Convolutional Neural Network-Based Fall Detection for the Elderly Person Monitoring

Kishanprasad G. Gunale [1], Prachi Mukherji [2], and Sumitra N. Motade [1,*]

[1] School of Electronics and Communication Engineering, Dr. Vishwanath Karad MIT World Peace University, Pune, India; Email: kishanprasad.gunale@mitwpu.edu.in (K.G.G.)

[2] Department of Electronics and Telecommunication Engineering, Cummins College of Engineering for Women, Savitribai Phule Pune University, Pune, India; Email: prachi.mukherji@cumminscollege.in (P.M.)

*Correspondence: sumitra.motade@mitwpu.edu.in (S.N.M.)

***Abstract*—The purpose of this paper is to present a generalized human fall detection system for elderly assistance. The human population above the age of 60 is constantly growing. India is the world's second-most peopled country, having 76.6 million individuals aged 60 and above, occupying more than 7.7% of the total population. Falls are considered a significant problem among the elderly. Falls are a significant source of mortality among the elderly. As a result, instant medical attention is required following a fall. When related to wearable sensors, vision-based fall detection is a more appropriate scheme for supervising old persons. Since a variety of backgrounds and scenes are available, fall event detection necessitates an intelligent approach for extracting the relevant feature. Deep learning algorithms have demonstrated very excellent classification performance in recent years. Compared to traditional techniques, fall detection systems using Convolution Neural Networks (CNN) are highly competent in detecting fall occurrences. The value of the proposed research lies in the CNN-based Fall Detection (FD) system which is data-independent. CNN is infamous for being difficult to tune and data-intensive. There is a risk of overfitting with only a few constructive cases of anomalies among hours of footage. The value of the proposed system lies in combination of diverse datasets from various human fall scenes of office, home, coffee room and lecture room for extraction of novel feature sets to be input to CNN model. The proposed system addresses vital issues the healthcare underthought by automated human fall detection with an accuracy of 93.81% by combining the FDD, MCFD, SDU, and URFD.**

***Keywords*—Convolution Neural Networks (CNN), computer vision, deep learning, elderly fall detection, healthcare**

I. Introduction

This study is especially motivated to contribute to the healthcare industry by developing an elderly assistance system. The world's population aging began in the last century with developed nations and surrounding developing nations. India is by no means exclusive to these phenomena. Population structure has evolved throughout time and continues to change in the future. The creation of elderly assist devices enhances the quality of life of older people who desperately need assistance. Falls account for more than 90% of hip fractures in those over 70, associated with higher mortality [1]. The country must be well prepared to face the difficulties of an older population. These incidents have driven research toward intelligent Fall Detection (FD) for the elderly needing immediate help after a fall. The majority of current research is focused on sensor-based FD. These methods are probably subject to constraints and lack resilience in vision-based conventional feature-based approaches. With advancements in the Internet of Things (IoT) and Deep Learning (DL), the possibility is to create a robust automatic FD scheme that can operate in dynamic scenes.

This work utilizes vision-based techniques for FD. Data monitoring for real-time application abnormalities is a crucial part since cameras are deployed in every part of daily life, resulting in massive data availability. Cameras are being put in elderly care centers, allowing us to monitor elderly people as supportive instruments for giving medical aid to them. These methods are beneficial to the elderly in the healthcare industry [2]. DL and available processing power has altered the Computer Vision (CV) environment. This study develops a novel Convolution Neural Networks (CNN) system, which can automatically extract the feature from the input image. It minimizes the dependency on handcrafted feature extraction requirements. The research objectives of this research are as given below.

(1) To combine diverse datasets for developing generalized HFD.
(2) To design a generalized human fall detection system and fine-tune parameters for improvement in the accuracy of the proposed method.

The structure of this paper is as follows. The performance of the current systems is described in Section II along with their methodology. The concluding part of this section presents the research gap and our contributions. The proposed approach for human FD is presented in Section III along with a full description of CNN's architecture for the FD system. Results and discussion of the proposed system using qualitative and quantitative analysis are presented in Section IV. The final section of

Manuscript received March 16, 2023; revised May 5, 2023; accepted May 31, 2023; published November 6, 2023.

doi: 10.12720/jait.14.6.1169-1176

this paper concludes the proposed approach with the future direction of the research.

II. Literature Review

This review serves two purposes: first, it will identify computer vision-based approaches for detecting human falls and second, it will highlight the gaps in previous research to support the suggested study.

The authors' strategy is to use a recent review [3] as a springboard for the literature review. In this work, the authors looked at non-intrusive (vision-based) DL-based HFD techniques' most recent (as of 2014) advancements. The importance of utilizing technology to assist the elderly is highlighted by the fact that there are an increasing number of older citizens in the world. A complete review of impending challenges, benchmark data, and DL methods in HFD has been provided by the authors.

Nouredanesh and McCormick [4] suggested mobile vision-based FD. Data is collected using a GoPro camera placed on the chest. This technique is wearable sensor-based, and a CNN extracts the features automatically for classification. The Mean Square Error of the system is 8%.

