WB-KNN for emotion recognition from physiological signals*

XIE Weilun (谢伟伦)[1,2] **and XUE Wanli** (薛万利)[1,2]**
1. School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China
2. Engineering Research Center of Learning-Based Intelligent System, Ministry of Education, Tianjin 300384, China

(Received 13 July 2020; Revised 28 August 2020)

K-nearest neighbor (KNN) has yielded excellent performance in physiological signals based on emotion recognition. But there are still some issues: the majority vote only by the nearest neighbors is too simple to deal with complex (like skewed) class distribution; features with the same contribution to the similarity will degrade the classification accuracy; samples in boundaries between classes are easily misclassified when k is larger. Therefore, we propose an improved KNN algorithm called WB-KNN, which takes into account the weight (both features and classification) and boundaries between classes. Firstly, a novel weighting method based on the distance and farthest neighbors named WDF is proposed to weight the classification, which improves the voting accuracy by making the nearer neighbors contribute more to the classification and using the farthest neighbors to reduce the weight of non-target class. Secondly, feature weight is introduced into the distance formula, so that the significant features contribute more to the similarity than noisy or irrelevant features. Thirdly, a voting classifier is adopted in order to overcome the weakness of KNN in boundaries between classes by combining different classifiers. Results of WB-KNN algorithm are encouraging compared with the traditional KNN and other classification algorithms on the physiological dataset with a skewed class distribution. Classification accuracy for 29 participants achieves 94.219 2% for the recognition of four emotions.
Document code: A **Article ID:** 1673-1905(2021)07-0444-5
DOI https://doi.org/10.1007/s11801-021-0118-2

The purpose of emotion recognition is to build a harmonious human-machine environment by giving the machine the ability to recognize human emotion. It has been widely applied in various fields, such as disease diagnosis, criminal investigation and distance education. At present, emotion recognition based on physiological signals has attracted increasingly attention, since physiological signals are not controlled by participants' subjective consciousness compared with facial expressions and speech, and they are convenient to record as long as the participant wears bio-sensors.

Physiological signals are characterized by weakness, instability and randomness, making it difficult to identify emotions from them. Many classification methods have been introduced to overcome the difficulty. As a simple, nonlinear, nonparametric but effective pattern classifier, K-nearest neighbor (KNN)[1] has been extensively used in physiological signals based emotion recognition.

But there are still some issues needed to be improved. On one hand, the basic majority vote strategy is too simple to deal with the skewed class distribution that samples of a more frequent class tend to dominate the prediction of the test example. Because they tend to be common among the k nearest neighbors due to their large number, and the vote only by the nearest neighbors is a little limited. On the other hand, the accuracy of the KNN algorithm can be severely degraded by the presence of noisy or irrelevant features, because all the features contribute the same to the similarity (sample distance). Feature selection is commonly used to deal with this problem, but may cause the loss of feature information. Besides, the samples in the boundaries between classes are easily misclassified when k is larger.

To address these issues, the WB-KNN algorithm is proposed in this paper, which takes into account the weight of features and neighbors and the boundaries between classes based on the traditional KNN algorithm. Specially, the WB-KNN algorithm includes three major improvements: The WDF weighting method is based on the distance and farthest neighbors are proposed to weight the classification. On the one hand, this method employs a standard normal distribution function with excellent distribution characteristics to weight the neighbors according to their distance to the test sample, in order that the nearer neighbors have the higher voting weight. On the other hand, considering the farther

* This work has been supported by the National Natural Science Foundation of China (Nos.61906135 and U1509207), and the National Key R&D Program of China (No.2018YF-B1305200).
** E-mail: xuewanli@email.tjut.edu.cn

neighbors are less likely to be the same class as the test sample, the farthest neighbors are used to improve the correction rate of the voting result by reducing the weight of non-target class.

The standard Euclidean distance adopted in traditional KNN algorithm is replaced with feature-weighted distance, so that the significant features contribute more to the similarity than noisy and irrelevant features, the feature weight is expressed by its importance evaluated using a tree-based ensemble method.

