Automatic Classification of Road Safety Level Using Model Transfer and Data Augmentation

Zhihua Hu, Xiaoming Zhu

***Abstract*—In this paper, we study the application of deep learning in highway road safety prediction. The road safety model is implemented by LSTM and transfer learning. First, using the existing highway road data with known safety levels the training data set is established for supervised learning. We further proposed a novel anisotropic cost function in order to reduce the risks in misclassification. Second, GAN is used to expand the data set in order to provide enough data to train deep model and improve the accuracy of the safety level classification. Third, in order to adapt to different road situations, transfer learning framework is used to share the common knowledge in road safety prediction, which improves the generalization ability of the proposed classification method. Finally, the experimental results show that the proposed method outperformed the conventional method and has a good practical value to improve traffic safety.**

***Index Terms*—road safety level, GAN, transfer learning.**

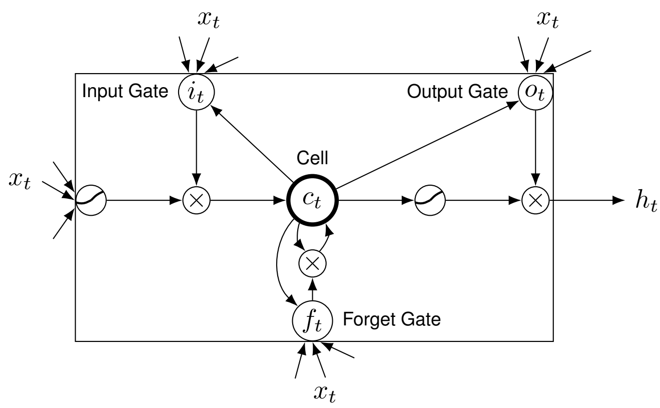

Fig. 1. Structure of LSTM network

I. Introduction

With the increase of highway and urban road construction scale, the problem of road traffic safety has become increasingly important. The systematic improvement of road safety level is of great significance. Conventionally, the road safety assessment depends on the experience of human experts. Recently, the automatic road safety level classification using machine learning algorithms have shown promising results[1]. Machine learning technologies can help to reduce the number of road traffic accidents, reduce economic losses, and improve the passenger transportation safety.

Automatic road safety level classification has drawn wide attentions. In Europe, in order to reduce the traffic accidents, UK has made specific regulations on the scope, steps and processes of road risk assessment. In New Zealand, the safety assessment has been established for each stage of road planning, design, construction, operation and maintenance. The automatic road safety evaluation system can assess the risks of the road through on-site inspection according to the characteristics of the road, and can regularly track the safety status[2], [3].

At present, road risk data is still difficult to achieve and traffic accident data can not be published publicly. Databases are indispensable to many machine learning studies[4], [5], and small data size is often challenging for data modeling[6]. Therefore studying the methods of database enhancement and simulation is of great importance to automatic road safety classification[7].

In order to better solve the road safety problem in highway, we use GAN to enhance the dataset. Transfer learning is used to train a novel highway risk classification model which is based on the similarity of feature structure from the existing highway data. Deep neural network is adopted to bring enhanced performance with a novel anisotropic cost function that we proposed for the road safety classification. The rest of the paper is organized as follows. In Sec.2, the algorithms used in our safety level classification is described. In Sec.3, the experimental results are given and the classification results are analysed and compared. Finally, the conclusions are given in Sec.4.

Manuscript received May 7, 2020; revised October 19, 2020. This work was supported by the Project of Scientific and Technological Innovation Team for Outstanding Young and Middle-aged Students in Hubei of China (Grants No. T201924).

Zhihua Hu and Xiaoming Zhu are with School of Computer Science, Huanggang Normal University, Huanggang, Hubei, China e-mail:huzhihua123@126.com, zhuxm_74@yahoo.com.

II. Algorithm description

The highway road safety level depends on two types of factors. First factor is that the facilities and road conditions of a certain section of highway. Second factor is that the neighbouring road sections that may influence the road safety. Therefore, this paper regards the road safety level classification as as sequential modeling problem. Various neural network and deep learning methods have been proposed with successful applications [8], [9], [10], [11], [12], [13]. Bustami et al.[12], applied a fundamental neural network to water level prediction. Hashida et al.[13], proposed to use convolutional neural networks to model natural language in social media. On sequential modeling problem, LSTM (Long-short term memory) network is a fundamental approach to model the contextual information[14].

A. LSTM

LSTM is an improvement of recurrent neural network(RNN). The difference between RNN and traditional deep neural network is that the neurons in the hidden layer are connected, and each neuron has two inputs: input of input layer and output of previous sequence. This structure enables RNN to better process and predict sequential data. The structure of LSTM gates is shown in Fig.1.

The problem of gradient disappearance and gradient explosion exists in RNN network, which makes it difficult to deal with long-term dependence in practical application. Long-short-term memory (LSTM) network is a special RNN network, LSTM solves the long-term dependence problem that RNN is difficult to handle, long term dependency can be learned from training data [15], [16], [17].

