J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 4931--4947

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

A Study on Emotion Recognition Model Incorporating Random Forest Algorithm and Its Improvement on Information System Security Performance

Gaofeng Huang[1,✉], Xiangjun Xu[1]

[1] *School of Electronic and Information Engineering, Wuhan Donghu University, Wuhan, Hubei, 430212, China*

ABSTRACT

The article uses the appropriate equipment for research data, designing the face and physiological signal emotion recognition network respectively, and putting its recognition features into the random forest classifier for training in order to realize the construction work of emotion recognition model. In-depth interpretation of the random forest algorithm based emotion recognition model in the application of information systems, combined with the research data, respectively, the emotion recognition model and system safety performance testing assessment. The emotion recognition model of this paper based on the 25% retention method has a recognition rate of 96.16% for the 14-dimensional B emotion features, which has the highest recognition efficacy and can well meet the system emotion recognition needs. The experimental group is found to be significantly different from the control group, and it is concluded that by introducing the emotion recognition model into the traditional information system, all three security performance indicators of the system are significantly improved.

Keywords: random forest; facial features; physiological signals; emotion recognition model; information system; safety performance

1. Introduction

With the rapid development of artificial intelligence technology, emotion recognition technology is becoming more and more mature, and the emotion recognition model based on physiological signals has become one of the hot spots in current research. Physiological signals are the response mechanisms within the human body, such as electrophysiological signals, physiological changes and biochemical signals, etc. These signals are often closely related to the generation of emotions [1-4]. Random forest, on the other hand, is a classical integrated learning algorithm, which is characterized by its ability to deal with high-dimensional data, solve the overfitting problem, and be able to evaluate

✉ Corresponding author.
E-mail address: hgf001x@163.com (G. Huang).
Received 12 January 2024; Revised 25 March 2024; Accepted 25 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-278

the importance of features [5-6]. The development of emotion recognition models incorporating random forest algorithms has great application prospects in areas such as information system security performance improvement.

In today's digital era, the security of information systems has become a great concern for organizations and individuals. With the popularization of the Internet and the increase of information collection and storage, information systems are facing more and more threats and risks [7-10]. In order to protect the security of information systems, a series of safeguards are essential. First of all, access control is the foundation of information system security, which verifies the identity of users and manages their access rights to system resources [11-14]. Proper access control prevents unauthorized personnel from accessing the system and sensitive data. Common access control measures include strong password policies, multi-factor authentication, role and permission management [15-16]. Secondly, data encryption is an important means of data security, which ensures the confidentiality of data by converting it into a form that cannot be read by unauthorized parties. Common data encryption methods include symmetric encryption, asymmetric encryption, and digital certificates [17-20].

Literature [21] points out the prevalence of cyber-attacks in the era of Industry 4.0 and examines the impact of information system security practices on risk management and supply chain performance in online supply chains, revealing that operations have a direct or indirect impact on supply chain performance. And governance directly affects supply chain agility and indirectly influences supply chain performance. Literature [22] emphasized the importance of information security and illustrated that transformational leadership of information security managers can improve the effectiveness of information security. It also suggests the need for leadership education programs and points out the important role of job training for information security managers. Literature [23] identifies deficiencies in current university information systems and suggests feasible recommendations to mitigate information security risks, including. Conducting training, developing, implementing and enforcing anti-virus security policies, etc. to ensure data protection. Literature [24] examined the impact of specific phases of information security management system on information security performance. It also described the impact of security PCDA cycle model on information security performance. Literature [25] explored effective methods to measure the level of information security and cyber security performance, especially the use of private cloud and its recommendations. Recommendations are also made based on the current shortcomings. It provides reference to the application of information security in universities. Literature [26] emphasizes the development of emotion recognition technology, its application, and the risk of privacy and public security it raises. It also systematically studied the implications of emotion recognition technology and discussed the delicate balance between improving public safety and protecting individual privacy.

The above studies emphasize the importance of information security and propose appropriate measures based on the deficiencies of the current information security system, but the improvement of information system security performance by incorporating the emotion recognition model of random forests is not reflected in these studies. This indicates that although the security of information systems has been emphasized, advanced algorithms and techniques have not been used in the protection of information security.

The article sets the sampling frequency of the data acquisition equipment, and the subjects are tested using the equipment to realize the work of emotion feature acquisition. The face emotion recognition network and physiological signal emotion recognition network are designed respectively, and the random forest algorithm is used as the feature classifier, and the recognized features are put into the classifier for training to complete the task of constructing the emotion recognition model, and in order to improve the processing speed of the model, the recognized features are screened. In response to the existence of limitations of traditional information systems. In this regard, the security improvement of information system by integrating emotion recognition model is proposed. The experimental samples are selected, groups are set up, and the security performance indexes of the two groups are compared and analyzed to confirm the role of the emotion recognition model in improving the security of the information system.

2. Emotion recognition model construction and its application

2.1. Emotion Recognition Model

In this paper, the overall flow of the model of the bimodal emotion recognition method combining facial expressions and physiological signals If the two results are the same, the classification results will be used as the final emotion classification results; if the results of the two sensors are different, then firstly, feature dimensionality reduction is performed on the facial expression features, so that the number of facial expression features is the same as the number of physiological signal features, and finally, through the Random Forest, the facial expression signal features are classified with the physiological signal features for the feature fusion classification, to get the final emotion classification results of the different samples of the two sensors.

2.1.1. Data acquisition. The data source for this study was the face recognition data collected through the emotion experiment, and before the emotion data was collected (in order to avoid contamination by the latter), 4-minute EEG data was collected at hourly intervals as a third type of neutral data. The entire experiment was conducted using equipment for face recognition image acquisition with a sampling frequency of 256 Hz. as a result, a total of 24 minutes of face recognition images were recorded from the subjects with 138,120 data points.

2.1.2. Face emotion recognition network. VGG-16 network is a representative classical network model in CNN, and it is the most widely used among VGGNet networks.VGG-16 convolutional neural network mainly includes the following structures: convolutional layer, pooling layer, fully connected layer, and softmax classifier. The convolutional layer is the key structure of the convolutional neural network. When the data is input into the convolutional layer, the features in it are extracted by the convolutional operation of the convolutional kernel. Convolution operation is a linear, translation invariant operation, it will be all the input and output information for the local weighted combination of the process of feature extraction: the input image data as an n-dimensional matrix, and then use a $m \times m$-dimensional ($m < n$) convolution kernel, from the upper-left position of the image in accordance with the order of the slide from left to right, from the top to the bottom of the order of the slide, every slide, the corresponding elements of the corresponding elements will be multiplied together after the addition of the summation can be The new feature matrix can be regarded as a representation of the input features, and the new feature matrix is called feature map or feature mapping.