According to Ref. [5], a motion detection system built on passive RFID tags was developed, and the observed that a fast, static fall causes a change in Doppler frequency, which causes acceleration. These characteristics depict the condition of the old people.

Hsieh *et al.* [6] developed an optical flow feedback CNN utilizing rule-based filters preceding to input convolutional layer and saved optical flow to observe variation in optical flow. This system suggests an action rule for detecting fall events. The posture analysis for FD using DL is presented by Feng *et al.* [7]. To compare methods based on the Boltzmann machine with deep belief networks, the Support Vector Machine (SVM) technique is used.

Fouzi *et al.* [8] developed a multivariate exponentially weighted moving average monitoring method that successfully identifies falls due to minor changes in sensitivity. To correctly identify falls based on fall-like motions, SVM classification is utilized. This technique obtained an F-measure of 95%, comparable to other classifiers like K-Nearest Neighbors (KNN).

Dynamic Vision Sensor (DVS), based on an embedded FD system that operates in real-time, is recommended for use with conventional FD [9]. Falls are detected using an improved DL network on the DVS dataset with an F1-score of 95.5%. This system was built on the NVIDIA Jetson TX1 board.

Gunale and Mukherji [10] presents a fall detection technique for videos recorded with the help of depth camera. Feature extraction followed by SVM and Stochastic Gradient Descent (SGD) classifiers are used for the detecting the fall event. In this approach SGD classifier accuracy outperforms the SVM method on SDU fall dataset.

Min *et al.* [11] proposed employing scene analysis based on the DL approach quicker Region-based Convolutional Neural Networks (R-CNN) to identify human falls connected to sofas and furniture. This approach has a 95.5% accuracy rate in detecting falls from couches and sofas, whereas other systems had difficulty detecting such falls.

Nunez-Marco *et al.* [12] developed a vision-based FD system that uses Transfer Learning (TL) from action recognition to FD. The system utilizes publicly available datasets such as URFD, Multicam, and FDD. Rather than identifying optical flow and assembling them, this innovative technology detects the correlation between subsequent frames. It gets environmental independence by eliminating the characteristics that are dependent on aesthetics. To reduce reliance on feature engineering, CNN-specific features are employed. TL is used by the author to address the issue of fewer fall samples available in the dataset. On the collective dataset of URFD, MultiCam, and FDD, the proposed technique obtained sensitivity and specificity of 94% each.

For empirical investigations, building frameworks, and ensuring successful Lean Six Sigma (LSS) implementation, computer vision-based HFD systems need to be designed utilizing a lean six sigma-based strategy [13].

The proposed system contributes significantly in the following ways:

(1) The suggested system uses four distinct FD datasets for this method with various contexts, which aids in the generalization of the system.
(2) In order to boost the accuracy of fall detection, a unique feature set is extracted using CNN.

After a careful review of different approaches for elderly FD systems based on sensors and CV following observations are documented:

- Most of the systems were developed using sensors placed over the body of the elderly. However, it is difficult for the elderly to wear sensors on the body.
- FD systems implemented using a machine learning approach, require adequate feature extraction.
- The stated accuracy of HFD systems was on the lower side.

III. Methodology

The overall research methodology [14] employed for this study is shown in Fig. 1. After a careful literature review, it is seen that generalization of the HFD system is the major research area that needs to be addressed. A significant challenge in learning predictive models for AD is a scarcity of appropriate data. Anomalies by definition are rare events, so even if one has hours of videos, there might not be any abnormal event. Further, even when a video captures a rare anomaly, such events are often short-lived a person typically falls within a few milliseconds.

Conventional models ignoring this severe class imbalance will either produce a lot of false alarms or may not be able to detect an anomaly at all. Another challenge is the varying nature of environments.

For ML researcher's bias-variance trade-off has long been a dilemma and exciting research area. Further investigation for designing a robust and data-independent FD system encourages us to propose a CNN based approach for FD. CNN is known to be notoriously difficult to tune and data-demanding. With rare positive examples of anomalies in hours of videos, it is easy to get trapped into overfitting. To mitigate this risk, authors propose to combine various datasets, of different environments, and extract a novel set of features which are then fed to the deep models. These models have been shown to have accuracy at par with state-of-the-art models.