To overcome the weakness of KNN in boundaries between classes, an ensemble classifier (voting classifier)[2] is adopted to classify the samples in boundaries between classes by combining native Bayes (NB), decision tree (DT), support vector machine (SVM) and using the average predicted probabilities (soft vote), which can result in better overall performance than any of the individual classifiers in the ensemble.

Besides, due to the nonlinear separability of the physiological dataset, kernel principal component analysis (KPCA) with radial basis function (RBF) is adopted to map the data into a high dimension space, extract more distinguishable features, and then reduce the feature dimension.

Based on the analysis above, an emotion recognition system based on WB-KNN algorithm is designed, as shown in Fig.1, the system includes three parts as follows.

1. Data and feature: data acquisition, preprocessing and feature extraction.

2. Feature processing: KPCA is used to map the data into a high dimension space to extract more distinguishable features, then feature importance is evaluated by a tree-based estimator.

3. Modeling and classification: WB-KNN model is established to predict the emotion category of the test set. There are three steps in the WB-KNN model. First, calculate the nearest and farthest neighbors by the feature-weighted distance; Second, use the WDF weighting method to weight the classification of the test samples that are not in the boundary between classes; third, adopt the voting classifier to classify the test samples in the boundary between classes.

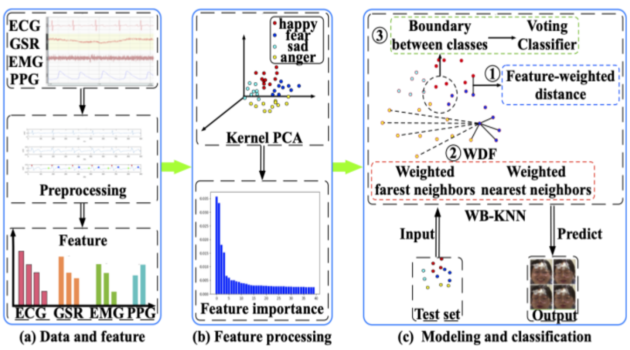

Fig.1 The emotion recognition system based on WB-KNN algorithm

Overall, our contributions are as follows:

1. The WDF weighting method based on distance and the farthest neighbors is proposed to improve the voting accuracy, by making the nearer neighbors contribute more to the voting decision and using the farthest neighbors to reduce the weight of non-target class.

2. Feature-weighted distance is used to measure the similarity of samples, which can eliminate the interference of noise or irrelevant features and can reduce the loss of feature information caused by feature selection.

3. The voting classifier combining different classifiers by the soft vote strategy is adopted to overcome the weakness of traditional KNN algorithm in boundaries between classes.

4. The effectiveness of the proposed WB-KNN algorithm has been proved on the physiological dataset, collected from 29 participants under four induced emotions.

Acquisition of a high quality physiological dataset is the first stage in physiological signals based emotion recognition. In this field, autonomic nervous system (ANS)[3] signals have been increasingly emphasized, such as electrocardiogram (ECG), galvanic skin response (GSR), electromyography (EMG), photoplethysmography (PPG), respiration (RSP), blood volume pulse (BVP), skin temperature (SKT), skin conductance (SC) and so on.

Kim et al[4] developed a novel emotion recognition system using short-term physiological signals (ECG, SKT, electrodermal activity). Li et al[5] proposed recognizing emotion using four physiological signals (ECG, SKT, SC and RSP). Chang et al[6] proposed an emotion recognition system with consideration of facial expression and physiological signals including SC, finger temperature and heart rate. Previous studies have shown that emotion recognition through the ANS signals is feasible and effective. Gou et al proposed a local mean representation-based k-nearest neighbor classifier (LMRKNN)[7]. However, these methods still do not solve this problem: when k is large, samples in the boundary between classes are easily misclassified. Our work is focused on this problem.

In this paper, four physiological signals (ECG, GSR, EMG and PPG) are collected under four induced emotions (happiness, fear, sadness and anger). The underlying correlations between emotions and the adopted signals are:

1. ECG and PPG are the digital representation of heart activity that can reflect a person's emotional state to a certain degree. The increase in heart rate under fear stimuli is a typical example.

