Ensemble Meteorological Cloud Classification Meets Internet of Dependable and Controllable Things

Jinglin Zhang, Pu Liu, Feng Zhang, Hironobu Iwabuchi, Antonio Artur de H. e Ayres de Moura, and Victor Hugo C. de Albuquerque, *Senior Member, IEEE*

***Abstract*—Advances in Internet of Things (IoT) and cloud/edge computing systems could precisely monitor the meteorological elements and environmental conditions. Remote automated observation system (RAOS) makes the full use of IoT to communicate with other sensors, enabling the active responses from passive devices for smart weather. Cloud observation and classification have been regarded as a successful application that could automatically perform emergency tasks in RAOS. However, with the increasing growth of resource exploitation, the performance of communications among the automatic observation platforms, and the efficiency of task allocation among them has become a critical challenge. In this article, an ensemble learning method and resource allocation scheme are proposed to realize the cloud observation and classification with the help of reliable and controllable infrastructures. On the one hand, several ensemble methods, like Bagging, AdaBoost, and Snapshot are selected as a base classifier to capture the cross-semantic and structure features of cloud, while applying them to the ensemble using convolutional neural networks with different base learners and residual neural networks with different depths. on the other hand, a particular cloud-edge distributed framework is proposed for cloud classification approach based on the intelligent network, to overcome the difficulty in the massive data transmission. The experimental results verify that the proposed ensemble approach achieves high accuracy of cloud classification, and effectively improves the number of allocated tasks. Ensemble methods can generate a more accurate prediction than any single classifier or the majority algorithms. It consistently yields lower error rates than single state-of-the-art models at no additional training cost.**

***Index Terms*—Cirrus cumulus stratus nimbus data set, convolutional neural networks, dependable and controllable things, ensemble learning (EL), meteorological cloud classification.**

Manuscript received February 20, 2020; revised July 18, 2020 and September 22, 2020; accepted October 21, 2020. Date of publication December 8, 2020; date of current version February 19, 2021. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFE0126100; in part by the National Natural Science Foundation of China under Grant 61702275, Grant 41675003, and Grant 41775008; and in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant SJKY19_0960. The work of Hironobu Iwabuchi was supported by the Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science under Grant Kiban B 17H02963. The work of Victor Hugo C. de Albuquerque was supported by the Brazilian National Council for Research and Development (CNPq) under Grant 304315/2017-6 and Grant 430274/2018-1. *(Jinglin Zhang and Pu Liu contributed equally to this work.) (Corresponding author: Jinglin Zhang.)*

Jinglin Zhang is with the School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210021, China (e-mail: jinglin.zhang@nuist.edu.cn).

Pu Liu is with the Key Laboratory of Meteorological Disaster, Ministry of Education, Nanjing University of Information Science and Technology, Nanjing 210021, China (e-mail: pu_liu@nuist.edu.cn).

Feng Zhang is with the Department of Atmospheric and Oceanic Sciences and Institute of Atmospheric Sciences, Fudan University, Shanghai 200438, China, and also with the Shanghai Qi Zhi Institute, Shanghai 200232, China (e-mail: fengzhang@fudan.edu.cn).

Hironobu Iwabuchi is with the Center for Atmospheric and Oceanic Studies, Tohoku University, Sendai 980-8577, Japan (e-mail: hironobu.iwabuchi.c4@tohoku.ac.jp).

Antonio Artur de H. e Ayres de Moura and Victor Hugo C. de Albuquerque are with the Graduate Program in Applied Informatics, University of Fortaleza, Fortaleza 60811-905, Brazil (e-mail: antonioartur@unifor.br; victor.albuquerque@unifor.br).

Digital Object Identifier 10.1109/JIOT.2020.3043289

I. Introduction

In THE era of computing paradigms, information communication technologies [1], have played an increasingly significant role in daily life. Internet of Things (IoT) aims to address the challenges in energy-related data collection, transportation, city infrastructure, healthcare, public safety, and support of smart systems. To achieve artificial intelligence-enabled Internet of dependable and controllable things, machine learning-based architecture should take into account the exceptional requirements for cloud resources, storage, and processing capacity [2], [3]. This provides the end-user with decisive computing, energy efficiency, mobility, location, and context awareness support [4], [5]. The integration of machine learning with IoT systems will create great opportunities for novel research like smart weather system which includes interdisciplinary efforts between atmospheric science and computer science to solve these challenges.

Remote automated observation system (RAOS) have some key benefits including infrared and panoramic cameras, gathering key weather elements and site information, using a time-lapse feature to monitor environmental changes, and request images on demand for weather event. In order to obtain accurate weather information, RAOS is significant as shown in Fig. 1. Especially, the cloud observation and classification has been regarded as a promising prototype to perform emergency tasks in RAOS [6], [7]. The architecture of Internet of dependable and controllable things for smart weather is illustrated in Fig. 2. Meanwhile, the information collected from various IoT devices (e.g., temperature and humidity sensor, multispectral camera and so on) in different locations is transferred to the remote servers.

Fig. 1. RAOS have some key benefits, including temperature and humidity sensor, rainfall recorder and anemometer, gathering key weather elements and site information, using a time-lapse feature to monitor environmental changes, and request images on demand for weather event.