Convolutional kernels are commonly used in the sizes of 3×3, 5×5 and 7×7. Larger convolutional kernels can obtain a larger receptive field, but replacing larger convolutional kernels with several smaller ones can reduce the parameters while obtaining the same receptive field, so most of the convolutional neural networks use smaller convolutional kernels.

The pooling layer reduces the spatial size of the feature mapping by downsampling operations. Neighboring pixels in an image often have similar values, so the feature mapping obtained through the convolutional layer usually has similar values, which can make the extracted features contain redundant information, to prevent model overfitting and improve the generalization ability of the model, a pooling layer is added after the convolutional layer in many convolutional neural networks. Average pooling and maximum pooling are more commonly used pooling functions, the output of average pooling is the average of all values in a region in the input feature mapping, corresponding to smoother features; while the output of maximum pooling is the maximum value in the region, corresponding to more significant features.

The main purpose of the fully connected layer (FC) in the deep network is to enable the network to carry out correct classification, usually placed at the end of the model structure, the input features to do feature weighting, that is, to do an integration of the features, and then the integrated feature image for downsizing operations.FC does not specify the size of the input image size, as long as the last layer of the network in the downsampling of the image size can be enough.

The softmax classifier is usually placed after the fully connected layer, and in the classification task, its role is to normalize the input vector and calculate the probability value of the corresponding category, so that the probability of each category is distributed between (0, 1), and the category with the highest probability value is the final prediction of the image category, the formula of softmax is as follows:

$$p(y|x)=\frac{e^{h(x,y_i)}}{\sum_{j=1}^{n}e^{h(x,y_i)}} \tag{1}$$

There are five convolutional layers in the VGG-16 network, two convolutional kernels of size 3×3 are stacked in both the first and second convolutional layers with channels 64 and 128 respectively, and the last three convolutional layers are stacked using three 3×3 convolutional kernels, all with channels 512.The adjacent convolutional blocks are connected using a maximal pooling layer, which reduces the spatial size of the feature mapping through the pooling operation to enhance the image feature invariance while reducing redundant information and computational parameters, enhancing the model generalization ability, and using the Re LU activation function to prevent gradient vanishing when training the model. The features obtained from the above steps are passed through three FC and one softmax layer to output the classification results.

2.1.3. Physiological signals emotion recognition network:

1) Feature extraction. The face video is converted into a continuous image sequence, and the face region of each image frame is extracted by the face detection algorithm. The forehead region of the face is selected as the ROI region to extract the IPPG signal, and the key point detection model in the Dlib library is used to obtain the key points of the forehead region of the face, and the rectangular box of the forehead region is drawn according to the key points. The color channel separation is performed on the forehead region image, and the images of three different channels, R, G and B, are obtained respectively.

Taken together, this paper uses eight feature parameters for recognizing emotions, five time domain feature parameters, SDNN, PNN50, PNN20, RMSSD, SDSD, and three frequency domain feature parameters, LF, HF, and LF/HF, respectively. The computation of HRV features is described in detail in the following. Before analyzing the heart rate variability, it is necessary to obtain the RR interval sequence, which is the time between two peaks. The IPPG signal can be obtained by the previous method and analyzed for heart rate variability. Detect the wave peaks of IPPG signal, record the value of the horizontal coordinate corresponding to the peak point, and calculate the time interval between the two peak points to obtain the RR interval. Based on the RR interval, time domain and frequency domain features were extracted.

2) Time domain features:

$$SDNN=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(RR_i-\overline{RR}\right)^2} \tag{2}$$

SDNN is the standard deviation of all RR intervals, which reflects the deviation of heart rate variability from the mean. where $\overline{RR}$ is the mean of the RR intervals and RR_i denotes the ird RR intervals. I.e:

$$PNN20 = \frac{NN20}{N} \times 100\% \tag{3}$$

$$PNN50 = \frac{NN50}{N} \times 100\% \tag{4}$$

PNN20 is the percentage of the number of RR intervals of 20 ms or more that reflect parasympathetic activity.The meaning and calculation of PNN50 is similar to that of PNN20.

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}\left(RR_{i+1} - \overline{RR_i}\right)^2} \tag{5}$$

RMSSD represents the root mean square of the difference between adjacent RR intervals, which can be used to measure the effect of the parasympathetic nervous system on heart rate. To wit:

$$SDSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}\left(RR_{i+1} - \overline{RR_i}\right)^2} \tag{6}$$

SDSD denotes the standard deviation of the RR interval difference, which more finely represents the overall change in HRV.

3) Frequency domain characterization. Frequency domain analysis of the IPPG signal using Welch power spectrograms was implemented based on the fast Fourier transform at the end of the averaging window. The signal of length N is divided into K intervals with M data present in each interval. The window function n is added to the time series and the periodogram for each part is:

$$I_i(\omega) = \frac{1}{U}\left|\sum_{n=0}^{M-1} x_i(n)\, w(n)\, e^{-i\pi fn}\right|^2, i = 1, 2, \cdots, M-1 \tag{7}$$

Among them:

$$U = \frac{1}{M}\sum_{N=0}^{M-1} \omega^2(n) \tag{8}$$

Treating the results of each component derived from the above calculations as uncorrelated yields the final power spectrum estimate:

$$P_\omega(f) = \frac{1}{K}\sum_{i=1}^{K} I_i(\omega) \tag{9}$$

2.1.4. Characterization. The merit of the classification model used in the emotion recognition process is directly related to the accuracy of the recognition results. Four machine learning models, Support Vector Machine, K-Nearest Neighbor Algorithm, Decision Tree and Random Forest, have been widely validated in the research related to emotion recognition, while Random Forest used in this paper is used as a classifier for emotion recognition.

