EasyChair Preprint

№ 10371

Weighted Fusion of Multiple Classifiers for Human Activity Recognition

Kezhu Zuo, Xinde Li, Jean Dezert and Yilin Dong

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

June 10, 2023

Weighted Fusion of Multiple Classifiers for Human Activity Recognition

Kezhu Zuo
School of Cyber Science and Engineering, Southeast University, Nanjing, China.
kezhuzuo@seu.edu.cn

Xinde Li
Corresponding Author
Key Laboratory of Measurement and Control of CSE, School of Automation, Southeast University, Nanjing, China.
Email: xindeli@seu.edu.cn

Jean Dezert
The French Aerospace Lab ONERA DTIS/MIDL Palaiseau, France.
Email:jean.dezert@onera.fr

Yilin Dong
College of Computer Science and Engineering Shanghai Maritime University Shanghai, China.
Email:yldong@shmtu.edu.cn

***Abstract*—Human Activity Recognition (HAR) based on wearable device has become a hot topic of research due to its wide range of applications in health-care, fitness and smart homes. However, the classification of some activities with similar sensor readings, such as standing and sitting, is usually more challenging for the design of efficient activity recognition algorithms. Considering the inconsistent performance of different classifiers, which can provide information complementary for individual classifier, we propose a novel multi-classifier fusion method based on belief functions (BFs) theory for HAR. Specifically, at first, four classifiers are trained using time-domain and frequency-domain features to obtain basic belief assignments (BBA) of activity, respectively. Then, three assessment criteria are utilized to evaluate the reliability of the classifiers and a scoring matrix is constructed. Next, the algorithm of Belief Function based the Technique for Order Preference by Similarity to Ideal Solution (BF-TOPSIS) is employed to calculate the weighting coefficients for each classifier. Finally, the discounting and Dempster's rules are adopted to combine the multiple classifiers and further decision making. Several experiments were conducted to illustrate the performance of the proposed method using the UCI smartphone dataset, and the results show that the proposed method is more accurate than the state-of-art methods.**

***Index Terms*—Belief functions theory, multiple classifiers fusion, BF-TOPSIS, human activity recognition.**

I. Introduction

With the booming development of micro-sensor technology, Human Activity Recognition (HAR) based on wearable sensors has become one of the hot research topics [1], [2]. Data of daily activities can be well collected in an all-round and non invasive discrete manner using accelerometers, gyroscopes and other such portable wearable devices, so as to accomplish the work of assisted living and health monitoring while effectively protecting the privacy of users [3]. Obviously, it has certain advantages compared to traditional vision-based methods. However, the accuracy of HAR based on wearable devices is affected by many factors, such as the number and the deployment location of sensors, the complexity of activities [4], and so on. Due to the uncertainty, diversity and individual differences of activities [5], many scholars took the perspective of multi-sensor information fusion to achieve higher accuracy of HAR. For example, Dong et al. [6] developed the kernel density estimation models to fit the multi-sensor data to obtain the basic belief assignments (BBAs), and then Dezert-Smarandache theory (DSmT) was adopted to combine the acquired BBAs. Uddin et al. [7] fused data from different multimodal sensors with statistical features of different orders and then trained a deep recurrent neural network (RNN) for activity recognition. Although they achieve good accuracy, it is still difficult to accurately identify some activities with high similarity of sensor readings such as sitting and standing. Furthermore, the reliability of activity recognition can be significantly compromised when sensor readings are missing or disturbed by noise without additional sensor information.

Recently, the multi-classifier fusion has been applied in pattern recognition [8], information fusion [9], [10] and other fields, especially for classification problems in complex environments. Different classifiers can learn different feature information, and multiple classifiers can provide complementary information compared with any individual classifier, which can help identify similar human activities such as sitting and standing. By using multi-classifier fusion, we expect the improvement of the classification accuracy, which brings the possibility of high precision HAR. On the other hand, multiple classifiers can be seen as multiple sources of evidence, and we fuse the basic belief assignments (BBAs) of the human activity categories output by the classifiers.

The multi-classifier fusion usually consists of generating membership classifiers, applying combination rules, and make a decision about the positioning of the patient. Various approaches have been proposed for membership classifier generation, for example, using different training samples, different features and different types of classifiers [11]. Common classifier fusion methods include voting method [12], naive Bayes [13], Dempster-Shafer (DS) rule in Dempster-Shafer theory (DST) [14], and so on. In the fusion process, the classifiers may have different reliabilities (weights) and their decision results may be contradicting, which inevitably brings conflict issues. In order to improve classification accuracy, it becomes particularly important to evaluate the reliability of classifiers before combining them. For instance, Liu et al. employed contextual reliability evaluation based on inner reliability and relative reliability concepts [10]. Dong et al. [15] took two classes of criteria into account to evaluate the classifiers. The first class is the conflict between the classifiers and the second class is the imprecision of the information provided by each classifier. The effective evaluation of the reliability of multiple classifiers and their fusion is a challenging problem for HAR tasks.

In this article, we propose a novel Weighted Fusion of Multiple Classifiers (WFMC) method for HAR based on BFs theory. Our main contributions are summarized as follows:

- Four classifiers including support vector machine (SVM), random forest (RF), multi-layer perceptron (MLP) and logistic regression (LR) are trained by same training dataset for acquiring BBAs of human activities. To improve the multi-classifier fusion accuracy, Belief Jensen–Shannon (BJS) divergence, Interval distance function and belief entropy are considered to measure the reliability of the classifiers and a scoring matrix is constructed.
- The BF-TOPSIS[1] multi-criteria decision-making algorithm is employed to calculate the weighting coefficients for each classifier, and multiple classifiers are fused using discounting technique and DS rule in this work, the final decision is made based on the maximum belief mass of all involved single focal elements.
- We evaluate the performance of our proposed method on the widely used UCI Smart-phone public dataset.

