

Identification of Similar Air Traffic Scenes with Active Metric Learning

CHEN Haiyan^{1,3*}, HOU Xiaye¹, YUAN Ligang², ZHANG Bing¹

1. College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P.R. China;

2. College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P.R. China;

3. Collaborative Innovation Center of Novel Software Technology and Industrialization, Nanjing 210023, P.R. China

(Received 10 June 2021; revised 30 July 2021; accepted 10 August 2021)

Abstract: The rapid growth of air traffic has continuously increased the workload of controllers, which has become an important factor restricting sector capacity. If similar traffic scenes can be identified, the historical decision-making experience may be used to help controllers decide control strategies quickly. Considering that there are many traffic scenes and it is hard to label them all, in this paper, we propose an active SVM metric learning (ASVM²L) algorithm to measure and identify the similar traffic scenes. First of all, we obtain some traffic scene samples correctly labeled by experienced air traffic controllers. We design an active sampling strategy based on voting difference to choose the most valuable unlabeled samples and label them. Then the metric matrix of all the labeled samples is learned and used to complete the classification of traffic scenes. We verify the effectiveness of ASVM²L on standard data sets, and then use it to measure and classify the traffic scenes on the historical air traffic data set of the Central South Sector of China. The experimental results show that, compared with other existing methods, the proposed method can use the information of traffic scene samples more thoroughly and achieve better classification performance under limited labeled samples.

Key words: air traffic; similar scene; active learning; metric learning; SVM

CLC number: TP391 **Document code:** A **Article ID:** 1005-1120(2021)04-0625-09

0 Introduction

Air traffic system is highly dynamic and complex due to the influence of weather and traffic flow. Air traffic controllers often need to judge emergencies and release control strategies quickly to avoid conflicts, congestion, delays and other situations. With the continuous growth of airspace traffic, controllers often fall into overloaded work conditions, which brings potential dangers to air traffic. Senior controllers can usually rely on their experience to make quick decision under historically similar weather and traffic flow conditions. For novice controllers, they cannot easily make control strategies due to the lack of experience in handling similar scenes.

Therefore, the automatic identification of similar air traffic scenes and the automatic extraction of corresponding traffic control strategies can assist the controllers to make quick decisions and reduce their workload. In 2013, Grabbe et al.^[1] proposed a similar scene-based strategy-assisted decision-making idea for air traffic control. It first finds out the historical days that are most similar to the reference day by measuring a set of given air traffic characteristics, and then analyzes the actual control strategies in the historical days (such as ground delay waiting procedures, TMIs, etc.) and performance evaluation indicators after implementation (such as average delay time, number of cancelled flights, etc.). Based on the analysis, recommendations of control strate-

*Corresponding author, E-mail address: chenhaiyan@nuaa.edu.cn.

How to cite this article: CHEN Haiyan, HOU Xiaye, YUAN Ligang, et al. Identification of similar air traffic scenes with active metric learning[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2021, 38(4):625-633.

http://dx.doi.org/10.16356/j.1005-1120.2021.04.009

gies on the reference day are given. Grabbe's idea is widely accepted by scholars, and a lot of research work has been carried out, mainly focusing on the feature selection, traffic scenes extraction and clustering of similar scenes^[2-3]. These methods eliminate the influence of correlation and redundancy between features by feature selection or extraction and then cluster the traffic scenes by measuring their distance. So that some different types of similar traffic scenes can be obtained. However, because the traffic scene data do not have labels and cluster algorithms cannot give the accurate number of categories, the clustering results only give a rough category division that is also hard to be explained. This obviously cannot bring direct help to the controllers in making current decision.

In order to classify and identify similar traffic scenes more accurately, we adopt the idea of semi-supervised learning to propose a method for identifying similar traffic scenes in sectors based on active SVM metric learning (ASVM²L). In this method, the senior controllers (Oracle) first mark some traffic scene samples. For a large number of unlabeled traffic scene samples, an active learning strategy is used to pick up the most difficult samples for metric learner and submit them to Oracle to get labels. Then the labeled samples are used to train a SVM metric learner. The process will be repeated until the best metric matrix is obtained, thereby obtaining the best classification of similar traffic scenes. At last, we verify the effectiveness of the proposed ASVM²L algorithm on standard data sets, and the performance of the proposed similar traffic scene identification method based on ASVM²L on the real traffic scene data set.

