2021 Third International Conference on Transdisciplinary AI (TransAI) | 978-1-6654-3412-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/TransAI51903.2021.00018

FallWatch: A Novel Approach for Through-Wall Fall Detection in Real-Time for the Elderly Using Artificial Intelligence

Aditya Chebrolu
Independence High School
Frisco, TX, USA
aditya.chebrolu.081@K12.friscoisd.org

***Abstract*—Falls are the leading cause of fatal injury in the elderly. Presently available fall-detection devices have many drawbacks including potential blind spots and low lighting, lack of privacy, and the need for the elderly to operate these devices despite cognitive decline. Radio-frequency (RF) imaging presents a promising solution as it is able to traverse through most materials while remaining highly reflective off of humans. FallWatch was designed as an artificial intelligence model to detect falls in real-time in spite of visual obstruction using RF signals while overcoming the drawbacks of RF including low resolution imaging and body-part specularity. Using an RF antenna array, multiple fall and non-fall examples were captured through several mediums of obstruction in cross-person and cross-environment settings. The data obtained was trained on a deep learning model consisting of: 1) Convolutional Neural Network to extract relevant information and capture spatial relationships, 2) Attention Mechanism to allow generalization to new people and environments, and 3) Recurrent Neural Network with Long Short-Term Memory to capture temporal relationships between RF frames. FallWatch was successful in detecting falls not only in through-wall scenarios, but also in cross-person and cross-environment settings while surpassing the performance of other fall detection systems. In conclusion, FallWatch presents a novel end-to-end approach for fall detection in the elderly and enables their monitoring in multiple care settings.**

***Index Terms*—Fall detection, deep learning, Radio-frequency imaging, Convolutional Neural Network, Attention Mechanism, Recurrent Neural Network with Long Short-Term Memory**

I. Introduction

Falls are the leading cause of injury and death in adults 70 and above [1]. In fact, in the US alone, an elderly individual falls every second, amounting to 36 million falls, 32,000 deaths, and $50 billion in costs annually [2]. Many of these falls may remain undetected for many hours as more than 12 million elderly individuals live alone [3]. One of the most serious consequences of undetected falls is the ”long lie”, in which the elderly individuals remain on the ground for more than one hour [4], [5]; a study of 125 adults over 65 years found that 50% died within 6 months of sustaining a long lie [1]. As such, there is a pressing need for early and effective fall detection for the elderly.

Current fall detection devices primarily consist of contact devices or non-contact devices. Contact devices commonly come in the form of wearable devices and sensors, such as wrist bands or pendants. These devices commonly monitor parameters such as heart rate variability, electrocardiogram, pulse oximetry, and kinematic attributes measured by accelerometers, gyroscopes, and magnetometers [6]. This data is then fed to threshold-based systems to detect falls. However, elderly individuals may forget to wear or charge these devices due to memory loss and decline of cognitive function. Additionally, many elderly individuals may be encumbered or constrained by these devices.

Non-contact devices, on the other hand, rely on the real-time capturing of images or video frames by a camera around the region of interest [6]. This data is then given to an image processing mechanism that is used to determine if a fall has occurred. Such devices, nevertheless, pose issues of lack of privacy, especially if an individual is to be monitored in secluded areas such as the bathroom. Additionally, non-contact devices may overlook blind spots, are unable to function in poor lighting conditions, and are hindered by visual obstruction.

Radio-frequency (RF) imaging presents an interesting solution for fall detection as it is able to:

- Traverse materials of low reflective indexes such as walls and furniture, while remaining highly reflective off humans, thereby overcoming blind spots and low lighting conditions [7].
- Retain privacy of the elderly individual as the acquired images are vague and an individual’s identity cannot be determined from a RF image with the naked eye.
- Operate wirelessly and completely independently of the elderly adult.

The above features make RF an ideal imaging wave to harness for the domain of fall detection as it transcends limitations present in both contact and non-contact devices. Hence, this study aims to design a device to detect falls in real-time for the elderly with RF imaging that can work despite of visual obstruction.

II. Related Works

Several researchers have proposed systems similar to FallWatch’s intended design, in which a low-power signal is

978-1-6654-3412-6/21/$31.00 ©2021 IEEE
DOI 10.1109/TransAI51903.2021.00018

Authorized licensed use limited to: Princeton University. Downloaded on July 08,2022 at 16:43:17 UTC from IEEE Xplore. Restrictions apply.

transmitted to the monitored individual, and the captured reflection intensities are used to infer falls [8]–[11].