The rest of this article is organized as follows. Section II presents the basic concepts of BFs theory, discounting technique and pignistic probability transformation. Section III provides a detailed description of the new proposed multi-classifier fusion strategy for HAR. Section IV presents the detailed experimental results and discussions. The final section V gives concluding remarks with some perspectives of this work.

II. Preliminaries

A. Belief Functions Theory

BFs theory (known also as DST) has been widely used in multi-sensor information fusion due to its ability to deal with uncertain and imprecise information [17]. The basic concepts are introduced in this section based on [14]. Let Θ be a finite set of elements denoted by

$$\Theta = \{\theta_1, \theta_2, ..., \theta_n\}. \tag{1}$$

The set Θ is called a frame of discernment (FoD), which consists of exhaustive and exclusive hypotheses. Information sources distribute mass of belief to elements of the power set of the FoD, denoted by 2^Θ. For example, if $\Theta = \{\theta_1, \theta_2\}$, then

$$2^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2\}. \tag{2}$$

A BBA, called a mass function, is defined by the mapping $m(\cdot) : 2^\Theta \mapsto [0, 1]$, which satisfies $m(\emptyset) = 0$ and

$$\sum_{A \in 2^\Theta} m(A) = 1. \tag{3}$$

For a proposition $A \subseteq \Theta$, the belief function is defined as:

$$Bel(A) = \sum_{B \subseteq A, B \in 2^\Theta} m(B). \tag{4}$$

The plausibility function is defined as:

$$Pl(A) = \sum_{B \cap A \neq \emptyset, B \in 2^\Theta} m(B). \tag{5}$$

If the focal elements of BBA are all singletons, the BBA is called Bayesian BBA [14]. In pattern classification, $m(A)$ represents the support degree of the object associated with class. For example, if A is a set of classes (e.g., $A = \{\theta_1, \theta_2\}$), $m(A)$ denotes the possibility of classification among the class θ_1 and θ_2 with respect to the object. In DST, the classical Dempster's rule (also called Dempster-Shafer rule, or just DS rule) is used to combine two (or more[2]) independent Sources of Evidence (SoEs), which is denoted as $m_1 \oplus m_2$ and defined as follows [14]: for $\forall A \in 2^\Theta, A \neq \emptyset$,

$$(m_1 \oplus m_2)(A) = \frac{1}{1-k} \sum_{B,C \in 2^\Theta | B \cap C = A} m_1(B) m_2(C) \tag{6}$$

with

$$k = \sum_{B,C \in 2^\Theta | B \cap C = \emptyset} m_1(B) m_2(C) \tag{7}$$

where k represents the total conflict degree. If $k = 1$, it implies that the two SoEs are in total conflict, and the DS rule cannot be applied because of division by zero.

B. Classical Discounting Technique

The SoEs may have varying degrees of reliability due to their different abilities of classification. The discounting operations are frequently conducted by using a discounting factor α for each source of evidence. A particular discounting operation has been introduced by Shafer [14] for the combination of SoEs with different degrees of reliability. Shafer discounts the masses of all focal elements by a discounting (weighting) factor $\alpha \in [0, 1]$ to the total ignorance. Each discounted BBA characterizing each discounted source of evidence is used in the fusion process. More precisely, for $\forall A \in 2^\Theta \backslash \{\Theta\}$, the discounted mass of discounted source of evidence is defined as follows:

$$\begin{cases} m^\alpha(A) = \alpha \cdot m(A) \\ m^\alpha(\Theta) = 1 - \alpha + \alpha \cdot m(\Theta) \end{cases} \tag{8}$$

where $\alpha = 1$ means that the SoE is completely reliable, and $\alpha = 0$ means that the SoE is completely unreliable.

C. Pignistic Probability Transformation

When multi-source information is combined, there may be disjunctive focal elements with strictly positive mass of belief. It is worth noting that the final decision is made only among singleton focal elements. Classically, a BBA is usually transformed into a (possibly subjective) probability measure

[1]BF-TOPSIS is an extension of the technique for order preference by similarity to ideal solution (TOPSIS) based on belief functions (BF) [16].

[2]To keep the presentation as simple as possible, we present DS rule for only two BBAs, see [14] for its generalization.

for decision making. The Pignistic Probability Transformation (PPT, or $BetP$ transform) proposed by Smets in [18], [19] is generally considered as a reasonable in-between decisional attitude between the max of $Bel(.)$ (pessimistic attitude) and max of $Pl(.)$ (optimistic attitude). The betting probability $BetP(\theta_i)$ of any singleton focal element θ_i of the FoD is defined by

$$BetP\left(\theta_i\right) = \sum_{\theta_i \in X, X \in 2^{\Theta}} \frac{m\left(X\right)}{|X|} \tag{9}$$

where $|X|$ refers to the cardinality of a subset X. One clearly sees tat the BetP transform evenly distributes the belief assignment of disjunctive focal element to the singleton focal element it contains.

III. Weighted fusion of multiple classifiers

A. Classifiers for HAR

In this article, we use classical machine learning classifiers [20] such as SVM, RF, MLP and LR to generate BBAs of human activity, and these classifiers can only give the mass of belief for singleton focal elements (i.e. we work with Bayesian BBAs). In the fusion of multiple classifiers, a BBA can be represented by the output of each classifier. It is worth noting that we should choose different types of classifiers as far as possible. In general, when the diversity between the multiple classifiers is larger, the advantages will be more obvious. At the same time, we need to guarantee the individual prediction accuracy of each classifier, which is the basis for the high accuracy of our WFMC algorithm. Furthermore, we train each classifier separately using the same training dataset. Once multiple classifiers are trained, we can obtain the corresponding BBAs for each category of human activity.

