The First International Conference on
**Machine Learning and Knowledge Discovery (MLKD 2024)**
Amirkabir University of Technology, December 18-19, 2024

# Anomaly Detection using ConvLSTM Autoencoder in Smart Home Environments

Zahra Atashgahi* Mohammadreza Jafaei† Ehsan Nazerfard‡ Alireza Nadali§

## Abstract

As the population ages, there is a growing concern for the safety and well-being of elderly individuals living independently. However, with the emergence of Ambient Intelligence, independent living for the elderly is no longer an impossible feat. Smart homes equipped with advanced technology offer a cost-effective solution to this problem. In this paper, we propose a novel approach to address the challenges of anomaly detection in the daily routines and behaviors of elderly individuals. We introduce a ConvLSTM Autoencoder model for processing spatiotemporal data, which is well-suited for identifying rare and sparse anomalies that are difficult to reproduce in normal behavior. To validate our proposed method, we utilize two datasets from the WSU CASAS smart home project and compare it with other state-of-the-art approaches. Our results demonstrate the effectiveness of our model in accurately detecting anomalies in the behavior of elderly individuals living in smart homes, paving the way for improved safety and quality of life for this vulnerable population.

**Keywords:** Machine Learning, Smart Homes, Anomaly Detection, ConvLSTM, Autoencoder

## 1 Introduction

By 2030, it is projected that more than 3% of the population will be aged between 65 and 74, with dementia being a common cognitive impairment in the elderly. With no known cure or treatment to slow its progression, detecting the early symptoms of dementia is crucial for improving the quality of life for those affected. Early-stage dementia can make everyday activities and tasks challenging, and individuals with Alzheimer's disease, which accounts for over 80% of dementia cases worldwide, often exhibit behavior anomalies due to short-term memory loss.

Anomaly or outlier detection is a vital tool for identifying observations that deviate significantly from the majority of a dataset. In the case of detecting dementia, the behavior of individuals with dementia often has characteristic patterns that differ significantly from those without. However, acquiring labeled datasets for supervised learning is often not feasible, making unsupervised learning methods such as Deep Anomaly Detection techniques a powerful alternative.

In recent years, Deep Learning has revolutionized Artificial Intelligence and has led to significant advances in anomaly detection. One of the most successful Deep Anomaly Detection methods is One-Class Support Vector Machine (OCSVM), which combines uncertainty and importance information to detect anomalous data.

In this study, we propose a novel method for detecting early symptoms of dementia in the elderly using Deep Learning techniques. Specifically, we introduce a Convolutional Autoencoder and Convolutional LSTM Autoencoder for spatiotemporal data processing, which can detect subtle anomalies in daily routines and behaviors. We compare our results with OCSVM and demonstrate significant improvements, especially in the context of Ambient Intelligence.

In summary, our work provides a valuable contribution to the field of anomaly detection and early detection of dementia in the elderly. The remainder of the paper is organized as follows: in section 2, we review related work on anomaly detection in various fields. Section 3 provides an overview of the proposed method, while section 4 presents the architecture in detail. Experimental results are presented in section 5, followed by concluding remarks in section 6.

## 2 Related work

Anomaly Detection is particularly useful in industry, since labeled data is unavailable and the exact model of a system is not accurate due to aging and small differences in manufacturing. In financial domain, Anomaly detection is mainly used for fraud detection in different areas such as Healthcare, Financial Services and manufacturing.

Anomaly detection is still at early stages in smart homes and much more developed in other areas as we mentioned above, yet it has been the center of many

*Department of Computer Engineering, Amirkabir University of Technology, zahraatashgahy@aut.ac.ir
†Department of Computer Engineering, Amirkabir University of Technology, mr.jafaei@aut.ac.ir
‡**Corresponding author**: Department of Computer Engineering, Amirkabir University of Technology, nazerfard@aut.ac.ir
§Department of Computer Engineering, Amirkabir University of Technology, a_nadali@aut.ac.ir

studies. Shin et al. [24] showed promising results using Support Vector Data Description in detecting anomalies in living patterns for elderly people who live alone. Also there are studies which utilized different machine learning methods such as Naïve Bayes [6], Restricted Boltzmann Machines (RBMs) [12] and Markov logic networks [13]. In recent years, there has been a growing interest in Deep Learning techniques and methods, mostly due to the abundance of data and how well these methods perform when there is sufficient data to train, however, there is a lack of deep learning based anomaly detection in smart homes.

