Sensor Fusion for Identification of Freezing of Gait Episodes Using Wi-Fi and Radar Imaging

Syed Aziz Shah, Ahsen Tahir, Jawad Ahmad, Adnan Zahid, *Graduate Student Member, IEEE*, Haris Pervaiz, Syed Yaseen Shah, Aboajeila Milad Abdulhadi Ashleibta, Aamir Hasanali, Shadan Khattak, and Qammer H. Abbasi

Abstract—**Parkinson's disease (PD) is a progressive and neurodegenerative condition causing motor impairments. One of the major motor related impairments that present biggest challenge is freezing of gait (FOG) in Parkinson's patients. In FOG episode, the patient is unable to initiate, control or sustain a gait that consequently affects the Activities of Daily Livings (ADLs) and increases the occurrence of critical events such as falls. This paper presents continuous monitoring ADLs and classification freezing of gait episodes using Wi-Fi and radar imaging. The idea is to exploit the multi-resolution scalograms generated by channel state information (CSI) imprint and micro-Doppler signatures produced by reflected radar signal. A total of 120 volunteers took part in experimental campaign and were asked to perform different activities including walking fast, walking slow, voluntary stop, sitting down & stand up and freezing of gait. Two neural networks namely Autoencoder and a proposed enhanced Autoencoder were used classify ADLs and FOG episodes using data fusion process by combining the images acquired from both sensing techniques. The Autoencoder provided overall classification accuracy of ∼87% for combined datasets. The proposed algorithm provided significantly better results by presenting an overall accuracy of ∼98% using data fusion.**

Index Terms—**Radar sensing, Wi-Fi sensing, deep learning, FOG detection.**
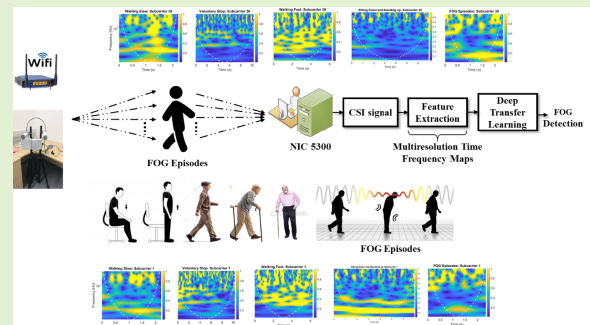
I. INTRODUCTION

PARKINSON'S disease (PD) is a progress and neurodegenerative condition causing motor impairments. There are more than 8.4 million Parkinson's patients around the globe and majority of the individuals are aged 50+ [1]. Parkinson's disease slowly grows and the efforts in managing the motor impairments lasts for longer period. Due to the severity of this

TABLE I
RADAR SENSOR DETAIL

Radar Model	SDR 580AD2
Waveform	FMCW
Operating Frequency	5.6 to 6.0 GHz
Bandwidth	400 MHz
Transmitting Power	20 dBm
Maximum Detectable Range	120 meters

disease, the patients' health get badly affected and because of enduring ailment, coerce them to get more frequent sick leaves and eventually it come to stage, where they have no choice but to consider early retirement from work. In this way, they get sufficient time to properly look after health maintain their health in more effective way and prevent any fatal and sudden disorders. keeping their health ample time to preserve usually take early retirement that eventually leads to loss in productivity along extreme healthcare and other societal expenditures.

In medical healthcare, one of the major motor related impairments in that present biggest challenge is freezing of gait (FOG) in Parkinson's patients. In FOG episode, the patient is unable to initiate, control or sustain a gait. The FOG events are usually confined to short period of time followed by regaining control and continue regular walking. Statistic

indicate that half of the Parkinson's patients experience these episodes twice a month and third experiences on day to day basis [2]. This results in limited mobility, deterioration in mental well-being and adverse effect on activities of daily livings and the patients are at high risk of falling causing fatal injuries [3]. Critical events such as falls make 20% to 30% of injuries among elderly people [4].

These signs make the occurrence of FOG episodes as serious health concerns for patients. It is thus highly important to deliver means for identifying and reducing the effect of FOG episodes. Rhythmic auditory stimulation (RAS), such as playing marching music and dance therapy, has been shown to be a safe, inexpensive, and effective method in resuming normal gait in PD patients [5].

The non-invasive wireless sensing leveraging wireless channel information (CSI) and micro-Doppler signatures in medical healthcare are two of the most promising solutions to detect the occurrence of FOG episodes in real-time and deliver feedback cue in an attempt to restore the normal gait alert the patient and reduce the risk of critical events such as falls. Providing timely cue to the Parkinson's disease patients has proved to decrease the severity of particular episodes [6] and reduce the overall duration of the episodes by 35% [7]. Several researchers have designed FOG classifiers based on either software devices. The software-based FOG detector [8] presents several constraints such as mapping machine learning algorithms on hardware. It is thus evident that the FOG detection in PD patients that provide hardware implementation is obtrusive and bulky to be worn on body. There are several case studies that have mapped real-time, complex hardware implementation of machine learning classifiers for wearable healthcare applications. The Field Programmable Gate Array (FPGA) based systems were proposed for heartbeat monitoring [9], [10], seizure episodes in epilepsy disease [11]–[13] and activities of daily living [14], [15]. In light of these advances in healthcare sector, this article presents a non-invasive, low-cost and easily deployable FOG detector based on the data fusion of scalogram obtained using Wi-Fi sensing and spectrograms acquired using radar sensor. A novel deep neural network, namely enhanced Autoencoder is proposed that classifies activities of daily living and detects FOG episodes with high accuracy.

The paper is organized as follows: Section II discusses the related work done on freezing of gait detection, section III explains the signal processing for Wi-Fi sensing, generating scalograms from CSI and producing spectrograms from received signals through radar sensor. Section III gives a brief introduction to the neural network used and discusses the proposed enhanced Autoencoder. Section IV provides the experimental setup and how data were acquired, section V provides details about results obtained and section VI concludes the paper.

II. RELATED WORK

Several researchers have addressed the application of software-based systems and wearable devices in healthcare sector to recognize human activities [16], detect inertial movements [17]. An extensive work has been done in research published work in [18], [19] that leverage wireless sensing

for medical healthcare application. In addition, several recent research articles have been published on patient monitoring, quantification of Parkinson's disease motor impairments using wearable sensors only. The main issues with these articles are the locations of sensors placed on bodies, the experimental setup and achieved accuracy. A comprehensive review of medical healthcare detection systems have been discussed in [20], [21]. There are several systems that exploit data mining and machine learning algorithms to determine the severity of particular disease. Arora *et al.* [19] recruited ten clinically diagnosed PD patients and ten healthy volunteers. Android smartphones were deployed on each subject's body and were asked to perform activities such as finger tapping, finger to nose test, and walk back-and-forth. The authors claim that this method could easily discriminate PD patients from healthy volunteers using machine learning algorithms, however due to the limited number of subjects and low classification accuracy, this solution is not suitable for FoG gait detection. Numerous wearable sensors have been for objective assessment of FoG episodes, yet there is small or no agreement with regard to location, number of participants, experiment setup and data-processing techniques. Stand-alone tri-axial accelerometer is widely used as in [22], magnetometer or combination of accelerometer and gyroscopes. Wearable sensors non-wearable sensors are used for cardiovascular activity, gait identification and activities of daily living [23]–[27].

The data-processing and classification algorithms are different in all aforementioned systems. Algorithms based on threshold are easily implemented and deliver adequate performance [28], however, these require optimum threshold levels to be tuned on all patients. Support vector machine (SVM), Naïve Bayes (NB), random forest (RF) are commonly machine learning algorithms used for classification tasks. NB and RF are less computional complex as compared to SVM algorithm, but the latter presents higher classification accuracy and robustness against a large amount of data. Capecci *et al.* [29] used smartphone worn on waist that executes gait test where accelerometer data is used leveraging threshold-based method, obtaining accuracy of 84.4%. Rodríguez *et al.* [30] used SVM algorithm with data obtained from 21 participants equipped with wearable sensor to determine FoG and provides specificity and sensitivity of 79.0% and 74.7%, respectively. In reference [31], the authors have used a model on convolution neural network using accelerometer data on 21 volunteers providing an accuracy of 92.3%.

