Research Article

Redundancy Removal Adversarial Active Learning Based on Norm Online Uncertainty Indicator

Jifeng Guo ⓘ**, Zhiqi Pang** ⓘ**, Wenbo Sun** ⓘ**, Shi Li, and Yu Chen** ⓘ

College of Information, Northeast Forestry University, Harbin 150040, China

Correspondence should be addressed to Yu Chen; 1491228435@qq.com

Active learning aims to select the most valuable unlabelled samples for annotation. In this paper, we propose a redundancy removal adversarial active learning (RRAAL) method based on norm online uncertainty indicator, which selects samples based on their distribution, uncertainty, and redundancy. RRAAL includes a representation generator, state discriminator, and redundancy removal module (RRM). The purpose of the representation generator is to learn the feature representation of a sample, and the state discriminator predicts the state of the feature vector after concatenation. We added a sample discriminator to the representation generator to improve the representation learning ability of the generator and designed a norm online uncertainty indicator (Norm-OUI) to provide a more accurate uncertainty score for the state discriminator. In addition, we designed an RRM based on a greedy algorithm to reduce the number of redundant samples in the labelled pool. The experimental results on four datasets show that the state discriminator, Norm-OUI, and RRM can improve the performance of RRAAL, and RRAAL outperforms the previous state-of-the-art active learning methods.

1. Introduction

In recent years, image processing tasks based on deep learning [1–3] have achieved great success, but they mainly rely on a large number of labelled datasets. Although supervised learning has better performance than semi-supervised learning [4–10] and self-supervised learning [11], it is highly dependent on labelled data. In reality, it is very difficult or even unrealistic to obtain a large number of labelled datasets in many fields, and it inevitably consumes many resources [12, 13]. To mitigate the impact of such problems, some researchers have proposed active learning [14, 15]. The process of active learning is to select or synthesize the most useful samples from the unlabelled samples for model training, then use Oracle to label the selected samples, and finally add the labelled samples to the labelled pool to update the task model for training. The process is repeated until the performance of the task model meets the requirements or the label budget is exhausted. At present, active learning has been widely used in image classification [16–18] and segmentation [19, 20] tasks and has made some achievements.

The recently proposed SRAAL method [21] uses annotated information and labelled/unlabelled state information to select samples and has achieved competitive performance. SRAAL inherited VAAL's idea of adversarial learning [22]; that is, SRAAL uses the adversarial approach [23, 24] to learn the feature representation of labelled samples and unlabelled samples and selects samples that can increase the diversity of the labelled pool according to the distribution of samples. In addition, SRAAL sets up an online uncertainty indicator (OUI) for the unlabelled samples to calculate the contribution of this sample to the model. The OUI considers the influence of the maximum element and variance in the category vector on the uncertainty. In summary, SRAAL comprehensively considers the diversity and uncertainty of samples.

We performed a visual analysis of the samples selected by SRAAL [21] and found that some samples have extremely high similarity. In this paper, we refer to similar samples as redundant samples. Redundant samples increase the annotation cost but contribute little to improvements in model performance. To solve this problem, we designed the redundancy removal module (RRM), which defines the

threshold value for the feature distance of samples, effectively avoiding the influence of redundant samples.

In addition, we found in the experiment that the OUI used the whole category vector to calculate the uncertainty score, and the score was positively correlated with the variance of the vector. This is not a good option. The reasons are as follows: due to the introduction of softmax, most elements in the category vector of the sample are close to zero, and we call these elements "tiny values." For a dataset with a small number of categories, these tiny values can seriously affect the variance. For example, for a dataset with 10 categories, the variances of the vectors $[0.6, 0.4, 0, 0, 0, 0, 0, 0, 0, 0]$ and $[0.5, 0.5, 0, 0, 0, 0, 0, 0, 0, 0]$ are 0.42 and 0.40, respectively, which are very similar numerically, while in fact, their uncertainties are quite different. To remedy this drawback, we designed a new OUI named Norm-OUI, which no longer relies on the variance to calculate the uncertainty but uses the p-norm of the vectors. Norm-OUI is more sensitive to the uncertainty of vectors.

The main contributions of this paper are summarized as follows:

(1) We propose a redundancy removal adversarial active learning (RRAAL) method based on norm online uncertainty indicator, which fully considers the diversity, uncertainty, and redundancy of samples

(2) We design a sample discriminator to improve the representational learning ability of the generator and proposed a Norm-OUI based on the p-norm to calculate the uncertainty score of the samples

(3) We design an RRM to remove redundant samples and thus reduce inefficient labelling

2. Related Work

Current mainstream active learning methods can be divided into synthesis-based methods [25, 26] and pool-based methods [27–29]. The method in this paper is a pool-based method. Pool-based methods can be divided into uncertainty-based and distribution-based methods.

Uncertainty-based methods [30–33] select the most uncertain samples for the model in each iteration. For example, in the realm of Bayesian frameworks, Gaussian processes [30, 31] are used to assess the uncertainty of samples. In addition, Bayesian optimization [34, 35] has many application scenarios. In the realm of non-Bayesian frameworks, the distance from the decision boundary [32] and expected risk minimization [33] are used to assess the uncertainty of samples. Yoo et al. proposed a method based on a loss prediction module (LPM) [36] to predict the sample uncertainty. Uncertainty-based methods often depend on the performance of the task model, and the samples selected are directly related to the task model.