Initial approaches to fall detection involved rule-based algorithms in which several RF sensors were deployed across the perimeter of a room. *Mager et al* attempted to find the relationship between the attenuation of RF signals based on a function of a person's horizontal and vertical position to determine fall state [8]. A two-level array of RF sensor nodes were deployed around the room, and the changes in the received signal strength indicator from such sensors were analyzed to infer falls. However, such systems are cumbersome to set up and are limited to only very basic examples and activities, i.e. fall, walk, sit, and bend. Such approaches additionally are unable to generalize to new environments and people.

The simplicity of rule-based algorithms motivated research utilizing artificial intelligence in fall detection. Of particular interest is the use of machine learning for fall detection, in which systems are able to learn from data to leverage patterns and hence act without being explicitly programmed. Recent solutions involving machine learning can be divided into either Doppler radar [9] and Channel State Information (CSI) based solutions [10], [11]. Both methods attempt to measure the velocity of motion as a function of time. Doppler-based approaches directly measure motion velocity through its relationship with the Doppler frequency. CSI-based approaches attempt to measure fast changes in RF signals with the wavelet transform or Short Time Fourier Transform. However, such models are unable to learn complex patterns and dependencies in RF images as they are not explicitly designed for the RF domain. As a result, these models are unable to adapt to diverse activities. Additionally, such models are unable to scale and adapt to new environments and people.

In summary, previous RF-based fall detection designs were found to have several drawbacks that are listed below:

- *Physical Challenges of RF:* RF images captured at frequencies designed for penetrative applications have several physical shortcomings, namely an intrinsically low resolution and the issue of specularity in humans [12]. Previous designs consisted of architectures that are too simplistic to leverage relationships in RF images and solve such problems.
- *Generalization to New People & Environments:* Previous designs are unable to separate reflections from different sources in space. When the environment or person changes, the received RF signatures may alter when compared to the training environment, hence modifying how fall and non-fall patterns appear. This renders devices unable to generalize to new environments and people.
- *Diverse Dataset:* No previous study detailing an RF-based fall detection design has been explicitly trained for through-wall scenarios. Attenuation and noise associated with wave propagation in complex media must be addressed by a robust RF fall detection device. Additionally, the list of activities in previous studies is relatively simple and these activities have to be performed in a specific way with a particular orientation to the radio, greatly hindering the scope of a fall detection device.

In this paper, a novel approach is presented to overcome the aforementioned shortcomings of the previous works. As such, this paper introduces FallWatch, a novel artificial intelligence based real-time methodology that allows fall detection in the elderly in spite of visual obstruction using RF imaging. A RF transceiver sends a low power wireless signal to the given environment and captures the reflection intensities of the received signals. Using a deep learning model, FallWatch then analyzes the reflection intensities to detect falls in real-time.

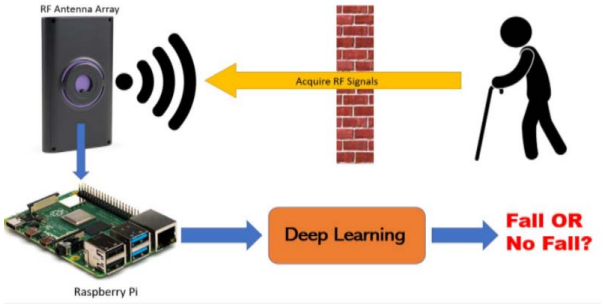

Fig. 1: Intended design of FallWatch.

III. Procedures

A. Methodology Overview

This paper proposes a system in which RF signals are computationally analyzed using a deep learning model to detect falls in real time in spite of the presence of visual obstruction.

An analysis of the problem leads to the formulation of a six-step pipeline: 1) Data Acquisition, 2) Data Pre-Processing, 3) Design Considerations, 4) Deep Learning Model Architecture, 5) Model Training Process, and 6) Real-Time Setup & Execution. Each section will explain the implementation and the relation behind the design of this device.

B. Data Acquisition

1) RF Signals: An RF signal is an electromagnetic wave which can be sampled to record its amplitude and phase. The phase of an RF signal is linearly related to its traveled distance.

