



# An Enhanced Fall Detection Approach in Smart Homes Using Optical Flow and Residual Autoencoder

Faten A. Elshwemy<sup>1\*</sup>, Reda Elbasiony<sup>2</sup>, Mohamed Talaat Saidahmed<sup>3</sup>

<sup>1\*,2,3</sup>Computers and Control Dep., Faculty of Engineering, Tanta University, Egypt

<sup>1\*</sup>fatel\_elshwemy@f-eng.tanta.edu.eg

## ABSTRACT

When focusing on the elderly people, falling is considered a major health problem that can lead to serious injuries; sometimes it can cause the death of them. So, fall detection is an important service for the elderly healthcare to improve the daily life of them and decrease the costs of monitoring these individuals. The fall detection is considered as a problem of detecting anomalies because falling is an abnormal activity. To deal with such issue, we present (OFSRAE) which is unsupervised fall detection framework based on deep learning techniques to detect the elderly people falls. Our proposed framework consists of two stages: the data preprocessing and the deep learning model. We applied the dense optical flow in the first stage to extract the information of the motion and direction of moving objects in the foreground. The deep learning model in the second stage is based on the convolutional long short term memory autoencoder (ConvLSTMAE) network, and the residual connections to extract spatial and temporal features of videos that are captured from thermal and depth cameras. The reconstruction error of an autoencoder is used to identify falls as anomalies recorded in such videos. We experimentally evaluated OFSRAE framework on the publicly available UR and thermal fall datasets which conserve the elderly privacy that is an important issue during monitoring. The experimental results show that the proposed framework detects falls with high performance compared to other deep learning models in the literature.

**Key words:** Anomaly detection, Fall detection, Deep learning, Convolutional autoencoder, Residual network, Optical flow.

## 1. INTRODUCTION

The population of the elderly people is predicted to rapidly grow over the world. Most elderly people need special care because they mostly suffer from various diseases [1]. The presence of internet of things (IoT) and smart homes technologies are used to provide healthcare monitoring for the elderly who are living alone. There are different sensors that can be used to monitor the elderly people without interfering

in their daily life. A major health problem that affects the quality of elderly life is human falls. Falling of the elderly causes serious health problems that can in sometimes cause death of these people. So, fall detection is an important service for the elderly healthcare to enhance the daily life of the elderly people and decrease the elderly people monitoring costs. Various fall detection approaches have been proposed to detect falls. These approaches use different methods to detect falls such as machine learning and deep learning techniques and use different devices to monitor the daily life of the elderly people. These devices can be divided into ambient devices, wearable sensors and vision devices [2]. The ambient devices are used to capture the sound and vibration of the elderly to detect their position and motion. The problem of these devices is low accuracy and causing many false alarms. While the wearable sensors such as accelerometers and gyroscopes are used to collect the movement information of the elderly people [3]. These sensors cause the inconvenience for these people because they have to wear them all the time. Furthermore, they may forget to wear these sensors. In vision devices, cameras are used to capture videos for the elderly people during their daily life. The camera provides more information about the elderly and the environment surrounding them [4]. However, camera violates the privacy of people. To overcome this problem, different types of cameras can be used such as depth and thermal cameras which they conserve the people privacy. Also, these cameras can capture the images in the dark places and during the night. Therefore, in this paper, we focus on videos captured from depth and thermal cameras to detect the elderly people falls. The main challenge is developing a robust fall detection system to avoid false positive and obtain high accuracy for decreasing the post-effects of the falling. In this paper, we introduce Optical Flow Spatiotemporal Residual AutoEncoder (OFSRAE) framework to detect the elderly people falls. This framework is composed of two steps. The first step is data preprocessing, in which the optical flow maps of video frames are extracted. In fall detection approaches, the feature extraction plays very important role to detect different forms of falls with high accuracy. In the second step of (OFSRAE) framework, the deep learning model is used optical flow maps to extract spatial and temporal features to identify falls as anomalies in unsupervised manner. Our deep learning model is trained in unsupervised manner because

collecting fall data is difficult and may put life of people in danger. We use the normal daily living activities (ADL) only to train our model. The model is built using the convolutional autoencoder network, ConvLSTMnetwork, and the residual connections. This model consists of encoder and decoder based on 2D convolutional layers, ConvLSTM and residual connections between encoder and decoder layers. In encoder part, the first convolutional layer is followed by tanh activation function. The rest convolutional layers are followed by batch normalization (BN) and relu activation function. In the final layer, one ConvLSTM layer is used. While the decoder part is built in a symmetrical way of the encoder structure. The residual connections are added between some encoder layers and their corresponding decoder layers to facilitate convergence and improve the result [5]. Training of our model aims to reduce the reconstruction error of the autoencoder between input and output. The OFSRAE framework is trained and tested on the depth dataset that captured from depth camera and the infrared thermal dataset that captured from thermal camera. After the training phase is successfully finished, the proposed framework is tested on both normal daily living activities (ADL) and fall data.