Random forest is constructed by multiple different decision numbers, and the decision trees are independent of each other. K samples are randomly selected from a dataset, and M decision trees are generated by sample features and information gain, and the final output is determined by the output category with the largest number of output categories among the M decision trees. For test sample a, the number of output categories is c, then the output $F(a)$ of the random forest recognition model can be expressed as equation (10):

$$F(a) = \frac{\arg\max}{i = 1, 2, \cdots, c}\left\{\sum_{i=1}^{M}\left[f_m(a) = i\right]\right\} \tag{10}$$

where $f_m(a)$ the output category of the m nd decision tree.

2.1.5. Mathematical modeling. For the data whose facial expression and physiological signal classification results are different, it is also necessary to use the classifier to classify their emotions. As can be seen from the above, the classifier is mainly used for emotion classification when the classification results of the two sensors are not the same, and at this time, the decision-making results of facial expression may not be good enough, thus, in this paper, we choose the random forest classification model to classify the feature fusion of facial expression and physiological signal features.

Random forest (RF) is an integrated machine learning method, which is a combination of random repeated sampling statistical method bootstrap and node random splitting method to construct multiple decision trees, and then get the classification results by voting [27-29]. Random forest can analyze complex features with good robustness and faster speed, so in this paper, random forest is chosen as the classifier for dual sensor feature fusion. Random forest is a classifier integrated by many decision trees, the specific steps to generate random forest are:

Step 1: Extract k samples from the original training set to form a subsample set by bootstrap method, i.e., random sampling method with put-back.

Step 2: Generate a decision tree for the extracted sample set. Randomly and unrepeatedly select no more than m features out of the total number of features, and among these features select the most suitable for classification for node splitting and let each tree grow freely.

Step 3: Repeat steps 1 and 2 n times to obtain n decision trees.

Step 4: Combine the n decision trees obtained in Step 3 to become a random forest, use this forest to predict the classification of the test sample, and decide the prediction result by ballot method.

2.1.6. Characterization. The total number of data points for all faces is 138,120, which corresponds to 340 features. Obviously, not all of the features are effective for the classification result, and this kind of features is called "redundant features". The existence of "redundant features" will not only make the classification results bad, but also lead to slow computing speed, so this subsection uses the feature selection algorithm based on the importance of the features to filter the features, so the random forest classifier is chosen for the Wrapper-style feature filtering.

Feature importance assessment is to look at the contribution made by each feature in each tree in the forest from the random forest, followed by taking the average value that is the feature importance of the feature. In this paper, Gini coefficient is used as the evaluation index of feature importance, FI is used to denote feature importance and GI is used to denote Gini coefficient, and the number of features is M in total, then the steps for solving the importance of the j st feature in the random forest FI_j are:

Step 1: Find the Gini coefficient of node q GI_q :

$$GI_q = \sum_{c=1}^{C}\left(\sum_{C' \neq c} p_{qc} p_{qc'}\right) = 1 - \sum_{c=1}^{C} p_{qc}^2 \tag{11}$$

where C denotes the total number of categorized categories and p_{qc} denotes the probability that the categorization result on node q is c. Its physical meaning is the probability that the category labeling results of two samples randomly drawn from node q are inconsistent.

Step 2: Find the importance FI_{jq} of the j th feature at node q, that is, find the amount of change in the Gini coefficient at node q before and after the branch:

$$FI_{jq} = GI_q - GI_{ql} - GI_{qr} \tag{12}$$

Step 3: Solve for the importance of the j st feature on the i nd decision tree FI_{ji} :

$$FI_{ji} = \sum_{q \in Q} FI_{jq} \tag{13}$$

Where Q is the set of nodes in which feature j appears in that decision tree.

Step 4: Find the importance of the j rd feature in the random forest FI_j:

$$FI_j = \sum_{i=1}^{N} FI_{ji} \tag{14}$$

and normalize it with the final result:

$$FI_{jl} = \frac{FI_j}{\sum_{i=1}^{M} FI_j} \tag{15}$$

where M is the total number of features.

2.2. Emotion Recognition Modeling in Information Systems

In the information system operation and maintenance management, the organization can try to take the state secret algorithm of face recognition identity authentication and access control technology, using the built-in or external camera, boot start to take the local face plus fingerprints dual biological factor authentication, connecting to the network to take face recognition authentication and local face authentication key verification, refresh, through the camera to continuously extract the user's face information during the use of the information equipment Perceptionless background authentication is carried out at any time to ensure that the user's identity and access privileges are legal and compliant throughout the process, and imposters are eliminated. In addition, face recognition can also be applied in other scenarios, for example, if external personnel enter a classified place without a bystander accompanying the personnel, the face recognition system will alarm and capture; and conduct full monitoring and auditing of other equipment requiring identity identification and access control, such as cryptographic file cabinets.

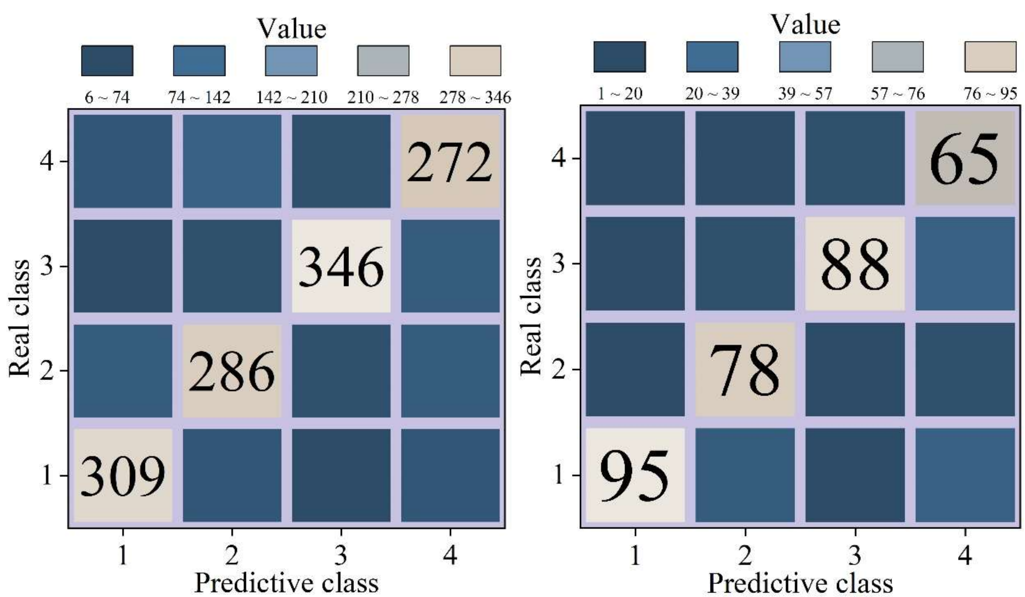
The traditional information system identity identification and access control measures only focus on one-time password authentication of the equipment. In this process, the confidentiality measures can be taken to stay in the “control their own passwords, control their own U-key” and other educational level, there is no direct and effective means of human management in technology. Theoretically, as long as the impostor illegally obtains access to the information system once, he can illegally access or even steal confidential and sensitive information for a long time.