2) Antenna Arrays: Antenna arrays can be utilized to determine the spatial direction from which an RF signal arrives with far greater accuracy than with simply one antenna. The larger an antenna array is, the better it is able to estimate spatial direction. As such, the angular resolution of a given antenna array can be expressed as follows:

Authorized licensed use limited to: Princeton University. Downloaded on July 08,2022 at 16:43:17 UTC from IEEE Xplore. Restrictions apply.

$$\Delta\theta = \frac{0.886\lambda}{L} \tag{1}$$

where L is the length of the antenna array and λ is the wavelength of the signal.

3) Frequency Modulated Carrier Wave (FMCW): FMCW is a special technique that can measure the depth of an RF reflector. This is done by transmitting a "frequency chirp", or a periodic RF signal whose frequency increases with respect to time. The RF transceiver can measure the time-of-flight between the transmission and receival of the RF signal to infer depth. This is done by measuring the change in frequency between the transmission and receival of the RF signal. A frequency chirp of slope k can be used to compute the reflection intensity P from a depth r as:

$$P(r) = \left| \sum_{t=1}^{T} s_t e^{j2\pi \frac{kr}{c} t} \right| \tag{2}$$

where s_t is the baseband time signal, c is the speed of light, and T is the duration of each chirp.

4) Data Acquisition: The Walabot Developer, a commercially available, FCC-compliant radio, is utilized as the RF signal transceiver in this study. It consists of an FMCW antenna array whose API returns a 3-Dimensional (3D) distribution, where each point represents the RF reflection intensity at the corresponding voxel. The Walabot has a frequency range of 3.3-10 GHz and has a transmission power that is roughly 1/1000 the power of Wi-Fi.

3,150 training samples of actions were collected in multiple indoor environments from participants aged 10-50 years by the researcher. Data was also *explicitly* collected through obstructive media such as wood, brick, drywall, cloth, in addition to open air environment to simulate through-wall scenarios. Several variations of fall and non-fall examples were collected; a full list of collected fall and non-fall examples is provided below:

Fall Examples	Non-Fall Examples
Falling from a chair	Standing
Losing consciousness	Sitting
Tripping and falling	Walking
Slipping and falling	Exercise & Yoga
Losing balance	Eating & Drinking
	Sleeping & Laying Down
	Picking up an object
	Attempting to sit in a chair
	No person

Non-fall examples highlighted in red represent adversarial examples designed to fool the neural network since they have motion similar to that of a fall. Additionally, scenes with no people were collected for the model to avoid mistaking ambient RF signatures for human falls.

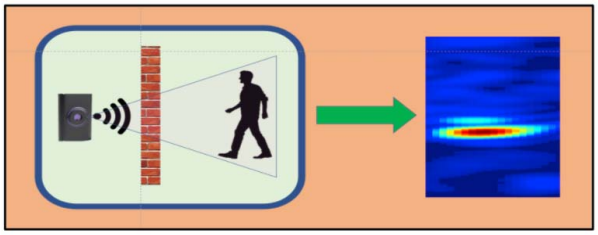

Fig. 2: Data Collection Setup

C. Data Pre-Processing

1) Cartesian Coordinate Conversion: The Walabot returns image data for a given fixed space in spherical coordinates. However, convolutional neural networks (CNNs) leverage spatially linked correlations in data to reduce the total number of parameters to be learned. Hence, to attain a more 3D representation of the data, spherical coordinates are converted to Cartesian coordinates using the following formulae:

$$\begin{cases} X = r * \sin\theta \\ Y = r * \cos\theta * \sin\phi \\ Z = r * \cos\theta * \cos\phi \end{cases} \tag{3}$$

2) Dimensionality Reduction: The processing of a 3D point cloud per each timestep is computationally expensive. Hence, the acquired RF data is reduced from a 3D point cloud to two 2-dimensional (2D) heatmaps that represent the vertical and horizontal projections of the reflected RF signals. This is done by summing the values of the reflection intensity over 2 planes, as shown below:

$$\begin{cases} R_{horz} = \sum_{y=y_{min}}^{y_{max}} R(X,Y,Z) \\ R_{vert} = \sum_{z=z_{min}}^{z_{max}} R(X,Y,Z) \end{cases} \tag{4}$$

Thus, the input data takes the form of these 2D heatmaps. This process not only allows for data compression but also retains important features to be learned by the deep learning model.