III. PRELIMINARIES
A. Wireless Wi-Fi Signals & Channel State Information Extraction

The Wi-Fi signals driven by Orthogonal Frequency Division Multiplexing (OFDM) such as IEEE 802.11 a/a/ac can efficiently and effectively overcome the multipath frequency selective fading experienced due in an indoor environment. The OFDM frequency spectrum is split into several orthogonal frequency carriers where the data that is to be transmitted is encoded and mapped using same modulation scheme. The received signal at the receiving side is down-converted into the baseband signal. The sub frequency channels can be converted
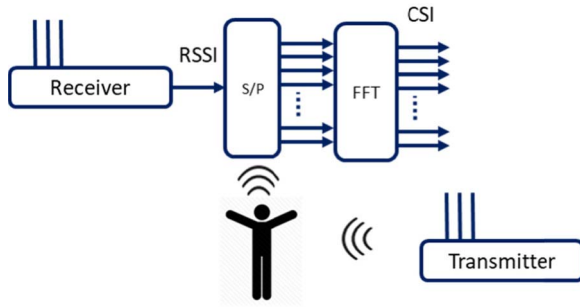
Fig. 1.   Frequency Carrier Conversion: Time Domain to Frequency Domain using RSSI/CSI by applying serial to parallel conversion and Fast Fourier Transform.

from time domain into frequency domain using serial-to-parallel signal converter and then applying Fast Fourier Transform (FFT) on all frequency channels, as shown in figure 1. The input serial data stream is formatted into the word size required for transmission, and shifted into a parallel format. The data is then transmitted in parallel by assigning each data word to one carrier in the transmission. The reason for converting serial into parallel is due to the fact that all subcarriers are transmitted at once in once burst rather than one after another, as in case for serial transmission. So that each frequency channel carry adequate information. The operating frequency of Wi-Fi router is set to 2.4 GHz.

The commodity devices such as Intel 5300 and Atheros ar5b225, are open-source device drivers that allow recording the channel state information of each frequency channel (carrier) representing the fine-grained physical layer channel measurements, consisting of wireless channel characteristics such as power distortion, multipath fading and shadowing effect.

Let H_i denote the channel state information values of frequency channel i, which is a complex value and is given as:

$$H_i = |H^i|e^{j\angle H_i} \tag{1}$$

Here $|H^i|$ and $\angle H_i$ denote the amplitude and phase information of ith subcarrier, respectively. The phase of individual subcarrier i, $\angle H_i$ is written as follows:

$$H_i = \angle \ddot{H}_i + (\lambda_p + \lambda_s)m_i + \lambda_c + \beta + Z, \tag{2}$$

where β indicates the initial phase offset of ith subcarrier of the phase-locked loop and mi is the subcarrier index of ith frequency channel. The environmental noise is represented in terms of Z, and λ_p, λ_s, and λ_c are the phase errors, sample frequency offset and central frequency offset, respectively. The CSI phase information is not adequate due to the random noise in the radio frequency channels because of using off-the-shelf Intel 5300 network interface. Hence, in this article, we have only considered the amplitude inform retrieved using channel state information.

B. Radar Technology & Micro-Doppler Signatures

The radar sensors use radio signals to identify and locate a target. The traditional radar sensor system consists of a trans-

mitter and receiver and a typical signal processing functional unit. The sensor when operating, continuously radiates electromagnetic signals and a target within area of interest reflects back the signal that is received by the radar system. Frequency modulated continuous wave (FMCW) radar uses transmission frequency that varies linearly across the waveform resulting in overcoming range-profile issues and is now widely used short-range wireless sensing applications, including activities of daily living, fall detection in healthcare and so on [23]. The FMCW radar operating at 5.8 GHz is extremely robust against interferences from RF signal sources and has the capacity to record human micro-Doppler signatures while maintaining high resolution. The specifications of an FMCW radar sensors used in this work is presented in table I. The main reason for using radar in the region of 5.6 – 6 GHz with operating frequency of 5.8 GHz (ISM, unlicensed band) is that the Wi-Fi transmitter was working at 2.4 GHz (ISM band, unlicensed band). In order to avoid co-channel interference, we opted for a Wi-Fi router and radar sensor that should work in an unlicensed band and should operate at different frequencies.

The RF signals transmitted by an FMCW radar sensor can be mathematically denoted as follows:

$$x(t) = \sum_{i=0}^{N-1} x^{(t)}(t - iT_F) \tag{3}$$

Here, T_F is the total duration of a frame as indicated in Figure 1, N_F represents the total number of transmitted frames. The transmitted FMCW signal comprising L number of chirps at the ith frame can be written as:

$$x_o = e(f_o t + \frac{\mu}{2}t^2) \quad \text{for } 0 \leq t \leq T \tag{4}$$

Here, f0 is the operating frequency, μ denotes a change of instantaneous frequency of an FCMW chirp signal. The value of μ can be determined bandwidth (B) divided by the duration of a frame (TF).

$$r_{l,k}^{(i)}(t)$$
$$= \sum_{m=1}^{M} a_m^{\sim(i)} x_0\left(\tau_m^{(i)}\right) e^{\left(2\pi f_{D,m}^{(i)} j \times (Tl+(i-1)T_F)\right)} e^{j\frac{2\pi}{\lambda}d_s k \sin\theta_m^{(i)}}$$
$$+ w_{l,k}^{(i)}(t), \tag{5}$$

In equation 6, $a_m^{\sim(i)}$ indicates the amplitude information in terms of complex values of the reflected signal of the mth target for i^{th} frame, d_s is the distance between the corresponding arrays, $\tau_m^{(i)}$, $\theta_m^{(i)}$, λ are the round-trip time delay, direction-of-arrival (DOA) and wavelength of the carrier frequency. The expression $f_{D,m}^{(i)}$ indicate the Doppler frequency induced due to the speed of the moving object (person in our case). In addition, $w_{l,k}^{(i)}(t)$ is the additive white gaussian noise (AWGN) signal for the kth array, the ith frame, and l^{th} chirp. Multiplying the conjugated components of FMCW transmitted signal $x_0^*(t)$ by $r_{l,k}^{(i)}(t)$ and considering $d_s = \frac{\lambda}{2}$, the beat signal $y_{l,k}^{(i)}(t)$ is described as the product of Doppler information, range profile and DOA and written as in equation 6: Here the expression $\left(e^{(-j2\pi f_{b,m}^{(i)}t)}\right)$, $\left(e^{j2\pi f_{D,m}(Tl+T_F(i-1))}\right)$, $\left(e^{j\pi \sin\theta_m^{(i)}}\right)$ and $w_{l,k}^{(i)}(t)$ represent the range, Doppler information, and direction-of-arrival information, respectively.

TABLE II
OPTIMIZED PARAMETERS FOR AUTOENCODER - (SCALOGRAMS/ WI-FI SENSING)

#	Width	Depth	Accuracy	#	Width	Depth	Accuracy
1	10	1	71.2	7	10-25-50-100	4	83.2
2	25	1	74.7	8	15-30-60-200	4	81.9
3	50	2	75.4	9	30-60-120-240	5	83.2
4	25-50	2	79.0	10	40-80-240-300	5	82.7
5	50-100	3	82.0	11	15-30-45-90-200-400	6	81.4
6	**25-50-100**	**3**	**84.5**	12	50-100-200-400-800	6	80.2

TABLE III
OPTIMIZED PARAMETERS FOR AUTOENCODER - (SPECTROGRAMS / RADAR SENSING)

#	Width	Depth	Accuracy	#	Width	Depth	Accuracy
1	10	1	73.5	7	10-25-50-100	4	84.1
2	25	1	77.2	8	15-30-60-200	4	85.1
3	50	2	81.4	9	30-60-120-240	5	82.7
4	25-50	2	83.7	10	40-80-240-300	5	83.5
5	50-100	3	82.0	11	15-30-45-90-200-400	6	83.4
6	**25-50-100**	**3**	**85.5**	12	50-100-200-400-800	6	84.1

IV. THE PROPOSED FOG DETECTION SYSTEM

The proposed freezing of gait detection system architecture is presented in figure 2 that resolves a multi-classification problem to identify the particular episodes and detect ADLs. The proposed system involves a low-cost commercially available 2.4 GHz WiFi router, a 5.8 GHz Ancortek radar sensor. Each body movement induces a unique CSI represented in terms of scalogram extracted from variances of amplitude information and a micro-Doppler signatures generated by received signal reflected by the subject's body. The multi-resolution time-frequency scalograms are obtained using continuous wavelet transform and micro-Doppler are acquired using short-time Fourier transform. The Wi-Fi-based and radar-based signatures are used image object to train and test the proposed enhanced classifier to determine the FOG episodes.