The main contributions of this paper aim to:

- Propose a fall detection framework called (OFSRAE) based on the optical flow and deep learning techniques and train it in unsupervised manner,
- Test the proposed framework on the publicly available depth URFD and thermal fall datasets where depth and thermal cameras conserve the person privacy, and
- Compare the proposed framework with convolutional autoencoder (CAE), denoising autoencoder (DAE), convolutional long short term memory autoencoder (CLSTM-AE) models, and spatiotemporal residual autoencoder (SRAE).

The remainder of this paper is organized as follows: section 2 presents an overview of related work in fall detection. Section 3 introduces an overview of the proposed framework. Section 4 reports the experiments and analyzes the results. Section 5 overviews the conclusion of the paper and the future work.

## 2. RELATED WORK

Various approaches have been proposed for fall detection in smart houses. Authors in [6] propose a human fall detection approach using deep learning technique to detect falls in videos captured by depth cameras. This approach is built using multi class LSTM. Authors in [7] propose a deep learning model to detect human falls called CNN-3B3Conv using accelerometer data. This model is based on a convolutional neural network. It consists of three blocks. The first and second blocks structure is the same which consists of three convolutional layers with different kernel sizes and one maxpooling layer. While the last block is composed of three fully connected layers. In [8], an enhanced threshold based

fall detection mechanism is proposed. It uses the smartphone accelerometers to collect data for identifying the daily activities and falling action. In this mechanism, three values of threshold are used to differentiate between different actions. A two-step fall detection approach is proposed in [9] based on smartphone sensors to detect falls. It consists of two steps: the first one is performing the multi-class classification and the second step is converting the multi-class classification to the binary classification. In [10], the ConvLSTM-AE model is used to detect falls from videos captured from thermal cameras. This model is built using two parts spatial and temporal autoencoder. The spatial encoder consists of two convolutional layers. While, the temporal encoder uses one ConvLSTM layer. In [11], the human fall detection method is proposed using convolutional autoencoder. It uses the signals that are collected by floor acoustic sensors such as voices, music, footsteps for training the convolutional autoencoder. When the different sound occurs which is not in the training dataset, it is recognized as fall. In [12], the unsupervised fall detection approach using residual autoencoder called SRAE is proposed. It uses the autoencoder model based on the convolutional layers, ConvLSTM layers and residual connections. The residual connections are used between the encoder and decoder layers. For training the autoencoder, the videos captured from thermal cameras are used. It trained in unsupervised manner by using only the normal ADL. The fall detection system based on multiple kernels learning called FallDroid is proposed in [13]. This system is composed of two steps: threshold based method and multiple kernels learning support vector machine (MKL-SVM). The threshold based method (TBM) is used first to discard most of ADLs according to threshold value. In the second step, MKL support vector machine is used to detect falls. In [14], the authors used CNN-AlexNet model for classification the normal and anomaly events. They focus on the daily human activities and fall actions.

## 3. METHODOLOGY

In this section, we present our proposed framework (OFSRAE) which used for fall detection problem. The proposed framework (OFSRAE) as shown in Figure 1 consists of two steps. The data preprocessing is the first step of our framework. In this step, the dense optical flow algorithm is applied to video frames of the datasets to extract the dense optical map of these frames. While in the second step, the residual convolutional LSTM autoencoder is trained on the dense optical maps to extract the spatial and temporal features.

### 3.1 Data Preprocessing

The data preprocessing is the first step in the proposed framework (OFSRAE) which aims to produce the optical flow maps of input video frames. There are two methods of optical flow: dense optical flow and sparse optical flow. In Dense optical flow, the flow vectors are created for all pixels of the frame. While the sparse optical flow creates the flow vectors

of some important pixels of the entire frame.

