

Received August 22, 2020, accepted September 1, 2020, date of publication September 10, 2020, date of current version September 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3022971

Sensor Location Analysis and Minimal Deployment for Fall Detection System

HIRAM PONCE¹, (Member, IEEE), LOURDES MARTÍNEZ-VILLASEÑOR, (Member, IEEE), AND JOSÉ NUÑEZ-MARTÍNEZ

Facultad de Ingeniería, Universidad Panamericana, Ciudad de Mexico 03920, Mexico

Corresponding author: Hiram Ponce (hponce@up.edu.mx)

This work was supported by the Universidad Panamericana through the grant “Fomento a la Investigación UP 2018,” under project code UP-CI-2018-ING-MX-04.

ABSTRACT Human falls are considered as an important health problem worldwide. Fall detection systems can alert when a fall occurs reducing the time in which a person obtains medical attention. In this regard, there are different approaches to design fall detection systems, such as wearable sensors, ambient sensors, vision devices, and more recently multimodal approaches. However, these systems depend on the types of devices selected for data acquisition, the location in which these devices are placed, and how fall detection is done. Previously, we have created a multimodal dataset namely UP-Fall Detection and we developed a fall detection system. But the latter cannot be applied on realistic conditions due to a lack of proper selection of minimal sensors. In this work, we propose a methodological analysis to determine the minimal number of sensors required for developing an accurate fall detection system, using the UP-Fall Detection dataset. Specifically, we analyze five wearable sensors and two camera viewpoints separately. After that, we combine them in a feature level to evaluate and select the most suitable single or combined sources of information. From this analysis we found that a wearable sensor at the waist and a lateral viewpoint from a camera exhibits 98.72% of accuracy (intra-subject). At the end, we present a case study on the usage of the analysis results to deploy a minimal-sensor based fall detection system which finally reports 87.56% of accuracy (inter-subject).

INDEX TERMS Ambient assisted living, fall detection, health monitoring system, human activity recognition, machine learning, sensor fusion.

I. INTRODUCTION

Falls are considered as an important health problem worldwide [1]. Fall detection systems can alert when a fall occurs reducing the time in which a person obtains medical attention. Patients, especially elders, often remain laying on the floor worsening the psychological and physical harm caused by the fall. This problem has gained a lot of attention from the research community since the phenomenon of population ageing is occurring around the world [2]. Fall detection systems can help to provide rapid provision of assistance when fall occurs [3].

Fall detection systems have been developed with wearable sensors, ambient sensors and vision devices [5] using one of these three approaches or, recently, a combination of

multimodal devices. Each of these modalities have advantages and drawbacks. Wearable sensors are portable, affordable, easy to install, reliable, robust, and can be embedded in commonly used devices as smart phones and smart watches [6]. They keep on tracking the subject's activities in different places. The main disadvantages are that sensors have to be worn sometimes for a long time and they are obtrusive. Vision devices do not have to be worn (except for motion capture sensors), but they have a limited range of vision. Multi-camera approaches can solve this problem with the drawback of increasing the cost. Cameras have also privacy issues, occlusion, and the quality of data is sensitive to environmental changes [6], [7].

Therefore, fall detection systems vary, first of all, in the types of sensors selected for data acquisition, the location in which the sensors are placed, and how detection is done. There are two main classification approaches for fall

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo¹.

detection: threshold based or machine learning based. Fall detection systems can be very different also depending on which activities and types of fall are selecting for experimentation and data.

Studies in human activity recognition commonly focus on the usage of one sensor modality, the features extracted from the data, and the classifiers that are sometimes ineffective when dealing with complex activities [8]. Works on fusion strategies have recently proposed for effective activity recognition [8] and fall detection [9]. Yang and Hu [10] presented the concept of multi-sensor fusion and grouped it into three levels of fusion: data, feature and decision level. According to the authors [10], direct data fusion is useful to overcome limitations of each single elements when poor selective sensors are used. Fusion at feature level entails the integration of feature sets corresponding to different sensors. Features are extracted from sensor data to integrate multi-dimensional feature vectors from which classification is made. Decision level fusion is based on the decisions of multiple sensors to improve detection accuracy.

In a previous work [11], we developed a multimodal fall detection system based on machine learning models. For this system, we first collected an extensive dataset, namely UP-Fall Detection, from 17 subjects performing different falls and activities of daily living (ADLs). Then, we analyzed different combinations of sensors (modalities) to perform a fall detection classifier. We achieved an F-score metric of 70.44% using five inertial measurement units (IMUs), one brainwave sensor and two RGB-cameras. Context sensors did not influence in the detection process. If we implemented this system in real-world, it would require many sensors. Thus, a proper selection of the minimum number of sensors is important to deploy a realistic system. In addition, in our previous work, the combination of sensors was made by group type. It prevents us to decide which sensors within the groups are better for our fall detection system.

Thus, this work aims to determine the minimal number of sensors required for developing an accurate fall detection system, using the UP-Fall Detection dataset [11]. In order to accomplish this goal, we analyze in this work the five IMUs and the two RGB cameras for possible inclusion into our system. First, we study each IMU and camera separately. Then, we combine one IMU and one camera in a feature level. We evaluate the performance of those combinations to finally select the minimal number of sensors. At the end, we present a case study on the preliminary results of deployment a fall detection system using an inter-patient scheme. It is worth-noting that context sensors are not considered because they did not influence the performance of our previous analysis.

It is important to reduce the number of sensors in order to simplify the system to decrease the obtrusiveness and facilitate ergonomic use specially among elders. Furthermore, a multimodal sensors approach is needed with the aim of improving reliability and precision of fall

detection systems. Regarding doppler radar and Kinect sensors, Cippitelli *et al.* [12] say that finding how to select the best type of sensor and its location for a specific scenario remains an open research question. Establishing the best location to deploy the radar sensor to avoid the attenuation of the radar Doppler signature is still an issue. In addition, the more suitable a wearable sensor placement is known, the better design of fall detection systems and experiments is, as noticed in [13]. Thus, an optimum combination of multimodal sensors in the optimal placement improves fall detection systems efficiency and precision.

Although several works reviewed in [8] performed data fusion and multiple classifiers systems for activity recognition, they try to solve the issue of sensor placement with various methods for example Kalman filtering. Nevertheless, very few works can be found with the aim of determining the best sensor placement for fall detection solutions. For instance, the work in [13] is based only on one sensor, one tri-axial sensor (accelerometer, gyroscope and magnetometer), that is placed in different parts of the body (right wrist, right thigh, right ankle, chest, waist and head). Thus, an improved analysis on sensor selection and placement is necessary for better understanding the performance of fall detection systems with minimal sources of information.

From early reviews of fall detection systems like Noury *et al.* [4] to more recent surveys namely Igual *et al.* [3], researchers point out the need of focusing fall detection systems in real conditions deployment. Most of the works related to fall detection have good performance when developing technology and testing in datasets build in laboratory conditions. Nevertheless, there is little use in daily geriatric practice for this systems due to the high number of false alarms caused by uncontrolled factors of real-world scenarios [3]. For this reason, an extra effort must be made to prove our model in a practical app that could be used in real scenarios.

The contributions of this paper to address all the above mentioned issues are:

- A methodological analysis to determine the minimal number of sensors required and their best placement for developing a robust fall detection system.
- A mobile application deployment that can be used in real scenarios built with the selected minimal sensors.
- A data set and online testing to evaluate the classification performance through the inter-patient paradigm in realistic conditions.

The paper is organized in the following parts. Section II gives an overview of practices in multimodal sensor location for fall detection. Section III describes the UP-Fall Detection dataset employed and the comparative analysis implemented for this study. Section IV presents the experiments and results of the comparative analysis. Then, Section V shows a preliminary case study on the deployment of a fall detection system using the results obtained in this work. Section VI discusses the comparative analysis and the findings of the case study. Finally, we conclude the paper.

TABLE 1. Related work comparing performance of wearable sensor placement for fall detection.

Ref.	Sensors	Location	Combination	Falls, ADL	Algorithms	Best performance	Conclusion
[14]	Accelerometer from smartphone	Trousers pocket, belt	Single	10 falls, 9 non-fall scenarios	DT, KNN, NB	Sensitivity belt 97%, specificity pocket 98.6%, accuracy belt 97.6%	Very similar performance in both locations
[15]	Accelerometer from smartphone	Left ankle, right ankle, chest, waist	Single	1 fall, 7 ADLs	FSM, SVM	Accuracy chest 98.23%	Best accuracy in chest
[13]	6 MTw Sensors	Head, chest, waist, right wrist, right thigh, right ankle	31	20 falls, 16 ADLs	KNN, BDM, SVM, LSM, DTW, ANN	Waist accuracy 99.87%, sensitivity 99.96%, specificity 99.76%	Best performance in waist
[16]	6 MTw Sensors	Head, chest, waist, right wrist, right thigh, right ankle	31	20 falls, 16 ADLs	KNN, J48-DT, SVM, RF, RC	Waist accuracy 99.28%, thigh accuracy 99.48%	Best performance in waist and thigh
[17]	Smartphone 4 wearable sensors	Thigh pocket, chest, waist, right wrist, ankle	31	3 falls, 11 ADLs	SVM, KNN, NB, DT	Sensitivity 95%, specificity 95%	Best performance waist and chest single or combined
[18]	Smartphone 4 wearable sensors	Thigh pocket, chest, waist, right wrist, ankle	31	3 falls, 7 ADLs	4 threshold algorithms	Sensitivity 60%, specificity 95%	Best performance wrist and wrist/waist combination

DT: decision trees, KNN: k -nearest neighbors, NB: naïve Bayes, FSM: finite state machine, BDM: Bayesian decision making, LSM: least squares method, DTW: dynamic time warping, ANN: artificial neural networks, SVM: support vector machines, RF: random forest, RC: random committee.

