2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Improving PGF Retrieval Effectiveness with Active Learning

Jingwei Qu*, Xiaoqing Lu*†, Songping Fu*, Zhi Tang*†
*Institute of Computer Science & Technology, Peking University, Beijing, P.R.China
†State Key Laboratory of Digital Publishing Technology, Beijing, P.R.China
Email: {qujingwei, lvxiaoqing, fusongping, tangzhi}@pku.edu.cn

***Abstract*—Multimedia education is playing a significant and increasing role for education purposes, thus leading to a large number of electronic documents. Plane geometry figures (PGFs), as important components of these documents, are regarded as very helpful information to most retrieval systems in the field of mathematics education. However, the burdensome work of annotation has become one of the chief obstacles to improve the efficiency of retrieval systems. In this paper, we introduce an active learning-based frame to select candidate instances for training the classifiers in retrieval systems, which are an emerging non-text-based information systems. In addition, an enhanced uncertainty measure and the selection of specific features of PGFs are proposed for our active learning algorithm. Comparative experiment results indicate that the proposed method effectively improves the performance of the PGF retrieval system and reduces the burdensome annotation workload.**

I. Introduction

In the field of mathematics education [1], graph-based document retrieval is attracting more attention than traditional keyword-based retrieval with substantial electronic documents. In fact, graph-based document retrieval has gradually become a topic of concern in a number of domains, such as instructional figures, architectural drawings, circuit diagrams, engineering designs, chemical formulas, business charts and trademarks, among others. However, the burdensome dataset annotation is an obstacle to the performance improvement of most machine learning-based algorithms in these retrieval systems. Consequently, the means to select the effectual data to be labeled for training has become a critical problem. Plane geometry figures (PGFs) are very important components of mathematics documents, as shown in Figure 1, the retrieval of PGFs is an emerging non-text-based information system in the mathematics education field. In this paper, we focus on PGF retrieval and propose an active learning-based frame to elaborately select candidate PGFs for training the classifiers in our PGF retrieval system. Several methods have been proposed to retrieve geometric contents directly via PGFs rather than traditional keywords. In these retrieval systems, shape features are adopted to achieve highly accurate search results. However, finding distinguishable features and efficient classifiers depends on exhaustive annotation and selection of effectual instances. Manually constructing a training database is an arduous task; it even results in a formidable obstacle for further research. To obtain a desired training dataset via active learning model, we are faced with at least

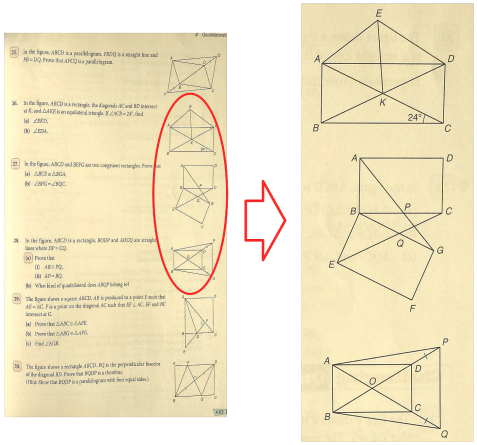


Fig. 1. Some examples of PGFs in a geometry textbook (Mathematics in Action, Copyright ©1997-2015 Pearson Education Asia Limited).

three challenges: 1) The annotation of PGFs involves several aspects for consideration, including the dominant elements, graph structure, educational purposes, and related knowledge points. 2) The limited types of geometry elements, such as triangles, rectangles, and circles, result in a highly similar PGF for human visual perception. Selecting the proper label for these almost indistinguishable PGFs requires deliberation. 3) Relatively compact PGF structures, which involve various complicated spatial relationships among geometric elements, present deep-seated obstacles to the development of effective classification methods for instance selection.

Our approach is based on an active learning-based model combining with multi-label classification. An enhanced uncertainty measure and the selection of specific features of PGFs are proposed for our active learning algorithm. A multi-class hierarchical label system rather than a simple label set is adopted for multi-label classification. The experiments demonstrate that the high-quality examples obtained with the proposed approach significantly improve the efficiency of training classifiers with relatively few annotations.