D. Design Considerations

Signals from the RF spectrum have significantly different properties from those of the visible light spectrum [7]. This leads to several challenges detailed below:

- *Low Spatial Resolution:* Penetrative RF applications have an extremely low spatial resolution and cannot provide high detail of the surrounding environment.
- *Body-Part Specularity:* According to the law of reflection, the angle of incidence is equivalent to the angle of

Authorized licensed use limited to: Princeton University. Downloaded on July 08,2022 at 16:43:17 UTC from IEEE Xplore. Restrictions apply.

reflection for a surface in which the wavelength is smaller than the roughness of the surface. As such, only signals that fall normal to a given surface are reflected back to the RF transceiver. However, when the wavelength is larger than the roughness of the surface, the reflected signals are spread out or "scattered" in various directions [12]. The wavelength of the RF transceiver used in this study is 5 cm, and hence people function as reflectors. However, the human body has an irregular surface, causing many signals to scatter and not be reflected back to the transceiver. Moreover, as a person moves, different body parts may reflect signals towards the RF transceiver or away from it, causing body parts to become invisible to view at different timesteps.

The above physical properties of RF imaging must be taken into consideration for designing an accurate fall detection system.

E. Deep Learning Model Architecture

Fall detection is formulated as a binary classification problem, i.e. fall or non-fall. This section details how FallWatch explicitly resolves the aforementioned design criteria in its deep learning design.

1) Inception-Based Convolutional Neural Network: CNNs have been a recent workhorse in deep learning models for image processing [13]. In contrast to traditional multilayer perceptron models, where each neuron is connected to all the neurons in the next layer, each neuron is only locally connected to a few neurons in the next layer in CNNs. This preserves spatial information within the data while greatly reducing the number of input parameters, allowing CNNs to learn complex features within images. As the CNN acquires relevant information from input images, it is able to overcome the low spatial resolution barrier of RF imaging.

CNNs utilize convolution operations with a matrix of numbers (referred to as a filter or kernel) that slides across the entire image and transforms the image based on the values from the filter. Mathematically, this operation can be defined as follows for a filter with size (k, k):

$$(w * x)_{i,j} = \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} w_{l,m} * x_{i-l,j-m} \tag{5}$$

where the input image is denoted by x represented as a matrix and the kernel by w.

Inception modules are also added within the CNN [14]. Inception modules contain filters of multiple sizes. This allows for the multiscale processing of data and hence the acquisition of a wider range of information. The neural network is able to learn which convolutions are more relevant per each module.

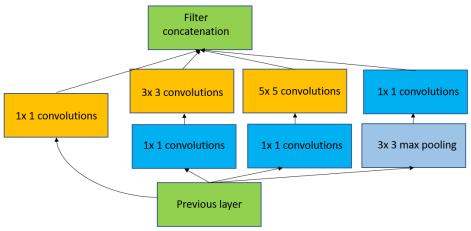

Fig. 3: Architecture of an inception module.

2) Soft-Attention Based Mechanism: Attention in neural networks is inspired by how humans do not focus at an entire given scene at once; rather, they selectively focus on different parts of the scene [15]. Attention-based mechanisms can allow the neural network to focus on more relevant information in a given frame, such as a person, and ignore the stochastic noise due to RF reflection intensities of the surrounding environment. As such, attention can allow the neural network to adapt to new people and environments where RF signatures vary from the training dataset. Attention can be implemented with either soft attention, which is deterministic and can be trained via backpropagation [16], or hard attention, which consists of a stochastic sampling model and can be trained using the Monte Carlo method [15].

In this work, soft attention is implemented at a frame level (i.e., within a single frame). The feature maps (representing relevant information per spatial regions of an image) of the last convolutional layer of the CNN are extracted. Given a set of feature maps $(x_1, x_2, x_3 \dots x_i)$ and a context vector h_{t-1} representing accumulated information of feature maps from previous timesteps, a score s_i can be calculated to measure the probability by which the model believes the region in the given frame is important:

$$s_i = tanh(W_c h_{t-1} + W_x x_i) \tag{6}$$

The score s_i is then passed to a softmax for normalization to compute the weight a_i given to this particular region:

$$a_i = \frac{exp(s_i)}{\sum_i exp(s_i)} \tag{7}$$

A weighted average can then be computed for all $x_1, x_2, x_3 \dots x_i$:

$$Z = \sum_i a_i x_i \tag{8}$$

This average Z is then passed to the LSTM detailed in the following section.