Distribution-based methods [22, 37] tend to select samples that increase the diversity of the labelled pool. By taking advantage of the image distance, a core-set approach [37] can select a set of data points from an unlabelled dataset and obtain a result that a model learned from the selected subset that is competitive for the remaining data points. VAAL [22] uses the adversarial learning [23, 24] of a variational autoencoder (VAE) [38] and discriminator to learn the feature representations of labelled samples and unlabelled samples and then uses the difference between them to make a sample selection. In essence, the method selects samples based on their diversity, which is not equal to the amount of information contained in a sample, so the results of the method may be unreliable. SRAAL [21] uses annotated information and labelled/unlabelled state information to select samples and fully considers the distribution and uncertainty of the samples. Our method also takes into account the uncertainty and diversity of the samples. In addition, we also consider the redundancy of the samples. The experimental results verify that RRAAL is superior to the existing pool-based methods.

The purpose of the synthesis-based methods is to synthesize the most useful samples for the model by using the generated model [24, 39]. The idea was first proposed in GAAL [25], which uses a GAN to generate samples closer to the decision boundary than the existing samples. BGADL [26] combines BDA [40] and BALD [41] to perform iterative training on the task model and the generated model, thus improving the performance of the task model. Similarly, ARAL [42] also uses generated images to update the task model. Synthesis-based methods have higher complexity, and their performance depends on the performance of both the generation model and the task model.

3. Method

In this section, we describe the RRAAL model presented in this paper. RRAAL selects the unlabelled sample with the most information based on the uncertainty, distribution, and redundancy of the sample, and its overall architecture is shown in Figure 1. RRAAL is composed of a representation generator (Section 3.2), a state discriminator (Section 3.3), and an RRM (Section 3.4). The representation generator is used to learn the feature representations of both labelled and unlabelled samples. The state discriminator predicts the state value of the sample according to the concatenated feature vector. The RRM selects a set of samples with the lowest redundancy on this basis. Section 3.4 introduces the sampling strategy based on the above three modules.

3.1. Unified Representation Generator. The unified representation generator of SRAAL includes an encoder, an unsupervised image reconstructor (UIR), and a supervised target learner (STL). The UIR learns the feature representation of the sample by reconstructing the sample, while the STL is used to embed the annotation information of the sample into the representation. The UIR is composed of transposed convolutional layers, and the STL is similar in structure to the task model. To improve the reconstruction ability of the UIR and further improve the ability of the encoder and UIR to learn the sample representation, we added a sample discriminator D_1 after the UIR to guide the reconstruction process of the encoder and UIR, as shown in Figure 1. The optimization objective of the sample discriminant D_1 is defined as follows:
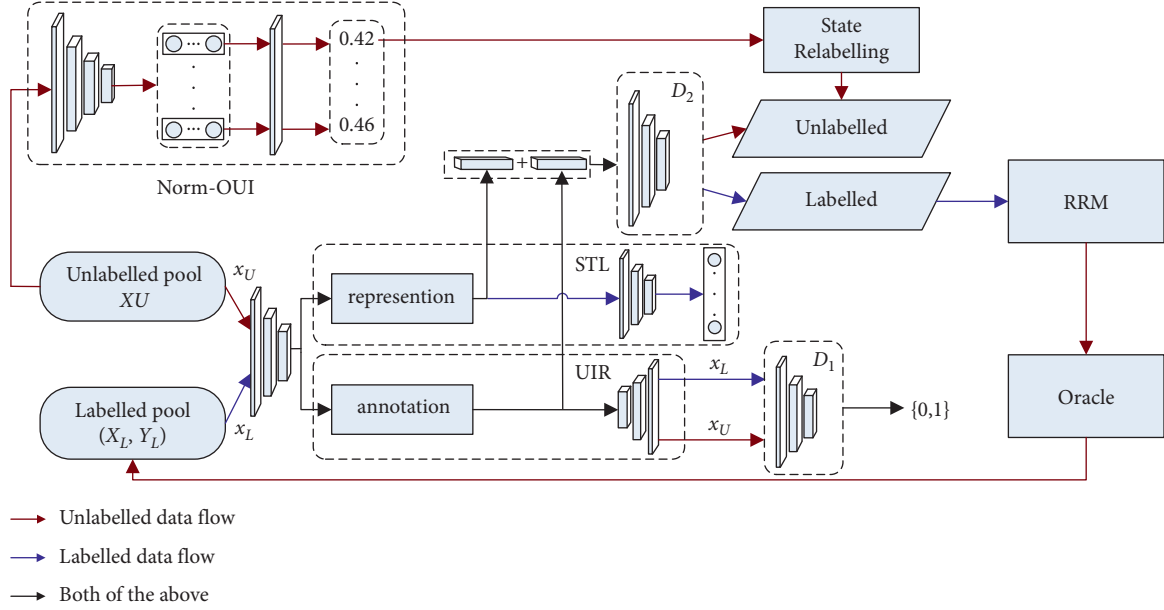
FIGURE 1: Network architecture of RRAAL. X_U represents the unlabelled pool, and (X_L, Y_L) represents the labelled pool, where Y_L is the label set. We used Oracle to label the unlabelled samples.

$$L_{D_1} = L_{D_1}^U + L_{D_1}^L,$$
$$L_{D_1}^U = -E[\log D_1(x_U)] - E[\log(1 - D_1)(x_U')], \quad (1)$$
$$L_{D_1}^L = -E[\log D_1(x_L)] - E[\log(1 - D_1)(x_L')],$$

where x_L and x_U are real labelled samples and unlabelled samples, respectively, while x_L' and x_U' are the generated labelled samples and unlabelled samples, respectively.