The rest of this paper is organized as follows: Section 2 discusses existing methods of PGF retrieval and active learning. Section 3 introduces the framework of the proposed approach. Section 4 describes the active learning-based selection algorithm in particular. Section 5 presents our experiments and evaluation results. Section 6 concludes this paper.

978-1-5090-4846-5/16/$31.00 ©2016 IEEE

II. Related Work

A. Shape-based Image Retrieval

Shape-based image retrieval has gradually attracted considerable attention from many fields, such as electronic circuits [2], floor plans [3], trademarks [4], symbols [5], sketch [6], and document classification [7]. The graph recognition and retrieval problem involve different shape descriptors based on skeleton, boundary, local region, or whole image. Shape description techniques generally are categorized into two groups [8]: contour-based [9] and region-based [10] shape descriptors, which represent shapes based either on interior region or on boundary. Several symbol retrieval systems are similar to PGF retrieval. However, these symbol retrieval methods ignore the information of basic primitives.

The PGF retrieval is a very efficient approach to search related contents in the field of mathematics education. [11] proposes a method to extract rectangles, parallelograms, and trapezoids in PGFs. A PGF retrieval method based on bag of shapes is described in [12]. Seo *et al*. [13] present a method for diagram understanding method that identifies the visual elements in a diagram while maximizing the agreement between textual and visual data. Some techniques are available for the detection of individual geometric shapes, including triangle, rectangle, and circle. Most triangle detection methods are applied in traffic sign detection [14]. Circles detectors [15] are used for natural images and document graphs. Rectangle detection methods are generally primitive-based [16] or Hough/Radon transform-based techniques [17]. However, the lack of spatial relation between basic primitives is common among these methods.

B. Active Learning

Settles [18] provides a general introduction to active learning and a review of several active learning algorithms. Tong and Koller [19] propose an active learning algorithm, called SIMPLE, with SVMs, which uses the current SVM classifier to query the instance closest to the decision hyperplane (in kernel space). Zhang *et al*. [20] propose a novel online semi-supervised active learning framework for object classification in traffic scene surveillance, which combines the online, active, and semi-supervised learning. An active learning method [21] automatically balances between the exploitation and the exploration trade-off, and measures the effect of a new example on the current model by deriving model changes of Gaussian process models in closed form. Liang and Grauman [22] explore active learning strategies for training relative attribute-ranking functions and introduces a novel criterion that requests a partial ordering for a set of examples. These methods mainly adopt uncertainty sampling as active learning strategy to choose the instance with the most uncertainty to label. Although this strategy works well in many conditions, it only focuses on the informativeness of a candidate instance over the current classifier and ignores the relationship between the candidate instance and the remaining unlabeled instances.

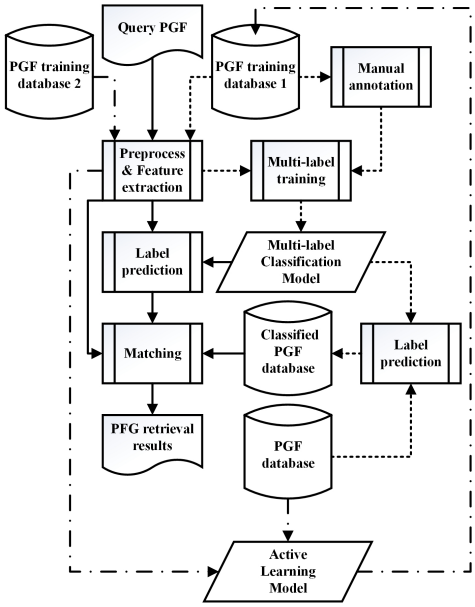


Fig. 2. The proposed framework for PGF retrieval system.