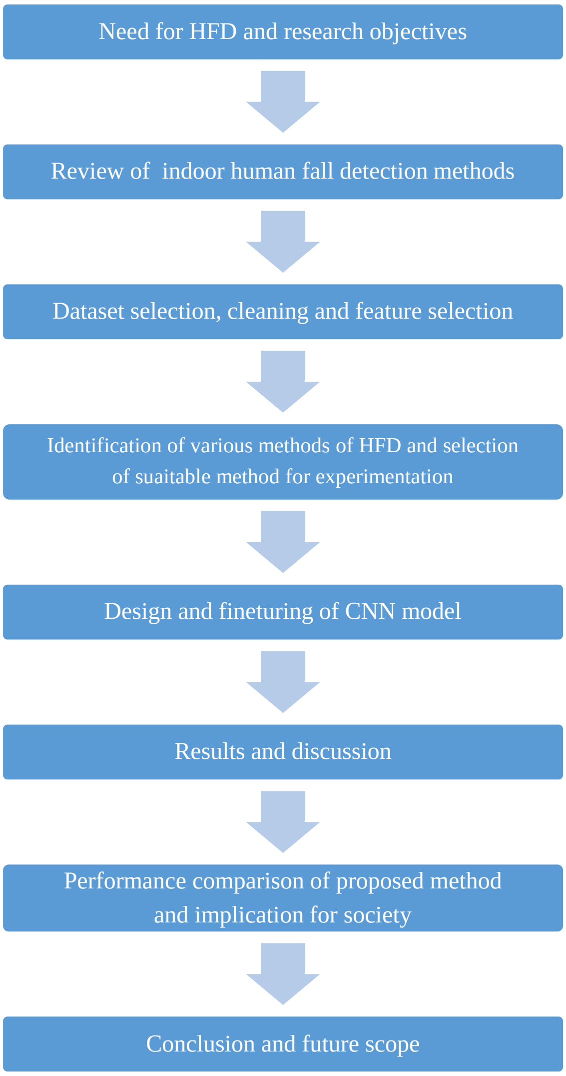


Figure 1. Methodology.

IV. Proposed CNN-Based FD

For early assistance, a vision-based method based on traditional feature engineering approaches is utilized generally in FD research. These techniques are not robust and are vulnerable to the limits of handmade feature extraction.

Fall event detection necessitates the use of an intelligent approach capable of automatically extracting FD characteristics in varying backgrounds and environments. Compared to traditional systems, CNN-based FD methods offer an effective identification of fall occurrences. With the advancement of the DL and IoT, it is now feasible to develop a reliable automatic FD system that can work in changing environments. DL and accessible computational power have altered the CV environment [15]. CNN's self-engineered capacity is used to produce a one-of-a-kind design. It automatically learns features, reducing the limits of hand-engineered feature extraction approaches. Fig. 2 illustrates the block diagram.

A. Input Fall Detection Dataset

This method uses four distinct indoor fall databases as input: FDD [16], MCFD [17], SDU [18], and URFD [19]. These datasets include both fall and no-fall sequences. Each database image is 320×240 in size and has three color planes (RGB). This system solely examines images that contain humans for further processing.

1) Le2i database

The videos in this collection were taken with a single camera in various interior places like an office, coffee room, lecture room, and home to analyze FD activities. 25 frames per second is the frame rate at which videos are recorded, and have 320×240 pixels resolution. The collection comprises 191 videos, each with its ground truth [13]. For training proposed system uses 91 videos from home and coffee room environments and 32 movies from both datasets are used to test the system.

2) Multiple fall database

From the viewpoint of the camera, the MCFD dataset is realistic that represents multiple falls in various directions and the many activities that occur in the everyday lives of older persons. The dataset is communal and challenging [17]. The camera captures numerous regular events from various angles, as illustrated in Fig. 3. This dataset contains critical events involving fall-like conditions, such as purposefully laying down. For analysis, this system uses MCFD with 22 falls and 5 non-fall videos.

3) SDU

SDU video database utilizes for the proposed approach for FD [18]. This dataset includes both female and male subjects taken using a Kinect camera with a video frame size of 320×240 at 30 frames per second.

4) University of Rzeszow (URFD)

It is a labeled dataset with fall events (30 videos) and everyday life activities (40 videos) [19]. Five people have recorded their fall occurrences, including a drop from a chair and a stand-up pose. It includes stooping, resting, scooping up an object off the ground, and reclining on a sofa.

B. Person Detection Using TensorFlow Object Detection Application Programming Interface (API)

This technique enhances system performance, reduces handcrafted feature engineering throughout the feature extraction process, and generalizes the system. The person detection in the scene uses TensorFlow object detection Application Programming Interface (API) [20]. This phase prevents the algorithm from being dependent on the look of the image. Human detection is dependent on backdrop intensity, and indoor and outdoor environments, and CV is too tough to detect the person. This method uses TensorFlow object detection API to recognize real-time objects in images. Huang *et al.* [20] offer a library that may trade off accuracy for speed and memory.

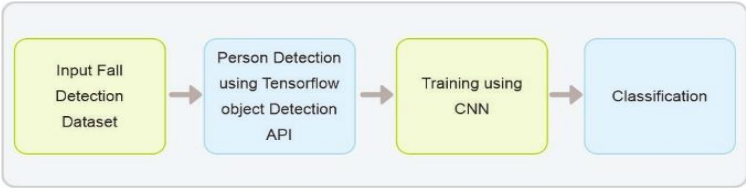


Figure 2. The CNN-based architecture of the FD system.

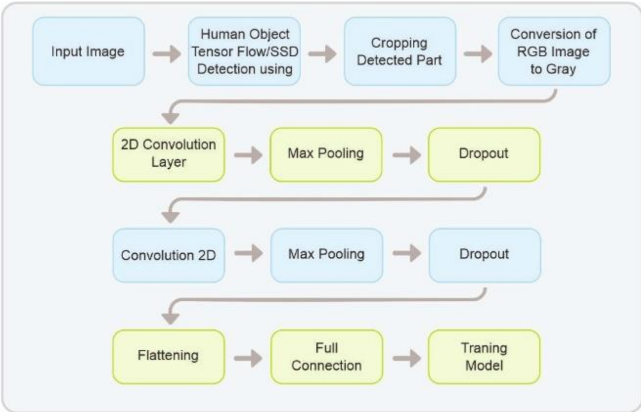


Figure 3. Architecture: CNN-based FD.