The optimization objective of the UIR is defined as follows:

$$L_{UIR} = L_{UIR}^U + L_{UIR}^L,$$
$$L_{UIR}^U = E\left[\log\left[p_\phi(x_U|z_U)\right] - D_{KL}\big(q_\theta(z_U|x_U)\|p(z)\big)\right] - L_{D_1}^U,$$
$$L_{UIR}^L = E\left[\log\left[p_\phi(x_L|z_L)\right] - D_{KL}\big(q_\theta(z_L|x_L)\|p(z)\big)\right] - L_{D_1}^L,$$
$$(2)$$

where L_{UIR}^U is the objective function of the unlabelled sample, L_{UIR}^L is the objective function of the labelled sample, z is the feature representation, ϕ parametrizes the decoder p_ϕ, and θ parametrizes the encoder q_θ.

Finally, the UIR reconstructs the image under the guidance of D_l and learns the feature representation of the sample. Previous experiments have indicated that adding annotation information can improve the performance of active learning models [29]. We use the same STL as in SRAAL, whose objective function is defined as follows:

$$L_{STL} = E\left[\log\left[p_\phi(y_L|z_L)\right] - D_{KL}\big(q_\theta(y_L|x_L)\|p(z)\big)\right]. \quad (3)$$

Because of the dependency on the label of the sample, the STL can only be trained using the labelled sample. Finally, the feature representation learned by the UIR and the

annotation information learned by the STL are concatenated as the final sample representation.

3.2. State Discriminator and State Relabelling.
Considering that the uncertainty score calculated by the OUI will be affected by a tiny value in the category vector, we designed a new indicator function: the Norm-OUI. In this method, the variance of the vector is no longer used to calculate the uncertainty score, but the p-norm of the vector is calculated. Therefore, we redefined the uncertainty score function as follows:

$$\text{score}(x_U) = 1 - \frac{\min \text{norm}(V)}{\text{norm}(V)} \times \max(V), \quad (4)$$

where V is the category vector and $\max(V)$ is the largest element in vector V. Min norm (V) is defined as follows:

$$\min \text{norm}(V) = \left[(\max(V))^p + (C-1)\left(\frac{1 - \max(V)}{1 - C}\right)^p\right]^{1/p}. \quad (5)$$

By definition, min norm (V) is the minimum p-norm for all the vectors whose largest element is $\max(V)$. The objective function of the state discriminator D_2 is as follows:

$$L_D = -E\left[\log\left(D_2\big(q_\theta(z_L|x_L)\big)\right)\right] - E\left[\log\left(\text{score}(x_U) - D_2\big(q_\theta(z_L|x_L)\big)\right)\right], \quad (6)$$

where score (x_U) is the new state value of unlabelled sample x_U. The objective function of the representation generator in the adversarial learning process with D_2 is as follows:

$$L_{adv} = -E\left[\log\left(D_2\big(q_\theta(z_L|x_L)\big)\right)\right] - E\left[\log\left(D_2\big(q_\theta(z_U|x_U)\big)\right)\right]. \quad (7)$$

```
(i)  Input: Initialized selected samples X_S =, unlabelled pool [x_1, x_2, x_3, . . .].
(ii) Hyperparameters: size K, j = 2, k = 0, distance d.
(1)  repeat
(2)    for i = 1 to (j − 1) do
(3)      if S(x_i, x_j) ≤ d then
(4)        k = 1
(5)      end if
(6)      i + +
(7)    end for
(8)    if k == 0 then
(9)      X_S ⟵ X_S ∪ x_j
(10)   end if
(11)   k = 0; j + +
(12) until size(X_S) = K
(13) return the selected samples X_S.
```

ALGORITHM 1: Redundancy removal strategy.

The total objective function for the representation generator is as follows:

$$L_G = \lambda_1 L_{UIR} + \lambda_2 L_{STL} + \lambda_3 L_{adv}, \tag{8}$$

where λ_1, λ_2, and λ_3 are hyperparameters that control the ratio of the function.

3.3. RRM. The purpose of an RRM is to remove redundant samples based on the state value predicted by the state discriminator and the feature distance of the image to reduce the cost of labelling. In this paper, the unlabelled samples are first arranged in descending order according to the predicted values as $[x_1, x_2, x_3, . . .]$, and then the feature representations learned by the representation generator are normalized. Then, the normalized representations are used to calculate the similarity between samples. The similarity between a pair of samples (x_i, x_j) is defined as follows:

$$S(x_i, x_j) = \sqrt{\sum_{k=1}^{n} \left(N(x_i)_k - N(x_j)_k \right)^2}, \tag{9}$$

where $N(x_i)$ and $N(x_j)$ represent the normalized feature representations of samples x_i and x_j, respectively. $N(x_i)_k$ and $N(x_j)_k$ represent the k-th element of the feature representations of $N(x_i)$ and $N(x_j)$, respectively. The RRM is based on the greedy algorithm for redundancy removal, and the specific steps are shown in Algorithm 1. The hyperparameter d is set in the algorithm to control the feature distance between the two samples, and X_S, which is finally returned, is the sample that needs Oracle labelling.

3.4. Sampling Strategy in Active Learning. The algorithm for training the RRAAL algorithm at each iteration is shown in Figure 1. In each iteration, the sampling process is divided into two phases. In the first phase, the generator generates feature representations for each sample, and the state discriminator D_2 predicts the state value of samples under the guidance of the Norm-OUI. In the second stage, we arrange

the unlabelled samples as $[x_1, x_2, x_3, . . .]$ in descending order according to the predicted value, input this sequence into the RRM for sample selection, and finally obtain the samples that need to be labelled. After each iteration, we need to update the task model and the entire active learning model.

4. Experiment

In this section, we evaluate the RRAAL algorithm in both classification and segmentation tasks.