II. MULTIMODAL SENSOR LOCATION FOR FALL DETECTION

This section presents a review of the related work focusing on defining the best sensor position for fall detection. This analysis describes approaches using wearable sensors and camera viewpoints for fall detection. There are many works in fall detection, but we only include those that perform an explicit analysis for selecting the best sensor combination and/or placement.

A. LOCATION ANALYSIS OF WEARABLE SENSORS FOR FALL DETECTION

We discuss the related work that aim to determine the best placement for wearable sensors. Table 1 summarizes the following description.

A comparison of wearable sensor placement on the human body for fall detection is discussed in [13]. The author focused on sensor positioning analyzing 31 combinations from locations like in right wrist, right thigh, right ankle, chest, waist and head. Fourteen subjects performed 20 simulated falls and 16 ADLs. From these exhaustive combination experiments, the waist location using a single sensor was found to be suited placement. A similar analysis based on single positions was presented by [16] using the same dataset [13]. In this case, the authors implemented different machine learning classifiers: J48 decision trees (DT), k -nearest neighbors (KNN), random forest (RF), random committee (RC) and support vector machines (SVM). They found waist and thigh positions as the best locations. Using the UMAFall dataset [18], [19], Santoyo *et al.* [17] presented a systematic assessment evaluating the importance of location of five sensors for fall detection. They compared the performance of combinations of these sensors using SVM, KNN, naïve Bayes (NB) and DT learning models. They concluded that location of wearable sensors is more influential than the type of classifier implemented. The best result was obtained when sensors were placed on chest, waist, or these two locations combined [18].

Aguiar *et al.* [14] presented a fall detection system based on the accelerometer embedded on a smart-phone located in trousers' front pocket (side is not determined) and the belt. They compared sensitivity, specificity and accuracy metrics using three algorithms: DT, KNN and NB. Their results showed no significant difference in performance regarding both smart-phone locations. Kau *et al.* [15] proposed a fall detection system based also on the accelerometer of a smart-phone using a cascade classification architecture. Regarding the location, this work compared the performance of the system in terms of accuracy, detection rate and false alarm rate of four different positions namely left ankle, right ankle, chest and waist. Five subjects performed one type of fall and nine ADL. Although other metrics like sensitivity and specificity were used, the only conclusion regarding the accelerometer position was that chest location obtained the best accuracy for fall detection in their experiments.

As noticed in Table 1, fall detection performance depends on sensor location and selection of machine learning models, but there is not a clear pattern on how to select them properly. In this regard, the selection of sensor placements might depend on the configuration of each fall detection system.

In a related survey, Rucco *et al.* [20] provided a literature review regarding the sensor technologies for fall risk assessment, fall prevention and fall detection. They present an overview of the type and location of wearable sensors for the monitoring and assessment of falls during static and dynamic tasks. Their analysis shows that the methodologies in general consider two sensors at maximum and accelerometers and gyroscopes are the most commonly used. They found that the trunk is the most used for positioning the sensors for dynamic stability.

B. CAMERA PLACEMENT FOR FALL DETECTION

According to [21], fall detection systems based on a single RGB camera are often viewpoint-dependent. A new dataset is needed when the camera is moved to a different viewpoint and in particular at different height. Therefore, different camera

viewpoints in a dataset collection can help identify whether a given algorithm have the viewpoint-independent property. This arises on the necessity of reliable camera-based fall detection systems regardless the position of the subject during a fall.

The research efforts found are more in the sense of achieving the cooperation of multiple cameras to make the system robust and to avoid occlusion [11], [22]. Only a few vision-based approaches for fall detection address, in some way, the issue of camera placement [23]–[25], but none of them show results that could be compared.

A recurrent neural network with long short term memory architecture that models the temporal dynamics of the 2D pose information of a fallen person was developed by Hasan *et al.* [26] in order to address the environmental problems of visual based approaches. Huang *et al.* [27] also proposed a novel 2D video-based fall detection pipeline with human pose estimation as the method of feature augmentation. They used OpenPose to extract the coordination of human body keypoints in raw RGB data which are input of a convolutional network for feature extraction. Binary classification is afterwards achieved with high sensitivity and specificity performance.

In [23], the authors discussed that a monocular computer vision system is efficient only if the camera is placed sideways. Occlusion is the most common reason for failure in this kind of system. In order to minimize occlusion with furniture, Nait-Charif *et al.* [28] set-up ceiling mounted wide-angle cameras. They acquired information tracking a person in two semantic zones for two days, with different light conditions. Mounting the cameras in the ceiling avoided occlusion, but they were not able to capture the vertical motion as when a camera is sideways, which it is very useful for fall detection.

A multi-view approach is presented in [24]. In this study, posture classification for fall detection was performed merging the decision provided by independently processing cameras. Through theoretical analysis, the authors determined the minimum number of cameras and their placement for detecting falls. Authors stated that placing two cameras in orthogonal viewing direction to the motion provides a 100% true negative detection rate.

A statistical method for fall detection based on Kinect cameras was proposed in [25]. In this work, authors combined viewpoint-invariance collecting all training data from one viewpoint and all test data were collected from another point of view with the purpose of evaluating the robustness of the system to camera displacements.

Similarly to sensor location, the viewpoint dependence of cameras is highly related to the setup of the fall detection system. Thus, a case-by-case basis analysis for best location and selection of sources of information is required for fall detection.

III. ANALYSIS FOR SENSOR LOCATION

This section presents the methodology for the analysis of sensor selection and placement in our fall detection system.

First, we present the UP-Fall Detection dataset to understand the configuration setup of our system. Then, we describe the methodology for the analysis on selecting the minimum number of sources of information in our system.

This study was approved and regulated by the Research Committee of the School in Engineering at Universidad Panamericana (Mexico). All the subjects in the study filled out an agreement considering the regulations and data policies pertinent. During the whole procedure, the subjects retained their right to participate voluntarily in the experiments.

A. UP-FALL DETECTION SYSTEM CONFIGURATION

In our previous work [11], we developed a multimodal fall detection system. We used multiple sources of information: five IMUs as wearable sensors, one helmet to extract an average brainwave signal, six ambient infrared sensors and two RGB cameras. We placed five Mientlab MetaSensor wearable sensors that collected raw data from 3-axis accelerometer, 3-axis gyroscope and one ambient light transducer. These sensors were positioned in the body at neck, waist, left wrist, right pocket's trousers and left ankle, as the most preferred locations based on literature [3], [18], [19]. The helmet was an electroencephalograph (EEG) NeuroSky MindWave helmet that measured the average EEG signals using a sensor located at the forehead. We placed 6 infrared sensors around the scenario as contextual data retrieval, so a person located inside this mesh could be detected. We also placed two Microsoft LifeCam Cinema cameras at 1.82m above the floor, one in lateral view and the other one in front view to the actions performed. The whole system was built in a controlled laboratory room keeping the same positions of ambient sensors and cameras. Also, a mattress was placed at the center of the laboratory for secure purposes during falls.

TABLE 2. Types of falls and ADLs reported in the UP-Fall Detection data set, adapted from [11].

Action ID	Type	Description	Time (s)
1	Fall	Go forward using hands until laying	10
2		Go forward using knees until laying	10
3		Go backwards until laying	10
4		Go side-wards until laying	10
5		As sitting in an empty chair until laying	10
6	ADL	Walking	60
7		Standing	60
8		Sitting	60
9		Picking up an object	10
10		Jumping	30
11		Laying	60
20	Unknown	Otherwise, falling until knees	—

We collected a large data set namely UP-Fall Detection [11], using the system described above, and it was publicly released for the community.¹ This data set recorded 17 healthy young adults without impairment performing 5 types of falls and 6 different simple ADLs that were repeated 3 times. Table 2 summarizes the actions (falls and activities) collected in the fall detection system. Specifically,

¹<https://sites.google.com/up.edu.mx/har-up/>

if a subject remained in knees after falling, then we considered this as a particular case. All types of falls and ADLs were considered independent and non-overlapping. Notice that all actions related to falls are actually a sequence of activities such that every time the subjects start in standing up position, then the subject falls and lastly the subject remains laying on the mattress (or in knees). For more information about this data set, see reference [11].

B. METHODOLOGY FOR SENSOR LOCATION ANALYSIS

Throughout this work, we consider the usage of the five IMU wearable sensors and the two RGB cameras, as shown in Figure 1. Infrared sensors (context) were not taken into account since, in our previous work, we determined that those did not impact to the performance of the fall detection system, significantly [11].

