IEEE Access
Multidisciplinary Rapid Review Open Access Journal

Received September 16, 2019, accepted December 17, 2019, date of publication December 30, 2019, date of current version January 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2019.2962813

# WiVi: A Ubiquitous Violence Detection System With Commercial WiFi Devices

**LEI ZHANG**[1,2,3], (Member, IEEE), **XIN RUAN**[1,2], AND **JU WANG**[4]
[1]College of Intelligence and Computing, Tianjin University, Tianjin 300350, China
[2]Tianjin Key Laboratory of Advanced Network Technology and Application, Tianjin University, Tianjin 300350, China
[3]Key Laboratory of Grain Information Processing and Control, Henan University of Technology, Ministry of Education, Henan 450001, China
[4]School of Biomedical Engineering, Tianjin Medical University, Tianjin 300070, China

Corresponding author: Ju Wang (wangju@tmu.edu.cn)

This work was supported by the NSFC under Grant No. 61772364. This work was also supported by National Instrument Program under Grant 2013YQ030915.

**ABSTRACT** School violent behaviors is a serious problem that greatly affects the healthy growth of the youth and the children. Current prevention measures mainly depend on propaganda and self-report. So far, there is still no effective solution that can automatically detect the violent behaviors. In this paper, we take the first attempt to build a ubiquitous passive violence detection system, WiVi, based on the commercial WiFi infrastructure. To capture the patterns of the complicated violent behaviors actions, besides the time-series features used in existing activity recognition works, WiVi also leverages the correlated features extracted from combined subcarriers, to take fully advantages of Channel State Information. WiVi integrates a feature fusion method to select the appropriate features for the classification model in different scenarios. We implement and evaluate WiVi in various real-world environments. The experimental results show that the recall and specificity that WiVi can achieve 93.46% and 93.57% respectively.

**INDEX TERMS** Violent behaviors, ubiquitous detection, channel state information (CSI).

## I. INTRODUCTION

School violent behaviors has becoming a major problem that seriously affects the physical and mental health of the youth all over the world. According to the National Center for Education Statistics of USA, 28% of the students in graders 6-12 experienced bullying [1]. Violent behaviors usually involves violence that hurts the youth's physical health. Research statistics [2] show that 32.4% of the middle school students experienced bullying with pushing or shoving, and 29.2% of them experienced hitting, slapping, or kicking. Violent behaviors can also cause problems such as emotional loss and insomnia, and even depression and suicide. The study [3] shows that the probabilities of depression and suicide for the victims being bullied are 4.8 times and 18.5 times larger than the ordinary people. Given the serious impacts of bullying, many countries have enacted the law of anti-bullying. However, few of the victims being bullied report to the school due to the fear of retaliation [4]. Hence, a timely and automatic violent behaviors detection method is urgently needed to prevent the youth from suffering from bullying.

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

However, detecting the violent behaviors is not an easy task because it can happen anywhere at any time. Monitoring of administrators such as safeguards and teachers can only cover a very limited areas in a very limited time. With the development of video analysis techniques and the widespread use of surveillance video, activity recognition approaches based on computer vision techniques are proposed to detect the violent incidents [5]. However, due to the cost and privacy issues, there are many blind spots of surveillance video such as the bathrooms, where 12.2% of violent behaviors happen [2]. Wearable sensor-based sensing technique is an alternative solution for violence detection [6]. But it requires the compliance or cooperation of users. Due to the privacy, it is non-travail to obtain the sensing data of the students' devices. Even the monitoring system can obtain the data, the students who bully others can easily evade inspections by taking off the wearable devices during bullying. Ambient environment sensor-based sensing technique can also detect the violent behaviors by analyzing ambient environmental information such as sound [7] and temperature [8]. But due to the limited sensing range of environmental sensors such as microphone, we have to densely deploy the dedicated sensors to build the special infrastructure. The deployment and maintenance costs are too high to achieve ubiquitous violent behaviors detection.

In recent years, the extensively deployed WiFi facilities boosts the advances of commercial WiFi based sensing techniques that can recognize specific human actions [9], [10] or gestures [11], [12]. Specifically, Channel State Information (CSI) provided by the physical layer reveals the multipath features of the channel, at the granularity of OFDM subcarriers. When human actions and gestures alter the multipath of signal propagation, the amplitude and phase of the received CSI will changes. If there are distinguishable signal change patterns during the action or gesture, we can recognize the human activity by identifying the corresponding patterns. The ubiquitous wireless signal sheds the light on using WiFi sensing techniques to detect the violent behaviors events.

Even though many efforts have been made on CSI-based activity recognition, existing approaches cannot be directly used in our scenario. First, most of the existing works focus on recognizing pre-defined actions and gestures with features extracted from the CSI time series of several conspicuous subcarriers. The features are usually manually defined or selected by time series processing techniques under a certain circumstance. However, violent behaviors is a more complicated activity that involves the whole body and will not leave distinguishable patterns on any subcarrier. Therefore, existing methods that rely on time series features cannot achieve high accuracy. How to recognize the violent activities without significant time series features is still an open problem. Second, existing approaches usually focus on relatively stable operating scenarios However, violent behaviors can happen anywhere in the campus. How to select and integrate different effective features in different environments is challenging. Besides, when operating environment changes, how the model adapt to the new environment in an appropriate manner still need to address.

To address aforementioned issues, we propose WiVi, a novel passive violent behaviors detection system that is able to accurately detect the complicated violent activities even when the operating environment changes. Our insight is that CSI is underutilized by existing methods because they treat CSI of different subcarriers as individual time series data and extract only time-series features. Actually, more features can be extracted from the combined subcarries, which are regarded as correlated features in this paper. For the complicated activities like violent behaviors, all the subcarries can be influenced but experience quite different patterns because many parts of the human body affects different subcarriers. Correlated features can capture the relationship between the change patterns of different subcarriers and then provide much more information than separately considering the patterns on individual subcarriers.

WiVi exploits both the time-series features extracted from individual CSI subcarriers, and the correlated features extracted from the combination of different subcarriers. Inspired by the feature extraction in computer vision area, we propose an Gabor-filter based feature extraction method to extract the correlated features from combined subcarries. To automatically integrate the time-series features and correlated features in different environments, we leverage Principle Component Analysis (PCA) technique and propose the PCA-based feature fusion method to automatically select the effective features. To achieve adaption for environment changes, we design a feedback adjustment method that adjusts the model parameters and even retrains the model when the system performance degrades to a user-defined threshold.

We implement a prototype of WiVi on commercial WiFi devices and evaluate its performance. The experimental results show that WiVi can detect the violent activities with an accuracy of 93.46%, along with a false alarm rate of 6.43%. The contributions of this paper are summarized as follows:

- We propose WiVi, a ubiquitous passive violent behaviors detection system using the commercial CSI infrastructure. To capture the patterns of complicated violent actions, WiVi exploits both time-series features and correlated features to fully explore the information contained in CSI.
- We design a Gabor-filter based feature extraction method to automatically extract the correlated features from the combined subcarriers. To accurately utilized the features, we also propose an PCA-based feature fusion method to integrate the time-series features and correlated features. Our PCA-based feature fusion method can efficiently obtain the suitable features that characterize the violent activities.
- We implement a prototype of WiVi with commercial WiFi devices and evaluate its performance in various environments. The experimental results show that WiVi can achieve the recall of 93.46% and the specificity of 93.57%.