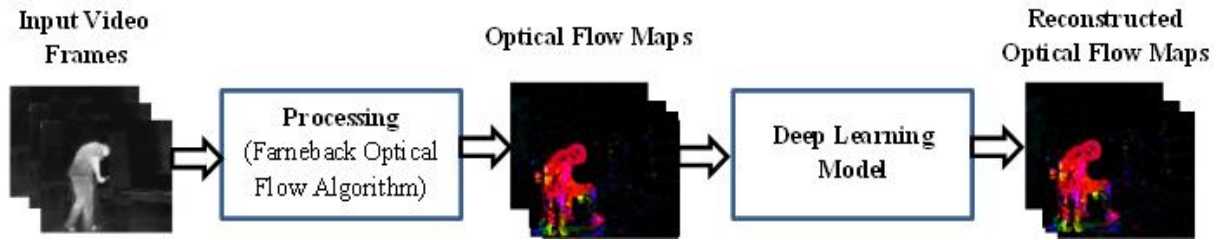


Figure 1: Proposed framework (OFSRAE)

In this paper, we use dense optical flow because the accuracy of the dense optical flow is higher than the sparse optical flow. We use farneback optical flow algorithm to extract dense optical flow maps of each frame. The farneback optical flow algorithm is proposed in [15] which use polynomial expansion transform to approximate each neighborhood of each pixel of frame with polynomial. It creates the dense optical flow objects by estimating the motion and direction of moving objects between two consecutive frames. The quadratic polynomials in the farneback optical flow algorithm represented in the local coordinate system [15] as in equation (1).

$$f(x) \sim x^T A x + b^T x + c \quad (1)$$

where  $A$  denotes to a symmetric matrix,  $b$  denotes to a vector, and  $c$  stands for a scalar. The coefficients are computed from a weighted least squares fit to the signal values in the neighborhood.

The first frame polynomial can be represented as in equation (2) which is obtained at time  $t$ .

$$f_1(x) \sim x^T A_1 x + b_1^T x + c_1 \quad (2)$$

The consecutive frame polynomial can be computed from the first frame polynomial with the displacement  $d$  at time  $(t + dt)$  as shown in equations (3), (4), and (5) by substituting  $x$  by  $(x-d)$  in equation (2).

$$f_2(x) = f_1(x - d) \quad (3)$$

$$f_2(x) = x^T A_1 x + (b_1 - 2A_1 d)^T x + d^T A_1 d - b_1^T d + c_1 \quad (4)$$

$$f_2(x) = x^T A_2 x + b_2^T x + c_2 \quad (5)$$

By equating the coefficients of equations (4) and (5), the equations (6), (7), and (8) are obtained as the following:

$$A_2 = A_1 \quad (6)$$

$$b_2 = (b_1 - 2A_1 d) \quad (7)$$

$$c_2 = (d^T A_1 d - b_1^T d + c_1) \quad (8)$$

The displacement vector can be obtained from equation (7) as shown in equation (9).

$$d = 0.5 A_1^{-1} (b_1 - b_2) \quad (9)$$

At the end, we obtain the dense optical flow maps of the consecutive frames of the thermal and depth videos which are used as inputs to the next step of our proposed framework (OF-SRAE).

### 3.2 Deep Learning Model

The deep learning model is the second step of the proposed framework after the data preprocessing step. The deep learning model which is used in this paper is proposed in [12] as shown in Figure 2. This model is based on the convolutional autoencoder, ConvLSTM and the residual connections to detect falls as anomalies. This model is composed of two networks: the encoder network and the decoder network. As shown in Figure 2, the encoder network consists of four convolutional layers and one ConvLSTM layer. The purpose of 2D convolutional layers is to extract the spatial features of the frames. The ConvLSTM layer is used to extract spatial and temporal correlations. The batch normalization layer (BN) and rectified linear unit (ReLU) activation function are used after each convolutional layer. Except, the first layer is followed by only TANH activation function. The benefits of the batch normalization layer (BN) are making training process faster and decreasing the overfitting. While the ReLU activation function is used to avoid the exploding and vanishing gradient problems. The stride is applied to each convolutional layer to perform the down sampling. The structure of the decoder network is built symmetrically of the structure of the encoder. It consists of one ConvLSTM and four transposed convolutional layers. The stride is applied to each transposed convolutional layer to perform the up sampling to increase the resolution of the output. To facilitate convergence the residual connection is used between encoder layers and decoder layers. For training the model, the dense optical flow maps of eight contiguous video frames that are obtained in the preprocessing step are used as input to the encoder. We trained the deep learning model in unsupervised manner by using the only normal activities of daily living (ADL) of datasets. The encoder encodes the dense optical flow maps and the decoder produces the reconstructed optical flow maps.

