FULL PAPER

ADVANCED
INTELLIGENT
SYSTEMS
Open Access

www.advintellsyst.com

Fusion of Wearable and Contactless Sensors for Intelligent Gesture Recognition

Xiangpeng Liang, Haobo Li, Weipeng Wang, Yuchi Liu, Rami Ghannam, Francesco Fioranelli, and Hadi Heidari*

A novel approach of fusing datasets from multiple sensors using a hierarchical support vector machine (HSVM) algorithm is presented. The validation of this method is experimentally carried out using an intelligent learning system that combines two different data sources. The sensors are based on a contactless sensor, which is a radar that detects the movements of the hands and fingers, as well as a wearable sensor, which is a flexible pressure sensor array that measures pressure distribution around the wrist. A HSVM architecture is developed to effectively fuse different data types in terms of sampling rate, data format, and gesture information from the pressure sensors and radar. In this respect, the proposed method is compared with the classification results from each of the two sensors independently. Herein, datasets from 15 different participants are collected and analyzed. The results show that the radar on its own provides a mean classification accuracy of 76.7%, whereas the pressure sensors provide an accuracy of 69.0%. However, enhancing the pressure sensors' output results with radar using the proposed HSVM algorithm improves the classification accuracy to 92.5%.

1. Introduction

During the past decade, novel multi-sensor data fusion mechanisms have been gaining attention due to the increased capabilities of sensing technologies and intelligent systems.[1–3] Generally, multi-sensor fusion improves a system's accuracy as a result of increased complexity.[1,2] Compared with single sensor systems, a multi-sensor system can observe an object from more than one perspective.[2–4] To accurately describe an object with multiple sensors, the fusion process aims to combine the strengths of each sensor and compensate for their relative weaknesses. In this respect, various data fusion strategies, mostly related to

machine-learning algorithms, have been proposed for different sensing purposes such as gesture, fault detection, intelligent robot, and health monitoring.[5,6]

In this article, the "objects" to be recognized are human gestures. The hand gesture is a natural way to interact between people, especially among those who have difficulty in speaking or hearing. Hand gestures are also important in human–computer interaction, particularly in situations where it is inconvenient to use speech or typical input devices.[6,7]

Using only wearable static sensors attached to the human body is unlikely to detect the full spectrum of hand gestures and might be perceived as uncomfortable. To address this limitation, a contactless sensor such as radar can be exploited as an enhancer to improve recognition accuracy and movement information. Combining these types of on-body and contactless sensors enables new methods in multisensory data fusion to emerge. In this article, a hybrid static and dynamic sensor system is proposed as a novel gesture-recognition approach. Here, "static" refers to gestures where a person's fingers are kept in specific positions, whereas "dynamic" refers to gestures involving transitions between two static gestures. In this regard, a hybrid intelligent system comprises of wrist-worn pressure sensors, with the addition of radar sensing, introduced and fused with the former to improve the overall recognition accuracy. Both sensors return time-dependent signals. However, for a natural sequence of human gestures, pressure sensor data are meaningful in a static state (fingers kept still), whereas radar data are more meaningful in a dynamic state (transition between static states).[7,8] These natural differences lead to the incompatibility of simultaneously fusing features extracted from the two sensors.[7,9] For the first time, a hierarchical support vector machine (HSVM) architecture was proposed to combine these features at different layers. Differently from our previous work that fuses results at decision level,[10] this article presents a purposely developed implementation of a multi-layer SVM classifier to incorporate together the diverse data from the two sensors by taking the confidence levels and the prediction labels from radar layer as an enhancer to improve the final result.

A schematic diagram depicting the fusion process and the potential applications is shown in **Figure 1**. In this architecture, the radar sensor acts as an enhancer instead of being used in

X. Liang, H. Li, W. Wang, Y. Liu, Dr. R. Ghannam, Dr. F. Fioranelli, Dr. H. Heidari
School of Engineering
University of Glasgow
Glasgow G12 8QQ, UK
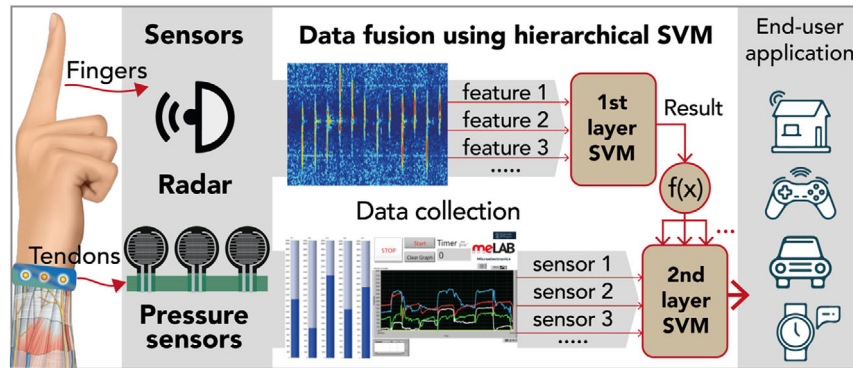E-mail: hadi.heidari@glasgow.ac.uk

Figure 1. The conceptual schematic of data fusion for gesture sensing with HSVM. The sensors are radar and resistive pressure sensor array, both with their own data acquisition systems. The proposed HSVM fuses completely different types of data sources to improve accuracy.

parallel with the pressure sensors because it only responds at the transition regions between gestures. The result of the first layer (radar) is fed to the second (pressure sensors) layer to improve its accuracy, which is where data fusion is achieved. In this procedure, the properties of linear SVM are fully used to optimize the training and recognition processes.

This article is organized as follows: Section 2 provides current state-of-the-art in multi-sensor fusion. Section 3 introduces our data acquisition methods and our experimental setup, as well as the preprocessing and feature extraction methods of radar data. Section 4 provides the building blocks of HSVM including SVM and directed acyclic graph (DAG) SVM. In Section 5, the HSVM is proposed and its performance is primarily tested. Next, the results of the HSVM architecture are presented in Section 6 and its enhancement to the original system is highlighted. Finally, concluding remarks are provided in Section 7.

2. Advances in Sensors' Fusion

State-of-the-art in the field of multi-sensor fusion demonstrated the feasibility of a complex system in achieving higher accuracy using data from multiple sensors.[1,4,11]. For the case of gesture recognition, vision-based sensors are known to be sensitive to background lighting and color, whereas movement-based sensors can be complementary as it is more immune to this problem. Combining these two types of sensors can therefore increase the overall accuracy of the gesture recognition.[12]

A fusion method to increase the accuracy of classification has already been proposed for a radar sensor and an inertial sensor in the context of detecting falls and classifying other human indoor activities. The fusion was carried out at different levels using SVMs and K-nearest neighbor (KNN). At the feature level, data from sensors were combined into a common feature vector sample. At the decision level, three approaches were used combining partial decisions and confidence level from different sensors, namely logarithmic opinion pool (LOGP), fuzzy logic, and a voting system. LOGP fusion cumulatively adds the confidence levels from different classifiers and converts them to posterior probability through a nonlinear logarithmic function. The final output is the class yielding the highest posterior probability. On the contrary, fuzzy logic first compares the confidence matrix of each

input and chooses the lower one from each class, then it selects the best number from this "worse" confidence matrix. Additionally, an election system that combines the outputs of four classifiers is proposed to provide subsequent improvement. When the decision clash happens, LOGP is used to mix the confidence level of radar and inertial sensors to generate an alternative prediction label. The accuracy improved significantly after the feature level and reached a maximum of ≈97.8% after decision fusion.[9]

Control through gestures has been used for playing video games using a combination of multi-channel electromyography (EMG) sensors and 3D accelerometer. To improve the accuracy, they segmented the EMG and accelerometer stream and extracted their feature for data fusion.[13] The two-stream hidden Markov model (HMM) was used for classification of the data from these two sensors. The probability of a pair of data is a combination of the probability from each sensor with weight factors assigned. The result of recognition was also determined by the maximum combination probability. The overall accuracy improved from 85.5% to 91.7%. Their manuscript presented a soft voting mechanism that is commonly used for data fusion.[13]

As two widely used deep-learning network, convolutional neural network (CNN) and long short-term memory (LSTM) have been applied to the depth sensor and the inertial data for action detection and recognition.[14,15] Actions such as stand, sit, and fall are captured by the multi-sensory system, followed by the classification using the deep learning-based fusion approach. Finally, the accuracy improved from 79.1% to 92.8%.[14]

Similarly, the CNN has been used for another multi-sensor system consisting of an optical sensor, a depth camera, and radar for developing a user interface during driving. A classifier was created by CNN after feature extraction. The performance of the sensor improved to ≈94%.[12] They also showed that the performance of SVM was not as good as the DNN method when the optical sensor provided unreliable data.[12]