V. PRE-PROCESSING

This section discusses the pre-processing for CSI data obtained using Wi-Fi signals and micro-Doppler signatures received using the radar sensor. The channel state information is extracted from the packets retrieved from Internet Control Messages Protocol (ICMP) packets. Consequently, in theory, the total number of CSI packets received are the same as ICMP packets. However, it was observed that received CSI packets were slightly lesser than the number of ICMP transmitted packets. To synchronize and calibrate the frequency of the data collected, we performed the linear transformation on the raw CSI. The IEEE 802.11n Wi-Fi networks use multiple data sub frequency carriers or subcarrier that transmit Wi-Fi signals simultaneously exploiting orthogonal frequency division multiplexing. In principle, each subcarrier carrier independent information. However, each adjacent subcarriers carry similar information at times. For this purpose, we have used principal component analysis to obtain an independent

dataset for each human activity. The CSI data packets can be integrated together into multiple independent principle components.

A. Scalogram

Time-frequency scalograms are extracted from CSI amplitude information from each received packet and are used as features to detect FoG episodes. The extracted features are a multiresolution energy density function obtained from 5000 packets using Continuous Wavelet Transform (CWT). The energy density function $E(t, f)$ is obtained by squaring the amplitude information of the CWT function $C_d (t, f)$ of a discrete sequence. The time-frequency scalogram can be computed from the CWT $C_c(t,s)$ of a continuous-time signal $x(t)$, represented in terms of time t and scale s as in equation 7.

$$C_c(\text{t, s}) = \int_{-\infty}^{\infty} x(\text{v}) \frac{1}{\sqrt{s}} \psi(\frac{v-t}{s}) \text{dv} \qquad (6)$$

Here $\psi(v - \frac{t}{s})$ denote the dilation of the wavelet $\psi(t)$. We have scaled the term $v - t = \tau$ and scale the value of s as a function of frequency f provided $s = g_1(w) = g_2(f)$. The continuous wavelet transform of discrete and continuous signals are represented in equation 8 and 9 as follows:

$$C_d(\text{t, f}) = \text{T} \sum_k x(\text{t} + \text{kT}) \psi(\text{kT, f}) \text{d}\tau \qquad (7)$$

$$C_c(\text{t, f}) = \int_{-\infty}^{\infty} x(\text{t} + \text{kT}) \psi(\text{kT, f}) \text{d}\tau \qquad (8)$$

Here x(KT) is a discrete sequence of samples having a period T = 1/F, where F is the sampling frequency. The CWT of a discrete signal can be obtained when x(kT) is replaced with

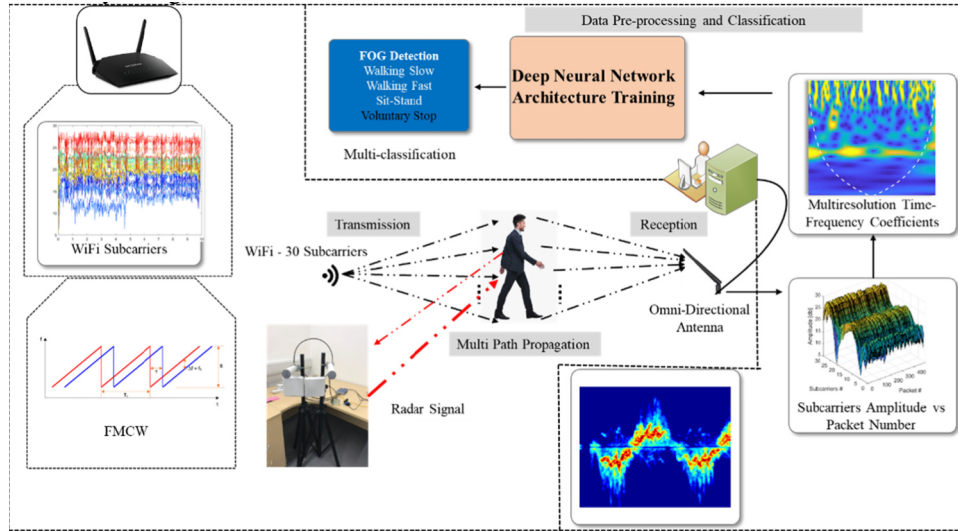| #	Width	Depth	Accuracy	#	Width	Depth	Accuracy
1	10	1	81.2	7	10-25-50-100	4	83.1
2	25	1	83.3	8	15-30-60-200	4	84.0
3	50	2	83.6	9	30-60-120-240	5	81.7
4	25-50	2	84.1	10	40-80-240-300	5	84.5
5	50-100	3	85.9	11	15-30-45-90-200-400	6	81.4
6	**25-50-100**	**3**	**87.1**	12	50-100-200-400-800	6	85.1

Fig. 2. Sensor fusion based architecture for FOG Detection.

$CSI^{SC}(kT)$, expressed in equation 10.

$$C_d(t, f) = T \sum_k CSI^{sc} x(t + kT) \psi(kT, f) d\tau \quad (9)$$

In equation 10, F = 60 Hz is the sampling frequency of CSI amplitude information and T = 0.02 seconds. The mother wavelet used in this work is the "morse" wavelet. The scalogram E(t,f) can further be described mathematically as follows:

$$E = C_d(t, f) \times C_d^*(t, f) \quad (10)$$

$$E = T^2 \sum_{k_1} \sum_{k_2} CSI^{sc} x(t + kT) CSI^{sc^*}(t + k_2T)$$

$$\times \psi(k_1T, f) \psi^*(k_2T, f) \quad (11)$$

The continuous wave transform-based scalograms deliver multiple resolution time-frequency analyses that are primarily ependent on the window size, resulting in various dilation of the mother wavelet. The scalograms provide adequate transient changes in the amplitude information of the CSI packets due to the human body movements and present high resolution as smaller window durations at higher frequency are used. Moreover, the scalograms have the potential to detect smooth features in waveforms due to larger windows durations at lower frequency. Figure 3(a) shows the variations in amplitude information for single-frequency carrier obtained using Wi-Fi sensing. A clear transition in each human activity can be seen, implying each body motion induces a unique imprint in terms of amplitude variation in dB. The scalograms presented in
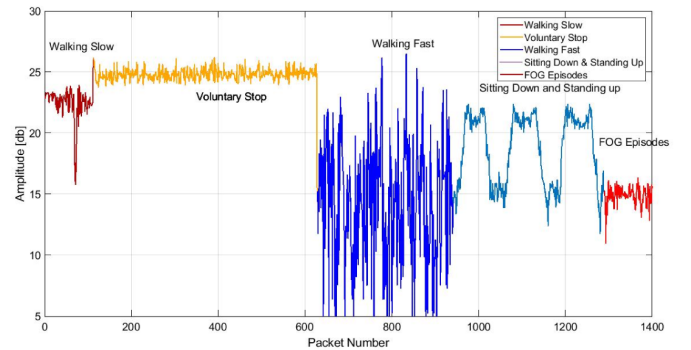

Fig. 3. Packet # for five human activities including FOG episodes.

figure 3(b) are plotted against the logarithmic scale (frequency domain). The white dotted line denotes the cone of influence that splits the region that where edge effects are significant to classify ADLs.