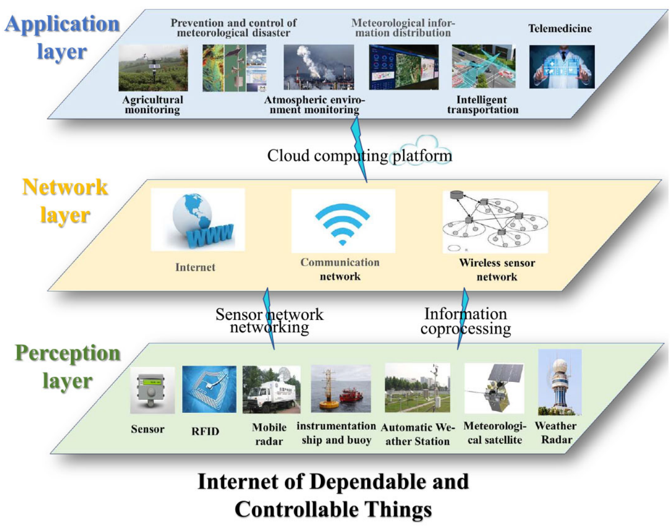


Fig. 2. Architecture of dependable and controllable things for smart weather. The information is gathered from various IoT devices in perception layer (e.g., temperature and humidity sensor, multispectral camera, rainfall recorder, anemometer, and mobile radar) in different locations. Then, the observed data is transferred to the remote servers.

As in meteorology, cloud shows in variety of forms and its feedbacks are the most uncertainty factors in climate model or credible predictions of climate change. Moreover, cloud-type variations are also main influences on earth radiation budget and energy balance [8]–[11]. Therefore, it helps to understand these process with an effective method, which classifies cloud categories automatically with high accuracy. In addition, with the increasing number of ground-based all-sky camera instruments, the tremendous amount of data demand a suitable image processing procedures to be fully exploited [12]–[15]. Thus, the automatic classification of ground observation's cloud data is critical in meteorological applications [16]. However, there still exist some common problems in the traditional methods that attempt to fully subtract the cloud texture, color and structure features. On the one hand, they cannot cover all the typical cloud classes, which are crucial for cloud observation. For examples, Singh and Glennen [17] and Dev *et al.* [18] only considered five classes conditions. Heinle *et al.* [19] and Calbo and Sabburg [15] classified cloud into seven and eight types, respectively. On the other hand, due to the complexity of algorithms, it is sometimes difficult to implement it to get a high classification result as a whole or a single class.

Recently, convolution neural networks (CNNs) have attracted many attentions due to machine learning challenges in image classification [20], image detection [21], localization [22], and so on. Deep neural networks have revealed the high variance and low bias due to their high capacity and flexibility, showing its potential in the improvement of predictive performance by modeling the average of multiple stochastically trained networks [23]. The goal of conventional machine learning approaches is to learn one hypothesis from training data, while ensemble learning (EL) methods use a set of hypotheses that are integrated to solve the same issue [24]. EL combines weak learners to work as a strong learner, and a single learner generalization ability is usually much weaker than that of ensemble [25]. The EL methods have the capability to combine several training baseline models results so as to provide better prediction performance in practice. The application of EL in the classification of ImageCLEF 2016 medical image public data set demonstrates that the fine-tuned CNNs ensemble achieves higher scores than new designed CNNs [26]. By combining the full advantages of different classifiers, the EL model wbased on deep learning is useful for cancer treatment [27].

EL methods have been successfully applied for remote sensing image classification. Using EL methods to classify remote sensing data with a few labeled training samples to achieve a high classification accuracy [28]. Due to the high dimensionality of hyperspectral remote sensing image, identified it is more difficult than other remote sensing imagery. Xia *et al.* [29] first proposed an EL approach with Rotation Forest, and applied them to hyperspectral remote sensing image classification, getting excellent effect in generating the classifier for the classification of hyperspectral image. Implemented EL methods in hyperspectral image classification also obtain satisfied results [30]. Li *et al.* [31] developed an enhanced EL method called multiclass boosted rotation forest (MBRF), which consistently exceeds the other referenced top classification methods. Although EL methods have been well applied in solving remote sensing imagery classification, the behavior of EL to classify meteorological cloud with deep neural networks needs to be updated. This article will investigates and compares different EL methods using CNNs with different structures and residual neural networks (RNNs) with different depths to find the best one, which is suitable for ground-based cloud classification.

Meanwhile, RAOS-assisted communication networks are faced with several challenges, such as inefficient communication, lack of resource allocation and incentive scheme. In this article, we propose a resource allocation scheme for RAOSs for meteorological cloud classification, to satisfy the demand of Internet of dependable and controllable things for smart weather. But there are still several challenges.

1) EL methods can not provide enough accuracy in remote sensing data classification, while the behavior of EL in classifying meteorological cloud with deep neural networks is still not fully resolved.
2) The characteristics of RAOSs are neglected and the RAOS-assisted communication systems in IoT are not involved.

To address these challenges, the main motivation of our work is to propose an ensemble framework for meteorological cloud images classification, which is flexible and efficient to be deployed on the cloud devices in Internet of dependable and controllable things. Meanwhile, the seamless integration of machine learning into IoT systems creates opportunities to solve these challenges. The remainder of this article is structured as follows. Section II describes the Bagging, AdaBoost, SAMME, stacking, super learner, and snapshot ensemble methods. Section III designs the experiments, including the details of the cirrus cumulus stratus nimbus (CCSN) data set and experimental settings. Section IV describes the performance of the ensemble methods on CCSN data sets. Finally, the conclusions are summarized in Section V.

II. Methodology

A. Bagging, AdaBoost, SAMME, Stacking, and Super Learner

Bootstrap aggregating, also called Bagging, [32] trains separate base learner and then combines them parallel to generate classifier. A base learner is trained by subsampling the training data set with replacement. AdaBoost [33] served as the well-known variants of Boosting is used to improve the final performance sequentially via selection of the training set and redistribution of weight iteratively. Stagewise additive modeling (SAMME) which was proposed by Hastie *et al.* [34] used a multiclass exponential to extend AdaBoost algorithm in binary classification to the multiclass case. The algorithm revealing satisfying effectiveness and reproducing ability has been proven.