Dataset. The datasets we selected in the classification experiment include CIFAR-10 [43], CIFAR-100 [43], and Caltech-101 [44]. Both CIFAR-10 and CIFAR-100 contain 60,000 images, of which 50,000 are training images and 10,000 are test images. CIFAR-10 has 10 categories with 6,000 images per category, while CIFAR-100 has 100 categories with 600 images per category. Caltech-101 contains 101 image categories and a background category, with a total of 9,146 images, with 40 to 800 images per category. The dataset we selected in the segmentation experiment is Cityscapes [45]; its training set contains 2,975 images, the verification set contains 500 images, and the test set contains 1,525 images.

For each dataset, we randomly sampled $M = 10\%$ samples from the entire dataset as the initial labelled pool X_L, and the remaining 90% of samples formed the initial unlabelled pool X_U. We select $K = 5\%$ samples from the unlabelled pool for labelling each iteration and then position these samples in the labelled pool until the labelled samples reach 40%. For each active learning method, we repeated the experiment five times with a different initially labelled pool and reported the average performance.

Task Model. The task model we used in the image classification experiment is ResNet-18 [46], and the task model we used in the segmentation experiment is a DRN [47]. We compared the average accuracy of the task model in the five experiments.
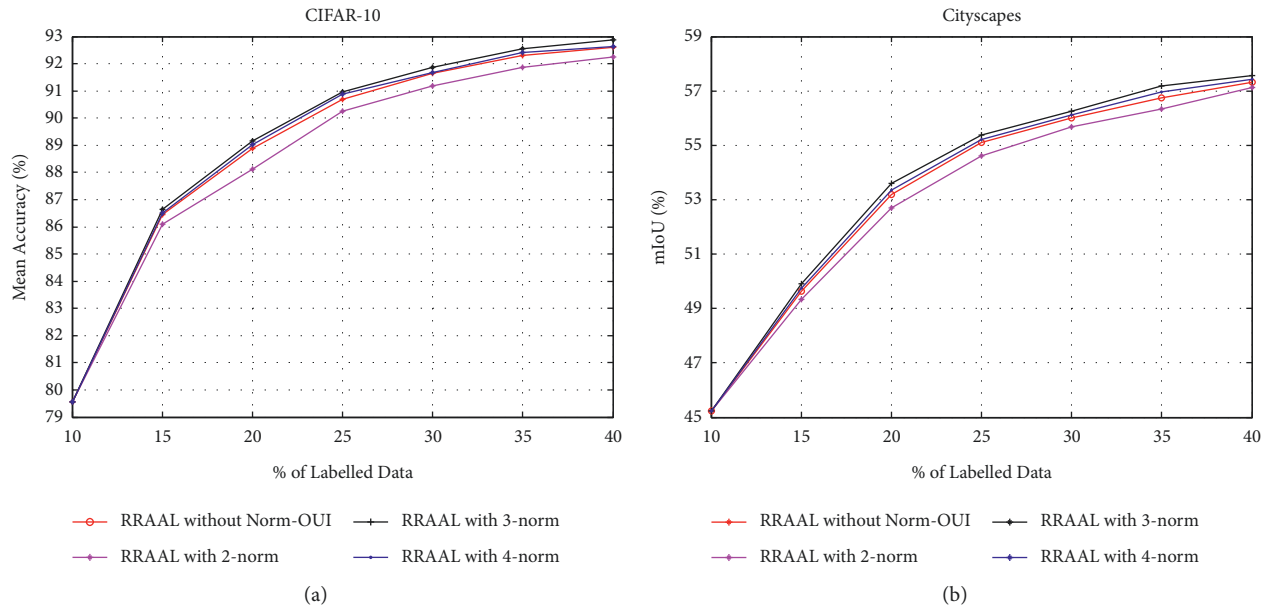
Figure 2: Basic rocket ship design. The rocket ship is propelled with three thrusters and features a single viewing window. The nose cone is detachable upon impact. (a) CIFAR-10. (b) Cityscapes.

4.1. Parameter Analysis. To explore the influence of the *p*-norm on the model performance, we conducted a parameter analysis experiment on the CIFAR-10 and Cityscapes datasets. This section compares the model performance with the 2-norm, 3-norm, and 4-norm in the Norm-OUI. To see the effect of the p-norm on the performance more clearly, we added RRAAL without the Norm-OUI (using the OUI) as a reference. The experimental results of the parameter analysis are shown in Figure 2.

As seen from the experimental results, the performance of RRAAL with the 2-norm is the worst on the two datasets and is lower than that of RRAAL without the Norm-OUI. The performance of RRAAL with the 3-norm and RRAAL with the 4-norm is better than that of RRAAL without the Norm-OUI, and RRAAL with the 3-norm achieves the optimal performance. Therefore, the proposed RRAAL finally uses 3-norm.

4.2. Ablation Study. To evaluate the contribution of the Norm-OUI, RRM, and sample discriminator D_1 introduced in RRAAL, we conducted an ablation study on the CIFAR-10 and Cityscapes datasets. The compared models include RRAAL, RRAAL without the Norm-OUI (using the OUI), RRAAL without the RRM, and RRAAL without both, and SRAAL. It is worth noting that the difference between RRAAL without both and SRAAL is that the former has a sample discriminator D_1.

Figure 3 shows the results of the ablation study. On the CIFAR-10 and Cityscapes datasets, the experimental results show that the overall performance of RRAAL is always better than that of the other methods, and the performance of RRAAL without both is slightly better than that of SRAAL and lower than that of the other three methods.

The experimental results show that (1) the Norm-OUI, RRM, and D_1 can improve the performance of SRAAL; (2) the performance is optimized when the three components are combined.