Multiple SVMs have been used to classify and fuse the data from synthetic aperture radar imagery and the optical images.[4] After two SVMs were used to classify the two data sources separately, another SVM-based decision fusion generates the final result. The results show that the additional SVM method outperforms other classifiers and fusion methods, but the accuracy was not improved too much compared with the single SVM.[4]

In this article, an innovative HSVM approach that uses a multi-layer SVM structure to exploit the relations between the different sensors sources is proposed. In particular, the radar and the pressure sensors produce different features that cannot be combined in a conventional feature fusion approach, as the former detects transitions/changes of gestures, and the latter detects static gestures (i.e., the gesture at the end of the transition). Furthermore, this method presents an approach to precisely allocate the weight of each sensor source for final decision-making. This method is particularly good at dealing with situations where the classes of the two sensor sources are partially in agreement but cannot be directly used together as parallel and simultaneous inputs of classifiers. The proposed HSVM architecture is still able to fuse pressure sensors and radar data and perform a good improvement (from 69.0% to 92.5%). The training and testing are all based on multiple linear classifiers of SVM, which is less computationally intensive compared with other calculations such as convolutions. A comparison between state-of-the-art multi-sensory data fusion is summarized in **Table 1**. The proposed HSVM yielded a relatively high accuracy and improvement.

3. Data Acquisition and Experimental Setup

A measurement setup comprising of a set of five flexible resistive pressure sensors and an ultra-wideband (UWB) pulse-Doppler radar was developed as a proof-of-concept data collection platform. The experimental setup and graphical user interface (GUI) are shown in **Figure 2**a and Figure 2b. As two different data sources, pressure sensors and radar require their specific readout circuits, data acquisition tools, and GUI, which will be introduced in this section.[7] Their data will be then processed in MATLAB to verify the performance of the proposed classification and data fusion architecture.

3.1. Wearable Resistive Pressure Sensor Array

According to the literature, gesture recognition can be achieved by monitoring tendon movements around the wrist using an array of pressure sensors.[7,16,17] The five pressure sensors based on force-sensitive resistors are mechanically supported by a purposely designed wristband to make sure the subtle movement is detectable by the pressure sensors. It is worth noting that there are many factors that could affect the overall quality of the data, such as thickness, size, and flexibility of the sensors. The impact

of these factors can be attenuated by ensuring that the sensors are worn consistently across all gestures and participants in the data collection, and by developing machine-learning and multi-sensor fusion algorithms capable of capturing the subtle tendon differences encoded in the data. The commercially available FSR402 provides proper characteristics to meet the requirements of these experiments. The pressure sensors were embedded in EcoFlex flexible substrate together with fixing tape. For force-to-voltage conversion, each pressure sensor was integrated into a simple voltage divider with an additional 10 kΩ resistor, according to its datasheet.[7,17,18] Afterward, the voltage variation was captured by a microcontroller. The data-acquisition platform achieved a sampling rate of 278 Hz for each sensor, and 51 Hz was chosen for this work. The data received from the five sensors were subsequently processed in SVM as a five-dimensional vector in the next step.

In our previous work, a real-time wrist-worn gesture-capacitive pressure-sensing system with SVM was demonstrated.[7] The capacitive-based pressure sensors were chosen because the interaction between skin and the capacitive sensor provided an enhancement to the capacitive output.[16,19] However, capacitive pressure sensors are not stable enough for durable and long-term measurements because the voltage range drifts over time when it is attached to the skin. Thus, instead of capacitive sensors, resistive pressure sensors were used because of their stability, as well as simple readout circuit with a higher sampling rate.

3.2. Contactless Radar Sensing and Preprocessing

An off-the-shelf UWB pulse-Doppler radar (X4M300) was used in capturing range and velocity changes relevant to finger movement. The center frequency of the radar transmitter was 7.29 GHz with ≈1.5 GHz bandwidth at −10 dB. The integrated microstrip antenna transmitted the radar signals with a pulse repetition frequency (PRF) equal to 200 Hz, with simultaneous reception and digitization of amplitude and phase components at the receiver. In the experiment, the radar chip was connected to a laptop to acquire the data. This UWB-Doppler radar had a resolution in the range of centimeters, sufficient for detecting fine and subtle movements such as hand gestures, as opposed to human macro-movements or movements of man-made objects such as cars.[20] Furthermore, hands are difficult targets to detect due to their small size and typical weak reflections, which can be easily mixed with background noise.

Prior to feature extraction, raw radar data were filtered to remove the static clutter and emphasize moving targets. After

Table 1. A comparison of recent advances in multi-sensor fusion.

| Data source	Object	Fusion method	Accuracy before fusion [%]	Final accuracy [%]	Improvement [%]	Computational burden	Reference
EMG + accelerometer	Gesture	HMM	85.5	91.7	6.2	Medium	[13]
RGBD + Inertial	Human action	CNN + LSTM	79.1	92.8	13.7	Intensive	[14]
Optical sensor + depth camera + radar	Gesture	CNN	90.9	94.1	3.2	Intensive	[11]
Radar + optical imagery	Land cover	Multi-SVM	68.9	77.1	8.2	Low	[2]
Inertial + radar	Human activity monitoring	SVM + KNN	89.3	97.8	8.5	Intensive	[9]
Pressure sensors + radar	Gesture	HSVM (linear)	69.0	92.5	23.5	Low	This work
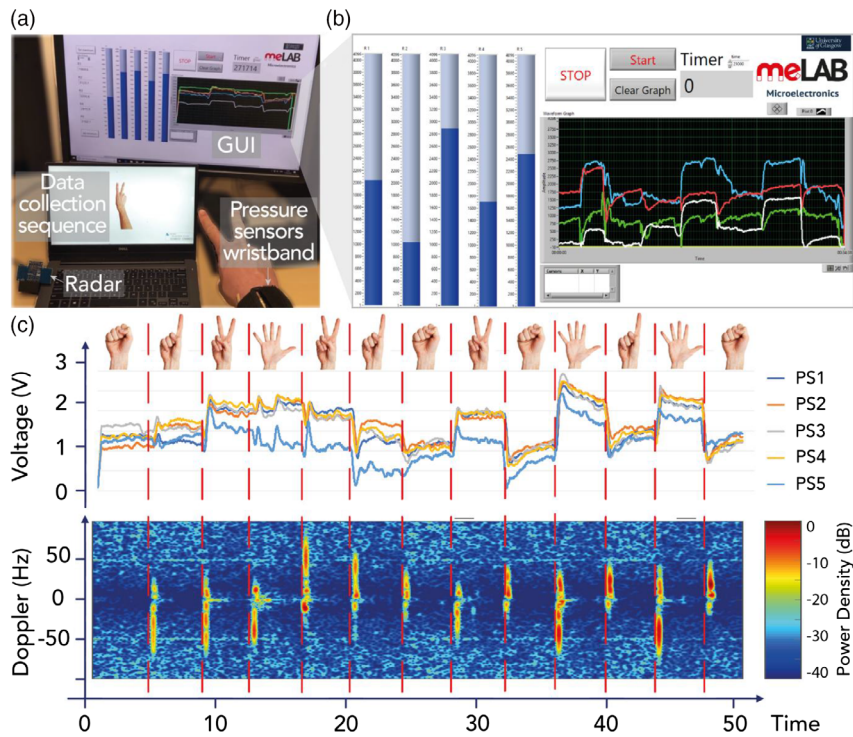
Figure 2. Experimental setup for the data collection. a) The experiment setup: at the same time, a radar sensor captures the motion of gesture. b) GUI developed in LabVIEW to visualize pressure sensors' data in real-time. c) Participant 5's training gestures and data over time as an example ("PS" in the figure denotes pressure sensor). The participant was required to show a sequence of gestures and keep fingers still for 4 s for each. It can be seen that different voltage combinations are measured for different gestures, but there are also similarities that may make the classification problem challenging. For the radar part, the Doppler signature corresponding to different transition parts are shown, where the positive Doppler corresponds to the movement approaching toward the radar, and vice versa. The color indicates the strength of the gesture movement with respect to different parts of the hand. The main misclassification takes places between those classes for which the directions of the movement are very similar, and the amplitudes of the recorded signals are also similar (e.g., 1 to 5 and 2 to 5).