III. Framework of The PGF Retrieval

As shown in Figure 2, the proposed model mainly consists of two modules, namely, classification and matching. The classification of PGFs involves preprocessing, feature extraction, and multi-label training of a classifier. An effective multi-label learning method and supporting features are explored and adopted. The model is trained with ground truth and then used to assign suitable labels automatically for all the PGFs in the database offline. For a query PGF, through online prediction of its labels, the relative candidate PGFs can be filtered according to at least one common label shared by the PGFs based on the multi-label classification results. To quickly respond to the PGF query in the first workflow, the PGF matching is performed. PGFs are retrieved among the candidate PGFs rather than in the entire database. Elaborate features are used to accurately compare the query PGF and every candidate PGF. These various features are extracted by different descriptors, and all these descriptors are represented as a unified form of vectors. The differences of the feature vectors are calculated using cosine similarity. Additional details about the classification part of our system can be found in [23].

In PGF retrieval fields, a large quantity of unlabeled data exists. If we manually construct a training database, then the annotation task can be arduous or even impossible for further research. Active learning is helpful to select the most representative or informative unlabeled examples for labeling and training. Therefore, in our improved system, an active learning model is adopted to automatically obtain additional effectual training data for classification. Several special problems should be considered when we combine active learning with the PGFs, including the selections of sample uncertainty measures, batch selection strategies, and domain-specific factors.

IV. Active Learning-based Model

A. Feature Selection of Active Learning

In order to analyze PGFs, following features need to be considered: basic geometric primitive feature, dual-primitive structure binary feature, main primitive feature, global feature, boundary feature, and Zernike moments feature. Some of them are firstly introduced in [12], and global feature, boundary feature, and Zernike moments feature are adopted in the multi-label classification [23].

We select basic geometric primitive feature ($FV_{primitive}$), dual-primitive structure binary feature ($FV_{structure}$), and boundary feature ($FV_{boundary}$) as the active learning features according to their characteristics in describing PGFs, which are explained as following:

- A PGF is composed of several basic geometric primitives, including triangles, circles, rectangles, parallelogram, trapezoid, line segments, and arcs, denoted as $\{S_1, S_2, \ldots, S_k\}$. In our work, the basic geometric primitives are the most fundamental elements; they reveal the intrinsic properties of PGFs. Therefore, they can be used to effectively distinguish between PGFs and other geometric figures. $FV_{primitive}$ is built in term frequency-inverse document frequency form:
$$FV_{primitive}(i,j) = tf(i,j) * idf(j) = \frac{n(i,j)}{\sum_k n(i,k)} * \log(\frac{N}{df(j)}) \quad (1)$$
where one PGF is a document, and each primitive S_j is a term. For an image I_i, N is the total number of I, $n(i,j)$ is the number of term S_j in I_i, and $df(j)$ is the number of images that contain term S_j.
- The strong spatial correlation of the three main shape types (i.e., triangle, circle, and rectangle) may be overlapping or tangent (inside, outside), including sharing ending points. Based on our previous work [12], 41 strong-correlated compound shapes are selected as the dual-primitive structures, as shown in Figure 3. The dual-primitive structure binary feature ($FV_{structure}$) is built as their frequency histogram (scaled to 0∼1).
- The boundary of one PGF image is extracted by using the envelope extraction method [24] and converted into a curvature description. All curvatures are then aligned from the curvature peak index bin and then scaled into the longest curvature length. Besides, we use three other curvature metrics: mean, standard deviation, and peak number. And the boundary features are clustered, the cluster number is then recorded as a binary feature. The boundary feature is expressed as
$$FV_{boundary} = \Big[curvature, mean, std, peakNum, clusterNum\Big] \quad (2)$$

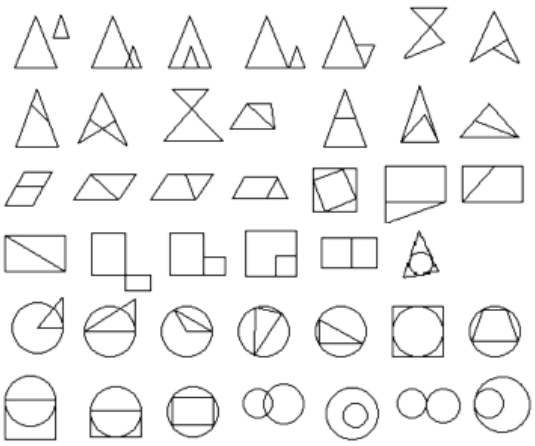


Fig. 3. Local structure samples.