The major difference between SAMME and AdaBoost is that, α_m in SAMME is changed into

$$\alpha^{(m)} = \log \frac{1 - \mathrm{err}^{(m)}}{\mathrm{err}^{(m)}} + \log(K - 1) \tag{1}$$

where K is the number of categories. When $K = 2$, SAMME equals to AdaBoost.

Stacking [35], [36] gathers strong and diverse sets of learners together, which are used to judge whether training data have been appropriately learned. These separate learners are then merged by a second-level learner (metalearner)

$$H_{\mathrm{Ensemble}} = h^{\mathrm{new}}(h_1(x), h_2(x), \ldots, h_T(x)) \tag{2}$$

where h^{new} is new classifier, which is generated based on the novel constructed data sets $h_1(x), h_2(x), \ldots, h_T(x)$.

The Super learner [37] is the development of stacking, which depends on ensemble framework with cross-validation by minimizing cross-validated risk for the combination.

$$\hat{\Psi}_{\mathrm{SL}}(W) = \sum_{k=1}^{K} \hat{\alpha}_k \hat{\Psi}_k(W) \tag{3}$$

where α represents the value that minimizes the cross-validated the candidate estimator's risk, and $\hat{\Psi}_k(W)$ is estimated by each algorithm in algorithm library of the entire data set.

B. Snapshot Ensemble, CNNs, and RNNs Ensembles

Snapshot ensemble was developed by Huang *et al.* [38]. The core of Snapshot ensemble is an optimization process which goes through several local minima before acquiring a final result. The purpose of this method is that to combine multiple neural networks without producing additional training cost and use the cyclic learning rate schedules [39] to get convergence rapidly. The output prediction is the average of m model's scores

$$h_{\mathrm{Ensemble}} = \frac{1}{m} \sum_{n=0}^{m-1} h_{M-i}(x) \tag{4}$$

where x is a sample for test and the $h_{M-i}(x)$ is the softmax result of snapshot i.

Furthermore, the general CNN architecture is used as base learner as shown in 3 and original RNNs configurations [21], [40] with depth of 22, 32, 44, and 56 are combined in the experiments. Fig. 3 shows the overall CNN architecture which includes two stages: 1) feature extraction and 2) classification. The RNN [21] is utilized to solve the problem of degradation, gradient vanishing and exploding by the batch normalization [41], to a large extent. CNNs with different base learner and RNNs with different depths are deployed to construct ensembles.

III. Experiment

A. Data Set

Three public computer vision data sets have been applied to evaluate the performance of the meteorological cloud classification methods.

1) *Cirrus Cumulus Stratus Nimbus (CCSN):* The CCSN data set is introduced in our previous work based on the meteorological criteria [42]. Table I presents the details descriptions of the cloud categories in the CCSN database. The data set contains only 2543 unique cloud images with 256 × 256 pixels in the JPEG format and includes most concern 10 cloud forms in the cloud observation. Fig. 4 displays some cloud images in the data set. Compared with other data sets, such as ImageNet [43], the quantity of training samples in CCSN is unrepresentative and negligible. However, the number of samples is three times larger than the available public cloud data set.
2) *SWIMCAT:* the SWIMCAT data set [18] consists of 784 sky/cloud patch images with 125 × 125 pixels. The data set is divided into five distinct categories: a) clear sky; b) patterned clouds; c) thick dark clouds; d) thick white

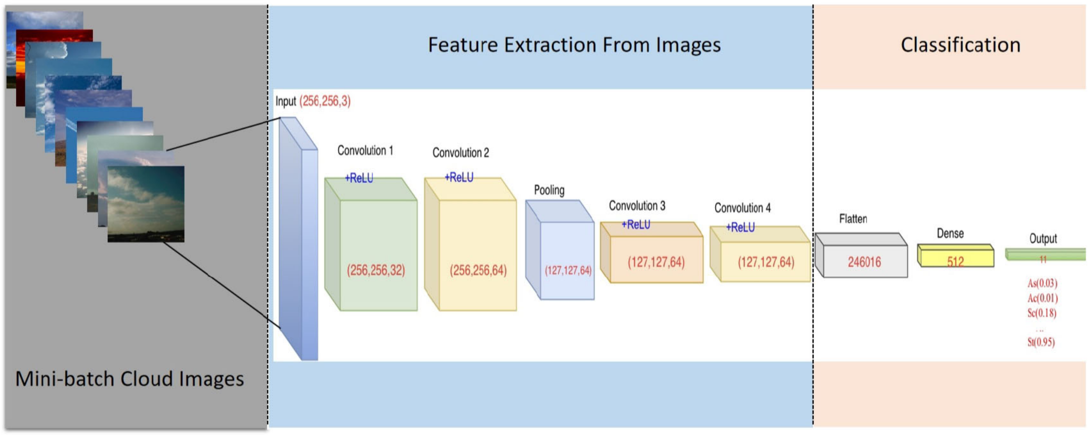


Fig. 3. General CNN model structure adopted in this article, it consists two stages: 1) feature extraction and 2) classification.