Recently, researchers all around the world have shown great interest in Convolutional Neural Networks (CNNs) [27, 28, 20] and Recurrent Neural Networks (RNNs) [18] in human activity recognition. Some of these studies take advantage of the inherent properties of CNNs and Long Short Term Memories (LSTMs) in activity recognition [20]. Moreover, deep convolutional and recurrent architectures have been used in order to analyze and process movement data with wearable sensors [18]. Although CNNs have achieved great success in activity recognition, there are not many studies that focus on activity detection in smart homes. Arifoglu et al. [5] employed these architectures in a supervised manner in activity and abnormal behavior detection. Ahmadet al. [1] introduces a new deep learning framework named ConvLSTM, which is based on recurrent neural networks (RNNs). The proposed architecture provides a more comprehensive and detailed forecast of NDVI compared to the existing methods. In [16], the ConvLSTM combined learning algorithm has been used for flood prediction, which has provided a much higher detection power than its previous modes. In this work, we employ Autoencoder together with ConvLSTM cells in order to perform anomaly detection the Aruba and Kyoto daily activity datasets, both from WSU CASAS smart home project [25].

The authors in [8] propose an load monitoring method that aims to develop an activity recognition system based on IoT architecture. They have employed three different classifier models are tested using real data: feed-forward neural network, LSTM, and support vector machines. The developed activities of daily living (ADLs) algorithm maps each ADL to a set of criteria depending on the appliance used. A sensitivity analysis is also carried out to study the impact of the group size on the classifier accuracy. Furthermore, the researchers in [7] propose a smart home monitoring system for living-alone senior citizens, relying on carefully designed, low-cost infrared sensor devices, as well as a cloud-based data processing and anomaly detection platform. For privacy preservation, they encrypt collected data and store data indices in a blockchain system, to achieve efficient data access control and auditing. For motion anomaly detection, the authors propose a simple but effective environment adaptation method to work with the one-class support vector machine method.

Also, Gulati and Kuar in [11] study the significance of deploying socially enabled IoT systems in Ambient Assisted Living (AAL) environment by proposing a robust social IoT based AAL system for elderly people. The proposed system is capable of providing assistance to the elderly staying in smart home environment. In case of emergency, the system automatically generates alerts intimating about the situation to the concerned entities. They have utilized two machine learning models; Naive Bayes (NB) and Random Forest (RF) to analyze the data in order to predict the well being of the elderly person. Moreover, a research group in [9] propose an approach to recognize the activities performed in a smart home. The proposed method separates the normal from the anomalous activities. Also, they identify the anomalous days based on the number of activities performed in a day. They perform activity recognition by applying probabilistic neural network on the pre-segmented activity-data obtained from the sensors deployed at different locations in a smart home. They use an autoencoder to identify the anomalous from the normal instances of activities. We further categorize the anomalies based on the criteria such as missing or extra sub-events, and unusual duration of activity. Since the ground truth of the anomalies is unavailable, they generate the ground truth using the boxplots of the duration, and the number of sub-events in an activity. A comprehensive evaluation of the proposed approach on two publicly available CASAS smart home datasets demonstrates its ability in the activity recognition and the correct identification of anomalies.

Wang et al. in [26] present a comprehensive and in-depth review, focusing on the techniques that profile ADLs and detect abnormal behaviour for healthcare. In particular, they discuss the definitions and examples of abnormal behaviour in the healthcare of elderly people. Another group in [19] design a location-based tracking system for a four-story nursing home. The main challenge is to identify the group activity among the nursing home's residents and to detect if they have any deviated activity behavior. We propose a location-based deviated activity behavior detection system to detect deviated activity behavior by leveraging data fusion technique. In order to compute the features for data fusion, an adaptive method is applied for extracting the group and individual activity time and generate daily hybrid norm for each of the residents. Next, deviated activity behavior detection is executed by considering the difference between daily norm patterns and daily input data for each resident. Lastly, the deviated activity behavior among the residents are classified using a rule-based classification approach. Also, the researchers in

[22] propose an explainable LSTM-based framework to classify the activities of daily life and detect anomalies within a fog-enhanced smart home. Data from sensors in a smart home are forwarded to fog nodes where the classification and anomaly detection tasks are carried out. They have evaluated the proposed approach on a standard dataset to demonstrate its application and feasibility in real-world applications.

Last, Alaghbari et al. in [2] propose a unified deep learning model for monitoring elderly in execution of ADLs. The proposed approach consists of three stages which are activity recognition, anomaly detection and next activity prediction. Such a system can provide useful information for the elderly, caregivers and medical teams to identify activities and generate preventive and corrective measures. The performance of the proposed unified approach has been evaluated on real smart home datasets to demonstrate its ability to recognize activities, detect anomalies and predict the next activity.

## 3 Preliminary results

In this section, we provide some preliminaries to make it easier to follow the description of our proposed model.