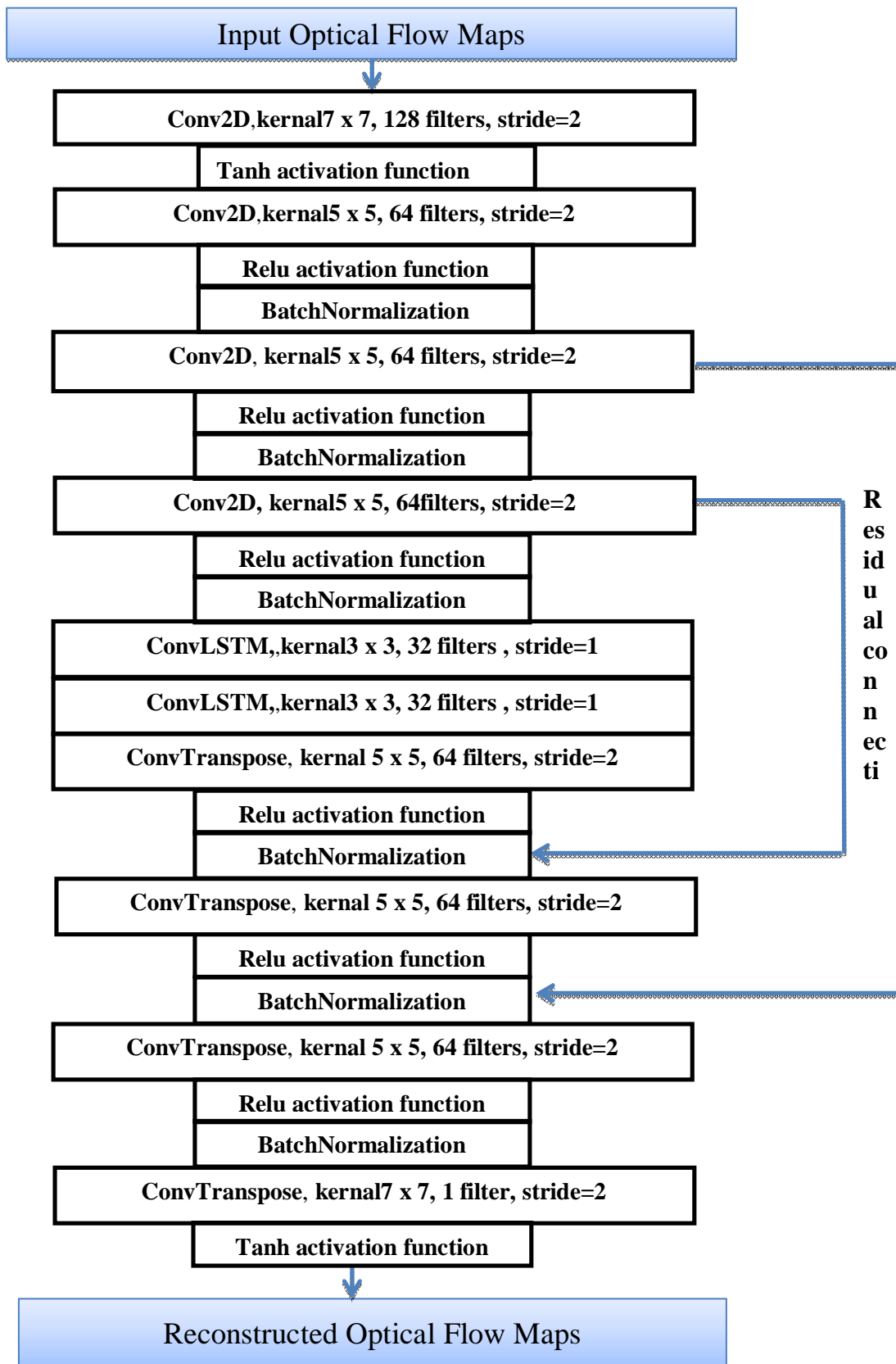


Figure 2: The Deep learning model of the proposed framework

The mean squared error loss is used to compute the reconstruction error between the input optical flow maps and reconstructed optical flow maps as shown in Eq. (10). The objective function is to reduce the reconstruction error.

$$MSE_i = \frac{1}{n} \sum_{i=1}^n (I_i - O_i)^2 \quad (10)$$

where  $MSE$  is a reconstruction error,  $n$  is the number of training samples,  $I$  is the input dense optical flow maps of eight contiguous video frames and  $O$  is the reconstructed optical flow maps.