The rest of this paper is organized as follows. We introduce our detailed designs of WiVi in Section II. Then, we evaluate WiVi and present the experimental results in Section III. We summary the related work in Section IV. Finally, we conclude our work in Section V.

## II. METHODOLOGY

In this section, we first present the overview of WiVi and then present the design details of WiVi by elaborating the four components of WiVi.

### A. OVERVIEW OF WiVi

WiVi tries to recognize violent activities in time with only commercial WiFi devices. The overview of WiVi is shown in Figure 1. WiVi mainly consists of four components: data collection, preprocessing, feature extraction, and classification. The data collection component extracts CSI samples from the commercial WiFi devices. In preprocessing component, the CSI samples are firstly calibrated by the data calibration module to remove the noise. Then the calibrated

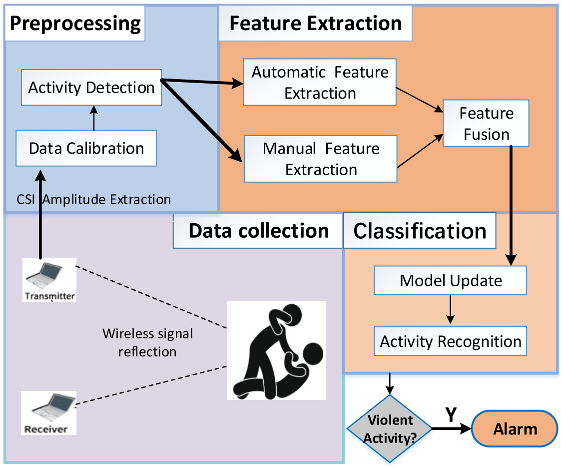



FIGURE 1. System overview.

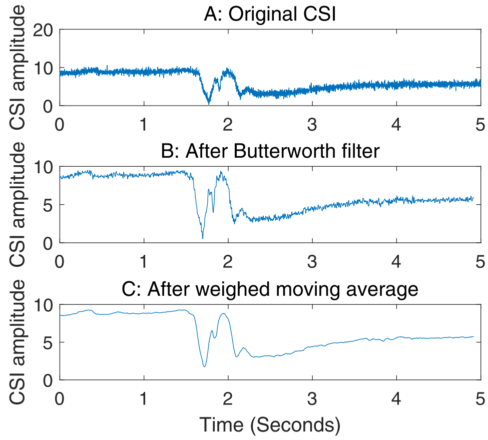



FIGURE 2. Data calibration.

CSI trace is processed by activity detection module to decide whether there is a possible activity in the CSI trace. If there is an activity, the module will detect the start and end of the activity. Then the extracted CSI segments are fed into feature extraction component. In WiVi, the CSI segments are fed into automatic and manual feature extraction modules at the same time to extract the correlated and time-series features respectively. Then WiVi integrates a feature fusion module to fuse the effective features in an appropriate way. The fused features are then used by the activity recognition module to decide whether the activity is an violent behavior. If detected, WiVi will send an alarm to inform the security guards. In the classification component, the extracted features are also used to update the classification model.

## B. DATA COLLECTION

WiVi collects CSI samples from the received WiFi packets on the receiver. CSI is an indicator that depicts the signal propagation in the multipath environment.

Suppose $X(f,t)$ is the frequency domain component of the transmitted signals with carrier frequency $f$ at time $t$, and $Y(f,t)$ is the corresponding frequency domain component received by the receiver. The relationship between $X(f,t)$ and $Y(f,t)$ is $Y(f,t) = H(f,t) \times X(f,t)$,where $H(f,t)$ is the complex valued channel frequency response (CFR). In practice, there are many propagation paths between transceivers.

$$H(f,t) = \sum_{k=1}^{N} \alpha_k(f,t) e^{-j2\pi f \tau_k(t)} \tag{1}$$

where $N$ is the total number of multipaths, $\alpha_k(f,t)$ and $\tau_k(k)$ represent the attenuation factor and propagation delay for the $k^{th}$ path respectively.

In WiVi, we use the laptop with Intel 5300 NIC to collect the CSI samples. The transmitter uses a single antenna and the receiver uses three antennas. For each transmitting and receiving antenna pair, there are 30 OFDM subcarriers. Hence, WiVi can obtain 90 CSI values from each received 802.11n frame. In this paper, the CSI sequences for each subcarrier with a given pair of sending/receiving antenna is called one CSI stream. The sampling rate of WiVi is 1000Hz.

## C. PREPROCESSING

The collected CSI streams are firstly processed by data calibration module to remove the noise. Then the activity detection module uses the calibrated CSI streams to decide whether there is a possible activity. If there is an activity, WiVi segments the CSI streams to locate the start and end of the activity for further processing. Otherwise, the CSI streams are discarded.

### 1) DATA CALIBRATION

The CSI provided by commercial WiFi devices can be extremely noisy. The major source of noise is the internal state transitions between the sender and receiver of WiFi NICs, such as transmission rate adaptation, transmission power changes and CSI reference level variations. Those internal state transitions can cause amplitude impulses and burst variations in CSI streams, as shown in Figure 2A.

In WiVi, we adopt Butterworth filter and weighed moving average methods to suppress the noise signal components. Specifically, Butterworth filter is leveraged to eliminate the out of-band noise. The CSI stream calibrated by Butterworth filter is shown in Figure 2B. The results show that Butteworth filter can only smooth out irrelevance in a certain range. There are still amplitude impulses and burst variations in the CSI stream. Hence, we further utilize the weighted moving average to process the CSI stream. Specifically, in a CSI steam, threr are $C_1, C_2, \ldots, C_t$. The CSI value at time $t$ is averaged by the previous CSI values. The latest CSI value has the highest weight $m$, so we have:

$$C_t = \frac{1}{m + \cdots + 2 + 1} \times (m \times C_t + (m-1) \times C_{t-1} + \cdots + 1 \times C_{t-m+1}) \tag{2}$$

where $C_t$ is the averaged new value. The value $m$ decides how many historical data are considered when averaging the current value. The improved result after the weighted moving average is shown in Figure 2C. The result shows that the noisy impulses are smoothed.

### 2) ACTIVITY DETECTION

Before further analyzing whether there is a violent activity, WiVi firstly decides whether there is a possible activity in the collected CSI stream to avoid unnecessary analysis.

We propose a Principal Component Analysis (PCA) based method to detect the existence of an activity in a given CSI stream. The major insight is the noise of CSI streams is random when there is no human activity but experience correlated variations when there human activities exist. Similar to CARM [13], WiVi leverages the second principal component to extract features for deciding the existence of human activities. This is because in the first component, some noise are mixed with the CSI variations caused by human activities. The second component also contains the activity information but has much less noise.