that, short-time Fourier transform (STFT) with 2.5 s window size and 95% overlapping was applied on the range–time matrix to map the information into the Doppler-time domain. In this gestures' recognition scenario, characterizing fingers' trajectory was of interest rather than the static position of palm and fingers. This was done by exploiting the micro-Doppler effect, visible from the result of the STFT.[21] The equation of STFT is derived below

$$\text{STFT}\{x(n)\}(m,\omega) = \sum_{n=-\infty}^{\infty} x(n)\omega(n-mR)e^{-j\omega t} \tag{1}$$

where $x(n)$ is the input signal, $\omega(n)$ is the chosen window function (Hamming window in this case), R is the hop size, also known as an overlapping factor between successive fast Fourier transforms (FFTs). The result of the STFT operation was a Doppler-time matrix, whose absolute value is usually referred to as spectrogram and was used to extract features, i.e., significant parameters to represent relevant information for the classification process. The features extracted from the radar data are listed in **Table 2**, including those related to the centroid, bandwidth, and singular value decomposition (SVD) of the micro-Doppler matrix resulting from the STFT. The centroid of the Doppler signatures represents the center of the mass

of the movement over time. The bandwidth aims to find the extent of the energy around the Doppler centroid. In previous works, SVD was used to transform the original Doppler-time matrix into three individual matrices U, S, and V, where U and V are the left and right eigenvectors' matrix of the original micro-Doppler signature resulting from the STFT.[22]

However, some of the features may be redundant and cost extra computation loads. To select the optimal features' sets, sequential forward selection (SFS) was used on the original features' set to evaluate the classification performance of different feature combinations through an SVM classifier.[23] The SFS approach started from single features and progressively selected additional features among the possible combinations to maximize the classification performances; the algorithm stopped when there was no longer any improvement after adding more features. The SFS algorithm selected features on the basis of classification performance. More features were progressively selected from possible combinations to maximize performance. Naturally, no features were added when the classification performance plateaued. Five out of the 20 extracted features were selected by the SFS algorithm to construct an optimal feature set. This included the kurtosis and the mean of the Doppler bandwidth, the skewness of the Doppler centroid, the sum of all pixels of the right matrix derived from the SVD of

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

Table 2. Feature selection of radar data.

Radar features	Number of features
Mean, standard deviation, skewness, and kurtosis of the centroid of the Doppler spectrogram	4
Mean, standard deviation, skewness, and kurtosis of the bandwidth of the Doppler spectrogram	4
Two-dimensional mean, standard deviation, skewness, and kurtosis of the whole segment of the spectrogram	4
Mean and standard deviation of the first left and right eigenvector of the SVD of the spectrogram	4
Sum of pixels of the entire left and right matrices	2
Mean of the diagonal of the left and right matrices	2

the spectrogram, and the mean of the first column (eigenvector) of this matrix.

3.3. Data Collection

With the abovementioned experimental setup, the data of pressure sensors and radar were collected simultaneously from 15 individual participants. All the volunteers who took part in the experiments were given a comprehensive prior description of the experimental procedure and objectives, and their explicit consent was obtained prior to data collection. The number of gestures was four since three gestures were ideally recognizable (over 90% accuracy) in the previous work of this article.[7] In this study, the participants were required to perform four different static gestures, namely the numbers 0, 1, 2, and 5 with one hand, and the related transitions between these gestures, as shown in Figure 2c. The participants wore the wristband and performed the specific gestures in front of the radar at a distance of ≈ 50 cm. It is worth noting that the inherent different nature of radar and pressure sensors array make pressure sensors' data meaningful only in the static state when the fingers are kept still and the tendons create a pressure stimulus, whereas the radar returns almost blank response as it cannot detect well the fingers' reflections. In contrast, the radar will capture accurate information related to the movement of the fingers, in this case, the transition movements between static states.

Therefore, not only each static state needed to be collected but also each different transition. For example, the transition of gesture 1 (G1) to gesture 3 (G3) and G3 to G1 gave us a completely different response. In other word, the data of G1–G3 differed from G2–G3 even if both of them stopped at G3. Taking this into account to obtain training data for the four desired gestures, performing at least $A_N^2|_{N=4} = 12$ transitions was necessary. To cover all 4 static gestures and 12 dynamic transitions, the participants were required to perform 13 static states and keep their fingers still for 4 s for each state. In their second trial, where the data were used to test the accuracy of classification algorithms in the recognition step, participants performed 7 gestures and kept their fingers still for 4 s in each case.

An example (for Participant 5 in Section 5) of the training data set is shown in Figure 2c. Given a certain sequence of changing gestures to classify, the pressure sensors' data recorded the voltage levels for each finger, whereas the radar

data were processed to extract the micro-Doppler signatures through STFT. Features were then extracted from each transition region between static gestures.

4. Multi-Class Support Vector Machine

The SVM machine-learning algorithm is used for gesture recognition. First invented by Vapnik, SVM is a binary classifier that finds an optimal hyperplane to separate two groups of multi-dimensional data.[7,24] Later on, Platt, a researcher of Microsoft, USA, proposed a sequential minimal optimization (SMO) for efficiently training SVM, which simplifies the implementation.[25–27] Since a single SVM is a binary classifier, the DAG can be used to organize the relationship between each binary classifier to achieve multi-classification.[28,29] An example of SVM-based classification is visualized at the end of this section (**Figure 3**). Instead of using available SVM libraries, a purposely developed SVM package described in this section was implemented specifically for the experimental work in this article.

4.1. Linear Support Vector Machine

Assume that a data set $(\vec{x_1}, y_1), \ldots, (\vec{x_i}, y_i)$, $x \in R$, $y \in [+1, -1]$ is the input of an SVM, where $\vec{x_i}$ is the multi-dimensional input data and y_i is their label. Assume that $\vec{x_i}$ is a n-dimensional input vector, the aim of SVM is to find a n-dimensional hyperplane to separate the two different groups.[24] The hyperplane can be described by

$$W^T \vec{x_i} + b = 0 \tag{2}$$

where W^T is the normal and b denotes the bias. With label y_i, this hyperplane should satisfy

$$1 - y_i(W^T \vec{x_i} + b) \leq 0 \quad \text{for } i = 1, 2, 3 \ldots i \tag{3}$$

The points on the hyperplane $W^T \vec{x_i} + b = \pm 1$ are support vectors in Figure 3a. Suppose the margin between the two hyperplanes ($= \pm 1$) is $\rho = \frac{2}{||W||}$.[28] Thus, the problem is to minimize $||W||$ for the optimal hyperplane with largest margin ρ, which can be described as

$$\min_{w,b} \theta(w) = \frac{1}{2}||w||^2 = \frac{1}{2}W^T W \tag{4}$$

subject to Equation (3).

Here, Lagrange multipliers α_i are introduced to solve this quadratic programming problem.[28]

$$L(W, b, \alpha) = \frac{1}{2}W^T W - W^T \sum_{i=1}^{p} \alpha_i y_i \vec{x_i} + \sum_{i=1}^{p} \alpha_i \tag{5}$$

For the Lagrangian multiplier with inequality constraint, the Karush–Kuhn–Tucker (KKT) can be used as optimality conditions to solve it further.[24,25,28] After calculating partial derivative of the equation with respect to W and b, the dual problem of linearly separable samples is
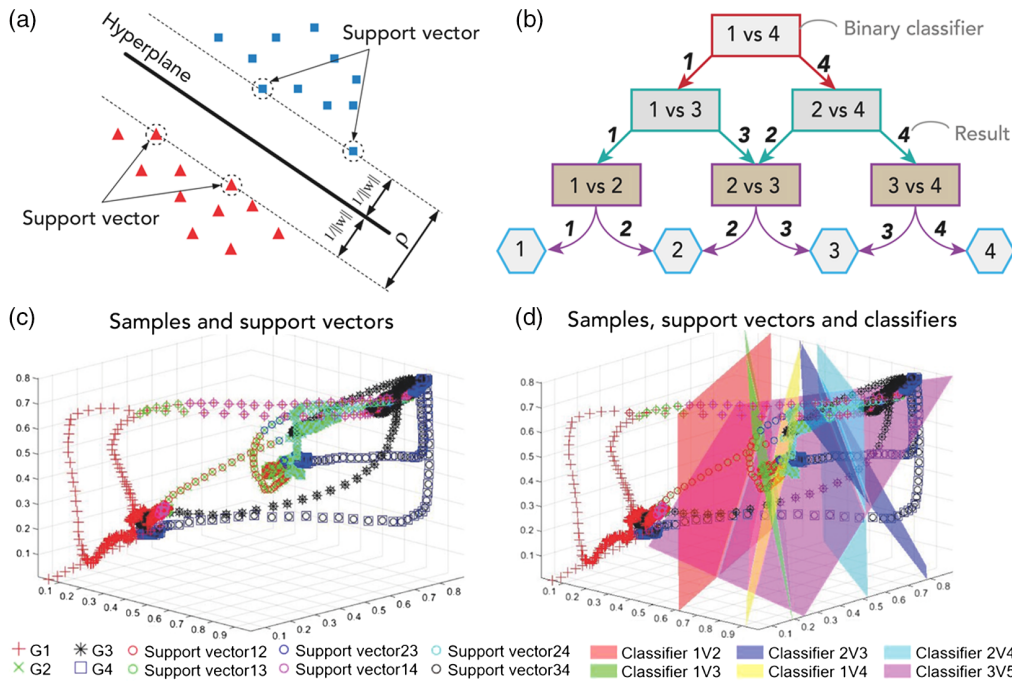
Figure 3. a) An example of 2D linear support vector machine including support vectors, samples, and the hyperplane separating them. b) DAG SVM consists of several "one-against-one" classifiers for multi-class classification. There are six classifiers for four classes classification. c) 3D data of pressure sensors of Participant 5. Note that the actual dimension of all data is 5, but this would not be displayable. After training, the support vectors for creating the classifiers are also highlighted. d) The classifiers are flat planes because they are derived from linear SVM. The classifiers roughly separate different groups of gesture data.