In the testing phase, the model is tested on both the normal activities and falls. To detect falls, we compute the reconstruction error for the dense optical flow maps. When the reconstruction error is low, this indicates the normal activity and when the reconstruction error per frame is high, this indicates the occurrence of a fall.

## 4. EXPERIMENTAL METHODOLOGY AND RESULTS

### 4.1 Dataset

To test the proposed (OFSRAE) framework, we use the following two datasets to detect falls. These datasets contain videos captured from depth and thermal cameras.

#### A. Thermal Dataset

The thermal fall dataset [16] consists of normal activities of daily living (ADL) and fall videos. A FLIR ONE thermal camera is used to capture these videos with a spatial resolution of  $640 \times 480$ . In a room with a single view this thermal camera is mounted on an Android phone. This dataset has 44 videos which divided into 9 ADL videos and 35 videos contain both falls and normal ADL. The falls in this dataset have occurred in different positions such as falling from sitting, falling from chair, and falling from standing. Figure 3 shows samples of frames containing normal activities of daily living (ADL) and fall.

#### B. UR Fall Detection Dataset

The URFD dataset [17] contains 40 videos with only normal ADL and 30 videos with falls and also normal activities. Two Microsoft Kinect cameras are used to capture these videos with resolution ( $640 \times 480$ ). One of the cameras is sitting parallel to the floor and other is mounted on the ceiling. Falls were performed from standing and sitting on the chair. Figure 4 shows samples of normal ADL and falls from URFD dataset.

### 4.2 Experimental Setup and Results

To evaluate the performance of the proposed OFSRAE framework to detect falls, we compare it with Spatio-temporal Residual AutoEncoder (SRAE), convolutional long short term memory autoencoder (CLSTMAE), convolutional autoencoder (CAE), and denoising autoencoder (DAE) models. Our experiments are executed on a processor core i7 (4GHz, 8

M cache) with GPU TITAN X 12 GB and 16 GB RAM running 64-bit Windows 10. All models and training procedures are implemented in Keras with Tensorflow backend. The configurations of the encoding and decoding phases for the DAE, CAE, CLSTMAE, and SRAE models are illustrated in Table 1 [12]. We use the same results for the compared models as in [12]. The proposed model is trained on the two datasets: thermal fall dataset and URFD dataset for 30 epochs. While the DAE and CAE models are trained on Thermal dataset for 200 epochs and CLSTMAE model is trained for 50 epochs. The DAE, CAE, and CLSTMAE models are trained on URFD dataset for 500 epochs. The SRAE model is trained on URFD dataset for 200 epochs and on thermal fall dataset for 30 epochs. The batch size was set as 32 for all models. We use the Adadelta optimizer in training phase for all models. The frames of both datasets are resized to  $64 \times 64$ . All models are trained in unsupervised manner.

Only normal ADL videos are used for training phase. While, in the testing phase, fall videos that contains of both normal and fall frames are used. A reconstruction error is calculated that can be used as anomaly score to identify a fall frame as an anomaly. We train all models on the URFD and thermal fall datasets for 100 epochs to show the convergence rate for each model. The convergence curves for all models are shown in Figure 5 and Figure 6. As shown in these figures, the proposed OFSRAE framework has fastest convergence rate than other compared models. To compare the performance of the models, we used ROC AUC (Receiver Operating Characteristic area under curve) and PR AUC (Precision Recall area under curve). The ROC AUC and PR AUC are the area under the ROC curve and the PR curve respectively. A ROC curve is a graphical plot between true positive rate (TPR) as shown in Eq. (11), and false positive rate (FPR) as shown in Eq. (12). The model is better when the ROC AUC is high. While, a PR curve is a graphical plot between precision rate as shown in Eq. (13) and recall rate as shown in Eq. (14). The model is better when the PR AUC is high.

$$TPR = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Negatives})} \quad (11)$$

$$FPR = \frac{\text{False Positives}}{(\text{False Positives} + \text{True Negatives})} \quad (12)$$

$$\text{Precision} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Positives})} \quad (13)$$

$$\text{Recall} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Negatives})} \quad (14)$$

Table 2 shows The ROC AUC and PR AUC results for all models: DAE, CAE, CLSTMAE, SRAE and the proposed OFSRAE on the thermal dataset and URFD dataset. The ROC AUC and PR AUC results are the average of AUC across all test videos, with standard deviation in brackets.