When training a model from scratch, a large dataset and time are required. The use of the TL idea along with a pre-trained model related to the TensorFlow API [21] helps to circumvent this issue. TL is an ML technique in which the advanced function's pre-trained model serves as the foundation for the next function. The benefit of adopting a pre-trained model is that the system may be taught using a pre-trained model for a comparable problem instead of starting from scratch. TL utilizes when there is insufficient data for the neural network to process new data.

The individual is initially recognized using the Tensor Flow Object Detection API in this function. TL employs the VGG16 model [22]. The database contains clipped and recorded identified individuals' regions of interest. The following phase involves manually sorting fall and no-fall human images in the database and training them using CNN.

C. Training Using CNN

Fig. 3. illustrates the architecture of the FD system based on CNN.

1) Input image

The images used as input are FDD [16], MCFD [17], SDU [18], and URFD [19]. These datasets include both fall and no-fall events. Each database image is 320×240 in size and has three color planes (RGB).

2) 2D convolution

The image is convolved with a feature detector as shown in Eq. (1), resulting in a 64×64 feature map.

$$(f \times g) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau) g(t-\tau) \tag{1}$$

where f represents the image, g represents the feature detector, and dummy variable τ indicates one function shift relative to the other. Following the creation of the feature map, the rectifier function is applied over the feature detector to boost the non-linearity of the function. The mentioned method utilizes Rectified Linear Unit (ReLU) and is formally represented by Eq. (2).

$$O_i = Re\, LU(X_J) = \max(0, X_J) \tag{2}$$

where X_J is the input, if $X_J > 0$ then $O_x = 0$ and if $X \geq 0$ then $O_X > X_j$.

3) Max pooling

An effective neural network would be invariant in terms of size and rotation. This system uses maximum pooling to make data scalable and scene invariant. With this technique, a mask with size $m \times n$ is created and convolved with the convolved image. The overall value of the mask is pooled to generate a new pooled featured map.

4) Flattening

Individual features are accepted by the Artificial Neural Network (ANN) classifier in the same manner that all classifiers do. In a word, ANN gets 1D data. Convolutional data is turned into a 1D feature vector by the process of flattening.

5) Fully connected layers

A fully connected layer consists of two interconnected layers. The first layer extracts 128 features. Overfitting may be avoided by carefully selecting features. Similarly, two out of 128 features indicating the feature's class (fall and No-fall) are evaluated in the second layer. This system defines a 20% dropout between two fully connected layers.

$$P(H) = \frac{e^H}{\sum_{c=1}^{C} e^H} \tag{3}$$

Eq. (3) provides the activation function's probability for target events and the last hidden layer is denoted by the letter H.

TABLE I. CNN LAYER DETAILS

Layers (Type)	Output Shape	Parameter
Input1	(None, 100, 100, 3)	0
Block1	(None, 100, 100, 64)	1792
Block1 conv2	(None, 100, 100, 64)	36928
Block1 pool1	(None, 50, 50, 64)	0
Block2 conv1	(None, 50, 50, 128)	73856
Block2 conv1	(None, 50, 50, 128)	147584

TABLE II. CNN LAYER 2 DETAILS

Layers (Type)	Output Shape	Parameter
block2 pool	(None, 25, 25, 128)	0
conv2d	(None, 23, 23, 128)	147584
max pooling2d	(None, 11, 11, 128)	0
Flatten	(None, 15488)	0
batch normalization	(None, 15488)	61952
dense 1	(None, 256)	3965184
dropout 1	(None, 256)	0
dense 2	(None, 128)	32896
dense 3	(None, 64)	8256
dense 4	(None, 1)	65

The constraints of the ML-based method include proper feature selection and optimization. The CNN-based FD technique collects FD features automatically and is appropriate for boosting performance. Tables I and II have information on CNN's layers.

V. DISCUSSION OF RESULTS

The primary goals of the proposed system design are to provide a solid database-independent system and reduce the hand-engineered feature extraction procedures. To accomplish these goals, the FD System employs a CNN-based automatic feature extractor. It performs generic feature extraction through suitable network parameters and training scheme adjustment. The FD System utilizes a CNN-based automated feature extractor to achieve these goals. The performance of the system is assessed separately on each dataset in this CNN-based technique. Table III summarizes the dataset distribution for the datasets utilized in this system.