1 Related Work

To identify similar traffic scenes, a lot of researches have been conducted in feature selection and extraction, similarity measurement and clustering. Kuhn et al.^[4] selected some features based on expert knowledge, such as the distance from the airport to the center of heavy rainfall, the strength of

the crosswind on the runway, the number of scheduled flights, etc., to calculate the similarity between the traffic scenes. Christien et al.^[5] used features such as the number and type of aircraft, the number of aircraft conflicts, and static sector indicators as input for traffic scene classification. In 2016, Grabbe et al.^[6] used the Yakkad coefficient to calculate the similarity of traffic scenes, and then applied it as the input of similar scene clustering and classification. Hu et al.^[7] extracted discrete features and time series features of traffic scenes by reducing the 382-D traffic features, and used these features to cluster similar traffic scenes.

It can be seen that the current related researches mostly use clustering to obtain the categories similar traffic scenes. By calculating the Euclidean distance between the samples, the similar scenes are grouped into a cluster and regarded as the same category. Through these studies, we can initially obtain several categories of traffic scenes consistent with the controller's cognition. This provides an opportunity for semi-supervised metric learning algorithms to be applied to measure the similarity of traffic scenes. Metric learning is an important research area in machine learning. Finding appropriate metrics between samples by metric learning can help subsequent learning algorithms achieve better performance. As the most common metric learning method, the Mahalanobis distance metric uses the original label information to generate side information, which is constructed into pairwise constraints or triple constraints acting on the loss function to implement metric learning. Xing et al.^[8] firstly proposed a convex optimization method using pairwise constraints to learn Mahalanobis distance metric, and then proposed the large margin nearest neighbor algorithm (LMNN)^[9] based on the K -nearest neighbor (K -NN) idea. The training goal of LMNN is to keep the K nearest neighbor samples of each sample in the same class, and the samples of different categories far apart. Neighborhood component analysis^[10] (NCA) obtains the transformation matrix in the Mahalanobis distance by randomly selecting the

cross-validation results of the nearest neighbor optimization leave-one-out method on the training set. Unlike other methods, the NCA method is non-parametric and does not make any assumptions about the shape and boundary of the class distribution. It is both a metric learning method and a linear dimensionality reduction method. Information theory metric learning (ITML)^[11] obtains the metric matrix by minimizing the differential relative entropy under the pairwise constraint of inequality. The maximally collapsing metric learning^[12] (MCML) constructs a convex optimization problem, which generates a metric matrix by folding samples of the same class to a point and pushing away samples of other classes infinitely. Wang et al.^[13] proposed a kernel classification framework for learning the Mahalanobis distance metrics in the original feature space by combining with SVM. SVM metric learning (SVM²L) supports both pairwise constraints and triple constraints. As a kernel metric learning algorithm, SVM²L is not only more efficient than other metric learning algorithms, but also more suitable for applications with a small number of useful samples.

In some research work, metric learning algorithms have been combined with active learning to solve semi-supervised metric problems. The key idea of active learning is to select the sample with the maximum information from the unlabeled samples and submit it to Oracle for labeling and then use them for algorithm training. It can achieve better accuracy with few labeled samples. Pasolli et al.^[14] proposed the active metric learning algorithm for remote sensing spectral image classification. By combining LMNN metric learning and 1-NN clustering, unlabeled samples were evaluated, and the most uncertain samples were selected, labeled and added to training set. Repeat the training process until a better metric result is obtained. Priyadarshini et al.^[15] proposed the batch decorrelation for active metric learning for object recognition. This algorithm uses the metric matrix calculated by deep neural network to evaluate the set of constraints and select a batch

of triplets with large amount of information to train better model.

Inspired by these works, we propose a metric learning algorithm that combines the idea of active learning with SVM²L to solve the problem of air traffic scenes.

2 Identifying Similar Traffic Scenes

2.1 Related definitions

In order to facilitate the introduction of the algorithm proposed in this article, the following four definitions are given.

Definition 1 The definition of labeled sample set is

$$D = \{(x_i, y_i) | x_i \in \mathbf{R}^d, y_i = 1, 2, 3, \dots\}$$

where $x_i \in \mathbf{R}^d$ indicates that the feature number of the i -th sample is d and y_i is the label of the i th sample.

Definition 2 The definition of triple constraint set is

$$TC = \{(x_1, x_2, x_3)^n | y_1^l = y_2^l \neq y_3^l, 1 \leq l \leq n\}$$

where n is number of triple constraints. For each triple constraint $(x_1, x_2, x_3)^l$, ensuring that sample x_2^l has the same label with sample x_1^l , and sample x_3^l is different.