To find the start and end of an activity, WiVi leverages the amplitude changes. When the activity starts, the amplitude will have a sudden change and return to noise floor when the activity ends. To deal with the changes of noise floor, a dynamic threshold is leveraged to check the noise level in WiVi system. Specifically, we leverage an exponential moving average algorithm to update the noise floor, $L_t$, during the static period.

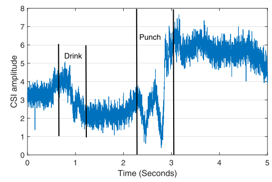
$$L_t = (1-\tau)L_{t-1} + \tau \times var_t \tag{3}$$

where the coefficient $\tau$ is set to be 0.06 empirically. And the variance $var_t$ for $i$-th sliding window with 100 samples is calculated. If $var_t$ is three times larger than the noise floor $L_t$ in a sliding window, the start point of an activity is detected. Similarly, the end point of the activity can also be acquired.

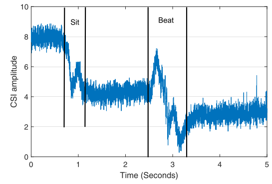
## D. FEATURE EXTRACTION

To distinguish the violent activities and daily activities, two obvious characteristics are observed from a large amount of data [14]. (1) Fierceness: The actions of assailant to attack the victim are always extremely fast and powerful. And the victim must respond quickly to avoid injury. As shown in Figure 3(a), this fact leads to a rapid rise or fall of the CSI, while the daily activities are different. (2) Range of motions: Aggressive actions generally have a large range of motion. Due to the fast and rapid features of violent activity, there will inevitably be a large range of motion. As shown in Figure 3(b), a large range of motion leads to a large variation of CSI amplitude. Even though there are many actions in daily activities, most of them are well distinguished from violent activities because they do not have the above two obvious characteristics. However, Some daily actions are also characterized by rapid and large-scale features, such as running and jumping in sports. In this paper, what we consider more in our experiments is the distinction between daily activities which are characterized by these two characteristics and violent activities.

For violent behaviors detection, feature extraction is the crucial component to achieve a satisfied detection accuracy. For the complicated violent activity, existing methods with only time-series features are suspected to have low detection accuracy, which is confirmed in our evaluation. In our work, to fully extract the information from the CSI streams for accurate violent behaviors detection, WiVi extracts features from two dimensions, the time-series and the correlated subcarriers. The time-series features are extracted manually like existing works do. Inspired by the automatic feature extraction method in computer vision area, automatic correlated feature extraction algorithm is designed to extract features from the CSI streams. Then, we design a PCA-based feature fusion method to integrate the features from two dimensions and select the effective features.




(a) Two different activities: drinking and punching




(b) Two different activities: sitting and beating

**FIGURE 3. Different types of action have different effects on CSI amplitude.**

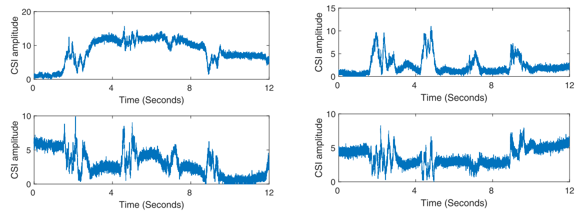
### 1) MANUAL FEATURE EXTRACTION

The preprocessing component decides the start and end of the interested activity. The manual feature extraction module is responsible for extracting the time-series features. Similar to existing works [9], the time-series features used in WiVi are 1) activity duration, 2) standard deviation, 3) median absolute deviation, 4) mean absolute deviation, 5) interquartile range, 6) average CSI, 7) maximum CSI, and 8) minimum CSI.

The disadvantages of manual feature extraction can be summarized as follows. First, existing works usually select the time-series features from several conspicuous subcarriers individually. The manually selected features are based on human experience, which is not able to make the best of CSI streams. Second, the environment changes cause the changes of effective features, which cannot be decided in advance. Therefore, it is necessary to find a counterpart method compensating for these defects.

### 2) AUTOMATIC FEATURE EXTRACTION

For the complicated actions like violent behaviors, all the subcarries can be affected but experience quite different patterns because different parts of the human body influences different subcarriers. Hence, the correlated features that reveal the relationship between the change patterns of different subcarriers can provide much more information than the time-series features extracted from individual subcarriers. For example, in Figure 4, we present the CSI streams on two subcarriers when continuously performing four actions. We can find that both subcarriers #7, #15 and #6, #20 experience amplitude changes during the actions but the patterns are quite different. Hence, only extracted from the individual time-series may lose information and even deduce contradictory results. And, time series based method can not handle this situation appropriately, but correlated feature can.




(a) Four actions sequence: squat, jump, punch, kick

(b) Four actions sequence:run, squat, slap, punch

**FIGURE 4. Two different subcarriers in one stream with the same activity.**

To extract the correlated features from multiple CSI streams, we propose a Gabor filter based feature extraction method. Gabor filter [15] is a linear filter for edge extraction in computer vision techniques. Its frequency and direction expression is similar to human visual system. Hence, it has superiority for texture analysis because it can provide good direction selection and scale selection characteristics. Gabor filter is leveraged to extract the correlated features because the relationships between multiple subcarriers are similar to the texture. We first transform the amplitude matrix of the whole CSI streams into a picture and then extract the features from the picture by Gabor filter.

The two-dimensional Gabor filter matrix is used to extract features from CSI pictures. The two-dimensional Gabor wavelet filter is defined as follows.

$$g(x, y, \lambda, \theta, \psi, \sigma, \gamma) = e^{-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}} e^{i(2\pi \frac{x'}{\lambda}+\psi)} \tag{4}$$

where $x$, $y$ is the coordinates of the pixels in the matrix, and wavelength $\lambda$, rotation angle $\theta$, phase $\psi$, aspect ratio $\sigma$, and bandwidth $\gamma$ are five parameters for Gabor filter. Eq. (4) can be divided into the form of real and imaginary parts as follows.

$$\begin{cases} g_{real}(x, y, \lambda, \theta, \psi, \sigma, \gamma) = e^{-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}} \cos(2\pi \frac{x'}{\lambda} + \psi) \\ g_{imag}(x, y, \lambda, \theta, \psi, \sigma, \gamma) = e^{-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}} \sin(2\pi \frac{x'}{\lambda} + \psi) \end{cases} \tag{5}$$

The traditional Gabor filter is slower when processing pictures. To provide the quick response, WiVi uses the Memetic algorithm [16] to design an improved Gabor filter, M-Gabor. Memetic algorithm is a hybrid computing intelligent optimization method framework. By combining the global search of the population with the local optimization of the individual, the Memetic algorithm can effectively improve the iterative process and the optimization efficiency, and avoid falling into premature convergence.

Ispired by Memetic algorithm, M-Gabor in WiVi uses a comprehensive learning particle swarm optimization algorithm [17] with good global search performance and an adaptive intelligent single-particle optimization algorithm. With the local optimization ability for co-evolution process, M-Gabor can obtain more representative feature data in a shorter time, and effectively improve the feature extraction efficiency.