The circulating neuron of LSTM is also called cell unit, and its internal structure is shown in Fig.1, different from the internal structure of RNN, LSTM adds a "processor" to determine whether the information is useful, the processor contains three gates that control information is retained and forgotten, they are called input gate, forgetting gate and output gate respectively.

(1) Forgetting gate The forgetting gate of LSTM determines which historical information is forgotten from cell state. The output of forgetting gate is shown in Eq.1

$$f_t = \theta(W_f x_t + U_f h_{t-1} + b_f) \tag{1}$$

where, W is the weight parameter matrix between the input layer and the hidden layer, U is the self cycling parameter matrix between two hidden layers, and b is the offset parameter, f_t is the output value of forgetting gate, activation function is sigmoid function.

(2) Input gate The input gate of LSTM determines what kind of new information is updated in the cell state. The implementation consists of two parts. The information to be updated is determined by the sigmoid layer.

$$i_t = \theta(W_i x_i + U_i h_{t-1} + b_i) \tag{2}$$

The other part is the candidate updates generated by tanh layer.

$$\tilde{C}_t = tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{3}$$

As shown in Fig.1, the forgetting gate and the input jointly control the update of cell state C_T. When the forgetting gate f_t comes to space, it forgets how much historical information C_T is saved by the input gate.

The input sequence is composed of index parameters of highway, including speed limit, slope, curvature of curve, height of isolation belt, number of lanes, width of road, whether or not to fork road, etc. according to the coordinates of each road section, the surrounding road condition index parameters are determined to form the observation feature sequence. The output sequence is the safety level of the highway, which is divided into five levels corresponding to discrete values.

(3) Output gate The output gate o_t of the LSTM determines which part of the cell state C_T information to output. The cell state C_T is processed through the tanh layer before passing through the output gate.

$$O_t = \theta(W_o x_t + U_o h_{t-1} + b_o) \tag{4}$$

$$h_t = o_t tanh(C_t) \tag{5}$$

TABLE I
EVALUATION OF MISCLASSIFICATION CONSEQUENCE (COST) BY LINGUISTIC VALUES.

Risk Levels	Predict 1	Predict 2	Predict 3	Predict 4	Predict 5
Level 1	VL	L	M	H	VH
Level 2	H	VL	L	M	H
Level 3	VH	H	VL	L	M
Level 4	VH	VH	H	VL	L
Level 5	VH	VH	VH	H	VL

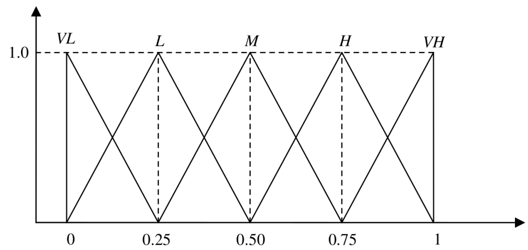

Fig. 2. Triangular membership functions for linguistic values.

B. Anisotropic cost function for road safety

The classification of road safety (risk) level is considered as a typical classification problem. The metrics of risk level classification are defined upon the measurement of an error:

$$Error = Prediction(obs_{road}) - Label(obs_{road}) \tag{6}$$

The observation obs_{road} of a certain location of a road, consists of various factors, such as speed limit, curvature, etc. The error can be defined as the difference between the label (known ground-truth) and the prediction of a classifier.

However the above definition in Eq.6 has not take the nature of traffic accidents into account. A misclassification of high risk road to low risk one can be very dangerous. On the other hand, prediction of higher level of risk (e.g. misclassify risk level 3 to risk level 4[1]) is in fact less dangerous.

In this paper, we propose a new anisotropic cost function (ACF) for our road risk level classification model.

$$Error_{aniso} = F\{Prediction(obs_{road}) - Label(obs_{road})\} \tag{7}$$

The function F is defined as a mapping from conventional $Error$ to anisotropic $Error_{aniso}$. When $Error$ is positive, the mapping decreases the cost. When $Error$ is negative, the mapping increases the cost in order to prevent high risk level being misclassified into low risk level. When $Error$ is zero, the output remains zero.

The degree of increasement or decreasement in our anisotropic cost function is based on fuzzy set theory to process the knowledges from road safety experts. The consequences of misclassifications are evaluated by human experts, by linguistic values, including very high (VH), high(H), medium(M), low(L), very low(VL) and zero(Z). An example of such evaluation is shown in Tab.I

Taking the triangular membership function, we can map the linguistic values to fuzzy members, as shown in Fig.2.

[1]the higher the number of the risk level is, the higher the road risk is, e.g. risk level 5 is the highest risk road, and risk level 1 is the lowest risk road.

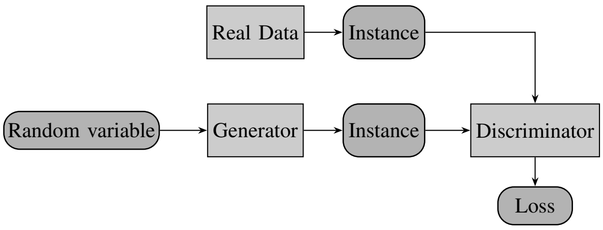

Fig. 3. GAN algorithm flow chart

The mean values of membership function are taken as the costs, $\{0, 0.25, 0.5, 0.75, 1\}$, for VL, L, M, H, VH.