Definition 3 Committee set is

$$V = \{V_i | i = 1, 2, \dots\}$$

where V represents the committee set and V_i is the i th committee member which can be machine learning algorithm, evaluation indicator, etc.

Definition 4 Given a positive semidefinite matrix M , the definition of distance between two samples is

$$d_M(x_i, x_j) = (x_i - x_j)^T M (x_i - x_j)$$

2.2 Active SVM²L

2.2.1 Metric learning based on SVM

SVM²L uses triple constraints T generated by the train set to learn a metric matrix M , which can satisfy most of the triple constraints in the set T . The problem can be expressed as the following optimization problem

$$\begin{aligned} & \min_{\mathbf{M}, \boldsymbol{\xi}} \frac{1}{2} \|\mathbf{M}\|_F^2 + C \sum_l \xi_l \\ \text{s.t. } & (\mathbf{x}_1^l - \mathbf{x}_3^l)^T \mathbf{M} (\mathbf{x}_1^l - \mathbf{x}_3^l) - (\mathbf{x}_1^l - \\ & \mathbf{x}_2^l)^T \mathbf{M} (\mathbf{x}_1^l - \mathbf{x}_2^l) \geq 1 - \xi_l \\ & \xi_l \geq 0 \quad \forall l \end{aligned} \quad (1)$$

Eq.(1) can be rewritten by Lagrange remainder as

$$\begin{aligned} L(\mathbf{M}, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = & \frac{1}{2} \|\mathbf{M}\|_F^2 + C \sum_l \xi_l - \\ & \sum_l \alpha_l [(\mathbf{x}_1^l - \mathbf{x}_3^l)^T \mathbf{M} (\mathbf{x}_1^l - \mathbf{x}_3^l) - (\mathbf{x}_1^l - \\ & \mathbf{x}_2^l)^T \mathbf{M} (\mathbf{x}_1^l - \mathbf{x}_2^l)] + \sum_l \alpha_l - \sum_l \alpha_l \xi_l - \sum_l \beta_l \end{aligned} \quad (2)$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are Lagrange multipliers and satisfy $\forall l, \alpha_l \geq 0, \beta_l \geq 0$. In order to find the minimum, let the partial derivatives of \mathbf{M} and $\boldsymbol{\xi}$ to be 0, then we can get

$$\begin{aligned} \frac{\partial L(\mathbf{M}, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \mathbf{M}} = 0 \Rightarrow \\ \mathbf{M} - \sum_l \alpha_l [(\mathbf{x}_1^l - \mathbf{x}_3^l)(\mathbf{x}_1^l - \mathbf{x}_3^l)^T - \\ (\mathbf{x}_1^l - \mathbf{x}_2^l)(\mathbf{x}_1^l - \mathbf{x}_2^l)^T] = 0 \end{aligned} \quad (3)$$

$$\begin{aligned} \frac{\partial L(\mathbf{M}, \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \xi_l} = 0 \Rightarrow \\ C - \alpha_l - \beta_l = 0 \Rightarrow 0 \leq \alpha_l \leq C \quad \forall l \end{aligned} \quad (4)$$

Putting Eqs.(3,4) into Eq.(2), we can get the dual problem as

$$\begin{aligned} \max_{\boldsymbol{\alpha}} - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j K_T(t_i, t_j) + \sum_i \alpha_i \\ \text{s.t. } 0 \leq \alpha_i \leq C, \forall l \end{aligned} \quad (5)$$

where t_i is the i -th triple constraint. $K_T(t_i, t_j)$ is the kernel function for triple constraints.

$$\begin{cases} K_T(t_i, t_j) = \text{tr}(T_i, T_j) \\ T_i = (\mathbf{x}_1^i - \mathbf{x}_3^i)(\mathbf{x}_1^i - \mathbf{x}_3^i)^T - (\mathbf{x}_1^i - \mathbf{x}_2^i)(\mathbf{x}_1^i - \mathbf{x}_2^i)^T \\ T_j = (\mathbf{x}_1^j - \mathbf{x}_3^j)(\mathbf{x}_1^j - \mathbf{x}_3^j)^T - (\mathbf{x}_1^j - \mathbf{x}_2^j)(\mathbf{x}_1^j - \mathbf{x}_2^j)^T \end{cases} \quad (6)$$