The emotion recognition model has the following advantages over traditional identity identification and access control measures. First, it can effectively prevent the situation of impostor in the intermediate links. For example, if a person is replaced in the middle of the terminal use process, the information system of human-fused emotion recognition model can immediately recognize the violation, lock and record the violation log. Secondly, it can make it easier and faster for legitimate users to log in and access the system with a better experience. For example, users do not need to enter user names and passwords, and can directly brush their faces to log in senselessly. Third, it has anti-snooping and anti-screen photo functions. The information system integrating the emotion recognition model can capture the face of peeping tom, cell phone camera taking pictures and other violations throughout the whole process, and lock the screen alarm in time, leaving the violation log.

3. Emotion Recognition Modeling and System Safety Performance Testing

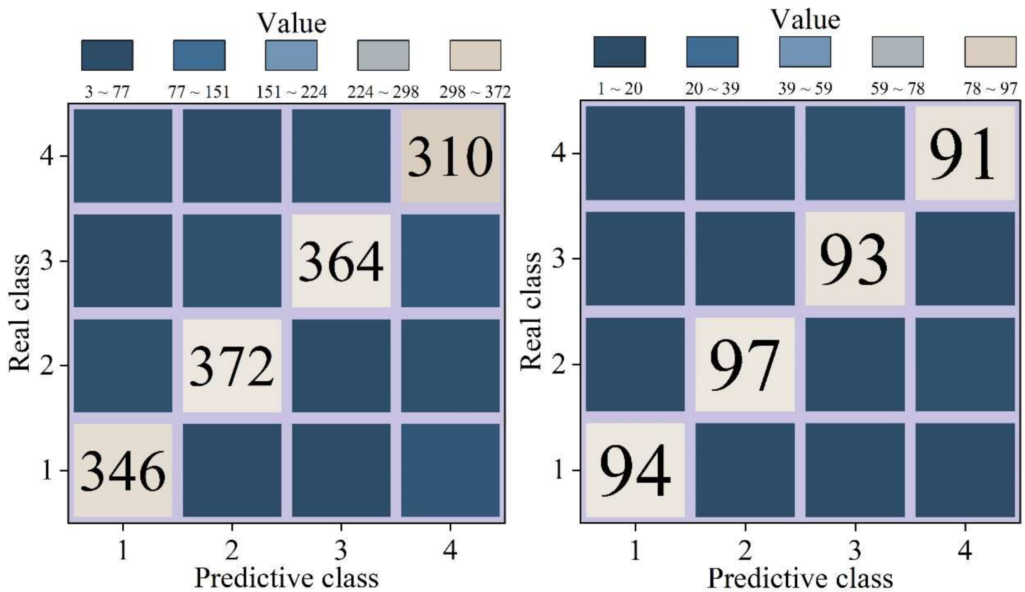
3.1. Emotion Recognition Model Detection

3.1.1. Model testing and results. According to the principle of decision tree (DT) and the model in this paper, the corresponding Matlab program was written for algorithm simulation and the 25% reservation method and 5-fold cross-verification were used to classify and identify the A and B features (7-dimensional and 14-dimensional) under the four emotional states (happy, angry, sad and happy) of the dataset, and the confusion matrix of face feature classification accuracy of the decision tree algorithm based on 5-fold cross-validation and 25% reservation method verification is shown in Fig. 1(a) and (b), respectively, based on 5-fold cross-validation and 25% The 7-dimensional B-feature classification accuracy confusion matrix of the decision tree algorithm verified by the retention method is shown in Fig. 2(a) and (b), respectively, and the 14-dimensional B-feature classification accuracy confusion matrix of the decision tree algorithm verified by the 5-fold cross-validation and 25% retention method is shown in Fig. 3(a) and (b), respectively. The confusion matrices of the A-feature classification accuracy and confusion matrices of the proposed model based on 5-fold cross-validation and 25% retention method are shown in Fig. 4(a) and (b), respectively. The 7-dimensional B-feature classification accuracy confusion matrices of the proposed model based on 5-fold cross-validation and 25% retention method are shown in Fig. 5(a) and (b), respectively. Based on the 5-fold cross-validation and 25% retention method, the 14-dimensional B-feature classification accuracy confusion matrix of the proposed model is shown in Fig. 6(a) and (b), respectively. The F1-Score results of A-feature classification for each emotional state based on the decision tree and the model in this paper are shown in Fig. 7, the F1-Score results of B-feature classification for each emotional state based on the decision tree algorithm are shown in Fig. 8, the F1-Score results of B-feature classification for each emotional state based on the model in this paper are shown in Fig. 9, and the accuracy of the 7-dimensional A, 7-dimensional B, and 14-dimensional B feature classification of the decision tree based on 5-fold cross-validation and 25% retention method verification and the sentiment recognition model in this paper is shown in Table 1.



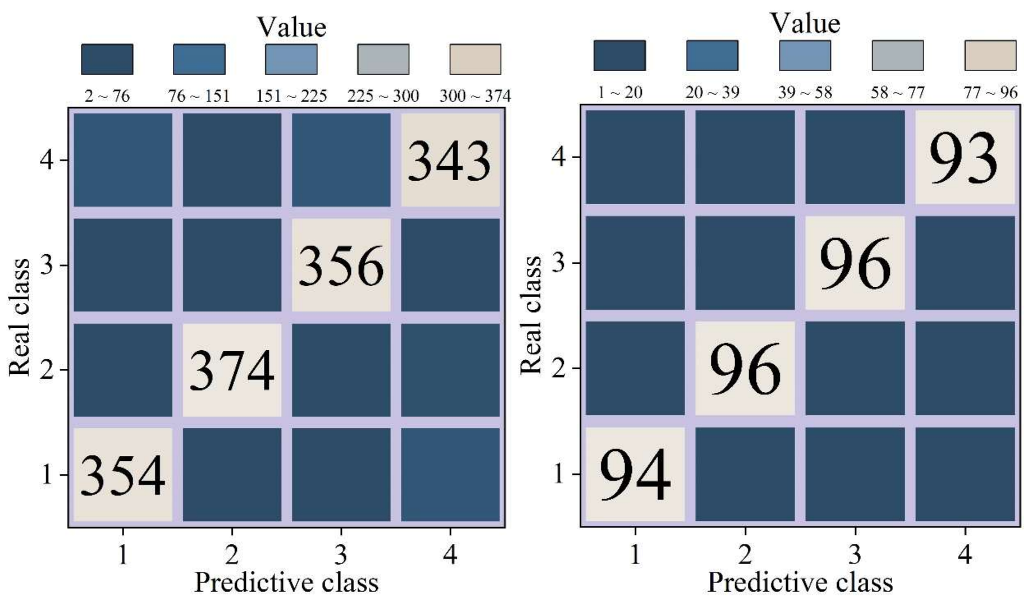
(a)5 fold cross verification (b)25% set asiB for forensic testing