B. Active Learning Algorithm

An active learner is established based on adaptive active learning, as shown in Figure 4. In this model,we adopt the similar approach used in [25] that combines a most uncertainty measure and an information density measure together. Compared to those methods mentioned in Section 2, the proposed method can select the candidate instance that is not only most uncertain to classify based on the current classifier, but also very informative about the remaining unlabeled instances. Some definitions we used are listed as following:

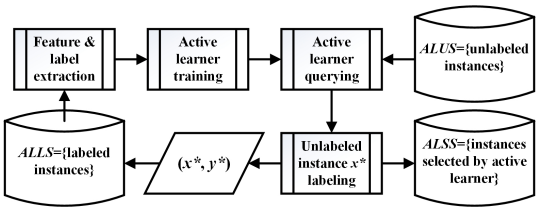


Fig. 4. Workflow of our active learning-based selection procedure.

- $ALLS$. The set of labeled instances already known to an active learner.
- $ALUS$. The set of unlabeled instances where an active learner queries the classification of an instance.
- $ALUS_x$. The set of unlabeled instances after removing an instance x from $ALUS$, such that $ALUS_x = ALUS - \{x\}$.
- $ALSS$. The set of instances, labeled with label "1", that an active learner recommends.
- LS. The label set and $LS = \{\pm 1\}$; for $y \in LS$, $y = 1$ means that an image is a suitable PGF with enough uncertainty, and $y = -1$ indicates that it is an unfit or extraordinary geometric figure, such as an analytic geometric figure.
- θ_{ALLS}. The classification model trained over $ALLS$.
- f. The uncertainty measure. It is defined as the conditional entropy of the label variable Y given the candidate instance x in probabilistic classification models:
$$f(x) = H(Y|x, \theta_{ALLS}) = -\sum_{y \in LS} P(y|x, \theta_{ALLS}) \log P(y|x, \theta_{ALLS}) \quad (3)$$

- d. The information density measure. It is defined as the mutual information between the candidate instance and the remaining unlabeled instances. For a candidate instance $x \in ALUS$, the information density measure, based on the mutual information, is defined as

$$d(x) = I(x, ALUS_x) = H(x) - H(x|ALUS_x) \quad (4)$$

where the entropy terms are computed in a Gaussian Process framework with multivariate Gaussian distributions. The Gaussian kernel function is adopted

$$k(x, x') = \exp\Big(-\frac{||x - x'||^2}{\tau^2}\Big) \quad (5)$$

and $\sigma_x^2 = k(x, x)$. Then the entropy term and the conditional entropy term

$$H(x) = \frac{1}{2}\ln(2\pi e\sigma_x^2) \quad (6)$$

$$H(x|ALUS_x) = \frac{1}{2}\ln(2\pi e\sigma_{x|ALUS_x}^2) \quad (7)$$

The information density measure defined in Equation 4 is computed by

$$d(x) = \frac{1}{2}\ln\Big(\frac{\sigma_x^2}{\sigma_{x|ALUS_x}^2}\Big) \quad (8)$$

- h_β. The adaptive combination of the uncertainty measure and the information density measure

$$h_\beta(x) = f(x)^\beta d(x)^{1-\beta} \quad (9)$$

where $0 \le \beta \le 1$ denotes a tradeoff controlling parameter over the two terms. The algorithm adaptively uses the β value from a set of pre-defined candidate values, such as $B = [0.1, 0.2, \ldots, 0.9, 1]$. For each different β value of B, it can select one instance to compose a set S. Then we try to select the best β value by minimizing the expected classification error on the unlabeled instances, which means selecting the most informative instance from S.
- B. The tradeoff controlling parameter β candidate values set.
- S. The instances that are selected with the corresponding β values of B.