In the M-Gabor algorithm, suppose the filter matrix size is $m \times n$ pixels, the coordinate of the pixel in the image is $(u, v)$, and the weighting factor of the extracted features is $w$, then the M-Gabor filter can be uniquely determined by ten parameters, $\lambda$, $\theta$, $\psi$, $\sigma$, $\gamma$, $m$, $n$, $u$, $v$, $w$. Then M-Gabor tunes the position vector of each particle by using the traditional trending random search process as a global optimization strategy and adding a localized optimization operation with clear purpose in each iteration. In M-Gabor, the parameters are sequentially arranged and connected end to end. Constructing the position vector of the Memetic algorithm optimization particle is as follows.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda_1$ | $\theta_1$ | $\psi_1$ | $\varphi_1$ | $\gamma_1$ | $m_1$ | $n_1$ | $u_1$ | $v_1$ | $w_1$ | ... |

When optimizing the parameters, M-Gabor uses CLPSO [17] as the global search strategy. The update formula is:

$$N_i^{k+1} = \omega_k \times N_i^k + c \times r^k \times (qbest_i^k - q_i^k) \tag{6}$$
$$q_i^{k+1} = q_i^k + N_i^{k+1} \tag{7}$$

where $N$ is particle velocity vector, $q$ is position vector, $i$ is the current update particle number, $k$ is iteration coefficient, $c$ is set of parameters, $r$ is a uniform distribution's random value in [0, 1], and $qbest$ is optimal position of individual particles. By introducing innovative learning strategies for particle velocity vector updates, CLPSO can achieve stronger global search capabilities than traditional particle swarm improvement methods.

For local search, M-Gabor leverages the fast converging AdpISPO. AdpISPO constructs the input parameters as the optimal particles and composes the small-scale particle population. Therefore it can adaptively adjust to the optimization process and accelerates the convergence speed of the particle optimization.

### 3) FEATURE FUSION

The time-series features and the correlated features should be appropriately integrated to automatically obtain the effective features. WiVi adopts a PCA-based feature fusion method. PCA, known as K-L transform, is one of the commonly used transforms in signal processing. Denote $x = \{x_1, x_2, \ldots, x_n\}$ as a $n_x$ dimensional input sample vector, the purpose of PCA is to generate $y = A^T x$, which meets $E[y_i y_j] = 0, i \neq j$. If $x$ is a normalized matrix, then $E[x] = 0$, and $E[y] = 0$, so covariance matrix can be expressed as:

$$R_y \equiv E[yy^T] = E[A^T xx^T A] = A^T R_x A \tag{8}$$

where $R_x$ is covariance matrix. For multiple training sets $X = \{x_1, x_2, \ldots, x_{n_x}\}$, The training set covariance matrix is:

$$R_x \approx \frac{\sum_k (x_k x_k^T)}{n_x} \tag{9}$$

Since $R_x$ is symmetrical, its eigenvectors are orthogonal to each other. The orthogonal eigenvectors $\{a_1, a_2, \ldots, a_{n_x}\}$ of $R_x$ is composed into transformation matrix $A$, then $R_y$ is the diagonal matrix, which meets the optimal unrelated requirements. Generally, the dimensions in $y$ are arranged in descending order of variance, and the larger variance of the preceding dimension is the principal component. And the PCA-based feature fusion steps are as follows:

(1) Feature merging: Combining the extracted features into a feature set in order $f = \{f_1, \ldots, f_\alpha\}$. For a training set $X = \{x_1, \ldots, x_n\}$ of multiple samples, a matrix $F = \{f_1, \ldots, f_\alpha\}$ of $n$ rows and $\alpha$ columns is obtained, where $f_i$ is the vector of the $i^{th}$ eigenvalues of $n$ different training samples, $f_i = \{f_{i,1}, f_{i,2}, \ldots, f_{i,n}\}^T$.

(2) Normalization: Given $N$ training samples, the mean value $\mu = \{\mu_1, \mu_2, \ldots, \mu_\alpha\}$ and the standard deviation $s = \{s_1, s_2, \ldots, s_\alpha\}$ can be obtained, where, for each feature dimension $i \in \{1, 2, \ldots, \alpha\}$,

$$\mu_i = \frac{\sum_{j=1}^{n} f_{i,j}}{n}, \quad s_i = \sqrt{\frac{\sum_{j=1}^{n} (f_{i,j} - \mu_i)^2}{n-1}} \tag{10}$$

Then we can normalize them as follows and get the the normalized feature matrix $G = \{g_1, g_2, \ldots, g_\alpha\}$.

$$g_{i,j} = \frac{f_{i,j} - \mu_i}{s_i} \tag{11}$$

(3) Feature transformation (PCA processing): Using the transformation matrix $A$, $G$ can be converted to the transform domain. Then, a matrix $H = F \cdot A$ with less correlation between different dimensions and descending according to the variance of each column is obtained, where $H$ is a matrix with $n$ rows and $\alpha$ columns. Then feature $f$ transforms into principal component $h = \{h_1, h_2, \ldots, h_\alpha\}$. The appropriate transformation matrix $A$ can be obtained through the training process.

(4) Dimensionality reduction: The first $d$-dimensional principal component (Pre-$d$ column data) of the matrix $H$ is the feature obtained by this method, where $d$ can be determined by cross-validation to determine the specific value, or the empirical value $d = {}^{\alpha}/_{3}$. The PCA-based feature fusion method effectively combines the features extracted by the two ways.

### E. CLASSIFICATION

#### 1) MODEL UPDATE

Least Square Support Vector Machine (LSSVM) is chosen as the classification model because of its lower computational complexity. The specific LSSVM calculation process is as follows:

(1) Creating the classification function. The training set is $(x_i, y_i), i = 1, 2, \ldots, l, x \in R^d, y \in R$. The offset is $b$, and the weight vector is $w$. Then LSSVM linear function of the high dimension is:

$$f(x) = w^T \cdot \varphi(x) + b \tag{12}$$

(2) Establishing the objective function according to the principle of structural risk minimization. Given the classification error is $\xi$, the regularization parameter is $r$. The objective function is as follows.

$$\min J(\omega, \xi) = \frac{1}{2}\|\omega\|^2 + c\sum_{i=1}^{l} \xi_i^2 \tag{13}$$

The limitation factor is:

$$y_i = \varphi(x_i)\omega + b + \xi_i, \quad i = 1, 2, \ldots, l \tag{14}$$

where $\|\omega\|^2$ is used to control the complexity of the model; $c$ is the regularization parameter; $\xi_i$ is relaxation factor.