From Eq.(3), we can get the formula relationship between metric matrix \mathbf{M} and Lagrange multiplier $\boldsymbol{\alpha}$, shown as

$$\begin{aligned} \mathbf{M} = \sum_l \alpha_l [(\mathbf{x}_1^l - \mathbf{x}_3^l)(\mathbf{x}_1^l - \mathbf{x}_3^l)^T - \\ (\mathbf{x}_1^l - \mathbf{x}_2^l)(\mathbf{x}_1^l - \mathbf{x}_2^l)^T] \end{aligned} \quad (7)$$

2.2.2 Query strategy for active learning

Semi-supervised learning has gradually become a hot issue in machine learning. In semi-supervised learning, unlabeled data are large and easy to obtain, but it is difficult to get their labels. The main

idea of active learning^[16] is that if the algorithm is allowed to select samples to train the model, it can achieve better accuracy with fewer labelled samples. Usually, the active learner selects samples from unlabeled samples and submit them to Oracle for labeling. Therefore, the key problem of active learning is how to select the samples to be labeled, that is, the query strategy. There are six kinds of query strategies: Uncertainty sampling, query-by-committee, expected model change, expected error reduction, variance reduction, and density-weighted methods^[17].

In this paper, we design a query-by-committee strategy for active learning. According to Definition 3, we use three K -NN classifiers as committee members $V' = \{V_1(1\text{-NN}), V_2(3\text{-NN}), V_3(5\text{-NN})\}$. For unlabeled sample x_u , these three classifiers will give three labels $\{\bar{y}_1, \bar{y}_2, \bar{y}_3\}$. Then there will three cases of the three labels:

(1) $\{\bar{y}_1, \bar{y}_2, \bar{y}_3\}$ contains three different labels, then the sample gets a difference value of three.

(2) $\{\bar{y}_1, \bar{y}_2, \bar{y}_3\}$ contains two different labels, then the sample gets a difference value of two.

(3) $\{\bar{y}_1, \bar{y}_2, \bar{y}_3\}$ are all same, then the sample gets a difference value of one.

According to these three cases, active learning selects the samples with the largest difference value and submits them to Oracle. Oracle gives true labels of the samples, and adds them to training set to learn a new matrix. It should be noted that samples with the same difference value have the same amount of information, and the value for metric learning is also the same. Therefore, if there are many samples with the largest difference value, we can randomly select a batch size of samples and submit them to Oracle.

2.2.3 ASVM²L algorithm

Combining SVM²L algorithm with active query strategy, we propose an active metric learning based support vector machines (ASVM²L). The process of ASVM²L is described as Algorithm 1. In Step 2, the algorithm selects the nearest neighbor samples of the same label and ones of different la-

bels for each sample to construct the triple constraints. Steps 3—5 solve SVM metric learning. Step 6 is the termination condition of the algorithm. Steps 7—8 use the active learning algorithm to select the samples with the largest difference value for labeling.

Algorithm 1 ASVM²L

Initial set: Committee set V' , Oracle O

Input: Unlabeled sample set X , the batch of selection batch, constraint parameter C , and initial parameters n

Output: metric matrix M

- (1) Select n samples to Oracle O for labeling. This labeled set is denoted as L_0 , and the rest of unlabeled set is named as U_0 .
- (2) Construct a triple constraint set TC_i based on the current labeled set L_i .
- (3) Solve Eq.(2) with the triple constraint TC_i , to get Lagrange multiplier α .
- (4) Solve M_i according to Eq.(7).
- (5) Project M_i into the positive semi-definite space to obtain the metric matrix M_i' .

(6) If $U_i = \emptyset$, the algorithm terminates.

(7) The unlabeled sample set S_i with the largest difference value in batch size is selected according to the designed active selection strategy.

(8) Let oracle O label for S_i , $L_{i+1} = L_i + S_i$, $U_{i+1} = U_i - S_i$, $i = i + 1$, back to Step 2.

2.3 Classify traffic scenes based on ASVM²L

Fig.1 is the framework of identifying similar traffic scenes based on ASVM²L. The left part represents the data pre-processing process. We extract 382 traffic features from original data and use principal component algorithm (PCA) to reduce the dimensionality to 60-D principal components. The middle part is the process of ASVM²L. According to the interaction between the metric learning model and Oracle, the optimal metric matrix is learned from the fewest labelled samples. The right part is the process of identifying similar traffic scenes. The new traffic scenes are classified by the existing labelled scenes and metric matrix, and the samples of the same class are similar scenes.