The main steps of the algorithm are listed as the following:

1. Initialize *ALLS* by certain labeled instances with a positive or negative label (± 1), which are selected according to manual annotation results. And *ALUS* is initialized with unlabeled instances. $B = [0.1, 0.2, \ldots, 1]$;
2. At the start of each trial, the active learner already holds a probabilistic classifier θ_{ALLS} that is trained using features $(FV_{primitive}, FV_{structure}, FV_{boundary})$ and labels of the current training set $ALLS$;
3. Consider each unlabeled instance $x \in ALUS$:
   (1) calculate $f(x)$ according to Equation 3;
   (2) calculate $d(x)$ according to Equation 4;
   (3) calculate $h_\beta(x)$ with different $\beta \in B$ according to Equation 9.
4. Initialize S as an empty set;
5. Consider each $\beta \in B$:
   (1) select an instance $x = \arg\max_{x \in ALUS} h_\beta(x)$;
   (2) add x into S, $S = S \cup \{x\}$.
6. Consider each $x \in S$, try each possible label, $y \in LS$, with probability $P(y|x, \theta_{ALLS})$, and add the instance-label pair (x, y) to the training set $ALLS$,

$$ALLS^+ = ALLS \cup \{(x, y)\} \quad (10)$$

   (1) re-train a new classifier θ_{ALLS^+} over $ALLS^+$;
   (2) measure the prediction loss of the new classifier on all unlabeled instances. Calculate the expected loss of the instance x as a weighted sum of the prediction loss obtained using all possible labels y under the distribution $P(y|x, \theta_{ALLS})$;
   (3) select the instance x^* from S according to

$$x^* = \arg\max_{x \in S} \sum_{y \in LS} P(y|x, \theta_{ALLS}) \Big(\sum_{x \in ALUS} \big(1 - P(\hat{y}|x, \theta_{ALLS^+})\big)\Big) \quad (11)$$

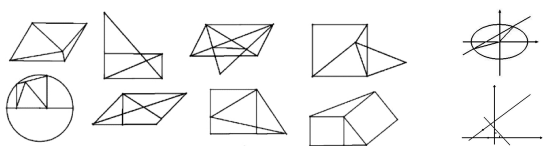
   where $\hat{y}$ is the predicted label for instance x.
7. The instance x^* is chosen as the recommended instance;
8. We label the instance x^* with a label y^* (± 1). Then, (x^*, y^*) is added to $ALLS$, and x^* is removed from $ALUS$. If $y^* = 1$, then x^* is added to $ALSS$. If $|ALSS| < M$ (a threshold), then the procedure iterates to step 2; otherwise the procedure exits.

V. Experiment

Because no authorized database for PGF evaluation is available, we have established two PGF databases, PKU-PGF-A and PKU-PGF-B. PKU-PGF-A contains 267 labeled PGFs extracted from digital PDF documents. The size ranges from 60×96 to 400×96, as shown in Figure 5(a). PKU-PGF-B contains 1,030 unlabeled images collected from math-learning websites. Apart from normal PGFs, it also includes several special geometric figures, as shown in Figure 5(b).

To demonstrate the performance of our active learning-based method, we first compare the predictive results based on two different training sets, $ALSS$ (images selected by the active learner from PKU-PGF-B) and RSS (images selected randomly from PKU-PGF-B), in the multi-label classification experiment with Mulan [26]. Then, we evaluate PGF retrieval quality, namely, precision and mean average precision (MAP) against recall between $ALSS$ and RSS. Lastly, the annotation workload reduction experiment is conducted.

Our experiments are conducted on a computer with a 3.2GHz Intel Core i5 CPU and a 16 GB memory using MATLAB R2014b and Eclipse.

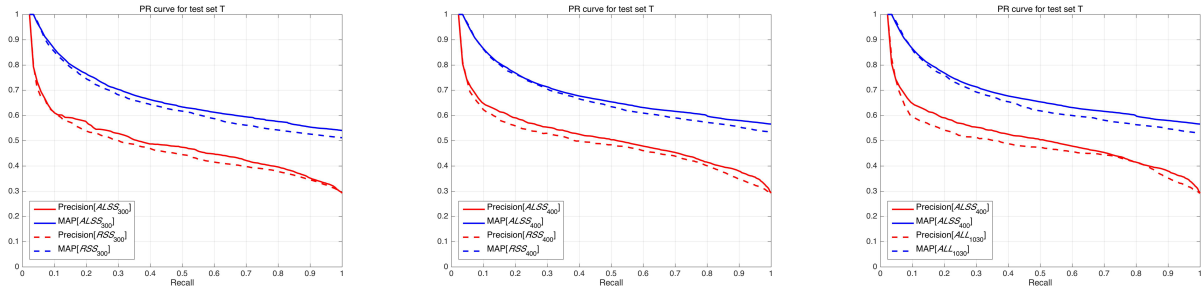


(a) (b)

Fig. 5. Some examples in PKU-PGF-A and PKU-PGF-B.