$$\max_{\alpha} W(\alpha) = L(W, b, \alpha) = \frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_i \alpha_j \gamma_i \gamma_j (\vec{x}_i)^{\mathrm{T}} \vec{x}_j \quad (6)$$

Satisfying

$$\alpha_i \geq 0 \quad (7)$$

$$\sum_{i=1}^{p} \alpha_i \gamma_i = 0 \quad (8)$$

The support vectors in Figure 3a are those samples with $\alpha_i > 0$ and they will be used to determine the position of the hyperplane (W and b).[24,28]

However, the collected samples are not supposed to be ideally separable by a linear hyperplane. The slack variable $\xi = (\xi_1, \xi_2, ..., \xi_m)$ and cost factor C are introduced to Equation (4) to allow for some data points located at unexpected locations, and avoid no solution scenario.[24,28] Afterward, the Equation (4) and Equation (3) are

$$\min_{w, \xi, b} \theta(W) = \frac{1}{2} ||w||^2 = \frac{1}{2} W^{\mathrm{T}} W + C \sum_i \xi_i \quad (9)$$

subject to

$$1 - \xi_i - \gamma_i (W^{\mathrm{T}} \vec{x}_i + b) \leq 0 \text{ and } \xi_i \geq 0 \quad (10)$$

After the same procedure as above, the $\max_{\alpha} W(\alpha)$ is remained but the constrains now are

$$0 \leq \alpha_i \leq C \quad (11)$$

$$\sum_{i=1}^{p} \alpha_i \gamma_i = 0 \quad (12)$$

4.2. Sequential Minimal Optimization

SMO is an algorithm that efficiently solves the Lagrangian multipliers of a quadratic programming problem by updating two of the multipliers continuously according to a certain procedure.[26,27,30] While other α_i are fixed, the two multipliers (α_1 and α_2) can be randomly chosen and optimized at each iteration. Therefore, Equation (6) can be written as

$$\max_{\alpha} W(\alpha) = L(W, b, \alpha) = \alpha_1 + \alpha_2 + \text{constant1}$$

$$- \frac{1}{2} \left(\alpha_1^2 \gamma_1^2 (\vec{x_1})^{\mathrm{T}} \vec{x_1} + \alpha_2^2 \gamma_2^2 (\vec{x_2})^{\mathrm{T}} \vec{x_2} + 2\alpha_1 \alpha_2 \gamma_1 \gamma_2 (\vec{x_1})^{\mathrm{T}} \vec{x_1} \right.$$

$$\left. + 2 \left(\sum_{i=3}^{m} \alpha_i \gamma_i (\vec{x_2})^{\mathrm{T}} \right) (\alpha_1 \gamma_1 \vec{x_1} + \alpha_2 \gamma_2 \vec{x_2}) + \text{constant2} \right) \quad (13)$$

Applying partial derivative to Equation (13) and equal to 0, the new α_2 can be obtained

$$\alpha_2^{\text{new}} + \alpha_2^{\text{old}} + \Delta 2 = \alpha_2^{\text{old}} - \frac{\gamma_2 (E_1 - E_2)}{\eta} \quad (14)$$

where prediction error $E_i = W^{\mathrm{T}} X + b_{\text{old}} - \gamma_i$ and $\eta = 2\vec{x_1} \vec{x_2} - \vec{x_1}^2 - \vec{x_2}^2$. Also, the feasible range of α_2^{new} needs to be

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

taken into consideration before updating it. Assume L and H are minimum and maximum, the feasible value of α_2^{new}, 1) if $\gamma_1 \neq \gamma_2$, $L = \max(0, a_2^{old} - a_1^{old})$, $H = \min(C, C + a_2^{old} - a_1^{old})$, 2) else $L = \max(0, a_2^{old} + a_1^{old} - C)$, $H = \min(C, a_2^{old} + a_1^{old})$.[7,25,26] Finally, $\alpha_2^{new, \text{ clipped}}$ and α_1^{new} can be obtained

$$\alpha_2^{new, \text{ clipped}} = \begin{cases} H, & \text{if} \quad H > \alpha_2^{new} \\ \alpha_2^{new} & \text{if} \quad L \leq \alpha_2^{new} \leq H \\ L, & \text{if} \quad \alpha_2^{new} < L \end{cases} \tag{15}$$

$$\alpha_1^{new} + \alpha_1^{old} + \Delta\alpha = \alpha_1^{old} - \gamma_1\gamma_2\Delta\alpha_2 \tag{16}$$

After that, the normal W can be updated. However, there are two new b due to the updated two KKT constraints. For rapidly updating, here the average number of them is calculated

$$\Delta b = \frac{(b_1 + b_2)}{2} = \frac{(E_1 + \gamma_1\Delta\alpha_1(\vec{x_1}^T)\vec{x_1}) + (E_2 + \gamma_2\Delta\alpha_2(\vec{x_1}^T)\vec{x_2})}{2} \tag{17}$$

One SMO loop is finished after updating W and b. This procedure is repeated until all samples satisfy KKT constraints and the optimal hyperplane can be obtained.[25–27] The incoming data can be classified by the sign of the classifier formula

$$f(\vec{x}) = \sin(W^T \vec{x_{in}} + b) = 0 \tag{18}$$

Instead of the linear classifier described above, the SVM has a superior performance in nonlinear classification by mapping both of the training and incoming data ($\vec{x_i}$ and $\vec{x_{in}}$) into higher dimensional space using kernel functions such as polynomial kernels ($K(\vec{x_i}, \vec{x_j}) = (\vec{x_i}, \vec{x_j + 1})^d$) and the Gaussian radial basis function $\left(K(\vec{x_i}, \vec{x_j}) = \exp\left(-\frac{||\vec{x_i}, -\vec{x_j}||^2}{2\sigma^2}\right)\right)$.[24,28] However, in this article, the linear SVM classifier is first analyzed because it suits well the HSVM where the distance between samples and hyperplanes is involved.

4.3. Directed Acyclic Graph SVM

A single SVM is a binary classifier. It is too complicated and computationally intensive to achieve multi-class classification with one SVM classifier.[7,24] A frequently used solution is the DAG SVM algorithm that consists a chain of binary SVM classifiers.[29,31] The architecture for the four gestures' problem is shown in Figure 3b. Each hyperplane is a "one-against-one" classifier.[29] Under this structure, $N(N-1)/2$ classifiers are required for multi-class classification, where N is the total number of categories.[7,31] Compared with "one-against-all" classifier in the literature, DAG is faster and has no overlap and unclassified situation.[29]

An example of the pressure sensors' training data of one of the participants has been shown in Figure 3c to demonstrate the details of the developed SVM implementation (namely, data from Participant 5 as later described in Section 6). Note that since the five-dimensional data are not displayable, three of them are taken in this section to show the sample distribution and the six classifiers generated by them.[7] After SMO, the support vectors of each classifier are also labeled. The hyperplanes are shown in

Figure 3d where the four different groups of the samples are roughly separated by each "one-against-one" classifier. The hyperplane is flat because the classifier is a linear SVM without kernel function. The same technique has also been applied on radar's data with respect to different features, number of classes (12), and number of binary classifiers (66).

5. Multi-Layer HSVM for Data Fusion

This section aims to present our proposed data fusion method for different types of data sources. As mentioned in Section 1, the data from the pressure sensors and the radar cannot be directly merged because their most significant data imply different gesture status (4 static states and 12 transition states, respectively). The classification results of pressure sensor and radar can be obtained by using separate SVM classifiers. Afterward, the prime difficulty is how to take the radar's results into the decision-making steps of pressure sensor data. Previous researches on multi-sensor fusion have been discussed in Section 2. For example, they addressed this issue by building a voting mechanism where several different machine-learning algorithms, such as SVM, neural network, KNN, etc., are applied on both the two types of data to find a final result using majority voting.[9] Implementing a voting mechanism for decision fusion brings the complication of finding the optimized structure and values of the weights of each individual classifier, especially if they provide different labels with different confidence levels. Furthermore, it is difficult to develop an implementation based on majority-voting that can suit different scenarios, where the correct label might be produced by a minority of classifiers.

