95]. The system developed with LSTM along with CNN, RNN and RCN which are reviewed in this article are described as follows:

#### 1) LSTM WITH 3D CNN

A fall detection system based on LSTM with 3D CNN is developed by Lu *et al.* [96] by overcoming the problems of

existing methods like segmentation, image noise, variation, occlusion, and illumination. The system trained an automatic feature extractor only using kinetic data. For focusing on the key region, a spatial visual attention scheme based on LSTM was then incorporated. To train the 3D CNN, the proposed system used Sports-1M datasets and fall dataset of Multiple Cameras are used to train the visual attention model. The extracted Spatio-temporal features are effectively used to recognize unforeseen motions. On the dataset, a typical split (80-20) % was performed. The sensitivity, specificity, and accuracy are 98.13%, 99.91% and 99.84% with 1 frame overlapping on multiple cameras fall datasets. The developed system was not tested in a real-life environment. In real-life scenarios, the performance of the proposed system would vary.

## 2) LSTM WITH RECURRENT NEURAL NETWORK (RNN) AND CNN

A fall detection system has been developed in [97] based on 2D skeletons extraction from an RGB camera. For humans' 2D skeletons extraction, CNN is adopted and then RNN with LSTM is employed to classify actions into standing, walking, lying, rising and falling categories. In the first stage, the developed system used a pre-built model named Deeper Cut (152 layers). The learning rate of the system is 0.0001 for training 500 epochs. The system was trained on their own dataset. A (60-40) % split was performed on the dataset to get the training and testing sets. The system achieved a moderate performance with an average accuracy of 88.9%. The system has an on-line version with some drawbacks and can achieve 8 frames/second processing speed with GPU model GTX 1060 6G providing one output per frame. Abobakr *et al.* [98] proposed an efficient, integrable and privacy-preserving fall detection system based on CNN and RNN. The system used images taken from Kinect RGB-D camera. Deep Convolutional LSTM with ResNet for visual features extraction, LSTM for sequence modeling and logistic regression for fall action recognition, is used in the proposed system. The proposed system used URFD (UR Fall Detection) public dataset for training and evaluating the system performance. A typical (80-20) % split was made on the dataset. The developed system achieved an average accuracy of 98% along with precision, sensitivity, and specificity of 97%, 100%, and 97% respectively on the validation set. The proposed convolutional LSTM model can provide real-time fall event detection with a processing speed of 15 sequences (each sequence contains 80 frames) per second on a GPU modeled NVIDIA TITAN-X. An IoT enabled fall detection technique has been proposed by Xu *et al.* [99] using the excellent feature extraction capability of CNN and LSTM for powerful processing ability of time series. The proposed algorithm takes acceleration data as input taken from a low-cost three-axis accelerometer. The system performed better than the Support Vector Machine (SVM) and CNN based combined method. The training and testing were performed on the MobiAct datasets containing fall events with the other 6 types of ADLs. On the dataset, a standard (80-20) % split was made. The average accuracy of the classification reached 98.98%. The detection specificity and precision is 99.76% and 98.61%. The developed system can be deployed and used as an AI-enabled IoT application for nursing homes as well as for collecting a significant amount of data for academic research.

The characteristics of body posture and human biomechanics equilibrium can be analyzed for fall prediction before hitting the wall or floor which is presented by Tao and Yun [100]. The developed system employed RNN and LSTM taking the input from depth camera data of Kinect sensor and analyzed three-dimensional skeleton joints for determining the ADLs and falls by computing center of mass (COM) positions and the region of base support. The proposed system has been evaluated on an open database generated by

Rougier and Meunier [101] which contains 4 types of ADLs and 4 types of fall events of people aged from 22 to 39 of different weight and height. The system achieved an average prediction accuracy of 91.7% and can predict falls before 333ms of hitting the wall or floor. The proposed method lays a foundation for fall protection studies. Ajerla *et al.* [102] proposed a framework for fall detection based on LSTM network that used edge devices like a laptop for computing rather than sending raw data to the cloud for real-time prediction of the fall events. The proposed framework used a cheap MetaMotionR sensor from MbientLab for three-axis accelerometer raw data, a streaming engine Apache Flink which is an open-source software through which data analytics has been streamed. A subset of the MobiAct public dataset has been used for training and testing the architecture. The developed system determined the waist as the best position for sensors placement. The proposed framework can predict fall events from real-time fall data with an accuracy of 95.8%. The uses of multiple sensors and multiple data streams can achieve better performance. A fall detection system architecture with health monitoring facilities has been proposed by Queralta *et al.* [103]. The system used both Edge computing and Fog computing with a compression algorithm to transfer data using low power enabled low-power wide-area network (LPWAN) technology. This improved the system latency. LSTM along with RNN networks has been implemented on the edge computer for detecting falls from the received data. From this edge gateways, real-time notifications and alerts are sent and raw data are sent to the cloud for online analysis. The proposed framework can operate in areas with poor network connectivity and increases battery life. An average accuracy of 95% and a precision of 90% have been achieved in predicting fall events by the proposed system operating on the MobiAct dataset. The system performance can be improved by tweaking a little bit and combining other methods.

Luna-Perejon *et al.* [104] introduced a fall detection method based on Gated RNN architectures along with LSTM and Gated Recurrent Units (GRU) taking accelerometer data as input which detects fall in real-time. The input is sent to a batch normalization layer, a dense end layer with a RNN layer for the prediction of fall events. The public SisFall dataset was used to train and test the system. A standard (80-20)% split was performed on the dataset. The performance of the proposed technique achieved F1-scores above 0.85 and 0.98 respectively. Mean F1-score was obtained 0.73 and 0.75 respectively for the GRU and LSTM versions. Common performance metrics such as accuracy, precision, sensitivity, etc. were not measured. Torti *et al.* [105] presented a method for implementing RNN architecture along with LSTM for fall detection which can be embedded on a Micro-Controller Unit (MCU) with Tri-axial accelerometers. The implementation was developed in Tensorflow for reference. The trained model was tested with specially annotated SisFall dataset. A (80-20)% split was performed on the dataset. The proposed system was able to achieve an average accuracy of 98% in prediction of fall events. A run-time detection



**TABLE 2.** Summary of the LSTM based fall detection systems.

Authors	Year	Device Category	Capturing Device	Used Method	Detection	Used Dataset
Lu <i>et al.</i> [96]	2018	Vision	Multiple Cameras	3D CNN, LSTM	Fall	Sports-1M datasets and Multiple Cameras fall dataset
Lie <i>et al.</i> [97]	2018	Vision	RGB camera	CNN, RNN, LSTM, Deeper Cut	Fall	Own dataset
Abobakr <i>et al.</i> [98]	2019	Vision	Kinect RGB-D camera	CNN, RNN, Deep Convolutional LSTM with ResNet, recurrent LSTM, logistic regression module	Fall	UR Fall Detection Dataset
Xu <i>et al.</i> [99]	2019	Sensor	Three-axis accelerometer	CNN and LSTM	ADL, Fall	MobiAct dataset
Tao and Yun [100]	2017	Vision	Kinect depth camera	RNN and LSTM	ADL, Fall	Open database created by Rougier and Meunier [101]
Ajerla <i>et al.</i> [102]	2019	Sensor	Three-axis accelerometer	LSTM	Fall	MobiAct dataset
Queralt <i>et al.</i> [103]	2019	Sensor	MPU9250 with 3-axis accelerometer, 3-axis gyroscope, and 3-axis magnetometer	RNN and LSTM	ADL, Fall	MobiAct Dataset
Perejon <i>et al.</i> [104]	2019	Sensor	Accelerometer	Gated RNN with LSTM and GRU	Fall	SisFall dataset
Torti <i>et al.</i> [105]	2018	Sensor	Tri-axial accelerometers	RNN and LSTM	Fall	SisFall dataset
Theodoridis <i>et al.</i> [106]	2018	Sensor	Accelerometer	RNN and LSTM	Fall	URFD dataset
Lin <i>et al.</i> [107]	2018	Vision	Kinect RGB-D camera	RNN and LSTM	ADL, Fall	Own dataset
Ge <i>et al.</i> [108]	2018	Vision	4 Kinect sensors (used RGB only)	RCN, RNN, LSTM, KERAS Library with Tensor-Flow	Fall	ACT4 <sup>2</sup> open video Dataset [9]

and highly-optimized module has also been implemented in SensorTile device and validated with the results obtained with TensorFlow installed on a workstation. The developed device is capable of operating in very low power consumption; thus, long battery life can be achieved and it is shown that the system can be run about 20 hours without recharging.

A RNN based fall detection method is proposed in [106] by Theodoridis *et al.* The proposed networks are capable of process and encode the sequential data taken from acceleration measurements from sensors worn in the body. The system used URFD dataset for training and testing their publicly available method. Additionally, the developed system used augmented data taken from random 3D rotations to benefit those networks in the training phase. The system used 4 different networks and methods to compare their result, namely, LSTM-Acc, LSTM-Acc Rot, Acc + SVM-Depth, UFT. A standard (90-10) % split was performed on the dataset. Among them, LSTM-Acc Rot method has performed the best which gives an achieved of 98.57%, precision, sensitivity, and specificity are of 100%, 96.67%, and 100% respectively, using 10-fold cross-validation. Hsiu-Yu *et al.* [107] proposed a deep learning method based on RNN along with LSTM which continuously takes sequential input images from Microsoft Kinect camera and classifies consecutive images for recognizing posture types as well as falls. The developed system is compared with Image-net in [109] which is a pure CNN model whereas the proposed system has LSTM in the seventh layer. The framework captured the correlation

taken from consecutive input images and use the previous iteration value to determine the continuous falling action. The system used fusion images as input extracted from the high-resolution RGB images applying Gaussian Mixture Model (GMM) for body shape and optical flow calculation, along with the depth images. As a result, the proposed system performs better than the existing Image-net model though it is slower than Image-net in terms of training the model.