B. Assessment Criteria

After acquiring multiple BBAs of the human activity to be identified, we can use DS rule to fuse these BBAs and further make decisions. In this article we work with DS rule mainly because of its simplicity even we are aware of its well-known disputable dictatorial behavior in some cases, and that is why we use discounting techniques. We will evaluate the performances of alternative fusion rules in our future works. From the perspective of conflicts between multiple BBAs or uncertain information, the reliability of multiple BBAs should be evaluated before combination, its goal is to eliminate and reduce the negative influence of unreliable BBAs on the final recognition accuracy. For this reason, the appropriate assessment criterion need to be chosen in advance. In this article, we have selected three assessment criteria, described as follows:

a) Divergence degree: The Belief Jensen–Shannon (BJS) divergence measure was presented by Xiao [21] to measure the divergence between belief functions in DST. It is the generalization of the Jensen-Shannon divergence [22] where the probability distribution is replaced with belief mass functions.

Let m_1 and m_2 be two BBAs on the same FoD, containing n mutually exclusive and exhaustive hypotheses. The BJS divergence between m_1 and m_2 is denoted as:

$$BJS\left(m_1, m_2\right) = \frac{1}{2}\left[\sum_{i=1}^{2^n-1} m_1(A_i)\log\left(\frac{2m_1(A_i)}{m_1(A_i)+m_2(A_i)}\right) + \sum_{i=1}^{2^n-1} m_2(A_i)\log\left(\frac{2m_2(A_i)}{m_1(A_i)+m_2(A_i)}\right)\right] \tag{10}$$

where A_i is a non empty element of the power-set 2^{Θ}, and $\sum_{i=1}^{2^n-1} m_1(A_i) = 1, \sum_{i=1}^{2^n-1} m_2(A_i) = 1$. The lower and upper bounds of the BJS divergence measure are respectively equal to zero and one. When m_1 has the same BBAs as m_2, the BJS divergence between m_1 and m_2 is 0. When two BBAs are completely different, the BJS divergence value is 1. In this article, the average BJS divergence of a BBA can be calculated by

$$\widetilde{BJS}(m) = \frac{1}{N-1}\sum_{j=1}^{N} BJS(m, m_j) \tag{11}$$

where N indicates the number of classifiers.

b) Distance degree: The smaller the distance between a pair of BBAs, the closer their belief values are, and the better for our decision-making. In this article, the interval distance [23] is an excellent metric, as it considers the belief intervals using the belief and plausibility functions of each focal element to describe the closeness between BBAs. The interval distance is defined as follows:

$$d_{BI}^{Ec}\left(m_1, m_2\right) \triangleq \sqrt{\frac{1}{2^{n-1}} \cdot \sum_{i=1}^{2^n-1}\left[d_{BI}\left(BI_1\left(A_i\right), BI_2\left(A_i\right)\right)\right]^2} \tag{12}$$

with

$$BI\left(A_i\right) = \left[Bel\left(A_i\right), Pl\left(A_i\right)\right] \tag{13}$$

$$d_{BI}\left(\left[a_1, b_1\right], \left[a_2, b_2\right]\right) = \sqrt{\left[\frac{a_1+b_1}{2} - \frac{a_2+b_2}{2}\right]^2 + \frac{1}{3}\left[\frac{b_1-a_1}{2} - \frac{b_2-a_2}{2}\right]^2}\ . \tag{14}$$

The average interval distance of one set of BBAs can be calculated by

$$\widetilde{d_{BI}^{Ec}}(m) = \frac{1}{N-1}\sum_{j=1}^{N} d_{BI}^{Ec}(m, m_j) \tag{15}$$

where N indicates the number of classifiers. The larger the value of the interval distance, the greater the degree of conflict between the current BBA and other BBAs, the less reliable it will be, and vice versa.

c) Uncertain degree: A novel effective measure of uncertainty (i.e. entropy) of BBAs is proposed by Dezert [24], this new continuous measure is effective in the sense that it satisfies a small number of very natural and essential desiderata. The new entropy measure is defined by

$$U(m)=\sum_{X\in 2^{\Theta}} s(X) \tag{16}$$

with

$$s(X)\triangleq -(1-u(X))m(X)\log(m(X))+u(X)(1-m(X)) \tag{17}$$

$$u(X)\triangleq Pl(X)-Bel(X). \tag{18}$$

$s(X)$ is the uncertainty contribution of X in $U(m)$. This measure of uncertainty coincides with Shannon entropy for any Bayesian BBA, it can be also interpreted as an effective generalization of Shannon entropy. We always have $U(m)\geq 0$, and $U(m)<U(m_v)$ if the BBA $m(.)$ is different of the vacuous BBA $m_v(.)$ defined by $m_v(\Theta)=1$. It is worth noting that it is possible that a non-Bayesian BBA can have an entropy value $U(m)$ smaller than the maximum of Shannon entropy given by $\log(|\Theta|)$. When X is a single focal element and satisfies $m(X)=1$, $U(m)$ has a minimum value of 0, which indicates that the source of evidence is completely certain and it plays an important role in the final combination.