C. Data augmentation on small sample

The number of road condition samples related to safety level is limited. Due to the confidentiality of traffic accidents data, the road information and accident information disclosed in practice are not enough to train a deep network. In machine learning applications, database is fundamental for solving learning and modeling problems[18], [19], [20], [21]. At present, a lot of work is devoted to solve the problem of small samples, that is, using limited and rare data to train an efficient model. In this paper, we use Generative adversarial networks (GAN) to generate additional training data.

The generative adversarial network (GAN) is a kind of deep learning model [22]. GAN consists of the generative network G (generator) and the discriminant network D (discriminator). The structure of the countermeasure network is shown in Fig.3. In the iterative training process, the goal of generative network G is to cheat network D, while the goal of discriminant network D is to distinguish the authenticity of the input data. Through training, G and D finally reach the Nash equilibrium point. At this time, the trained generative network G can transform a random vector to a certain distribution into a sample similar to the training set.

In the current highway data analysis, in order to train the safety level model of a specific road section, it is necessary to collect the highway index data of the road section, such as speed limit, slope, etc. In the actual road safety data analysis, we often do not have enough training samples to complete the complex and large-scale deep network training. Therefore, we adopt the CycleGAN[23] for sample enhancement. By using this data enhancement method, we can generate a large number of reliable and high-quality training data to improve the accuracy of the model.

The principle of CycleGAN can be summarized as follows: there are two sample spaces X and Y, which map the samples in X space to the samples in Y space. The goal of CycleGAN is to learn the mapping from X to Y. CycleGAN structure is shown in Fig.4. In the road safety data, X represents the sample of safe road, Y represents the sample of risky road section. Through CycleGAN, safe samples and risky samples can be transformed into each other, so that a large number of safe samples can be used to generate specific risky samples to make up for the deficiency in data quantity.

In Fig.4, the mapping from X to Y is F, which can convert the data x in X space to the data $F(x)$ in Y space; The mapping from Y to X is G, which can convert the data y in Y space to the data $G(y)$ in X space. CycleGAN learns the mapping of F and G at the same time, and requires

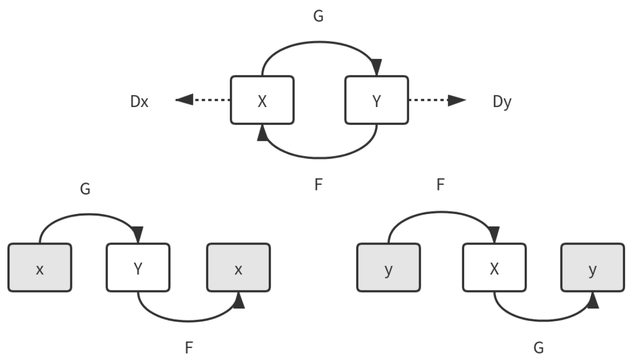

Fig. 4. Depiction of CycleGAN structure

$G(F(x)) \approx x$, and $F(G(y)) \approx y$, that is, after the data in X space is converted to the data in Y space, it can also be converted back. According to $G(F(x)) \approx x$ and $F(G(y)) \approx y$, cycle consistency loss can be defined. We can define CycleGAN loss according to Eq.8.

$$\begin{aligned} L_{CycleGAN}(G, D_x, X, Y) = & L_{GAN}(F, D_Y, X, Y) + \\ & L_{GAN}(G, D_X, X, Y) \\ & + \lambda L_{cycle-consis}(F, G, X, Y) \end{aligned} \tag{8}$$

It should be pointed out that the number of data generated by CycleGan is limited. Because the input and output of the generation model trained by CycleGAN is a 1:1 mapping relationship, the number of input data of CycleGAN keeps a 1:1 relationship with the number of its generation. CycleGAN needs more human intervention to generate road data, such as adjusting models, manually filtering data, etc., in order to obtain satisfactory data sets.

D. Transfer Learning

The goal of transfer learning(TL) is to use the knowledge learned from one domain to help solving tasks in a new domain[24], [25], [26]. In our road safety analysis, TL helps to transfer the established model from old road database to new roads. Conventionally researchers adopt various optimization methods to adapt the feautures and model parameters[27], [28], [29], [30].

There are several advantages to apply transfer learning. First, it can reuse the existing models. Second, it doesn't require collecting and labelling a huge amount of data. Third, it can quickly apply the learned knowledge to a new field and save time and costs.

Transfer learning can be applied to supervised learning, unsupervised learning or reinforcement learning[31], [32]. We can achieve the shared representation through multi-task learning. Through the joint learning of multiple source domain tasks, we can extract the shared knowledge suitable for other similar tasks. The general knowledge extracted in the process of multi-task learning in various fields can not be achieved by single task learning. The generalization ability is also improved.