### 3.1 Recurrent Neural Networks

Recurrent Neural Network (RNN) is a type of artificial neural networks that is used to process sequential data, which allows modeling temporal dynamic behavior. In Artificial Neural Networks, one may presume that data instances are independent but in activity detection, this presumption does not hold true. RNNs have internal states which makes it possible to process sequential data. They are typically as follows (see Fig. 1):

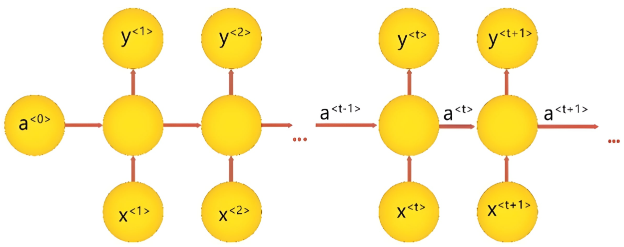



Figure 1: Schematic representation of a traditional RNN

For each timestep $t$, the activation $a^{<t>}$ and the output $y^{<t>}$ are expressed in Eq. 1:

$$\begin{aligned} a^{<t>} &= g_1(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a) \\ y^{<t>} &= g_2(W_{ya}a^{<t>} + b_y), \end{aligned} \tag{1}$$

where $W_{aa}$,$W_{ax}$,$W_{ya}$,$b_a$,$b_y$ denote coefficients that are shared temporally and $g_1$,$g_2$ are activation functions.

### 3.2 ConvLSTM

Nowadays it is quite common to find spatiotemporal data, like videos, satellite pictures or security cameras. The major downside of using LSTM for processing spatiotemporal data is redundancy and lack of leveraging the spatial information. ConvLSTM [23] is a recurrent layer, but the inputs, cell outputs and hidden states and gates are 3D tensors, which is depicted in Fig. 2.

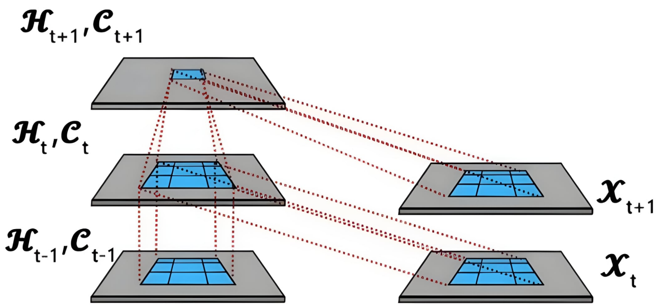


Figure 2: Schematic representation of a traditional RNN

Equations governing the ConvLSTM cells are shown in Eq. 2, where '$*$' denotes the convolution operator and '$\circ$' element-wise product.

$$\begin{aligned} i_t &= \sigma\left(W_{xi} * \chi_t + W_{hi} * \mathcal{H}_{t-1} + W_{ci} \circ \mathcal{C}_{t-1} + b_i\right) \\ f_t &= \sigma\left(W_{xf} * \chi_t + W_{hf} * \mathcal{H}_{t-1} + W_{cf} \circ \mathcal{C}_{t-1} + b_f\right) \\ \mathcal{C}_t &= f_t \circ \mathcal{C}_{t-1} + i_t \circ \tanh\left(W_{xc} * \chi_t + W_{hc} * \mathcal{H}_{t-1} + b_c\right) \\ o_t &= \sigma\left(W_{xo} * \chi_t + W_{ho} * \mathcal{H}_{t-1} + W_{co} \circ \mathcal{C}_{t-1} + b_o\right) \\ \mathcal{H}_t &= o_t^{\circ} \tanh\left(\mathcal{C}_t\right) \end{aligned} \tag{2}$$

### 3.3 Autoencoders

An autoencoder (AE) is a neural network that is trained to reproduce the input as its output [10]. Autoencoders have a hidden layer $h$ that tried to represent input in another domain. Autoencoders consist of two parts: an encoder function $h = f(x)$ and a decoder that reconstructs $r = g(h)$. Encoder finds a new domain to map the input data, which generally have a smaller dimension than $x$.

Reproducing inputs as output may seem futile, but traditionally, AEs have been used for dimensionality reduction. That being said, AEs can be utilized for Anomaly detection. We presume that Encoder learned an accurate mapping function for inputs and decoder reproduces the input, so instances with high reconstruction error are believed to be anomalies, since they cannot be mapped and then reproduced like normal instances.

### 3.4 Data sets

This study utilized two smart home datasets, namely the Aruba dataset [25] and the Kyoto dataset [20] from the WSU CASAS testbeds[1], to evaluate the proposed method. The Aruba dataset was collected from an off-campus smart apartment and includes sensor data from a volunteer adult's home, who regularly received visits from their children and grandchildren. The apartment is equipped with 35 binary sensors (4 doors/cabinet and 31 motion sensors), and the dataset includes a list of (sensor, timestamp) sensor readings, as well as 11 daily activities that occurred over a period of more than 7 months. These activities include Cooking, Relaxing, Eating, Work, Sleeping, Washing the dishes, Bed to toilet transition, Entering home, Leaving home, Housekeeping, and Respirate. As the activities in the Aruba dataset are typically performed, we modified some of these activities to reflect abnormal behavior. Fig 3 illustrates the layout of the Aruba apartment [17]. The proposed method was evaluated on these datasets, and the results showed its effectiveness in detecting abnormal behavior related to dementia in smart homes.