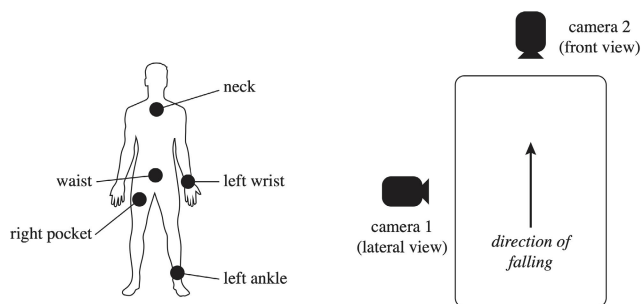


FIGURE 1. Layout of the sensors and camera viewpoints comprising the analysis of the fall detection system in this work.

In this work, we propose to analyze each IMU and camera separately in order to determine the predictability power of the system depending on the single location of IMUs into the body or the viewpoint of the cameras. Then, we propose to explore the predictability power of the system using a multi-sensor feature level basis approach. This sensor fusion considers the combination of features from one IMU sensor with features from one camera viewpoint. At the end, a statistical analysis is proposed to determine the minimal sensor (possibly a suitable combination of features) scheme for our fall detection system. To achieve this analysis, we propose the next workflow: (i) temporal segmentation, (ii) feature extraction and labeling, (iii) sensor modality selection, (iv) building machine learning models and (v) evaluation. The methodology follows in this work is presented in Figure 2. Details are described below.

1) TEMPORAL SEGMENTATION

First, we isolate the raw data from IMU sensors and RGB cameras. We performed a temporal segmentation, i.e. windowing, to the data. This procedure is typically implemented in fall detection systems for temporal abstraction [29]. Then, we selected three different window lengths of 1 second, 2 seconds and 3 seconds size to evaluate the impact of these windows to the overall fall detection system. Segments were obtained through the raw data with 50% of overlapping.

2) FEATURE EXTRACTION AND LABELING

We extracted different temporal and frequential features from windows, but it depended on the sensor type. Within each window frame of IMU sensors, 12 temporal features and 6 frequential features from the fast Fourier transform of the raw signals were extracted, as shown in Table 3.

TABLE 3. Temporal and frequential features extracted from window frames of raw data.

Type	Feature	Reference
temporal	mean	[29]
	standard deviation	[30]
	root mean square	[31]
	maximal amplitude	[32]
	minimal amplitude	[32]
	median	[33]
	number of zero-crossing	[31]
	skewness	[34]
	kurtosis	[34]
	first quartile	[33]
third quartile	[33]	
autocorrelation	[32]	
frequential	mean	[31]
	median	[31]
	entropy	[33]
	energy	[33]
	principal frequency	[33]
	spectral centroid	[33]

In addition, we extracted motion features from temporal window frames of RGB cameras. For each pair of consecutive images, we computed the relative motion of pixels using the Horn-and-Schunck optical flow method [35]. This results in two matrices of values, with the same size as the images, that represents (i.e. pixel-wise) the magnitude of displacement of pixels in the horizontal and vertical axes, namely u and v , respectively. From that, we computed the resultant relative motion of pixels as the Euclidean displacement, d , from u and v components. We treated formed a matrix with the d values per pixel, and we treated it as an image. Thus, we resized it to a fixed size of 20×20 pixels. We then computed the mean values of these images within a temporal window frame. We represented this mean matrix as a feature vector of 400 elements by unrolling it for further experimentation.

To this end, each IMU comprised 126 features (12 temporal and 6 frequential features) per channel (3-axes accelerometer, 3-axes gyroscope and 1 light intensity values). Each camera comprised 400 visual features as described above. For this analysis, we use the full feature set per device; thus no feature selection was done.

Furthermore, the action labels, included in the UP-Fall Detection dataset, are twelve numeric values representing the classes of action, as shown in Table 2. However, in this work, we considered for classifying a *fall* or *no-fall* action within a window frame. Thus, we changed the tags of actions from ADLs (6-11 IDs) and the unknown actions (ID 20) as class-0 (*no-fall*). As stated before, fall types are actually sequences of

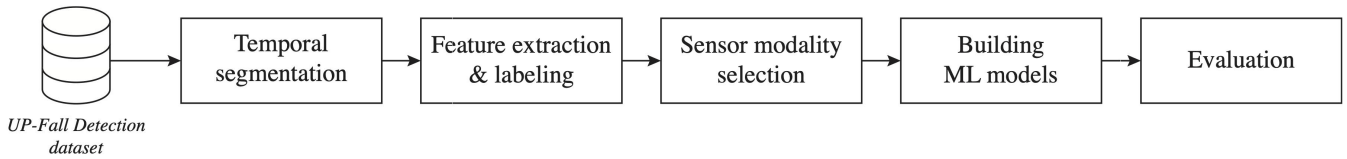


FIGURE 2. Methodology for minimal sensor analysis. Data from IMU sensors and camera viewpoints are retrieved from the UP-Fall Detection dataset. Then, data is segmented using windowing. Feature extraction and labeling is done within each window frame. After that, a selection among three different modalities of sensors is done. With these data, machine learning models are built. Finally, an evaluation on the different sensor modalities is performed.

three actions: while the subject is standing up (tagged here as *no-fall*), while the subject is falling down (tagged as class-1 or *fall*), and while the subject remains laid on the mattress or in knees (tagged as *no-fall*). Lastly, we tagged the most frequent class within a window frame.

3) SENSOR MODALITY SELECTION

We conducted two main experiments for determining the influence of single sensor location or viewpoint from cameras into the predictability power of the fall detection system. Then, we conducted a third experiment for determining the influence of combining one IMU sensor with one camera viewpoint. These experiments are listed below:

- *Experiment 1 – Single IMU sensor.* We took the feature set corresponding to an IMU sensor for building a machine learning model. We repeated this experiment for each IMU sensor location.
- *Experiment 2 – Single camera viewpoint.* We took the feature set corresponding to a camera viewpoint for building a machine learning model. We repeated this experiment for each camera viewpoint namely lateral view (camera 1) and front view (camera 2).
- *Experiment 3 – Feature sensor fusion.* We took the feature sets from the best ranked IMU sensor and the best ranked camera viewpoint to combine them and to build a machine learning model. Figure 3 shows this combination procedure.

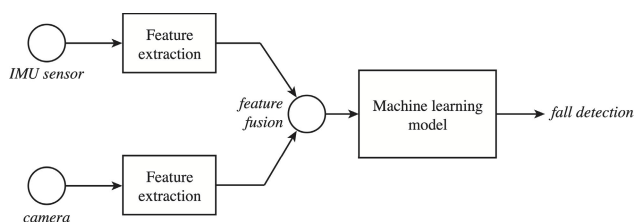


FIGURE 3. Proposed feature fusion with data retrieved from a single IMU sensor and a camera viewpoint.

4) BUILDING MACHINE LEARNING MODELS

We built four supervised machine learning classifiers to detect actions as fall or no-fall using the feature sets as inputs. A brief description of these methods are presented following:

- *Random forest (RF).* This is one of the most used methods in fall detection and human activity recognition [36].

It implies an ensemble of decision trees aiming to process the inputs into them, and computing the output class as the most frequent solution of the given trees.

- *Support vector machines (SVM).* It is also a popular classifier for fall detection [3]. The underlying of SVM is to transform the inputs into another space easily separable by hyper-planes. The latter are built over kernels that are trained to fulfill the classification task.
- *Multi-layer perceptron (MLP).* This is a classical artificial neural network using perceptrons as activation functions. It is typically employed for general nonlinear classification [37].
- *k-nearest neighbors (KNN).* This is an instance-based method that seeks the k-nearest neighbors of training points and compares them with an input data point. The output response is based on the most frequent class observed in the neighbors [37].

Since the data set is unbalanced (i.e. more no-fall than fall tags), we balanced the data set doing an over-sampling in the minority class (fall) by doubling the samples. We also performed a sub-sampling of the majority class (no-fall) to one third. Lastly, we split the data as follows: we retrieved the feature data associated with each subject, then these feature data per subject were split in 70% for training and 30% for testing. This allows us to have 70% training data of each subject, and the remaining data of each subject for testing.

Training parameters in the classification models were set as follows. RF were implemented with 2 estimators and one bootstrap. SVM were set with a radial basis function (RBF) kernel with coefficient $c = 1$ and tolerance value of 0.001. MLP were implemented with rectified linear units (ReLU) as activation functions in a one hidden layer of 100 neurons; for training we used a penalty parameter of 0.0001, a batch size of 200 with shuffling, a tolerance value of 0.0001, a regularization coefficient of 1×10^{-9} and a stochastic gradient method over 10 epochs. KNN was used with 5 neighbors with the Euclidean distance metric. Then, we conducted ten repetitions for each classifier model and we reported its average performance. Notice that this analysis was done for each feature data set at different window lengths (i.e., 1-sec, 2-sec and 3-sec).

5) EVALUATION

We evaluated the performance of the different configurations (window lengths, sources of information and machine learning models) of the fall detection system in terms of

four metrics: accuracy (1), precision (2), sensitivity (3) and F-score (4), where TP , TN , FP and FN represent true positive, true negative, false positive and false negative values.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$precision = \frac{TP}{TP + FP} \quad (2)$$

$$sensitivity = \frac{TP}{TP + FN} \quad (3)$$

$$F\text{-score} = 2 \times \frac{precision \times sensitivity}{precision + sensitivity} \quad (4)$$

IV. RESULTS OF THE SENSOR LOCATION ANALYSIS

This section reports the results of our analysis over the UP-Fall Detection system in order to determine the minimum number of sensors or possibly a minimal combination of them. Results are ordered as the experiments described in Section III-B3.