TABLE I
DETAILS OF CCSN DATABASE

Category	Number of Images	Type
Ci	139	Cirrus
Cs	287	Cirrostratus
Cc	268	Cirrocumulus
Ac	221	Altocumulus
As	188	Altostratus
Cu	182	Cumulus
Cb	242	Cumulonimbus
Ns	274	Nimbostratus
Sc	340	Stratocumulus
St	202	Stratus
Ct	200	Contrails
Total	2543	

Fig. 4. Cloud samples in the CCSN database followed by Ac, As, Sc, Ns, Ci, Cc, Cu, Cb, St, Ct, Cs. Each class is illustrated 11 samples.

clouds; and e) veil clouds. There are several researches with the SWIMCAT data set, where some necessary cloud categories are missing, making it insufficient for meteorological research and applications.

3) *Huazhong University of Science and Technology (HUST) Data Set:* The HUST contains 1231 cloud images, that are classified into six classes by the color census transform (CCT). The preprocessing CCT is applied on color sky images, which converts RGB values to opponent color space with census transform. Thus, the HUST is made up of the six-class database [44] followed by Ac, As, Cc, Clear sky, Cs, Cu, Sc, St, and Ci.

B. Experimental Settings

All the experiments are conducted using Keras [45] (version 2.2.4) with Tensorflow [46] (version 1.12) as backend running on NVIDIA GeForce GTX1080Ti. Table II presents the detailed parameter settings in the experiments. Bagging (R + MV)/(R + WA) stands for Bagging method with randomly picked training set and uses majority vote (weighted average) to combine base learners. Bagging (A + MV)/(A + WA) stands for Bagging method with the same training set and uses majority vote (weighted average) to combine base learners. Snapshot ensemble is trained in the CNNs ensembles using $100 \times L$ epochs (L is the number of base learners) with $\alpha_0 = 0.0001$. But, using $200 \times L$ epochs with $\alpha_0 = 0.0002$ train in the RNNs ensembles. SAMME and AdaBoost both employ $3m$ ($m = 2030$ is the size of the training set) training samples for each training stage. Moreover, both of them uses 50 epochs (80 epochs in the RNNs ensembles) to train a base learner. The rest of ensemble methods use 100 epochs (200 epochs in the RNN ensembles) to train a base learner.

C. Evaluation Indices

So as to evaluate the proposed approach and other comparative ensemble methods, true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs) are applied to calculate four evaluation indices: precision, recall, accuracy and F-measure. The precision, recall and accuracy are defined as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{5}$$

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{6}$$

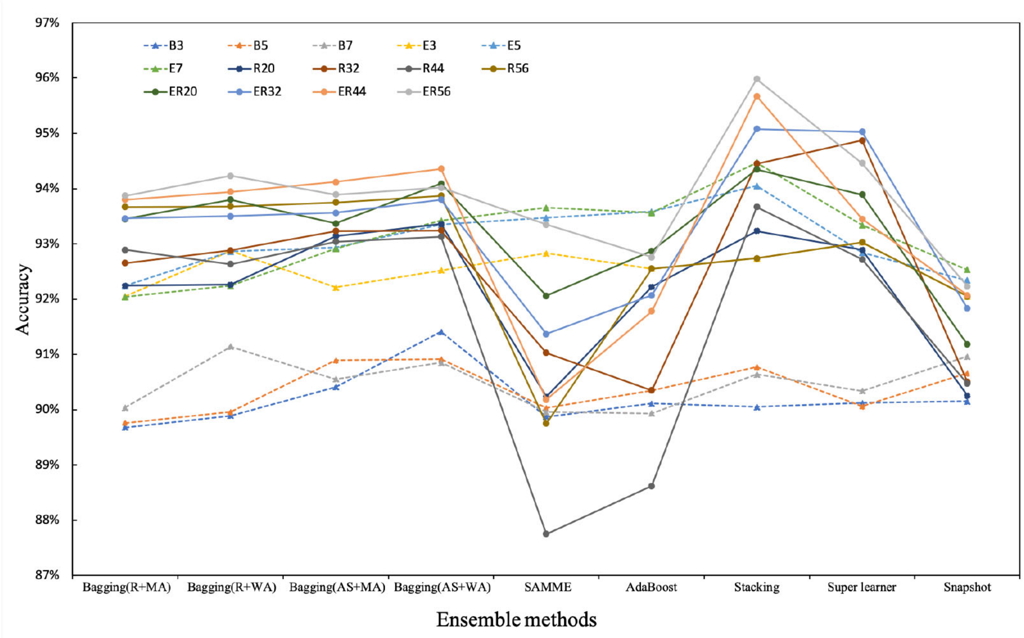


Fig. 5. Performance of ensemble different depths and different numbers base learner. We can see that the weighted averaging performs better than majority voting as an approach for combining base learners. Stacking outperforms all the other ensemble methods.

TABLE II
PARAMETER SETTINGS IN EXPERIMENT

Ensemble method	Training strategy	CNN ensembles Epochs	CNN ensembles α_0	RNN ensembles Epochs	RNN ensembles α_0
Bagging	R+MV	100		200	
	R+WA	100		200	
	A+MV	100		200	
	A+WA	100		200	
SAMME	3m	50		200	
AdaBoost		50		200	
Stacking		100		200	
Super learner		100		200	
Snapshot		$100 \times L$	0.0001	$200 \times L$	0.0002

Note. R = Random; WA = Weighted Average; MV = Majority Vote; A = All Samples. L is the number of base learners; m = 2,030 is the size of training set;

TP is positive true value when the model correctly predicts the positive class. TN indicates value of TN when the model correctly predicts the negative class. FP represents the value of FP when the model incorrectly predicts the positive class. FN stands for the value of FN when the model incorrectly predicts the negative class. Then the F-measure is the harmonic mean of the precision and recall which is used to measure the accuracy. F-measure is formulated as follows:

$$\text{F-measure} = \frac{\alpha^2 + 1 \times R \times P}{R + \alpha^2 \times P}. \tag{7}$$