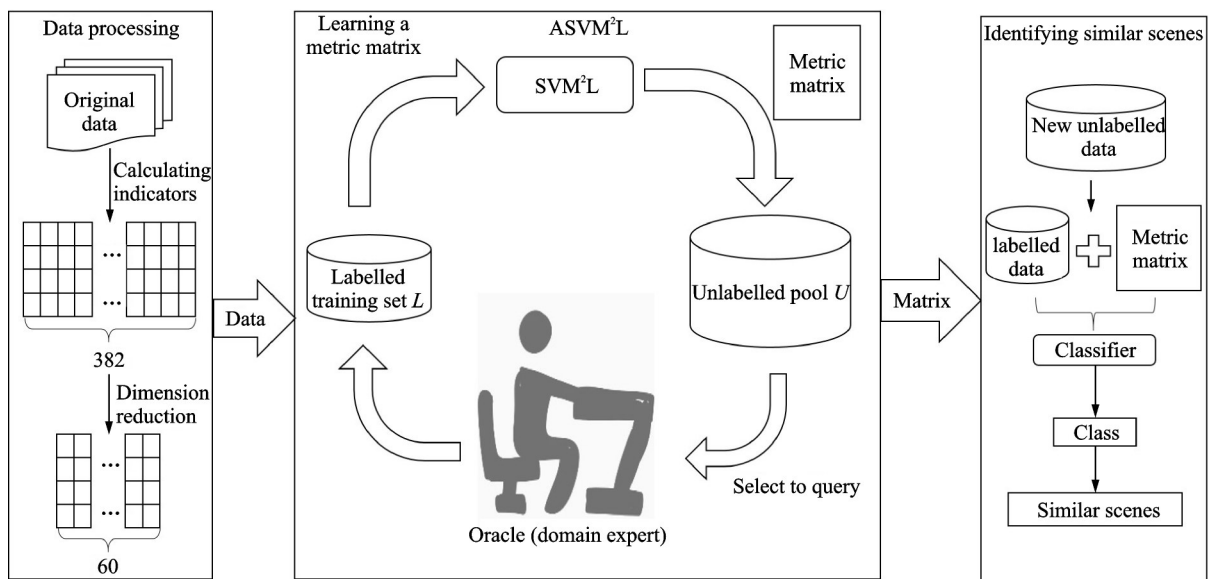


Fig.1 Framework of identifying similar traffic scenes based on ASVM²L

3 Experiments and Result Discussion

3.1 Validation of ASVM²L algorithm

To verify the effectiveness of ASVM²L algo-

rihm, we compare ASVM²L with other metric learning algorithms, LMNN, NCA, ITML, MC-ML, and SVM²L, on five UCI standard data sets shown in Table 1.

For each data set, 70% of the data are used for

Table 1 Standard data sets

Data set	Feature	Sample	Class
Madelon	50	2 600	2
DNA	180	2 000	3
Sat	36	6 435	6
Usps	50	9 298	10
Letter	16	15 000	26

metric matrix learning, and the remaining 30% are used for K -NN classification to verify the effect of metric learning. We calculate the 3-NN classification accuracy after metric learning for all the six algorithms on each data set. The experimental results are shown in Table 2. It can be seen that ASVM²L achieved the highest accuracy on two data sets and the second highest accuracy on three data sets. It can be concluded that ASVM²L has more stable and superior metric performance.

Table 2 Performance comparison of metric learning algorithms

Data set	LMNN	NCA	ITML	MCML	SVM ² L	ASVM ² L
Madelon	71.74	68.72	60.51	68.72	71.81	72.45
DNA	91.33	76.17	85.50	74.67	91.87	91.75
Sat	81.50	80.10	81.24	80.10	82.37	82.79
Usps	97.74	97.25	97.29	97.25	97.20	97.30
Letter	93.47	91.85	90.64	91.84	94.89	94.80

To verify the effectiveness of active learning query strategy, we compare ASVM²L with SVM²L using random sampling strategy on the five standard data sets. The learning curves of the two algorithms are shown in Fig.2. The red curve represents the 3-NN classification accuracy after ASVM²L, and the blue curve represents the 3-NN classification accuracy after random-sampling SVM²L. It can be seen that by introducing the active learning strategy, ASVM²L can achieve the better classification accuracy than the random-sampling SVM²L. The results on the five data sets all show that when the number of training samples reaches half of the size of all samples, the ASVM²L algorithm has obtained the learning performance similar to supervised metric learning, which means ASVM²L can greatly take advantage of the label information of limited labelled samples, or it can use fewer labelled samples to get better performance.