B. Micro-Doppler Signature

Any subject that is moving within radar range has mechanical rotation/vibration along with its bulk translation, it leads to the generation of frequency modulation on the reflected lectromagnetic signal that produces sidebands around the subject's Doppler frequency shift known as the micro-Doppler

effect [24], [25]. Any object in motion at distance in P moves with frequency f_v and displacement D_v, with function displacement value of, $D(t) = D_v sin2\pi f_v t cos\beta cos\alpha_p$. Let us assume R_0 to be the distance between the radar sensor and the moving target initial position *O,* the total range between the two will constantly change with respect to time, the target object's micro-movements are expressed as represented as $R(t) + D(t)$. The reflected signal received can be written as:

$$s(t) = \rho \left(e^{(2\pi f0t + 2\pi \frac{R(t)}{\lambda})} \right) = \left(\rho e^{(2\pi f_0 t + \phi(t))} j \right) \quad (12)$$

The value of f_0 in equation 7 is the carrier frequency, λ is its wavelength and ρ is the backscattering coefficient. Putting $R(t)$ in equation 7, the received signal can be mathematically described as follows:

$$s(t) = \rho \left(e^{\left(2\pi f0t + 2\pi \frac{R(t)}{\lambda} \right)} \right) = \left(\rho e^{(2\pi f_0 t + \phi(t))} j \right) \quad (13)$$

$$s(t) = \rho \left(e^{(2\pi f_0 t + 2\pi \frac{R(t)}{\lambda})} \right) \times e j 2\pi f_0 t$$
$$+ D_v \sin(w_v t) \cos\beta \cos\alpha_p \frac{2\pi}{\lambda} \quad (14)$$

Here $w_v = 2\pi f_v$, the derivative of the second phase component provides micro-Doppler shift expression which is expressed as:

$$s(t) = \rho \left(e^{(2\pi f_0 t + 2\pi \frac{R(t)}{\lambda})} \right) \times e j 2\pi f_0 t$$
$$+ D_v \sin(w_v t) \cos\beta \cos\alpha_p \frac{2\pi}{\lambda} \quad (15)$$

VI. AUTOENCODER AND PROPOSED ENHANCED AUTOENCODER FOR FOG DETECTION

One of the major challenges researchers face is the application of deep neural networks when classifying RF signals (Wi-Fi and Radar in our case) is the limited size of the available dataset. RF data acquisition is consumed a lot of time, is expensive and involves a large number of volunteers [26]. It is highly unlikely to record hundreds of thousands of Wi-Fi and radar data and hence novel algorithms have to be introduced so as to avoid underfitting and overfitting problems. In this context, a 14-layer deep convolution neural network architecture has been proposed to classify human gait classification. In this work, we consider a conventional Autoencoder and compare the results with our proposed three-layer enhanced Autoencoder.

A. Auto-Encoder

An Autoencoder neural network that reproduces the input values at the output side with specific limitations. For instance, given

$$z_i = \sigma(\hat{W} e_i + b^\sim) \quad (16)$$

The values of $W^\sim$ and $b^\sim$ show the weights and biases at the decoder side, respectively. During the unsupervised pretraining process, the neural network tried to reduce the reconstruction error to its minimum value as:

$$J(\theta) = \frac{1}{N} \sum_{i=1}^{N} (x_i - z_i)^2 \quad (17)$$

To prevent the Autoencoder neural network, cost function with a sparsity parameter is applied to force the network to learn the correlation among the given input data. With the addition of sparsity parameters, the cost function can thus be written as:

$$\text{gmin}_\theta J(\theta) = \frac{1}{N} \sum_{i=1}^{N} (x_i - z_i)^2 + \beta \sum_{i=1}^{N} KL(p || p_l) \quad (18)$$

where h denotes the number of hidden neurons, β is the sparsity proportion and KL describes Kullback-Leibler divergence and can be expressed as follows:

$$KL(p||p_l) = p \log\left(\frac{p}{p_j}\right) + (1 - p) \log\left(\frac{1 - p}{1 - p_j}\right) \quad (19)$$

where p_j represents the activation function for j^{th} hidden neuron and p denote the value for activation function, h is the number of hidden neurons. After pretraining the network, the decoder is removed from the network and encoder values are placed for training, using a supervised learning method by adding the SoftMax classifier with six neurons after the encoder. An input vector of K is fed as an input to the SoftMax function that primarily normalizes it into a probability distribution consisting of K probabilities, $P(y_k||x_i)$ for $k = 1, 2, \ldots K$, proportional to the exponentials of the input numbers [28], [29]. In probability distribution function, the input value x_i corresponds to labels class y_k. The probability class p_k can be mathematically expressed as:

$$P\left(y = k|x_i = \frac{e^{\theta_k x_i}}{\sum_{k=1}^{K} e^{\theta_k x_i}} \right) \quad (20)$$

an input matrix x, the network aims to estimate $h_w(x) \approx$ x. The unsupervised algorithm was introduced in order to initialize the weights and biases of an Autoencoder which was extremely efficient effective when a limited number of training data were available. The Autoencoder neural network implements unsupervised pretraining processing by encoding decoding the Input data, respectively. The Autoencoder approximates a nonlinear mapping on the input data matrix x as follows:

$$e_i = \sigma(Wx_i + b) \quad (21)$$

where σ describes the nonlinear activation function, *W* and *b* are the weights and biases at the encoder side, respectively. The features that are encoded in the network are then decoded in order to reconstruct the particular input matrix x using the following function [27].

The optimum values of weights and biases are obtained when the cost function is minimized as:

$$J(\theta) = \sum_{i=1}^{N} \sum_{k=1}^{K} \{y^i = k\} \log \frac{e^{\theta_k x_i}}{\sum_{k=1}^{K} e^{\theta_k x_i}} \quad (22)$$

The 'fine-tuning' technique that works on the gradient-based technique, is applied to resolve the values as indicated in
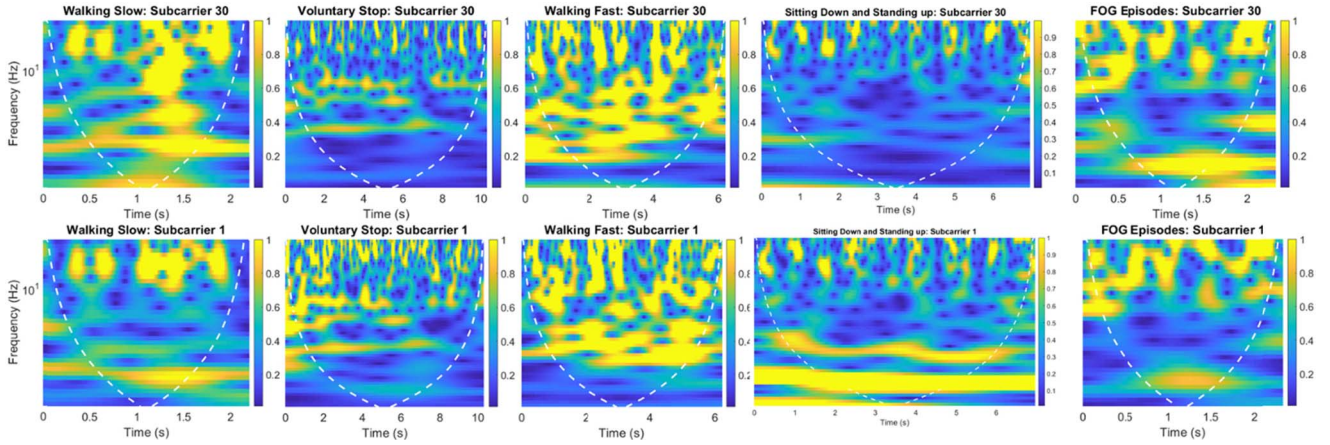
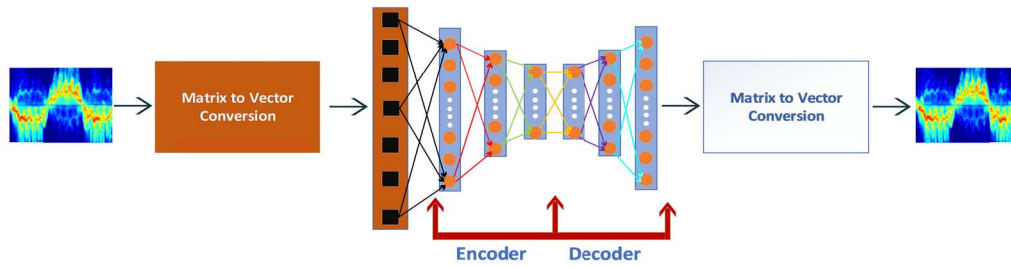Fig. 4. Scalgrams generated for different human activities.