**Figure 3:** Thermal dataset – ADL and fall frames.



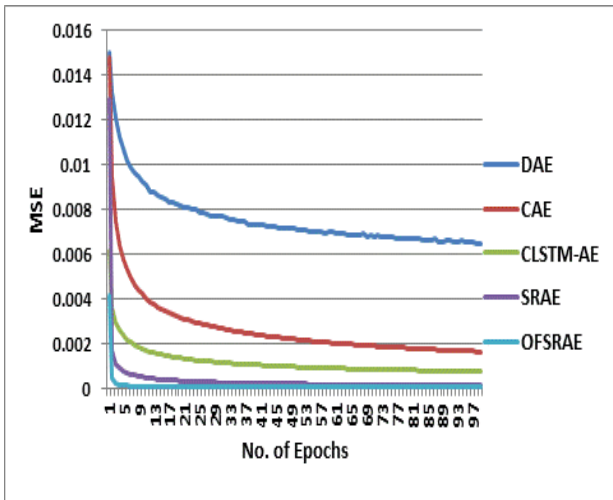
**Figure 4:** URFD dataset – ADL and fall frames.

**Table 1:** Configurations of DAE, CAE, CLSTMAE, and SRAE models

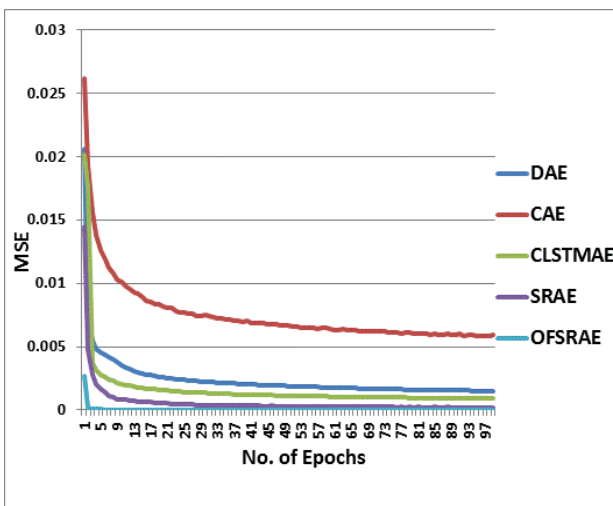
DAE	CAE	CLSTMAE	SRAE
Dense (4096)	2D Convolution (kernel=3 x 3, filters=16)	2D Convolution (kernel=11 x 11, filters=128, stride=4)	2D Convolution( kernel =7 x 7, filters=128 , stride=2)
Dense (1500)	2D Max_pooling (2,2)	2D Convolution ( kernel =5 x 5, filters =64 , stride =2)	2D Convolution (kernel=5 x 5, filters =64 , stride =2)
Dense (1000)	2D Convolution( kernel=3 x 3, filters =8)	2D ConvLSTM (kernel =3 x 3, filters =64)	2D ConvLSTM( kernel 3 x 3, filters=32)
Dense (500)	2D Max_pooling (2, 2)	2D ConvLSTM (kernel =3 x 3, filters =32 )	2D ConvLSTM (kernel 3 x 3, filters=32)
Dense (500)	2D Convolution( kernel =3 x 3, filters =8)	2D Convolution Transpose( kernel=5 x 5, filters =128 , stride=2)	2D Convolution Transpose (kernel =5 x 5, filters=64 , stride=2)
Dense (1000)	2D Max_pooling (2, 2)	2D Convolution Transpose( kernel =5 x 5, filters =64 , stride=2)	2D Convolution Transpose (kernel =5 x 5, filters=64 , stride=2)
Dense (1500)	2D Convolution( kernel =3 x 3, filters =8)	2D Convolution Transpose( kernel =11 x 11, filter=1 , stride =4)	2D Convolution Transpose( kernel =5 x 5, filters=64 , stride =2)
Dense (4096)	2D Max_pooling (2,2)		2D Convolution Transpose( kernel=7 x 7, filter=1 , stride =2)
	2D Convolution Transpose( kernel =3 x 3, filters =8)		(With residual connection between convolution and convolution transpose layers)
	2D Up_sampling (2, 2)		
	2D Convolution Transpose ( kernel =3 x 3, filters =8 )		
	2D Up_sampling (2, 2)		
	2D Convolution Transpose (kernel =3 x 3, filters =16)		
	2D Up_sampling (2, 2)		
	2D Convolution Transpose (kernel =3 x 3, filter =1 )		