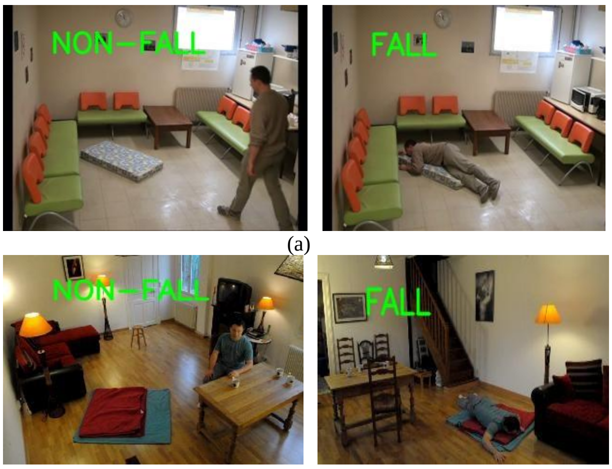
TABLE III. DATABASE INFORMATION USED FOR CNN-BASED FD

Database	Environment	Total Frames		Training		Testing	
		Fall	No-Fall	Fall	No-Fall	Fall	No-Fall
FDD [16]	Coffee Room	3989	21728	3752	20883	237	845
	Home	668	9137	590	8189	78	948
	Lecture Room	667	3880	257	2555	410	1325
	Office	1061	3465	385	3405	676	60
Multicam [17]	–	30290	135848	25002	119909	5288	15939
SDU [18]	–	1953	20342	1323	14203	630	6139
UR fall [19]	–	566	1737	453	1389	113	348

The proposed system evaluates the system's performance using qualitative and quantitative analyses. The mentioned experiment uses sensitivity, specificity, and accuracy as performance metrics since the parameters are appropriate for unbalanced class allocations. Because fall is a rare occurrence, there are fewer fall samples available related to no-fall activities.

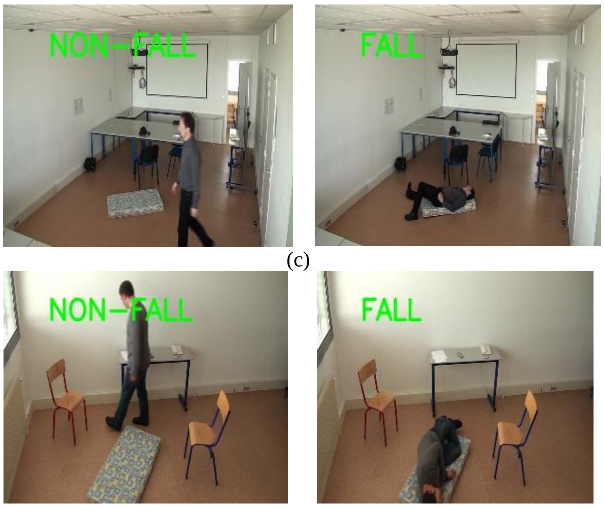
D. CNN-Based FD on Le2i/FDD

Fig. 4 depicts the qualitative analysis of the proposed CNN-based FD system on each Le2i dataset.



(a)

(b)

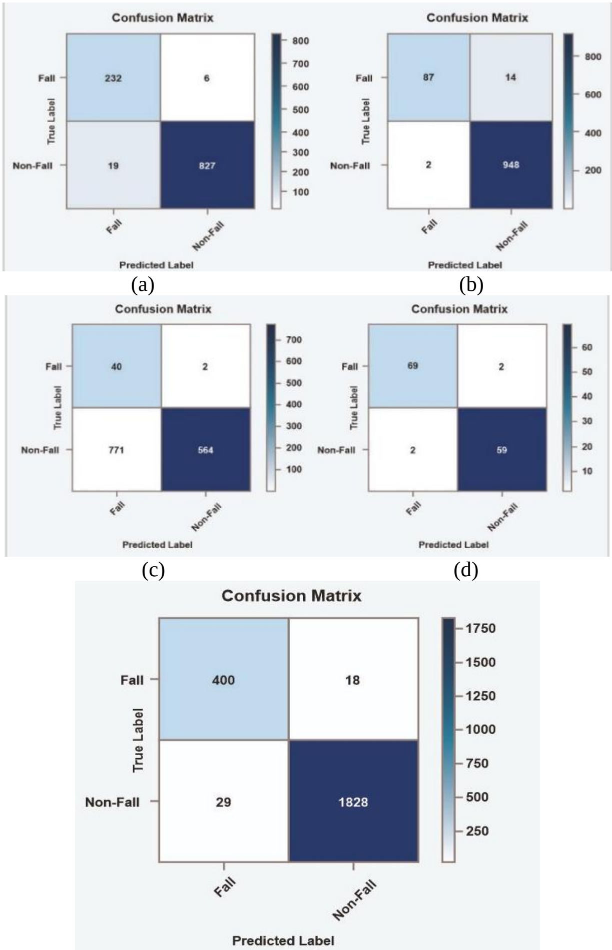


(c)

(d)

Figure 4. Qualitative analysis Le2i dataset of the environment. (a) Coffee room, (b) Home, (c) Lecture room. (d) Office.

Fig. 5 shows the confusion matrix of FDD environments. It also shows that false alarm in FDD, Coffee room, Home, and Office is significantly less, proving the system's robustness. CNN-based approach achieved above 99% of accuracy on different environments of Le2i. As the lecture room is a complex environment, accuracy is less.



(a) (b)

(c) (d)

(e) Combined all Le2i environments.