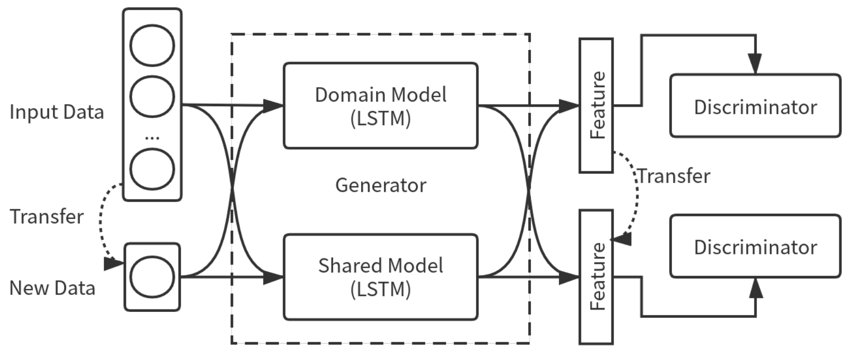

Fig. 5. Framework of shared knowledge learning

Current transfer learning approaches can be grouped into four categories: feature-based transfer learning, model-based transfer learning, instance-based transfer learning and relation-based transfer learning. Among them, feature-based transfer learning and model-based transfer learning can be combined with deep learning model.

Feature-based Transfer Learning refers to the transfer in feature space, that is, learning a good feature representation through the source domain, and then transferring knowledge from the source domain to the target domain through feature coding, so as to improve the performance in the target domain. The common method is to observe the common features between the source domain and the target domain, then project the features of the source domain and the target domain into the same feature space, and finally transfer the common features automatically.

Model-based Transfer Learning , also known as parameter transfer learning, refers to the application of the whole model in the source domain to the target domain. Model-based transfer can be used to apply the original model to a new domain. In the new domain, only a small amount of annotated data is needed to fine tune the parameters, which can achieve improved results.

Instance-based Transfer Learning is simply to evaluate the samples and assign a larger weight to the more important samples. First, find the data similar to the target domain in the data set of the source domain, and then match the data of the target domain.

Relation-based Transfer Learning refers to the knowledge transfer between related fields, which assumes that the data in the source and target fields have the same relationship.

In the classification road safety level, it is often necessary to consider the differences of geographical factors. In different provinces and different geographical environments, there may be great differences in highway safety characteristics. For example, in mountainous areas, the slope is large, and in urban areas, traffic congestion and lane change bring risks. Therefore, in the actual deployment of safety classification model, we need to improve the generalization ability of the model.

The method proposed in this paper falls into the model-transfer category, which improves the road safety prediction. First, using the road section with known risks, a specific domain model can be trained. Second, using a shared model, it can be transferred to the target road with unknown factors.

E. The shared knowledge learning for traffic data

Based on the method of model transfer, we adopt the shared knowledge learning proposed in [33] to train the model $S(x_i;\theta_s)$ as shared knowledge and transfer to new domain. The output of the new domain data in the shared model is z_{new}, as shown in Eq.9.

$$z_{new} = S(x_{new};\theta_s) \tag{9}$$

where x_{new} represents the input data of the new domain, and θ_s represents the parameters of the shared model. As shown in Fig.5, the original domain model is trained on a large collection of traffic data with labels indicating the road safety levels. LSTM is used as generator and the multi-layer neural network as discriminator. The output of the shared knowledge learning model is used as a part of the feature vector combined with domain-specific features. By learning the common knowledge for traffic data, the transferred model is used for the new input data. With a small amount of fine tuning data the model can be applied to new domains.

$$h(z_{new}) = \theta(w_s z_{new} + b) \tag{10}$$

where z_{new} represents the output vector of the new domain data in the shared model, and w_s represents the shared parameter. The cross entropy is used as loss function:

$$L_i = -log(\frac{e^{y_{(i)new}}}{\sum_j e^{h(z_{(j)new})}}) \tag{11}$$

where y_{new} indicates the risk level label of the new domain data x_{new}.

III. Experimental Results

A. Experiment setup

In this experiment, three-layer LSTM network structure is adopted, five neuron nodes are set in the output layer, and the probability distribution of output samples in five risk levels is achieved through softmax function.In order to improve the generalization ability of the network, the dropout layer is acquired in each layer of LSTM network layer. The parameter is 0.3. The neural network initialization adopts the gloot uniform distribution initialization method, the loss function adopts the cross entropy loss function, and the training adopts the Adam optimization algorithm, $learning\_rate = 0.01$, $batch\_size = 128$, $epochs = 2000$.

