MAKING THE TORCH LIGHTER: A REINFORCED ACTIVE SAMPLING FRAMEWORK FOR IMAGE CLASSIFICATION

Peng Liu, Zhipeng Ye, Xianglong Tang, Wei Zhao

School of Computer Science and Technology
Harbin Institute of Technology, Harbin 150001, PR. China
{yezhipeng, pengliu, tangxl, zhaowei}@hit.edu.cn

ABSTRACT

In this paper, we aim to construct a more reasonable and effective active sampling model, named as reinforcement uncertainty sampling with bag-of-visual-words (RUSB). Compared with traditional active sampling strategy based on uncertainty, both certainty metric and sample post-processing are introduced for better performance. The certainty metric is measured by the bag-of-visual-words (BoVW) classification model in order to entirely evaluate samples, and the post-processing module is driven by the Q-learning method to construct a compact and efficient training set for the BoVW module. The performance of BoVW is used to initialize and determine the status of the post-processing module during the process of iteration. Meanwhile, the weight of the measurement is associated with each iteration instead of being set manually. Experimental results on real world datasets show the effectiveness of the proposed framework.

***Index Terms*—** Image classification, sample selection, bag-of-visual-words, entropy-based sampling, Q-learning

1. INTRODUCTION

Image classification is a fast developing field, and is commonly solved by the technique of machine learning. Although large amount of data is available in the age of big data, there still exists an inevitable problem that unlabeled data is abundant but labeling is expensive in many situations. Thus active learning is proposed to deal with this problem. The active sampling engine selects unlabeled instances according to a certain strategy. After decades of development, uncertainty sampling based on entropy[1] has been proven an effective strategy. However, it is reported that traditional sampling methods do not reflect the desired training distribution, leading to the result that additional labeling work should be done[2]. Thus the NUSB active sampling strategy is proposed to solve this problem by introducing metric based on category distribution for better sampling results[3]. There are three shortcomings of this model. First of all, the weight of the measurement is not associated with the learning process, reducing the adaptivity of the strategy. In the second place, the importance of a selected sample is ignored, and the selected samples will always be added to the labeled set to retrain the BoVW-based classifier. This will introduce redundancy to the labeled set. Last but equally important, the negative-accelerated principle is not able to model the performance variation during the learning process since it is based on the assumption that the performance of learning increases with iteration. Thus in this paper we propose an improved learning framework to further improve the performance and adaptivity of the NUSB sampling framework by introducing a post-processing module based on Q-learning to schedule the the weight of metric modules. Also the importance of the selected samples are evaluated from the perspective of cognition to compact the scale of the training set for the visual module. The structure of the proposed framework is summarized in **Fig. 1**.

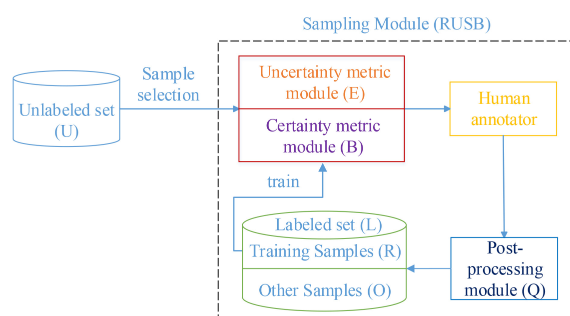

Fig. 1: Components of the proposed active sampling framework.

Having a brief introduction of this research, before delving into details of our method, we highlight the structure of the paper. In the next section the proposed framework is described. Experimental results and analysis are shown in Section 3. Conclusions are presented in Section 4.

This research is supported by National Science Foundation of China (Grant No. 61671175) and Natural Science Foundation of Heilongjiang Province, China (Grant No. F201012).

2. REINFORCEMENT SAMPLING FRAMEWORK FOR IMAGE CLASSIFICATION

In this section, we will respectively introduce each module through Section 2.1∼2.3, then summarized the whole learning framework in Section 2.4.