A. RESULTS OF EXPERIMENT 1 – SINGLE IMU SENSOR

We implemented an independent learning model for each window-length feature data set related to a single IMU sensor. Table 4 summarizes the results for each single IMU sensor located at left ankle, waist, neck, right pocket and left wrist. It also reports the mean and standard deviation (in parenthesis) of all metrics evaluated over 10 repetitions. Bold numbers represent the best values (in window length) per classifier.

In Table 4, it can be observed that RF performed the best in terms of accuracy and F-score, independently on the location of the IMUs. In contrast, MLP and KNN performed the worst. The two metrics, mentioned earlier, are valuable in the sense that accuracy corresponds to the ratio of the total number of correct predictions and the number of samples, while F-score measures the harmonic average of the precision and sensitivity. In terms of precision, SVM always computed the best. On the other hand, window lengths of 2-sec and 3-sec allowed, in almost all the cases, the best estimations in fall detection, as seen in Table 4. This result can be explained in the sense that falls have mean time duration of 2 seconds, and these window lengths can capture the falls better than partitions of 1-sec lengths.

In Table 5, it is summarized the ranking of the IMU sensors per performance classifier, based on the F-score, obtained from the previous results. It shows single IMU sensors sorted left-to-right. As observed, the best performance is obtained using the single IMU sensors: waist, neck and right pocket (i.e. the IMU types in the shadowed region). Also, ankle and left wrist IMU sensors performed the worst. Lastly, Table 6 shows the window length preference per single IMU sensor as the best performance of each classifier. Notice that the most frequent window lengths per classifier are: 3-sec for RF, 3-sec for SVM, 2-sec for MLP and 2-sec for KNN.

To this end, we conclude that the top three locations of IMU sensors are waist, neck and right pocket using RF and

SVM over features extracted in windows of three seconds (and overlapping of 1.5 seconds).

B. RESULTS OF EXPERIMENT 2 – SINGLE CAMERA VIEWPOINT

We conducted a similar analysis for each camera viewpoint. We implemented an independent learning model for each window-length feature data set. Table 7 summarizes the results for each camera viewpoint: lateral view (camera 1) and frontal view (camera 2). It reports the mean and standard deviation (in parenthesis) of all the metrics evaluated over the 10-fold cross-validation approach, and bold numbers represent the best values (in window length) per classifier.

From Table 7, it can be observed that RF performs the best based on accuracy and F-score, independently on the camera viewpoint. SVM, KNN and MLP reached poor performance. For camera 1 (lateral view), MLP was the worst model classifier; SVM was the worst for camera 2. In addition, Table 8 summarizes the ranking of the most suitable camera viewpoint depending on the classifier, sorted by the F-score. Notice that the lateral view (camera 1) performed the best. Also, RF classifier outperformed in this modality against to the other learning models. Moreover, Table 9 shows the window length preference per camera viewpoint to get the best performance in each classifier.

Thus, the best location of a camera is the lateral view, with reference to the human body, using RF in 3-sec window length segments with overlapping of 1.5 seconds.

C. RESULTS OF EXPERIMENT 3 – FEATURE SENSOR FUSION

From the above results, we considered two possible positions of IMU sensors, i.e. waist and right pocket. These locations represent good choices by elderly people since those are the most frequent acceptable wearable locations and, also, these are the most correlated body locations reported in literature [13]. In addition, we considered the lateral view (camera 1) as the preferred camera viewpoint.

In that sense, we combined each selected IMU sensor with camera 1. This means that all features from both the IMU sensor and the camera were joint together for training purposes (see Figure 3). Table 10 reports the results in performance of these combinations using the four classifiers with the 3-sec window length, as defined before. As it can be observed, RF is the learning model that best performed in terms of accuracy and F-score in both wearables and cameras. In terms of the modalities, the combination using the waist wearable and the lateral view (camera 1) performed the best.

D. ANALYSIS OF THE RESULTS

In Table 11 we summarized the results obtained in the above analysis based on the RF model and 3-sec window length. As shown, accuracy and F-score values are very similar, thus a statistical analysis was conducted. We calculated the non-parametric statistical Wilcoxon test [38] for comparing the five modalities and determining the significant difference

TABLE 4. Metrics of the IMU sensors. It includes the mean (standard deviation) values. Numbers in bold represent the best metrics per classifier.

IMU type	Classifier	Window	Accuracy (%)	Precision (%)	Sensitivity (%)	F-score (%)
Left Ankle	RF	1-sec	97.86 (0.21)	93.74 (0.98)	92.05 (1.53)	92.87 (0.72)
		2-sec	98.42 (0.32)	93.94 (1.82)	96.27 (1.82)	95.06 (1.00)
		3-sec	98.19 (0.44)	91.35 (1.92)	96.97 (1.42)	94.07 (1.43)
	SVM	1-sec	95.17 (0.48)	100	68.15 (3.16)	81.02 (2.25)
		2-sec	95.29 (0.69)	100	70.24 (4.37)	82.44 (3.06)
		3-sec	95.69 (0.41)	100	70.92 (2.79)	82.96 (1.90)
	MLP	1-sec	72.90 (21.60)	42.11 (15.52)	46.57 (21.82)	37.56 (6.85)
		2-sec	82.13 (2.95)	40.90 (6.71)	24.17 (24.13)	24.59 (14.43)
		3-sec	83.12 (1.74)	37.09 (11.17)	17.11 (10.06)	21.87 (8.63)
	KNN	1-sec	87.31 (0.52)	62.03 (2.47)	41.97 (2.70)	50.04 (2.52)
		2-sec	87.33 (0.83)	62.81 (3.71)	49.37 (2.84)	55.20 (2.47)
		3-sec	88.56 (0.71)	63.53 (3.53)	53.95 (3.32)	58.26 (2.54)
Waist	RF	1-sec	97.96 (0.23)	94.07 (1.22)	93.29 (1.21)	93.21 (0.75)
		2-sec	98.50 (0.37)	94.45 (1.35)	96.19 (1.95)	95.30 (1.19)
		3-sec	98.41 (0.43)	91.60 (2.15)	98.36 (1.26)	98.36 (1.35)
	SVM	1-sec	95.65 (0.29)	100	71.29 (1.91)	83.22 (1.29)
		2-sec	95.38 (0.46)	100	70.79 (2.88)	82.87 (1.98)
		3-sec	95.65 (0.69)	100	70.66 (4.64)	82.72 (3.20)
	MLP	1-sec	79.62 (4.84)	39.65 (6.26)	48.51 (20.41)	39.54 (12.34)
		2-sec	73.49 (13.06)	32.49 (8.51)	46.59 (16.79)	36.24 (8.16)
		3-sec	78.31 (4.14)	31.03 (8.23)	34.87 (15.68)	31.14 (9.36)
	KNN	1-sec	89.98 (0.36)	74.16 (2.22)	52.21 (1.47)	61.25 (1.29)
		2-sec	91.20 (0.53)	78.46 (2.29)	61.23 (2.95)	68.73 (2.13)
		3-sec	91.84 (0.59)	80.20 (2.15)	59.67 (3.46)	68.38 (2.74)
Neck	RF	1-sec	97.76 (0.38)	92.26 (1.60)	93.09 (1.46)	92.67 (1.23)
		2-sec	98.28 (0.48)	93.87 (1.39)	95.36 (2.16)	94.60 (1.52)
		3-sec	98.70 (0.31)	92.81 (1.70)	98.95 (0.79)	95.77 (0.98)
	SVM	1-sec	95.43 (0.38)	100	69.84 (2.51)	82.22 (1.71)
		2-sec	95.32 (0.40)	100	70.40 (2.56)	82.60 (1.76)
		3-sec	95.73 (0.54)	100	71.18 (3.65)	83.11 (2.58)
	MLP	1-sec	83.70 (4.94)	54.28 (11.86)	46.20 (22.20)	44.27 (9.61)
		2-sec	83.13 (7.64)	56.72 (13.55)	43.81 (17.15)	44.93 (8.37)
		3-sec	72.81 (23.41)	42.16 (13.48)	50.99 (31.94)	36.13 (15.04)
	KNN	1-sec	90.43 (0.37)	77.12 (1.99)	52.51 (1.67)	62.46 (1.54)
		2-sec	90.80 (0.49)	81.49 (2.83)	54.21 (2.38)	65.06 (2.03)
		3-sec	91.31 (0.60)	84.55 (2.15)	50.53 (3.59)	63.20 (3.16)
Right Pocket	RF	1-sec	97.89 (0.28)	93.46 (0.86)	92.61 (1.70)	93.02 (0.99)
		2-sec	98.47 (0.43)	94.77 (1.51)	95.63 (1.55)	95.20 (1.37)
		3-sec	98.57 (0.33)	92.43 (1.71)	98.48 (1.21)	95.35 (1.08)
	SVM	1-sec	95.52 (0.50)	100	70.52 (3.34)	82.66 (2.26)
		2-sec	95.13 (0.41)	100	69.20 (2.65)	81.77 (1.87)
		3-sec	95.76 (0.50)	100	71.44 (3.38)	83.30 (2.29)
	MLP	1-sec	85.78 (3.57)	60.93 (12.73)	52.57 (21.70)	51.46 (6.79)
		2-sec	86.16 (2.66)	57.95 (8.44)	61.23 (15.53)	57.67 (7.47)
		3-sec	83.55 (2.82)	46.45 (6.41)	47.96 (17.58)	45.31 (7.58)
	KNN	1-sec	91.76 (0.31)	76.96 (1.31)	65.22 (1.85)	70.59 (1.29)
		2-sec	91.96 (0.61)	77.49 (2.31)	69.40 (2.89)	73.19 (2.18)
		3-sec	91.32 (0.69)	72.75 (3.68)	66.71 (2.35)	69.52 (1.89)
Left Wrist	RF	1-sec	97.17 (0.33)	93.66 (1.07)	87.21 (1.46)	90.32 (1.16)
		2-sec	98.32 (0.52)	95.20 (1.44)	94.17 (2.87)	94.66 (1.69)
		3-sec	97.88 (0.34)	91.61 (1.58)	94.34 (2.26)	92.93 (1.16)
	SVM	1-sec	95.40 (0.33)	100	69.68 (2.19)	82.11 (1.52)
		2-sec	95.37 (0.42)	100	70.71 (2.64)	82.82 (1.83)
		3-sec	95.24 (0.34)	100	67.89 (2.29)	80.86 (1.62)
	MLP	1-sec	77.20 (11.96)	41.05 (13.39)	39.88 (19.60)	34.33 (6.83)
		2-sec	76.69 (18.29)	45.47 (13.88)	42.86 (18.32)	39.06 (4.57)
		3-sec	76.35 (19.19)	37.20 (14.16)	42.17 (22.21)	35.59 (12.32)
	KNN	1-sec	87.32 (0.35)	64.10 (2.05)	39.39 (2.53)	47.18 (2.12)
		2-sec	87.19 (0.58)	64.00 (2.49)	43.37 (3.81)	51.63 (3.12)
		3-sec	87.47 (0.94)	63.22 (5.99)	37.30 (4.69)	46.76 (4.57)