The Lagrangian function corresponding to LSSVM optimization is:

$$L(\omega, b, \xi, a) = \frac{1}{2}\|\omega\|^2 + c\sum_{i=1}^{l} \xi_i^2 - \sum_{i=1}^{l} a_i(\varphi(x_i)\omega + b + \xi_i - y_i) \tag{15}$$

where $a_i (i = 1, 2, \ldots, l)$ is Lagrange multiplier. To obtain the partial derivative from the above equation and set the result to zero. Then, High-dimensional calculation can be transformed into problems for solving linear equations.

$$\begin{bmatrix} 0 & e^T \\ e & \Omega + \frac{1}{2}c * I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \tag{16}$$

where, $e = [1, 1, .., 1]$, $Y = [y_1, y_2, \ldots, y_l]$, $\Omega_{ij} = K(x_i, y_j)$. Finally, the regression model $f(x)$ can be derived.

$$f(x) = \sum_{i=0}^{l} a_i K(x, x_i) + b \tag{17}$$

(3) Choosing the kernel function. The radial basis is used as the kernel function. The classification decision function of LSSVM is:

$$f(x) = sgn(\sum_{i=1}^{n} a_i \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2}) + b) \tag{18}$$

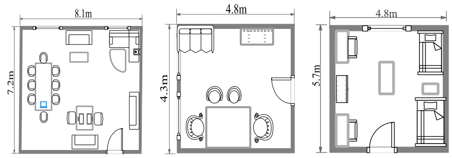
The width parameter of the radial basis kernel function is $\sigma$.

#### 2) ACTIVITY RECOGNITION

In the process of model update, violent activities and daily activities are detected and labeled in the continuously captured WiFi wireless signal streams. Then the fusion features along with the corresponding labels are fed into the LSSVM classifier to build the classification model. When detecting violent behaviors in practice, the LSSVM classifier makes the decision based on the observed CSI streams. If any violent behaviors is detected, WiVi will send the alarms to the security guards.

#### 3) THE METHOD OF MODEL UPDATE

To enable the function adaption for environment changes, we propose a feedback mechanism. The method allows WiVi to adaptively adjust the model for the given expected accuracy, which ensures that WiVi always works with high accuracy. Set the User-defined accuracy to $K$, and WiVi initial reset coefficient $R$ to 1. The accuracy of WiVi's real-time detection is $A$, when $A$ is less than $K$, the reset coefficient



(a) Laboratory (b) Office (c) Dorm room

**FIGURE 5.** Experimental scenarios.

is setted to $-1$. When the reset coefficient $R$ is $-1$, WiVi considers that the environment has changed. Then we re-collect data, train the model, and form a new model. The specific mathematical expression is:

$$R = \begin{cases} 1, & K > A \\ -1, & K \leq A \end{cases} \tag{19}$$

## III. EVALUATION

In this section, we first introduce our implementation and evaluation methodology. Then we present the evaluation results of WiVi in various scenarios.

### A. EVALUATION METHODOLOGY

#### 1) IMPLEMENTATION

We implement a prototype of WiVi on commercial WiFi devices. We leverage two Think-Pad T-series laptops equipped with Intel 5300 wireless NICs as the WiFi transmitter and receiver. Specifically, one laptop with an antenna works as the transmitter, and the other one with three antennas works as the receiver. Therefore the receiver can provide three CSI streams when operating on IEEE 802.11n monitor mode. The CSI samples are collected by the CSI Tool on Ubuntu desktop. The transmission rate is set to be 1KHz.

#### 2) SETTING

As shown in Figure 5, WiVi is evaluated in three scenarios, including a 8.1m $\times$ 7.2m laboratory where there are tables, chairs and computers in the environment, an 4.8m $\times$ 4.3m office where there are chairs, tables and bookcase, and a 4.8m $\times$ 5.7m dorm room where there are beds, tables and chairs. Our dataset is collected in these three environments. And we verify if WiVi is able to recognize the violent behaviors in different places.

Literature [18] referred to design our action set, which includes: walking, running, jumping, sitting, throwing, squatting, kicking, punching, slapping, stabbing, pushing, beating, and strangling. These actions involve many parts of the body. Eight volunteers, including four males and four females, contribute the experimental data. The volunteers vary in age (18-40 years old), height (155-183cm), and weight (45-85kg). During the experiments, the volunteers perform all kinds of actions and repeat 100 times for each action. So we obtain $100 \times 13 \times 8 \times 3$ traces for eight volunteers in the three scenarios.

**TABLE 1.** Overall performance of WiVi in three environments.

| Scenarios | Recall | Specificity |
|---|---|---|
| Office | 93.23% | 94.58% |
| Dorm room | 95.36% | 94.62% |
| Laboratory | 91.78% | 91.47% |

#### 3) EVALUATION METRICS

To analyze the performance of WiVi, we adopt the following two standard metrics, recall and specificity. Recall is defined as the percentage of correctly detected violent activities.

$$recall = \frac{TP}{TP + FN} \tag{20}$$

And specificity is defined as the percentage of correctly detected daily activities.

$$specificity = \frac{TN}{TN + FP} \tag{21}$$

where TP, TN, FP, and FN stand for the number of true positives, true negatives, false positives and false negatives, respectively. Recall and specificity demonstrate the accuracy of detecting violent activities as well as the capability of distinguishing the violent activities from daily activities.

### B. EVALUATION RESULTS

#### 1) OVERALL PERFORMANCE

The experimental results of detecting violent activities for all the volunteers in the environment of office, dorm room and laboratory are summarized in Table 1. WiVi achieves the recall rate of 93.46% and specificity rate of 93.57% on average. The recall and specificity in laboratory are 91.78% and 91.47%, the lowest in all the three environments. This is because the layout of laboratory is more complicated than the other two places and multipath interference is more serious. Among the three environments, WiVi in the dorm room achieves the best performance because the environment of dorm room is relatively more stable than the other two environments.

#### 2) IMPACT OF DIFFERENT FEATURES

WiVi considers not only the time-series features but also the correlated features to take full advantage of the CSI traces. To validate the effectiveness of different features, we conduct experiments with different features to recognize the violent behaviors. We consider the manually extracted time-series features, like existing works use for activity recognition, the automatically extracted correlated features, and the fusion features used in WiVi that combine both time-series features and correlated features by PCA. We plot the Receiver Operating Characteristic (ROC) curves of three different sets of features, as shown in Figure 6. The ROC curve can depict the tradeoff between TPR and FPR over various settings. From the results in Figure 6, we can find that the fusion features provide the best performance and the time-series features achieve the worst performance. We further study the performance of using different sets of features in different environments. The experimental results are shown

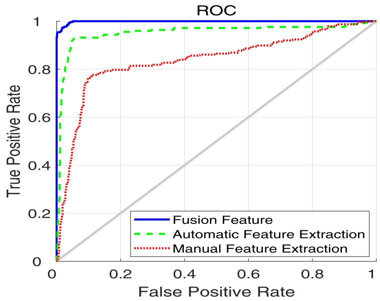



**FIGURE 6.** ROC curves of WiVi with fused features, correlated features and time-series features.

in Figure 7. We can find that the experiments show similar results in all the three environments. Fusion features can achieve 93.4% TPR in all three environments, which is the best. The correlated features can achieve the TPR of 87.2%, 90.8%, 90.3% in the lab, office, and dorm respectively. The time-series widely adopted in existing works can achieve the TPR of 82.4%, 86.1%, 85.6% in the lab, office, and dorm respectively. The results demonstrate that both time-series and correlated features contribute to the violent behaviors detection and they should be fused together to achieve the better performance.