2.1. Uncertainty sampling strategy

Traditional sampling criterion based on uncertainty selects samples that the learning module do not have sufficient knowledge on, which is helpful to choose to be labeled. The entropy-based sampling strategy is given by eq. (1),

$$D_e = -\sum_{i=1}^{n} P(y_i|z) \log P(y_i|z) \tag{1}$$

where $P(y_i|z)$ is the estimated probability of sample y_i given prior z, and n is the scale of unlabeled pool.

2.2. BoVW classification method

The certainty measurement is determined by the category of samples. Here one of the most famous learning techniques, bag-of-visual words (BoVW), is introduced to estimate the category of samples. Nowadays, BoVW has become one of the most commonly used approaches in image classification, whose simplicity and effectiveness has been tested during these years. In BoVW, images are represented as visual words obtained by clustering local features such as DoG or SIFT, to train the classifier. The certainty measurement is given by eq. (2),

$$\begin{aligned} c &= \arg\max_{i=1}^{n} P_i^A \\ D_s &= \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(Y - Y_i^c\right)^2} \end{aligned} \tag{2}$$

where P_i^A is the response of the classifier according to certain bag-of-features, c is the estimated category of the testing sample, m is the number of samples in c, Y_i^c is the i-th labeled sample of c. Here D_s corresponds to the mean square error (MSE)[4] between current sample and the samples of corresponding category in the labeled set. By taking MSE of labeled set into consideration, the representativeness is introduced into the model to improve the effectiveness of the proposed measurement.

2.3. Sample post-processing

According to the learning principle[5], the performance of a subject will not increase until samples with new knowledge are proceeded, including those that cannot be handled at first. Therefore for the sample selection problem, it is also important to retrain the classifier with samples that are previously incorrectly classified, and ignore those correctly classified since they make little contribution to improving the performance of the classifier. This phenomenon was ignored by the NUSB sampling framework and the selected samples will always be added to the training set, introducing redundancy and lower the efficiency of the learning progress. Meanwhile, the decaying rate of the weight of the metric for sample selection is independent from the status of learning. It is manually set the learning process begins, which is not appropriate in practice. Thus in this paper, we introduce the reinforcement learning (RL) based sample post-processing module Q to solve these problems.

The aim of RL is to model optimization problems for artificial systems, and learn to predict and optimize their behavior in environments by transition of states. It provides both qualitative and quantitative frameworks to model decision-making problems under the constraint of rewards[6]. The classical reinforcement learning algorithm considers the problem of an agent interacting with its environment while trying to maximize total reward accumulated over time. The environment is modeled as a Markov decision process, including the set of states S, the set of actions a, the reward function R, and the transition function T[7]. Among the learning methods, Q-learning[8] has been successfully utilized in many fields such as robot path planning[9] and visual recognition[10]. The algorithm runs iteratively, aiming to choose the state of next iteration from S according to the value of Q-table and the reward of taking an action. The Q-learning rule is:

$$Q(s,a) = R(s,a) + \gamma \cdot \max_{\tilde{a}}\{Q(\tilde{s},\tilde{a})\} \tag{3}$$

, where $\tilde{s}$ and $\tilde{a}$ are the state of next iteration and corresponding action, γ is the learning rate.

To solve the problems mentioned above, inspired by the work of Epshteyn[11], the proposed framework allows domain experts to specify possibly inaccurate models of the world, and uses it as a blueprint for exploration. The certainty metric module is naturally treated as the inaccurate model, and the planning of RL module is made according to the observation received from the certainty metric module. Based on this principle, the variation of performance is modeled in correspondence with the status set S of the Q-learning method. The state-action table is defined in **Table 1**, where $S = \{S_1, S_2\}$. S_1 and S_2 respectively correspond to the rise and drop of the performance of the certainty metric module B. The labeled set, L, is divided into two parts, including the training samples R and other samples O. R is used to retrain B after each iteration. Action a_1 and a_2 are designed to accept and reject the selected samples to be added to the training set R according to whether they are correctly or incorrectly classified. The rejected samples will be simply added to O, and the incorrectly classified samples will be added to R. Then the classifier of B is retrained with R at the end of each iteration. Parameters of Q will also be refined with the iteration proceeds.