TABLE 5. Ranking (left-right) of the best IMU sensor per classifier, based on the F-score (in parenthesis). Shaded region represents the top three classifiers performing fall detection.

Classifier	IMU type				
	Rank # 1	Rank # 2	Rank # 3	Rank # 4	Rank # 5
RF	(98.36) Waist	(95.77) Neck	(95.35) Right Pocket	(95.06) Ankle	(94.66) Left Wrist
SVM	(83.30) Right Pocket	(83.22) Waist	(83.11) Neck	(82.96) Ankle	(82.82) Left Wrist
MLP	(57.67) Right Pocket	(44.93) Neck	(39.54) Waist	(39.06) Left Wrist	(37.56) Ankle
KNN	(73.19) Right Pocket	(68.73) Waist	(65.06) Neck	(58.26) Ankle	(51.63) Left Wrist

TABLE 6. Preferred window length in single IMU sensor per classifier.

Classifier	Window length				
	Left Ankle	Waist	Neck	Right Pocket	Left Wrist
RF	2-sec	3-sec	3-sec	3-sec	2-sec
SVM	3-sec	1-sec	3-sec	3-sec	2-sec
MLP	1-sec	1-sec	2-sec	2-sec	2-sec
KNN	3-sec	2-sec	2-sec	2-sec	2-sec

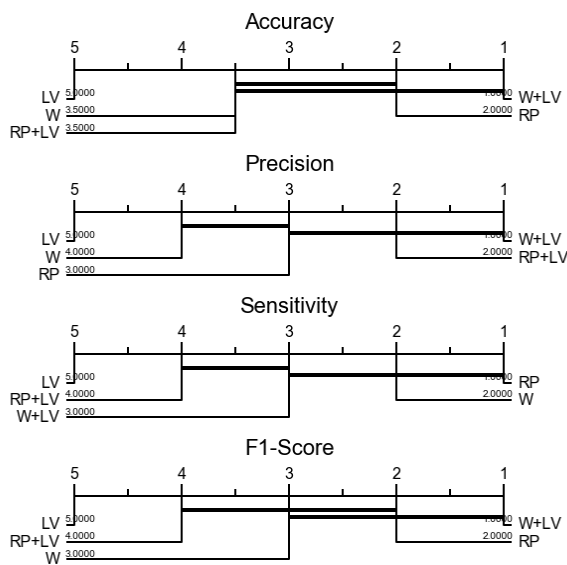


FIGURE 4. CD diagrams of the statistical analysis conducted in the best sensor modalities. W: Waist, RP: Right Pocket, LV: Lateral View.

among them, using a confidence level of $\alpha = 0.05$. Table 12 shows the p -values among the different modalities in terms of accuracy, precision, sensitivity and F-score. Figure 4 shows a graphical representation of this statistical analysis using critical difference (CD) diagrams.

It can be observed that the lateral viewpoint is significantly different from the other combinations in terms of the accuracy and F-score. In addition, waist and right pocket modes are similar, as well as, the combinations are similar among themselves. From this analysis, it can be concluded that using a single camera with lateral view will not be able to get an accurate performance, and all of the other modes can be used indistinguishably. In this regard, it is possible to say that a minimal configuration of a single sensor is preferable than those with multiple sources.

V. CASE STUDY: MINIMAL SENSOR-BASED UP-FALL DETECTION SYSTEM

After the analysis for sensor location, we conducted a preliminary deployment of a minimal sensor-based system for fall

detection using the configuration setup of our UP-Fall Detection system. This case study aims to evaluate the classification performance through the inter-patient paradigm [39].

From Table 11, we decided to use only the IMU sensor related to the right pocket location. This decision came from the fact that a device located in the trouser’s pocket is easier to adopt than any other wearable device. In addition, there is no strong evidence that using a combination of a single IMU sensor and a camera viewpoint will increase the accuracy of the fall detection system, as demonstrated in the statistical analysis. Four key factors influenced our decision: (i) the sensor analysis considers indistinguishable using a single sensor or a combination of sources, (ii) a minimal configuration using a single sensor is preferable than those with multiple sources, (iii) easiness in smartphone adoption from the user’s point of view [6], and (iv) cameras are associated to privacy concerns that may increase difficulties for adoption in users [7]. Thus, we decided to implement the fall detection system in a smartphone device located in the right pocket of the subjects. To validate our minimal sensor-based fall detection system, we deployed a machine learning model (only using information from the UP-Fall Detection dataset) in a mobile application, and then we tested the system under realistic conditions and with new and previously unseen subjects to the model.

In activity recognition and fall detection sensor orientations are commonly considered as fixed across acceleration signal sequences, however this assumption cannot be held in real world applications [40]. The acceleration signals describe three orthogonal acceleration components in the sensor reference system which is not always well-aligned with the body reference system [41]. The orientation varies depending on how the subject carries the smartphone. We cannot assume the landscape orientation in a pocket of the subject when falling [42]. The effects of different placements and orientations of the device on the accuracy must be studied in [40], [41]. In order to consider realistic conditions, several orientations of smartphone placement must be considered to simulate different natural settings.

A. MOBILE APPLICATION DEPLOYMENT

We developed a mobile application for the operating system iOS v13.0. The mobile application was able to collect data from the embedded sensors in an iPhone XS smartphone, specifically from a three-axis accelerometer and a three-axis gyroscope. Feature extraction and machine learning model prediction were executed locally and in real-time on

TABLE 7. Metrics of the camera viewpoints. It includes the mean (standard deviation) values. Numbers in bold represent the best metrics per classifier.

Camera	Classifier	Window	Accuracy (%)	Precision (%)	Sensitivity (%)	F-score (%)
Lateral View	RF	1-sec	84.30 (0.38)	48.95 (1.16)	50.73 (2.15)	49.80 (1.41)
		2-sec	86.02 (0.62)	55.59 (1.76)	61.76 (2.89)	58.50 (2.09)
		3-sec	87.83 (0.68)	59.24 (2.07)	65.75 (4.35)	62.27 (2.78)
	SVM	1-sec	84.56 (0.10)	17.24 (18.22)	0.45 (0.56)	0.88 (1.08)
		2-sec	83.82 (0.24)	47.42 (6.38)	6.96 (1.74)	11.97 (2.64)
		3-sec	84.70 (0.44)	50.92 (4.20)	16.09 (2.27)	24.25 (2.05)
	MLP	1-sec	70.79 (27.71)	3.07 (6.14)	20.0 (40.0)	5.33 (10.65)
		2-sec	70.42 (27.23)	3.19 (6.39)	20.0 (40.0)	5.51 (11.01)
		3-sec	84.69 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
	KNN	1-sec	83.24 (0.26)	42.01 (1.44)	24.10 (1.36)	30.62 (1.39)
2-sec		83.08 (0.60)	45.53 (2.64)	29.90 (2.37)	36.05 (2.23)	
3-sec		84.23 (0.59)	48.05 (2.66)	36.65 (3.15)	41.52 (2.65)	
Frontal View	RF	1-sec	82.65 (0.44)	43.59 (1.35)	44.03 (1.76)	43.79 (1.31)
		2-sec	85.25 (0.59)	53.50 (1.72)	58.20 (3.62)	55.71 (2.23)
		3-sec	85.45 (0.92)	52.32 (2.81)	58.32 (4.66)	55.06 (2.85)
	SVM	1-sec	84.64 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
		2-sec	83.98 (0.15)	2.0 (6.0)	0.10 (0.31)	0.20 (0.59)
		3-sec	84.69 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
	MLP	1-sec	63.86 (31.75)	4.61 (7.04)	30.0 (45.83)	7.99 (12.20)
		2-sec	77.23 (20.42)	1.60 (4.79)	10.0 (30.0)	2.75 (8.26)
		3-sec	50.0 (34.69)	7.66 (7.66)	50.0 (50.0)	13.28 (13.28)
	KNN	1-sec	82.64 (0.25)	38.43 (1.16)	21.61 (1.39)	27.63 (1.25)
		2-sec	82.41 (0.35)	40.47 (2.05)	21.59 (1.88)	28.13 (1.96)
		3-sec	82.72 (0.63)	37.08 (3.67)	18.21 (1.73)	24.40 (2.20)

TABLE 8. Ranking (left-right) of the best camera per classifier, based on the F-score (in parenthesis). Shaded region represents the top classifiers performing fall detection.