Fig. 5. Architecture for autoencoder used for classifying spectrograms.

equation 22, where the neural network is trained by a supervised machine learning technique algorithm. The Autoencoder neural network architecture that processes the micro-Doppler signatures as images are shown in figure 3.

B. Proposed Enhanced Autoencoder

The enhanced Autoencoder integrates the advantages of convolutional filtering in convolutional neural networks having an unsupervised pretraining of an Autoencoder. On the contrary to the topology for an Autoencoder network, instead of the fully connected layers, the encoder of the proposed neural network consist of convolutional layers and the decoder side have deconvolutional layers. The deconvolutional filters are primarily the transposed copies of the convolutional layers as done in this work and are learned from scratch. Moreover, each and every deconvolutional filer is followed up by an unpooling filter [30]. The unpooling function is operated by saving the locations of the largest values when the pooling process occurs that inherently preserve the values while zeroing the remaining ones. The proposed system is ten times faster than conventional Autoencoder and can provide adequate results when limited number of data (observations) are available for training, specifically in the case of RF sensing system such as ours. The spatial locality of the neural network is preserved by accommodating a convolutional function at each neuron. Hence for a specific input value, P, the encoder calculates the values as follows:

$$e_i = \sigma \left(P * F^n + b \right) \tag{23}$$

Here σ indicates the values for activation function, $*$ is the 2-D convolution, F^n is the nth 2-D convolutional layer filter and b represents the basis of the encoder. The unsupervised pretraining process that is applied to the neural network aiming to reduce the subsequent following expression:

$$E(\theta) = \sum_{i=1}^{m} (x_i - z_i)^2 \tag{24}$$

After performing the unsupervised pretraining process, the decoding part is eliminated and the softmax classifier and fully connected layers are included at the end of the neural network. Next, the neural network is fine-tuned by optimization function as done in the convolution neural network. Similarly, the ADAM algorithm in conjunction with the ReLU activation function for optimization of the two fully connected layers having 150 hidden neurons each. The optimization of the hyperparameters of the proposed enhance Autoencoder was performed by grid search method and is discussed in subsequent section. The overall architecture of the proposed enhanced Autoencoder is presented in figure 6.

The deep network architecture was implemented using 14 hidden neurons. Selecting the optimum number of hidden neurons involved rigorous experimentations. Three methods were used namely fixed, constructive and destructive. In the fixed approach, a group of neural networks with different numbers of hidden neurons were trained and evaluated on the test available dataset using a different number of randomly selected starting weights. The increment in the number of hidden neurons one, two or more depending on the computational
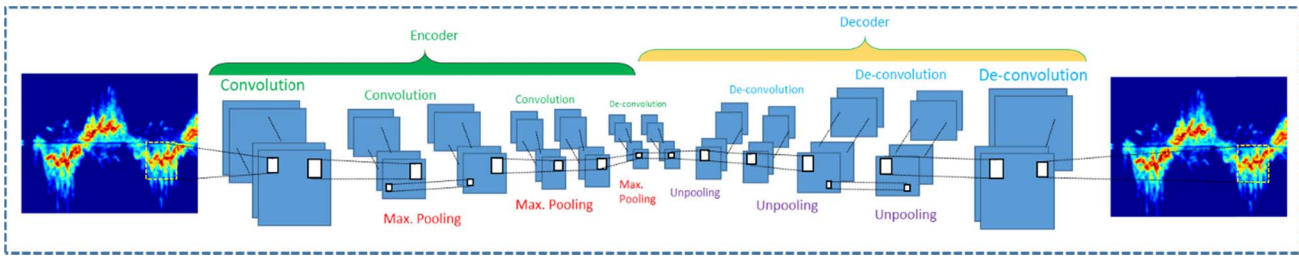
Fig. 6. Proposed enhanced Autoencoder architecture. Two layers namely convolutional and deconvolutional are shown. After the unsupervised pretraining process, the decoder section is eliminated where softmax and two fully connected layers are introduced terminal of encoder side.
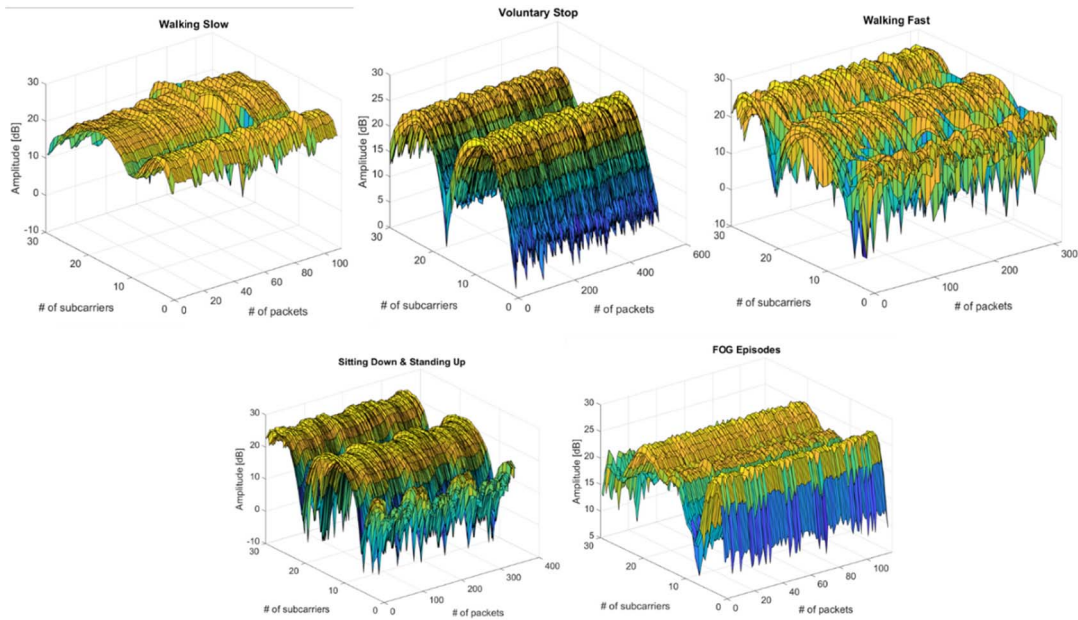

Fig. 7. Variances of amplitude information against subcarrier index and number of packets received over a period of time.

resources available. Plotting the evaluation criterion (e.g. sum of squared errors) on the test set as a function of the number of hidden neurons for each neural network generally produces a bowl-shaped error graph. The network with the least error found at the bottom of the bowl was selected because it was able to generalize best. This approach was time consuming, but generally worked very well. The constructive and destructive approaches involve changing the number of hidden neurons during training rather than creating separate networks each with a different number of hidden neurons, as in the fixed approach. The constructive approach involves adding hidden neurons until network performance starts deteriorating. The destructive approach is similar except that hidden neurons are removed during training.

Fig. 8. Experimental setup for data acquisition.

VII. EXPERIMENTAL SETTINGS AND DATA ACQUISITION

The data were recorded from two different sources, i.e. from radar sensors and Wi-Fi router. The experiments were performed at a large room at University of Glasgow as shown in figure 8. The radar sensor antennas were kept at a distance of 0.25 meter as shown in figure 8. All the activities were performed with aspect angel parallel to the radar sensor and Wi-Fi transmitter.