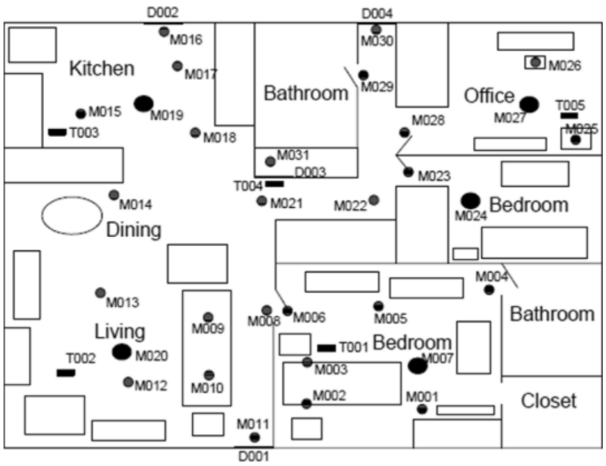



Figure 3: The sensor layout for the Aruba apartment, where circles represent motion/area sensors, and bars represent door/cabinet sensors.

The Kyoto dataset [18] contains data collected from 20 students performing 5 daily activities, both in adlnormal and adlerror versions. The adlnormal version consists of activities performed normally, while the adlerror version includes errors commonly observed in everyday functional independence of elderly people with cognitive impairments. The activities included in this dataset are Make a phone call, Wash hands, Eat, Clean and Cook. Fig. 4 illustrates the sensor layout of the apartment. The sensors can be categorized into motion sensors (M01-M26), item sensors for oatmeal, raisins, brown

[1] https://casas.wsu.edu/datasets


Figure 4: The sensor layout of the Kyoto apartment (left), and the apartment cabinet detail (right)

sugar, bowl, and measuring spoon (I01-I05), medicine container sensor (I06), pot sensor (I07), phone book sensor (I08), cabinet sensor (D01), water sensors (AD1-A, AD1-B), burner sensor (AD1-C), and phone usage (*).

First and foremost, time-period batches are extracted without any preprocessing using sliding window approach [25]. Length of the chosen window is set to 30 seconds empirically. Then this batch is mapped into last-fired representation, meaning the sensor that its binary value changed last in the window shall output 1 until another sensor changes. It is shown that the representation presented in [4] results in better activity recognition compared to the other representations.

### 3.5 Generating anomalies pertaining to dementia

In this research, we seek to distinguish three abnormal behaviors from normal activities. These activities that usually are symptoms of dementia are: Repetition, Disturbance in sleep, and Confusion [25].

Repetition: Elderly people who suffer from dementia are prone to forgetting whether they already performed a particular task or not. Time sensitive activities such as having a meal or drink, taking medicine, personal hygiene, etc. are such activities that only the number of incidents is of importance in medical assessments. For instance, one might forget to have lunch, have multiple lunches [16], may forget to have dinner and start preparing it in a completely irrelevant time.

In the interest of replicating this behavior, we synthesized it by manually inserting a particular set of actions in a random time of normal activities. Hence, it will result in multiple reappearance of that anomaly, happening in different periods of time. As an example, the following instance was generated presuming one forgot to have dinner, brushing teeth, sleeping, preparing dinner, eating, sleeping. Let us assume that $S$ is a sequence of activities $S = d_1.d_2.\ldots.d_n$ that have occurred in a day such as here each $d$ is a period of time pertaining to a certain task, then we choose an arbitrary time period (denoted as $x$) to inset the anomalous data, so $S = d_1.d_2....d_x.a_1.a_2....a_m.d_(x+1)....d_n$. Anomalous data here could be any task that occurs in a specific time, like preparing dinner, brushing teeth, personal hygiene and etc. Thus we have multiple occurrences of any given activity in the sequence.

Disturbance in sleep: One of the symptoms of dementia in late stages is disruption in their sleeping schedule. The older adults suffering from dementia may have significant changes in their sleeping habits and in some cases, they cannot sleep at all. They may wake up numerous times to go to the lavatory [21]. To simulate this behavior, we insert particular activities during normal night-time sequence. These activities are Eating, Bed to toilet transition and Sleeping in an arbitrary area. Again let us assume $S = n_1.n_2....n_n$ sequence is a sequence of night-time activities, time-slice instances of preparing dinner are injected into this sequence so $S = n_1.n_2....n_x.p_1.p_2....p_m.d_(x+1)....d_n$. Hence, we have a cooking and eating dinner activity in middle of the night.