### 3) LSTM WITH RECURRENT CONVOLUTIONAL NETWORK (RCN) AND RNN

Ge *et al.* [108] proposed a co-saliency-enhanced RCN architecture based fall detection scheme for detecting fall events from videos. The system takes video clips as input, enhanced by co-saliency enhancement and then RCN is applied followed by a RNN with LSTM to label the output. The proposed system used KERAS library with Tensor-Flow backend for experiment. The experiments were applied on an open dataset from multiple camera videos at different views. A (50-50)% split was performed on the dataset. The experimental results of testing have shown performance with accuracy of 98.96%.

The summary of the reviewed LSTM based systems is shown in Table 2, organized with the following criteria: the category of the systems, devices used for acquisition of raw data, deep learning methods and other techniques used for the systems, whether a system only detects fall or ADLs, used

**TABLE 3.** Summary of the auto-encoder based fall detection systems.

Authors	Year	Device Category	Capturing Device	Used Method	Detection	Used Dataset
Jokanovic <i>et al.</i> [116]	2016	Sensor	Radar	Two stacked auto-encoders (AE) and softmax layer (regression classifier)	ADL, Fall	Own dataset
Droghini <i>et al.</i> [117]	2019	Sensor	Presonus AudioBox 44VSL sound card	Deep neural network Auto-encoder (AE)	ADL, Fall	Own dataset
Zhou and Komuro [118]	2019	Vision	RGB cameras	Variational Auto-encoder (VAE)	ADL, Fall	HQFD and Le2i fall dataset
Seyfioglu <i>et al.</i> [119]	2018	Sensor	Radar micro doppler	Auto-encoder combining CNN	ADL, Fall	Own dataset

datasets, etc. Table 2 presents LSTM with CNN, RNN and RCN based developed systems in a short view.

The majority of the systems detect fall events instead of ADL. Almost, all of them are in the vision device category. URFD [98], [106], open video dataset [100], [108], sports-1M dataset [96] and multiple cameras fall dataset [96], Mobi-Act dataset [99], [102], [103], SisFall dataset [104], [105] are used and besides own dataset is used in [97], [107].

### C. AUTO-ENCODER (AE) BASED FALL DETECTION SYSTEMS

Auto-encoder is an unsupervised artificial neural network [110], [111]. Being an unsupervised neural network, the model is trained without giving the labeled dataset [112]. There are four basic parts considered to develop a model, those are an encoder, bottleneck, decoder, and reconstruction loss [113]–[115].

Fall detection systems that are developed based on Auto-encoder techniques and reviewed in this article are described as follows:

Jokanovic *et al.* [116] proposed a fall detection system using radar based on deep neural network (DNN) where time-frequency (TF) analysis provides distinguishable components of human motion. DNN consists of stacked auto-encoders (two stacked) and a softmax layer (regression classifier) output. Four human activities or motions like falling, walking, bending/straightening, sitting are considered in the proposed system. The performance of the model is evaluated by constructing a confusion matrix of the proposed system along with the conventional approached. Approximately 87% success rate is achieved comparing a 78% success rate of the conventional approach. Droghini *et al.* [117] proposed an unsupervised method for acoustic fall detection by a deep convolutional-auto-encoder using a DT (downstream threshold) classifier. The model is trained as a novelty detector through the end-to-end strategy gathering signals of Floor Acoustic Sensor, corresponding to commonly occurring sounds. The model differentiates sound events generated by common indoor human activity like footsteps, speech and human falls event through fall event sound. The datasets of the proposed system is collected in a rectangular room 7 x 2 m using a Presonus AudioBox 44VSL sound card and FAS, positioned on the floor. The two types of sound are fall event sound and normal activities of human. The result is

evaluated in two categories one is clear sound environment, around 94.61% accuracy is achieved for the proposed system and another is a noisy sound environment, around 95.02% accuracy is achieved for the proposed system.

Variational Auto-encoder (VAE) with 3D convolutional residual block and a region extraction technique for enhancing accuracy is used to detect fall actions by Zhou and Komuro [118]. The proposed system is a kind of unsupervised learning method which used reconstruction error to detect fall actions as well as ADL. High-Quality Simulated fall dataset (HQFD) and Le2i fall dataset (Le2i) are used on the proposed system. The developed system achieved accuracy of 88.7% through the unsupervised Res-VAE model. Three-layer Deep Convolutional Auto-encoder (CAE) is proposed by Seyfioglu *et al.* [119]. The three layers are a combination of CNN and Auto-encoders that utilized unsupervised pre-training to initialize the weights in the subsequent convolutional layers. The architecture is more effective than other conventional classifiers. The proposed system used radar as a sensing device. For generating the dataset, 11 different people are participated to collect 1007 gait samples of 12 different classes. The accuracy obtained by the system is 94.2% for discriminating 12 indoor activity classes involving aided and unaided human motion.

The summary of the reviewed Auto-encoder based systems is shown in Table 3, organized with the following criteria: the category of the systems, devices used for acquisition of raw data, deep learning methods and other techniques used for the systems, whether a system only detects fall or ADLs, used datasets, etc. Table 3 presents an overview of the Auto-encoder based reviewed systems in a brief. Most of the proposed systems based on Auto-encoder sensor-based, one of them used auto-encoder with the CNN model. Vision-based proposed system in [118] used HQFD and Le2i datasets. But other systems used own datasets. Some of them used radar [116], [119]. Using own dataset, HQFD and Le2i proposed systems detect fall events as well as ADL. Some other systems have also proposed an approach that used auto-encoder which is available in [120], [121].

## III. DISCUSSIONS AND FUTURE DIRECTIONS

### A. DISCUSSIONS

In this review, we investigate the most recent deep learning based fall detection systems. The overall review is

**TABLE 4. Summary of the deep learning methods used for fall detection.**

Authors	CNN	LSTM	RNN	RCN	Auto-encoder
Adhikari et al.[63]	✓	×	×	×	×
Li et al. [64]	✓	×	×	×	×
Yhdego et al. [65]	✓	×	×	×	×
Yu et al. [66]	✓	×	×	×	×
Shen et al. [67]	✓	×	×	×	×
Santos et al. [68]	✓	×	×	×	×
Zhou et al. [69]	✓	×	×	×	×
He et al. [70]	✓	×	×	×	×
Sadrezami et al. [71]	✓	×	×	×	×
Tasoulis et al. [72]	✓	×	×	×	×
Lu. et al [73]	✓	×	×	×	×
Wang et al. [74]	✓	×	×	×	×
Zhang and Zhu [75]	✓	×	×	×	×
Cameiro et al. [76]	✓	×	×	×	×
Casilari et al. [77]	✓	×	×	×	×
Espinosa et al. [78]	✓	×	×	×	×
Cho and Yoon [79]	✓	×	×	×	×
Tsai and Hsu [80]	✓	×	×	×	×
Rahnemoonfar and Alkittawi [81]	✓	×	×	×	×
Li et al. [82]	✓	×	×	×	×
Hwang et al.[83]	✓	×	×	×	×
Kasturi et al. [84]	✓	×	×	×	×
Hsieh and Jeng [85]	✓	×	×	×	×
Lu et al. [96]	✓	✓	×	×	×
Lie et al. [97]	✓	✓	✓	×	×
Abobakr et al. [98]	✓	✓	✓	×	×
Xu et al. [99]	✓	✓	×	×	×
Tao and Yun [100]	×	✓	✓	×	×
Ajerla et al. [102]	×	✓	×	×	×
Queralta et al. [103]	×	✓	✓	×	×
Perejon et al. [104]	×	✓	✓	×	×
Torti et al. [105]	×	✓	✓	×	×
Theodoridis et al. [106]	×	✓	✓	×	×
Lin et al. [107]	×	✓	✓	×	×
Ge et al. [108]	×	✓	✓	✓	×
Jokanovic et al. [116]	×	×	×	×	✓
Droghini et al. [117]	×	×	×	×	✓
Zhou and Komuro [118]	×	×	×	×	✓
Seyfioglu et al. [119]	✓	×	×	×	✓

summarized in Table 4. In Tables 1, 2 and 3, we categorized all the systems that work with vision data and with various sensors data.

Vision based systems work with the data from RGB and Depth camera of Kinect shown in [63], [64], [80], [81], [98], [100], [107], [108], RGB Cameras introduced in [67], [69], [76], [78], [82], [85], [96], [97], [108]. Sensor based systems work with accelerometer data shown in [70], [72], [74], [77], [79], [99], [102]–[106], gyroscope data depicted in [70], [104], radar data demonstrated in [69], [71], [116], [119]. Comparing both categories, vision based systems provide better result than sensor based system.