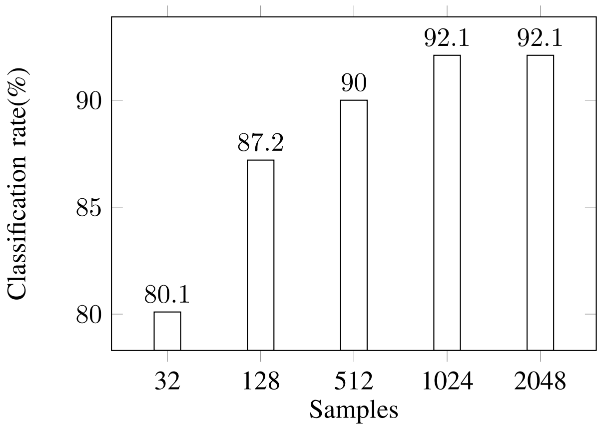

Fig. 6. Search for optimal data set expansion

The experimental data set adopts the road data set collected locally, including 10 roads, 1200 km road section length, 21 public reported traffic accidents, and 12 road condition parameters. After the data enhancement of the generated countermeasure network, the above seed data is expanded to a larger training and verification data set, including comparison of data volume expansion for different scales.

B. Data expansion and comparison experiment

In this experiment, CycleGAN is used to expand the data set. Here, test set and training set are expanded in different scales to get the effect of expansion scale on accuracy and find the most suitable expansion scale.In this experiment, the neural network model of pre-training weight is used for transfer learning. Figure 6 shows the effect of different expansion scales on the experimental results. It can be seen from the figure that the accuracy of the data expansion scale is significantly improved when sample size grow from 32 times to 1024 times. When the data is expanded to 2048 times, its accuracy rate is stable (around 92.1%).

C. Results and analysis

In this experiment, accuracy is used as the evaluation standard of domain discriminator. The accuracy rate represents the ratio between the number of correct samples in the prediction results and the number of all test samples. The larger the ratio is, the higher the accuracy rate is.

Table II is the classification accuracy results of our algorithm and Gaussian mixture models(GMM)[34], [35], standard LSTM[14], [36] without transfer learning, SVM[37], XGBoost[38], AlexNet[39], AutoEncoder[40] and DBN[41]. In GMM training, the mixture number is adjusted and set to 16. The initialization algorithm for EM parameter training is K-means with K=16. In SVM experiment, the kernel is tested among “RBF”, “Linear”, “Poly” and “Sigmod”, and the first kernel RBF(Gaussian Kernel) has the best result on road safety data. For XGBoost algorithm, we set tree depth to 5, and learning rate to 0.02. AlexNet is a typical example of convolutional neural netowrk, we use “ReLU” as its activation function. It has five convolutional layers and three fully connected layers. AutoEncoder is used to encode the input road parameters and the load safety level, as described in reference[40], the predicted safety level can be obtained by minimizing a fitness function of the network. DBN is used as a simple example of deep neural network. Restricted Boltzmann Machine (RBM) is used for its initialization.

As can be seen from the table, the accuracy rate of the proposed method is higher than that of other conventional methods. We also plotted the confusion matrix in Fig.7 for further comparison and analysis. It can be concluded from the results in the table that the proposed model in this paper is superior to the other traditional algorithms.

In conventional neural network structures, the cost function (loss function) is defined based on the error between the predictions and the labels(known ground-truth). The error form is generally isotropic, which means whether the prediction is larger or smaller than the label is identical for error measurement. However, if we consider the actual risk of misclassification in road safety, we can see that anisotropic error form has a clear advantage in practice. As we have discussed in Sec.II-B, we extend our methods to the case of ACF cases and use a more accurate metric to evaluate the experimental results for road safety classification.

Mean Anisotropic Errors(MAEs) with ACF[2] over different algorithms are shown in Fig.8 together with Mean Isotropic Errors(MIEs). Compared with various existing algorithms, our proposed algorithm outperform the rest at a considerable margin.

In order to exam the significance of our proposed model, T-Test is carried out on the classification results by comparing each pair of algorithms. The significance (p-value) of our proposed algorithm is shown in Tab.III. “ACF” stands for anisotropic cost function and “ICF” stands for isotropic cost function. From the results we can see that the proposed model is significantly better than the conventional models in road safety classification.

IV. Conclusion

In this paper, the application of deep neural network in road safety data modeling is studied. CycleGAN is used to enhance the data set, and a large-scale data sample expansion is carried out on a small amount of actual road accident data, so as to meet the training requirements of deep network. Due to the complexity of highway safety problems, the method of transfer learning is adopted to improve the generalization ability on different highway conditions. The experimental results show that the proposed anisotropic cost function is suitable for road safety classification and the proposed model can be adapted to various highway safety data and achieves satisfactory results.

References

[1] T. Jinhua, Research on road risk management, Beijing: Tsinghua University, 2011.

[2] S. P. Miaou, J. J. Song, “Bayesian ranking of sites for engineering safety improvements: Decision parameter, treatability concept, statistical criterion, and spatial dependence,” *Accident Analysis and Prevention*, vol.37, no.4, pp.699-720, 2005.

[3] E. Hermans, F. Van den Bossehe, G. Wets, “Combining road safety information in a performance index,” *Accident Analysis and Prevention*, vol.40, no.4, pp.1337-1344, 2008.

[4] Y. Jin, Y. Zhao, C. Huang and L. Zhao, “The design and establishment of a Chinese whispered speech emotion database,” *Technical Acoustics*, vol.29, no.1, pp.63-68, 2010.