For training validation and testing, the data was acquired for 120 participants over the course of 20 days with age range from 30 to 76 years. The five human activities obtained were walking quickly, walking slowly, sitting down on chair, standing up from chair, voluntary stop and freezing of gait. Pair of human activities such as fast walk & slow walk, sitting down & standing up, voluntary stop & freezing episodes were purposely selected as they similar in pattern and would chal-
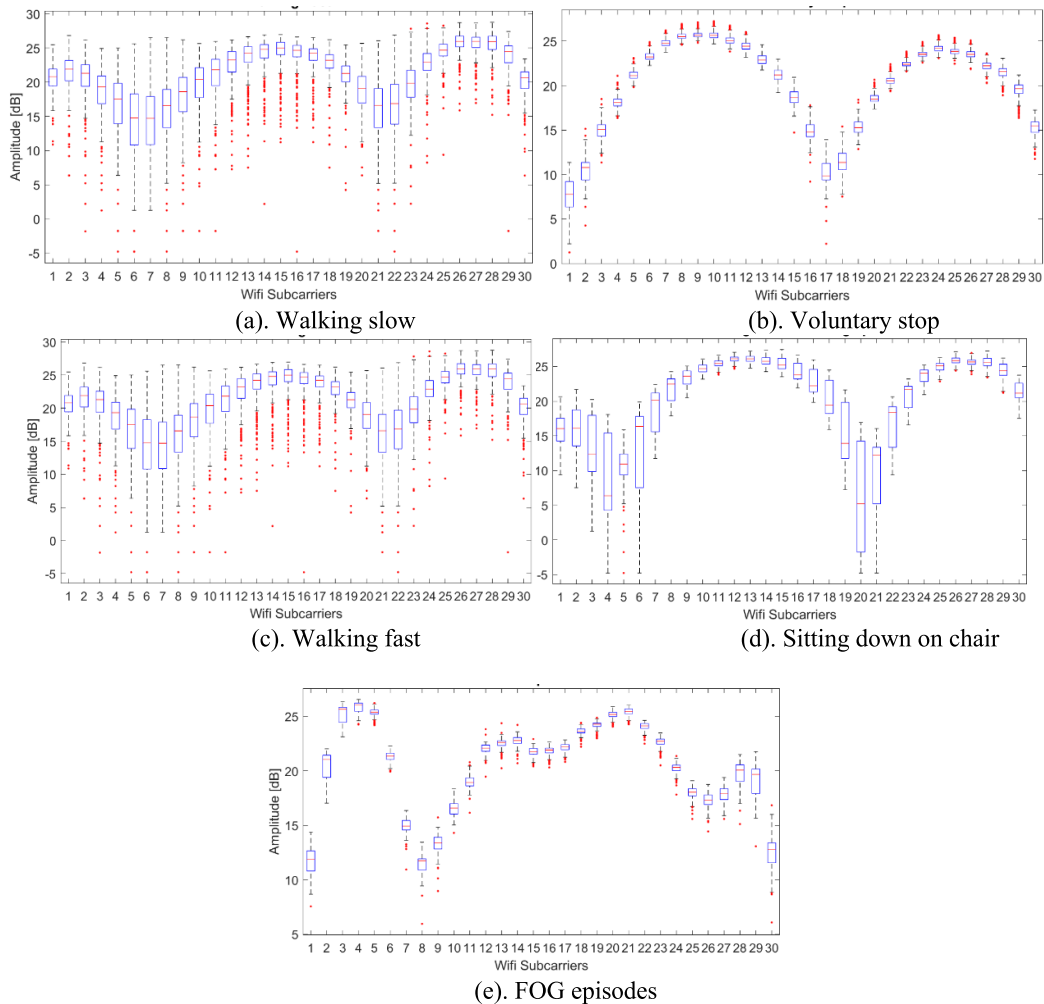
(a). Walking slow   (b). Voluntary stop

(c). Walking fast   (d). Sitting down on chair

(e). FOG episodes

Fig. 9. Box plot of CSI amplitude information to detect FOD episode.

lenge the classifier. Accuracy classification of FOG episodes is vital having minimum false alarms and low missed detection because inaccurate detection of the particular events can have extremely adverse effects on the patients. In this context, all of the 120 volunteers were asked to repeat a single activity more than six times, resulting in 4320 observations in total. The duration of each activity was 5 seconds with the exception of walking activity that was recorded for 10 seconds.

A. Data Acquisition Using Wi-Fi

The experimental design to detect FOG episodes in indoor settings (10 meters by 12 meters room at University of Glasgow, UK) is shown in figure 6. The transmitter, in this case, is a Wi-Fi router that operates at 2.4 GHz and is deployed 8 meters (in line-of-sight) away from the receiving antenna. The receiver is an Omni-directional antenna wired with the Intel 5300 network interface card installed in the PCIe slot of a Dell Inspiron desktop computer, Intel® Core™ i7-9700 Processor, 8GB RAM). The experimental procedure involved the acquisition of multiple data frequency carriers where 30 OFDM subcarriers were obtained from each CSI id. These packets contain the variances of amplitude and phase information of specific human activity which was obtained

using signal processing techniques and was performed in Matlab 2019 tool in order to get the time-frequency scalograms. Figure 7 shows the perturbations of amplitude information obtained for five human activities obtained from channel state information using Wi-Fi sensing. Figure 7(a) shows the box plot (amplitude level against group of 30 subcarriers) for person walking fast in indoor settings. Huge variation can be seen, where maximum fluctuation occurs between 12 dB and 25 db. As the person stops walking, the change in amplitude level decreases dramatically and slight variations can be observed as in figure 7(b). There was a sudden increase in variances of amplitude information when person start walking very fast as in figure 7(c). For sitting down on chair, the CSI signatures are distinct from rest of the four imprints as shown in figure 7(d). When the person was experiencing FOG episodes, implying that the feet were felt as if glued to the ground and upper body was trying to move. Hence slight variations can be seen in figure 7(e).

The channel state information measurements for activities of daily living and FOG episodes show distinguishable variances in the amplitude information against individual subcarriers as shown in figure 8. The box plot for CSI data against individuals subcarriers is shown in figure 9, indicating the first

TABLE V
OPTIMIZED PARAMETERS FOR PROPOSED AUTOENCODER (SCALOGRAMS/ WI-FI SENSING)

Filter Size	Width	Depth	Accuracy (%)	Width	Depth	Accuracy (%)
9 x 9 – 3x3	10	1	85.1	10	3	91.1
	20	1	86.0	20	3	90.4
	100	1	86.3	100	3	88.1
	10	2	90.8	10	4	91.0
	20	2	91.8	20	4	90.9
	100	**2**	**92.3**	100	4	90.1

TABLE VI
OPTIMIZED PARAMETERS FOR PROPOSED AUTOENCODER (SPECTROGRAM/ RADAR SENSING)

Filter Size	Width	Depth	Accuracy (%)	Width	Depth	Accuracy (%)
9 x 9 – 3x3	10	1	90.8	10	3	89.1
	20	1	91.2	20	3	91.2
	100	1	88.5	100	3	92.6
	10	2	91.0	10	4	93.0
	20	2	93.1	20	4	92.5
	100	**2**	**94.7**	100	4	93.1

TABLE VII
OPTIMIZED PARAMETERS FOR PROPOSED AUTOENCODER (DATA FUSION | RADAR SENSING + WI-FI SENSING)

Filter Size	Width	Depth	Accuracy (%)	Width	Depth	Accuracy (%)
9 x 9 – 3x3	10	1	93.3	10	3	910
	20	1	94.2	20	3	93.1
	100	1	91.8	100	3	94.8
	10	2	93.4	10	4	95.3
	20	2	95.0	20	4	91.0
	100	**2**	**98.1**	100	4	94.4

quartile $Q1\_CSI^{sc}$, median Md_CSI^{sc}, third quartile $Q3\_CSI^{sc}$, interquartile range IQR, minimum value min_CSI^{sc} and maximum values max_CSI^{sc} as shown in equations 25 to 29.

$$Q1_{CSI^{sc}} = |\mathrm{H_{sc}}(t + \frac{N+1}{4} \times T)| \qquad (25)$$

$$Md_{CSI^{sc}} = |\mathrm{H_{sc}}(t + \frac{N+1}{4} \times T)| \qquad (26)$$

$$Q3_{CSI^{sc}} = |\mathrm{H_{sc}}(t + 3 + \frac{N+1}{4} \times T)| \qquad (27)$$

$$IQR = Q3_{CSI^{sc}} - Q1_{CSI^{sc}} \qquad (28)$$

$$\mathrm{min\_CSIsc} = \mathrm{Q1\_CSI^{sc}} - 1.5(\mathrm{IQR}) \qquad (29)$$

The box plot as in figure 8, primarily presents the statistical information regarding the variances of amplitude information of CSI packets across all 30 subcarriers. It is evident that the box plot that statistical features for individual's frequency carrier including quartiles, interquartile ranges and median do not provide adequate information to distinguish activities of daily living and may be similar for specific subcarriers.