**TABLE 5. Performances of the reviewed systems developed for fall detection.**

Authors	Accuracy (%)	Sensitivity (%)	Specificity (%)	Processing Speed
Adhikari et al. [63]	74	99	N/A	N/A
Li et al. [64]	99.98	100	99.98	0.0234s per image
Yhdego et al. [65]	96.43	N/A	N/A	N/A
Yu et al. [66]	96.88	N/A	N/A	N/A
Shen et al. [67]	98.05	N/A	N/A	N/A
Santos et al. [68]	93.8	N/A	N/A	N/A
Zhou et al. [69]	99.85	N/A	N/A	N/A
He et al. [70]	98.61	98.62	99.80	0.13s per image [GTX 1080]
Sadrezami et al. [71]	92.72	93.44	91.67	N/A
Tasoulis et al. [72]	96.79	N/A	N/A	N/A
Lu. et al [73]	94	94	N/A	62.6 fps [Tesla K40]
Wang et al. [74]	98.61	97.58	99.58	N/A
Zhang and Zhu [75]	91.5	N/A	N/A	N/A
Cameiro et al. [76]	98.77	N/A	N/A	N/A
Casilari et al. [77]	99.22	98.64	99.63	N/A
Espinosa et al. [78]	95.64	97.95	83.08	N/A
Cho and Yoon [79]	75.65	N/A	N/A	N/A
Tsai and Hsu [80]	99.2	98.9	99.1	15 frames/second in NVIDIA Jetson TX2
Rahnemoonfar and Alkittawi [81]	97.58	N/A	N/A	7 fps
Li et al. [82]	100	N/A	N/A	0.5s of falling based on 225 testing trials from 5 trainers.
Hwang et al.[83]	96.9	N/A	N/A	N/A
Kasturi et al. [84]	100	N/A	N/A	N/A
Hsieh and Jeng [85]	94.95	N/A	N/A	25 frames/sec
Lu et al. [96]	99.73	96.65	99.85	N/A
Lie et al. [97]	88.9	N/A	N/A	8 frames/sec
Abobakr et al. [98]	98	100	97	[GTX 1060 6G GPU] 120 frames/sec [NVIDIA TITAN-X GPU]
Xu et al. [99]	98.98	N/A	99.76	N/A
Tao and Yun [100]	91.7	N/A	N/A	N/A
Ajerla et al. [102]	95.8	N/A	N/A	N/A
Queralta et al. [103]	95	N/A	N/A	N/A
Torti et al. [105]	98	N/A	N/A	N/A
Theodoridis et al. [106]	98.57	96.67	100	N/A
Ge et al. [108]	98.96	N/A	N/A	N/A
Jokanovic et al. [116]	87	N/A	N/A	N/A
Droghini et al. [117]	94.61	N/A	N/A	N/A
Zhou and Komuro [118]	88.7	N/A	N/A	N/A
Seyfioglu et al. [119]	94.2	N/A	N/A	N/A

\* N/A = Not Appropriately defined

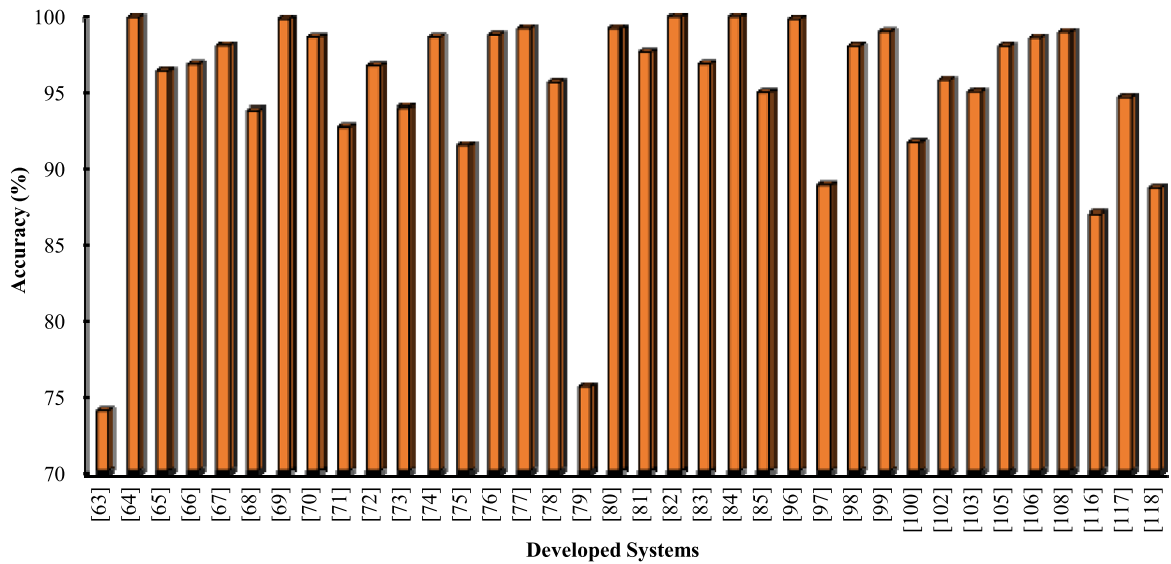


FIGURE 3. Accuracy of the reviewed fall detection systems.

Most of the researchers have used CNN for developing the fall detection system. 1D-CNN, 2D-CNN, 3D-CNN are described on Table 1. 3D-CNN gives better performance in [96]. Some of them have used LSTM, RNN and RCN, Auto-encoder with three layers, regression, and logistic classifier.

Most of the auto-encoder architectures are used by sensor based system using radar data illustrated in [116], [119]. The proposed system in [119] used auto-encoder along with convolutional neural network for better result. Some of the systems have used combination of methods for accurate result [113], [116]. The used deep learning method by various systems has been shown in Table 4 in a short view.

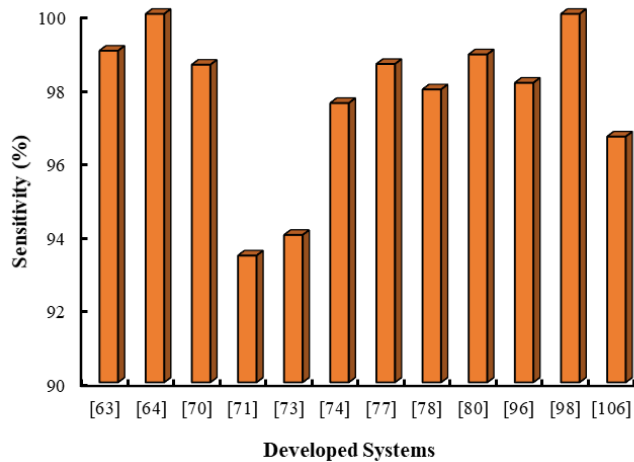


FIGURE 4. Sensitivity of the developed fall detection systems.

The performance, specifically the accuracy, sensitivity, specificity and processing speed of the systems obtained by experimenting on different datasets is demonstrated in Table 5 and described in Figures 3-5, depending on some criteria.

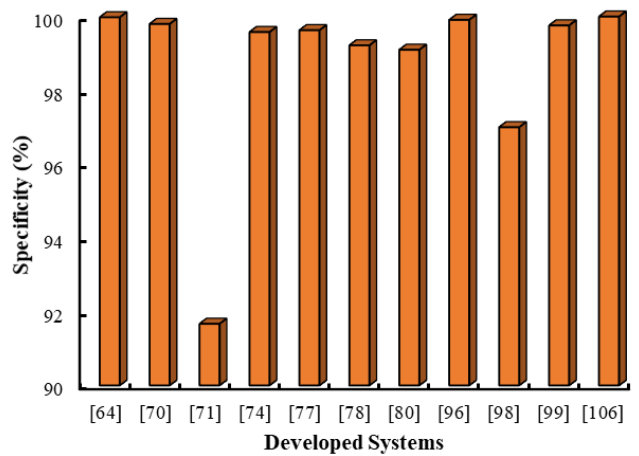


FIGURE 5. Specificity of the developed fall detection systems.

We considered the criteria that the deep learning methods used, datasets used for training and experiment, accuracy, sensitivity and specificity of the systems.

Almost all of the systems measured the average accuracy of the proposed systems. However, very few literatures mentioned the calculated sensitivity and specificity score. The processing speed are not clear in all of the system. Mentionable that in [82] has 0.5s delay to response, though it is based on 225 testing trials from 5 trainers.

Figure 3 shows the accuracy comparison between the reviewed fall detection systems. However, as different systems used different datasets, the systems are not really comparable. The system proposed in [82] has achieved the highest accuracy of 100% using 3D CNN where datasets were collected by 10 trainers of 25 trials of normal movement and 25 falling movement. The system developed in [84] has also achieved the highest accuracy of 100% using 3D CNN.

The systems in [64], [69], [74], [76], [77], [80], [82], [96] have also achieved high levels of accuracy using various deep learning methods. The highest accuracy achieved 99.98% in [64] using CNN with 10-fold cross-validation, 99.85% in [69] using three CNN based on Alex-Net and SSD-Net, 99.84% in [96] using 3D CNN along with LSTM. The lowest accuracy obtained 74% in [63] using CNN. The moderate accuracy of the reviewed system is 96.43% found in [65].

As sensitivity is one of the criteria to evaluate any working model, some of the reviewed systems [63], [64], [70], [71], [73], [74], [77], [78], [80], [96], [98], [106] have mentioned it. From Figure 4, the highest sensitivity of the reviewed system is 100% found in [64], [98]. CNN with 10-fold cross-validation is used in [64] and Microsoft Kinect cameras is used for capturing fall event. CNN, LSTM with ResNet, recurrent LSTM, logistic regression module is used in [98]. Both of the proposed system [64], [98] used same dataset named as UR fall detection dataset. The lowest sensitivity obtained 93.44% in [71] which used only CNN for training, radar is used for sensing the fall event. The moderate sensitivity of the reviewed systems are 98.64% [77] and 98.13% [96].

Figure 5 shows the found specificity of the reviewed fall detection systems. The highest specificity is 99.98% in [64], the lowest one is 83.08% in [78] and the moderate specificity of the reviewed system is 97% in [71]. Though there are differences in the performance of those systems, the systems are not comparable, due to mostly varied datasets.

Sensitivity and specificity both are very important for any healthcare system, especially fall detection systems. However, sensitivity and specificity are often inversely related. That means an increase in sensitivity often results in a decrease in specificity and vice versa. Sensitivity is also known as true positive rate and indicates the ability of the system to correctly identify a fall event. Specificity is known as true negative rate and indicates the ability of the system to correctly identify a non-fall event. High sensitivity is often required in tests or systems that identify a serious but treatable disease or condition. As falls are life-threatening events whose damage can be mitigated easily if detected early, maximizing sensitivity should be a higher priority than maximizing specificity. Maximizing sensitivity would also increase the false positive rate that means the rate of detecting non-fall events as fall events would increase. However, it is better to be safe than risking a person's life.