With the proposed HSVM, this issue is avoided with the implementation of the second layer for fusion. The main idea of HSVM is to create two layers of SVM. In the second layer, the data of pressure sensors and weighted radar result are fused to get the final result.

5.1. Principle

In the first layer, as in **Figure 4**a, the extracted radar's features are the input, whereas the output is the classification results along with their scores. Each class has a unique score, which represents the confidence level of choosing this class when the classifier makes a decision. In this article, the confidence level is a $12 \times 6 \times 5$ cell array, where 12 represents the number of transition classes (all the possible classes for the radar defined at training stage), 6 corresponds to the number of observations (the 6 transitions observed in each testing dataset), and 5 indicates the number of participants. The formula for calculating the confidence level is shown below

$$f(x) = \sum_{j=1}^{n} a_j\gamma_j G(x_j, x) + b \tag{19}$$

where $f(x)$ is a function related to confidence level, a and b are the estimated SVM factors, $G(x_j, x)$ is the product of the predictor and the support vectors for jth class. The confidence level of each class in the form of the loss function is always <0; however, the

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
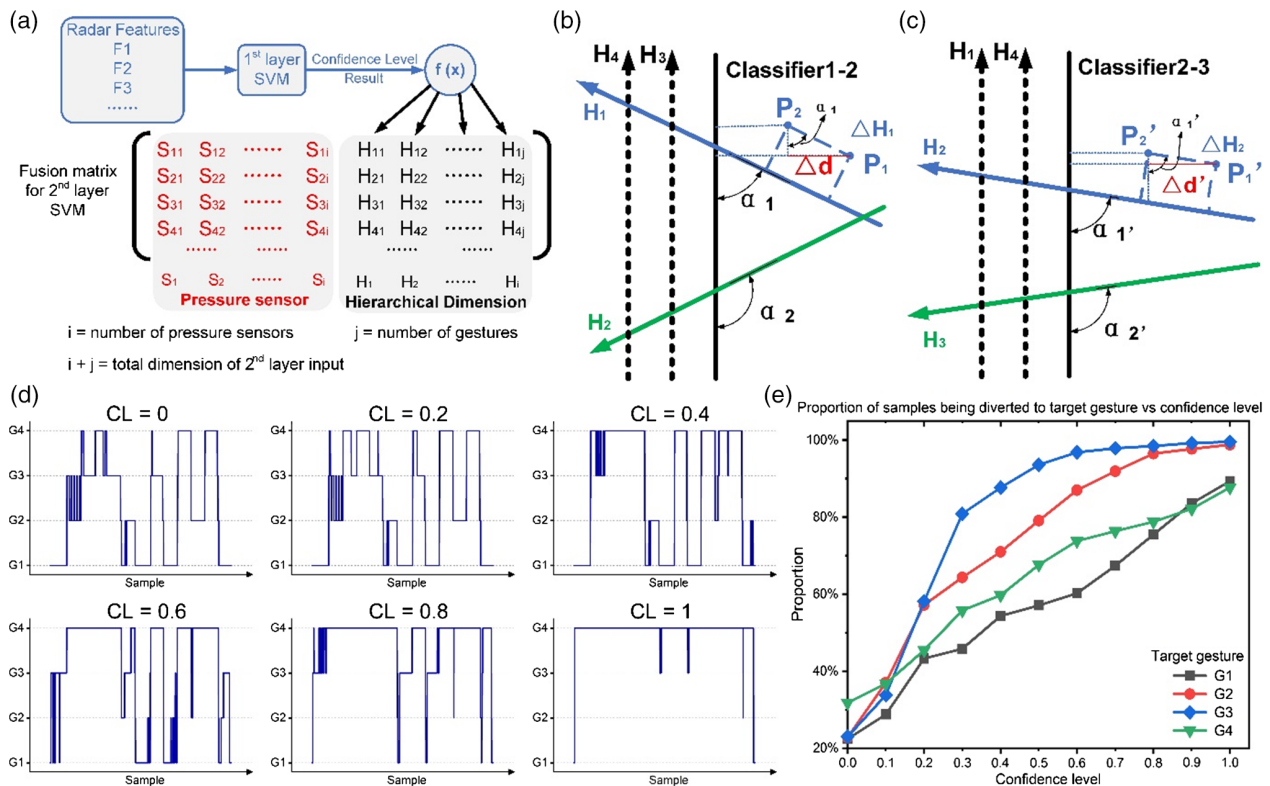SYSTEMS**
Open Access
www.advintellsyst.com

Figure 4. a) The input of second layer SVM is a $i + j$ dimensional vector. The schematic shows the elements of this matrix. The number of dimensions depends on the number of gestures and the number of sensors in the pressure sensor array. In addition to the pressure sensors' data, another j-dimensional hierarchical vector will be obtained from the result of radar. b,c) Two examples of the relative position between classifier and hierarchic dimensions in the case of four gesture recognition. After training, the angle between classifiers and hierarchic dimensions is different. This phenomenon leads to a different Δd. The α_1 in (b) is smaller than the α_2 in (c) and therefore $\Delta d < \Delta d'$ if $\Delta H1 = \Delta H2$. With this concern, the angle α should be taken into consideration when calculating the ΔH for uniform Δd. d) The performance evaluation of proposed HSVM. In the first graph, the result is exactly the same as the result without HSVM if CL = 0. Afterward, the CL value is increased by the step 0.2, where the ground truth target gesture is G4. As it can be observed, the result is gradually led to G4 and over 90% of the result is G4 when CL = 1. e) The proportion of correct target gestures over increased values of the parameter CL is also demonstrated. As expected, the four target gestures present rather similar but not ideally matched tendency. The reason is the locations of each data group are different. The closer the data group to the hyperplane, the easier the points of this group to be led to another side of hyperplane, which will change the predicted result.

class with confidence level closest to 0 means that it is highly possible to choose this class as the predicted label.

$$\mathrm{CL}(x) = 1 - |f(x)| \tag{20}$$

Using 1 to subtract the absolute value of the confidence level, the old confidence level is turned into new scores with its value normalized between 0 and 1.

The input matrix of the second layer SVM is the key element of data fusion. The proposed idea is shown in Figure 4a. First of all, the data of pressure sensors are on the left side. In this experiment, five resistive pressure sensors ($i = 5$) are embedded into the wristband and the corresponding data sets are S1...S5. In addition, there are four hierarchical dimensions on the right side for four recognizable gestures. By transforming the values in this hierarchical domain, this structure can exploit the radar's result as a weighted factor to affect the result of pressure sensors data. In other words, for example, if the first layer (radar) claims that the result is gesture 2 with very high (or low) confidence, the final result of the second layer will be

mostly diverted to gesture 2 (or remain to the result of pressure sensor).

The data operation of training and recognition will be introduced as follows.

5.2. Training

The two SVM layers are trained separately. The first layer follows the normal SVM procedure as described. SVM receives radar's features to train the classifiers. In the second layer, the training matrix consists of not only the data of pressure sensors but also the hierarchical dimensions. The values of the hierarchy dimension only depend on their label during training. The H_N is set to "1" scaled by a certain number if the label is N, whereas the other dimensions remain "0". The reason is that "virtual features" are desired for each gesture and, more importantly, this space will be filled in by CL values in testing. For example, if the label of pressure sensor data is gesture 3 ($H_3 = 1$) for four gesture recognition, the input matrix is

**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

$$[\, S_1 \quad S_2 \quad S_3 \quad S_4 \quad S_5 \quad 0 \quad 0 \quad 1/\text{scale} \quad 0 \,] \qquad (21)$$

where $S_1 \ldots S_5$ are pressure sensors' data. The scale value depends on the range of data set and cost factor of SVM. The choice of scale value will be discussed.

5.3. Testing for Recognition

The incoming data flow for recognition and input matrix of the second layer is shown in Figure 4a. First, the first SVM layer processes the radar data and outputs the results and confidence level. Assume that the number of static gestures is 4 and the result of the first layer is a transition from gesture x to gesture y ($G_x \to G_y$) with confidence level CL $= z$. Second, G_x is negligible because it is the past status, whereas G_y is kept as the final gesture state after a transition. G_y is also what we aim to recognize. All H values in the matrix are initialized to "1" according to the principle that if all the values in hierarchical dimension are set to be the same number as what it is in training ($H_N = 1$), the final result will not be influenced by the radar sensor as all its CL values will be the same. Third, when the confidence levels at the testing stage are received from the first layer SVM, H_y is expected to be close to 1 while the other values dropped. In other words, the yth dimension will be kept to the value closer to the training value "1/scale," whereas the others are dragged away from this number. Therefore, the final result will be diverted to G_y.