Classifier	Camera viewpoint	
	Rank # 1	Rank # 2
RF	(62.27) Lateral View	(55.71) Front View
SVM	(24.25) Lateral View	(0.20) Front View
MLP	(13.78) Front View	(5.51) Lateral View
KNN	(41.52) Lateral View	(28.13) Front View

TABLE 9. Preferred window length in camera view per classifier.

Classifier	Window length	
	Lateral View	Front View
RF	3-sec	2-sec
SVM	3-sec	2-sec
MLP	2-sec	3-sec
KNN	3-sec	2-sec

the device. Figure 5 shows the workflow for the design of this fall detection system.

We took the feature data set extracted with window frames of 3-seconds length related to the right pocket IMU sensor (see Section III-B2). Then, we made a feature selection using a method by Witten and Frank [43]. This method selects features by combining subsets of attributes and evaluating them in a classifier, and then ranking the most powerful attributes found in each subset. The evaluation of subsets was implemented with three ranking methods for attribute

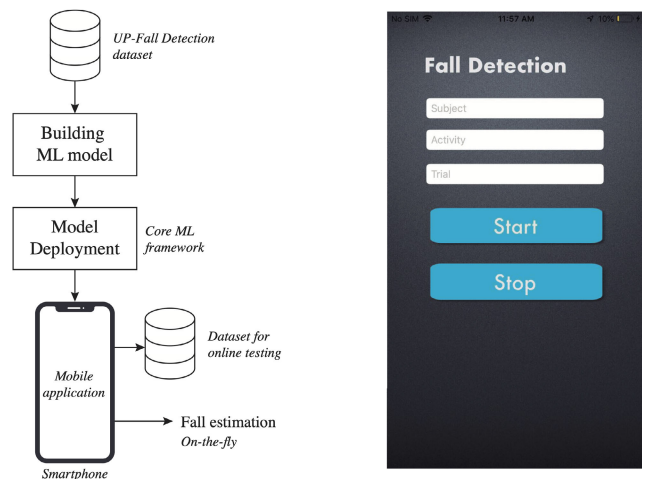


FIGURE 5. Mobile application design. (left) Workflow of the fall detection system on the device, and (right) a screenshot of the application for online testing purposes.

correlation (Pearson’s correlation), relief and classification (ZeroR and Decision Tables). The 10 top ranked features were finally selected, as summarized in Table 13.

From our sensor analysis, we determined that RF machine learning model was a suitable option. Thus, we trained an RF model using the selected features from the right pocket IMU sensor. Again, we balanced the feature set doing an over-sampling in the minority class by doubling the samples and a sub-sampling of the majority class to one third. We split the feature set in 70% for training and 30% for validation.

TABLE 10. Metrics of the combined IMU sensors and Camera 1 using 3-sec window length. It includes the mean (standard deviation) values. Numbers in bold represent the best metrics per modality.

Metric	Waist + Lateral View				Right Pocket + Lateral View			
	RF	SVM	MLP	KNN	RF	SVM	MLP	KNN
Accuracy (%)	98.72 (0.35)	95.59 (0.40)	77.67 (11.04)	91.71 (0.61)	98.41 (0.49)	95.79 (0.58)	84.92 (2.98)	91.71 (0.58)
Precision (%)	94.01 (1.51)	100	33.73 (11.69)	77.90 (3.33)	93.64 (1.46)	100	55.70 (11.36)	73.63 (3.19)
Sensitivity (%)	97.63 (1.56)	70.26 (2.71)	37.11 (26.74)	61.64 (3.68)	95.79 (2.65)	71.58 (3.91)	48.29 (25.11)	68.95 (2.73)
F-score (%)	95.77 (1.15)	82.51 (1.85)	29.81 (12.81)	68.73 (2.58)	94.69 (1.67)	83.38 (2.64)	45.21 (14.19)	71.13 (1.69)

TABLE 11. Comparison of the most suitable multimodal devices using RF and 3-sec window length. It includes the mean (standard deviation) values. Numbers in bold represent the best metrics per modality.

Multimodal	Accuracy (%)	Precision (%)	Sensitivity (%)	F-score (%)
Waist	98.41 (0.43)	91.60 (2.15)	98.36 (1.26)	98.36 (1.35)
Right Pocket	98.57 (0.33)	92.43 (1.71)	98.48 (1.21)	95.35 (1.08)
Lateral View	87.83 (0.68)	59.24 (2.07)	65.75 (4.35)	62.27 (2.78)
Waist + Lateral View	98.72 (0.35)	94.01 (1.51)	97.63 (1.56)	95.77 (1.15)
Right Pocket + Lateral View	98.41 (0.49)	93.64 (1.46)	95.79 (2.65)	94.69 (1.67)

TABLE 12. Pairwise p -values among modalities for all metrics.

Metric	p-values	RP	LV	W + LV	RP + LV
Accuracy	W	0.125047	0.005062	0.137254	0.72058
	RP	-	0.005062	0.506795	0.342829
	LV	-	-	0.005062	0.005062
	W + LV	-	-	-	0.012616
Precision	W	0.202622	0.005062	0.046853	0.028417
	RP	-	0.005062	0.059336	0.139414
	LV	-	-	0.005062	0.005062
	W + LV	-	-	-	0.260393
Sensitivity	W	0.888366	0.005062	0.539233	0.021824
	RP	-	0.005034	0.219467	0.032201
	LV	-	-	0.005062	0.005005
	W + LV	-	-	-	0.068364
F-score	W	0.168807	0.005062	0.168807	0.721277
	RP	-	0.005062	0.575062	0.284503
	LV	-	-	0.005062	0.005062
	W + LV	-	-	-	0.015156

W: Waist, RP: Right Pocket, LV: Lateral View.

TABLE 13. Features extracted and selected from sensor data.

Sensor-axis	Feature
Accelerometer: z-axis (g)	Standard deviation
Accelerometer: y-axis (g)	Third quartile
Gyroscope x-axis (deg/s)	Skewness
Accelerometer: y-axis (g)	Kurtosis
Gyroscope z-axis (deg/s)	Median
Accelerometer: z-axis (g)	Maximal amplitude
Gyroscope z-axis (deg/s)	Autocorrelation
Gyroscope z-axis (deg/s)	Kurtosis
Accelerometer: y-axis (g)	Maximal amplitude
Accelerometer: x-axis (g)	Autocorrelation

We did ten repetitions for training the RF model, and the best model obtained was selected. Then, we deployed this RF model into the smartphone device using the Core ML framework [44] to integrate our machine learning model into the mobile application. For testing purposes, we included a data acquisition and collection procedure on the device that captures raw signals from the sensors. These data and their timestamps were stored in files to create a new data set (see Section V-B). The mobile application also included a simple graphical user interface (Figure 5) to get information about the subject ID, action ID and number of repetition

(trial) performed. Notice that data acquisition, feature extraction, model prediction and fall estimation run in real-time on the device.

B. DATA SET FOR ONLINE TESTING

We created a new data set with different subjects, in contrast to the UP-Fall Detection dataset, for inter-patient evaluation scheme. We recruited 5 young subjects without any impairments (1.70 ± 0.05 m height and 63.8 ± 6.3 kg weight), 2 males and 3 females, ranging from 19 to 24 years old. They performed the same 11 actions as summarized in Table 2, three trials each. The same laboratory conditions and instrumentation were used, as described in Section III-A. In addition, we used an iPhone XS smartphone and the mobile application presented above. We collected raw data from the three-axis accelerometer and the three-axis gyroscope at 100 Hz. We used the lateral view camera (camera 1) for manual inspection and tagging of sequences of actions during falls.

For this new data set, we changed the orientation of the smartphone in each trial. In the first trial (*orientation 1*), the smartphone was oriented upside down and with the screen flipped to the body of the subject. In the second trial (*orientation 2*), the smartphone was oriented upright and with the screen flipped to the body of the subject. In the third trial (*orientation 3*), the smartphone was oriented upright and with the screen flipped out the body of the subject.

C. RESULTS OF ONLINE TESTING

This section reports the preliminary results of the deployed fall detection system in the smartphone, following the analysis of sensor location proposed in Section III.