Table 6 summarizes the keywords "Fall detection", "Fall recognition", "Fall Localization", "Fall Rescue", and "Real-Time" used in the reviewed deep learning based fall detection systems to describe their proposed system and method. It is shown that almost all systems have used the key term "Fall detection" for describing their methods and systems. A few number of systems have also used the key term "Fall recognition" [74], [77]–[79], [85], [100], [102] and "Fall localization" [69], [73], [76], [80], [83], [85], [97], [100], [118] along with the key term "Fall detection". Hence, detection and recognition can be interchangeably used in

terms of fall events. After detecting a fall incident, it is very necessary to contact emergency services, as most fatal or life threatening events do not result from the falls themselves, but from delayed treatments or responses. Table 6 shows whether systems are explicitly connected to any fall response services such as emergency contacts, ambulance, or hospital services via online messages or SMS. Out of all the reviewed systems, only the systems proposed in [67] and [70] are explicitly connected to fall rescue and emergency healthcare services. Almost all of the reviewed systems have experimented on different methods to discriminate among fall and ADL events, however, very few systems have proposed a real-life prototype or a complete fall detection and rescue solution.

Real-time monitoring is another important factor in fall detection systems, as fall events occur in a very short period of time and quick response after fall events are paramount to patient survivability. Out of the reviewed systems, the systems proposed in [64], [67], [70], [72], [80]–[82], [85], [98], [100], [102], [104], and [105] have reported real-time performance.

Computational power and storage are also important considerations for fall detection systems. Developing deep learning systems require significant amount of storage to store the large datasets and powerful computational devices such as multicore CPUs, multiple GPUs, TPUs, and large amounts of RAM. However, due to the advent of cheap cloud computing platforms, and the ever-decreasing costs of computing devices, researchers no longer need to concern themselves with storage and computational power restrictions while developing deep learning systems. Recently, embedded AI computing devices such as the NVIDIA Jetson TX2 have enabled the researchers to deploy computation-heavy deep learning solutions very easily in real-world scenarios. However, the prices of such devices are still very high. Most of the reviewed fall detection systems were trained in online cloud computing platforms, or in high-configuration personal computers. As most of the reviewed systems did not propose real-world prototype of the entire fall detection and rescue systems, power consumption was not reported in the most cases.

The sensing devices in the reviewed systems were just used to record the data and sent the data to servers. While deploying on real-world scenarios, most of the reviewed fall detection systems would need to employ variants of client-server architecture due to the computation heavy nature of the models. The sensing devices would then need Internet connectivity capabilities

Most of the commonly used open datasets used in the reviewed systems have been summarized in Table 7. Information regarding the datasets such as year of publication, number of subjects along with male and female ratio, age range of the volunteers, scenario of the data collection zone, number of types in the dataset along with fall and ADL type information, number of samples in the dataset along with ADL and fall ratio, used sensors in the dataset, number of

**TABLE 6.** Keyword summary of the reviewed deep learning based fall detection systems.

Authors	Fall Detection	Fall Recognition	Fall Localization	Fall Rescue	Real-Time
Adhikari et al.[63]	✓	×	×	×	×
Li et al. [64]	✓	×	×	×	✓
Yhdego et al. [65]	✓	×	×	×	×
Yu et al. [66]	✓	×	×	×	×
Shen et al. [67]	✓	×	✓	✓	✓
Santos et al. [68]	✓	×	×	×	×
Zhou et al. [69]	✓	×	✓	×	×
He et al. [70]	✓	×	×	✓	✓
Sadrezami et al. [71]	✓	×	×	×	×
Tasoulis et al. [72]	✓	×	×	×	✓
Lu. et al [73]	✓	×	✓	×	×
Wang et al. [74]	✓	✓	×	×	×
Zhang and Zhu [75]	✓	×	×	×	×
Cameiro et al. [76]	✓	×	✓	×	×
Casilari et al. [77]	✓	✓	×	×	×
Espinosa et al. [78]	✓	✓	×	×	×
Cho and Yoon [79]	✓	×	×	×	×
Tsai and Hsu [80]	✓	✓	✓	×	✓
Rahneemofar and Alkittawi [81]	✓	×	×	×	✓
Li et al. [82]	✓	×	×	×	✓
Hwang et al.[83]	✓	×	✓	×	×
Kasturi et al. [84]	✓	×	×	×	×
Hsieh and Jeng [85]	✓	✓	✓	×	✓
Lu et al. [96]	✓	×	×	×	×
Lie et al. [97]	✓	×	✓	×	×
Abobakr et al. [98]	✓	×	×	×	✓
Xu et al. [99]	✓	×	×	×	×
Tao and Yun [100]	✓	✓	✓	×	✓
Ajerla et al. [102]	✓	✓	×	×	✓
Queralta et al. [103]	✓	×	×	×	×
Perejon et al. [104]	✓	×	×	×	✓
Torti et al. [105]	✓	×	×	×	✓
Theodoridis et al. [106]	✓	×	×	×	×
Lin et al. [107]	✓	×	×	×	×
Ge et al. [108]	✓	×	×	×	×
Jokanovic et al. [116]	✓	×	×	×	×
Droghini et al. [117]	✓	×	×	×	×
Zhou and Komuro [118]	✓	×	✓	×	×
Seyfioglu et al. [119]	✓	×	×	×	×

sensing points in the dataset, and the positions of the sensing points have been summarized.

### B. RECOMMENDATIONS FOR FUTURE RESEARCH

Various systems have been developed and proposed for recognition of fall events based on deep learning. Though it is a difficult task to develop an elderly fall dataset to train the system effectively, it is still a very important task to do so. Majority of the systems [63]–[75] used only CNN based approach where 3D CNN or CNN with 10-fold cross-validation or hybrid method can be used for better performance. Other CNN architectures such as R-CNN, Faster

R-CNN can be explored. Furthermore, LSTM can be a choice, it can contribute to increase performance of overall fall detection system. Though LSTM is used combined with RNN, RCN in [96]–[98], [102], [103], [108] but performance is not better than CNN based proposed system. Hence for development and implementation of fall detection system, LSTM combined with other deep learning method as well as RNN, RCN can be a future research. The auto-encoder based systems in [116]–[119] didn't exhibit substantial performance increase. Hence fall detection with auto-encoder approach needs to be more effectively developed.

**TABLE 7. Summary of the open fall detection datasets used in reviewed systems.**

Dataset	Year	Number of Subjects (M/FM)	Age Range	Scenario	Number of Types (ADLS/Falls)	Number of Samples (ADLS/Falls)	Used Sensors	Number of Sensing Points	Positions of Sensing Points
MultiCAM	2010	1(N/A)	N/A	Home	7/4	192(N/A)	C	8	Ce
FDD	2012	N/A	N/A	Home, Office, Coffee Room, Lecture Room	6/3	250(57/193)	C	1	Ce
MobiFall	2013	24(17/7)	[22-47]	Gym Hall	9/4	630 (342/288)	A, G, O	1	T
SDUFall	2014	10	N/A	Lab	1/5	1800(1500/300)	K	1	Ce
TST	2014	11 (N/A)	[22-39]	Office, Lab	4/4	264(132/132)	A	2	Wa, Wr
URFD	2014	6(6/0)	>26	Office, Home	5/4	70(40/40)	A	1	Wa
MobiAct	2016	57(42/15)	[22-47]	Gym Hall	9/4	2526(1879/647)	A, G, O	1	T
UMAFall	2016	17(10/7)	[18-55]	Home	8/3	531(322/209)	A, G, M	5	An, Ch, T, Wa, Wr
NTU	2016	40(N/A)	[10-35]	Lab	1/59	56000(N/A)	K	3	Ce
SisFall	2017	38(19/19)	[19-75]	Gym Hall	19/15	4505(2707/1798)	A, A, G *	1	Wa
UniMiB SHAR	2017	30(24/6)	[18-60]	N/A	9/8	7013(5314/1699)	A	1	T
UP-Fall	2019	17(9/8)	[18-24]	Lab	6/5	561(306, 255)	IR, C, A, G, L, E	14	H, N, Wa, Wr, An, F

Note: N/A: Not Appropriately defined, C: RGB camera, A: Accelerometer, G: Gyroscope, O: Orientation Measurements, K: Kinect sensor, M: Magnetometer, IR: InfraRed sensor, L: Luminosity sensor, E: Electroencephalography (EEG) headset, Ce: Ceiling (mounted high above from ground), T: Thigh (left or right trouser pocket), Wa: Waist, Wr: Wrist, An: Ankle, Ch: Chest, H: Head, N: Neck, F: Floor, \* The employed sensing unit contained two accelerometers from two different vendors

A total of three sensing modalities have been discussed in this review which are as: Camera, Accelerometer, and Radar.

Two different types of cameras have been used in the reviewed systems such as: general RGB camera, and Kinect sensor. RGB camera based systems are low-cost. Also, they do not hinder the user's movement unlike wearable fall detection systems. However, RGB camera based systems are not privacy-preserving, hence they cannot be freely used in certain areas such as: bathrooms, toilets, etc. which are environments with significant fall risk. In addition, RGB camera based systems that are connected to the Internet are prone to hackers. These disadvantages can be countered by using Kinect sensor or depth camera based systems. Depth cameras only capture the depth information about the user's location. This depth information needs further processing before use. Hence, this information is not readily exploitable. Thus, depth camera based systems can be used in areas such as the toilets, washrooms, etc. too. Kinect sensors based systems are often coupled with other sensors such as accelerometer, gyroscope to increase the overall performance of the system. Vision based systems are very background specific. Changing the background greatly affects the overall performance of the system negatively. Most of the datasets used by vision based

systems are recorded in closed environments. In real-life scenarios, the background recorded by the devices would be very diverse, which would make most of the vision based systems less-effective.

Accelerometer is a very commonly used sensor in fall detection systems. They promote the development of wearable or embedded fall detection systems. Accelerometers are currently a standard sensor in all smartphones. This integration tremendously helps developing a fall detection and recognition system in a single device. The main advantages of using an accelerometer as a fall detection sensing modality are: they are very cheap, generated data requires little to no pre-processing, very energy efficient, and ease of use, etc. However, there's a social stigma related to wearable fall detection systems. Although, due to the ever-decreasing size of embedded systems, many accelerometer based fall detection systems can be implemented in very small form factor, which makes them barely noticeable.