4.3. Classification Experiment. We compare RRAAL with the current mainstream methods: SRAAL [21], LLAL [36], core-set [37], Monte Carlo dropout (MC dropout) [48], VAAL [22], and Random. Figure 4 shows the experimental results of our proposed RRAAL and other methods on the three datasets. On the CIFAR-10 dataset, RRAAL outperformed the other methods throughout the process. When the data rates were 20%, 30%, and 40%, the mean accuracies of RRAAL were 0.83%, 0.71%, and 0.62%, respectively, higher than those of the second best method (SRAAL). The experimental results show the superiority of RRAAL in datasets with a small number of categories.

The number of categories in the CIFAR-100 dataset is 10 times larger than that in CIFAR-10, which makes the dataset more challenging. On the CIFAR-100 dataset, RRAAL is obviously superior to VAAL, core-set, MC dropout, and Random and slightly superior to SRAAL and LLAL. When the data rates are 20%, 30%, and 40%, the mean accuracies of RRAAL are 0.98%, 1.01%, and 0.98% higher than those of SRAAL and 1.20%, 1.50%, and 1.31% higher than those of LLAL, respectively. Thus, it can be seen that RRAAL still has advantages.

We calculated the final performance of each method on the three datasets, and the results are shown in Table 1. As can be seen from Table 1, compared with other methods, RRAAL achieves the best performance on all three datasets.

In addition, we calculated the computational costs on three datasets. The computational cost of Random is the smallest, so the computational cost of Random is taken as

(a)   (b)
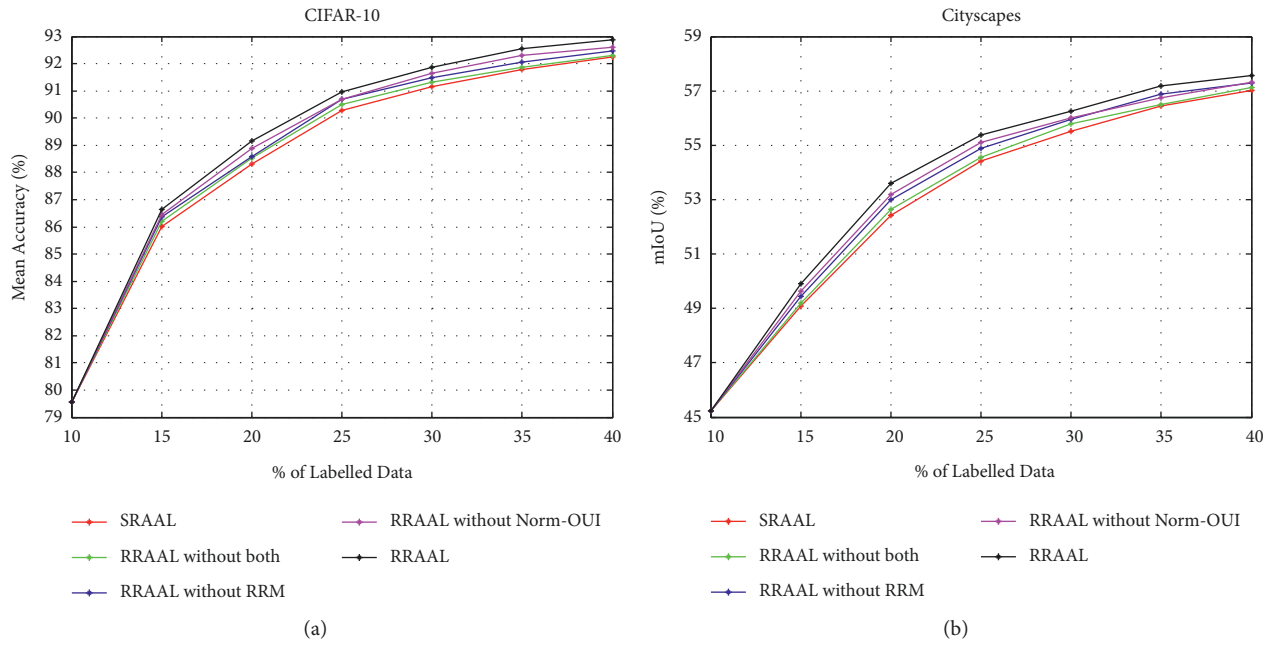
FIGURE 3: Basic rocket ship design. The rocket ship is propelled with three thrusters and features a single viewing window. The nose cone is detachable upon impact. (a) CIFAR-10. (b) Cityscapes.
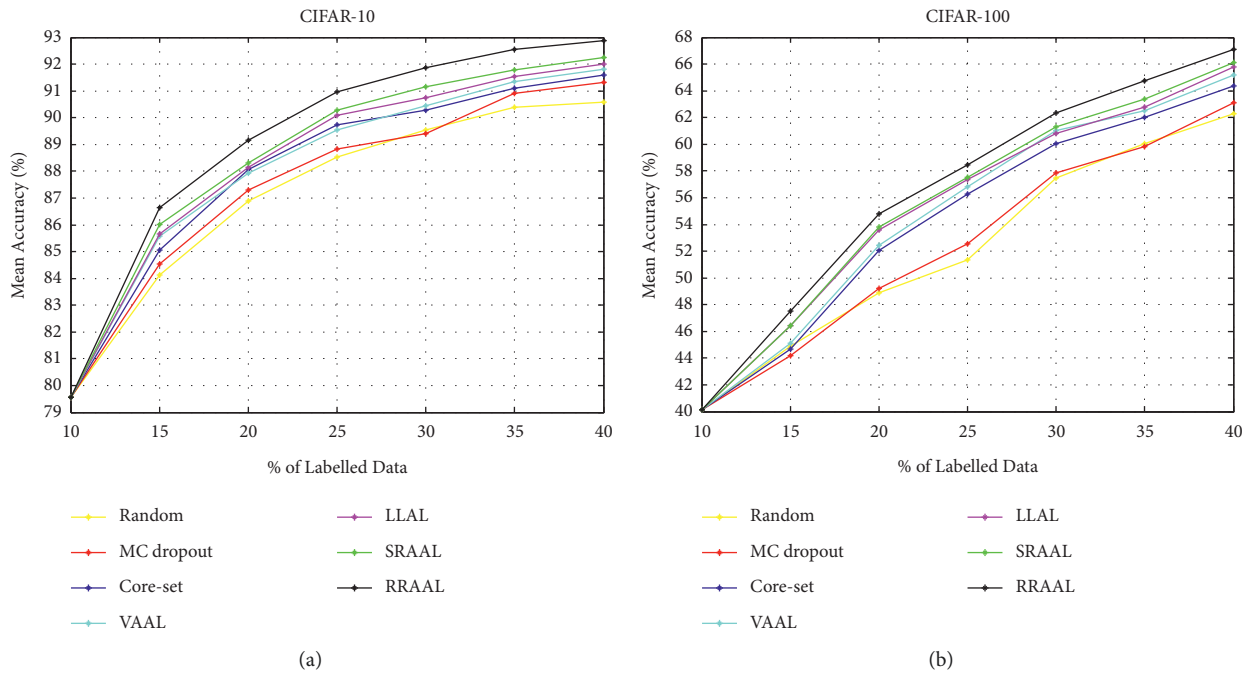