Table 14 summarizes the results of the inter-patient online testing in the three orientations of the smartphone. The overall performance obtained was 87.56% of accuracy and 69.79% of sensitivity. It can be observed that the performance of the deployed system depends on the orientation of the smartphone. For instance, both *orientation 2* and *orientation 3*,

TABLE 14. Inter-patient performance of the online testing for each orientation of the smartphone in the right pocket. It reports the accuracy and sensitivity metrics for each subject, and the average.

Subject	Orientation 1		Orientation 2		Orientation 3	
	accuracy (%)	sensitivity (%)	accuracy (%)	sensitivity (%)	accuracy (%)	sensitivity (%)
1	70.94	85.71	92.94	80.0	92.42	83.33
2	80.20	75.0	93.51	100.0	95.09	100.0
3	72.25	66.67	87.14	50.0	85.71	37.50
4	75.59	71.43	91.63	75.0	95.26	33.33
5	83.17	28.57	98.56	100.0	98.09	80.0
Average	75.79	57.78	91.79	69.23	92.55	57.50

where the smartphone was oriented upright, performed better than *orientation 1* based on the accuracy metric. It is remarkable to say that the IMUs in the UP-Fall Detection dataset were randomly oriented [11], thus orientation in the online testing gives insights on the preferable orientation of the smartphone (i.e., upright or upside-down). In terms of the sensitivity metric, all the orientations performed similar results, with some preference on *orientation 2* (the screen of the smartphone is flipped to the body of the subject).

Thus, our proposed deployed fall detection system based on a smartphone outperforms the work reported in [45].

VI. DISCUSSION

Our results showed that the three best wearable sensor locations are at waist, neck (near chest) and right pocket (thigh). These results are consistent with the top locations (waist, chest and thigh) reported in the related work (Table 1). Regarding the camera, when a fall occurs in real life situations, we can not know if the person is seen by the camera from a lateral or frontal viewpoint. Under simulated conditions in our experiments, the lateral viewpoint shows slightly better performance. This result is consistent with [23] which states that sideways view present better results for fall detection. If vision devices are only used in a fall detection system, achieving the cooperation of multiple cameras to gain robustness and avoid occlusion is suggested.

From the results of the analysis of sensor placement, it is evident that there is a decreasing in the predictive power of this deployed fall detection system (overall accuracy of 87.56%) when comparing with the binary classifier model trained directly from the UP-Fall Detection dataset (overall accuracy of 98.57%, see Table 11). Differences in sensors, e.g. sampling rate, resolution, orientation, limited the performance of the deployed model. Furthermore, the deployed model was a direct translation from the one obtained from the UP-Fall Detection dataset. Thus, advanced techniques, e.g. transfer learning, should be applied for improving the performance of the deployed fall detection system.

In terms of our case study using a smartphone, we consider this preliminary experiment successful in the sense that it is possible to minimize the number of sensors and features to achieve a fall detection system even in an inter-patient scheme. Two reasons affected the decision to use an smartphone in this study, to say, the easiness of adoption and the built-in IMU sensors providing the same channels information as the ones used in the UP-Fall Detection dataset. Due to all the simplification procedure and the different factors identified, the minimal sensor-based fall detection system still reaches a significant accuracy (87.56%). This also shows that the proposed methodological analysis for sensor location can be applied on a realistic fall detection system.

Lastly, we aware that the dataset contains data from falls simulated by young healthy adults without any impairments. This can cause some differences with real falls in elderly.

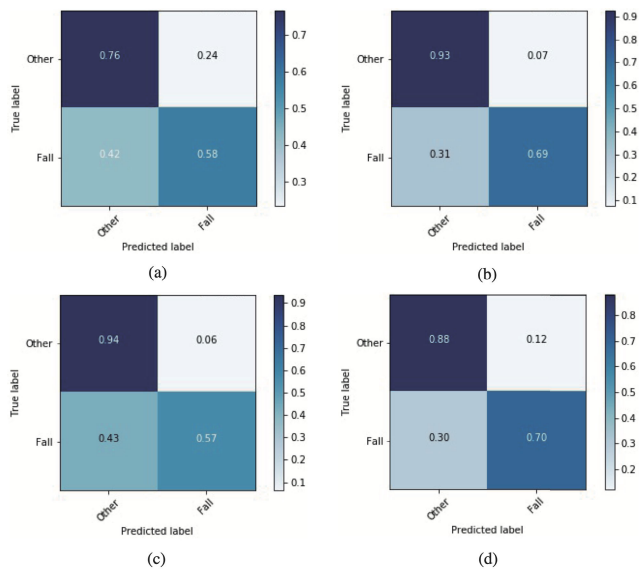


FIGURE 6. Normalized confusion matrices of the average performance in the inter-patient evaluation: (a) *orientation 1*, (b) *orientation 2*, (c) *orientation 3*, and (d) overall performance (87.56% of accuracy and 69.79% of recall).

Figure 6 shows the average normalized confusion matrices of the three orientations and the overall performance summarizing all the orientations. The confusion matrices of the orientations (Figure 6(a)–(c)) show that fall actions are estimated less accurate than no-falls. Furthermore, the confusion matrix of the overall performance reveals that the deployed binary classifier model estimates falls 70% of the times while the no-falls are detected 88% (Figure 6). For comparison purposes with the state-of-the-art, the work in [45] is one of the limited reported efforts regarding to holistic approaches of fall detection systems (from early stage design to deploy models). In that work, authors developed a single IMU-based system with 59% of precision and 44% of sensitivity.

However, we did not recruited elderly people mainly due to safety and healthy issues.

VII. CONCLUSION

In this work, we determined the minimal number of sensors required for developing an accurate fall detection system, using the configuration of the UP-Fall Detection dataset. After the analysis on single IMU sensors placed on different locations into the body and two camera viewpoints, we identified that the best combination was the waist IMU sensor and the lateral viewpoint of the camera, reaching 98.72% of accuracy. Moreover, we found from a statistical analysis that this combination is significantly similar to use the right pocket IMU sensor in both single or combined with a camera in lateral viewpoint. From this methodological analysis, we then implemented a minimal sensor-based fall detection system using a smartphone. Due to the several simplifications and differences in technical specifications of the sensors encountered in our experimentation, the smartphone-based fall detection system achieved an overall accuracy of 87.56%.

For future work, we are considering to use transfer learning for improving the performance of the smartphone-based system. Also, we are investigating the creation of a dedicated distributed hardware for fall detection system.