2. Facial EMG is a signal that can be easily measured from the face surface. It is caused by the activity of facial muscle contraction and reflects the state of the nerve muscle. It reflects the emotion change from facial expression.

3. GSR is another signal that can be measured on the body surface and reflects the electrical conductivity of

the skin. The skin conductance level is closely linked to emotion and attention.

WB-KNN is proposed to overcome the drawbacks of traditional KNN. Algorithm 1 presents the pseudo code of the WB-KNN algorithm. To understand the WB-KNN algorithm better, the implementation of the pseudo code is described in detail below.

Algorithm 1 Procedure of the WB-KNN algorithm

REQUIRE: The $N \times M$ training set with S classes, a test sample $x_i \in R^M$, weight vector of features α.

ENSURE: The label of x_i, $y(x_i)$.

1: Calculate the distance $d(x_i, x_j)$ between x_i and the training sample x_j;

2: **for** j=1 to N **do**

$d(x_i, x_j)$=*Distance* (α, x_i, x_j)

end for

3: Sort x_j by $d(x_i, x_j)$ to find the k nearest neighbors $x_i^{(n)}$ and k farthest neighbors $x_i^{[n]}$ of x_i, n=1,...k;

4: Judge whether x_i is in the boundaries between classes according to $x_i^{(n)}$;

5: **if** $x_i \in$ *boundary* **then**

$y(x_i)$=*VotingClassifier*(x_i);

6:**else**

Compute the weight of class $s(\omega_s)$ voted by $x_i^{(n)}$ and $x_i^{[n]}$ using the WDF weighting method;

for s=1 to S

ω_s=*SumWeight*$\{x_i^{(n)} \in s\}$− *SumWeight*$\{x_i^{[n]} \in s\}$;

end for

$y(x_i)$=*argmax*$_s\{\omega_s\}$;

end if

7: **return** $y(x_i)$.

First of all, k nearest neighbors $x_i^{(n)}$ and farthest neighbors $x_i^{[n]}$ of the test sample x_i are computed by the feature-weighted distance formula:

$$d(x_i, x_j) = \sqrt{\sum\nolimits_{m=1}^{M} [\alpha_m (x_{im} - x_{jm})]^2} \ , \tag{1}$$

where α_m represents the weight of feature m.

Then, use the label information of $x_i^{(n)}$ and $x_i^{[n]}$ for the classification of x_i. In traditional KNN algorithm, x_i is classified by a majority vote of its nearest neighbors $x_i^{(n)}$. But the simple majority voting classification has a drawback when the class distribution is skewed. That is, samples of a more frequent class tend to dominate the prediction of the test sample, because they tend to be common among the k nearest neighbors due to their large number, x_i will be misclassified to S_2. An effective approach to overcome the drawback is to weight the classification according to the distance from the test sample to each of its k nearest neighbors.

To overcome the shortcomings of these traditional weighting methods, the WDF weighting method based on the distance and farthest neighbors is proposed. On the one hand, this method assigns the weight to the neighbors based on their distance to the test sample using a standard normal distribution function. The voting weight of x_i in the classification of x_i is υ_{xi}. It is easy to know that the maximum value and the attenuation rate of the function are controlled by σ, which can be adjusted for the dataset with different distribution easily. On the other hand, considering the farthest samples are less likely to the same class as the test sample, the farthest neighbors $x_i^{[n]}$ are used to reduce the weight of non-target class. The label of x_i by the WDF weighting method is denoted by

$$y(x_i) = \arg\max_s \{\sum\nolimits_{n=1}^{k} (\upsilon_{x_i^{(n)}} - \upsilon_{x_i^{[n]}}) \Big| y(x_i^{(n)}) = s \vee y(x_i^{[n]}) = s\} \ . \tag{2}$$