Confusion: Older adults who are suffering from dementia are prone to getting confused during activities. For instance, they may leave the burner on after cooking. In order to test our proposed method against these kind of abnormal behavior, we took advantage of the Kyoto dataset. The adlerror activities in this dataset consist of confusions such as forgetting the medication with the meal or leaving the water running after washing hands, so we assumed that the Confusion abnormal behavior is already reflected in the Kyoto dataset.

## 4 Proposed Method

We have introduced a novel model for anomaly detection in daily activities, which we refer to as ConvLSTMAE, short for ConvLSTM Autoencoder. Unlike traditional autoencoders that use convolutional layers for encoding and decoding, our model utilizes ConvLSTM layers as the encoder and decoder (see Fig. 5). This allows our model to capture both spatial and temporal dependencies in the input data, making it better suited for detecting anomalies in time-series data such as daily activity patterns.

The dataset used in this research is spatiotemporal, meaning that it contains both spatial and temporal correlations. In contrast to the normal Convolutional Layer, our proposed model, called ConvLSTMAE, employs ConvLSTM layers for anomaly detection in daily activities (see Fig. 5). This is because Convolutional Layers in ConvLSTM are capable of processing spatially correlated data, while LSTM layers can process temporally correlated data. Our inspiration for this research comes from the LSTM-based Encoder-Decoder model developed by Malhorta et al. [14], but their model assumes that data instances are not topologically correlated, which is not true for the Aruba dataset. Therefore, we chose to employ ConvLSTM layers in our Autoencoder. Our Encoder comprises two ConvLSTM layers, with 100 and 50 filters, a kernel size of 3x3, and a window size of 30 seconds. The Decoder is an LSTM layer, which achieved similar results to more complex architectures and ConvLSTM cells. We decided not to employ deeper and more complicated models since only two layers of LSTM cells were sufficient to achieve satisfactory results, and more computationally expensive models did not significantly improve performance.

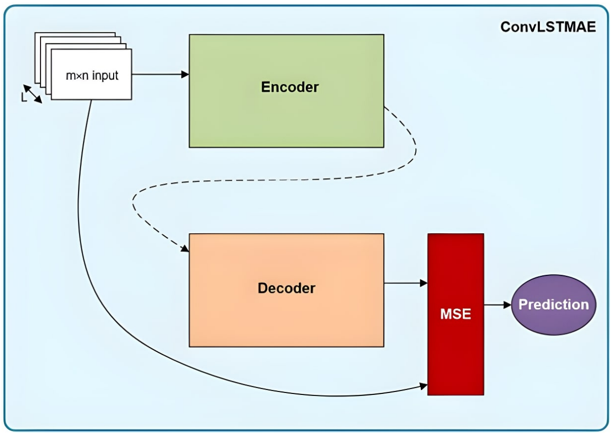


Figure 5: The ConvLSTMAE architecture

Let us assume we have a sequence $X = \{X^{(1)}.X^{(2)}.\cdots.X^{(L)}\}$, where each instance is an M×N matrix which contains reading of each sensor with their corresponding location in the home. The Aruba dataset consists of temporal sequences only, mapping these sequences to the accompanying location, hence forming the matrix, makes it spatiotemporal. Network is first trained on normal data, minimizing the reconstruction error using Stochastic Gradient Descent algorithm.

We assumed that data is presented by Tensor $X \in R^{P\times M\times N}$, which there are $P$ measurements in an M×N grid. As it was mentioned, granular-level patterns do exist in daily behavior. Zhang et al. [29] describes these patterns as movement segments, which can be represented by cumulative movement vector.

Till now, to the best of our knowledge, ConvLSTM-based architectures have been used to analyze and perform anomaly detection on videos and images with great success [15] and it is the first time these cells have been leveraged for anomaly detection in behavior.

## 5 Experimental Results

In this section, we present the experimental results of the ConvLSTMAE model and compare it to other state-of-the-art methods. Before getting into the results, we describe our Experimental Setup.

### 5.1 Experimental setup

For the sake of evaluation, we split the datasets into train and test sets, however it is divided with a fixed time period like day preserves the spatiotemporal characteristics in the data. We also partitioned the Aruba layout, illustrated in Fig. 3, into a 10×8 grid, such that each resulting cell hosts a single sensor. In order to implement proposed architecture, we have deployed Keras Deep Learning library's implementation of the CNNs and LSTM.