C. Reliability Evaluation of Classifiers

In this article, each classifier can be regarded as a evidence source. We obtain the reliability of one classifier by evaluating its output, as follows:

a) Construction of scoring matrix: Supposing that there exists N classifiers over the same FoD, and their BBAs composition are as follows:

$$\begin{array}{c|cccc} & A_1 & A_2 & \cdots & A_M \\ \hline C_1 & m_1(A_1) & m_1(A_2) & \cdots & m_1(A_M) \\ C_2 & m_2(A_1) & m_2(A_2) & \cdots & m_2(A_M) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_N & m_N(A_1) & m_N(A_2) & \cdots & m_N(A_M) \end{array} \tag{19}$$

where $A_i\in 2^{\Theta}$, and $C_j, j=1,2,...,N$ represents the jth classifier. Then we calculate the scores of each classifier according to the assessment criteria $Crit_\eta$, $\eta=1,2,...,q$ and the scoring matrix S can be generated as follows:

$$\begin{array}{c|cccccc} & C_1 & C_2 & \cdots & C_j & \cdots & C_N \\ \hline Crit_1 & S_{11} & S_{12} & \cdots & S_{1j} & \cdots & S_{1N} \\ Crit_2 & S_{21} & S_{22} & \cdots & S_{2j} & \cdots & S_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ Crit_q & S_{q1} & S_{q2} & \cdots & S_{qj} & \cdots & S_{qN} \end{array} \tag{20}$$

In this article, $q=3$, that is: $Crit_1\triangleq BJS(\cdot)$, $Crit_2\triangleq d_{BI}^{Ec}(\cdot)$ and $Crit_3\triangleq U(\cdot)$.

b) Construction of local BBAs for classifiers: Considering the assessment criteria and their corresponding evaluation vectors, we can calculate the positive support degree $Sup_\eta(C_j)$ and negative support degree $Inf_\eta(C_j)$ for each classifier by the following equations (see [16] for details)

$$Sup_\eta(C_j)\triangleq \sum_{\kappa\in\{1,\cdots,N\}|S_{\eta\kappa}\leq S_{\eta j}} |S_{\eta j}-S_{\eta\kappa}|. \tag{21}$$

$$Inf_\eta(C_j)\triangleq -\sum_{\kappa\in\{1,\cdots,N\}|S_{\eta\kappa}\geq S_{\eta j}} |S_{\eta j}-S_{\eta\kappa}|. \tag{22}$$

Then, the maximum value $C_{\max}^{\eta}$ and minimum value $C_{\min}^{\eta}$ of the classifier C_j under the assessment criteria $Crit_\eta$ can be obtained by the following equations.

$$C_{\max}^{\eta}\triangleq max_j Sup_\eta(C_j) \tag{23}$$

$$C_{\min}^{\eta}\triangleq min_j Inf_\eta(C_j). \tag{24}$$

Next, the construction of local BBAs is based on the method presented in [16] and defined as follows:

$$\begin{cases} m_{j-\eta}(C_j)\triangleq Bel_\eta(C_j) \\ m_{j-\eta}(\overline{C}_j)\triangleq 1-Pl_\eta(C_j) \\ m_{j-\eta}(C_j\cup\overline{C}_j)\triangleq Pl_\eta(C_j)-Bel_\eta(C_j) \end{cases} \tag{25}$$

with

$$\begin{cases} Bel_\eta(C_j)\triangleq \dfrac{Sup_\eta(C_j)}{C_{\max}^{\eta}} \\ Bel_\eta(\overline{C}_j)\triangleq \dfrac{Inf_\eta(C_j)}{C_{\min}^{\eta}} \\ Pl_\eta(C_j)\triangleq 1-\dfrac{Inf_\eta(C_j)}{C_{\min}^{\eta}} \end{cases} \tag{26}$$

where $m_{j-\eta}(C_j)$, $m_{j-\eta}(\overline{C}_j)$ and $m_{j-\eta}(C_j\cup\overline{C}_j)$ respectively represent the positive support belief, negative support belief and uncertainty belief of the classifier C_j based on the assessment criteria $Crit_\eta$.

c) Calculation of weight factors: We employ the BF-TOPSIS algorithm [16] to calculate the weight factors for each classifier and the specific steps are as follows.

- Step 1 Calculate the local BBAs $m_{j-\eta}(C_j)$, $m_{j-\eta}(\overline{C}_j)$ and $m_{j-\eta}(C_j\cup\overline{C}_j)$ of each classifier according to the scoring matrix.
- Step 2 For each classifier, calculate $d_{BI}^{Ec}(m_{j-\eta}, m_\eta^{best})$ and $d_{BI}^{Ec}(m_{j-\eta}, m_\eta^{worst})$, m_η^{best} and m_η^{worst} represent the best and the worst ideal BBAs based on the assessment criteria $Crit_\eta$, respectively, where $m_\eta^{best}(C_j)=1$ and $m_\eta^{worst}(\overline{C}_j)=1$.
- Step 3 Calculate the weighted average distance $d^{best}(C_j)$ and $d^{worst}(C_j)$ of classifier, where

$$d^{best}(C_j)\triangleq \sum_{\eta=1}^{N} \upsilon(Crit_\eta)\cdot d_{BI}^{Ec}(m_{j-\eta}, m_\eta^{best}) \tag{27}$$

$$d^{worst}(C_j)\triangleq \sum_{\eta=1}^{N} \upsilon(Crit_\eta)\cdot d_{BI}^{Ec}(m_{j-\eta}, m_\eta^{worst}) \tag{28}$$

where $\upsilon(Crit_\eta)$ represents the weight of assessment criteria $Crit_\eta$. In this article, $\upsilon(Crit_1)=\upsilon(Crit_2)=\upsilon(Crit_3)=1/3$.

- Step 4 The final weight of the classifier C_j is defined as follows:

$$\omega\left(C_j\right)=\frac{d^{worst}\left(C_j\right)}{d^{worst}\left(C_j\right)+d^{best}\left(C_j\right)} \tag{29}$$

In the proposed WFMC algorithm, when a classifier is in complete conflict with other classifiers, it will be supported to a small degree. According to the reliability evaluation algorithm, the classifier will receive a small weighting factor, which discounts the masses of all focal elements to the total ignorance. This reduces the total conflict between classifiers in the fusion process, making the total conflict in the proposed WFMC algorithm always less than 1, thus improving the reliability of the fusion results.