In this process, the angle between each classifier and hierarchical axis needs to be taken into consideration because the gradually diverted result is expected to perform coherently for every target gesture, i.e., given similar confidence levels, the results should not be drastically diverted. If the H value is reduced or increased by a consistent number, the classifiers that have larger angle with hierarchical axis would be more sensitive to the confidence level, whereas the similar sensitivity of every classifier is desired to unify the effect of confidence level for different gestures. This problem can be compensated by taking the angles into calculation. To explain this phenomenon, Figure 4b,c presents in a simplified sketch two examples of the relative position of classifiers (1–2 and 2–3) and four hierarchical axes after training. In the nine-dimensional space, all axes are orthogonal. As can be seen in Figure 4b, a stand-up side view of the classifier, an implication of the parallelism of the classifier and two H axes is their property that the fluctuation on H_3 and H_4 axes will not affect the result of Classifier 1–2. As an example, the angles between classifiers and axes after training Participant 5 in Section 6 are calculated and provided in **Table 3**.

Assume that P_1 is an incoming data that need to be classified. During the fusion with radar's result, the point is dragged along H_1 axis by ΔH_1 to P_2. The point gets closed to Classifier 1–2

$$\Delta d = \Delta H_1 \sin \alpha_1 \qquad (22)$$

For another classifier in Figure 4c, the α_1' is larger than α_1 when $\Delta H_1 = \Delta H_2$, which leads to a greater $\Delta d'$. However, greater Δd makes the data point easier to enter another side of the classifier. Therefore, Δd should be uniformed by changing ΔH. Furthermore, this value can be the CL from radar under the

Table 3. Angles between hierarchic axis and classifiers (Participant 5).

	H1 [°]	H2 [°]	H3 [°]	H4 [°]
Classifier 1–2	14.15	−14.15	0	0
Classifier 1–3	7.67	0	−7.67	0
Classifier 1–4	10.22	0	0	−10.22
Classifier 2–3	0	17.69	−17.69	0
Classifier 2–4	0	18.22	0	−18.22
Classifier 3–4	0	0	39.15	−39.15

condition that the uniform extent scale is obtained. The actual ΔH is

$$\Delta H_n = \frac{\text{CL}_n}{\sin \alpha_x} \qquad (23)$$

In summary, assuming that the target gesture is G3, the input vector of second layer SVM for nth classifier in the testing is

$$[\, S_1 \quad S_2 \quad S_3 \quad S_4 \quad S_5 \quad 1-\Delta H_n \quad 1-\Delta H_n \quad 1 \quad 1-\Delta H_n \,]$$
$$(24)$$

where the value in the third hierarchical axis remains while the others are reduced.

5.4. Performance Evaluation

The representative performance of the proposed HSVM is evaluated as following. The training process has been done by the training data of Participant 5, followed by manually giving the target gesture, which hints the result of radar, and varying CL to observe the tendency of the recognition result.

As can be seen in Figure 4d, at CL $= 0$ and target gesture $=$ G4, the result presents the same as the result without HSVM. In other word, in the case of CL $= 0$, the HSVM classifier considers only pressure sensor data since the radar's counterpart is not confident. What can be clearly seen in the figures is the gradual increase in the number of target gesture (G4) with increasing CL. The proportion of samples being diverted to target gesture for each target gesture is summarized in Figure 4e. It is getting more samples located in target gesture when CL is increased. The four tendencies are not ideally uniform due to their different initial coordinates. Some data points are closer to a classifier, but some points could be far away.

5.5. Fusion Parameters

In SVM, a crucial parameter is the cost factor in Equation (9). Also, scale value is introduced to HSVM in Equation (21). Note that a high-cost factor and scale value results in the situation that the location of each point will be very close to the hyperplane at the testing/recognition step. In this case, a minor difference in ΔH (CL) will lead to the class diverted. To avoid this situation, the cost factor and scale value should be kept within a reasonably small number. After all data of pressure sensors are normalized, the accuracy and improvement over varying cost factor and scale
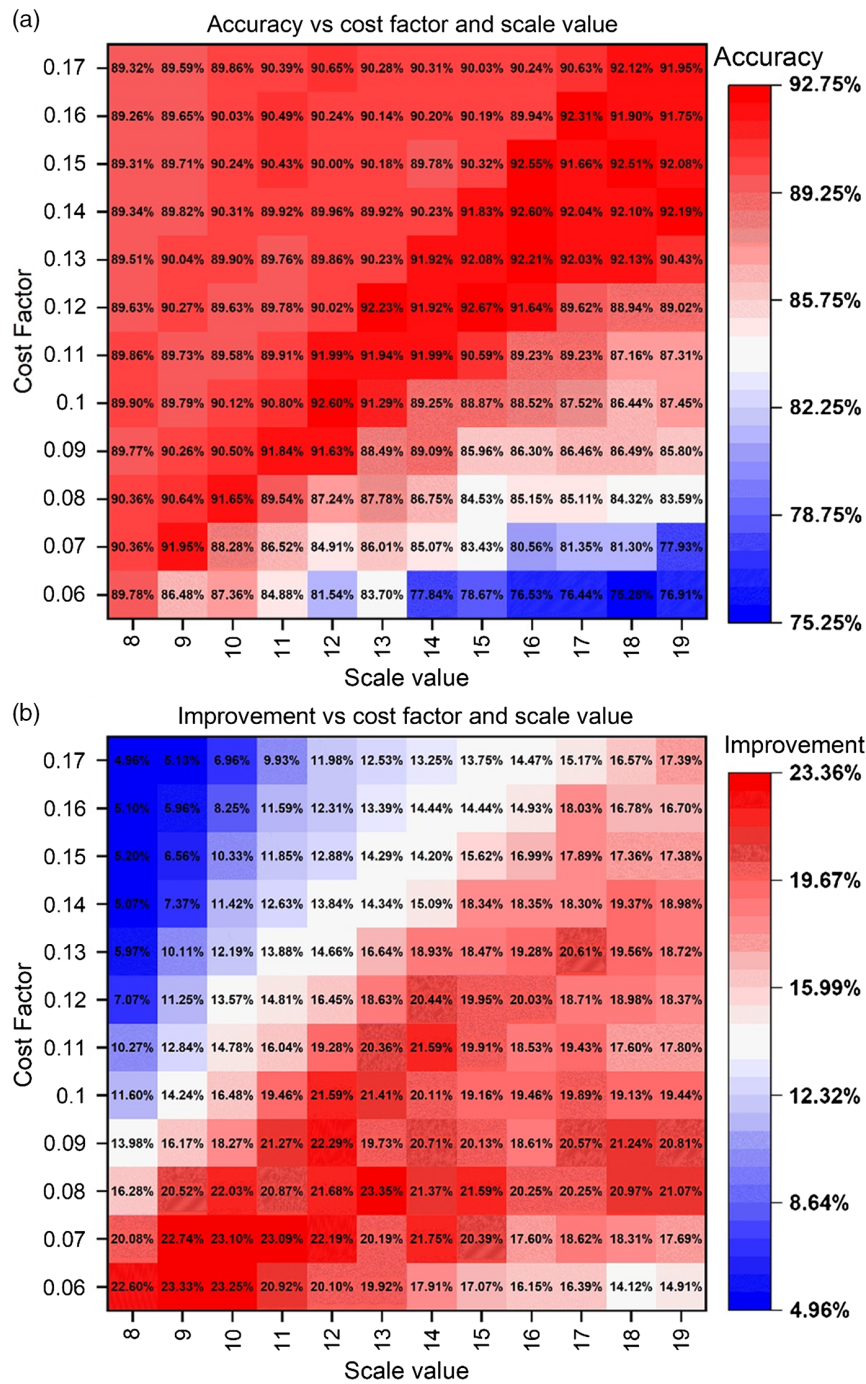
1900088 (9 of 13)

Figure 5. Results for (a) improvement of accuracy versus scale and cost factor as parameters and (b) final accuracy versus scale and cost factor. The highest values are highlighted. The values are the average number of 15 times simulation to improve reliability.

value are computed and shown in **Figure 5**a. From the graph, the highest accuracy (92.67%) occurs when scale = 15 and $C = 0.12$, while highest improvement is 23.35% at scale = 13 and $C = 0.08$. With the purpose of maximizing the accuracy, the parameters of the highest accuracy in Figure 5b are chosen for next step. The suitable situation of this scale and cost factor values should satisfy the following three conditions: 1) all data are normalized; 2) five data axes; and 3) four hierarchical axes.