TABLE I
COMPARISON RESULTS OF PREDICTIVE PERFORMANCE BASED ON $ALSS$ VS. RSS

	Example-Based			**Label-Based**				
	Recall	F_1	Accuracy	Micro-Recall	Micro-F_1	Macro-Precision	Macro-Recall	Macro-F_1
RSS_{50}	0.5617	0.4508	0.3590	0.6108	0.5259	0.4398	0.5025	0.4241
$ALSS_{50}$	0.5658	0.4885	0.4081	0.5918	0.5297	0.4489	0.5805	0.4720
RSS_{100}	0.6108	0.4724	0.3745	0.6614	0.5641	0.4805	0.5752	0.4651
$ALSS_{100}$	0.7050	0.5307	0.4196	0.7532	0.5554	0.4724	0.7188	0.4768
RSS_{200}	0.4608	0.4908	0.4175	0.4842	0.5730	0.4818	0.3528	0.3792
$ALSS_{200}$	0.6446	0.5460	0.4492	0.6804	0.5636	0.4576	0.6378	0.4687
RSS_{300}	0.3933	0.4152	0.3488	0.4114	0.5179	0.3782	0.3102	0.3337
$ALSS_{300}$	0.7296	0.5441	0.4354	0.7753	0.5652	0.5424	0.7435	0.5001
RSS_{400}	0.3700	0.4367	0.3650	0.3829	0.5193	0.3226	0.2368	0.2563
$ALSS_{400}$	0.6383	0.5440	0.4533	0.6646	0.5512	0.3661	0.6202	0.4364
RSS_{ave}	0.4793	0.4532	0.3730	0.5101	0.5400	0.4206	0.3955	0.3717
$ALSS_{ave}$	**0.6567**	**0.5307**	**0.4331**	**0.6931**	**0.5530**	**0.4575**	**0.6602**	**0.4708**



(a) (b) (c)

Fig. 6. Precision and MAP curve: (a) $ALSS_{300}$ vs. RSS_{300}; (b) $ALSS_{400}$ vs. RSS_{400}; (c) $ALSS_{400}$ vs. ALL_{1030}.

A. Multi-label Classification Experiment

1) Multi-label Set and Evaluation Metrics: We consider the dominant primitives and their spatial relationship in PGFs as labels to form the multi-label set, *MLS* = {*circle*, *triangle*, *quadrangle*, *others*, *intersection relation*, *inclusion relation*}. Different classes are generated by diverse label values. Then, two types of metrics are adopted to evaluate the multi-label classification, namely, example-based metrics and label-based metrics, which include recall, precision, F measure, etc. Specific formulas can be found in [27].

2) Multi-label Learning Results: We separately set $ALSS$ and RSS as a training set to compare their multi-label classification predictive performance on the test set $T \subset$ PKU-PGF-A. In our experiment,

- $|ALSS_{50}| = 50$, $|ALSS_{100}| = 100$, $|ALSS_{200}| = 200$, $|ALSS_{300}| = 300$, $|ALSS_{400}| = 400$;
- $|RSS_{50}| = 50$, $|RSS_{100}| = 100$, $|RSS_{200}| = 200$, $|RSS_{300}| = 300$, $|RSS_{400}| = 400$;
- $|T| = 200$.