[2]Error is defined in Eq.7

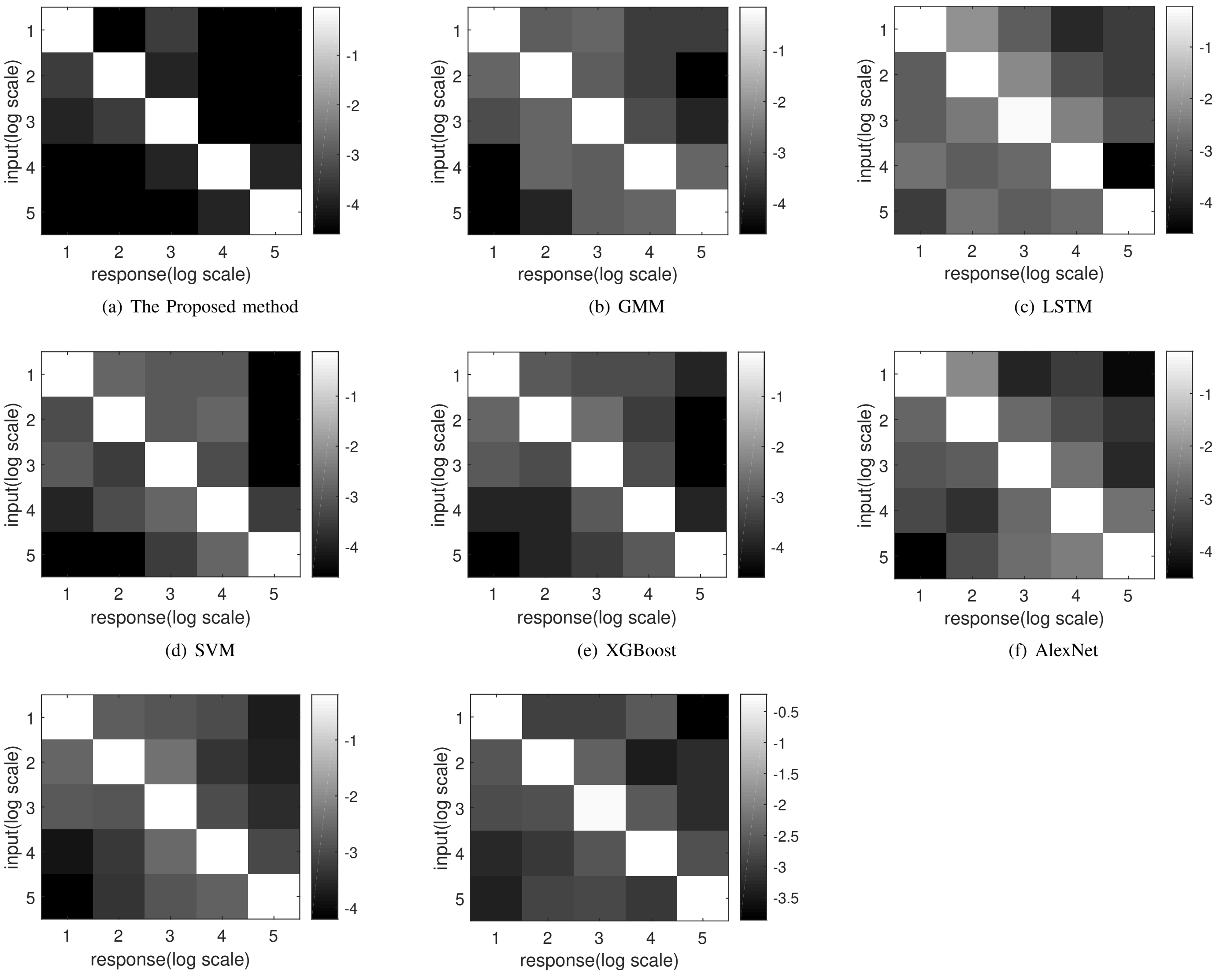

Fig. 7. Results on the confusion matrics using the proposed method and conventional methods

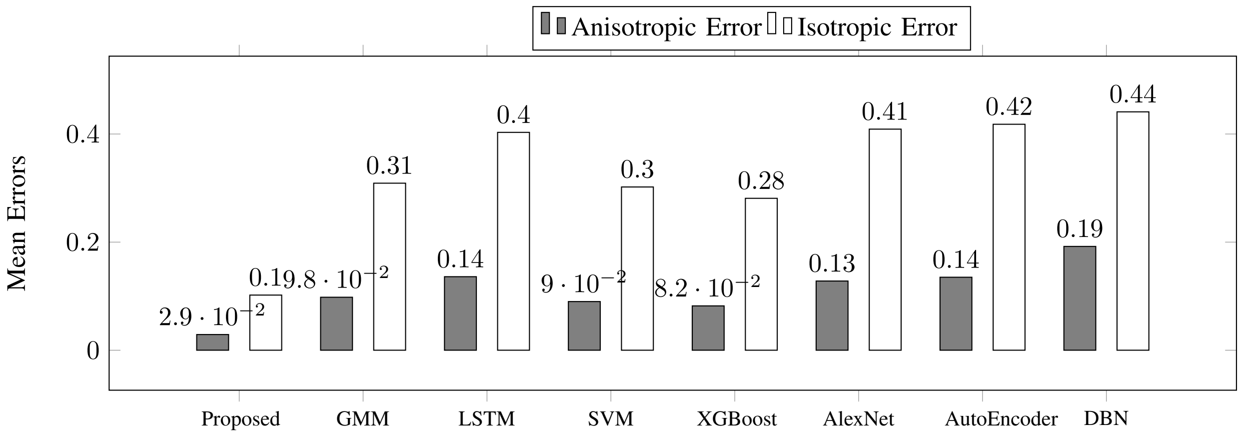

Fig. 8. Comparing mean anisotropic errors and mean isotropic errors over different algorithms.