6. Data Set and Experimental Results

6.1. Data Set

Section 3 introduced the methodology followed for the data collection, whereas Section 4 and 5 provided the data training and recognition techniques. Based on them, the data were collected with 15 participants, of which some were involved in previous

similar experiments (Participant 1, 3, 4, 7, and 10) and others were totally new to this. For each participant, two sets of data were collected for training and recognition, respectively, to mimic calibration and operation step in real applications. The participants were required to show 13 static states, containing 12 transitions in between, as training data, and 7 static states (with 6 transitions) in testing data. Every gesture was held for 4 s. To facilitate data processing, the sequence of all five participants was "G1-G2-G3-G4-G3-G2-G1-G3-G1-G4-G2-G4-G1" in training and "G1-G4-G3-G1-G3-G2-G4" in testing. An example of training data has been visually shown in Figure 2c. All data were labeled for supervised learning in training and result verification in testing.

To see the enhancement, SVM with same parameters is first applied on data from individual sensors separately, and then, HSVM is used to fuse them together.

6.2. Results of Pressure Sensor array

We used pressure sensors' data with a DAG SVM without hierarchical dimension. The total dimension was five for five pressure sensor inputs. In each 4 s gesture, the first 1 s (transition region) is removed to keep the data being processed stable. Six hyperplanes were created in a five-dimensional space as shown in Figure 3d. To be coherent with radar, the data were analyzed starting from the second static state gesture because radar has no response for the first one. The overall accuracy of pressure sensors' data is 69.0% on average. By combining all subjects' data and taking the average, the confusion matrix of the four gestures is shown in **Figure 6**a. The classifiers precisely separate G1 from the others but are often confused by G3 and G4.

6.3. Radar Results

From the classification results of radar in Figure 6b, some significant errors occur between Class 1 and 5, 2 and 3, 3 and 5, due to the very small range and velocity difference between each transition. The average results of "training and testing on the same group of participants" are ≈66.1%. A further + 10.6% improvement is possible using suitable feature selection techniques. Therefore, only using radar data, the result can yield an accuracy of 76.7%.

6.4. Fusion Results

6.4.1. Fixed CL Value (First Trial)

Fusion of the two sensors' data was tested by two methods. In the first trial, the results of the first layer was passed to next layer without their CL. Instead, the CL was manually set within the second layer from 0 to 1 by steps of 0.1. In addition to analyzing the maximum accuracy, the parameter of maximum improvement (scale = 13 and $C = 0.08$) was also tested and discussed. The case of the highest improvement shows that the HSVM can bring the overall accuracy to over 90% even when the result of only pressure sensors is low. The improvement of accuracy compared with the result of pure pressure

sensor array and the total accuracy can be observed in Figure 6c. The parameters of scale = 15 and $C = 0.12$ were tested with the concern of maximizing accuracy. From the chart, it can be found that the highest accuracy and largest improvement occur when CL = 0.4. Before this point, radar's result increased the final accuracy, and subsequently, the fusion result followed closely the radar's result. The highest improvement and accuracy are 14.7% and 90.5%, respectively, at CL = 0.4. When scale = 13 and $C = 0.08$, it can be seen that the HSVM was able to reach a 20% improvement. However, due to its lower final accuracy compared with the scenario of scale = 15 and $C = 0.12$, the scenario of highest accuracy was tested in the second trial.

6.4.2. CL Generated from Radar Layer SVM (Second Trial)

In the second trial, the CL was received from radar and normalized to 0 to 1 as CL instead of manually setting the value. Figure 6d provides the confusion matrix of the final result. The bottom right of the matrix shows that the misclassification of G3 and G4 in the testing of only pressure sensors' data was significantly compensated by the radar enhancer. The radar sensor is more sensitive to the G4, corresponding to gesture "five" because more fingers move.

The result of each participant and each gesture in this trial is shown in **Table 4**. Most cases were improved to at least 90%, especially for Participant 1 and 5, whose data were drastically improved to 97.8% and 99.4%, respectively. Interestingly, the accuracy of enhancer of Participant 1 is the lowest one but yields the second largest enhancement, which is diametrically opposite to Participant 3. The main reason is the first layer of Participant 1 generates a lower accuracy, but more precise CL value compared with Participant 3. However, no increase in Participant 9 was found, and its accuracy of both before and after fusion yields an obviously low value, which was assumed caused by nonideal collection of the data. Overall, the average accuracy of all participants is 92.5%, which is slightly higher than the highest accuracy in the first trial. However, the CL value in the first trial is manually set and then the highest point was found. More significantly, taking the CL from the radar is an automatic operation without the need of analyzing the performance in advance or overfit the process to a specific scenario or group of users, which is more realistic and beneficial for practical applications of the system. Finally, Figure 6e,f shows the agreement between the result and ground truth before and after fusion. The result after fusion ideally matches the ground truth with a minor misclassification.

To summarize, the advances of this work compared with the other state-of-the-art multi-sensor fusion have been highlighted in Table 1. Although it is difficult to find an ideally comparable case (fusion of dynamic and static sensor for gestures), this work can provide a significant improvement and increase the accuracy by using only multiple linear classifiers of SVM.

7. Conclusion and Discussion

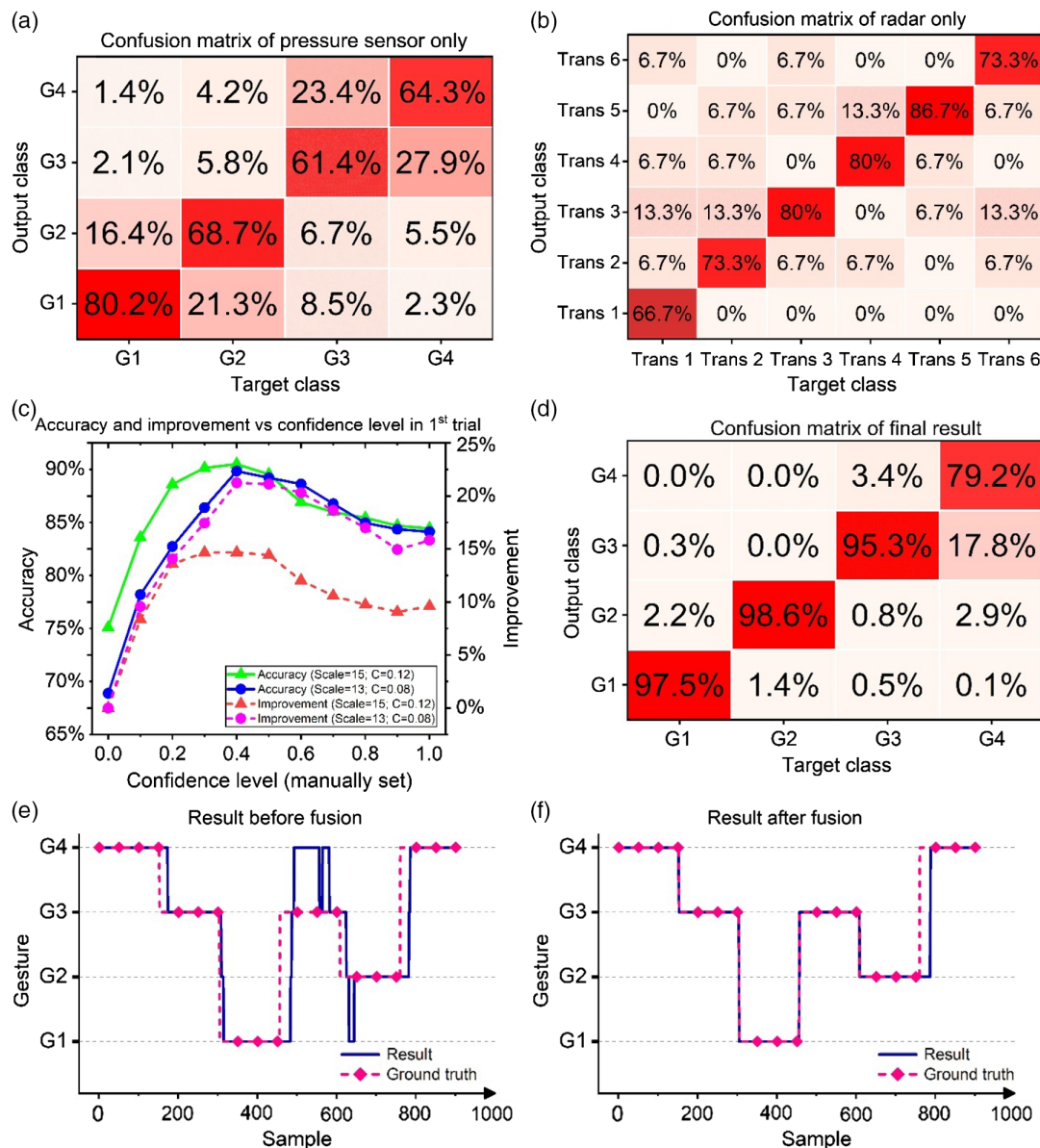This article demonstrates the performance of an HSVM fusion algorithm for the simultaneous combination of wearable and