### 3) THE ADJUSTMENT METHOD IN DIFFERENT ENVIRONMENTS

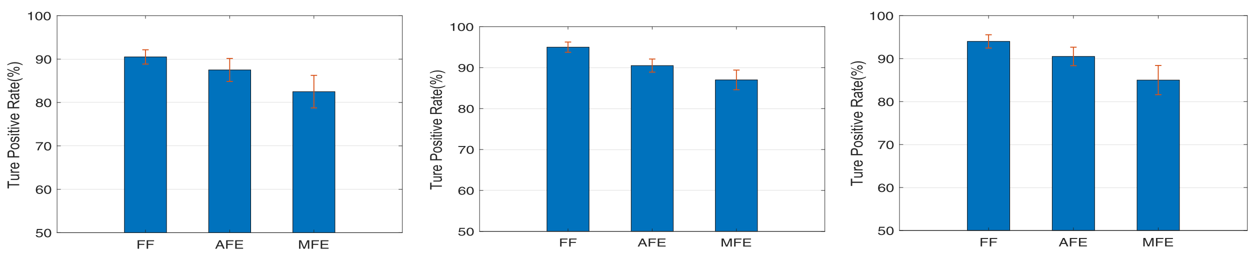
When the operating environments in which the data is collected changes, the signal propagation path also changes. The impacts of human activities on CSI can also change. Therefore, WiVi is designed with a feedback adjustment method to adapt to environment changes and adjust the classification model of WiVi in time. Specifically, WiVi maintains high recall and specificity in different environments. As shown in the Figure 8, the predefined system performance threshold $K$ is 80%, which can be decided by users according to the needs. In the initial laboratory environment, the accuracy of WiVi is above 90%. As time goes by, the environment changes. At 3.5 minutes, the recall and specificity of WiVi dropped to 75%, respectively, because the target moves from the laboratory to the office. The achieved performance is below the threshold $K$. Hence, the model need to be adjusted. After about 3 minutes, WiVi's performance return to over 90%. At about 7 minutes, the office environment is changed, by adding some tables and chairs. The performance of WiVi also drop, but do not fall below $K$. Then the model will be used without retraining. At 12 minutes, the operating environments restore. The accuracy also restores because we use the same classification model. Finally, at 14 minutes, WiVi's accuracy decreases significantly when the operating environment changes from the office to the restroom. After 4-minute adjustment, the performance of WiVi increase above the threshold. From the results, we can find that WiVi usually needs 3 to 5 minutes for adjusting the model from one environment to another. Once the model is trained, it can be used in the new environments with appropriate adjustments. The feedback-based adjustment method lets WiVi maintain a satisfied performance even when the environment changes.

### 4) THE PERFORMANCE IN DIFFERENT ENVIRONMENTS

Violent activity can occur in many places. Multi-scene experiments can demonstrate the system compatibility and applicability. Therefore, in addition to the previous three scenarios, we have also done some evaluation experiments in other scenarios. Several places are chosen where cameras or other detection devices are not easy to detect but violent behaviors often occurs in campus, such as the restroom,staircase corners. As shown in the Figure 9, five scenarios are used to do some evaluation experiments: classroom, restroom, staircase corner, gym and building corner. We can find that the performance of WiVi for these five scenarios is above 90%. Specifically, WiVi's overall performance in classroom, restroom, staircase corner and gym is similar. However, the recall and specificity of the building corner are 90% and 90.6% respectively, which is lower than other four. This may be the result of the instability environment when doing experiments outdoors. For example, someone suddenly makes a big movement in a relatively close distance, or the effects of wind and animal. From the experimental results, The adaptability and robustness to the different environment of WiVi are acceptable. Therefore, with the automatic adjustment method, the recognition accuracy for environmental change of the application scenarios is considerable.

### 5) IMPACT OF MULTIPLE PEOPLE

The principle behind action recognition is that the actions affect WiFi signal propagation, which is reflected by the variations of CSI. Therefore, if there are too many people moving




(a) Laboratory (b) Office (c) Dorm room

**FIGURE 7.** Feature of three different types' performance in three different scenarios. (FF, AFE and MFE which represent fusion feature, automatic feature extraction and manual feature extraction respectively).

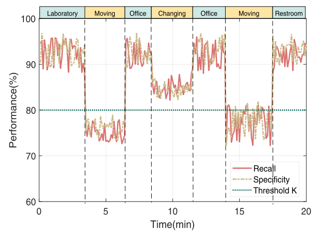



**FIGURE 8.** Changes in system performance over time and environment.

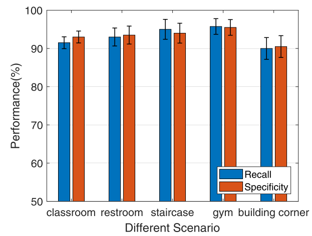
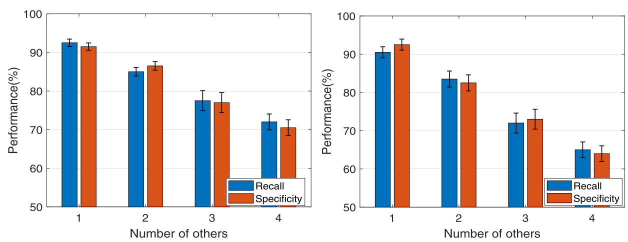



**FIGURE 9.** The performance of WiVi in different environments.




(a) Case 1 (b) Case 2

**FIGURE 10.** The Performance of WiVi in office with multiple people.

in the detecting area, the patterns of the impacts of the interested actions on WiFi signal can be corrupted by the actions that have larger impacts on CSI. The occurrence of violence is certainly not a person, sometimes there are several people.

To evaluate the robustness of WiVi in the environment with multiple people, we test its performance in two cases: (1) there is a volunteer who does the action of punching or other violent action and several (1-4) volunteers who remained still and watching the violent behaviors; (2) there is a volunteer who does the action of punching or other violent actions and several other (1-4) volunteers that have small movements, such as changing body postures and occasionally walking around in the environment. The experiments are conducted in the office.

The experimental results of case (1) and case (2) are shown in the Figure 10 (a) and (b). When there are two people in the environment, the performance of the recall and specificity of case (1) is 92.2%, 91.8%, and the case (2) is 90.5%, 92.4%. When the number of people increased to 5, the performance of case (1) recall and specificity is 72.6%, 70.7%, and the case (2) is 65.3%, 64.7%. For both cases, the performance decreases as the current people increases. However, the degradation of case (2) is much faster than case (1). The reason may be that the signals introduced by multiple volunteers doing actions mix together make the model unable to distinguish the type of the action.