Table 1: The state-action table of the RL module.

	a_1=accept	a_2=reject
S_1	$Q(S_1, a_1)$	$Q(S_1, a_2)$
S_2	$Q(S_2, a_1)$	$Q(S_2, a_2)$

2.4. The RUSB sampling model

The measurement of sample selection is given by eq. (4),

$$D = \beta D_e + (1 - \beta) D_s \tag{4}$$

where $\beta \in [0, 1]$ is utilized to adjust the importance of the two measurements. With the undergo of iterations, the performance of the certainty metric module B is boosted, and it should have a larger weight to influence the whole measurement. Thus β is a decaying weight, and the decaying speed is set according to the Q-table and current status of the learning module in each iteration according to the performance of the classifier instead of being manually set as a fixed value, which is one of the obvious difference between other active sampling framework such as NUSB. The adjustment of β is given in eq. (5).

$$\beta = \beta + \alpha\sigma \tag{5}$$

$$\alpha = \begin{cases} -1, \tilde{S} = S_1 \\ 1, \tilde{S} = S_2 \end{cases} \tag{6}$$

$$\sigma = \frac{1}{\ln(Q(\tilde{S}, \tilde{a}))} \tag{7}$$

The sampling algorithm corresponds to **Fig. 1** is summarized in **Algorithm. 1**. Compared with previous active sampling frameworks, combining certainty metric provided by visual classifier with entropy measurement entirely evaluate samples and lights the way for the sampling process. Meanwhile, the introduction of the sample post-processing module lighter the torch, and makes the learning process more reasonable and effective by dynamically adjust parameter and construct more compact training set with RL technique.

Algorithm 1 The RUSB active sampling framework

Require: Initial labeled set L', unlabeled set U.
Ensure: Labeled set L.
//Initialization:
1. Train C with L'.
//Sample selection & post-processing
2. Evaluate each sample of U by the metric given by eq.(4).
3. Select up to $k\%$ samples for the human annotator. Send the annotated samples to the RL module, and update the Q-table by eq.(3).
4. Update the weight β, and move the selected samples to either R or S.
5. Train C with R.
6. Goto step 3 until the stop criterion is satisfied.
Return L.

3. EXPERIMENTS AND ANALYSIS

In this section, we report the experimental results of every sampling strategies. The testing dataset is constructed from three popular visual datasets, including MSRC[12], Caltech-101[13] and PASCAL VOC 2007[14]. Also, images from the Google image search engine were supplied. The proposed framework is compared with relevant and representative strategies in this area, including traditional entropy-based sampling strategy[15], MLA[16], and the NUSB[3]. Constitution of the constructed dataset are shown in **Table 2**. The reasons of constructing the training and testing dataset from various datasets are[17]: (i) due to the limited availability of human labeled training data that stay in the same feature space or under the same distribution cannot be guaranteed to be sufficient enough to avoid the over-fitting problem. Although MSRC is a good training set, the remaining samples could not sufficiently reflect difference of performance between strategies due to the limitation on the number of samples. Other three image datasets contains much more various and richer samples than MSRC, which is an ideal supplement to the testing data, making it possible to have a more detailed test, and (ii) including related data in a different domain can expand the availability of our knowledge about the target future data. The stop criterion investigated by NUSB is adopted.

Table 2: Categories of MSRC and constitution of dataset for transfer test.