Our framework OFSRAE holds the highest AUC values. For thermal dataset, its improvements of ROC AUC and PR AUC are 33%, and 76% than DAE model, 22 %, and 74% than CAE model, 14 % and 24% than CLSTMAE model, 1% and 4% than SRAE, respectively. For URFD dataset, its improvements of ROC AUC and PR AUC are 51%, and 35% than DAE model, 53 %, and 41% than CAE model, 29 % and 24% than CLSTMAE model, 5% and 6% than SRAE, respectively. We noted that, CLSTMAE model performs better than DAE and CAE models because CLSTMAE model can extract the spatial and temporal features of video images. While SRAE model performs better than CLSTMAE model. This is because; besides SRAE extracts both the spatial and temporal features of video images, it also uses residual connections that facilitate convergence and improve performance. Interesting to see also that the proposed framework performs better than other compared models. This is because; our framework uses dense optical flow algorithm that extracts the motion and direction information of foreground objects.



**Figure 5:** Convergence curves for DAE, CAE, CLSTMAE, SRAE, and proposed OFSRAE on thermal fall dataset with 100 epochs.



**Figure 6:** Convergence curves for DAE, CAE, CLSTMAE, SRAE, and proposed OFSRAE on URFD dataset with 100 epochs.

**Table 2:** AUC values for different models for Thermal and URFD datasets

Models	Thermal dataset		URFD dataset	
	ROC AUC	PR AUC	ROC AUC	PR AUC
DAE	0.64 (0.15)	0.17 (0.15)	0.42 (0.3)	0.63 (0.24)
CAE	0.75 (0.16)	0.19 (0.18)	0.40 (0.33)	0.57 (0.29)
CLSTMAE	0.83 (0.11)	0.69 (0.15)	0.64 (0.15)	0.74 (0.23)
SRAE	0.97 (0.08)	0.93 (0.13)	0.87 (0.1)	0.92 (0.1)
<b>OFSRAE</b>	<b>0.98 (0.02)</b>	<b>0.97 (0.04)</b>	<b>0.92 (0.1)</b>	<b>0.98 (0.02)</b>

### 5. CONCLUSION AND FUTURE WORK

In this paper, a fall detection framework called OFSRAE is proposed to detect the elderly people falls as anomalies. The proposed framework is built using two stages: data preprocessing and deep learning model. The optical flow method is used in data preprocessing stage to extract the speed and direction features of moving objects in the videos of the daily activities of the elderly people. While in the next stage, the deep learning model based on the convolutional autoencoder, ConvLSTM and residual connections is used to extract the spatial and temporal features of the dense optical flow maps. This model is trained in unsupervised manner by using only the normal daily living activities (ADL). While in the testing phase, the proposed model is tested on videos contain both normal ADL and fall actions. To test the proposed framework performance, we used the infrared thermal fall dataset and URFD dataset. We compared the proposed framework with DAE, CAE, CLSTMAE, and SRAE deep learning approaches. The comparisons are made in terms of ROC AUC and PR AUC. From the comparison results, the proposed framework has high performance to detect falls than other compared models. For the thermal fall dataset, the proposed OFSRAE framework achieves the highest (ROC AUC 98 % and PR AUC 97 %) values than SRAE (ROC AUC 97 % and PR AUC 93 %), CLSTMAE (ROC AUC 83% and PR AUC 69%), CAE (ROC AUC 75 % and PR AUC 19 %), and DAE (ROC AUC 64 % and PR AUC 17 %) models. For the UR fall detection dataset, the proposed OFSRAE framework achieves the highest ROC AUC (92 %) and PR AUC (98 %) values than SRAE (ROC AUC 87 % and PR AUC 92 %), CLSTMAE (ROC AUC 64% and PR AUC 74%), CAE (ROC AUC 40 % and PR AUC 57 %), and DAE (ROC AUC 42 % and PR AUC 63 %) models. For future

work, we plan to apply the proposed framework to different anomaly detection applications.