(a)   (b)
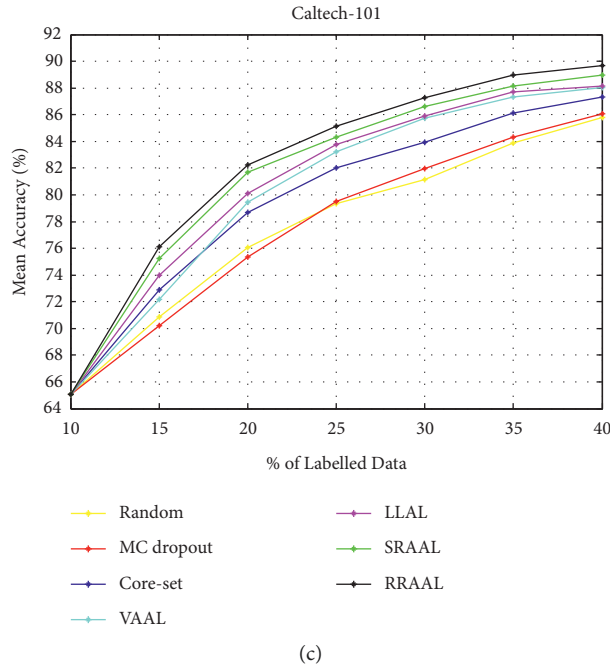
FIGURE 4: Continued.

(c)

FIGURE 4: Active learning results of image classification on the (a) CIFAR-10; (b) CIFAR-100; (c) Caltech-101 dataset.

TABLE 1: Experimental results of RRAAL and other methods on CIFAR-10, CIFAR-100, and Caltech-101.

Methods	CIFAR-10	CIFAR-100	Caltech-101
Random	90.59	62.29	85.79
MC dropout	91.33	63.13	86.09
Core-set	91.59	64.39	87.32
VAAL	91.81	65.21	88.06
LLAL	92.01	65.78	88.13
SRAAL	92.25	66.11	88.95
RRAAL	92.87	67.09	89.58

TABLE 2: Time cost of RRAAL and other methods on CIFAR-10, CIFAR-100, and Caltech-101.

Methods	CIFAR-10	CIFAR-100	Caltech-101
Random	1.00	1.00	1.00
MC dropout	1.05	1.06	1.06
Core-set	1.51	1.54	1.52
VAAL	2.07	2.11	2.03
LLAL	1.28	1.31	1.33
SRAAL	2.15	2.18	2.12
RRAAL	2.21	2.17	2.19

unit 1. The experimental results are shown in Table 2. As can be seen from Table 2, although RRAAL has a higher computational cost, it is very similar to SRAAL and VAAL and achieves higher performance than them. Because the computational cost is much less than the manual cost, RRAAL is still useful.

4.4. Segmentation Experiment. We compare RRAAL with the current mainstream methods: SRAAL [21], VAAL [22], core-set [37], query-by-committee (QBC) [49], MC dropout [48],

and Random. Image segmentation is more challenging than image classification. Figure 5 shows the experimental results of our proposed RRAAL method and the other methods on the Cityscapes dataset. RRAAL has the best performance, and SRAAL and VAAL rank second and third, respectively. The performance of core-set and QBC is similar, which is better than that of MC dropout and Random. When the data rates were 20%, 30%, and 40%, the mIoU of RRAAL was 1.16%, 0.73%, and 0.56 higher than that of SRAAL and 1.70%, 1.23%, and 0.69% higher than that of VAAL, respectively. This result fully verifies the superiority of RRAAL.
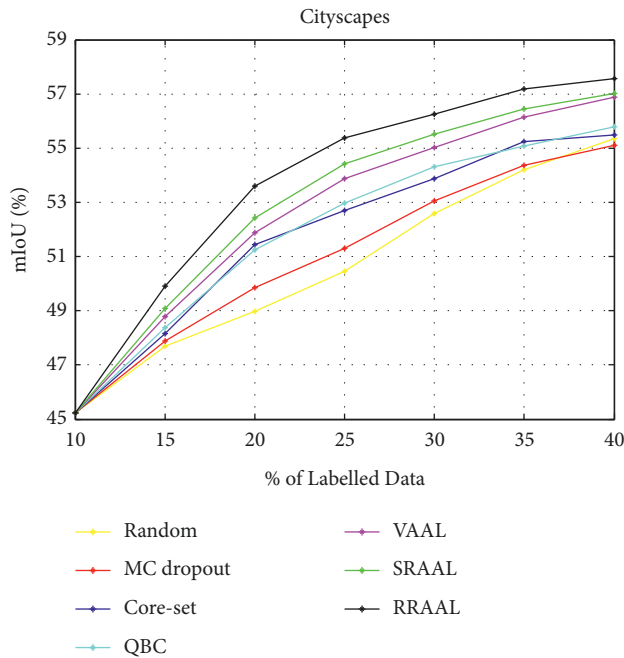
Figure 5: Active learning results of image segmentation on the Cityscapes dataset.