Fig. 1. A feature classification accuracy confusion matrix based on DT



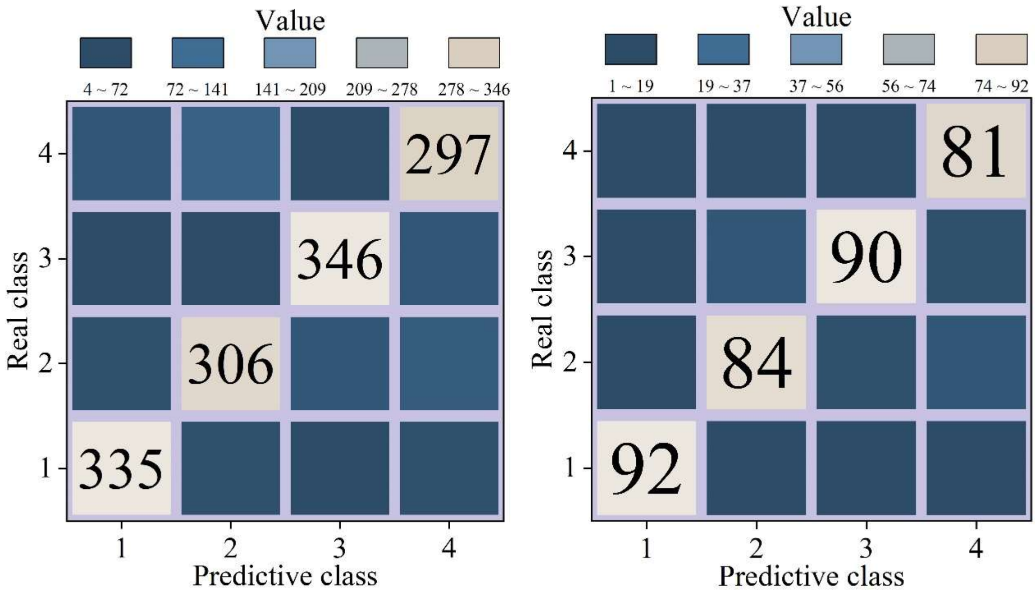
(a)5 fold cross verification (b)25% set asiB for forensic testing

Fig. 2. Accuracy confusion matrix of 7-D B feature classification based on DT



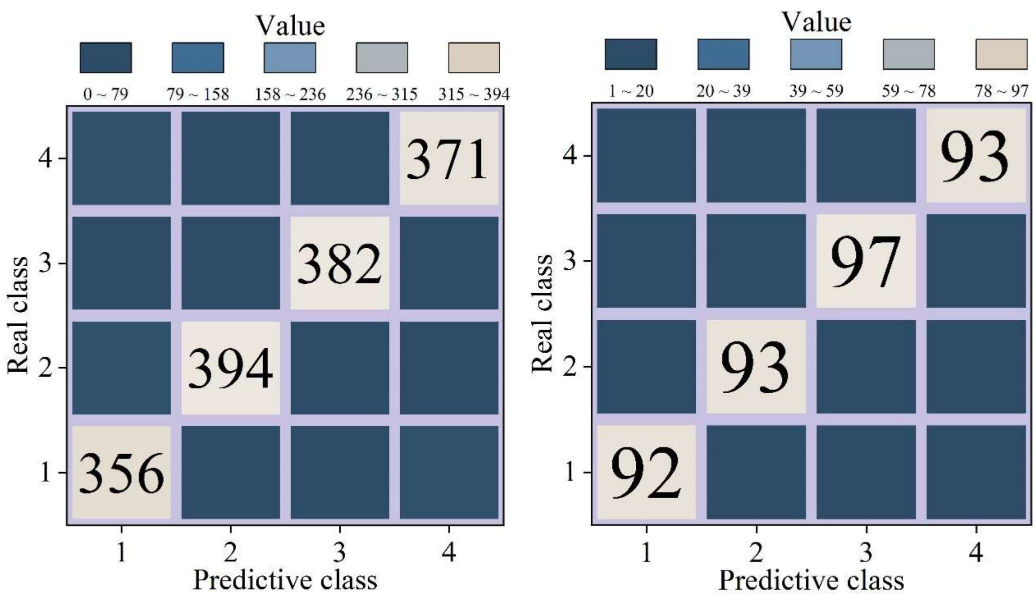
(a)5 fold cross verification (b)25% set asiB for forensic testing