After obtaining the weight factors for each classifier, multiple classifiers can be fused using the classical (i.e. Shafer's) discounting technique and DS rule, and decision can be made based on the maximum BetP probability value. For the convenience of implementation, the brief framework of the WFMC method is given in Fig. 1.

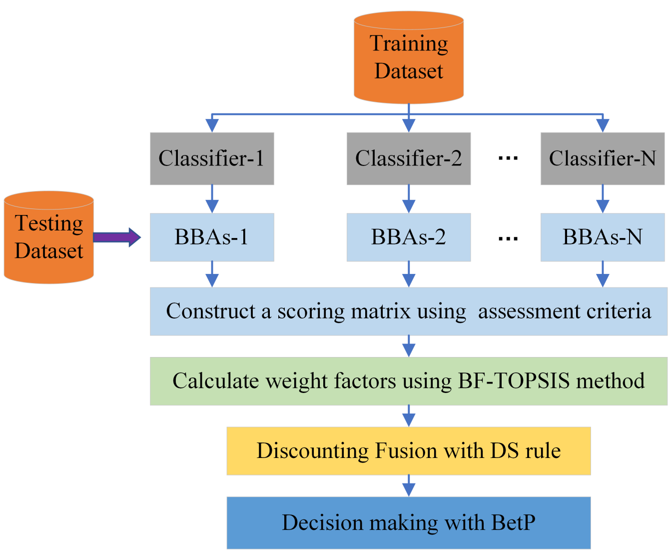

Fig. 1. The framework of WFMC method.

As we can see in Fig. 1, the proposed WFMC algorithm includes four main steps:

- **Step 1 (Classifiers trained):** Multiple classifiers of different types are trained by the same training dataset for acquiring BBAs.
- **Step 2 (Classifiers evaluation):** For each BBA generated by the classifier, a reliability evaluation is performed using three criteria and a scoring matrix is constructed.
- **Step 3 (Calculation of weight factors):** The BF-TOPSIS algorithm is employed to calculate the weight factors for each classifier based on the scoring matrix.
- **Step 4 (Discounting fusion):** Multiple classifiers are combined sequentially using the classical discounting technique and DS rule, the final decision can be made based on the maximum BetP probability.

IV. Experiments and Discussions

A. UCI Smartphone Dataset

In this article, the UCI Smartphone dataset is considered for experimental verification. In UCI Smartphone dataset, the experiments have been carried out with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed six activities (walking, walking upstairs, walking downstairs, sitting, standing and laying) wearing a smartphone on the waist. Three-axial linear acceleration and three-axial angular velocity at a constant rate of 50Hz were captured by using its embedded accelerometer and gyroscope. The sensor signals (accelerometer and gyroscope) were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50 percent overlap (128 readings/window). More descriptions of the UCI Smart-phone dataset can be found in [25].

B. Example

In order to show how our WFMC method works, an example is given to illustrate its specific procedures. Firstly, the focal element in BFs theory can be applied to mathematically represent human activities. Specifically $\theta_1 \triangleq$ walking, $\theta_2 \triangleq$ walking upstairs, $\theta_3 \triangleq$ walking downstairs, $\theta_4 \triangleq$ sitting, $\theta_5 \triangleq$ standing, $\theta_6 \triangleq$ laying. For the 480th sample data with a true label of (standing) in the test dataset, the corresponding BBAs generated by four classifiers are shown in Table I. According to the principle of maximum probability, it can be seen that SVM and RF support θ_4 (sitting) while MLP and LR support θ_5 (standing), which causes trouble to make decisions. We utilize DS rule to combine the four classifiers and the fusion results have the maximum belief value of 0.567 to support θ_4 (sitting), which is not what we want.

Next, we use WFMC algorithm for testing. The scoring matrix is acquired based on (9), (13) and (14), as shown in Table II. Then, we can get the positive support and negative support degree of each classifier according to (17) and (18), which are given in Table III and Table IV. It can be seen that $BJS(\cdot)$ has the highest support for RF, while $U(\cdot)$ has the highest support for SVM, and $d_{BI}^{Ec}(\cdot)$ supports both RF and MLP. After that, the derived local BBAs of each classifier can be also obtained using (21) shown in Table V, Table VI and Table VII. And then by using step 2 and step 3 in BF-TOPSIS algorithm, we can obtain distance $d^{best}\left(C_j\right)$ and $d^{worst}\left(C_j\right)$ of classifiers. The weight coefficients of each classifier can be further obtained based on (27), as shown in Table VIII. It can be seen that SVM acquires the smallest weighting factor, while MLP gets the largest weighting factor and RF has a similar weighting factor to MLP, which indicates that MLP has the highest reliability for the current activity. Finally, four classifiers are combined using DS rule (6) generalized[3] for four BBAs, and the probability values for each category of activity are obtained based on (9), as shown in Table IX. We

[3]Because DS rule is associative, the four BBAs can also be fused sequentially and the sequential order of DS fusion does not impact the final result.

can see that θ_5 (standing) has the maximum BetP probability value, which is consistent with the true label.

TABLE I
BBAs OF THE 480TH TEST SAMPLE.

	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	**Result**
SVM	0.0	0.0	0.0	0.844	0.156	0.0	θ_4
RF	0.0	0.0	0.0	0.573	0.427	0.0	θ_4
MLP	0.0	0.0	0.0	0.349	0.651	0.0	θ_5
LR	0.0	0.0	0.0	0.251	0.749	0.0	θ_5
DS rule	0.0	0.0	0.0	0.567	0.433	0.0	θ_4

TABLE II
SCORING MATRIX OF FOUR CLASSIFIERS.