Fall detection systems that used radars as sensing modalities have the same advantages and disadvantages as fall detection systems that use Kinect or depth cameras as sensors. Radars preserve user's privacy and can be mounted in areas such as washrooms, toilets, etc. However, the generated

**TABLE 8. Summary of the shortcomings of general deep learning based fall detection systems and recommendations.**

Type	Concerns	Recommendations
Technology-specific concerns	<ul style="list-style-type: none"> <li>▪ CNNs are prone to overfitting easily.</li> <li>▪ CNNs are very sensitive towards the training data. Thus, unforeseen changes in real-life data such as rotation, lighting, shape, etc. might reduce the performance of the CNN.</li> <li>▪ CNNs are open to adversarial attacks.</li> <li>▪ LSTMs are harder to train and deploy due to high memory requirements.</li> <li>▪ LSTMs require long training time and extensive hyperparameter tuning.</li> <li>▪ Deep learning based fall detection systems require specific hardware such as GPUs, multi-core CPUs, large amounts of ram, etc., thus making them harder to deploy in real-world scenarios.</li> <li>▪ Social stigma exists towards personalized fall detection systems in general.</li> </ul>	<ul style="list-style-type: none"> <li>▪ CNNs trained using 10-fold cross validation might be less prone to overfitting to training data and might perform better in real-world scenarios.</li> <li>▪ In video datasets, performance of object detectors such as context R-CNN, Faster R-CNN, YOLO, SSDs might be explored.</li> <li>▪ LSTMs can be combined with other high performing deep learning model to facilitate from the sequential and temporal nature of fall events.</li> <li>▪ Augmentation can be used to make the CNN based deep learning systems more generalized.</li> <li>▪ Research for employing deep learning solutions in embedded systems might be beneficial.</li> <li>▪ Fall detection systems can be provided as standard applications for smartwatches, smartphones, etc. due to advances such as Core ML, TFLite, ML Kit, etc.</li> <li>▪ Embedded AI systems such as the different variants of OpenCV AI Kit might be perfect for vision-based standalone fall detection solutions.</li> </ul>
Sensor-specific concerns	<ul style="list-style-type: none"> <li>▪ RGB cameras are not privacy-preserving.</li> <li>▪ Vision systems in general are in some cases background-specific and suffer from common pitfalls such as occlusion, scale variation, etc.</li> <li>▪ Radar based systems in general perform worse in environments that have multi static systems.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Floor sensors, IR arrays, RGB-Depth cameras or other privacy preserving sensors might be used in places such as washrooms, toilets, etc.</li> <li>▪ Custom fall detection system for each monitoring environment might counter the background-specific nature of vision systems.</li> <li>▪ Multi-camera setup might counter common problems such as occlusion and scale variation of vision systems.</li> <li>▪ Deep learning based fall detection systems employing multiple sensors such as wearable sensors, vision systems, and floor sensors might help in monitoring the entire living environment such as homes, hospitals, etc. and reduce the number of false negatives.</li> </ul>
Security and safety aspects	<ul style="list-style-type: none"> <li>▪ Deep learning based fall systems that depend on network connectivity and have client-server architecture are more prone to hackers.</li> <li>▪ Most fall detection systems have no failsafe protocols. So, if the primary fall detection system fails due to connectivity issues, power issues, etc., the entire fall detection system fails to provide safety for the patients.</li> <li>▪ Most deep learning based fall detection systems do not have any way to directly connect to the fall rescue operations such as emergency response services, healthcare workers, etc.</li> </ul>	<ul style="list-style-type: none"> <li>▪ End-to-end encryption schemes might be used while sending sensitive information such as RGB camera, radar outputs to server. However, this might increase the overall latency of the system.</li> <li>▪ Deep learning based fall detection systems must have failsafe protocols that would work when the primary detection systems fails due to power failure, network infrastructure failure, etc.</li> <li>▪ Fall detection systems must have ways to one-tap connect to emergency response services for minimizing health risk.</li> </ul>
Data-specific concerns	<ul style="list-style-type: none"> <li>▪ Fall detection datasets are mostly collected in controlled lab environments, thus they do not reflect the real-world fall data.</li> <li>▪ Deep learning systems in general need large amounts of data.</li> <li>▪ Fall detection systems in general are not person-specific.</li> </ul>	<ul style="list-style-type: none"> <li>▪ More effort should be put in collecting fall data in real-life environments.</li> <li>▪ Multiple similar datasets might be combined to build larger datasets for training complex deep learning systems.</li> <li>▪ Fall detection systems might be incorporated with person-specific data such as medical reports, history, etc. to provide a more custom safety experience.</li> </ul>

data requires further processing before it can be effectively used. Systems using radar as capturing device may exhibit degraded performance when multi-static systems are invoked

in the surrounding environment. Hence those situations and problems can be handled more efficiently in the future development of fall detection systems.



Security is also a big issue. Most of the reviewed systems did not mention anything about the security of the systems. As most of the devices used in the proposed systems are connected to the Internet for communication and data processing purposes, they are prone to intruders and hackers. Most of the public datasets were generated by recording ADL and fall data in closed protected environments. However, real-life environments are very diverse. More emphasis should be given on datasets that reflect sensor data from real-life scenarios. Similar kinds of public datasets can be combined to develop larger benchmark datasets. Testing the proposed systems on those benchmark datasets would make a comparison between methods easier, and thus help in developing more robust and accurate methods. Deep learning based fall detection systems are very heavy computing. Further research on making these systems and networks lightweight and deployable in embedded systems may be done. Most of the reviewed fall detection systems fail to detect falls in a multi-person fall scenario. These systems also have trouble with separating ADL events such as lying, crouching, etc. from fall events.

Almost all of the reviewed fall detection systems lack failsafe protocols. Most of the fall detection systems rely on network infrastructure, constant large amounts of power supply, etc. But, no failsafe protocols are mentioned that would still detect falls and provide minimal emergency response services in case of power failure, network dis-connectivity, etc. Surveillance based methods and wearable methods can be combined to develop more robust fall detection systems. Current fall detection systems in general reviewed in this literature are concerned with falls of the elderly, diseased, etc. However, falls and fall related casualties are not limited to these cases. Other types of falls such as falling from rock climbing sessions, fatal falls of babies and young children, falling in various sports competitions, etc. should also be taken into account, and systems should be developed for detecting falls in those scenarios. Current fall detection systems are general systems, they are not patient-specific. However, current advances in artificial intelligence enable us to develop system that cater to specific patients. These systems should be able to adapt to identify fall patterns of individuals based on their personal characteristics such as weight, age, height, illness, activity history, etc. Research should also be focused on fall prediction and fall protection systems that can predict potential falls, or mitigate the effects of the fall, respectively. Various neurological illness and certain diseases have strong correlation with fall incidents in patients. However, most of the reviewed fall detections systems are self-contained and thus have no prior information about the monitored individual or patient. Prior knowledge about the monitored individual's medical history might help smart fall detection systems to be more patient-specific and cater to the individual's specific needs. However, this comes with security and privacy concerns. Most of the fall detection systems have used different evaluation metrics to evaluate the performance of the systems. A major drawback of the fall

detection systems in general is the lack of a simplified standard evaluation procedure. This makes comparison among the fall detection systems very difficult, as different evaluations metrics are generally not comparable. Further research focusing on the development of a generalized standard metric for fall detection systems should be performed. As sensors are crucial parts for any fall detection system, the position of the sensors plays a huge role. However, different fall detection systems employ different sensor positions. Like the evaluation metrics, future standardization of the sensor positions is necessary, as it is useless to compare performance metrics from sensors placed in different positions. In the future, research should also be focused on detection falls in multi-person scenarios.

Deep learning based fall detection systems also share the general shortcomings of the principal methods used to infer the events. CNNs are not the ultimate model of the human visual system. CNN based systems are very specific with regards to rotation, orientation, lighting, subject shape, environment, etc. Thus, most of the CNN based fall detection systems are only applicable in real-life when the monitored environment is identical to the environment of the used dataset. This problem can be partially bypassed by using data augmentation as well as combinations of datasets representing various scenarios. Also, CNN based systems are open to adversarial attacks. However, it is still a relatively new field and a more research is required to develop a decent algorithm that can defend against adversarial attacks. LSTMs require longer training time and more rigorous hyper-parameter tuning. Also, LSTMs require more memory than similar methods. Thus, it is hard to deploy solutions that employ LSTMs. In recent years, research is being focused on optimizing deep learning hardware. These researches are helping to reduce the chip size or circuit size required to deploy proper deep learning solutions. These researches can help to pave a way towards deploying more exotic and memory or CPU intensive deep learning based fall detection systems. Auto-encoders are very data specific. The extent of an auto-encoders utility is limited by the extent of the diversity among the training data. Thus, auto-encoders based system can also benefit from mixed datasets that represent various environmental scenarios. The summary of the general shortcoming of deep learning based fall detection systems and future recommendations is presented in Table 8.

#### IV. CONCLUSION

This review presents an overview of recent advances in deep learning based fall detection systems. The reviewed systems have been categorized as: CNN based systems, LSTM based systems, and Auto-encoder based systems. The majority of the related works have used CNN for developing fall detection systems. 3D CNN or CNN with 10-fold cross-validation performed the best among the reviewed systems. Two types of hardware technology are prevalent: sensor based systems and vision based systems. The goal of the reviewed fall detection systems was to detect cases of elderly people falling using

the best deep learning methods and to inform a nearby nurse or support/medical staff within a short time. We hope that this review will be helpful for researchers who are interested in developing accurate and robust deep learning based fall detection systems in the future.