Fig. 3. Accuracy confusion matrix of 14-D B feature classification based on DT



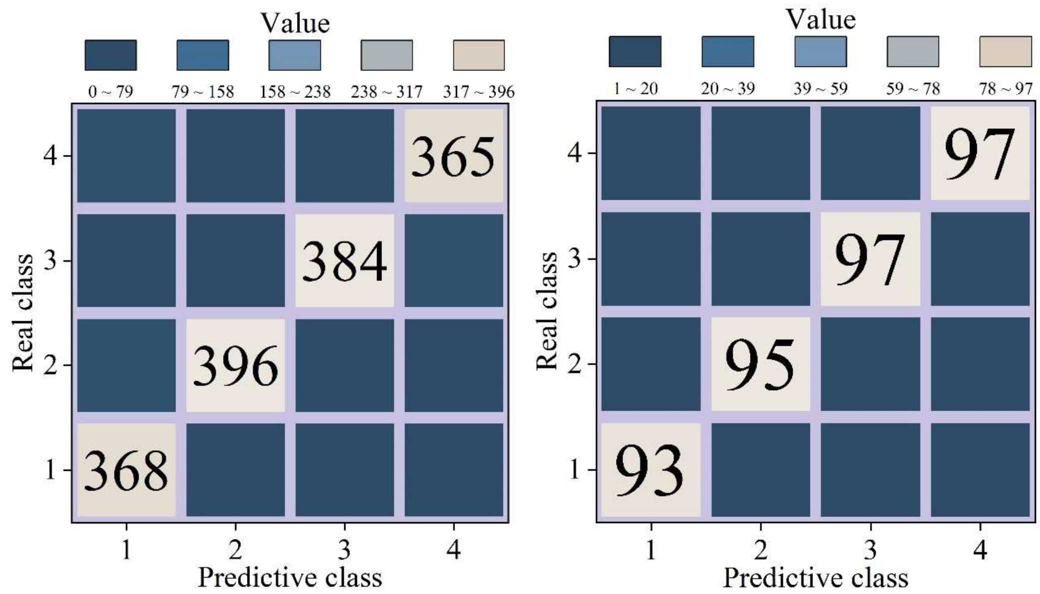
(a)5 fold cross verification (b)25% set asiB for forensic testing

Fig. 4. A feature classification accuracy confusion matrix based on Our model



(a)5 fold cross verification (b)25% set asiB for forensic testing

Fig. 5. Accuracy confusion matrix of 7-D B feature classification based on Our model



(a)5 fold cross verification (b)25% set asiB for forensic testing

Fig. 6. Accuracy matrix of 14-D B feature classification based on Our model

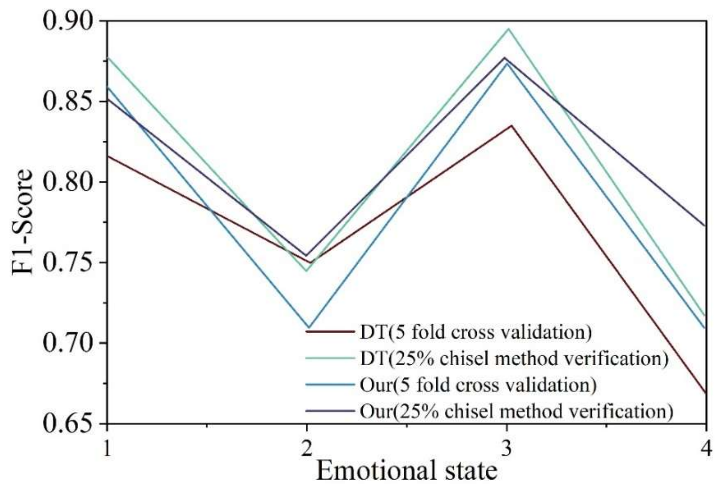


Fig. 7. Comparison of F1-Score results of DT and Our in A feature classification

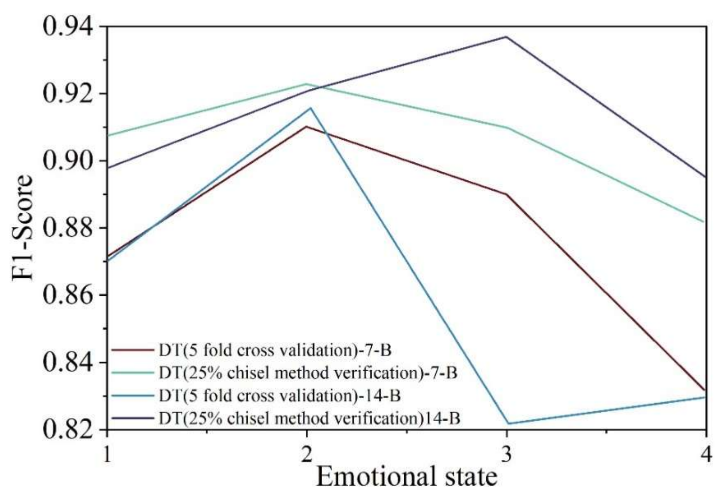


Fig. 8. DT-based algorithm of B feature classification F1-Score results

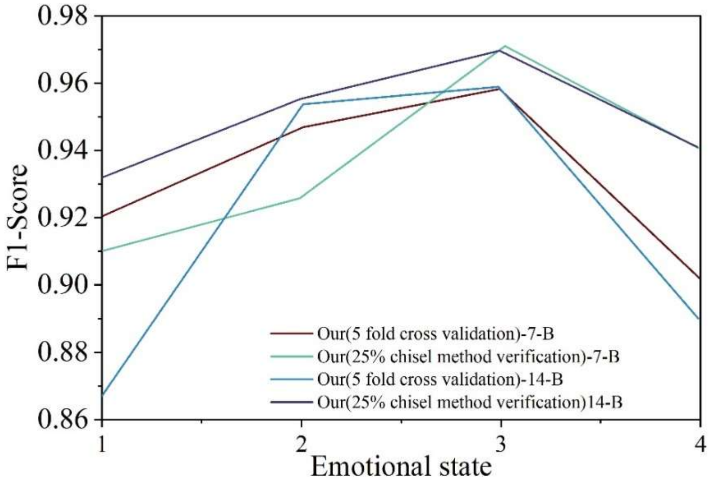


Fig. 9. Our-based algorithm of B feature classification F1-Score results

Table 1. Classification recognition accuracy based on DT and random Our model

Dimension	DT		Average classification accuracy of DT	Our		Average classification accuracy of Our model
	5 fold cross verification	25% set asiB for forensic testing		5 fold cross verification	25% set asiB for forensic testing	
A	76.62%	82.37%	79.50%	78.26%	82.49%	80.38%
7 dimensional B	88.56%	91.42%	89.99%	93.36%	94.56%	93.96%
14 dimensional B	89.64%	92.48%	91.06%	94.46%	96.16%	95.31%