Another improvement is intended to the boundaries between classes. Generally, larger value of k reduces effect of the noise on the classification, but makes boundaries between classes less distinct. A voting classifier is utilized to improve the classification accuracy of samples in the boundaries between classes by combining different base classifiers and using the average predicted probabilities. x_i is determined to be a sample in boundaries between classes, if $\exists S1, S2 \in S$,

$$|count(x_i^{(n)} \in S1) - count(x_i^{[n]} \in S2)| \le \delta, \tag{3}$$

where $count(x_i^{(n)} \in S1)$ represents the number of samples belonging to class $S1$ in $x_i^{(n)}$ and δ is a small integer. Then the label of x_i is:

$$y(x_i) = \arg\max_s \{\sum\nolimits_{l=1}^{L} \beta_l \Big| h_l(x_i) = s, s = 1 \ldots S\} \ , \tag{4}$$

where L is the number of base classifiers, $h_l(x_i)$ is the label predicted by base classifier h_l and β_l is the weight of h_l.

Based on the improvements above, a lot of comparative experiments have been conducted to verify the effectiveness of the WB-KNN algorithm.

To acquire a high quality database of physiological signals, a scientific and reasonable arrangement is made for the experiment from the participants, experimental instruments, emotional stimuli and experimental scene.

The participants include 29 students, 15 males and 14 females, which from our class. They are aged from 18 to 30 years (mean is 22.97, standard deviation is 2.83) and they can express emotions normally.

For emotional stimuli, representative video clips[8] are adopted, as video induction is a multimedia (audio, visual and cognitive) and convenient approach to evoke specific targeted emotions. Specially, we choose a movie clip of Diors Man for happiness, a movie clip of Grudge for fear, a movie clip of the Aftershock for sadness, a movie clip of Silenced for anger and one minute static picture as a transition. Scientific researches indicate that an adult's attention can only be highly concentrated for about 20 min. Considering the participants' patience, each movie clip plays about 4 min in happiness, fear, sadness and anger order. And the internal of each two clips plays one minute static picture to help the participants calm down.

The emotion induction experiment is arranged in a closed and quiet room. Four physiological signals (ECG, GSR, EMG and PPG) are acquired using BIOPAC MP150 system. The sampling rate is set at 200 Hz for all channels. In addition, we also use two computers. One is used to record signals from BIOPAC MP150 system. Another with a camera is used to play emotional stimuli and record the facial expressions of participants synchronously.

At the beginning of the experiment, lab assistant explains the experimental process for participants and help them wear bio-sensors. Then, about 20 min emotion induced video is played. After the experiment, participants are asked to fill out the feedback form about their emotional experience. In the subsequent steps, the reports of the participants' emotional experience and facial expressions will be used for the label of physiological signals.

Physiologically, the physiological signal is so weak that it is easily interfered by noise, the electromagnetic phenomenon of the experimental instrument, the power frequency and the action of participants. Therefore, it is necessary to preprocess the physiological signals in emotion recognition.

First of all, to reduce the required length of signals, each sample is intercepted into a signal segment of 20 s, according to the participants' emotional reports and facial expressions. Subsequent processing is all based on these signal segments.

For ECG and PPG signals, wavelet decomposition is used to remove baseline drift and detect characteristic waveform (R waveform in ECG and P waveform in PPG). Fig.2 shows the preprocessing results of ECG and PPG signals. For EMG signal, a Butterworth low-pass filter with 0.4 Hz is used to denoise. For GSR signal, a Butterworth low-pass filter with 0.3 Hz is employed for smoothing.