	SVM	RF	MLP	LR
$BJS(\cdot)$	0.218	0.043	0.056	0.088
$d_{BI}^{Ec}(\cdot)$	0.113	0.068	0.068	0.084
$U(\cdot)$	0.628	0.984	0.935	0.8162

TABLE III
POSITIVE SUPPORT DEGREE $Sup_\eta(\cdot)$ OF FOUR CLASSIFIERS.

$Sup_\eta(\cdot)$	SVM	RF	MLP	LR
$BJS(\cdot)$	0.0	0.232	0.194	0.130
$d_{BI}^{Ec}(\cdot)$	0.0	0.061	0.061	0.029
$U(\cdot)$	0.852	0.0	0.05	0.287

TABLE IV
NEGATIVE SUPPORT DEGREE $Inf_\eta(\cdot)$ OF FOUR CLASSIFIERS.

$Inf_\eta(\cdot)$	SVM	RF	MLP	LR
$BJS(\cdot)$	-0.466	0.0	-0.013	-0.077
$d_{BI}^{Ec}(\cdot)$	-0.119	0.0	0.0	-0.033
$U(\cdot)$	0.0	-0.575	-0.425	-0.188

TABLE V
LOCAL BBAs OF FOUR CLASSIFIERS ON $BJS(\cdot)$.

	SVM	RF	MLP	LR
$m_{BJS(\cdot)}(C_j)$	0.0	1.0	0.835	0.560
$m_{BJS(\cdot)}(\overline{C}_j)$	1.0	0.0	0.027	0.164
$m_{BJS(\cdot)}(C_j \cup \overline{C}_j)$	0.0	0.0	0.138	0.276

TABLE VI
LOCAL BBAs OF FOUR CLASSIFIERS ON $d_{BI}^{Ec}(\cdot)$.

	SVM	RF	MLP	LR
$m_{d_{BI}^{Ec}(\cdot)}(C_j)$	0.0	1.0	1.0	0.470
$m_{d_{BI}^{Ec}(\cdot)}(\overline{C}_j)$	1.0	0.0	0.0	0.273
$m_{d_{BI}^{Ec}(\cdot)}(C_j \cup \overline{C}_j)$	0.0	0.0	0.0	0.257

C. Measure of Performances

The classical Accuracy is applied to measure the performance of our proposed method. The specific definitions are as follows:

$$Accuracy = \frac{1}{n}\sum_{i=1}^{n}\frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{30}$$

TABLE VII
LOCAL BBAs OF FOUR CLASSIFIERS ON $U(\cdot)$.

	SVM	RF	MLP	LR
$m_{U(\cdot)}(C_j)$	1.0	0.0	0.059	0.337
$m_{U(\cdot)}(\overline{C}_j)$	0.0	1.0	0.740	0.327
$m_{U(\cdot)}(C_j \cup \overline{C}_j)$	0.0	0.0	0.202	0.336

TABLE VIII
WEIGHTED COEFFICIENTS OF FOUR CLASSIFIERS.

	$d^{best}(C_j)$	$d^{worst}(C_j)$	$\omega(C_j)$
SVM	0.471	0.236	0.333
RF	0.236	0.471	0.667
MLP	0.238	0.506	0.680
LR	0.334	0.461	0.580

TABLE IX
RESULTS OF THE WFMC METHOD.

	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	Θ
Weighted fusion	0.0	0.0	0.0	0.389	0.548	0.0	0.063
$BetP(\cdot)$	0.01	0.01	0.01	0.399	0.558	0.01	0.0

where i denotes class index and n is the number of classes. TP_i, TN_i, FN_i and FN_i are respectively True Positives, True Negatives, False Positives and False Negatives.

D. Experimental Results and Analysis

According to the specific steps described in Fig. 1, we first train four classifiers using 7352 samples, including a SVM, a RF, a MLP and a LR. For the parameters of SVM, the sigmoid function is selected as kernel function, and the penalty parameter is set to 1.0. For the parameters of RF, the number of trees in the forest is set to 150. For the parameters of MLP, the number of hidden layers is set to 300. For the parameters of LR, the penalty is set to L1. Default parameters are selected for the remaining parameters of four classifiers. In this article, features are extracted from raw sensor data for model training, including 11 time-domain and 6 frequency-domain features as shown in Table II. Then the trained four classifiers are employed to predict the testing dataset containing 2947 samples. Furthermore, we fuse the four classifiers using the DS rule and the proposed WFMC algorithm, respectively, the results are shown in the Table III, and the related confusion matrixs are shown in Fig. 2 and Fig. 3. We can find that LR has the highest accuracy among the individual classifier with 93.52%, which is weaker than the DS rule approach. It indicates that individual classifier has limited classification ability. Moreover, we can clearly see that the performance of the proposed WFMC method is significantly better than other mentioned method, which shows the effectiveness of our strategy.

Compared to the approach of traditional DS rule, the proposed method effectively improves the recognition accuracy. The misclassification where sitting was incorrectly recognized as standing is reduced from 12.4% to 8.4% and the misclassification where walking downstairs was incorrectly recognized

TABLE X
FEATURE EXTRACTION

Domain	Features
Time	Mean value, Standard deviation, Median absolute value, Maximum, Minimum, Signal magnitude area, Average sum of the squares, Interquartile range, Signal entropy, Autoregression coefficients, Correlation
Frequency	Largest frequency component, Weighted average, Skewness, Kurtosis, Energy of a frequency interval, Angle between two vectors

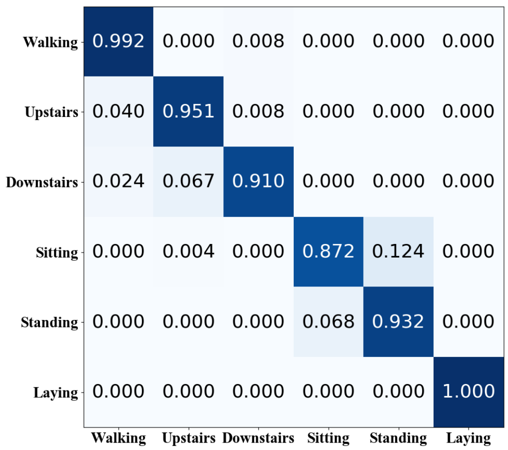

Fig. 2. Confusion matrix on UCI smartphone dataset by DS rule.