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
INTELLIGENT
SYSTEMS**
Open Access

www.advintellsyst.com

Figure 6. a) Confusion matrix of pressure sensors only. b) Confusion matrix of radar only. In confusion matrices, the diagonal elements are the events correctly classified, whereas the nondiagonal elements are the misclassifications. c) The result in first trial, both cases of the highest accuracy and improvement is tested. d) Confusion matrix after data fusion. Finally, the classification results of the Participant 1 (e) before and (f) after fusion are compared.

contactless sensors. Unlike other multi-sensory systems, the fusion takes place neither at the feature level nor the decision level, but in between. The overall accuracy of the proposed HSVM learning system was improved by adding a radar as an enhancer to the pressure sensors. This resulted in an average improvement of 23.5%, such that the classification accuracy reached 92.5%. To further demonstrate the capabilities of HSVM, the parameter of the highest improvement was tested because the parameter of highest accuracy requires greater computational time and processing power. The results showed that the highest improvement can reach 19.2%. Moreover, the HSVM was still able to improve its accuracy to >90%

when the classification accuracy of the pressure sensors' data was <75%.

The experimental analysis and results from 15 individual participants confirmed that the HSVM could be a promising approach in organizing different sources of data in a flexible and scalable multi-sensor intelligent system. Further investigations of the HSVM technique may involve 1) increasing the number of layers or sensors to combine more data sources; 2) analytically describing and modeling the effect of different parameters; 3) investigating the computational and implementation speeds; and 4) examining other machine-learning approaches in addition to SVM in a hierarchical structure.

Table 4. Accuracy for each participant in the testing trial.

	Result of pressure sensors only [%]	Result of fusion [%]	Accuracy improvement [%]
Participant 1	81.2	97.0	15.7
Participant 2	38.0	79.9	41.9
Participant 3	87.3	91.0	3.7
Participant 4	86.9	97.7	10.8
Participant 5	83.6	95.1	11.5
Participant 6	75.0	99.8	24.8
Participant 7	78.5	87.5	9.0
Participant 8	83.5	100.0	16.5
Participant 9	56.0	52.4	−3.5
Participant 10	86.9	100.0	13.1
Participant 11	37.2	99.6	62.4
Participant 12	78.2	90.2	12.0
Participant 13	59.0	97.5	38.6
Participant 14	52.4	100.0	47.6
Participant 15	51.3	99.8	48.6
Average	69.0	92.5	23.5

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Keywords

[1] M. Liggins I. I., D. Hall, J. Llinas, *Handbook of Multisensor Data Fusion: Theory and Practice*, 2, CRC press, Boca Raton, FL **2017**.
[2] B. Khaleghi, A. Khamis, F. O. Karray, S. N. Razavi, *Inf. Fusion* **2013**, *14*, 28.
[3] J. Liu, M. Gong, K. Qin, P. Zhang, *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 545.
[4] B. Waske, J. A. Benediktsson, *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3858.
[5] a) G. Fortino, S. Galzarano, R. Gravina, W. Li, *Inf. Fusion* **2015**, *22*, 50; b) W. Li, G. Wei, D. W. Ho, D. Ding, *IEEE Trans. Neural Netw. Learn. Syst.* **2018**; c) R. C. Luo, M.-H. Lin, R. S. Scherp, *IEEE J. Robot. Autom.* **1988**, *4*, 386; d) J. Hang, J. Zhang, M. Cheng, *IET Renew. Power Gener.* **2014**, *8*, 289.
[6] C. Tawk, M. in het Panhuis, G. M. Spinks, G. Alici, *Adv. Intell. Syst.* **2019**, *1*, 1900002.
[7] X. Liang, R. Ghannam, H. Heidari, *IEEE Sens. J.* **2019**, *19*, 1082.
[8] H. Li, A. Shrestha, H. Heidari, J. L. Kernec, F. Fioranelli, *IEEE Sens. J.* **2018**, 1.
[9] H. Li, A. Shrestha, H. Heidari, J. Le Kernec, F. Fioranelli, *IEEE J. Electromagn. RF Microwaves Med. Biol.* **2018**, *2*, 102.
[10] H. Li, X. Liang, A. Shrestha, Y. Liu, H. Heidari, J. Le Kernec, F. Fioranelli, *IEEE J. Electromagn. RF Microwaves Med. Biol.* **2019**.
[11] K. Liu, C. Chen, R. Jafari, N. Kehtarnavaz, *IEEE Sens. J.* **2014**, *14*, 1898.
[12] P. Molchanov, S. Gupta, K. Kim, K. Pulli, in *IEEE Int. Conf. Workshops Autom. Face Gesture Recognit.*, Vol. 1, IEEE, Ljubljana, **2015**, p. 1.
[13] X. Zhang, X. Chen, W.-H. Wang, J.-H. Yang, V. Lantz, K.-Q. Wang, in *Int. Conf. Intell. User Interface*, ACM, **2009**, p. 401.
[14] N. Dawar, N. Kehtarnavaz, *IEEE Sens. J.* **2018**, *18*, 9660.
[15] J. L. Kernec, F. Fioranelli, C. Ding, H. Zhao, L. Sun, H. Hong, J. Lorandel, O. Romain, *IEEE Signal Process. Mag.* **2019**, *36*, 29.
[16] X. Liang, H. Heidari, R. Dahiya, in *IEEE New Gener. Circuits Syst.*, Genova, Italy **2017**, p. 181.
[17] A. Dementyev, J. A. Paradiso, in *Proc. Annu. ACM Symp. User Interface Software Technol.*, ACM, Honolulu, Hawaii, USA **2014**, p. 161.
[18] H. Heidari, E. Bonizzoni, U. Gatti, F. Maloberti, R. Dahiya, *IEEE Sens. J.* **2016**, *16*, 8736.
[19] K. O. Htet, H. Fan, H. Heidari, in *IEEE Int. Symp. Circuits Syst. (ISCAS)*, Florence, Italy **2018**, pp. 1–5.
[20] a) R. Ricci, A. Balleri, *IET Radar Sonar Navig.* **2015**, *9*, 1216; b) F. Fioranelli, M. Ritchie, H. Griffiths, *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1933; c) A. Shrestha, J. Le Kernec, F. Fioranelli, E. Cippitelli, E. Gambi, S. Spinsante, in *Int. Conf. Radar Syst.*, IET, Belfast, **2017**, p. 6; d) M. S. Seyfioğlu, A. M. Özbayoğlu, S. Z. Gürbüz, *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 1709.
[21] V. C. Chen, W. J. Miceli, D. Tahmoush, *Radar Micro-Doppler Signatures: Processing and Applications*, The Institution of Engineering and Technology, Hertfordshire **2014**.
[22] J. J. M. De Wit, R. I. A. Harmanny, P. Molchanov, in *Int. Radar Conf.*, IEEE, Lile, **2014**, p. 1.
[23] S. Z. Gürbüz, B. Erol, B. Çağlıyan, B. Tekeli, *IET Radar Sonar Navig.* **2015**, *9*, 1196.
[24] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, NY **2000**.
[25] J. Platt, *Microsoft Research*, MSR-TR-98-14, April **1998**.
[26] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, K. R. K. Murthy, *Neural Comput.* **2014**, *13*, 637.
[27] F. Cai, V. Cherkassky, *IEEE Trans. Neural Netw. Learn. Syst.* **2012**, *23*, 997.
[28] N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines: and Other Kernel-Based Learning Methods*, Cambridge University Press, Cambridge, UK, **2000**.
[29] P. Chen, S. Liu, in *Fifth Int. Conf. Nat. Comput.*, Tianjin, China **2009**, pp. 460–462.
[30] J. F. Wang, S. Peng, J. C. Wang, P. C. Lin, T. W. Kuan, in *IEEE Int. Conf. Syst. Man Cybernetics*, IEEE **2011**, pp. 1621–1625.
[31] Y. Xue, Z. Ju, K. Xiang, J. Chen, H. Liu, *Appl. Sci.* **2017**, *7*, 358.