TABLE II
ACCURACY COMPARISON

Risk level	Proposed model	GMM	LSTM	SVM	XGBoost	AlexNet	AutoEncoder	DBN
1	0.953	0.831	0.778	0.839	0.851	0.823	0.802	0.796
2	0.945	0.849	0.771	0.835	0.833	0.801	0.778	0.773
3	0.944	0.843	0.744	0.871	0.856	0.803	0.798	0.753
4	0.937	0.823	0.807	0.854	0.891	0.792	0.812	0.776
5	0.970	0.860	0.792	0.891	0.889	0.786	0.823	0.801

TABLE III
P-VALUES OF LEFT-TAILORED T-TEST WITH 0.05 SIGNIFICANCE LEVEL

Models to Compare	with ACF	with ICF
(Proposed,GMM)	3.02e-205	1.01e-111
(Proposed,LSTM)	8.79e-299	6.34e-288
(Proposed,SVM)	5.61e-110	9.89e-56
(Proposed,XGBoost)	8.83e-38	6.34e-11
(Proposed,AlexNet)	6.90e-155	7.88e-89
(Proposed,AutoEncoder)	1.12e-164	1.76e-155
(Proposed,DBN)	1.56e-201	3.48e-301

[5] C. Huang and H. Jiang, "Image indexing and content analysis in children's picture books using a large-scale database," *Multimedia Tools and Applications*, vol.78, no.15, pp.20679-20695, 2019.

[6] Z. Qiao, L. Zhou, J. Z. Huang, "Sparse Linear Discriminant Analysis with Applications to High Dimensional Low Sample Size Data," *International Journal of Applied Mathematics*, vol.39, no.1, pp.48-60, 2009.

[7] A. Doniec, R. Mandiau, S. Piechowiak and S. Espie, "A behavioral multi-agent model for road traffic simulation," *Engineering Applications of Artificial Intelligence*, vol.21, no.8, pp.1443-1454, 2008.

[8] X. Wei, H. Xia and L. Zhou, "Broad Forest: A Non-Neural Network Style Broad Model for Streaming Video QoE Evaluation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.30, no.4, pp.632-645, 2020.

[9] J. Liu, Q. Zhang and Z. Luo, "Dynamical analysis of fuzzy cellular neural networks with periodic coefficients and time-varying delays," *IAENG International Journal of Applied Mathematics*, vol.46, no.3, pp.298-304, 2016.

[10] H. Zhang, M. Du and W. Bu, "Sliding mode controller with RBF neural network for manipulator trajectory tracking," *IAENG International Journal of Applied Mathematics*, vol.45, no.4, pp.334-342, 2015.

[11] P. Wanda and H. J. Jie, "DeepSentiment: Finding Malicious Sentiment in Online Social Network based on Dynamic Deep Learning," *IAENG International Journal of Computer Science*, vol.46, no.4, pp.616-627, 2019.

[12] R. Bustami, N. Bessaih, C. Bong, S. Suhaili, "Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River," *IAENG International Journal of computer science*, vol.34, no.2, pp.228-233, 2007.

[13] S. Hashida, K. Tamura and T. Sakai T, "Classifying Tweets using Convolutional Neural Networks with Multi-Channel Distributed Representation," *IAENG International Journal of Computer Science*, vol.46, no.1, pp.68-75, 2019.

[14] L. Zaman, S. Sumpeno, M. Hariadi, Y. Kristian, E. Setyati and K. Kondo, "Modeling Basic Movements of Indonesian Traditional Dance Using Generative Long Short-Term Memory Network," *IAENG International Journal of Computer Science*, vol.47, no.2, pp.262-270, 2020.

[15] K. Greff, R. K. Srivastava, J. Koutník, et al., "LSTM: A Search Space Odyssey," *IEEE Transactions on Neural Networks & Learning Systems*, vol.28, no.10, pp.2222-2232, 2015.

[16] F. A. Gers, N. N. Schraudolph and J. Schmidhuber, "Learning Precise Timing with LSTM Recurrent Networks," *Journal of Machine Learning Research*, vol.3, no.1, pp.115-143, 2003.

[17] J. Wang, Y. Yang, J. Mao,et al., "CNN-RNN: A Unified Framework for Multi-label Image Classification," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.2285-2294, 2016.