## REFERENCES

- [1] United Nations, "World population prospects: The 2017 revision, key findings and advance tables," Dept. Econ. Social Affairs, Population Division, New York, NY, USA, Working Paper ESA/P/WP/248, 2017.
- [2] World Health Organization. Accessed: Sep. 25, 2019. [Online]. Available: [http://www.searo.who.int/entity/health\\_situation\\_trends/data/chi/elderly-population/en/](http://www.searo.who.int/entity/health_situation_trends/data/chi/elderly-population/en/)
- [3] *Falls in Older Persons: Risk Factors and Prevention*. Accessed: Oct. 15, 2019. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK235613/>
- [4] E. Burns and R. Kakara, "Deaths from falls among persons aged  $\geq 65$  years—United States, 2007–2016," *Morbidity Mortality Weekly Rep.*, vol. 67, no. 18, p. 509, 2018.
- [5] S. R. Lord, H. B. Menz, and C. Sherrington, "Home environment risk factors for falls in older people and the efficacy of home modifications," *Age Ageing*, vol. 35, no. 2, pp. ii55–ii59, Sep. 2006.
- [6] P. Vallabh and R. Malekian, "Fall detection monitoring systems: A comprehensive review," *J. Ambient Intell. Humanized Comput.*, vol. 9, no. 6, pp. 1809–1833, Nov. 2018.
- [7] P. Wang, C.-S. Chen, and C.-C. Chuan, "Location-aware fall detection system for dementia care on nursing service in evergreen inn of Jianan hospital," in *Proc. IEEE 16th Int. Conf. Bioinf. Bioeng. (BIBE)*, Oct. 2016, pp. 309–315.
- [8] N. Doulamis, "Viusal fall alert service in low computational power device to assist persons' with dementia," in *Proc. 3rd Int. Symp. Appl. Sci. Biomed. Commun. Technol. (ISABEL)*, Nov. 2010, pp. 1–5.
- [9] O. Rawashdeh, W. Sa'deh, M. Rawashdeh, G. Qu, M. Ferrari, B. Harrison, R. Hammond, and M. Maddens, "Development of a low-cost fall intervention system for hospitalized dementia patients," in *Proc. IEEE Int. Conf. Electro/Inf. Technol.*, May 2012, pp. 1–7.
- [10] B. Homann, A. Plaschig, M. Grundner, A. Haubenhofer, T. Griedl, G. Ivanic, E. Hofer, F. Fazekas, and C. N. Homann, "The impact of neurological disorders on the risk for falls in the community dwelling elderly: A case-controlled study," *BMJ Open*, vol. 3, no. 11, Nov. 2013, Art. no. e003367.
- [11] S. Nooruddin, M. Milon Islam, and F. A. Sharna, "An IoT based device-type invariant fall detection system," *Internet Things*, vol. 9, Mar. 2020, Art. no. 100130.
- [12] T. Prajapati, N. Bhatt, and D. Mistry, "A survey paper on wearable sensors based fall detection," *Int. J. Comput. Appl.*, vol. 115, no. 13, pp. 15–18, Apr. 2015.
- [13] X. Yu, "Approaches and principles of fall detection for elderly and patient," in *Proc. HealthCom 10th Int. Conf. e-Health Netw., Appl. Services*, Jul. 2008, pp. 42–47.
- [14] S. Chaudhuri, H. Thompson, and G. Demiris, "Fall detection devices and their use with older adults: A systematic review," *J. Geriatric Phys. Therapy*, vol. 37, no. 4, pp. 178–196, 2014.
- [15] Y. Delahoz and M. Labrador, "Survey on fall detection and fall prevention using wearable and external sensors," *Sensors*, vol. 14, no. 10, pp. 19806–19842, Oct. 2014.
- [16] Z. Zhang, C. Conly, and V. Athitsos, "A survey on vision-based fall detection," in *Proc. 8th ACM Int. Conf. Pervas. Technol. Rel. Assistive Environ. (PETRA)*, 2015, pp. 1–7.
- [17] C. Debes, A. Merentitis, S. Sukhanov, M. Niessen, N. Frangiadakis, and A. Bauer, "Monitoring activities of daily living in smart homes: Understanding human behavior," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 81–94, Mar. 2016.
- [18] F. Erden, S. Velipasalar, A. Z. Alkar, and A. E. Cetin, "Sensors in assisted living: A survey of signal and image processing methods," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 36–44, Mar. 2016.
- [19] F. Bagalà, C. Becker, A. Cappello, L. Chiari, K. Aminian, J. M. Hausdorff, W. Zijlstra, and J. Klenk, "Evaluation of accelerometer-based fall detection algorithms on real-world falls," *PLoS ONE*, vol. 7, no. 5, May 2012, Art. no. e37062.
- [20] S. S. Khan and J. Hoey, "Review of fall detection techniques: A data availability perspective," *Med. Eng. Phys.*, vol. 39, pp. 12–22, Jan. 2017.
- [21] M. Mubashir, L. Shao, and L. Seed, "A survey on fall detection: Principles and approaches," *Neurocomputing*, vol. 100, pp. 144–152, Jan. 2013.
- [22] R. Igual, C. Medrano, and I. Plaza, "Challenges, issues and trends in fall detection systems," *Biomed. Eng. OnLine*, vol. 12, no. 1, p. 66, 2013.
- [23] M. Islam, N. Neom, M. Imtiaz, S. Nooruddin, M. Islam, and M. Islam, "A review on fall detection systems using data from smartphone sensors," *Ingénierie des systèmes d'Inf.*, vol. 24, no. 6, pp. 569–576, Dec. 2019.
- [24] G. Koshmak, A. Loutfi, and M. Linden, "Challenges and issues in multi-sensor fusion approach for fall detection," *J. Sensors*, vol. 2016, pp. 1–12, Dec. 2016.
- [25] T. Xu, Y. Zhou, and J. Zhu, "New advances and challenges of fall detection systems: A survey," *Appl. Sci.*, vol. 8, no. 3, p. 418, Mar. 2018.
- [26] M. A. Baba Khel and M. Ali, "Technical analysis of fall detection techniques," in *Proc. 2nd Int. Conf. Advancements Comput. Sci. (ICACS)*, Feb. 2019, pp. 1–8.
- [27] L. Ren and Y. Peng, "Research of fall detection and fall prevention technologies: A systematic review," *IEEE Access*, vol. 7, pp. 77702–77722, 2019.
- [28] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [29] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285–1298, May 2016.
- [30] J. Ahmad, H. Farman, and Z. Jan, "Deep learning: Methods and applications," in *SpringerBriefs in Computer Science*. Singapore: Springer, 2019, pp. 31–42.
- [31] L. N. V. Colon, Y. DeLaHoz, and M. Labrador, "Human fall detection with smartphones," in *Proc. IEEE Latin-America Conf. Commun. (LAT-INCOM)*, Nov. 2014, pp. 1–7.
- [32] S.-L. Hsieh, C.-T. Yang, and H.-J. Li, "Combining wristband-type devices and smartphones to detect falls," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 2373–2377.
- [33] A. Z. Rakhman and L. E. Nugroho, "Fall detection system using accelerometer and gyroscope based on smartphone," in *Proc. 1st Int. Conf. Inf. Technol., Comput., Electr. Eng.*, Nov. 2014, pp. 99–104.
- [34] B. Aguiar, T. Rocha, J. Silva, and I. Sousa, "Accelerometer-based fall detection for smartphones," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2014, pp. 1–6.
- [35] V. R. L. Shen, H.-Y. Lai, and A.-F. Lai, "Application of high-level fuzzy Petri nets to fall detection system using smartphone," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 3, Jul. 2013, pp. 1429–1435.
- [36] F. Wu, H. Zhao, Y. Zhao, and H. Zhong, "Development of a Wearable-Sensor-Based fall detection system," *Int. J. Telemedicine Appl.*, vol. 2015, pp. 1–11, Feb. 2015.
- [37] P. Vallabh, R. Malekian, N. Ye, and D. C. Bogatinoska, "Fall detection using machine learning algorithms," in *Proc. 24th Int. Conf. Softw., Telecommun. Comput. Netw. (SoftCOM)*, Sep. 2016, pp. 1–9.
- [38] A. Rahaman, M. M. Islam, M. R. Islam, M. S. Sadi, and S. Nooruddin, "Developing IoT based smart health monitoring systems: A review," *Revue d'Intelligence Artificielle, Lavoisier*, vol. 33, no. 6, pp. 435–440, 2019.
- [39] M. Luštrek and B. Kaluža, "Fall detection and activity recognition with machine learning," *Informatica*, vol. 33, no. 2, pp. 205–212, 2009.
- [40] K. de Miguel, A. Brunete, M. Hernando, and E. Gamba, "Home camera-based fall detection system for the elderly," *Sensors*, vol. 17, no. 12, p. 2864, Dec. 2017.
- [41] G. Rescio, A. Leone, and P. Siciliano, "Supervised machine learning scheme for electromyography-based pre-fall detection system," *Expert Syst. Appl.*, vol. 100, pp. 95–105, Jun. 2018.
- [42] A. Shahzad and K. Kim, "FallDroid: An automated smart-phone-based fall detection system using multiple kernel learning," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 35–44, Jan. 2019.
- [43] J. C. Dogan and M. S. Hossain, "A novel two-step fall detection method using smartphone sensors," in *Proc. IEEE Int. Conf. Smart Comput. (SMARTCOMP)*, Jun. 2019, pp. 434–438.