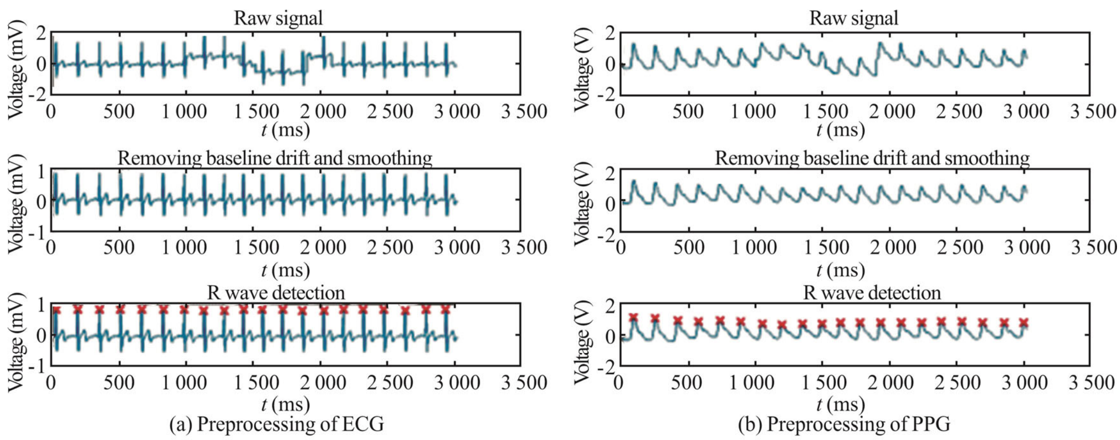

Fig.2 The preprocessing of (a) ECG and (b) PPG signals

We use Augsburg Biosignal Toolbox (AuBT) to process the data. Emotion-relevant features are extracted by combining these underlying features and their statistical features including mean, median, min, max, range and std. Removing meaningless features, there are total 136 features, include 80 ECG features, 20 EMG features, 17 GSR features and 19 PPG features. Then, the physiological dataset is constructed by these features and emotional labels. The number of samples on happiness, fear, sadness and anger are 105, 185, 168 and 165 respectively, this indicates that the class distribution of the physiological dataset is a little skewed, the number of samples on happiness is significantly less than the number of samples on fear.

As we know, the accuracy of the KNN algorithm can be severely degraded by the presence of noisy or irrelevant features, or if the feature scales are not inconformity with their importance. A popular approach is used by feature-weighted distance to express the similarity between samples.

Before calculating the feature weight, data is first normalized to eliminate the influence of feature scale. As physiological datasets are usually nonlinearly separable, KPCA with RBF is utilized to map the data into a high dimension space, extract nonlinear features and then reduce feature dimension, which can possibly generate more distinguishable features.

Generally, the weight of features is expressed by their importance in classification. After KPCA processing, a tree-based estimator is used to evaluate feature importance through Gini coefficient, which in turn can be used to discard irrelevant features. Then the features are ranked in descending order depending to their importance. Fig.3 shows the importance of the top 50 features after ranking. It can be observed that only the limited features have a high importance. One way to eliminate interference of irrelevant features is to only use the few important features for the classification, but this method will cause the loss of feature information. Therefore, feature weighting method is proposed by giving each feature a weight according to their importance in the classification, which can not only eliminate interference of irrelevant features, but also reduce the loss of feature information.

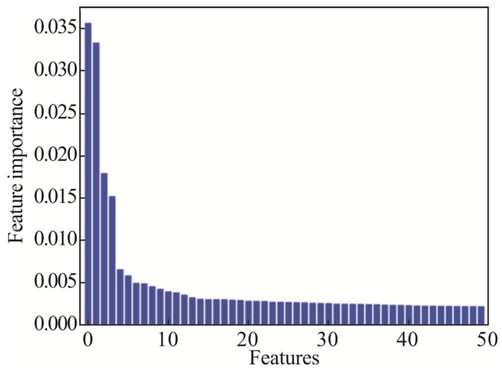

Fig.3 The importance of the top 50 features after ranking

To verify the effectiveness of the improvements to traditional KNN algorithm, a lot of contrast experiments are performed on the physiological dataset by 10-fold cross validation.

Three popular classification algorithms (NB, DT and SVM) are used to construct the voting classifier, since they all have a good performance on the physiological dataset. Their voting weights are assigned as 1, 1.5 and 1 respectively, according to their classification accuracies (87.480 8%, 91.807 5% and 85.558 1%). The voting classifier will be used to classify the samples in the boundaries between classes in the following experiments.