5. Conclusions

In this paper, we first analysed the problems existing in SRAAL, such as an impractical state indicator function and excessive redundancy, and then proposed RRAAL to solve these problems. RRAAL uses the distribution, uncertainty, and redundancy for sample selection and includes a representation generator, a state discriminator, and an RRM. First, we analysed the parameters of the Norm-OUI and selected the 3-norm. Then, we set up an ablation study to verify the contributions of the Norm-OUI, RRM, and sample discriminator D_1. Finally, we verified the effectiveness of RRAAL with classification and segmentation tasks. The performance of RRAAL is 0.62%, 0.98%, and 0.63% higher than that of the state-of-the-art method (SRAAL) in classification datasets and 0.56% higher than that of SRAAL in segmentation datasets. The experimental results show that the overall performance of RRAAL on the four datasets is better than that of the existing mainstream methods.

Data Availability

The data used to support this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] C. Li, J. Guo, and C. Guo, "Emerging from water: underwater image color correction based on weakly supervised color transfer," *IEEE Signal Processing Letters*, vol. 25, no. 3, pp. 323–327, 2018.

[2] B. P. Abbott, R. Abbott, T. D. Abbott et al., "Observation of gravitational waves from a binary black hole merger," *Physical Review Letters*, vol. 116, no. 6, Article ID 061102, 2016.

[3] C. Li, C. Guo, J. Guo, P. Han, H. Fu, and R. Cong, "PDR-Net: Perception-inspired single image dehazing network with refinement," *IEEE Transactions on Multimedia*, vol. 22, no. 3, pp. 704–716, 2019.

[4] D. P. Kingma, S. Mohamed, D. J. Rezende, and M. Welling, "Semi-supervised learning with deep generative models," *Advances in Neural Information Processing Systems*, vol. 1, pp. 3581–3589, 2014.

[5] Y. Ouali, C. Hudelot, and M. Tami, "An Overview of Deep Semi-supervised Learning," in *Proceedings of the International Conference on Machine Learning*, vol. 1, July 2020.

[6] N. Doulamis and A. Doulamis, "Semi-supervised deep learning for object tracking and classification," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pp. 848–852, Paris, France, October 2014.

[7] D. H. Lee, "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks," in *Proceedings of the 2013 ICML Workshop on challenges in representation learning*, vol. 3, no. 2, Article ID 896, Atlanta, USA, June 2013.

[8] G. Niu, W. Jitkrittum, B. Dai, H. Hachiya, and M. Sugiyama, "Squared-loss mutual information regularization: a novel information-theoretic approach to semi-supervised learning," in *Proceedings of the 2013 International Conference on Machine Learning*, pp. 10–18, Atlanta, USA, June 2013.

[9] M. Wang, W. Fu, S. Hao, D. Tao, and X. Wu, "Scalable semi-supervised learning by efficient anchor graph regularization," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 7, pp. 1864–1877, 2016.

[10] E. Protopapadakis, A. Doulamis, N. Doulamis, and E. Maltezos, "Stacked autoencoders driven by semi-supervised learning for building extraction from near infrared remote sensing imagery," *Remote Sensing*, vol. 13, no. 3, p. 371, 2021.

[11] X. Zhai, A. Oliver, A. Kolesnikov, and L. Beyer, "S4l: self-supervised semi-supervised learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1476–1485, Seoul, Korea, November 2019.

[12] A. Esteva, B. Kuprel, R. A. Novoa et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[13] Q. She, K. Chen, Z. Luo, T. Nguyen, T. Potter, and Y. Zhang, "Double-Criteria Active Learning for Multiclass Brain-Computer Interfaces," *Computational intelligence and neuroscience*, vol. 2020, 2020.

[14] J. Guo, Z. Pang, M. Bai, P. Xie, and Y. Chen, "Dual Generative Adversarial Active Learning," *Applied Intelligence*, vol. 51, pp. 5953–5964, 2021.

[15] W. H. Beluch, T. Genewein, A. Nürnberger, and J. M. Köhler, "The power of ensembles for active learning in image classification," in *Proceedings of the IEEE Conference on Computer*

Vision and Pattern Recognition (CVPR), pp. 9368–9377, Salt Lake City, UT, USA, June 2018.

[16] Y. Gal, R. Islam, and Z. Ghahramani, "Deep bayesian active learning with image data," in *Proceedings of the International Conference on Machine Learning*, pp. 1183–1192, Sydney, Australia, August 2017.

[17] Y. Sun, J. Zhang, and Y. Zhang, "Multi-sensor image classification based on active learning," in *Proceedings of the 2012 5th International Congress on Image and Signal Processing*, pp. 1290–1293, Chongqing, Sichuan, China, October 2012.

[18] D. Tuia, F. Ratle, F. Pacifici, M. F. Kanevski, and W. J. Emery, "Active learning methods for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2218–2232, 2009.