Figure 5. Confusion Matrix for FDD environments. (a) Coffee Room (b) Home (c) Lecture Hall (d) Office.

TABLE IV. COMPARISON OF CNN-BASED FD ON FDD

Proposal	**Sensitivity (%)**	**Specificity (%)**	**Accuracy (%)**
Charfi *et al*. [22]	98	99.6	–
Zerrouki and Houacine [23]	–	–	97.02
Nunez-Marcos *et al*. [12]	99	97	97
Our work	93	99	97.93

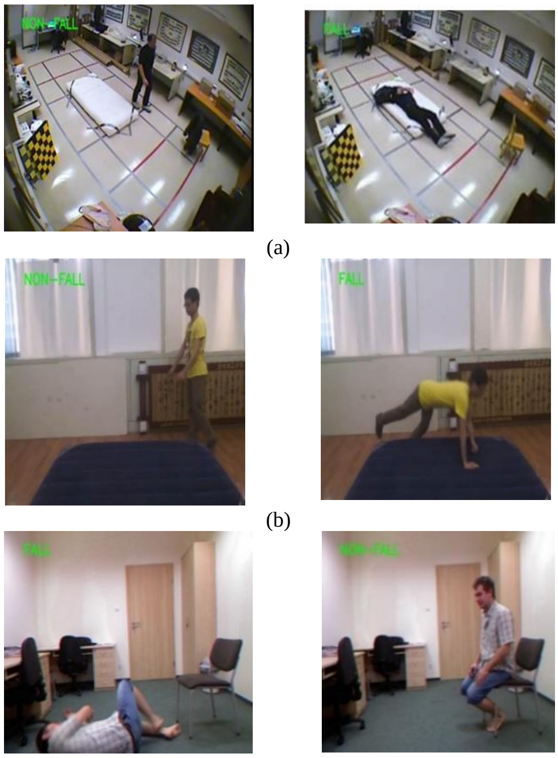
Table IV displays the results of FDD testing. FDD comprises four distinct habitats. For the system's robustness and generalization, this system merges datasets of four environments of the FDD dataset. This combination achieves an accuracy value of 97.93%, outperforming the current methods reported in [22–24].

E. CNN-Based FD on MCFD, SDU, and URFD

Fig. 6 depicts the qualitative analysis of the proposed CNN-based FD system on the MCFD, SDU, and URFD datasets [25–28].

In Fig. 7, each of these four dataset's confusion matrix values are displayed separately. According to the confusion matrix, CNN-based FD works well in various settings such as FDD, MCFD, SDU, and URFD.

Fig. 7 and Table V show qualitative and quantitative analyses of the proposed system utilizing the MCFD, SDU, and URFD datasets. MCFD uses several cameras to capture the fall activity, and the system reaches a specificity of 95%. SDU offers depth information for scenes with a 96.83% accuracy rating. When the scene is unstable, URFD takes a fall into account. This database has an accuracy value of 98.70%.

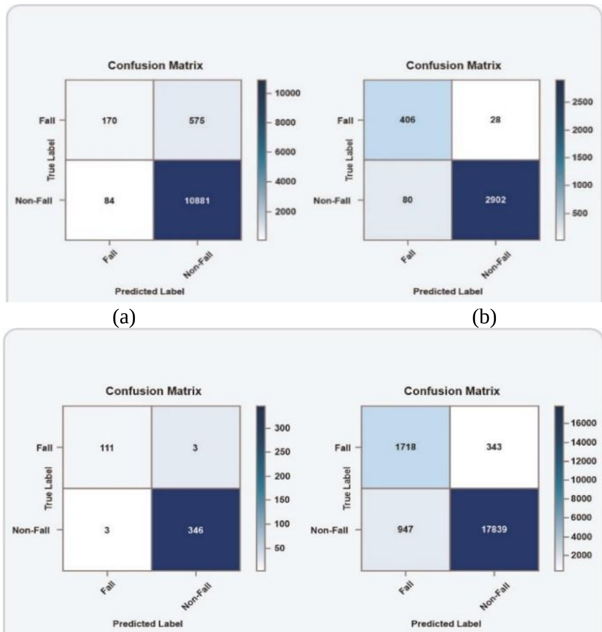


(a)

(b)

(c)

Figure 6. Qualitative analysis on MCFD, SDU, and URFD datasets.



(a) (b)

(c) (d)

Figure 7. Confusion Matrix for different environments. (a) MCFD. (b) SDU (c) URFD. (d) combined FDD, MCFD, SDU & URFD.

TABLE V. COMPARISON OF CNN-BASED FD

Dataset	Proposal	Sensitivity (%)	Specificity (%)	Accuracy (%)
MCFD	S. Wang *et al.* [25]	89.2	90.3	–
	K. Wang *et al.* [26]	93.7	92	–
	Our Work	66	95	–
SDU	Erdem *et al.* [27]	–	–	91.89
	Ma *et al.* [18]	–	–	86.83
	Aslan *et al.* [28]	–	–	88.83
	Our work	83.53	99	96.83
URFD	Zerrouk*i et al.* [24]	–	–	97.02
	Nunez-Marcos *et al.* [12]	100	92	95
	Our work	97.36	99.14	98.7
Combined URFD+MCFD+FDD+and SDU	Our work	64.46	91.21	93.81
Combined URFD+ MCFD+ FDD	Nunez-Marcos *et al.* [12]	94	94	–

Performance analysis builds a robust and generalized FD system by integrating FDD, MCFD, SDU, and UR datasets. The confusion matrix of CNN-based FD tested on combined FDD, MCFD, SDU, and URFD shows that, while the accuracy is reasonable, at 93.81%, the likelihood of false alarms is minimal. It highlights a useful practical quality of the advised method.