3.1.2. Results and analysis. It can be seen from the recognition accuracy confusion matrix and Table 1 that for the DT classification algorithm, the accuracy of the classification and recognition of sentiment using the B feature is 10.50% higher than that of the A feature with the same feature vector dimension, and the accuracy of the classification and recognition of emotion using the 14-dimensional B feature is higher, with an average classification accuracy of 91.06%, which indicates that the classification and recognition of emotion using the 14-dimensional B feature is the most effective. Meanwhile, the classification accuracies based on the 25% leave-out method are all 3.82% higher than the classification accuracies using the 5-fold cross-validation method on average, which indicates that the DT algorithm using the 25% leave-out method is more effective in recognizing the emotional state using the features. For the model in this paper, under the same dimension of feature vectors, the accuracy of classifying and recognizing emotions using B features is 13.59% higher than that using A features on average, and the accuracy of classifying and recognizing emotions using 14-dimensional B features is even higher, with an average classification accuracy of 95.31%, which indicates that the classification and recognition of emotions using 14-dimensional B features is the most effective. Meanwhile, the classification accuracy based on the 25% leave-out method is higher than that using the 5-fold cross-validation method by an average of 2.38%, thus indicating that this paper's model using the 25% leave-out method is more effective in recognizing emotion states using the features. In summary, the model of this paper validated using the 25% leave-out method is most effective in recognizing emotional states for B features.

3.2. Information system security performance testing

3.2.1. Experimental samples. According to the above subsection 2.2, it can be seen that the emotion recognition model can enhance the security performance of information systems, in order to further substantiate the above research point, this study covers a total of 50 cases, which are randomly divided into experimental and control groups, each containing 25 cases each. The experimental group received an information system that incorporated the emotion recognition model, while the control group continued to use an information system that did not incorporate the emotion recognition model, and the following is a comparison of the overall data between the experimental group and the control group.

3.2.2. Analysis of results. In this study, the information security risk of the experimental group and the control group was comprehensively assessed, and the observation indexes included the three dimensions of privacy, experience, and resistance to leakage, and the results of the information system security performance testing are shown in Table 2. SPSS statistical software was used to analyze the data in this study. For continuous variables, mean and standard deviation were used for description. For categorical variables, frequencies and percentages were used for description. Comparisons between the experimental group and the control group were made using independent samples t-test or non-parametric test, selected based on data normality and variance chi-square, and the significance level was set at $p<0.05$. In terms of the system privacy indexes, the experience indexes and the anti-leakage indexes, the mean values of the experimental group were significantly higher than those of the control group and the difference was significant ($p<0.05$), which indicates that by adding the emotion recognition to the traditional information management system model in traditional information management system, all security performance evaluation indexes of the system are improved. For example, there is a lack of sufficient security in some places, and there are considerable hidden dangers if users are coerced into identity authentication during identification. Therefore, when the user uses the identification system, the user's facial expression is detected to assist in coercion detection. When the system detects the user's fearful expression, the system issues an early warning to alert the background staff.

Table 2. Information system security performance test results

Case	Control group			Experimental group		
	Privacy	Sense of experience	Leak resistance	Privacy	Sense of experience	Leak resistance
1	2.96564	2.76318	2.40108	3.86737	3.40154	3.52309
2	2.77672	2.58884	2.89888	3.1284	3.29933	3.65423
3	2.32168	2.33807	3.16333	2.94415	3.43308	3.59385
4	2.32252	2.51455	2.32341	3.33044	3.62127	3.83797
5	2.03719	2.17124	2.76245	3.10931	2.50948	4.3129
6	2.76382	2.68936	3.21788	3.7309	3.21919	2.84073
7	2.76948	3.04927	2.66881	3.6612	3.21597	4.17887
8	2.02023	2.3251	2.54828	3.93463	3.83803	3.41165
9	2.53936	2.32252	2.84162	3.58692	3.87453	3.82018
10	2.10888	2.65052	3.04196	3.40643	3.86227	3.78479
11	2.71204	2.61475	2.60163	2.93803	3.83614	3.4046
12	2.08533	2.55192	3.23884	3.80876	3.49685	3.22177
13	2.87057	2.05983	2.82419	3.83154	3.31033	2.83529
14	2.35011	2.61221	1.98977	4.20148	3.56275	2.86016
15	3.3105	2.85333	2.68153	3.46729	4.39951	3.68843
16	2.52859	1.92611	2.82143	3.46448	3.75093	3.16969
17	2.10401	2.88532	2.85027	3.09474	4.02084	3.9585
18	2.56861	2.61407	2.58938	3.982	3.63763	3.41056
19	2.45177	2.47144	2.87123	3.7768	3.00307	4.82186
20	1.97467	2.13174	2.93214	3.2573	3.9168	3.40168
21	2.43294	2.5626	2.27374	3.16747	3.79724	4.01005
22	2.6145	3.17526	2.50097	4.50259	4.65813	3.75394
23	2.19007	3.10536	2.53292	3.31158	3.30415	3.44428
24	2.40757	2.55712	2.52073	3.95731	3.54781	4.3711
25	2.17206	2.74357	2.41348	2.67963	2.91891	3.75493
26	2.28128	2.53407	2.49716	3.38554	3.46472	3.59488
27	2.5791	2.35694	2.84151	3.74186	3.10118	4.67756
28	2.40233	2.39481	1.86912	3.50409	3.6059	3.77125
29	3.02386	2.3593	1.96738	3.64639	3.71887	3.34656
30	1.54686	2.47226	2.22599	3.39009	3.10802	3.814
Total	2.48±0.312	2.52±0.306	2.67±0.329	3.45±0.41	3.57±0.38	3.68±0.41

4. Conclusion

In this paper, we first take the data collection as the starting point, formulate the emotion recognition network of face and physiological signals respectively, use the random forest algorithm to complete the work of classification of emotional features, and finally complete the work of constructing the emotion recognition model, and then in order to enhance the security performance of the information system, we propose to improve and optimize the system by using the emotion recognition model, and based on the research dataset, we explore the emotion recognition model and the security performance of the information system.

1) Regardless of the kind of algorithm used, the recognition effect of B features is better than A features. Compared with the effect of emotion recognition based on decision tree algorithm (DT), the model of this paper has a higher priority in emotion recognition, and in addition the model of this paper, which is validated by adopting the 25% leave-out method, has the most significant effect of emotion recognition on B features.

2) Adding emotion recognition model on the basis of traditional information system, the system privacy index, experience index and anti-leakage index are significantly improved, while $P<0.05$, indicating that the emotion recognition model has a significant positive correlation on the security performance of information system.

Funding

This work was supported by Science and Technology Research Project of Hubei Provincial Department of Education (No.B2023292).