To evaluate the performance of our proposed method, ConvLSTMAE, we use standard metrics such as Precision, Recall, Specificity, and F-score. In addition, we plot the Receiver Operating Characteristic (ROC) curve and its area under the curve (AUC). Recall, also known as True Positive (TP) rate, measures the probability of correctly detecting abnormal instances, while Specificity, also known as True Negative (TN) rate, measures the probability of correctly identifying normal behavior. Precision indicates the number of correctly identified positive instances among all positive predictions. F-score is a single metric that balances Precision and Recall. The performance of anomaly detection methods is evaluated based on Recall and Specificity, as they measure the ability of the method to detect abnormal instances and identify normal behavior accurately.

### 5.2 Evaluation of ConvLSTMAE

In this section, we represent the ConvLSTMAE's performance and compare it to the other state-of-the-art methods including OCSVM [3], CNN-LSTM [5] and LSTMAE [14]. In both cases, networks are first trained with normal behavior data and then tested with anomalous data, instances with high reconstruction error (above a certain threshold that was set empirically) were considered abnormal. As we previously mentioned, two possible anomalies have been generated in order to determine ConvLSTMAE's capabilities. Results on the Aruba dataset are depicted in Table 1 and 2, while the results corresponding to the Kyoto dataset are provided in Table 3.

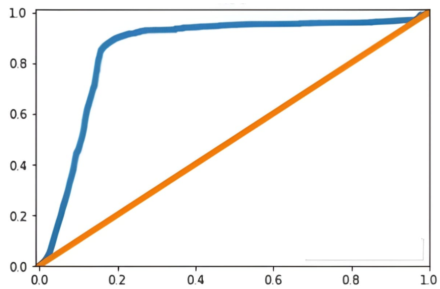
The results regarding the disturbance in sleep anomaly are demonstrated in Table 1. The results show that OCSVM (specificity of 0.91) outperforms all the other methods due to the fact that these anomalies are not temporal and completely out of place. However, since OCSVM is not fit to process time-series data, its performance on the repeat anomaly leaves much to be desired. In the disturbance in sleep anomaly, ConvLSTMAE (with specificity of 0.82) outperformed LSTM (specificity of 0.71) and LSTMAE (specificity of 0.73). The results suggest that ConvLSTMAE is capable of extracting meaningful features even in data that is not inherently spatiotemporal, compared to other state-of-the-art methods. These methods also fail in giving good precision results (0.035 and 0.03 respectively), and their f-score (0.068 and 0.060) is considerably lower than ConvLSTMAE (f-score of 0.095), indicating ConvLSTMAE is capable of differentiating between normal and abnormal behavior.

Table 1: The performance comparison among discussed models for the *disturbance in sleep* anomaly

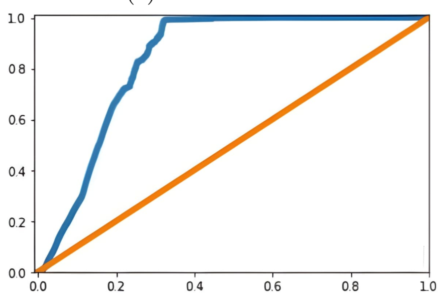
| Model | Recall | Specificity | Precision | F-score | AUC |
|---|---|---|---|---|---|
| CovLSTMAE | 0.83% | 0.82% | 0.050% | 0.095% | 0.862% |
| LSTMAE | 0.77% | 0.73% | 0.031% | 0.060% | 0.803% |
| LSTM | 0.93% | 0.71% | 0.035% | 0.068% | 0.83% |
| OCSVM | 0.94% | 0.91% | 0.114% | 0.204% | 0.93% |

Table 2 displays the results for the repetition anomaly. This type of anomaly is highly correlated with previous instances, and as expected, OCSVM performs poorly. Our proposed method achieved the highest f-score of 0.12, outperforming LSTMAE, LSTM, and OCSVM with f-scores of 0.073, 0.041, and 0.066, respectively. However, our method falls behind on recall compared to LSTMAE, ConvLSTMAE, LSTM, and OCSVM. Despite this, higher precision and specificity rates indicate significant improvement over other methods. OCSVM's low precision and recall values of 0.037 and 0.36, respectively, show that it fails to learn class-specific features to correctly differentiate between anomaly and normal behavior. In contrast, CNN can detect changes in feature patterns implicitly, as each activity is formed by steps captured by motion sensors throughout the home. For example, having dinner in the middle of the night requires going to the kitchen, starting the oven, and preparing food. LSTM also captures repetition-related activities. However, our proposed method cannot detect anomalies in specialized activities such as adding too much seasoning to food, having a bland meal due to a patient forgetting to add salt, confusing the location of a certain item, or leaving kitchen utilities on because there are no sensors pertaining to these activities in the Aruba dataset. Future research is expected to address this issue. Furthermore, our method may detect abnormal behavior with a delay if there is a gradual deterioration in the patient's health, as the reconstruction error needs to reach a minimum threshold. Fine-tuning the network for each individual can somewhat resolve this issue. The ROC curves corresponding the discussed models for the repetition anomaly is illustrated in Figure 6. For the confusion anomaly, the ROC curve represent pretty much the same pattern as of the repetition anomaly, so we just provided one of them.