Tab.1 shows the classification accuracy of KNN models with distinct improvements. It can be seen obviously that each improvement has improved the classification accuracy gradually, which validates the effectiveness of our WB-KNN algorithm.

Tab.1 The classification accuracy of KNN models with different improvements

n	KNN (%)	KNN-1 (%)	KNN-2 (%)	KNN-3 (%)
1	78.817 2	78.817 2	79.618 5	83.000 5
3	**89.738 9**	90.542 8	91.021 5	91.984 1
5	88.773 7	90.547 9	91.026 6	93.266 8
10	84.754 2	**91.820 3**	**93.259 1**	**94.219 2**
20	75.762 9	90.064 0	91.500 3	92.135 2
30	64.357 4	87.811 1	89.892 5	90.855 1
50	53.466 5	87.004 6	89.567 3	90.530 0

(KNN-1: feature-weighted KNN; KNN-2: on the basis of KNN-1, using the WDF weighting method to weight the classification (σ=0.000 1); KNN-3: on the basis of KNN-2, using the voting classifier to classify the samples in the boundaries between classes. All these KNN models have the same k value of 7.)

Tab.2 shows the classification accuracy of KNN-1 models with different weighting methods. As we can see, the DF weighting method based on the distance and farthest neighbors has a better performance than traditional weighting methods by tuning the hyper parameter σ, which proves its superiority and tenability.

Tab.2 The classification accuracy of KNN-1 models with different weighting methods

Weight	3NN (%)	4NN (%)	5NN (%)	6NN (%)	7NN (%)
σ=0.01	91.20	91.32	91.09	91.11	91.16
σ=0.001	91.59	91.72	92.13	92.10	91.17
σ=0.000 1	**92.09**	**92.17**	**93.03**	**92.86**	**92.95**

(d: the distance between the test sample and its neighbors; σ=0.01: the WDF weighting method with σ=0.01; 1NN: the feature-weighted KNN (KNN-1) with k=1. All these KNN models are established under n=10.)

In general, the WB-KNN algorithm improves emotion recognition performance over traditional KNN and other classification algorithm (NB, DT and SVM).

This paper proposed the WB-KNN algorithm to surmount the shortcomings of traditional KNN. The improvements mainly focused on three aspects: replacing the standard Euclidean distance with feature-weighted distance; using the WDF weighting method based on the distance and farthest neighbors to weight the classification; adopting the voting classifier to balance out the weakness of traditional KNN in the boundaries between classes. We collected four physiological signals from 29 participants under four induced emotions to investigate the performance of the WB-KNN algorithm. The results of emotion recognition from the physiological dataset with WB-KNN algorithm dramatically outperformed than traditional KNN and other classification algorithm (NB, DT and SVM). Classification accuracy was 94.219 2% on the four categories. As physiological signals are complex, more effective emotion-relevant features should be extracted from multiple physiological signals in the future emotion recognition system.

References

[1] Cover T and Hart P, IEEE Transactions on Information Theory **13**, 21 (1967).

[2] Pedregosa F, Varoquaux G and Gramfort A, Journal of Machine Learning Research **12**, 2825 (2011).

[3] Kreibig S.D, Biological Psychology **84**, 394 (2010).

[4] Kim K.H, Bang S.W and Kim S.R, Medical & Biological Engineering & Computing **42**, 419 (2004).

[5] Li L and Chen J.H, International Conference on Intelligent Information Hiding and Multimedia, 355 (2006).

[6] Chang C.Y, Tsai J.S, Wang C.J and Chung P.C, IEEE Conference on Computational Intelligence in Bioinformatics &Computational Biology, 278 (2009).

[7] Gou J.P, Qiu W.M, Yi Z, Xu Y, Mao Q.R and Zhan Y.Z, ACM Transactions on Intelligent Systems and Technology **10**, 1 (2019).

[8] Koelstra S, Muhl C, Soleymani M, Lee J.S, Yazdani A, Ebrahimi T, Pun T, Nijholt A and Patras I, IEEE Transactions on Active Computing **3**, 18 (2012).