[19] S. D. Jain and K. Grauman, "Active image segmentation propagation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2864–2873, Las Vegas, NV, USA, June 2016.

[20] A. Vezhnevets, J. M. Buhmann, and V. Ferrari, "Active learning for semantic segmentation with expected change," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3162–3169, Providence, RI, USA, June 2012.

[21] B. Zhang, L. Li, S. Yang, S. Wang, Z. J. Zha, and Q. Huang, "State-relabeling adversarial active learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8756–8765, Seattle, WA, USA, June 2020.

[22] S. Sinha, S. Ebrahimi, and T. Darrell, "Variational adversarial active learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5972–5981, Seoul, Korea, November 2019.

[23] I. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," in *Proceedings of the International Conference on Neural Information Processing Systems (NIPS)*, pp. 2672–2680, Montreal, Canada, December 2014.

[24] J. Guo, Z. Pang, F. Yang, J. Shen, and J. Zhang, "Study on the method of fundus image generation based on improved gan," *Mathematical Problems in Engineering*, vol. 2020, 2020.

[25] J. J. Zhu and J. Bento, "Generative adversarial active learning," in *Proceedings of the International Conference on Machine Learning*, Sydney, Australia, August 2017.

[26] T. Tran, T. T. Do, I. Reid, and G. Carneiro, "Bayesian generative active deep learning," in *Proceedings of the International Conference on Machine Learning*, Long Beach, CA, USA, June 2019.

[27] D. Wu, "Pool-based sequential active learning for regression," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 5, pp. 1348–1359, 2018.

[28] Y. Geifman and R. El-Yaniv, "Deep active learning over the long tail," *Computer Science Bibliography*, vol. 1, 2017.

[29] A. J. Joshi, F. Porikli, and N. Papanikolopoulos, "Multi-class active learning for image classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2372–2379, Miami, FL, USA, June 2009.

[30] A. Kapoor, K. Grauman, R. Urtasun, and T. Darrell, "Active learning with Gaussian processes for object categorization," in *Proceedings of the IEEE Conference on International Conference on Computer Vision*, pp. 1–8, Rio de Janeiro, Brazil, October 2007.

[31] H. Wang, X. Gao, K. Zhang, and J. Li, "Single-image super-resolution using active-sampling Gaussian process regression," *IEEE Transactions on Image Processing*, vol. 25, no. 2, pp. 935–948, 2015.

[32] K. Brinker, "Incorporating diversity in active learning with support vector machines," in *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 59–66, Washington, DC, USA, August 2003.

[33] Z. Wang and J. Ye, "Querying discriminative and representative samples for batch mode active learning," *ACM Transactions on Knowledge Discovery from Data*, vol. 9, no. 3, pp. 1–23, 2015.

[34] P. Chen, J. Xia, B. M. Merrick, and T. J. Brazil, "Multiobjective Bayesian optimization for active load modulation in a broadband 20-W GaN Doherty power amplifier design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 3, pp. 860–871, 2016.

[35] M. Kaselimi, N. Doulamis, A. Doulamis, A. Voulodimos, and E. Protopapadakis, "Bayesian-optimized bidirectional LSTM regression model for non-intrusive load monitoring," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2747–2751, Brighton, UK, May 2019.

[36] D. Yoo and I. S. Kweon, "Learning Loss for Active Learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 93–102, Convention center in Long Beach, CA, USA, June 2019.

[37] O. Sener and S. Savarese, "Active learning for convolutional neural networks: a core-set approach," in *Proceedings of the 2017 Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, July 2017.

[38] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," in *Proceedings of the 2017 Conference on Machine Learning*, Atlanta, USA, June 2013.

[39] M. Mirza and S. Osindero, "Conditional generative adversarial nets," in *Proceedings of the 2014 Conference on Machine Learning*, Beijing, China, June 2014.

[40] T. Tran, T. Pham, G. Carneiro, L. Palmer, and I. Reid, "A bayesian data augmentation approach for learning deep models," in *Proceedings of the International Conference on Neural Information Processing Systems (NIPS)*, pp. 2797–2806, Long Beach, CA, USA, December 2017.

[41] Y. Gal, R. Islam, and Z. Ghahramani, "Deep bayesian active learning with image data," in *Proceedings of the International Conference on Machine Learning*, pp. 1183–1192, Sydney, Australia, August 2017.

[42] A. Mottaghi and S. Yeung, "Adversarial representation active learning," in *Proceedings of the 2019 Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, June 2019.

[43] A. Krizhevsky and G. Hinton, *Learning Multiple Layers of Features from Tiny Images*, Technical Report-0, Toronto Univercity, Toronto, Ontario, 2009.

[44] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: an incremental bayesian approach tested on 101 object categories," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, p. 178, Washington, DC., USA, June 2004.

[45] M. Cordts, M. Omran, S. Ramos et al., "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3213–3223, Las Vegas, Nevada, USA, June 2016.

[46] H. Jung, M. K. Choi, J. Jung, J.-H. Lee, S. Kwon, and W. Y. Jung, "ResNet-based vehicle classification and localization in traffic surveillance systems," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*

Workshop (CVPRW), pp. 61–67, Honolulu, HI, USA, July 2017.

[47] F. Yu, V. Koltun, and T. Funkhouser, "Dilated residual networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 472–480, Honolulu, HI, USA, June 2017.

[48] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: representing model uncertainty in deep learning," in *Proceedings of the International Conference on Machine Learning*, pp. 1050–1059, New York, NY, USA, June 2016.

[49] W. Kuo, C. Häne, E. Yuh, P. Mukherjee, and J. Malik, "Cost-sensitive active learning for intracranial hemorrhage detection," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 715–723, Granada, Spain, September 2018.