Authorized licensed use limited to: Princeton University. Downloaded on July 08,2022 at 16:43:17 UTC from IEEE Xplore. Restrictions apply.

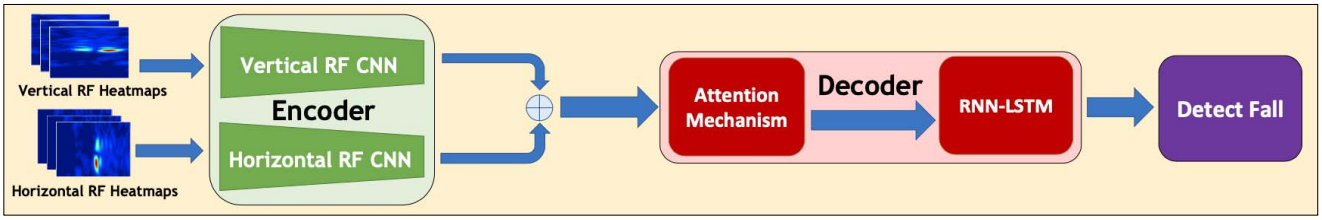

Fig. 4: Model architecture of FallWatch. The encoder takes a sequence of RF images as input and encapsulates it into a state with a fixed shape. The decoder then maps the encoded state to the model's prediction of a fall or no fall.

3) Recurrent Neural Network with Long-Short Term Memory: Recurrent Neural Networks (RNNs) are a variant of traditional neural networks; RNNs maintain an internal state that represents context information over a given sequence of data [17]. This makes RNNs useful for modeling temporal dependencies.

Long-Short Term Memory (LSTMs) have been a workaround for the vanishing gradient problem that plagues traditional RNNs [18]. Unlike traditional RNNs, LSTMs regulate the flow of information through their state through three gate matrices: an input gate that determines which information gets through, a forget gate that determines which information to forget, and an output gate that determines which information to output. These gate matrices are calculated for each timestep. Mathematically, the gates can be computed as follows,

$$\begin{aligned} i^t &= \sigma(W_i * [h^{t-1}, x^t] + b_i) \\ f^t &= \sigma(W_f * [h^{t-1}, x^t] + b_f) \\ o^t &= \sigma(W_o * [h^{t-1}, x^t] + b_o) \end{aligned} \tag{9}$$

where i^t is the input gate, f^t is the forget gate, and o^t is the output gate. Through leveraging temporal dependencies, LSTMs are able to aggregate RF information over time and can hence mitigate the issue of body-part specularity. As such, a robust fall detection system not only requires location information of a given person, but also the spatio-temporal information that embodies diverse fall and non-fall patterns. In comparison to *Y. Tian et al.* [19], who attempted to leverage spatio-temporal information with 3D convolutions for RF-based fall detection, the CNN spatial information extractor and the LSTM temporal associator are explicitly separated in FallWatch's design, as this has been shown to learn spatio-temporal relationships more effectively [20].

F. Model Training Process

The model is trained end-to-end to minimize a loss function that captures the difference between the model's predictions and the ground truth labels. All parameters of the neural network are optimized with a binary cross entropy loss:

$$\mathcal{L}(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(\hat{y}_i) + (1 - y_i) \cdot log(1 - \hat{y}_i) \tag{10}$$

However, to improve convergence and reduce potential overfitting, an L2 regularization term is introduced into the loss function. At the high level, the regularization term attempts to make the parameters of the model more homogeneously distributed by penalizing large weights [21]. This prevents the network from developing a small set of parameters with large values, hindering generalization to the test set. The L2 regularization term is given below:

$$\mathcal{R}(M) = \lambda \sum_{j=0}^{M} W_j^2 \tag{11}$$

Hence, the final loss function used to optimize the model is given below:

$$\mathcal{L} = \mathcal{L}(y, \hat{y}) + \mathcal{R}(M) \tag{12}$$

The model is trained for 50 epochs at a learning rate of 0.001. The Adaptive Moment Estimation (ADAM) optimizer is used to improve training efficiency and speed by maintaining a per-parameter learning rate that improves performance on problems with sparse gradients [22].