Categories of MSRC	VOC 2007	Caltech-101	Google Image
Sheep		√	
Human		√	
Dog		√	
Cat		√	
Tree		√	
Flower		√	
Chair		√	
Car		√	
Sign			√
House			√
Book			√
Plane	√		
Cow	√		
Boat	√		
Bicycle	√		
Bird	√		

The value of the Q-table are initialized to 100, and the reward of state transition is $R = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. One-versus-all strategy is utilized for training SVM classifiers of BoVW, and the vocabulary size is 1000. SIFT[18] is utilized as local descriptor. For each iteration, top $k\%$ samples are selected and sent to the post-processing module. Here $k = 20$. The performance of sampling strategies are evaluated by the classification results trained by the samples selected from different strategies. Experimental results are given below.

The effectiveness of sample selection strategies are reflected by the overall performance of the visual classifier trained with selected samples. BoVW is chosen as the visual classifier for all sampling frameworks. The results are given in **Fig. 2**. We can see that the performance of RUSB sampling framework outperforms other strategies and achieves better performance. With the undergo of iteration, the difference of performance between RUSB and traditional entropy-based strategy becomes larger. These are due to the fact that RUSB introduces certainty metric to entirely evaluate samples that makes the sampling process more effective. Also, there is a feedback between the visual classifier and the parameter adjustment of each iteration, making the learning process more reasonable. The results show the fact that combination of category distribution and sampling strategy is effective in improving precision of image classification, which is consistent with existing research[2]. Meanwhile, the RUSB sampling framework is independent from specific visual classifier due to its modular structure, and could achieve even better performance when the if more advanced classifier is adopted in the certainty metric modular.

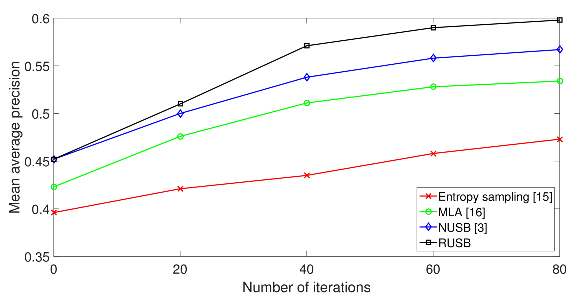

Fig. 2: The overall performance of all active sampling strategies.

Fig. 3 shows the number of selected samples with the proceeding of iterations, where $\frac{NSR}{NSS}$ is the ratio between the samples added to the training set R and the total number of selected samples. From the result we can see that the number of samples added to R decreases with the iteration proceeds. It indicates the fact that with the raise of the performance of the visual classifier in certainty metric module B, more correctly classified samples are added to O instead of R, thus the scale of R is compacted so that the efficiency of the learning process is improved.

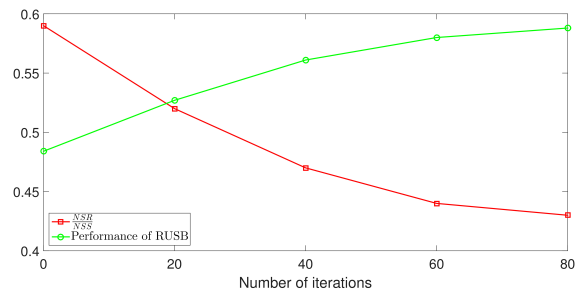

Fig. 3: The number of selected samples with the proceeding of iterations.

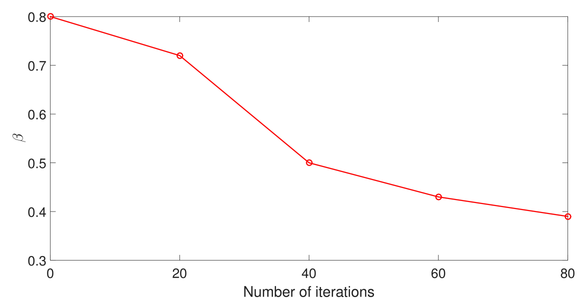

Fig. 4: Variation of β during the sampling process.