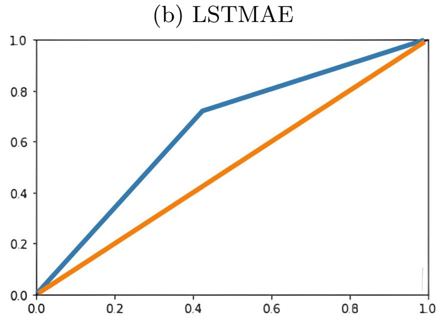
Finally, the results of the confusion anomaly are presented in Table 3. Our proposed method, ConvLST-



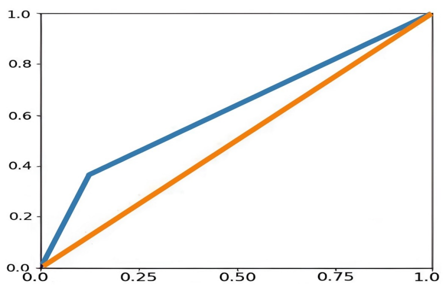

(a) ConvLSTMAE




(b) LSTMAE




(c) LSTM




(d) OCSVM

Figure 6: The ROC curves, corresponding the discussed models, for the *disturbance in sleep* anomaly. (The ROC curve plot the recall (vertical axis) against the false positive rate at each threshold setting.)

MAE, achieved the best result with an AUC of 0.903, but its performance was statistically similar to LSTMAE. The convolutional layer's contribution was more pronounced in the Aruba dataset compared to the Kyoto dataset, as our method relies on normal daily activities to detect anomalies based on the timing of their occurrence. Nevertheless, our method offers advantages, such as not requiring labeled data to distinguish anomalies (which is not always available in real-world datasets), and the floating threshold can be used for fine-tuning. In contrast, OCSVM requires the anomaly ratio to be specified beforehand, which is not a valid assumption in many cases.

Table 2: The performance comparison among discussed models for the *repetition* anomaly

| Model | Recall | Specificity | Precision | F-score | AUC |
|---|---|---|---|---|---|
| CovLSTMAE | 0.84% | 0.84% | 0.065% | 0.12% | 0.853% |
| LSTMAE | 0.98% | 0.67% | 0.038% | 0.073% | 0.836% |
| LSTM | 0.71% | 0.57% | 0.021% | 0.041% | 0.65% |
| OCSVM | 0.36% | 0.87% | 0.036% | 0.066% | 0.62% |

Table 3: The performance comparison among discussed models for the *confusion* anomaly

| Model | Recall | Specificity | Precision | F-score | AUC |
|---|---|---|---|---|---|
| CovLSTMAE | 0.957% | 0.793% | 0.538% | 0.688% | 0.903% |
| LSTMAE | 0.955% | 0.792% | 0.536% | 0.687% | 0.899% |
| LSTM | 0.81% | 0.78% | 0.48% | 0.61% | 0.81% |
| OCSVM | 0.81% | 0.58% | 0.33% | 0.47% | 0.70% |

## 6 Conclusions

The results presented in this paper demonstrate the effectiveness of the proposed ConvLSTMAE method in detecting abnormal behavior related to dementia in smart homes. Compared to state-of-the-art methods, our approach achieved significant improvements in performance metrics such as precision, recall, and F-score. Moreover, the use of AE-based methods provides a generalizable framework for detecting various abnormal behaviors in elderly individuals with dementia. However, there are still limitations in detecting anomalies related to specialized activities or gradual deterioration in the patient's health. Future research can address these challenges and refine the proposed method. Overall, our work presents a promising direction for using deep learning techniques to address important healthcare challenges.

## References

[1] R. Ahmad, B. Yang, G. Ettlin, A. Berger, and P. Rodríguez-Bocca. A machine-learning based convlstm architecture for ndvi forecasting. *International Transactions in Operational Research*, 30(4):2025–2048, 2023.

[2] K. A. Alaghbari, M. H. M. Saad, A. Hussain, and M. R. Alam. Activities recognition, anomaly detection and

next activity prediction based on neural networks in smart homes. *IEEE Access*, 10:28219–28232, 2022.

[3] S. Alam, S. K. Sonbhadra, S. Agarwal, and P. Nagabhushan. One-class support vector classifiers: A survey. *Knowledge-Based Systems*, 196:105754, 2020.

[4] D. Arifoglu and A. Bouchachia. Activity recognition and abnormal behaviour detection with recurrent neural networks. *Procedia Computer Science*, 110:86–93, 2017.