G. Real-Time Setup & Execution

The processing speed of the model was calculated to be 23 ms per image, verifying the real-time capability of FallWatch.

IV. Results & Discussion

A. Evaluation Setup

FallWatch's performance is evaluated by comparing its predicted output to the human labeled ground truth. The data is split into an 80:20 ratio for the training and testing set, respectively.

As the ratio of falls to non-falls in the dataset is not equally distributed, accuracy is not used to gauge FallWatch's performance. Instead, FallWatch is evaluated in terms of the sensitivity and specificity of fall detection. Through acquiring

Authorized licensed use limited to: Princeton University. Downloaded on July 08,2022 at 16:43:17 UTC from IEEE Xplore. Restrictions apply.

the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) examples, the following metrics can be calculated:

- *Precision (p):* The proportion of correctly identified falls out of all identified falls.
- *Recall (r):* The proportion of correctly identified falls out of all fall examples.
- *F1 Score (f):* The harmonic mean of precision and recall.

$$p = \frac{TP}{TP + FP} \qquad r = \frac{TP}{TP + FN} \qquad f = \frac{2 * p * r}{p + r}$$

FallWatch is also compared to other leading RF-based fall detection systems, namely FallDeFi [10] and RT-Fall [11]. Although previous studies are compared on a different dataset, FallWatch's dataset is larger and more diverse, as shown in the table below.

TABLE I: FallWatch Dataset Comparison to Prior Works

	FallWatch Dataset	FallDeFi & RT-Fall Dataset
Number of Falls	567	326
Number of Non-Falls	2583	744
Number of Fall Patterns	5	4
Number of Non-Fall Patterns	9	8
Number of People	5	3
Number of Environments	5	5

B. Overall Performance

FallWatch's overall performance is first tested by training and testing it on all data. FallWatch is able to detect 97.3% of the falls (recall) while delivering a precision as high as 96.4%. FallWatch surpasses previously published RF-based fall detection systems, especially in terms of precision.

TABLE II: FallWatch's Overall Performance

	Precision	Recall	F1 Score
FallWatch	0.964	0.973	0.969
FallDeFi	0.866	0.943	0.902
RT-Fall	0.800	0.852	0.825

C. Through-Wall Performance

FallWatch's performance behind obstructive mediums (through-wall) is compared to its performance in line of sight scenarios (open-air). The same fall detection model is evaluated without any new training. However, the test data is stratified into two groups: whether the monitored individual in the given sample is behind an obstructive medium or not.

As RF signals may be attenuated or scattered through obstructive mediums, a slight decrease in performance is to be expected as represented in the data for through-wall scenarios. Overall, however, the table demonstrates how FallWatch is able to work well in spite of the presence of visual obstruction.

TABLE III: FallWatch's Through-Wall Performance

	Precision	Recall	F1 Score
Open-Air	0.966	0.974	0.970
Through-Wall	0.923	0.917	0.920

D. Same vs. Cross Environment & Person Performance

FallWatch's performance in same vs. cross environments & people is analyzed. The test set of the model is stratified into three equal components: samples captured with the same people & environments as the training dataset, samples captured with the same people but different environments from the training dataset, and samples captured with different people & different environments from the training dataset.

There is only a marginal decline in terms of performance between cross person, cross environment settings and same person of 0.022, 0.050, and 0.03 in terms of precision, recall, and F1 score, respectively. This confirms that FallWatch is able to generalize well to other environments & people beyond the training dataset without much loss of performance.

TABLE IV: FallWatch's Performance Under Different Settings

	Precision	Recall	F1 Score
Same Person, Same Environment	0.978	0.984	0.980
Same Person, Cross Environment	0.967	0.951	0.959
Cross Person, Cross Environment	0.956	0.934	0.950

E. Cross Environment & Person Performance

Lastly, FallWatch's cross environment, cross person performance is compared to other published RF-based fall detection systems. FallWatch is able to detect 95.6% of the falls (recall) while delivering a precision as high as 93.4%. In comparison, the previous best performing design is able to detect 94.3% of falls but only delivers a precision of 76.1%. Hence, FallWatch's design significantly improves upon previously published devices in terms of generalization to new environments and people.