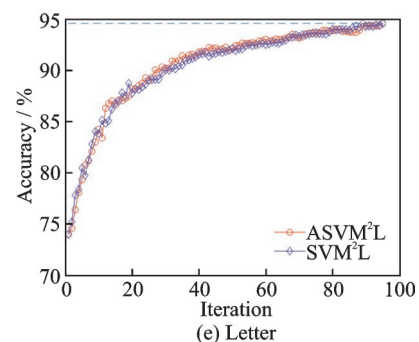
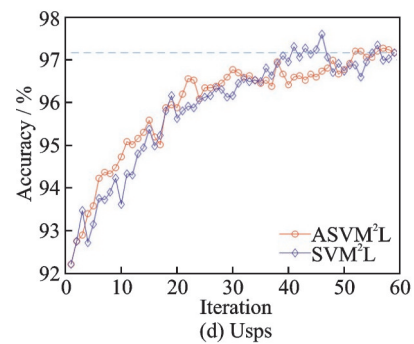
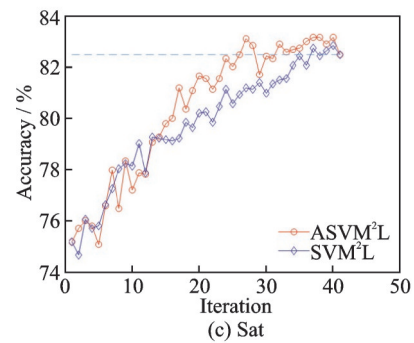
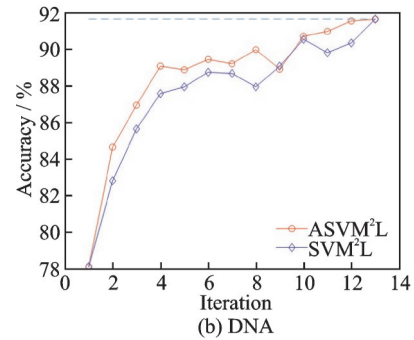
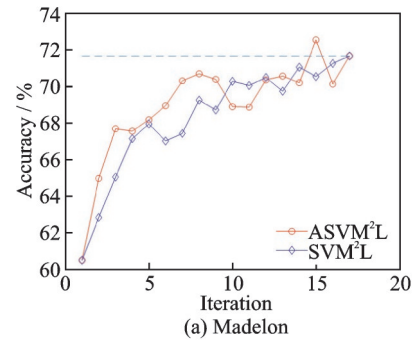


Fig.2 Learning curves of ASVM²L and random-sampling SVM²L on five data sets

3.2 Validation of ASVM²L for classifying traffic scenes

To verify the performance of ASVM²L for air traffic scenes classification, we use the real air traffic data including fly log data and fly plan data from of the No.5 regional sector in Central South of China. The data include 382 features, such as sector traffic flow, variance of heading, flight distance, flight time, average flight speed, and the other features of each altitude. According to the actual operation of the sector, the data are divided into two classes. The method of feature calculation is described in detail in Ref.[7].

Since the original air traffic data are large, and have high dimensionality and strong redundancy, we use PCA to reduce the dimensionality of the data by calculating the cumulative variance contribution rate. Fig.3 shows the change curve of cumulative variance contribution rate with the number of principal components. When we set the threshold to 0.85, we can reduce the dimensions of the original data from 382 to 60. That means the extracted 60 principal components can represent most of the information of the original data.

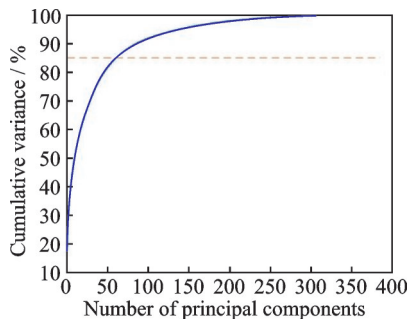


Fig.3 Change curve of cumulative variance

The feature correlation heat map after dimensionality reduction is shown in Fig.4. It can be seen that the extracted 60 principal components can be considered as having no redundancy.

After the data preprocessing, we use the first 70% of the data for metric matrix learning, and the last 30% for 3-NN to verify the ability of ASVM²L for improving the classification accuracy. We also apply six supervised metric learning algorithms to

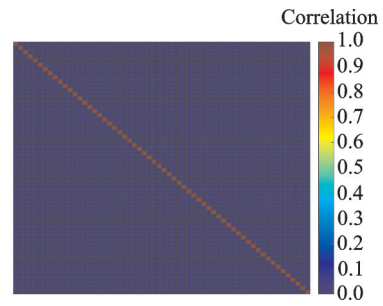


Fig.4 Feature correlation heat map after dimensionality reduction

conduct the same experiment for comparisons. The classification accuracies of the seven algorithms are listed in Table 3.