B. Data Acquisition Using Radar

The Ancortec radar sensor used to collect data transmitted signals with power approximately +20 dBm at 5. 8 GHz, having a bandwidth of 400 MHz bandwidth. Two Yaggi antennas were used as transmitter and receiver with gain equal to 10 dBi that had horizontal and vertical beamwidth of 60 degrees. The yaggi antenna is primarily manufactured conductors having high reflectivity potential. It has numerous applications due to its unidirectional pattern, high gain and broad bandwidth. It is designed by folding dipole directors where the dipole is electrically connected with the feeder. The dipole is used here is a resonant with the dipole length, that is half the wavelength (half of the wavelength of 5.8 GHz frequency range) and is fed to the feeder. The radiator is set to 5% less than the folded
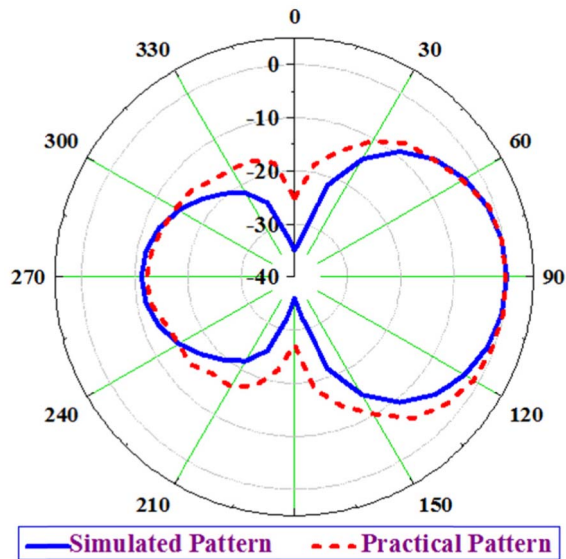
Fig. 10. Radiation pattern of yaggi antenna.

Fig. 11. Spectrograms obtained for human activities and FOG episodes.

dipole material and is deployed at the forefront to a distance of lower than quarter wavelength of operating waveform. The RF signals directors are designed lesser outwards in order to give yaggi antenna deployed at transmitter and receiver side taper in the radiation and reception direction that essentially makes the operating antenna extremely directional for detecting activities of daily living. The radiation pattern of yaggi antenna used in proposed system is presented in figure 10.

The radar sensor was powered by Universal Serial Bus (USB) cable since its power consumption limited to USB standards. In addition, by analyzing the life span and autonomy of the proposed system in real-world realistic deployment, it would always be connected to a computer for data acquisition and operational task, or to the electricity mains in indoor settings.

The FMCW radar sensor presents both a range profile and Doppler information. In this work, we have only considered the latter information due to computational purpose and ADL and FOD can be accurately identified merely using this information. For data acquisition, the radar was put on an in the same room where Wi-Fi system was deployed. The participants were asked to perform activities within a short range of 1-meter to about 3-meter range and the radar data were obtained with Wi-Fi data simultaneously. The two antennas were deployed in such a way that it would keep the torso of the volunteers in the center of radar beam enhance strength of reflected strength. The recorded data were processed using Short Time Fourier Transform (STFT) to get the spectrograms and produce micro-Doppler signatures.

The Range-Time-Intensity graph can be obtained by stacking up the reflected signals in a matrix format and applying Fast Fourier Transform (FFT) function along the fast-time direction to produce range cells of human action. Next, STFT algorithm is applied to the range profile that consists the target's, in our case people in action, to generate their micro-Doppler signatures [31]. The STFT algorithms apply
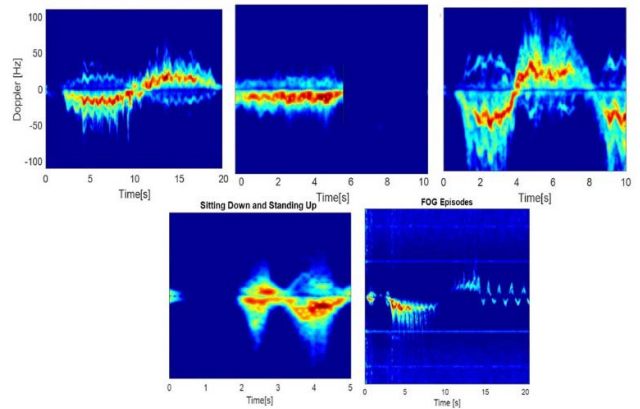
a sequence of FFTs algorithms with narrow, overlapping windows along the total duration of the collected data; the absolute squared value of the complex matrix is the known as the spectrogram, that is a plot of velocities of body parts in action (obtained through the Doppler effect) with respect to time. A notch MTI filter is applied to mitigate the contribution of static objects targets near 0 Hz such as walls, ceiling, floor, and furniture.

Figure 11 shows the micro-Doppler signatures activities of daily living and FOG episodes obtained at the room in University of Glasgow. The positive values of micro-Doppler components are movements towards the radar, while the negative values are the movements away.

VIII. RESULTS AND DISCUSSIONS

In this article, the neural network models are implemented in Python using Keras that exploits Tensorflow using tensor manipulation library. Both Autoencoder and proposed enhanced Autoencoder classification algorithms are tested using scalograms of Wi-Fi data and spectrograms of radar sensor leveraging tenfold cross-validation of both datasets as described in section earlier section. Both neural networks are trained for 300 epochs having a minibatch size of 80. The validation accuracy for both datasets is obtained by splitting 25% of the training data as the validation set and the evaluated the models after the completion of each iteration using the validation datasets.

The adaptive moment estimation algorithm is used to optimize the pretraining process and for fine-tuning having a learning rate of 0.001. Leveraging grid search technique, we have identified the optimum values for width and depth without overfitting as described in Table II (Wi-Fi sensing), Table III (radar sensor) and Table IV (data fusion; combing scalograms and spectrograms). We have opted to use a three-layer Autoencoder having layers of 100, 50 and 25, respectively. The best classification accuracy with optimized parameters are shown in red.

For all three scenarios (Wi-Fi sensing, radar sensing and data fusion: combined datasets), the best classification performance was obtained for 100,50 and 25 layers. The Wi-Fi sensing based on scalograms classification provided an

accuracy of 84.5%, the spectrogram classification for radar sensing delivered an accuracy of 85.5.%. When the two datasets were combined, the accuracy of 87.1% was achieved where a small improvement was observed.

We further examine the optimization of hyperparameters of proposed enhanced Autoencoder was performed using grid search, as presented in Table V to Table VII. The enhanced Autoencoder was implemented with three convolutional layers. Both convolutional and deconvolution layers were populated with 30 numbers of 9 by 9 and 3 by 3 concatenated filters. The best classification accuracy using scalograms with Wi-Fi Sensing, 92.3% was obtained with width of 100, depth of 2, having filter size of 9 by 9 and 3 by 3. In addition, the classification accuracy of radar sensing, 94.7%, was slightly better than Wi-Fi sensing as shown in Table VI. When the same network configurations were used, and the combined datasets of both devices were used, a significant change in the performance was observed. The accuracy of the system increased to 98.1% when classification all human activities. Table VII shows the highest classification accuracy (in red) obtained for width of 100 and depth of 2.

The confusion matrix of data fusion for ADLs recognition (walking slow, voluntary stop, walking fast, sitting down and standing up, and FOG episodes) and identifying FOG episodes for Autoencoder and proposed enhanced Autoencoder are shown in Table VI and Table VII, respectively. In Table VI, the Autoencoder provided adequate classification accuracy. However, similar activities such as walking fast & walking slow, voluntary stop and FOG episodes were deliberately chosen to confuse the classifier, consequently challenging its performance. For instance, the walking fast was classified with accuracy of around 90%, however it slightly misclassified with walking slow; indicating an error of ∼10%. A similar pattern was observed for another pair of activities namely voluntary stop and FOG episodes. The particular freezing of gait event was identified with an accuracy of 85% but the algorithm detected false negative rate of 14% for voluntary stop.