IV. Result and Discussion

A. Results for Ensemble of CNNs and RNNs on CCSN Data Set

In this section, the performances of different ensemble methods are evaluated under different numbers of CNNs base learners and different depth of RNNs. In Fig. 5, B3 (B5, B7) stands for the CNNs using the 3 (5, 7) base learners, and E3 (E5, E7) presents ensemble accuracy with the 3 (5, 7) base learners. Similarly, the R20 (R32, R44, R56) stands for the residual networks with depth equal to 20 (32, 44, 56) combining 3 base learners. The ER20 (ER32, ER44, ER56) stands for the ensemble of a residual network with depth equal to 20 (32, 44, 56) combining 3 base learners. The results are depicted in Fig. 5, which demonstrates the weighted averaging performs better than majority voting in combination of base learners, and stacking outperforms all the other ensemble methods. Besides stacking, the ensemble methods of super learner demonstrate the great potential for meteorological cloud classification. This means that more base learners may lead to better performance according to the results of combination with the different number of base learners in the stacking. Snapshot ensemble does not play provide similar performance as other ensemble methods owing to its relatively low classifier diversity which only offers a single training strategy to generate all the base learners. Furthermore, comparing the ensemble with different depths of RNNs, it can be found deeper RNNs gives higher predictive accuracy. Snapshot ensemble in different depths does not work well compared with other algorithms. The reason is that all the base learners are produced by single training process and its learning rate settings might be too simple for RNNs training.

B. Several Ensemble Methods Evaluation

The inconsistent performance between EL methods and base learner on CCSN data set are listed in Table III, to illustrate the effectiveness of ensemble different depths and different numbers of base learner comparing with each individual methods. In Table III, we observe that the combination of base learners with different structures or more base learners can lead to better performance. For example, the effectiveness of stacking method increases by a large margin with 3.83% in E7–C7. The performance of SAMME varies greatly in the ER56-E56, followed by that of E5–C5. These increased results serve as

TABLE III
EFFECTIVENESS OF ENSEMBLE DIFFERENT DEPTHS AND DIFFERENT NUMBERS BASE LEARNER COMPARING WITH EACH INDIVIDUAL METHOD. UNIT IS PERCENTAGE

Ensemble method	E3-C3	E5-C5	E7-C7	ER20-R20	ER32-R32	ER44-R44	ER56-R56
Bagging (R+MV)	2.37	2.48	2	1.21	0.81	0.91	0.2
Bagging (R+WA)	**2.98**	2.9	1.1	1.54	0.62	1.31	0.55
Bagging (A+MV)	1.8	2.04	2.36	0.23	0.33	1.08	0.14
Bagging (A+WA)	1.11	2.44	2.57	0.73	0.56	1.23	0.15
SAMME	2.96	**3.44**	3.69	**1.83**	0.34	2.43	**3.6**
AdaBoost	2.44	3.23	3.63	0.65	**1.72**	**3.16**	0.21
Stacking	2.69	3.28	**3.83**	1.12	0.63	2	3.24
Super learner	2.91	2.78	3	1	0.16	0.73	1.43
Snapshot	1.9	1.68	1.57	0.93	1.33	1.6	0.18

Note. R = Random; WA = Weighted Average; MV = Majority Vote; A = All Samples;

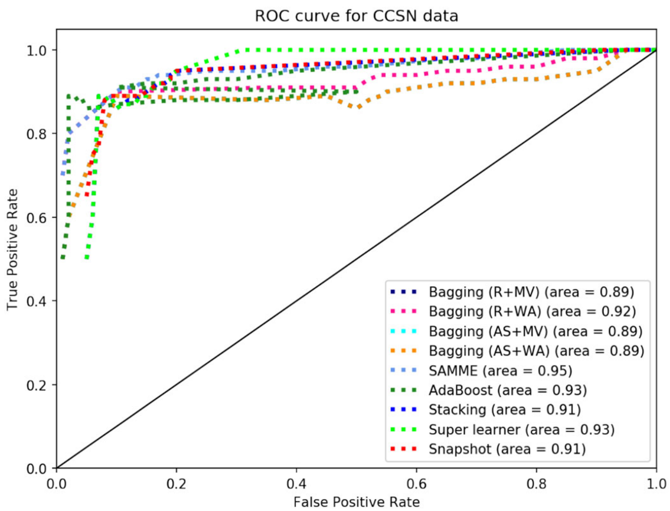


Fig. 6. ROC curves for several ensemble methods (e.g., Bagging, SAMME, AdaBoost, and super learner) with CCSN data set. The SAMME strategy gets higher AUC scores by combining multiple different classifiers, which obtains higher classification result (0.95) than individual classifier operating alone.

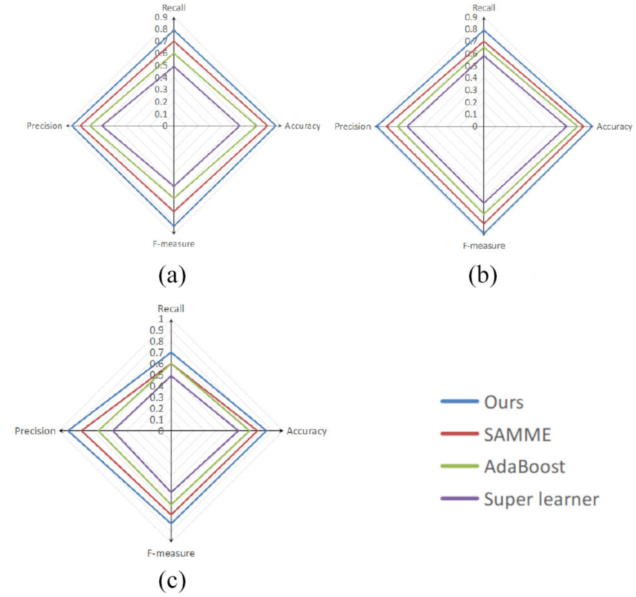


Fig. 7. Better performance of our approach than the comparative ensemble methods evaluated by four evaluation indices: 1) precision; 2) recall; 3) F-measure; and 4) accuracy. The three radar charts correspond to the performance on three data sets (CCSN, SWIMCAT, and HUST), respectively. The contour line displays the values ranged from 0 to 1. The closed curves with different colors represent different methods. The color legend is at the lower right corner of this figure.