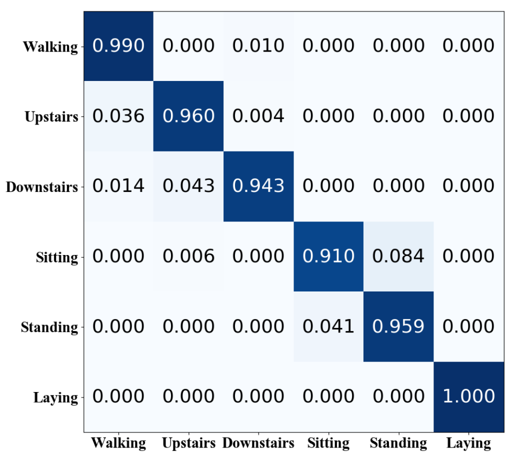

Fig. 3. Confusion matrix on UCI smartphone dataset by the WFMC method.

as walking upstairs is reduced from 6.7% to 4.3%. This is due to the fact that the three evaluation criteria we have given are a good measure of the conflict between multiple classifiers and their own uncertainty, and the BF-TOPSIS algorithm efficiently calculates the weight coefficients for each classifier, which improves the accuracy of the multi-classifier fusion.

Furthermore, we compare with some state-of-the-art approaches in literatures to demonstrate the superiority of our method, including Activity Graph Based Convolutional Neural Network [26], DSmT-Based Kernel Density Estimation [6], Sensor fusion and deep recurrent neural network-based [7], Two-stream Transformer Network [27], Hesitant Fuzzy Belief Structure Based Fused Extreme Learning Machine [28]. As we can see in Table IV, our method outperforms these state-of-the-art methods in terms of accuracy.

TABLE XI
COMPARISON OF WFMC METHOD WITH TRADITIONAL METHODS ON THE UCI SMARTPHONE DATASET.

Method	Accuracy	Time (s)
SVM	91.75%	9.08
RF	92.94%	2.06
MLP	92.53%	1.95
LR	93.52%	1.97
DS rule	94.43%	26.01
WFMC	**96.20%**	**33.10**

TABLE XII
COMPARISON OF WFMC METHOD WITH STATE-OF-THE-ART METHODS ON THE UCI SMARTPHONE DATASET.

Method	Accuracy	Time(s)
Activity Graph CNN-Based [26]	90.17%	11.34
DSmT-Based Kernel Density Estimation [6]	93.05%	24.46
Sensor fusion and deep RNN-based [7]	94.27%	15.95
Two-stream Transformer Network [27]	94.12%	20.19
Hesitant Fuzzy Belief Based ELM [28]	95.20%	23.78
WFMC	**96.20%**	**33.10**

In terms of time consumption, our WFMC method was programming in Python 3.7 with a hardware of Intel Core i7-8700 CPU at 3.20 GHz and 16 GB RAM. We use 2947 test samples and counted the total time consumed by each method. As can be seen, traditional machine learning algorithms have the advantage of being fast. As our WFMC algorithm is developed based on DST, it inevitably increases the computational burden. Nevertheless, the average elapsed time per test sample is about 11ms, which is sufficient for practical applications.

V. CONCLUSION

In this article, we have proposed a novel weighted fusion of multiple classifiers based on belief functions theory for human activity recognition. Firstly, we train four classical machine learning classifiers by using time-domain and frequency-domain features to obtain basic belief assignments of human activities. Secondly, we evaluate the outputs of four classifiers using three criteria and construct a scoring matrix. Thirdly, we use the multi-criteria BF-TOPSIS algorithm to calculate the weight coefficients of each classifier. Finally, we adopt a discounting technique and DS rule to combine the four classifiers, and make decisions thanks to the pignistic probability values. Several experiments have been conducted based on the UCI Smartphone dataset. The experimental results prove that our WFMC approach can significantly improve the classification accuracy with respect to several classical and state-of-the-art methods.

In our future works, we will evaluate a better measure of divergence between belief functions based on a more effective definition of relative entropy and cross-entropy. We will also explore the possibility to adapt our Stable Preference Ordering Towards Ideal Solution (SPOTIS) rank reversal multi-criteria method for HAR instead using the BF-TOPSIS method which is not robust to rank reversal. We will test and compare an other decision-making technique based on belief-interval distance, and work on how to reduce the complexity of multi-classifier fusion for HAR in order to apply it to an online real activity recognition system.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 62233003 and 62073072, and in part by the Key Projects of Key Program of Jiangsu Province under Grant BE2020006 and Grant BE2020006-1, and in part by Shenzhen Natural Science Foundation under Grant JCYJ20210324132202005 and JCYJ20220818101206014.

REFERENCES

[1] R. Gravina, P. Alinia, H. Ghasemzadeh, G. Fortino, *Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges*, Information Fusion, Vol. 35, pp. 68–80, May 2017.

[2] Y.J. Hao, R. Zheng, B.Y. Wang, *Invariant Feature Learning for Sensor-Based Human Activity Recognition*, IEEE Trans. on Mobile Computing, Vol. 21(11), pp. 4013–4024, 2022.