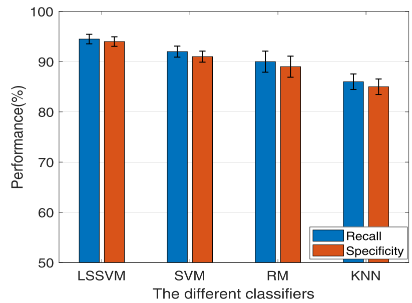



**FIGURE 11.** The performance of four classifiers.

#### 6) IMPACT OF DIFFERENT CLASSIFIERS

There are many classifiers that can be used as the recognition model, such as random forests, KNN, SVM, naive bayes classifiers, and so on. The choice of classifier also has an important impact on the overall performance of the system. Initially, the SVM classifier is leveraged to classify in WiVi. But, sometimes due to excessive data, the processing time is too long. LSSVM is faster than SVM. Hence, the LSSVM classifier is used to classify in our current implementation. The classifier achieves the best classification result by selecting the kernel function and adjusting the parameters. Then the performance of the random forest, KNN and SVM classifiers are compared. As the Figure 11 shown, when classifying with LSSVM and SVM, the performance are similar in terms of recall and specificity. That is, the two classifiers have the same effect in solving the classification problem of WiVi. However, when leveraging random forest and KNN classifiers, their performances are relatively low, and the performance of recall and specificity are between 90% and 80%. Overall, LSSVM classifier performs best, which is mostly above 90%.

#### 7) IMPACT OF DISTANCE

The distance between the target and the transmitter or receiver can significantly influence the detection accuracy because the signal attenuation and multipath efforts. We conduct experiments to evaluate the performance of WiVi, when the target has various distances to the receiver. Note that during our experiments, the data are collected both in line-of-sight (LOS) and non-LOS environments.We vary the target-to-device distances from 1m to 4m with step size of 1m. The experimental recall and specificity are shown in the Table 2. We can find that with the distance increase, both the recall and specificity decrease as expected. When the distance is 4m, the recall of WiVi is 80.6% and the specificity of WiVi is 81.7%. In violent behaviors detection application, the violent behaviors usually happen in the covert corners with limited size. Hence, the achievable performance of WiVi is acceptable. To obtain higher accuracy for a given user case, the system may need deploy more WiFi devices.

#### 8) IMPACT OF CSI SAMPLING RATE

Sampling rate is also important for activity recognition because it directly decides how much information the CSI data contain in a period. The accuracy of activity

**TABLE 2. The performance of WiVi under varying distances.**

| Distance | Recall | Specificity |
|---|---|---|
| 1m | 92.24% | 91.57% |
| 2m | 90.25% | 90.84% |
| 3m | 88.61% | 87.35% |
| 4m | 80.68% | 81.74% |

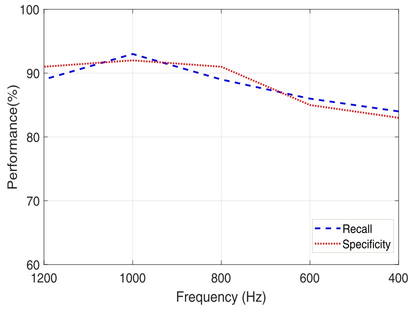



**FIGURE 12. The performance of different sampling rates.**

detection and extraction depends on the high precision of the CSI traces. A higher sampling rate is expected to provide more information and increase the feature extraction and classification accuracy. However, a higher sampling rate also causes more demands on computational cost. Therefore, it is necessary to find an appropriate sample rate. We evaluate WiVi with the sampling rate of 1200Hz, 1000Hz, 800Hz, 600Hz, and 400Hz. Figure 12 shows the experimental performance. We can find that with the decrease of sampling rate, the performance also decreases as expected. When the sampling is 1000Hz, the recall is 93.1% and the specificity is 92.5%. We can also find that a higher sampling rate than 1000Hz cannot bring much benefit and even have performance degradation because of the heavy computational cost. Hence, in our current implementation, we select 1000Hz as the default sampling rate.

#### 9) IMPACT OF THE NUMBER OF ANTENNAS

The antenna is a component used in a wireless device to transmit or receive radio waves. Wireless devices such as radars, communications and WiFi devices transmit information through radio waves, which all need receive and transmit radio waves. Antennas, like transmitters and receivers, are an important part of radio equipment. In our scenarios, when only one set of equipment is used, the transmitter is fixed with one antenna, and the receiver is equipped with 1,2 or 3 different number antennas. The number of antennas at the receiver determine the number of streams in data collected. As the Figure 13 shows, we evaluate the different number of antennas. From the figure we can find that the performance is all above 90% regardless of the number of receiving antennas, but the best performance is the three receiving antennas, which achieve the recall 93.9% and specificity 94.2% respectively. Therefore, our experimental data is collected with three receiving antennas.

## IV. RELATED WORK

Since violent behaviors is a kind of human activity, we summary the related works on human activity detection from computer vision, wearable sensor, ambient sensor, and Radio Frequency (RF) based approaches.

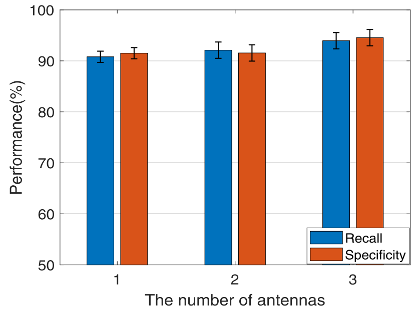



**FIGURE 13. The performance when using multiple antennas.**

### A. COMPUTER VISION BASED APPROACHES

Computer vision based methods leverage the surveillance cameras in the target areas to capture images or videos and use the image processing techniques to recognize specific activities [5], [19], [20]. Due to the rich information in the video, computer vision based approaches are usually accurate if clear line-of-sight (LOS) images are provided. However, it is not easy to capture clear images due to the non-LOS blockages and privacy concerns, resulting in a lot of blind spots such as the bathrooms.

### B. WEARABLE SENSOR BASED APPROACHES

A wearable sensor based activity detection system leverage the sensors embedded in the wearable devices such as belts and smart watches [21], [22]. Smartphones also provide a wide variety of sensors that can be used to recognize human activities [23], [24]. However, those systems usually require the users wear the dedicated devices and agree to share their sensing data. But in violent behaviors detection scenario, the user's compliance is not available.

### C. AMBIENT SENSOR BASED APPROACHES

An ambient sensor based action detection system provide a non-invasive detection solution by exploiting ambient sensors to monitor environment changes, such as temperature [8] and sounds [7]. The principle behind those ambient sensor based approaches is that different human activities cause different changes in ambient acoustic noise or temperature. However, due to the limited sensing range of those sensors, dense deployment is necessary to enable the ubiquitous violent behaviors detection. The deployment and maintenance costs are high.

### D. RF BASED APPROACHES

The influence of human activities on RF signals can be captured and used to infer the human activities. In recent years, many WiFi based activity recognition methods are proposed to recognize walking, falling, and even respiration [25], [26]. But most of the existing WiFi based activity sensing methods select a few distinctive subcarriers and use the time-series analysis techniques to extract the features individually on each selected subcarrier. However, violent behaviors is a complicated activity that involves a series of motions of whole body. Existing WiFi sensing based methods are hard to accurately recognize the violent behaviors based on manually selected time-series features.