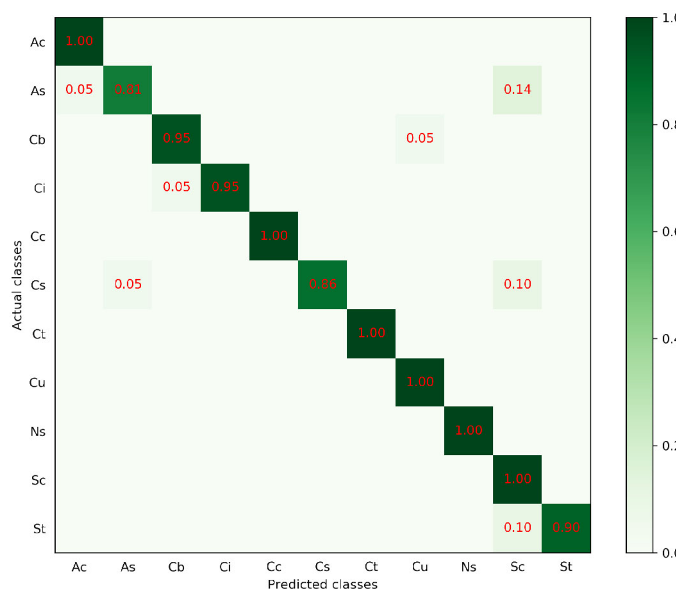


Fig. 8. Confusion matrix is used to evaluate the performance of EL methods.

good evidence to prove that the ensemble model is robust and effective. Moreover, the predictive results of different ensemble methods on the CCSN data set are compared by utilizing the receiver operating characteristic (ROC) curve. In addition, the area under the curve (AUC) is thought to be an essential evaluation for model comparison, which measures the relationship between precision and recall. The large area means both high precision and high recall. Fig. 6 demonstrates nine ROC curves with various ensemble schemes on the CCSN data set. By combining multiple different classifiers, the SAMME strategy gets higher AUC scores and obtains higher classification result (0.95) than individual classifier operating. Furthermore, the AUC scores of SAMME ensemble method obtain an area, which is higher than other methods on the CCSN data set, owing to its ability to learn and subtract semantic features automatically.

Fig. 7 illustrates the radar chart that corresponds to our approach performing better with four evaluation indices (precision, recall, F-meausre and accuracy) than the state-of-the-art methods. The vertex in the radar chart represents the evaluation index, and the contour line shows the values ranged from 0 to 1. The results display that the performance curve of our approach (blue curve) stays outermost in all of the three radar charts. Thus, the performance curve of our approach has the larger area than those of SAMME (red color), AdaBoost (green color), and super learner (purple color).

C. Effectiveness of the Stacking Methods

The confusion matrix is used to evaluate the performance of the stacking methods, and distinguish the type of cloud. The results shown in Fig. 8 are calculated with 21 training samples per class in CCSN data set. However, there are only 513 samples in validation data set, 231 cloud image in compute confusion matrix. The advantage of Stacking is more significant, especially acquiring perfect scores in Ac, Cc, Ct, Cu, Ns, Sc categories. The effectiveness of Stacking strategy is also verified in our experiments, by comparing with previous work [42]. Comparing with two confusion matrix, the accuracy of As and Cs are the most evidently improvement, with 0.39 and 0.19, respectively.

V. Conclusion and Future Work

Internet of dependable and controllable things is crucial in the smart climate. The numerous RAOS devices provide key information in the management of Internet of dependable and controllable things. In this article, EL methods are proposed to adopt the dependable and controllable infrastructures for cloud observation and classification in the smart climate application. EL approaches are proposed to increase the ground-based cloud classification accuracy when only limited labeled training samples are available, while the resource allocation scheme for RAOS improve the utilization of network resource and the efficiency of task allocation. The classification results demonstrate that ensembles methods can often outperform any individual classifier and so it can be considered as a excellent learning algorithm in contrast with base learners. Besides, weighted averaging is superior to majority voting in combination of base learners. Stacking usually has better performance than other ensemble strategies. As the metamodel training in stacking is much more efficient and robust than a base learner. Future work is to develop more sophisticated classification techniques, such as semisupervised learning or GAN methodologies, to improve cloud classification accuracy.

Acknowledgment

Their database examples are available at https://github.com/upuil/CCSN-Database. The code in the article was based in part on the source code of the Github page https://github.com/zhangyaqi1989/Ensemble-Methods-for-Image-Classification.