## REFERENCES

1. M. M. Islam, *etc.* **A Review on Fall Detection Systems Using Data from Smartphone Sensors**, *Ingénierie des Systèmes d'Information*, vol. 24, no. 6, pp. 569-576, 2019. <https://doi.org/10.18280/isi.240602>
2. A. Witte and R. Zarnekow. **Transforming Personal Healthcare through Technology - A Systematic Literature Review of Wearable Sensors for Medical Application**, in *Proc. of the 52nd Hawaii International Conference on System Sciences*, pp. 3848-3857, 2019. <https://doi.org/10.24251/HICSS.2019.466>
3. A. Geetha, M. Aravindan, J. Anand Jahan, and S. Aravindh. **CNN Based Fall Detection and Health Monitoring System Using IoT**, *International Journal of Advanced Science and Technology*, vol. 29, no. 3, pp. 8820 – 8831, 2020.
4. G. L. Xiong, E. Bayen, S. Nickels, R. Subramaniam, P. Agrawal, *etc.* **Real-Time Video Detection of Falls in Dementia Care Facility and Reduced Emergency Care**, *The American Journal of Managed Care*, vol. 25, no. 7, pp. 314-315, 2019.
5. J. R. Rosario. **Development of a Face Recognition System Using Deep Convolutional Neural Network in a Multi-view Vision Environment**, *International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE)*, vol. 8, no. 3, pp. 369-374, 2019. <https://doi.org/10.30534/ijatcse/2019/06832019>
6. A. Shojaei-Hashemi, P. Nasiopoulos, and J. J. Little. **Video-based Human Fall Detection in Smart Homes Using Deep Learning**, *International Symposium on Circuits & Systems*, IEEE, 2018
7. G. Santos, P. Endo, K. Monteiro, E. Rocha., I. Silva, and T. Lynn. **Accelerometer-Based Human Fall Detection Using Convolutional Neural Networks**, *Sensors (Basel)*, vol. 19, no. 7, pp. 1-12, 2019.
8. J. Lee, and H. Tseng. **Development of An Enhanced Threshold-based Fall Detection System Using Smartphones with Built-in Accelerometers**, *IEEE Sensors Journal*, vol. 19, no. 18, pp. 8293-8302, 2019. <https://doi.org/10.1109/JSEN.2019.2918690>
9. J. C. Dogan, and M. S. Hossain. **A Novel Two-Step Fall Detection Method Using Smartphone Sensors**, in *IEEE International Conference on Smart Computing (SMARTCOMP)*, pp. 434-438, 2019.
10. J. Nogas, S. Khan, and, A. Mihailidis. **Fall Detection from Thermal Camera Using Convolutional LSTM Autoencoder**, In *Proc. of the 2nd workshop on Aging, Rehabilitation & Independent Assisted Living, IJCAI Workshop, Sweden*, pp. 1-4, 2018.
11. 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 book: *Quantifying and Proc. Biomedical and Behavioral Signals*, pp. 185-196, 2019. [https://doi.org/10.1007/978-3-319-95095-2\\_18](https://doi.org/10.1007/978-3-319-95095-2_18)
12. F. A. Elshwemy, R. Elbasiony, and M. T. Saidahmed. **A New Approach for Thermal Vision based Fall Detection Using Residual Autoencoder**, *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 2, pp. 250-258, 2020. <https://doi.org/10.22266/ijies2020.0430.24>
13. A. Shahzad, and K. Kim. **FallDroid: An Automated Smart Phone based Fall Detection System using Multiple Kernel Learning**, *IEEE Transition on Industrial Informatics*, vol. 15, no. 1, pp. 35- 44, 2019. <https://doi.org/10.1109/TII.2018.2839749>
14. M. F. Abu Hassan, A. Hussain, M. H. Saad, and Y. Yusof. **Convolution Neural Network-based Action Recognition for Fall Event Detection**, *International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE)*, vol. 8, no. 1.6, pp. 466-470, 2019. <https://doi.org/10.30534/ijatcse/2019/6881.62019>
15. G. Farneböck. **Two-frame Motion Estimation based on Polynomial Expansion**, in *13th Scandinavian Conference on Image Analysis (SCIA 2003)*, Berlin, Germany: Springer, vol. 2749, pp. 363-370, 2003. [https://doi.org/10.1007/3-540-45103-X\\_50](https://doi.org/10.1007/3-540-45103-X_50)
16. V. S. Ganesan, O. R. Murthy, and A. Dhall. **Thermal Imaging based Elderly Fall Detection**, in *Proc. of ACCV Workshop*, Springer, pp. 541–553, 2016.
17. UR Fall Detection Dataset is available in: <http://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html>.