To train a CNN model for fall detection, labeled training data is typically used to teach the model to recognize specific visual features that distinguish a fall from a non-fall event or laid-down scene.

For example, a fall may be characterized by rapid movement or changes in posture, as well as a sudden change in acceleration or orientation. The CNN model may learn to recognize these features by analyzing labeled video data of falls and non-falls and identifying the patterns of visual features that are associated with each type of event.

During the training process, the CNN model learns to extract relevant features from the video data and use these features to predict whether a given event is a fall or a non-fall. The accuracy of the model depends on the quality and quantity of the training data, as well as the complexity and effectiveness of the features that the model has learned to recognize. Gunale and Mukherji [2] worked on handcrafted feature extraction-based fall detection method by combining features such as Motion History Image, orientation angle, and aspect ratio and discussed the importance of the MHI feature which helps in differentiating fall events and laid down scenes to correctly recognize the fall event.

Limiting false alarms is crucial in the AD system created utilizing CNN-based FD for recognizing actual abnormalities since false alarms would needlessly activate critical services such as police or medical aid.

VI. CONCLUSION AND RESEARCH IMPLICATIONS

The CNN-based FD system obtained state-of-the-art results in four freely existing FD datasets: FDD, MCFD, SDU, and URFD. In this system to test the generality of the algorithm four datasets were combined and tested for FD. When compared to the state-of-the-art, the suggested system’s accuracy on a variety of datasets for FD was enhanced. CNN learned common features of the problem domain as compared to the hand-engineered feature engineering approach. The proposed system achieved 97.03% accuracy on combined all Le2i environments, which is satisfactory, and it outperforms the existing methods. On another side, the system also achieved 95% specificity on the MCFD dataset, 96.83% accuracy on the SDU dataset, and 98.7% accuracy on the URFD dataset. Also, the system gains an accuracy of 93.81% on the combination of all four FDD, MCFD, SDU, and URFD datasets.

However, there is still research potential for improvement. The authors identify four possible research directions to bring vision-based FD to practical deployments.

- More research on TL using FD dataset is required to enhance general feature extractor.
- Although optical flow images have good representational power for motion, they also require a lot of computer effort during subsequent pre-processing frames, and the performance is impacted by changes in light. Following the end-to-end learning approach, the picture pre-processing stage can be omitted and processing can be done solely with raw images.
- More complicated network architectures need to be built to acquire comprehensive and hierarchical motion representations from raw images.
- As available datasets include only one actor per video, it is assumed that multi-person FD is the next step in FD. Authors believe that region-based CNNs might be a viable research path for this purpose, allowing us to automatically recognize various people in images and analyze those regions using our FD method.
- Quality assessment framework suggested in [29] may be carried out for developing a HFD as an elderly care product. Similarly, a detailed costing analysis while developing this will help in extending this work [30].

VII. LIMITATION OF THE WORK

The proposed approach is novel as it offers improved overall model performance by combining four diverse datasets FDD, MCFD, SDU, and URFD. Our study provides a more comprehensive evaluation of fall detection systems by using a larger and more diverse set of data than previous studies. However, as compared to state-of-the-art [12] the sensitivity value is lower as the proposed system has combined four datasets. Forth dataset used for the model training is MCFD, due to the addition of the MCFD dataset sensitivity values are lower. In this regard, model training and further fine-tuning are needed to improve the sensitivity values of the proposed model.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Kishanprasd G. Gunale is the researcher of this work. He has implemented the algorithms, prepared the dataset, analyzed data, and represented the results with the conclusion. Prachi Mukherji is the supervisor and she validated the results and data. Sumitra N. Motade has collected all required information and written in terms of the research paper. All authors approve the final version.

REFERENCES

[1] G. F. Fuller, "Falls in the elderly," *Am Fam Physician*, vol. 61, no. 7, pp. 2159–2168, 2000.

[2] K. G. Gunale and P. Mukherji, "Indoor human fall detection system based on automatic vision using computer vision and machine learning algorithms," *Journal of Engineering Science and Technology (JESTEC)*, vol. 13, no. 8, pp. 2587–2605, 2018.

[3] E. Alam, A. Sufian, P. Dutta *et al.*, "Vision-based human fall detection systems using deep learning: A review," *Computers in Biology and Medicine*, vol. 146, 2022.

[4] M. Nouredanesh, A. McCormick, S. L. Kukreja *et al*., "Wearable vision detection of environmental fall risk using Gabor Barcodes," presented at the 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob), Singapore, pp. 956–956, 2016.