REFERENCES

- [1] Organization WH, Ageing WHO, Unit LC, *WHO Global Report on Falls Prevention in Older Age*, World Health Organization, Geneva, Switzerland, 2008.
- [2] *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables*, United Nations, New York, NY, USA, 2017.
- [3] R. Igual, C. Medrano, and I. Plaza, "Challenges, issues and trends in fall detection systems," *Biomed. Eng. Online*, vol. 12, no. 1, p. 66, 2013.
- [4] N. Noury, A. Fleury, P. Rumeau, A. K. Bourke, G. O. Laighin, V. Rialle, and J. E. Lundy, "Fall detection-principles and methods," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 1663–1666.
- [5] M. Mubashir, L. Shao, and L. Seed, "A survey on fall detection: Principles and approaches," *Neurocomputing*, vol. 100, pp. 144–152, Jan. 2013.
- [6] Y. Delahoz and M. Labrador, "Survey on fall detection and fall prevention using wearable and external sensors," *Sensors*, vol. 14, no. 10, pp. 19806–19842, Oct. 2014.
- [7] X. Ma, H. Wang, B. Xue, M. Zhou, B. Ji, and Y. Li, "Depth-based human fall detection via shape features and improved extreme learning machine," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 6, pp. 1915–1922, Nov. 2014.
- [8] H. F. Nweke, Y. W. Teh, G. Mujtaba, and M. A. Al-garadi, "Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions," *Inf. Fusion*, vol. 46, pp. 147–170, Mar. 2019.
- [9] G. Koshmak, A. Loutfi, and M. Linden, "Challenges and issues in multi-sensor fusion approach for fall detection," *Journal of J. Sensors*, vol. 2016, Dec. 2015, Art. no. 6931789.
- [10] G. Z. Yang, J. Andreu-Perez, X. Hu, "Multi-sensor fusion," in *Body Sensor Networks*. London, U.K.: Springer, 2014, pp. 301–354.
- [11] L. Martínez-Villaseñor, H. Ponce, J. Brieva, E. Moya-Albor, J. Núñez-Martínez, and C. Peñafort-Asturiano, "UP-fall detection dataset: A multi-modal approach," *Sensors*, vol. 19, no. 9, p. 1988, Apr. 2019.
- [12] E. Cippitelli, F. Fioranelli, E. Gambi, and S. Spinsante, "Radar and RGB-depth sensors for fall detection: A review," *IEEE Sensors J.*, vol. 17, no. 12, pp. 3585–3604, Jun. 2017.
- [13] A. Özdemir, "An analysis on sensor locations of the human body for wearable fall detection devices: Principles and practice," *Sensors*, vol. 16, no. 8, p. 1161, Jul. 2016.
- [14] B. Aguiar, T. Rocha, J. Silva, and I. Sousa, "Accelerometer-based fall detection for smartphones," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2014, pp. 1–6.
- [15] L.-J. Kau and C.-S. Chen, "A smart phone-based pocket fall accident detection, positioning, and rescue system," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 1, pp. 44–56, Jan. 2015.
- [16] P. Ntanasis, E. Pippa, A. T. B. Barshan, and V. Megalooikonomou, "Investigation of sensor placement for accurate fall detection," in *Proc. Int. Conf. Wireless Mobile Commun. Healthcare*. Cham, Switzerland: Springer, 2016, pp. 225–232.
- [17] J. Santoyo-Ramón, E. Casilari, and J. Cano-García, "Analysis of a smartphone-based architecture with multiple mobility sensors for fall detection with supervised learning," *Sensors*, vol. 18, no. 4, p. 1155, Apr. 2018.
- [18] E. Casilari, J. A. Santoyo-Ramón, and J. M. Cano-García, "Analysis of a smartphone-based architecture with multiple mobility sensors for fall detection," *PLoS ONE*, vol. 11, no. 12, Dec. 2016, Art. no. e0168069.
- [19] E. Casilari, J. A. Santoyo-Ramón, and J. M. Cano-García, "UMAFall: A multisensor dataset for the research on automatic fall detection," *Procedia Comput. Sci.*, vol. 110, pp. 32–39, 2017.
- [20] R. Rucco, A. Sorriso, M. Liparoti, G. Ferraioli, and P. Sorrentino, "Type and location of wearable sensors for monitoring falls during static and dynamic tasks in healthy elderly: A review," *Sensors*, vol. 18, no. 5, p. 1613, May 2018.
- [21] Z. Zhang, C. Conly, and V. Athitsos, "A survey on vision-based fall detection," in *Proc. 8th ACM Int. Conf. Pervas. Technol. Rel. Assistive Environ. (PETRA)*, 2015, pp. 1–7.
- [22] R. Cucchiara, A. Prati, and R. Vezzani, "A multi-camera vision system for fall detection and alarm generation," *Expert Syst.*, vol. 24, no. 5, pp. 334–345, Nov. 2007.
- [23] C. Rougier, J. Meunier, A. St-Arnaud, and J. Rousseau, "Robust video surveillance for fall detection based on human shape deformation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 5, pp. 611–622, May 2011.
- [24] N. Thome, S. Miguet, and S. Ambellouis, "A real-time, multiview fall detection system: A LHMM-based approach," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 11, pp. 1522–1532, Nov. 2008.
- [25] Z. Zhang, W. Liu, V. Metsis, and V. Athitsos, "A viewpoint-independent statistical method for fall detection," in *Proc. 21st Int. Conf. Pattern Recognit. (ICPR)*, Nov. 2012, pp. 3626–3630.
- [26] M. Hasan, S. Islam, and S. Abdullah, "Robust pose-based human fall detection using recurrent neural network," in *Proc. IEEE Int. Conf. Robot., Automat., Artif.-Intell. Internet-Things (RAAICON)*, Dhaka, Bangladesh, Nov./Dec. 2019, pp. 48–51.
- [27] Z. Huang, Y. Liu, Y. Fang, and B. K. P. Horn, "Video-based fall detection for seniors with human pose estimation," in *Proc. 4th Int. Conf. Universal Village (UV)*, Boston, MA, USA, Oct. 2018, pp. 1–4.
- [28] H. Nait-Charif and S. J. McKenna, "Activity summarisation and fall detection in a supportive home environment," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, 2004, pp. 323–326.
- [29] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Comput. Surv.*, vol. 46, no. 3, pp. 1–33, Jan. 2014.
- [30] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, "Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey," in *Proc. 23th Int. Conf. Archit. Comput. Syst.*, Feb. 2010, pp. 1–10.
- [31] A. Phinyomark, A. Nuidod, P. Phukpattaranont, and C. Limsakul, "Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification," *Elektronika ir Elektrotechnika*, vol. 122, no. 6, pp. 27–32, 2012.
- [32] W. Dargie, "Analysis of time and frequency domain features of accelerometer measurements," in *Proc. 18th Int. Conf. Comput. Commun. Netw.*, Aug. 2009, pp. 1–6.
- [33] A. Rasekh, C.-A. Chen, and Y. Lu, "Human activity recognition using smartphone," 2014, *arXiv:1401.8212*. [Online]. Available: <http://arxiv.org/abs/1401.8212>
- [34] L. Atallah, B. Lo, R. King, and G.-Z. Yang, "Sensor placement for activity detection using wearable accelerometers," in *Proc. Int. Conf. Body Sensor Netw.*, Jun. 2010, pp. 24–29.
- [35] H. BK and S. BG, "Determining optical flow," *Proc. SPIE*, vol. 0281, pp. 319–331, Nov. 1981.
- [36] S. Kozina, H. Gjoreski, M. Gams, and M. Luštrek, "Efficient activity recognition and fall detection using accelerometers," in *Proc. Int. Competition Evaluating AAL Syst. Through Competitive Benchmarking*. Berlin, Germany: Springer, 2013, pp. 13–23.

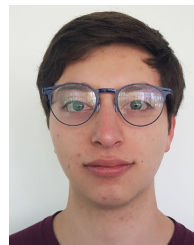
- [37] T. M. Mitchell, "The discipline of machine learning," School Comput. Sci., Mach. Learn., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-ML-06-108, 2006, vol. 9.
- [38] G. W. Corder and D. I. Foreman, *Nonparametric Statistics: A Step-by-Step Approach*. Hoboken, NJ, USA: Wiley, 2014.
- [39] G. Garcia, G. Moreira, D. Menotti, and E. Luz, "Inter-patient ECG heart-beat classification with temporal VCG optimized by PSO," *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, Dec. 2017.
- [40] S. Yu and C. Hsinchun, "Fall detection with orientation calibration using a single motion sensor," in *Proc. Int. Conf. Wireless Mobile Commun. Healthcare*. Cham, Switzerland: Springer, 2016, pp. 233–240.
- [41] A. Jain and V. Kanhangad, "Human activity classification in smartphones using accelerometer and gyroscope sensors," *IEEE Sens. J.*, vol. 18, no. 3, pp. 1169–1177, Feb. 2018.
- [42] S.-L. Hsieh, C.-T. Yang, and H.-J. Li, "Combining wristband-type devices and smartphones to detect falls," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 2373–2377.
- [43] I. H. Witten, E. Frank, H. A. Mark, and P. J. Christopher, *Data Mining: Practical Machine Learning Tools and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 2016.
- [44] Apple. (2020). *Framework Core ML: Integrate Machine Learning Models Into Your App*. Accessed: Apr. 3, 2020. [Online]. Available: <https://developer.apple.com/documentation/coreml>
- [45] K. Frank, M. J. Vera Nadales, P. Robertson, and T. Pfeifer, "Bayesian recognition of motion related activities with inertial sensors," in *Proc. 12th ACM Int. Conf. Adjunct Papers Ubiquitous Comput.-Adjunct*, 2010, pp. 445–446.



HIRAM PONCE (Member, IEEE) graduated in mechatronics engineering and received the master's degree in science engineering and the Ph.D. degree in computer science from the Tecnológico de Monterrey, Mexico. He is currently a full-time Professor and Researcher with the School of Engineering, Universidad Panamericana, Mexico. He is a member of the National System of Researchers rank level I of the National Council of Science and Technology, Mexico. He is an author of more than 70 international journal and conference publications, eight book chapters, and four books in the area of artificial intelligence and robotics. He is an executive board member of the Mexican Society of Artificial Intelligence, Mexico, a member of the Technical Committee on Neural Networks in the IEEE Computational Intelligence Society, a member of the Technical Committee on Robotics and Mechatronics in IFToMM, among other memberships. He was awarded with a Google Research Award for Latin America in 2017. He has served as a guest editor in different special issues of reputed journals. He is currently an Associate Editor of IEEE Access.



LOURDES MARTÍNEZ-VILLASEÑOR (Member, IEEE) received the Ph.D. degree in computational sciences from the Tecnológico de Monterrey, Mexico. She is a Computer Systems Engineer with the Tecnológico de Monterrey. She is currently a Research Professor in the area of artificial intelligence with Universidad Panamericana and the Head of the Computing Area at the Faculty of Engineering. She has the distinction of rank level I of the National System of Researchers of National Council of Science and Technology (Conacyt). She is an author of more than 35 international journal and conference publications, three book chapters, and five books in the area of artificial intelligence. Her main research interests are artificial intelligence applied to human activity recognition and user modeling, and ethics for artificial intelligence. She is a member of the board of the Mexican Society of Artificial Intelligence. She has been the guest editor in different special issues of international journals.



JOSÉ NUÑEZ-MARTÍNEZ is currently pursuing the B.S. degree in mechatronics engineering with Universidad Panamericana, Mexico, and he is scheduling to begin a Specialist degree, in July 2020. During his B.S. studies, he took an exchange semester with the Hochschule Schmalkalden University of Applied Sciences, Germany, during the Spring semester of 2019. He started his involvement with research projects involving Human Activity Recognition (HAR) during the Summer of 2018 as an intern under Ph.D. Lourdes Martínez-Villaseñor, and has been working in a Multimodal System for Fall Detection ever since. He has also taken part in several projects such as SAE AeroDesign, Space Missions Mexico, and Colibri Mission UP-MIT. His research interests include in the use of technology to help improve the quality of life for as many people as possible, while learning in the process.

...