[18] L. Zhao, C. Huang, "Key technologies in practical speech emotion recognition," *Journal of Data Acquisition and Processing*, vol.29, no.2, pp.157-170, 2014.

[19] C. Huang, Y. Jin, Y. Zhao, Y. Yu and L. Zhao, "Recognition of practical emotion from elicited speech," *In Proc. First International Conference on Information Science and Engineering*, pp.639-642, 2009.

[20] C. Huang and H. Jiang, "Image indexing and content analysis in children's picture books using a large-scale database,", *Multimedia Tools and Applications*, vol.78, no.15, pp.20679-20695, 2019.

[21] Y. Jin, Y. Zhao, C. Huang, et al., "The design and establishment of a Chinese whispered speech emotion database," *Technical Acoustics*, vol.29, no.1, pp.63-68, 2010.

[22] I. Goodfellow, "NIPS 2016 Tutorial: Generative Adversarial Networks.", *Conference and Workshop on Neural Information Processing Systems*, pp.1-10, 2016.

[23] J. Y. Zhu, T. Park, P. Isola, et al., "Unpaired image-to-image translation using cycle-consistent adversarial networks," *2017 IEEE International Conference on Computer Vision (ICCV)*, pp.2242-2251, 2017.

[24] I. Guyon, G. Dror, V. Lemaire, et al., "Unsupervised and transfer learning challenge," *International Joint Conference on Neural Networks (IJCNN)*, pp.793-800, 2011.

[25] J. Yosinski, J. Clune, Y. Bengio Y, et al., "How transferable are features in deep neural networks?" *International Conference on Neural Information Processing Systems*, pp.3320-3328, 2014.

[26] S. J. Pan, Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol.22, no.10, pp.1345-1359, 2010.

[27] X. Zhang, C. Huang, L. Zhao and C. Zou, "Recognition of practical speech emotion using improved shuffled frog leaping algorithm," *Chinese Journal of Acoustics*, vol.33, no.4, pp.441-441, 2014.

[28] C. Huang, Y. Jin, Y. Zhao, Y. Yu and L. Zhao, "Speech emotion recognition based on re-composition of two-class classifiers," *In Proc. 3rd International Conference on Affective Computing and Intelligent (ACII)*, pp.1-3, 2009.

[29] C. Zou, C. Huang, D. Han, et al. "Detecting practical speech emotion in a cognitive task," *In Proc. of IEEE 2011 20th International Conference on Computer Communications and Networks*, Lahaina, HI, USA, pp.1-5, 2011.

[30] C. Wu, C. Huang and H. Chen, "Text-independent speech emotion recognition using frequency adaptive features," *Multimedia Tools and Applications*, vol.77, no.18, pp.24353-24363, 2018.

[31] R. Liang,Z. Liang, J. Cheng, Y. Xie and Q. Wang, "Transfer learning algorithm for enhancing the unlabeled speech", *IEEE Access*, no.8, pp.13833-13844, 2020.

[32] X. Wei, J. Zhao, L. Zhou and Y. Qian, "Broad Reinforcement Learning for Supporting Fast Autonomous IoT," *IEEE Internet of Things Journal*, vol.7, no.8, pp.7010-7020, 2020.

[33] X. Duan, The research and application of sentiment classification based on transfer learning, Master's Thesis, School of Computer Science, Beijing University of Posts and Telecommunications, June 4, 2019.

[34] C. Huang, B. Song and L. Zhao, "Emotional speech feature normalization and recognition based on speaker-sensitive feature clustering," *International Journal of Speech Technology*, vol.19, no.4, pp.805-816, 2016.

[35] Chengwei Huang, et al., "Cascaded projection of Gaussian mixture model for emotion recognition in speech and ECG signals," *Journal of Southeast University (English Edition)*, no.03, pp.320-326, 2015.

[36] Yue Xie, Ruiyu Liang, Zhenlin Liang, Chengwei Huang, Cairong Zou, Björn Schuller, "Speech emotion classification using attention-based LSTM," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vo.27, no.11, pp.1675-1685, 2019.

[37] J. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol.9, no.3, pp.293-300, 1999.

[38] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp.785-794, 2016.

[39] S. Lu, et al., "Pathological brain detection based on AlexNet and transfer learning," *Journal of computational science*, no.30, pp.41-47, 2019.

[40] B. Betechuoh, T. Marwala and T. Tettey, "Autoencoder networks for HIV classification," *Current Science*, pp.1467-1473, 2006.

[41] G. Hinton, S. Osindero and Y. Teh, "A fast learning algorithm for deep belief nets," *Neural computation*, vol.18, no.7, pp.1527-1554, 2006.

Zhihua Hu received his Masters Degree and Ph.D. from Huazhong University of Science & Technology and Wuhan University. He is currently a professor in Huanggang Normal University. His research interests include information security and data mining.

Dr. Hu is a member of China Computer Federation and a member of Operations Research Society of China.

Xiaoming Zhu received his Bachelor's Degree and Ph.D. from Tsinghua University and The University of California, Berkeley, respectively. He is currently a professor in Huanggang Normal University. His research interests include wireless communication, data mining and artificial intelligence.