[5] L. Zhu, R. Wang, Z. Wang *et al.*, "TagCare: Using RFIDs to monitor the status of the elderly living alone," *IEEE Access*, vol. 5, pp. 11364–11373, 2017.

[6] 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.

[7] F. Pengming, Y. Miao, M. N. Syed *et al.*, "Deep learning for posture analysis in fall detection," in *Proc. 19th International Conference on Digital Signal Processing*, 2014, pp. 12–17.

[8] F. Harrou, N. Zerrouki, Y. Sun *et al.*, "Vision-based fall detection system for improving the safety of elderly people," *IEEE Instrumentation and Measurement Magazine*, vol. 20, no. 6, pp. 49–55, 2017.

[9] L. Hyunwoo, K. Jooyoung, Y. Dojun *et al*., "Embedded real-time fall detection using deep learning for elderly care," *Corr, Multiple Datasets*, 2017.

[10] K. Gunale and P. Mukherji, "An intelligent video surveillance system for anomaly detection in home environment using a depth camera," in *Proc Soft Computing: Theories and Applications: of (SoCTA)*, pp. 473-481, 2017.

[11] W. Min, H. Cui, H. Rao *et al*., "Detection of human falls on furniture using scene analysis based on deep learning and activity characteristics," *IEEE Access*, vol. 6, pp. 9324–9335, 2018.

[12] A. Núñez-Marcos, A. Gorka, and A. C. Ignacio, "Vision-based fall detection with convolutional neural networks," *Wireless Communications and Mobile Computing*, vol. 1, 2017.

[13] D. Seth and S. Rastogi, "Application of vendor rationalization strategy for manufacturing cycle time reduction in engineer to order (ETO) environment: A case study," *Journal of Manufacturing Technology Management*, vol. 30, no. 1, pp. 261–290, 2017.

[14] G. Yadav, D. Seth, and T. N. Desai, "Analysis of research trends and constructs in context to lean six sigma frameworks," *Journal of Manufacturing Technology Management*, vol. 28, no. 6, pp. 794–821, 2017.

[15] N. C. Thompson, H. Kristjan, K. L. Greenewald *et al.*, "The computational limits of deep learning," arXiv preprint, arXiv:2007.05558, 2020.

[16] I. Charfi, J. Miteran, J. Dubois *et al.*, "Optimised spatio-temporal descriptors for real-time fall detection: Comparison of SVM and AdaBoost based classification," *J. Electron. Imaging (JEI)*, vol. 22, no. 4, p. 17, 2013.

[17] E. Auvinet, C. Rougier, J. Meunier *et al.*, "Multiple cameras fall dataset," Technical report 1350, DIRO—Université de Montréal, July 2010.

[18] X. Ma, H. Wang, B. Xue *et al*., "Depth-based human fall detection via shape features and improved extreme learning machine," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 6, pp. 1915–1922, 2014.

[19] K. Bogdan and K. Michal, "Human fall detection on an embedded platform using depth maps and wireless accelerometer," *Computer Methods and Programs in Biomedicine*, vol. 117, no. 3, pp. 489–501, 2014.

[20] J. Huang, V. Rathod, C. Sun *et al*., "Speed/accuracy trade-offs for modern convolutional object detectors," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3296–3297.

[21] Transfer learning and fine-tuning. [Online]. Available: https://www.tensorflow.org/tutorials/images/transfer_learning

[22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint, arXiv:1409.1556, 2014

[23] I. Charfi, J. Miteran, J. Dubois *et al*., "Definition and performance evaluation of a robust SVM-based fall detection solution," in *Proc. the Eighth IEEE International Conference on Signal Image Technology and Internet Based Systems*, 2012, pp. 218–224

[24] N. Zerrouki and A. Houacine, "Combined curvelets and hidden Markov models for human fall detection," *Multimedia Tools and Applications*, vol. 77, no. 5, pp. 6405–6424, 2018.

[25] S. Wang, L. Chen, Z. Zhou *et al*., "Human fall detection in surveillance video based on PCANet," *Multimedia Tools and Applications*, vol. 75, no. 19, pp. 11603–11613, 2016.

[26] K. Wang, G. Cao, D. Meng *et al*., "Automatic fall detection of human in video using a combination of features," in *Proc. IEEE International Conference on Bioinformatics and Biomedicine*, 2016, pp. 1228–1233.

[27] E. Akaguunduz, M. Aslan, A. Sengur *et al*., "Silhouette orientation volumes for efficient fall detection in-depth videos," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 756–763, 2017.

[28] M. A. Aslan, Y. Sengur, H. Xiao *et al*., "Shape feature encoding via fisher vector for efficient fall detection in depth-videos," *Applied Soft Computing*, vol. 37, pp. 1023–1028, 2015.

[29] R. P. Mohanty, D. Seth, and S. Mukadam, "Quality dimensions of e-commerce and their implications," *Total Quality Management & Business Excellence*, vol. 18, no. 3, pp. 219–247, 2007.

[30] D. Seth and M. K. Pandey, "A multiple-item inventory model for a non-stationary demand," *Production Planning and Control*, vol. 20, no. 3, pp. 242–253, 2009.