Table 3 Classification accuracy of traffic scenes

Algorithm	Type	Size of label	Accuracy/%
Euclidean	Supervised	6 183	81.58
LMNN	Supervised	6 183	81.80
NCA	Supervised	6 183	81.85
ITML	Supervised	6 183	82.03
MCML	Supervised	6 183	81.60
SVM ² L	Supervised	6 183	83.25
ASVM ² L	Semi-supervised	1 861	84.34

From Table 3, we can get that the 3-NN classification accuracy of all metric learning algorithms is better than the Euclidean distance. ASVM²L as a semi-supervised learning algorithm performs the best in classify traffic scenes compared with other supervised algorithms. What's more noteworthy is that ASVM²L only uses 30% of the labelled samples to achieve the best accuracy, which means that the proposed method can make better use of the label information of traffic scene samples when solving this problem.

Fig.5 shows the change of classification accuracy

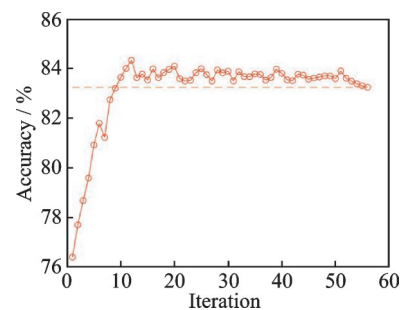


Fig.5 Learning curve of ASVM²L on air traffic data

cy of ASVM²L with iterations on the whole air traffic scene data set. It can be seen from Fig. 5 that when the ASVM²L runs to its 12th iteration, the classification accuracy reaches the highest 84.34%, and the number of labeled samples used for calculation at this time is 1 861. After that, the classification accuracy begins to decline slowly and gradually converges to 83.25%, which is exactly the classification accuracy of SVM²L on this data set. In other words, when we use the whole data set to learn ASVM²L, we just degenerate it into the supervised algorithm SVM²L, which completely fails to reflect the advantages of active learning in semi-supervised learning.

4 Conclusions

In order to classify air traffic scenes more accurately, in this paper, we propose a support vector machine-based active metric learning algorithm ASVM²L to measure the samples before classification. Considering that we currently do not have accurate labeled traffic scene samples, ASVM²L can combine knowledge of domain experts and machine learning results by using the active learning method. Through active query sampling, the most uncertain unlabeled samples with the most information are selected and labeled by domain experts, and then they can be used to train a better metric matrix, which will greatly help the subsequent classifier to improve the classification performance. Comparative experiments on five standard data sets prove that the ASVM²L algorithm proposed in this paper can make full use of the information of the limited labelled samples, and quickly obtain the best metric matrix. Comparative experiments on data of air traffic scenes also prove that as a semi-supervised metric learning algorithm, ASVM²L can achieve the best classification accuracy by much less labelled traffic scenes samples. Relying on the obtained classification model, more traffic scenes can be classified, and more similar scenes can be identified. Furthermore, control strategies in similar scene can be acquired and analyzed, and similar decision-making

rules can be found to help controllers make decisions quickly in the current scene.