Fig. 4 reflects the variation of β during the sampling process. From the result we can see that the dropping speed of β varies at different stages of the sampling process. At the early stage (the number of iterations is less than 20) and the end of the learning progress (the number of iterations is larger than 60), the speed is slow, while in the middle stage (the number of iterations is larger than 20 but less than 60), the weight drops quickly, indicating the fact that the framework will dynamically adjust the weight of the certainty measurement module in evaluation according to the status of the visual classifier reflected by its performance. This phenomenon is consistent with the result given in **Fig. 2**. Therefore associating the value of β in eq. (4) is effective in boosting the effectiveness of the sample selection strategy.

4. CONCLUSION

In this paper, a reinforced active sampling framework named RUSB is proposed from the perspective of cognition for sample selection. Besides traditional uncertainty metric based on entropy, certainty metric estimated from sample category information is introduced to entirely evaluate samples. A new sample post-processing module based on Q-learning is proposed to improve the flexibility and efficiency of the framework. Experimental results have proven the effectiveness of the proposed sampling framework.

5. REFERENCES

[1] Burr Settles and Mark Craven, "An analysis of active learning strategies for sequence labeling tasks," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2008, pp. 1070–1079.

[2] W. Wu, Y. Liu, and M. Guo, "Constructing training distribution by minimizing variance of risk criterion for visual category learning," in *IEEE International Conference on Image Processing*, 2012, pp. 101–104.

[3] Z. Ye, P. Liu, X. Tang, and W. Zhao, "May the torcher light our way: A negative-accelerated active learning framework for image classification," in *IEEE International Conference on Image Processing*, 2015, pp. 1658–1662.

[4] Richard Dosselmann and XueDong Yang, "A comprehensive assessment of the structural similarity index," *Signal, Image and Video Processing*, vol. 5, no. 1, pp. 81–91, 2011.

[5] Robert L Solso, M Kimberly MacLin, and Otto H MacLin, *Cognitive psychology*, Allyn & Bacon, 2005.

[6] Peter Dayan and Yael Niv, "Reinforcement learning: The good, the bad and the ugly," *Current Opinion in Neurobiology*, vol. 18, no. 2, pp. 185 – 196, 2008.

[7] Apostolos N. Burnetas and Michael N. Katehakis, "Optimal adaptive policies for markov decision processes," *Mathematics of Operations Research*, vol. 22, no. 1, pp. 222–255, 1997.

[8] Christopher J. C. H. Watkins and Peter Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3, pp. 279–292, 1992.

[9] A. Konar, I. Goswami Chakraborty, S. J. Singh, L. C. Jain, and A. K. Nagar, "A deterministic improved q-learning for path planning of a mobile robot," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, no. 5, pp. 1141–1153, 2013.

[10] Lucas Paletta, Gerald Fritz, and Christin Seifert, "Q-learning of sequential attention for visual object recognition from informative local descriptors," in *Proceedings of the 22Nd International Conference on Machine Learning*, 2005, pp. 649–656.

[11] Arkady Epshteyn, Adam Vogel, and Gerald DeJong, "Active reinforcement learning," in *Proceedings of the 25th International Conference on Machine Learning*, 2008, ICML '08, pp. 296–303.

[12] S. Savarese, J. Winn, and A. Criminisi, "Discriminative object class models of appearance and shape by correlatons," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, vol. 2, pp. 2033–2040.

[13] Li Fei-Fei, Rob Fergus, and Pietro Perona, "Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories," *Computer Vision and Image Understanding*, vol. 106, no. 1, pp. 59 – 70, 2007.

[14] Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman, "The pascal visual object classes (voc) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.

[15] A. Holub, P. Perona, and M. C. Burl, "Entropy-based active learning for object recognition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2008, pp. 1–8.

[16] Xin Li and Yuhong Guo, "Multi-level adaptive active learning for scene classification," in *Proceedings of European Conference on Computer Vision*, 2014, pp. 234–249.

[17] L. Shao, F. Zhu, and X. Li, "Transfer learning for visual categorization: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 5, pp. 1019–1034, 2015.

[18] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.