[5] D. Arifoglu and A. Bouchachia. Detection of abnormal behaviour for dementia sufferers using convolutional neural networks. *Artificial intelligence in medicine*, 94:88–95, 2019.

[6] D. J. Cook and M. Schmitter-Edgecombe. Assessing the quality of activities in a smart environment. *Methods of information in medicine*, 48(05):480–485, 2009.

[7] L. Fang, Y. Wu, C. Wu, and Y. Yu. A nonintrusive elderly home monitoring system. *IEEE Internet of Things Journal*, 8(4):2603 – 2614, 2021.

[8] P. Franco, Y. Martinez, JM. amd Kim, and M. Ahmed. Iot based approach for load monitoring and activity recognition in smart homes. *IEEE Access*, 9:45325–45339, 2021.

[9] L. Gillani Fahad and S. F. Tahir. Activity recognition and anomaly detection in smart homes. *Neurocomputing*, 423:362–372, 2021.

[10] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT press, 2016.

[11] N. Gulati and P. Kaur. Friendcare-aal: a robust social iot based alert generation system for ambient assisted living. *J Ambient Intell Human Comput*, 13:1735–1762, 2022.

[12] N. Hammerla, J. Fisher, P. Andras, L. Rochester, R. Walker, and T. Plötz. Pd disease state assessment in naturalistic environments using deep learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29, 2015.

[13] O. D. Lara and M. A. Labrador. A mobile platform for real-time human activity recognition. In *2012 IEEE consumer communications and networking conference (CCNC)*, pages 667–671. IEEE, 2012.

[14] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff. Lstm-based encoder-decoder for multi-sensor anomaly detection. *arXiv preprint arXiv:1607.00148*, 2016.

[15] J. R. Medel and A. Savakis. Anomaly detection in video using predictive convolutional long short-term memory networks. *arXiv preprint arXiv:1612.00390*, 2016.

[16] M. Moishin, R. C. Deo, R. Prasad, N. Raj, and S. Abdulla. Designing deep-based learning flood forecast model with convlstm hybrid algorithm. *IEEE Access*, 9:50982–50993, 2021.

[17] Y. Nawal, M. Oussalah, and B. Fergani. New incremental svm algorithms for human activity recognition in smart homes. *J Ambient Intell Human Comput*, 14:13433–13450, 2023.

[18] F. J. Ordóñez and D. Roggen. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1):115, 2016.

[19] B. Pik Lik Lau, Z. Koh, Y. Zhou, B. Kai Kiat Ng, C. Yuen, and M. Liang Low. Location-based activity behavior deviation detection for nursing home using iot devices. *Internet of Things*, page 100702, 2023.

[20] C. A. Ronao and S.-B. Cho. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert systems with applications*, 59:235–244, 2016.

[21] J. Saives, C. Pianon, and G. Faraut. Activity discovery and detection of behavioral deviations of an inhabitant from binary sensors. *IEEE Transactions on Automation Science and Engineering*, 12(4):1211–1224, 2015.

[22] M. Sharma and P. Kaur. Xlaam: explainable lstm-based activity and anomaly monitoring in a fog environment. *J Reliable Intell Environ*, Internet of Things, 2022.

[23] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28, 2015.

[24] J. H. Shin, B. Lee, and K. S. Park. Detection of abnormal living patterns for elderly living alone using support vector data description. *IEEE Transactions on Information Technology in Biomedicine*, 15(3):438–448, 2011.

[25] T. L. van Kasteren, G. Englebienne, and B. J. Kröse. Human activity recognition from wireless sensor network data: Benchmark and software. In *Activity recognition in pervasive intelligent environments*, pages 165–186. Springer, 2011.

[26] Y. Wang, X. Wang, D. Arifoglu, C. Lu, A. Bouchachia, Y. Geng, and G. Zheng. A survey on ambient sensor-based abnormal behaviour detection for elderly people in healthcare. *Electronics*, 12(7), 2023.

[27] J. Yang, M. N. Nguyen, P. P. San, X. Li, and S. Krishnaswamy. Deep convolutional neural networks on multichannel time series for human activity recognition. In *Ijcai*, volume 15, pages 3995–4001. Buenos Aires, Argentina, 2015.

[28] M. Zeng, L. T. Nguyen, B. Yu, O. J. Mengshoel, J. Zhu, P. Wu, and J. Zhang. Convolutional neural networks for human activity recognition using mobile sensors. In *6th international conference on mobile computing, applications and services*, pages 197–205. IEEE, 2014.

[29] T. Zhang, W. Fu, J. Ye, and M. Fischer. Learning movement patterns of the occupant in smart home environments: an unsupervised learning approach. *Journal of Ambient Intelligence and Humanized Computing*, 8:133–146, 2017.