The results are shown in Table I, $ALSS$ achieves superior predictive performance than RSS under all the considered metrics according to the overall results and average values, $ALSS_{ave}$ and RSS_{ave}. Especially, the predictive performance of $ALSS$ is obviously superior than RSS on each considered metric with scales 300 and 400. The metric values of $ALSS_{50}$, $ALSS_{100}$, and $ALSS_{200}$ are not very stable, which indicates that their scales are insufficient to stably predict the multi-label classification of T. We can find that RSS has better results than $ALSS$ with the scales 100 and 200 in metrics, Micro-F measure and Macro-Precision. It may be caused by the results, small amounts of positives (true positives and false positives) and high proportion of true positives in positives, that RSS predicts. This results in higher Macro-Precision of RSS than $ALSS$, which also has impact on Micro-F measure. However, $ALSS_{100}$ and $ALSS_{200}$ still outperform RSS on the same scale in most metrics, such as recall, F measure, accuracy, Micro-recall, Macro-recall, and Macro-F measure. And given that the multi-label classification results work as a filter to obtain PGF candidates, which we have mentioned in Section 3, we pay more attention on recall, Micro-recall, and Macro-recall metrics. The predictive performances under these three metrics show that $ALSS$ outperforms RSS. As for RSS, its predictive metric values on T are unstable in spite of varying scale. Therefore, $ALSS$ is obviously more effective than RSS as the training set.

B. PGF Retrieval Experiment

On the basis of the above multi-label classification results, PGFs are retrieved among candidate PGFs rather than in the entire database for a query PGF. For PGF matching, basic geometric primitive feature, dual-primitive structure binary feature, main primitive feature, and global feature are used to form the feature vectors of PGFs. $ALSS_{300}$ and $ALSS_{400}$, which have stable predictive performance, are separately se-

lected as the training set to compare PGF retrieval quality with RSS on the same scale over the test set T, as we have collected the ground-truth similarity scores for PKU-PGF-A. As shown in Figure 6(a) and 6(b), the precision (red line) and MAP (blue line) curve against recall of $ALSS$ (solid line) both outperform RSS (dash line) with the scales 300 and 400. These retrieval results exhibit the competitive performance of $ALSS$, which implies that a large number of relevant PGFs have been ranked before irrelevant ones.

C. Annotation Workload Reduction Experiment

Further more, we labeled all 1030 unlabeled images in PKU-PGF-B, thus getting a new training set, ALL_{1030}, $|ALL_{1030}| = 1030$. We compare PGF retrieval quality between ALL_{1030} and $ALSS_{400}$. According to the results shown in Figure 6(c), we can find that the retrieval performance of $ALSS_{400}$ (solid line) is better than ALL_{1030}'s (dash line). Therefore, the active learning method with 400 labeled instances obtains better PGF retrieval quality than the performance of the original classifier with 1030 labeled instances. It is obvious that our active learning-based frame contributes to reduce the annotation workload.

According to all the above experiment results, the active learning-based selection procedure effectively improves the selection of training instances for multi-label learning and PGF retrieval. Moreover, this procedure practically helps reduce the annotation workload. And it is an effective way to build PGF databases for training classifiers.

VI. Conclusion and Future Work

We focus on the approach combining active learning and multi-label classification to improve the efficiency of PGF retrieval with a feature descriptor and enhance the selection of training instances simultaneously. Besides, it contributes to reduce the burdensome work of annotation. Although the current database is limited, the comparative experiment continues to illustrate the effectiveness of the proposed model. This method also provides an effective means to extend graph databases. In the future, we will explore more different methods of uncertainty measure for unlabeled samples to improve the function and performance of active learning.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 61300061, No. 61472014, No. 61573028).

References

[1] R. Zanibbi and D. Blostein, "Recognition and retrieval of mathematical expressions," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 15, no. 4, pp. 331–357, 2012.

[2] P. De, S. Mandal, A. Das, and B. Chanda, "Detection of electrical circuit elements from documents images," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2015, pp. 94 020O–94 020O.

[3] L.-P. d. l. Heras, O. R. Terrades, and J. Lladós, "Attributed graph grammar for floor plan analysis," in *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*. IEEE, 2015, pp. 726–730.

[4] R. Jain and D. Doermann, "Logo retrieval in document images," in *2012 10th IAPR International Workshop on Document Analysis Systems*. IEEE, 2012, pp. 135–139.