References

[1] K. Kai, W. Cong, and L. Tao, "Fog computing for vehicular ad-hoc networks: Paradigms, scenarios, and issues," *J. China Univ. Posts Telecommun.*, vol. 23, no. 2, pp. 56–65, 2016.
[2] C. Chen, C. Wang, T. Qiu, Z. Xu, and H. Song, "A robust active safety enhancement strategy with learning mechanism in vehicular networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 12, pp. 5160–5176, Dec. 2020.
[3] F. Wang, D. Jiang, H. Wen, and H. Song, "Adaboost-based security level classification of mobile intelligent terminals," *J. Supercomput.*, vol. 75, no. 9, pp. 7460–7478, 2019.
[4] M. Zheng, S. Chen, W. Liang, and M. Song, "NSAC: A novel clustering protocol in cognitive radio sensor networks for Internet of Things," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5864–5865, Jun. 2019.
[5] Z. Li, R. Chen, L. Lu, and G. Min, "Dynamic resource discovery based on preference and movement pattern similarity for large-scale social Internet of Things," *IEEE Internet Things J.*, vol. 3, no. 4, pp. 581–589, Aug. 2016.
[6] S. Wang *et al.*, "A fast CP-ABE system for cyber-physical security and privacy in mobile healthcare network," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 4467–4477, Jul./Aug. 2020.
[7] J. Liang *et al.*, "Deepfuzzer: Accelerated deep greybox fuzzing," *IEEE Trans. Dependable Secure Comput.*, early access, Dec. 20, 2019, doi: 10.1109/TDSC.2019.2961339.
[8] D. P. Duda, P. Minnis, K. Khlopenkov, T. L. Chee, and R. Boeke, "Estimation of 2006 northern hemisphere contrail coverage using MODIS data," *Geophys. Res. Lett.*, vol. 40, no. 3, pp. 612–617, 2013.
[9] W. B. Rossow and R. A. Schiffer, "ISCCP cloud data products," *Bull. Amer. Meteorol. Soc.*, vol. 72, no. 1, pp. 2–20, 1991.
[10] T. Chen, W. B. Rossow, and Y. Zhang, "Radiative effects of cloud-type variations," *J. Climate*, vol. 13, no. 1, pp. 264–286, 2000.
[11] G. L. Stephens, "Cloud feedbacks in the climate system: A critical review," *J. Climate*, vol. 18, no. 2, pp. 237–273, 2005.
[12] J. E. Shields, M. Karr, T. Tooman, D. Sowle, and S. Moore, "The whole sky imager—A year of progress," in *Proc. 8th Atmospheric Radiat. Meas. (ARM) Sci. Team Meeting*, Tucson, Arizona, 1998, pp. 23–27.
[13] G. Pfister, R. McKenzie, J. Liley, A. Thomas, B. Forgan, and C. N. Long, "Cloud coverage based on all-sky imaging and its impact on surface solar irradiance," *J. Appl. Meteorol.*, vol. 42, no. 10, pp. 1421–1434, 2003.
[14] C. N. Long, J. M. Sabburg, J. Calbó, and D. Pages, "Retrieving cloud characteristics from ground-based daytime color all-sky images," *J. Atmos. Ocean. Technol.*, vol. 23, no. 5, pp. 633–652, 2006.
[15] J. Calbo and J. Sabburg, "Feature extraction from whole-sky ground-based images for cloud-type recognition," *J. Atmos. Ocean. Technol.*, vol. 25, no. 1, pp. 3–14, 2008.
[16] M. Azimi-Sadjadi and S. Zekavat, "Cloud classification using support vector machines," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. Taking Pulse Planet Role Remote Sens. Manag. Environ. (IGARSS)*, vol. 2, 2000, pp. 669–671.
[17] M. Singh and M. Glennen, "Automated ground-based cloud recognition," *Pattern Anal. Appl.*, vol. 8, no. 3, pp. 258–271, 2005.
[18] S. Dev, Y. H. Lee, and S. Winkler, "Categorization of cloud image patches using an improved texton-based approach," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2015, pp. 422–426.
[19] A. Heinle, A. Macke, and A. Srivastav, "Automatic cloud classification of whole sky images," *Atmos. Measurement Tech.*, vol. 3, no. 3, pp. 557–567, 2010.
[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
[22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Representations*, May 2015. [Online]. Available: arXiv:1409.1556.
[23] C. Ju, A. Bibaut, and M. van der Laan, "The relative performance of ensemble methods with deep convolutional neural networks for image classification," *J. Appl. Stat.*, vol. 45, no. 15, pp. 2800–2818, 2018.
[24] Z.-H. Zhou. *Ensemble Learning*. Accessed: Feb. 2020. [Online]. Available: http://cs.nju.edu.cn/zhouzh/zhouzh.files/publication/springerEBR09.pdf
[25] T. G. Ditterrich, "Machine learning research: Four current direction," *Artif. Intell. Mag.*, vol. 18, no. 4, pp. 97–136, 1997.
[26] A. Kumar, J. Kim, D. Lyndon, M. Fulham, and D. Feng, "An ensemble of fine-tuned convolutional neural networks for medical image classification," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 31–40, Jan. 2017.
[27] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "A deep learning-based multi-model ensemble method for cancer prediction," *Comput. Methods Programs Biomed.*, vol. 153, pp. 1–9, Jan. 2018.
[28] M. Han and B. Liu, "Ensemble of extreme learning machine for remote sensing image classification," *Neurocomputing*, vol. 149, pp. 65–70, Feb. 2015.
[29] J. Xia, P. Du, X. He, and J. Chanussot, "Hyperspectral remote sensing image classification based on rotation forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 11, no. 1, pp. 239–243, Jan. 2014.
[30] A. Samat, P. Du, S. Liu, J. Li, and L. Cheng, "E^2LMs: Ensemble extreme learning machines for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1060–1069, Apr. 2014.
[31] F. Li, L. Xu, P. Siva, A. Wong, and D. A. Clausi, "Hyperspectral image classification with limited labeled training samples using enhanced ensemble learning and conditional random fields," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2427–2438, Jun. 2015.
[32] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996.
[33] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997.
[34] T. Hastie, S. Rosset, J. Zhu, and H. Zou, "Multi-class adaboost," *Statist. Interface*, vol. 2, no. 3, pp. 349–360, 2009.
[35] D. H. Wolpert, "Stacked generalization," *Neural Netw.*, vol. 5, no. 2, pp. 241–259, 1992.