[3] Y. Tang, L. Zhang, F. Min, J. He, *Multiscale Deep Feature Learning for Human Activity Recognition Using Wearable Sensor*, IEEE Trans. on Indus. Elect., Vol. 70 (2), pp. 2106–2116, Feb. 2023.

[4] Y. Dong, R. Zhou, C. Zhu, L. Cao, X. Li, *Hierarchical Activity Recognition Based on Belief Functions Theory in Body Sensor Networks*, IEEE Sensors Journal, Vol. 22(15), pp. 15211–15221, Aug. 2022.

[5] X. Li, X. Huang, J. Dezert, F. Smarandache, *Enrichment of qualitative beliefs for reasoning under uncertainty*, in Proc. of 10th Int. Conf. on Information Fusion, Québec, Canada, July 2007.

[6] Y. Dong, X. Li, J. Dezert, M.O. Khyam, Md. Noor-A-Rahim, S.S. Ge, *Dezert-Smarandache Theory-Based Fusion for Human Activity Recognition in Body Sensor Networks*, IEEE Trans. on Industrial Informatics, Vol. 16(11), pp. 7138–7149, Nov. 2020.

[7] Md.Z. Uddin, M. Hassan, A. Alsanad, C. Savaglio, *A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare*, Information Fusion, Vol. 55, pp. 105-115, March 2020.

[8] Y. Zhao, X. Wang, Z. Lin, Z. Huang, *Multi-Classifier Fusion for Open-Set Specific Emitter Identification*, Remote Sensing, Vol. 14(9), May 2022.

[9] X. Li, W Yang, X Wu, and J. Dezert, *A fast approximate reasoning method in hierarchical DSmT(B)*, Acta Electronica Sinica, vol. 39, no. S1, pp. 31-36, 2011.

[10] Z.G. Liu, Q. Pan, J. Dezert, J.W. Han, Y. He, *Classifier Fusion With Contextual Reliability Evaluation*, IEEE Trans. on Cybernetics, Vol. 48(5), pp. 1605–1618, 2018.

[11] S. Liu, X. Li, C. Hu, J. Yao, Junping, X. Han, J. Wang, *Spammer detection using multi-classifier information fusion based on evidential reasoning rule*, Scientific Reports, Vol. 12(1), July 2022.

[12] Z.G. Liu, Q. Pan, J. Dezert, A. Martin, *Combination of Classifiers With Optimal Weight Based on Evidential Reasoning*, IEEE Trans on Fuzzy Systems, Vol. 26(3), pp. 1217–1230, June 2018.

[13] X. Li, J. Dezert, X. Huang, Z. Meng, and X. Wu, *A fast approximate reasoning method in hierarchical DSmT(A)*, Acta Electronica Sinica, vol. 38, no. 11, pp. 2566–2572, 2010.

[14] G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, USA, 1976.

[15] Y. Dong, X. Li, J. Dezert, S.S. Ge, *A Novel Multi-Criteria Discounting Combination Approach for Multi-Sensor Fusion*, IEEE Sensors Journal, Vol. 19(20), pp. 9411–9421, Oct. 2019.

[16] J. Dezert, D. Han, H. Yin, *A New Belief Function Based Approach for Multi-Criteria Decision-Making Support*, Proc. of 19th Int. Conf. on Information Fusion, pp. 782–789, Heidelberg, Germany, July 5–8, 2016.

[17] Z.G. Liu, X. Zhang, J. Niu, J. Dezert, *Combination of Classifiers With Different Frames of Discernment Based on Belief Functions*, IEEE Trans on Fuzzy Systems, Vol. 29(7), pp. 1764–1774, July 2021.

[18] P. Smets, *Constructing the pignistic probability function in a context of uncertainty*, Uncertainty in AI, Vol. 5, pp. 29–39, 1990.

[19] P. Smets, *Decision making in the TBM: the necessity of the pignistic transformation*, Int. Jour. of Approx. Reasoning, Vol. 38, 2005.

[20] L. Breiman, *Random forests*, Machine Learning, Vol. 45(1), pp. 5–32, 2001.

[21] F. Xiao, *Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy*, Information Fusion, Vol. 46, pp. 23–32, March 2019.

[22] https://en.wikipedia.org/wiki/Jensen-Shannon_divergence.

[23] D. Han, J. Dezert, Y. Yang, *Belief Interval-Based Distance Measures in the Theory of Belief Functions*, IEEE Trans. on SMCS, Vol. 48(6), pp. 833–850, June 2018.

[24] J. Dezert, *An Effective Measure of Uncertainty of Basic Belief Assignments*, in Proc. of 25th Int. Conf. on Information Fusion, Linköping, Sweden, July 4–7, 2022.

[25] D. Thakur, S. Biswas, *An Integration of feature extraction and Guided Regularized Random Forest feature selection for Smartphone based Human Activity Recognition*, J. of Network and Computer Applications, Vol. 204, Aug. 2022. https://doi.org/10.1016/j.jnca.2022.103417.

[26] P. Yang, C. Yang, V. Lanfranchi, F. Ciravegna, *Activity Graph Based Convolutional Neural Network for Human Activity Recognition Using Acceleration and Gyroscope Data*, IEEE Trans. on Industrial Informatics, Vol. 18(10), pp. 6619–6630, Oct. 2022.

[27] S. Xiao, S. Wang, Z. Huang, Y. Wang, H. Jiang, *Two-stream transformer network for sensor-based human activity recognition*, Neurocomputing, Vol. 512, pp. 253–268, Nov. 2022.

[28] Y. Dong, X. Li, J. Dezert, R. Zhou, C. Zhu, L. Wei, S.S. Ge, *Evidential Reasoning With Hesitant Fuzzy Belief Structures for Human Activity Recognition*, IEEE Trans. on Fuzzy Systems, Vol. 29(12), pp. 3607-3619, Dec. 2021.