References

[1] Canal, F. Z., Müller, T. R., Matias, J. C., Scotton, G. G., de Sa Junior, A. R., Pozzebon, E., & Sobieranski, A. C. (2022). A survey on facial emotion recognition techniques: A state-of-the-art literature review. Information Sciences, 582, 593-617.

[2] Saxena, A., Khanna, A., & Gupta, D. (2020). Emotion recognition and detection methods: A comprehensive survey. Journal of Artificial Intelligence and Systems, 2(1), 53-79.

[3] Egger, M., Ley, M., & Hanke, S. (2019). Emotion recognition from physiological signal analysis: A review. Electronic Notes in Theoretical Computer Science, 343, 35-55.

[4] Rim, B., Sung, N. J., Min, S., & Hong, M. (2020). Deep learning in physiological signal data: A survey. Sensors, 20(4), 969.

[5] Rigatti, S. J. (2017). Random forest. Journal of Insurance Medicine, 47(1), 31-39.

[6] Genuer, R., Poggi, J. M., Genuer, R., & Poggi, J. M. (2020). Random forests (pp. 33-55). Springer International Publishing.

[7] Jouini, M., Rabai, L. B. A., & Aissa, A. B. (2014). Classification of security threats in information systems. Procedia Computer Science, 32, 489-496.

[8] Rahman, M. M., Faraji, M. R., Islam, M. M., Khatun, M., Uddin, S., & Hasan, M. H. (2024). Gravitating towards information society for information security in information systems: a systematic PRISMA based review. Pakistan Journal of Life and Social Sciences (PJLSS), 22(1).

[9] Zafar, H. (2013). Human resource information systems: Information security concerns for organizations. Human Resource Management Review, 23(1), 105-113.

[10] Berdik, D., Otoum, S., Schmidt, N., Porter, D., & Jararweh, Y. (2021). A survey on blockchain for information

systems management and security. Information Processing & Management, 58(1), 102397.

[11] Tourani, R., Misra, S., Mick, T., & Panwar, G. (2017). Security, privacy, and access control in information-centric networking: A survey. IEEE communications surveys & tutorials, 20(1), 566-600.

[12] Qiu, J., Tian, Z., Du, C., Zuo, Q., Su, S., & Fang, B. (2020). A survey on access control in the age of internet of things. IEEE Internet of Things Journal, 7(6), 4682-4696.

[13] Ouaddah, A., Mousannif, H., Abou Elkalam, A., & Ouahman, A. A. (2017). Access control in the Internet of Things: Big challenges and new opportunities. Computer Networks, 112, 237-262.

[14] Ravidas, S., Lekidis, A., Paci, F., & Zannone, N. (2019). Access control in Internet-of-Things: A survey. Journal of Network and Computer Applications, 144, 79-101.

[15] Zhang, P., Liu, J. K., Yu, F. R., Sookhak, M., Au, M. H., & Luo, X. (2018). A survey on access control in fog computing. IEEE Communications Magazine, 56(2), 144-149.

[16] Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., & Koucheryavy, Y. (2018). Multi-factor authentication: A survey. Cryptography, 2(1), 1.

[17] Atadoga, A., Farayola, O. A., Ayinla, B. S., Amoo, O. O., Abrahams, T. O., & Osasona, F. (2024). A comparative review of data encryption methods in the USA and Europe. Computer Science & IT Research Journal, 5(2), 447-460.

[18] Barker, E., & Mouha, N. (2017). Recommendation for the triple data encryption algorithm (TDEA) block cipher (No. NIST Special Publication (SP) 800-67 Rev. 2 (Draft)). National Institute of Standards and Technology.

[19] Nurullaev, M., & ALOEV, R. D. (2020). Software, algorithms and methods of data encryption based on national standards. IIUM Engineering Journal, 21(1), 142-166.

[20] Verma, G., Liao, M., Lu, D., He, W., Peng, X., & Sinha, A. (2019). An optical asymmetric encryption scheme with biometric keys. Optics and Lasers in Engineering, 116, 32-40.

[21] Fernando, Y., Tseng, M. L., Wahyuni-Td, I. S., de Sousa Jabbour, A. B. L., Chiappetta Jabbour, C. J., & Foropon, C. (2023). Cyber supply chain risk management and performance in industry 4.0 era: information system security practices in Malaysia. Journal of Industrial and Production Engineering, 40(2), 102-116.

[22] Choi, M. (2016). Leadership of information security manager on the effectiveness of information systems security for secure sustainable computing. Sustainability, 8(7), 638.

[23] Kwaa-Aidoo, E. K., & Agbeko, M. (2018). An Analysis of Information System Security of a Ghanaian University. International Journal of Information Security Science, 7(2).

[24] Jang, S. S., Lee, S. J., & Noh, B. N. (2012). The effects of the operation of an information security management system on the performance of information security. Journal of The Korea Institute of Information Security & Cryptology, 22(5), 1123-1132.

[25] Computing, C. O. P. C. (2019). Measuring information security and cybersecurity on private cloud computing. Journal of Theoretical and Applied Information Technology, 96(1).

[26] Ortiz-Clavijo, L. F., Gallego-Duque, C. J., David-Diaz, J. C., & Ortiz-Zamora, A. F. (2023). Implications of Emotion Recognition Technologies: Balancing Privacy and Public Safety. IEEE Technology and Society Magazine, 42(3), 69-75.

[27] Mohammed Dabboor,Junye Xu,Maria Vakalopoulou,Stéphane Bélair,Jarrett Powers,Marco Carrera & Leqiang Sun. (2024). The RADARSAT Constellation Mission for Soil Moisture Retrieval of Bare Soil by Compact Polarimetry and Random Forest Regression. Canadian Journal of Remote Sensing(1),

[28] Wang Zhengqing,Yu Yan,Chen Yulin,Liu Zhongnan,He Haiyang,Guo Zhixin... & Pan Junhua. (2025).

Geohazard Sensitivity Evaluation in Xinning, Hunan, China, Using Random Forest, Artificial Neural Network, and Logistic Regression Algorithms. Natural Hazards Review(2),

[29] Rayees Ali,Haroon Sajjad,Tamal Kanti Saha,Md Hibjur Rahaman,Md Masroor,Roshani & Aastha Sharma. (2024). Assessment of climate change in Upper Jhelum Sub-catchment, India, using nonparametric methods and random forest moBl. Acta Geophysica(prepublish),1-20.