[36] L. Breiman, "Stacked regressions," *Mach. Learn.*, vol. 24, no. 1, pp. 49–64, 1996.

[37] S. E. Sinisi, E. C. Polley, M. L. Petersen, S.-Y. Rhee, and M. J. van der Laan, "Super learning: An application to the prediction of HIV-1 drug resistance," *Statist. Appl. Genet. Mol. Biol.*, vol. 6, no. 1, pp. 1–24, 2007.

[38] G. Huang, Y. Li, G. Pleiss, Z. Liu, J. E. Hopcroft, and K. Q. Weinberger, "Snapshot ensembles: Train 1, get M for free," 2017. [Online]. Available: http://arxiv.org/abs/1704.00109.

[39] I. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," 2016. [Online]. Available: arXiv:1608.03983.

[40] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 630–645.

[41] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015. [Online]. Available: arXiv:1502.03167.

[42] J. Zhang, P. Liu, F. Zhang, and Q. Song, "Cloudnet: Ground-based cloud classification with deep convolutional neural network," *Geophys. Res. Lett.*, vol. 45, no. 16, pp. 8665–8672, 2018.

[43] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2009, pp. 248–255.

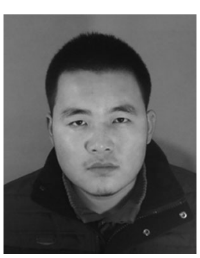
[44] W. Zhuo, Z.-G. Cao, and Y. Xiao, "Cloud classification of ground-based images using texture–structure features," *J. Atmos. Ocean. Technol.*, vol. 31, no. 1, pp. 79–92, 2014.

[45] F. Chollet *et al.* (2015). *Keras*. [Online]. Available: https://github.com/fchollet/keras

[46] M. Abadi *et al.* (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. [Online]. Available: https://www.tensorflow.org/

Jinglin Zhang received the M.S. degree in circuits and systems from Shanghai University, Shanghai, China, in 2010, and the Ph.D. degree in electronics and telecommunications from the National Institute of Applied Sciences-Rennes (INSA de Rennes), Rennes, France, in 2013.

Since 2014, he has been on the faculty of the School of Atmospheric Science, Nanjing University of Information Science and Technology, Nanjing, China. His current research interest includes computer vision, high-performance computing, interdisciplinary research with pattern recognition and atmospheric science.

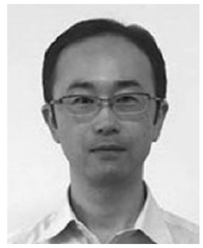


Pu Liu is currently pursuing the master's degree with the School of Atmospheric Science, Nanjing University of Information Science and Technology, Nanjing, China.

His research is mainly focused on the computer vision and deep learning in meteorological.

Feng Zhang received the M.S. degree in atmospheric physics from the Chinese Academy of Meteorological Sciences, Beijing, China, in 2010, and the Ph.D. degree in meteorology from the University of Chinese Academy of Sciences, Beijing, in 2013.

His current research interests include radiative transfer theory, remote sensing, and machine learning in atmospheric sciences.



Hironobu Iwabuchi received the Doctor of Science degree in geophysics from the Graduate School of Science, Tohoku University, Sendai, Japan, in 2001.

From 2002 to 2010, he worked as a Research Scientist with the Frontier Research System for Global Change and the Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan, and the Department of Atmospheric Sciences, Texas A&M University, College Station, TX, USA. Since 2011, he has been a Tenured Associate Professor with the Graduate School of Science, Tohoku University. His current research interests are on modeling of atmospheric radiative transfer, remote sensing of clouds and roles of atmospheric particles in the Earths climate system. The researches are conducted in his group and in collaboration with domestic and international research groups.

Antonio Artur de H. e Ayres de Moura received the bachelor's degree in mechanical engineering from the University of Fortaleza (UNIFOR), Fortaleza, Brazil, in 2019, a year of it done from Dalian University of Technology, Dalian, China. He is currently pursuing the master's degree in artificial intelligence with UNIFOR.

He is a Data Scientist with the Superintendence for Research and Public Safety Strategy of Cear State, UNIFOR. He is deeply involved in the field of artificial intelligence with experience in computer vision, natural language processing, and social network analysis. He is currently works with reinforcement learning applied to fleet management and first response on the state's emergency services, machine learning, deep learning, and artificial intelligence.

Victor Hugo C. de Albuquerque (Senior Member, IEEE) received the graduated in mechatronics engineering from the Federal Center of Technological Education of Ceara, Fortaleza, Brazil, the M.Sc. degree in teleinformatics engineering from the Federal University of Ceará, Fortaleza, and the Ph.D. degree in mechanical engineering from the Federal University of Paraiba, João Pessoa, Brazil.

He is a Professor and a Senior Researcher with the University of Fortaleza (UNIFOR), Fortaleza, where he is a Data Science Director with the Superintendency for Research and Public Safety Strategy of Ceara State. He is currently an Associate Professor of Graduate Program in Applied Informatics with UNIFOR and the Leader of the Industrial Informatics, Electronics and Health Research Group (CNPq). He is a specialist, mainly, in IoT, machine/deep learning, pattern recognition, and robotic.