TABLE V: Comparison of FallWatch's Performance in Different Environments & People to other Fall-Detection Systems

	Precision	Recall	F1 Score
FallWatch	0.956	0.934	0.950
FallDeFi	0.761	0.943	0.902
RT-Fall	0.566	0.628	0.577

V. Conclusion

In summary, FallWatch presents a novel end-to-end solution to detect falls even in the presence of visual obstruction while also allowing for generalization to new people and new environments. Using several state-of-the-art deep learning

Authorized licensed use limited to: Princeton University. Downloaded on July 08,2022 at 16:43:17 UTC from IEEE Xplore. Restrictions apply.

techniques in sequential synchrony, the model achieves a high performance that allows it to be an excellent non-intrusive approach for fall detection for the elderly and to monitor patients in any care settings. FallWatch is a paradigm shift in fall detection in the elderly that functions not only wirelessly and independently, but also maintains privacy, and overcomes the challenges of blind spots, visual obstruction, & low lighting.

A demonstration video is filmed to depict FallWatch's real-time RF image processing: **https://youtu.be/fNNcmG5gsFg**.

REFERENCES

[1] P. H. Millard, "How dangerous are falls in old people at home?" *British Medical Journal (Clinical research ed.)*, vol. 282, no. 6267, p. 905, 1981.

[2] L. Z. Rubenstein, "Falls in older people: Epidemiology, risk factors and strategies for prevention," *Age and ageing*, vol. 35, no. suppl_2, pp. ii37–ii41, 2006.

[3] R. Stepler, *Smaller share of women ages 65 and older are living alone*, Aug. 2020. [Online]. Available: https://www.pewresearch.org/social-trends/2016/02/18/smaller-share-of-women-ages-65-and-older-are-living-alone/.

[4] J. Fleming and C. Brayne, "Inability to get up after falling, subsequent time on floor, and summoning help: Prospective cohort study in people over 90," *Bmj*, vol. 337, 2008.

[5] E. J. Bisson, E. W. Peterson, and M. Finlayson, "Delayed initial recovery and long lie after a fall among middle-aged and older people with multiple sclerosis," *Archives of physical medicine and rehabilitation*, vol. 96, no. 8, pp. 1499–1505, 2015.

[6] A. Ramachandran and A. Karuppiah, "A survey on recent advances in wearable fall detection systems," *BioMed research international*, vol. 2020, 2020.

[7] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 6, pp. 1–13, 2015.

[8] B. Mager, N. Patwari, and M. Bocca, "Fall detection using rf sensor networks," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, IEEE, 2013, pp. 3472–3476.

[9] B. Jokanović and M. Amin, "Fall detection using deep learning in range-doppler radars," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 1, pp. 180–189, 2017.

[10] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "Falldefi: Ubiquitous fall detection using commodity wi-fi devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 4, pp. 1–25, 2018.

[11] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "Rt-fall: A real-time and contactless fall detection system with commodity wifi devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, pp. 511–526, 2016.

[12] P. Beckmann and A. Spizzichino, "The scattering of electromagnetic waves from rough surfaces," *Norwood*, 1987.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[14] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[15] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention," in *International conference on machine learning*, PMLR, 2015, pp. 2048–2057.

[16] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[17] M. I. Jordan, "Attractor dynamics and parallelism in a connectionist sequential machine," in *Artificial neural networks: concept learning*, 1990, pp. 112–127.

[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[19] Y. Tian, G.-H. Lee, H. He, C.-Y. Hsu, and D. Katabi, "Rf-based fall monitoring using convolutional neural networks," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–24, 2018.

[20] Z. Zuo, B. Shuai, G. Wang, X. Liu, X. Wang, B. Wang, and Y. Chen, "Convolutional recurrent neural networks: Learning spatial dependencies for image representation," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 18–26.

[21] A. Krogh and J. A. Hertz, "A simple weight decay can improve generalization," in *Advances in neural information processing systems*, 1992, pp. 950–957.

[22] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

Authorized licensed use limited to: Princeton University. Downloaded on July 08,2022 at 16:43:17 UTC from IEEE Xplore. Restrictions apply.