References

- [1] GRABBE S R, SRIDHAR B, MUKHERJEE A. Similar days in the NAS: An airport perspective[C]// Proceedings of Aviation Technology, Integration, & Operations Conference. [S.l.]: AIAA, 2013.
- [2] GRABBE S R, SRIDHAR B, MUKHERJEE A. Clustering days and hours with similar airport traffic and weather conditions[J]. Journal of Aerospace Information Systems, 2014, 11(11): 751-763.
- [3] SUN D, SRIDHAR B, GRABBE S R. Disaggregation method for an aggregate traffic flow management model[J]. Journal of Guidance Control & Dynamics, 2015, 33(3): 666-676.
- [4] KUHN K D. A methodology for identifying similar days in air traffic flow management initiative planning[J]. Transportation Research, 2016(69): 1-15.
- [5] CHRISTIEN R, BENKOUAR A, CHABOUD T, et al. Air traffic complexity indicators & ATC sectors classification[C]//Proceedings of Digital Avionics Systems Conference. [S.l.]: IEEE, 2002: 1-7.
- [6] GRABBE S R, SRIDHAR B, MUKHERJEE A, et al. Traffic flow management impact on delay and fuel consumption: An Atlanta case study[J]. Air Traffic Control Quarterly, 2016, 20(3): 203-224.
- [7] HU Minghua, ZHANG Xuan, YUAN Ligang, et al. Identifying similar operation scenes for busy area sector dynamic management[J]. Transactions of Nanjing University of Aeronautics & Astronautics, 2020, 37(4): 615-629.
- [8] XING E P, NG A Y, JORDAN M I, et al. Distance metric learning with application to clustering with side-information[C]//Proceedings of NIPS 2002. [S.l.]: NIPS, 2002.
- [9] LIU W, WEN Y, YU Z, et al. Sphereface: Deep hypersphere embedding for face recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2017: 212-220.
- [10] PLÖTZ T, ROTH S. Neural nearest neighbors networks[J]. Advances in Neural Information Processing Systems, 2018, 31: 1087-1098.
- [11] CHOI J, MIN C, LEE B. Mathematical analysis on information-theoretic metric learning with application to supervised learning[J]. IEEE Access, 2019, 7: 121998-122005.

- [12] WU W, TAO D, LI H, et al. Deep features for person re-identification on metric learning[J]. Pattern Recognition, 2021, 110: 107424.
- [13] WANG F, ZUO W, ZHANG L, et al. A kernel classification framework for metric learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 26(9): 1950-1962.
- [14] PASOLLI E, YANG H L, CRAWFORD M M. Active-metric learning for classification of remotely sensed hyperspectral images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 54(4): 1925-1939.
- [15] PRIYADARSHINI K, GORU R, CHAUDHURI S, et al. Batch decorrelation for active metric learning [C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence and Seventeenth Pacific Rim International Conference on Artificial Intelligence. [S.l.]: IJCAI-PRICAI, 2020.
- [16] ZHOU Z H. A brief introduction to weakly supervised learning[J]. National Science Review, 2018, 5(1): 44-53.
- [17] AGGARWAL C C, KONG X, GU Q, et al. Active learning: A survey [M]//Data classification: Algo-

rithms and applications. [S.l.]: CRC Press, 2014: 571-605.

Acknowledgements This work was supported by the National Natural Science Foundation of China (No.61501229) and the Fundamental Research Funds for the Central Universities(Nos.2019054, 2020045).

Author Dr. CHEN Haiyan received her M.S. and Ph.D. degrees in computer science from Nanjing University of Aeronautics and Astronautics(NUAA), Nanjing, China, in 2005 and 2012, respectively. She is currently a lecturer at the College of Computer Science and Technology, NUAA. Her research interests include machine learning and artificial intelligence.

Author contributions Dr. CHEN Haiyan designed the study, analyzed the result and revised the manuscript. Mr. HOU Xiaye designed the algorithm, implemented the model and wrote the manuscript. Dr. YUAN Ligang contributed to the background and data of this study. Mr. ZHANG Bing contributed to the data preprocessing. All authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing interests.

(Production Editor: SUN Jing)

基于主动度量学习的空中交通相似场景识别

陈海燕^{1,3}, 侯夏晔¹, 袁立罡², 张 兵¹

(1.南京航空航天大学计算机科学与技术学院,南京 211106,中国;

2.南京航空航天大学民航学院,南京 211106,中国; 3.软件新技术与产业化协同创新中心,南京 210023,中国)

摘要:空中交通的快速增长不断增加了管制员的工作负荷,这已成为制约部门运行的重要因素。如果能够识别出相似的交通场景,就可以利用历史决策经验帮助管制员快速决策控制策略。考虑到交通场景众多且难以标记所有样本,本文提出了一种主动支持向量机度量学习算法(Active SVM metric learning algorithm, ASVM²L)来度量和识别相似的交通场景。首先获得了一些由资深空中交通管制员标记的交通场景样本;接着设计了一种基于投票差异的主动查询策略来选择最有价值的未标记样本交予领域专家进行标记;然后,利用ASVM²L从所有标记样本中学习到一个度量矩阵,用于后续分类算法完成相似场景的分类。在标准数据集上验证了ASVM²L的有效性,然后在中国中南扇区的历史空中交通数据集上对交通场景进行了度量和分类。实验结果表明,与现有的其他方法相比,本文所提方法能够更彻底地利用样本的信息,在有限的标记样本下达到更高的分类精度。

关键词:空中交通;相似场景;主动学习;度量学习;支持向量机