[5] N. Nayef and T. M. Breuel, "Graphical symbol retrieval using a branch and bound algorithm," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 2153–2156.

[6] T.-A. Pham, M. Delalandre, S. Barrat, and J.-Y. Ramel, "Accurate junction detection and characterization in line-drawing images," *Pattern Recognition*, vol. 47, no. 1, pp. 282–295, 2014.

[7] E. Saund, "A graph lattice approach to maintaining and learning dense collections of subgraphs as image features," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 10, pp. 2323–2339, 2013.

[8] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern recognition*, vol. 37, no. 1, pp. 1–19, 2004.

[9] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 2, pp. 286–299, 2007.

[10] S. Tabbone, O. R. Terrades, and S. Barrat, "Histogram of radon transform. a useful descriptor for shape retrieval," in *19th International Conference on Pattern Recognition-ICPR 2008*, 2008.

[11] K. Li, X. Lu, H. Ling, L. Liu, T. Feng, and Z. Tang, "Detection of overlapped quadrangles in plane geometric figures," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 260–264.

[12] L. Liu, X. Lu, K. Li, J. Qu, L. Gao, and Z. Tang, "Plane geometry figure retrieval with bag of shapes," in *Document Analysis Systems (DAS), 2014 11th IAPR International Workshop on*. IEEE, 2014, pp. 1–5.

[13] M. J. Seo, H. Hajishirzi, A. Farhadi, and O. Etzioni, "Diagram understanding in geometry questions." in *AAAI*, 2014, pp. 2831–2838.

[14] W.-J. Kuo and C.-C. Lin, "Two-stage road sign detection and recognition," in *Multimedia and Expo, 2007 IEEE International Conference on*. IEEE, 2007, pp. 1427–1430.

[15] C. Akinlar and C. Topal, "Edcircles: a real-time circle detector with a false detection control," *Pattern Recognition*, vol. 46, no. 3, pp. 725–740, 2013.

[16] Y. Liu, T. Ikenaga, and S. Goto, "An mrf model-based approach to the detection of rectangular shape objects in color images," *Signal Processing*, vol. 87, no. 11, pp. 2649–2658, 2007.

[17] C. R. Jung and R. Schramm, "Rectangle detection based on a windowed hough transform," in *Computer Graphics and Image Processing, 2004. Proceedings. 17th Brazilian Symposium on*. IEEE, 2004, pp. 113–120.

[18] B. Settles, "Active learning literature survey," *University of Wisconsin, Madison*, vol. 52, no. 55-66, p. 11, 2010.

[19] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," *The Journal of Machine Learning Research*, vol. 2, pp. 45–66, 2002.

[20] Z. Zhang, J. Qin, Y. Wang, and M. Liang, "Object classification in traffic scene surveillance based on online semi-supervised active learning," in *2014 22nd International Conference on Pattern Recognition (ICPR)*. IEEE, 2014, pp. 3086–3091.

[21] A. Freytag, E. Rodner, P. Bodesheim, and J. Denzler, "Labeling examples that matter: Relevance-based active learning with gaussian processes," in *Pattern Recognition*. Springer, 2013, pp. 282–291.

[22] L. Liang and K. Grauman, "Beyond comparing image pairs: Setwise active learning for relative attributes," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 208–215.

[23] S. Fu, X. Lu, L. Liu, J. Qu, and Z. Tang, "A diagram retrieval method with multi-label learning," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2015, pp. 94 020N–94 020N.

[24] J. Song, X. Lu, H. Ling, X. Wang, and Z. Tang, "Envelope extraction for composite shapes for shape retrieval," in *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE, 2012, pp. 1932–1935.

[25] X. Li and Y. Guo, "Adaptive active learning for image classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 859–866.

[26] G. Tsoumakas, E. Spyromitros-Xioufis, J. Vilcek, and I. Vlahavas, "Mulan: A java library for multi-label learning," *The Journal of Machine Learning Research*, vol. 12, pp. 2411–2414, 2011.

[27] G. Tsoumakas, I. Katakis, and I. Vlahavas, "Mining multi-label data," in *Data mining and knowledge discovery handbook*. Springer, 2010, pp. 667–685.