A significant improvement in classification accuracy of individual activities can be seen in Table VII. The proposed algorithm detected the walking fast, sitting down and standing up activities with 100% accuracy. The two similar pair of activities, that were presented as to confuse the proposed algorithm, were also classified with accuracies of ∼99%

IX. CONCLUSION

This paper presented freezing of gait detection Parkinson's disease patients that primarily causes motor impairments leading to high risks of falls. In this context, activities of daily living namely walking fast, walking slowly, voluntary stop, sitting down on chair/standing up from a chair and freezing of gait episodes were monitored using two non-invasive sensing technique. The two methods for data acquisition were Wi-Fi-based sensing and radar-based sensing techniques. The multi-resolution scalograms were produced from CSI data and spectrograms were obtained from received signal using radar sensor. An extensive experimental campaign was conducted involving more than 100 volunteers with age range from 30 to 76 years. For classification purpose, two deep neural networks

such as an Autoencoder and proposed convolutional neural network based enhanced Autoencoder were used to classify activities of daily and detect FOG events. It was observed that the Autoencoder delivered overall classification accuracy of ∼87% for combined datasets, however the proposed algorithm outperformed the same classifier providing classification accuracy of 98.1% for the same datasets.

REFERENCES

[1] T. Pringsheim, N. Jette, A. Frolkis, and T. D. L. Steeves, "The prevalence of Parkinson's disease: A systematic review and meta-analysis: PD PREVALENCE," *Movement Disorders*, vol. 29, no. 13, pp. 1583–1590, Nov. 2014.

[2] I. Witt, H. Ganjavi, and P. MacDonald, "Relationship between freezing of gait and anxiety in Parkinson's disease patients: A systemic literature review," *Parkinson's Disease*, vol. 2019, pp. 1–24, Jul. 2019.

[3] S. A. Shah, X. Yang, and Q. H. Abbasi, "Cognitive health care system and its application in pill–rolling assessment," *Int. J. Numer. Model., Electron. Netw., Devices Fields*, vol. 32, no. 6, Jun. 2019, Art. no. e2632.

[4] W. Saadeh, S. A. Butt, and M. A. B. Altaf, "A patient-specific single sensor IoT-based wearable fall prediction and detection system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 5, pp. 995–1003, May 2019.

[5] A. Ashoori, D. M. Eagleman, and J. Jankovic, "Effects of auditory rhythm and music on gait disturbances in Parkinson's disease," *Frontiers Neurol.*, vol. 6, p. 234, Nov. 2015.

[6] L. Rosenthal, D. Sweeney, A.-L. Cunnington, L. R. Quinlan, and G. Ólaighin, "Sensory electrical stimulation cueing may reduce freezing of gait episodes in Parkinson's disease," *J. Healthcare Eng.*, vol. 2018, pp. 1–6, Aug. 2018.

[7] C. Punin *et al.*, "A non-invasive medical device for Parkinson's patients with episodes of freezing of gait," *Sensors*, vol. 19, no. 3, p. 737, Feb. 2019.

[8] L. Borzì *et al.*, "Home monitoring of motor fluctuations in Parkinson's disease patients," *J. Reliab. Intell. Environ.*, vol. 5, no. 3, pp. 145–162, 2019.

[9] M. Bhosale and A. Chitre, "Real time ECG acquisition and FPGA based QRS detection," in *Proc. 4th Int. Conf. Comput. Commun. Control Autom. (ICCUBEA)*, Aug. 2018, pp. 1–4.

[10] S. M. Abubakar, W. Saadeh, and M. A. B. Altaf, "A wearable long-term single-lead ECG processor for early detection of cardiac arrhythmia," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, Mar. 2018, pp. 961–966.

[11] H. G. Daoud, A. M. Abdelhameed, and M. Bayoumi, "FPGA implementation of high accuracy automatic epileptic seizure detection system," in *Proc. IEEE 61st Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2018, pp. 407–410.

[12] L. Liu, S. Shah, G. Zhao, and X. Yang, "Respiration symptoms monitoring in body area networks," *Appl. Sci.*, vol. 8, no. 4, p. 568, Apr. 2018.

[13] P. S. Thorbole, S. D. Kalbhor, V. K. Harpale, and V. K. Bairagi, "Hardware implementation of genetic algorithm for epileptic seizure detection and prediction," in *Proc. Int. Conf. Comput., Commun., Control Autom. (ICCUBEA)*, Aug. 2017, pp. 1–5.

[14] N. B. Gaikwad, V. Tiwari, A. Keskar, and N. C. Shivaprakash, "Efficient FPGA implementation of multilayer perceptron for real-time human activity classification," *IEEE Access*, vol. 7, pp. 26696–26706, 2019.

[15] P. Lei, J. Liang, Z. Guan, J. Wang, and T. Zheng, "Acceleration of FPGA based convolutional neural network for human activity classification using millimeter-wave radar," *IEEE Access*, vol. 7, pp. 88917–88926, 2019.

[16] X. Yang, D. Fan, A. Ren, N. Zhao, and M. Alam, "5G-based user-centric sensing at C-Band," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 3040–3047, May 2019.

[17] X. Wang, X. Wang, and S. Mao, "RF sensing in the Internet of Things: A general deep learning framework," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 62–67, Sep. 2018.

[18] X. Yang *et al.*, "Non-contact early warning of shaking palsy," *IEEE J. Transl. Eng. Health Med.*, vol. 7, no. 3, pp. 1–10, Mar. 2019.

[19] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, Apr. 2015, pp. 837–846.

[20] X. Yang *et al.*, "Sleep apnea syndrome sensing at C-band," *IEEE J. Transl. Eng. Health Med.*, vol. 6, no. 2, pp. 1–8, Jun. 2018.

[21] X. Yang *et al.*, "Wandering pattern sensing at S-band," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 6, pp. 1863–1870, Nov. 2018.

[22] S. Shah *et al.*, "Internet of Things for sensing: A case study in the healthcare system," *Appl. Sci.*, vol. 8, no. 4, p. 508, Mar. 2018.

[23] S. A. Shah and F. Fioranelli, "RF sensing technologies for assisted daily living in healthcare: A comprehensive review," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 34, no. 11, pp. 26–44, Nov. 2019.

[24] M. Friesen, "A review of emerging smart RFID in," Vancouver, BC, Canada, Tech. Rep., Oct. 2014.

[25] H. Lim and J. H. Park, "Joint time-frequency analysis of radar engine models," *IEEE J. Healthcare Technol.*, vol. 25, pp. 1069–1080, Jan. 2011.

[26] S. Tanoli *et al.*, "An experimental channel capacity analysis of cooperative networks using universal software radio peripheral (USRP)," *Sustainability*, vol. 10, no. 6, p. 1983, Jun. 2018.

[27] S. A. Shah, D. Fan, A. Ren, N. Zhao, X. Yang, and S. A. K. Tanoli, "Seizure episodes detection via smart medical sensing system," *J. Ambient Intell. Humaniz. Comput.*, pp. 1–13, 2018.

[28] D. Haider, S. A. Shah, S. I. Shah, and U. Iftikhar, "MIMO network and the alamouti, STBC (space time block coding)," *Amer. J. Elect. Electron. Eng.*, vol. 5, no. 1, pp. 23–27, 2017.

[29] K. Iwata, "Extending the peak bandwidth of parameters for softmax selection in reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 8, pp. 1865–1877, Aug. 2017.

[30] Q. Zhang, D. Haider, W. Wang, S. Shah, X. Yang, and Q. Abbasi, "Chronic obstructive pulmonary disease warning in the approximate ward environment," *Appl. Sci.*, vol. 8, no. 10, p. 1915, Oct. 2018.

[31] A. Zahid *et al.*, "Machine learning driven non-invasive approach of water content estimation in living plant leaves using terahertz waves," *Plant Methods*, vol. 15, no. 1, pp. 1–13, Dec. 2019.