## V. CONCLUSION

In this paper, we first design and implement a ubiquitous passive violent behaviors detection system called WiVi which leverages commercial off-the-shelf WiFi devices to detect the complicated violent actions. To fully extract the information in CSI and capture the accurate patterns of the complicated violent actions, WiVi utilizes both time-series features and correlated features. The correlated features extracted from combined subcarriers provide much more information to depict the violent actions. A PCA-based feature fusion method is leveraged to automatically combine the effective features together in WiVi. In practice, WiVi adopts a feedback mechanism that adapt to the operating environment changes by retraining the model when performance drops below a threshold. We implement a prototype of WiVi and evaluate its performance in various real environments. The experimental results demonstrate the effectiveness and robustness of WiVi.

## REFERENCES

[1] U.S. Department of Education. (2013). *Student Reports of Bullying and Cyber-Bullying: Results From the 2011 School Crime Supplement to the National Crime Victimization Survey*. [Online]. Available: https://nces.ed.gov/pubs2013/2013329.pdf
[2] C. P. Bradshaw, A. L. Sawyer, and L. M. O'Brennan, "Bullying and peer victimization at school: Perceptual differences between students and school staff," *School Psychol. Rev.*, vol. 36, no. 3, pp. 361–382, 2007.
[3] R. Veenstra, S. Lindenberg, A. J. Oldehinkel, A. F. De Winter, F. C. Verhulst, and J. Ormel, "Bullying and victimization in elementary schools: A comparison of bullies, victims, bully/victims, and uninvolved preadolescents," *Develop. Psychol.*, vol. 41, no. 4, pp. 672–682, 2005.
[4] UNESCO. (2017). *School Violence and Bullying—Global Status Report*. [Online]. Available: http://unesdoc.unesco.org/images/0024/002469/246970e.pdf
[5] E. B. Nievas, O. D. Suarez, G. B. García, and R. Sukthankar, "Violence detection in video using computer vision techniques," in *Proc. Int. Conf. Comput. Anal. Images Patterns*, 2011, pp. 332–339.
[6] K. Chang, "Wearable design for violent crime against children," in *Proc. Graduate School Creative Des. Eng.*, 2017.
[7] T. Giannakopoulos, D. Kosmopoulos, A. Aristidou, and S. Theodoridis, "Violence content classification using audio features," in *Proc. Hellenic Conf. Artif. Intell.*, 2008, pp. 502–507.
[8] J. M. Carlsmith and C. A. Anderson, "Ambient temperature and the occurrence of collective violence: A new analysis," *J. Pers. Social Psychol.*, vol. 37, no. 3, pp. 337–344, Jun. 2006.
[9] Y. Wang, K. Wu, and L. M. Ni, "WiFall: Device–free fall detection by wireless networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 581–594, Feb. 2017.
[10] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional neural networks based indoor Wi–Fi localization using channel state information," *IEEE Access*, vol. 5, pp. 18066–18074, 2017.
[11] F. Adib and D. Katabi, "See through walls with WiFi!" in *Proc. ACM SIGCOMM Conf. SIGCOMM*, 2013, pp. 75–86.
[12] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with Wi-Fi!" in *Proc. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 593–604.
[13] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 65–76.
[14] Z. Sun, S. Tang, H. Huang, L. Huang, Z. Zhu, H. Guo, and Y.-E. Sun, "iProtect: Detecting physical assault using smartphone," in *Proc. Int. Conf. Wireless Algorithms, Syst., Appl.* Qufu, China: Springer, 2015, pp. 477–486.
[15] J. R. Movellan. (2002). *Tutorial on Gabor Filters*. [Online]. Available: https://pdfs.semanticscholar.org/bae5/bae884633e7da2c1ec75d158f8849d2183d3.pdf
[16] L. M. P. M. C. Cotta and J. Gallardo, *Memetic Algorithms: A Contemporary Introduction*. Hoboken, NJ, USA: Wiley, 2016.
[17] J. Bosveld and D. Q. Huynh, "Boosted particle swarm optimization of Gabor filter feature vector," in *Proc. Int. Conf. Digit. Image Comput. Techn. Appl. (DICTA)*, Dec. 2012.
[18] Official Website of the Government of Newfoundland and Labrador. (2018). *Defining Violence and Abuse*. [Online]. Available: https://www.gov.nl.ca/VPI/types/#1
[19] J. Ha, J. Park, H. Kim, H. Park, and J. Paik, "Violence detection for video surveillance system using irregular motion information," in *Proc. Int. Conf. Electron., Inf., Commun.*, 2018, pp. 1–3.
[20] D. Wang, Z. Zhang, W. Wang, L. Wang, and T. Tan, "Baseline results for violence detection in still images," in *Proc. IEEE 9th Int. Conf. Adv. Video Signal-Based Surveill.*, Sep. 2012, pp. 54–57.
[21] S. Tao, M. Kudo, and H. Nonaka, "Privacy–preserved behavior analysis and fall detection by an infrared ceiling sensor network," *Sensors*, vol. 12, no. 12, pp. 16920–16936, Dec. 2012.
[22] Y.-L. Hsu, S.-C. Yang, H.-C. Chang, and H.-C. Lai, "Human daily and sport activity recognition using a wearable inertial sensor network," *IEEE Access*, vol. 6, pp. 31715–31728, 2018.
[23] Y. Ren, Y. Chen, M. C. Chuah, and J. Yang, "Smartphone based user verification leveraging gait recognition for mobile healthcare systems," in *Proc. Sensor, Mesh Ad Hoc Commun. Netw.*, 2013, pp. 149–157.
[24] Y. Chen and C. Shen, "Performance analysis of smartphone–sensor behavior for human activity recognition," *IEEE Access*, vol. 5, pp. 3095–3110, 2017.
[25] X. Zheng, J. Wang, L. Shangguan, Z. Zhou, and Y. Liu, "Smokey: Ubiquitous smoking detection with commercial WiFi infrastructures," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Apr. 2016, pp. 1–9.
[26] K. Qian, C. Wu, Z. Yang, Y. Liu, and K. Jamieson, "Widar: Decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi," in *Proc. ACM Int. Symp.*, 2017, pp. 1–10.

**LEI ZHANG** (Member, IEEE) received the Ph.D. degree in computer science from Auburn University, Auburn, AL, USA, in 2008. She was an Assistant Professor with the Computer Science Department, Frostburg State University, Frostburg, MD, USA, from 2008 to 2011. She is currently an Associate Professor with the College of Intelligence and Computing, Tianjin University, Tianjin, China, where she is also a member of the Tianjin Key Laboratory of Advanced Network Technology and Application. Her research interests include mobile computing, computer networks, and data mining. She is a member of the ACM.

**XIN RUAN** received the B.S. degree from the Tianjin University of Commerce, Tianjin, China, in 2017. He is currently pursuing the master's degree with Tianjin University, Tianjin. His research interests include wireless networks and data mining.

**JU WANG** is currently a Professor with the School of Biomedical Engineering, Tianjin Medical University. His research mainly focuses on the modeling of complex systems, especially the systems related to human.