- [44] M. Cheffena, "Fall detection using smartphone audio features," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 4, pp. 1073–1080, Jul. 2016.
- [45] N. Yodpijit, T. Sittiwanchai, and M. Jongprasithporn, "The development of artificial neural networks (ANN) for falls detection," in *Proc. 3rd Int. Conf. Control, Autom. Robot. (ICCAR)*, Apr. 2017, pp. 547–550.
- [46] E. Casilari-Pérez and F. García-Lagos, "A comprehensive study on the use of artificial neural networks in wearable fall detection systems," *Expert Syst. Appl.*, vol. 138, Dec. 2019, Art. no. 112811.
- [47] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, and T. Chen, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018.
- [48] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural Comput.*, vol. 29, no. 9, pp. 2352–2449, Sep. 2017.
- [49] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, Jul. 2018.
- [50] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Müller, "Deep learning for time series classification: A review," *Data Mining Knowl. Discovery*, vol. 33, no. 4, pp. 917–963, Jul. 2019.
- [51] G. Yao, T. Lei, and J. Zhong, "A review of convolutional-neural-network-based action recognition," *Pattern Recognit. Lett.*, vol. 118, pp. 14–22, Feb. 2019.
- [52] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–13, Feb. 2018.
- [53] S. Wang, L. Chen, Z. Zhou, X. Sun, and J. Dong, "Human fall detection in surveillance video based on PCANet," *Multimedia Tools Appl.*, vol. 75, no. 19, pp. 11603–11613, Oct. 2016.
- [54] Y. Fan, M. D. Levine, G. Wen, and S. Qiu, "A deep neural network for real-time detection of falling humans in naturally occurring scenes," *Neurocomputing*, vol. 260, pp. 43–58, Oct. 2017.
- [55] H. Sadreazami, M. Bolic, and S. Rajan, "TL-FALL: Contactless indoor fall detection using transfer learning from a pretrained model," in *Proc. IEEE Int. Symp. Med. Meas. Appl.*, Jun. 2019, pp. 1–5.
- [56] M. Gochoo, T.-H. Tan, S.-H. Liu, F.-R. Jean, F. Alnajjar, and S.-C. Huang, "Unobtrusive activity recognition of elderly people living alone using anonymous binary sensors and DCNN," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 2, pp. 693–702, Mar. 2019.
- [57] Z. Zhang, X. Ma, H. Wu, and Y. Li, "Fall detection in videos with trajectory-weighted deep-convolutional rank-pooling descriptor," *IEEE Access*, vol. 7, pp. 4135–4144, 2019.
- [58] Y. Zhang and H. Neumann, "An empirical study towards understanding how deep convolutional nets recognize falls," in *Proc. Eur. Conf. Comput. Vis.* in Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 11134, 2019, pp. 112–127.
- [59] J. Quero, M. Burns, M. Razaq, C. Nugent, and M. Espinilla, "Detection of falls from non-invasive thermal vision sensors using convolutional neural networks," *Proceedings*, vol. 2, no. 19, p. 1236, Oct. 2018.
- [60] Y. Tian, G.-H. Lee, H. He, C.-Y. Hsu, and D. Katabi, "RF-based fall monitoring using convolutional neural networks," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–24, Sep. 2018.
- [61] A. H. Fakhruddin, X. Fei, and H. Li, "Convolutional neural networks (CNN) based human fall detection on body sensor networks (BSN) sensor data," in *Proc. 4th Int. Conf. Syst. Informat. (ICSAI)*, Nov. 2017, pp. 1461–1465.
- [62] M. D. Solbach and J. K. Tsotsos, "Vision-based fallen person detection for the elderly," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 1433–1442.
- [63] K. Adhikari, H. Bouchachia, and H. Nait-Charif, "Activity recognition for indoor fall detection using convolutional neural network," in *Proc. 15th IAPR Int. Conf. Mach. Vis. Appl. (MVA)*, May 2017, pp. 81–84.
- [64] X. Li, T. Pang, W. Liu, and T. Wang, "Fall detection for elderly person care using convolutional neural networks," in *Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Oct. 2017, pp. 1–6.
- [65] H. Yhdego, J. Li, S. Morrison, M. Audette, C. Paolini, M. Sarkar, and H. Okhravi, "Towards musculoskeletal simulation-aware fall injury mitigation: Transfer learning with deep CNN for fall detection," in *Proc. Spring Simulation Conf. (SpringSim)*, 2019, pp. 1–12.
- [66] M. Yu, L. Gong, and S. Kollias, "Computer vision based fall detection by a convolutional neural network," in *Proc. 19th ACM Int. Conf. Multimodal Interact. (ICMI)*, 2017, pp. 416–420.
- [67] L. Shen, Q. Zhang, G. Cao, and H. Xu, "Fall detection system based on deep learning and image processing in cloud environment," in *Proc. Adv. Intell. Syst. Comput.*, vol. 772, 2019, pp. 590–598.
- [68] G. Santos, P. Endo, K. Monteiro, E. Rocha, I. Silva, and T. Lynn, "Accelerometer-based human fall detection using convolutional neural networks," *Sensors*, vol. 19, no. 7, p. 1644, Apr. 2019.
- [69] X. Zhou, L.-C. Qian, P.-J. You, Z.-G. Ding, and Y.-Q. Han, "Fall detection using convolutional neural network with multi-sensor fusion," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2018, pp. 1–5.
- [70] J. He, Z. Zhang, X. Wang, and S. Yang, "A low power fall sensing technology based on FD-CNN," *IEEE Sensors J.*, vol. 19, no. 13, pp. 5110–5118, Jul. 2019.
- [71] H. Sadreazami, M. Bolic, and S. Rajan, "Fall detection using standoff radar-based sensing and deep convolutional neural network," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 1, pp. 197–201, Jan. 2020.
- [72] S. K. Tasoulis, G. I. Mallis, S. V. Georgakopoulos, A. G. Vrahatis, V. P. Plagianakos, and I. G. Maglogiannis, "Deep learning and change detection for fall recognition," in *Engineering Applications of Neural Networks (Communications in Computer and Information Science)*, vol. 1000, Cham, Switzerland: Springer, 2019, pp. 262–273.
- [73] N. Lu, Y. Wu, L. Feng, and J. Song, "Deep learning for fall detection: Three-dimensional CNN combined with LSTM on video kinematic data," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 1, pp. 314–323, Jan. 2019.
- [74] L. Wang, M. Peng, and Q. F. Zhou, "Fall detection based on convolutional neural networks using smart insole," in *Proc. 5th Int. Conf. Control, Automat. Robot.*, 2019, pp. 593–598.
- [75] Q. Zhang and S. Zhu, "Real-time activity and fall risk detection for aging population using deep learning," in *Proc. 9th IEEE Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, Nov. 2018, pp. 1055–1059.
- [76] S. A. Cameiro, G. P. da Silva, G. V. Leite, R. Moreno, S. J. F. Guimaraes, and H. Pedrini, "Multi-stream deep convolutional network using high-level features applied to fall detection in video sequences," in *Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Jun. 2019, pp. 293–298.
- [77] E. Casilari, R. Lora-Rivera, and F. García-Lagos, "A study on the application of convolutional neural networks to fall detection evaluated with multiple public datasets," *Sensors*, vol. 20, no. 5, p. 1466, Mar. 2020.
- [78] R. Espinosa, H. Ponce, S. Gutiérrez, L. Martínez-Villaseñor, J. Brieve, and E. Moya-Albor, "A vision-based approach for fall detection using multiple cameras and convolutional neural networks: A case study using the UP-fall detection dataset," *Comput. Biol. Med.*, vol. 115, Dec. 2019, Art. no. 103520.
- [79] H. Cho and S. M. Yoon, "Applying singular value decomposition on accelerometer data for 1D convolutional neural network based fall detection," *Electron. Lett.*, vol. 55, no. 6, pp. 320–322, Mar. 2019.
- [80] T.-H. Tsai and C.-W. Hsu, "Implementation of fall detection system based on 3D skeleton for deep learning technique," *IEEE Access*, vol. 7, pp. 153049–153059, 2019.
- [81] M. Rahmehoonfar and H. Alkittawi, "Spatio-temporal convolutional neural network for elderly fall detection in depth video cameras," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2018, pp. 2868–2873.
- [82] S. Li, H. Xiong, and X. Diao, "Pre-impact fall detection using 3D convolutional neural network," in *Proc. IEEE 16th Int. Conf. Rehabil. Robot. (ICORR)*, Jun. 2019, pp. 1173–1178.
- [83] S. Hwang, D. Ahn, H. Park, and T. Park, "Maximizing accuracy of fall detection and alert systems based on 3D convolutional neural network," in *Proc. 2nd Int. Conf. Internet Things Design Implement.*, Apr. 2017, pp. 343–344.
- [84] S. Kasturi, A. Filonenko, and K.-H. Jo, "Human fall recognition using the spatiotemporal 3D CNN," in *Proc. IW-FCV*, 2019, pp. 1–3.
- [85] Y.-Z. Hsieh and Y.-L. Jeng, "Development of home intelligent fall detection IoT system based on feedback optical flow convolutional neural network," *IEEE Access*, vol. 6, pp. 6048–6057, 2018.
- [86] R. Miikkulainen, J. Liang, E. Meyerson, A. Rawal, D. Fink, O. Francon, B. Raju, H. Shahrzad, A. Navruzyan, N. Duffy, and B. Hodjat, "Evolving deep neural networks," in *Artificial Intelligence in the Age of Neural Networks and Brain Computing*. New York, NY, USA: Academic, 2019, pp. 293–312.

- [87] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019.
- [88] V. Veeriah, N. Zhuang, and G.-J. Qi, "Differential recurrent neural networks for action recognition," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 4041–4049.
- [89] Y. Zhao, R. Yang, G. Chevalier, X. Xu, and Z. Zhang, "Deep residual bidir-LSTM for human activity recognition using wearable sensors," *Math. Problems Eng.*, vol. 2018, pp. 1–13, Dec. 2018.
- [90] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," in *Proc. Int. Joint Conf. Artif. Intell. (IJCAI)*, Jan. 2016, pp. 1533–1540.
- [91] Y.-C. Chang, Y. H. Lai, and T.-C. Huang, "Inferential motion reconstruction of fall accident based on LSTM neural network," *J. Med. Biol. Eng.*, vol. 39, no. 3, pp. 315–328, Jun. 2019.
- [92] A. Manashty, J. Light Thompson, and H. Soleimani, "A concise temporal data representation model for prediction in biomedical wearable devices," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1438–1445, Apr. 2019.
- [93] C. Taramasco, T. Rodenas, F. Martínez, P. Fuentes, R. Munoz, R. Olivares, V. H. C. De Albuquerque, and J. Demongeot, "A novel monitoring system for fall detection in older people," *IEEE Access*, vol. 6, pp. 43563–43574, 2018.
- [94] S. Mahfuz, H. Isah, F. Zulkernine, and P. Nicholls, "Detecting irregular patterns in IoT streaming data for fall detection," in *Proc. IEEE 9th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Nov. 2018, pp. 588–594.
- [95] I. Wayan Wiprayoga Wisesa and G. Mahardika, "Fall detection algorithm based on accelerometer and gyroscope sensor data using recurrent neural networks," in *Proc. IOP Conf. Ser., Earth Environ. Sci.*, 2019, vol. 258, no. 1, Art. no. 012035.
- [96] N. Lu, X. Ren, J. Song, and Y. Wu, "Visual guided deep learning scheme for fall detection," in *Proc. 13th IEEE Conf. Autom. Sci. Eng. (CASE)*, Aug. 2017, pp. 801–806.
- [97] W.-N. Lie, A. T. Le, and G.-H. Lin, "Human fall-down event detection based on 2D skeletons and deep learning approach," in *Proc. Int. Workshop Adv. Image Technol. (IWAIT)*, Jan. 2018, pp. 1–4.
- [98] A. Abobakr, M. Hossny, H. Abdelkader, and S. Nahavandi, "RGB-D fall detection via deep residual convolutional LSTM networks," in *Proc. Digit. Image Comput., Techn. Appl. (DICTA)*, Dec. 2018, pp. 1–7.
- [99] J. Xu, Z. He, and Y. Zhang, "CNN-LSTM combined network for IoT enabled fall detection applications," *J. Phys., Conf. Ser.*, vol. 1267, Jul. 2019, Art. no. 012044.
- [100] X. Tao and Z. Yun, "Fall prediction based on biomechanics equilibrium using Kinect," *Int. J. Distrib. Sensor Netw.*, vol. 13, no. 4, Apr. 2017, Art. no. 155014771770325.
- [101] C. Rougier and J. Meunier, "Fall detection using 3d head trajectory extracted from a single camera video sequence," *J. Telemed. Telecare*, vol. 11, no. 4, pp. 2–3, 2005.
- [102] D. Ajerla, S. Mahfuz, and F. Zulkernine, "A real-time patient monitoring framework for fall detection," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–13, Sep. 2019.
- [103] J. P. Queralta, T. N. Gia, H. Tenhunen, and T. Westerlund, "Edge-AI in LoRa-based health monitoring: Fall detection system with fog computing and LSTM recurrent neural networks," in *Proc. 42nd Int. Conf. Telecommun. Signal Process. (TSP)*, Jul. 2019, pp. 601–604.
- [104] F. Luna-Perejon, J. Civit-Masot, I. Amaya-Rodriguez, L. Duran-Lopez, and J. Pedro Dominguez-Morales, "An automated fall detection system using recurrent neural networks," in *Artificial Intelligence in Medicine* (Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11526. Cham, Switzerland: Springer, 2019, pp. 36–41.
- [105] E. Torti, A. Fontanella, M. Musci, N. Blago, D. Pau, F. Leporati, and M. Piastra, "Embedded real-time fall detection with deep learning on wearable devices," in *Proc. 21st Euromicro Conf. Digit. Syst. Design (DSD)*, Aug. 2018, pp. 405–412.
- [106] T. Theodoridis, V. Solachidis, N. Vretos, and P. Daras, "Human fall detection from acceleration measurements using a recurrent neural network," in *Proc. IFMBE*, vol. 66, 2018, pp. 145–149.
- [107] L. Hsiu-Yu, H. Yu-Ling, and L. Wen-Nung, "Convolutional recurrent neural networks for posture analysis in fall detection," *J. Inf. Sci. Eng.*, vol. 34, no. 3, pp. 577–591, 2018.
- [108] C. Ge, I. Y.-H. Gu, and J. Yang, "Co-saliency-enhanced deep recurrent convolutional networks for human fall detection in E-Healthcare," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 1572–1575.
- [109] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 2, 2012, pp. 1097–1105.
- [110] G. Dong, G. Liao, H. Liu, and G. Kuang, "A review of the autoencoder and its variants: A comparative perspective from target recognition in synthetic-aperture radar images," *IEEE Geosci. Remote Sens. Mag.*, vol. 6, no. 3, pp. 44–68, Sep. 2018.
- [111] M. P. Kenning, X. Xie, M. Edwards, and J. Deng, "Local representation learning with a convolutional autoencoder," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 3239–3243.
- [112] D. Charte, F. Charte, S. García, M. J. del Jesus, and F. Herrera, "A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines," *Inf. Fusion*, vol. 44, pp. 78–96, Nov. 2018.
- [113] A. Rangamani, A. Mukherjee, A. Basu, A. Arora, T. Ganapathi, S. Chin, and T. D. Tran, "Sparse coding and autoencoders," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 36–40.
- [114] D. Charte, F. Herrera, and F. Charte, "Ruta: Implementations of neural autoencoders in R," *Knowl.-Based Syst.*, vol. 174, pp. 4–8, Jun. 2019.
- [115] F. Charte, A. J. Rivera, F. Martínez, and M. J. del Jesus, "Automating autoencoder architecture configuration: An evolutionary approach," in *Understanding the Brain Function and Emotions* (Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11486. Cham, Switzerland: Springer, 2019, pp. 339–349.
- [116] B. Jokanovic, M. Amin, and F. Ahmad, "Radar fall motion detection using deep learning," in *Proc. IEEE Radar Conf. (RadarConf)*, May 2016, pp. 1–6.
- [117] D. Droghini, D. Ferretti, E. Principi, S. Squartini, and F. Piazza, "An end-to-end unsupervised approach employing convolutional neural network autoencoders for human fall detection," in *Quantifying and Processing Biomedical and Behavioral Signals* (Smart Innovation, Systems and Technologies), vol. 103. Cham, Switzerland: Springer, 2019, pp. 185–196.
- [118] J. Zhou and T. Komuro, "Recognizing fall actions from videos using reconstruction error of variational autoencoder," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 3372–3376.
- [119] M. S. Seyfioglu, A. M. Ozbayoglu, and S. Z. Gurbuz, "Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1709–1723, Aug. 2018.
- [120] H. Alkittawi, "A deep-learning-based fall-detection system to support aging-in-place," Ph.D. dissertation, Dept. Comput. Sci., Texas A&M Univ.-Corpus Christi, Corpus Christi, TX, USA, 2017.
- [121] S. S. Khan and B. Taati, "Detecting unseen falls from wearable devices using channel-wise ensemble of autoencoders," *Expert Syst. Appl.*, vol. 87, pp. 280–290, Nov. 2017.



**MD. MILON ISLAM** received the B.Sc. and M.Sc. degrees in computer science and engineering (CSE) from the Khulna University of Engineering & Technology, Khulna, Bangladesh, in 2016 and 2019, respectively. In 2017, he joined the Department of Computer Science and Engineering, Khulna University of Engineering & Technology, as a Lecturer. He is currently working as an Assistant Professor in CSE with the Khulna University of Engineering & Technology. He is

also a Reviewer of several reputed journals and conferences. In addition, he has published more than 33 research papers in peer reviewed journals and conferences. His research interests include machine learning and its application, deep learning, intelligent systems design, health informatics, the Internet of Things (IoT), and to solve real life problems with the concept of computer science.



**OMAR TAYAN** received the B.Eng. degree and the Ph.D. degree in computer networks from the University of Strathclyde, Glasgow, U.K. He is currently a Professor with the NOOR Research Center, College of Computer Science and Engineering (CCSE), Taibah University, Saudi Arabia. He was a Consultant with the Strategic and Advanced Research and Technology Innovation Unit, Taibah University. He is one of the founding members of the NOOR Research Center, Taibah

University. He has about 70 journals, conference papers, technical reports, and invited talks to his credit, as well as a book publication in computer-networks. He has successfully completed about 10–15 research and development projects as a Principle-Investigator and a Co-Investigator in projects funded by the King Abdulaziz City for Science and Technology (KACST), Ministry of Higher Education (MoHE), and the Deanship of Research, Taibah University. His research interests include information security, image processing, modeling and simulation, computer networks and networks-on-chip (NoC), and wireless sensor networks for intelligent transportation systems, including hajj transportation systems and crowd management.



**SHEIKH NOORUDDIN** received the B.Sc. degree in computer science and engineering (CSE) from the Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh, in 2020. He has published several research articles in peer reviewed journals, including the *Internet of Things: Engineering Cyber Physical Human Systems* (Elsevier). His research interests include computer vision, machine learning, deep learning, embedded systems, and the Internet of Things (IoT).



**MD. REPON ISLAM** was born in Dinajpur, Bangladesh. He is currently pursuing the B.Sc. degree in computer science and engineering (CSE) with the Khulna University of Engineering & Technology, Khulna, Bangladesh. He has developed some interesting projects using C, C++, Python, Java, HTML, CSS, PHP, ASP.net, SQL, Android, and iOS. His research interests include computer vision, machine learning, deep learning, and data science.



**MUHAMMAD NOMANI KABIR** received the M.Sc. degree in computational sciences in engineering and the Ph.D. degree in computer science from the University of Braunschweig, Germany. He is currently a Senior Lecturer with the Faculty of Computer Systems and Software Engineering, University Malaysia Pahang, Malaysia. His research interests include issues related to modeling and simulation, computational methods, image processing, and computer networks.



**MD. SAIFUL ISLAM** was born in Barisal, Bangladesh. He is currently pursuing the B.Sc. degree in computer science and engineering (CSE) with the Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh. His research interests include computer vision, competitive programming, database, data science, machine learning, and deep learning. He has already worked on some interesting projects throughout his study.



**MD. RABIUL ISLAM** received the M.Sc. degree in electrical and electronic engineering (EEE) from the Khulna University of Engineering & Technology (KUET), Bangladesh, in 2019. He is currently working as an Assistant Professor with the Department of EEE, Bangladesh Army University of Engineering and Technology (BAUET). His research interests include machine learning, pattern recognition, signal processing, and human computer interfacing.

...