

Cross-Checking-Based Trademark Image Retrieval for Hot Company Detection

Hao Wu, Beijing Normal University, China

Zhiyi Zhang, Beijing Normal University, China

Zhilin Zhu, Shandong Technology and Business University, China*

ABSTRACT

A trademark is an essential symbol of a company, consisting of a semantically rich image under ordinary circumstances. The popularity of a company can be measured by the frequency of its trademark being used. Therefore, efficiently retrieving trademark images would directly contribute to the detection of popular companies. However, most mainstream retrieval methods are not especially pertinent to trademark image retrieval. To solve this problem, a combination of the ResNet50 network and Autoencoder with local sensitive hashing (LSH) is used to conduct full cross-checking, which significantly improves the effectiveness of trademark image retrieval. Meanwhile, image super-resolution-based sparse coding is also proposed to achieve high-precision trademark image retrieval and its effect is particularly significant for challenging trademark images. Finally, the authors conduct extensive experiments on a high-quality database to demonstrate the substantial effectiveness of the proposed methods.

KEYWORDS

Autoencoder with LSH, hot company detection, image super-resolution, ResNet, sparse coding, trademark image retrieval

1. INTRODUCTION

With rapid changes occurring in the global economy and ways of doing business, the fortunes of companies and industries are also changing rapidly. Researchers, investors, and policy-makers are keen to face these changes proactively. They invest a great deal of resources to collect and analyse data to understand business performance and, more importantly, to predict the future of a company. One important measurement of a company's performance and its potential is its popularity with the general public. In particular, if a company's trademark appears frequently, it can indicate that the company is highly popular. Consequently, retrieving trademark images efficiently and accurately is becoming increasingly important.

Image retrieval technology has gone through three stages of development: text-based image retrieval (TBIR), content-based image retrieval (CBIR), and semantic-based image retrieval. TBIR is

DOI: 10.4018/JOEUC.335455

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

known as “searching images by tags”. This method is simple but time-consuming and labour-intensive because tags and indices such as titles, authors, and other metadata attributes are added by manual annotation. There were enormous amount of trademarks registered worldwide (World Intellectual Property Organization, 2018). Since the volume of digital image data on the internet has increased rapidly, along with the number of trademark images, TBIR is unsuitable for trademark retrieval from the internet where images lack annotation.

In contrast to TBIR, CBIR uses features that can be extracted automatically to retrieve images, avoiding the subjectivity of manual description, and improving retrieval efficiency. Low-level visual features include colour, texture, shape, etc., and different feature representations require different similarity measurement methods. Colour is the most intuitive physical feature of colour images; the methods available to describe colour include colour histograms (Swain & Ballard, 1991), colour correlograms (Huang et al., 1997), and colour coherence vectors (Pass et al., 1997). Texture is a measurement of the relationship between pixels in a local area; its purpose is to describe the spatial distribution of grey levels in the neighbourhood of pixels. Shape descriptors are even more important than colour or texture descriptors and can be grouped into contour-based and region-based approaches. The former uses image boundary information, while the latter uses information on the grey distribution in a certain area. The Fourier descriptor (Del Vecchio & Salvini, 2000) is one of the most commonly studied and used contour-based shape descriptors. It is characterized by good computational performance and is easy to normalize. However, it is unable to capture the local representation of shapes and is sensitive to boundary noise and variations, leading to the Gibbs phenomenon when used to reconstruct complex trademarks.

In addition to low-level features, images can be analysed according to their high-level semantic content, i.e., what they conceptually represent. Machine learning and neural network models such as AlexNet (Krizhevsky et al. 2017), VGGNet (Simonyan & Zisserman, 2014), Inception V4 (Szegedy et al., 2017), ResNet (He et al, 2016), and DenseNet (Huang et al., 2017) have been widely used due to their strength in extracting highly semantic and abstract features and realizing nonlinear feature mapping (Perez et al., 2018). Some methods achieve improved performance through deep learning. An end-to-end model (Mafla et al., 2021) combines text and visual features to achieve fine-grained classification and image retrieval through a multimodal inference module. Recently, more novel deep learning models have been proposed. CVNet (Lee et al., 2022) adopts geometric verification after a global search with global descriptor matching and local feature matching. Global search quickly performs a rough search across the entire database, and geometric validation reorders the results of a rough search by precisely assessing only the candidates identified by the global search. ViT-Slim (Chavan et al., 2022) replaces the convolutional neural network in network slimming with a transformer to realize more flexible and efficient visual retrieval and classification. Zhao et al. drew on the idea of dense retrieval, discretized images and texts into tokens, and aligned them across modalities, greatly improving the efficiency of large-scale graphic retrieval (2023).

In the retrieval of trademark images, the characteristics of the trademark image should be fully considered. Traditional image retrieval methods are difficult to apply directly to trademark image retrieval. Although the improved sparse coding-based method (Sun et al., 2019), spatial pyramid matching-based method (Lazebnik et al., 2006), and GIST-based similarity calculation method (Hays & Efros, 2007) can achieve image retrieval with relatively few reference images, there are still certain limitations in their application. Therefore, we are attempting to draw on the advantages of traditional methods as much as possible and optimize them according to actual needs. Many practical methods (Bao et al, 2021; Zou et al., 2022; Trappey et al., 2021) have been proposed by experts and applied to solve practical problems, achieving convincing results in their fields of application. Although the above methods cannot be directly used to solve our problem, their entire strategy can serve as a reference.

As a classic network model, ResNet performs well for image feature extraction. Its small parameter count makes model loading and weight training fast. In addition, the pretrained ResNet50 model, trained on a large set of open-source image data, is recognized as a high-performing feature extraction

model for tasks such as image captioning. Thus, in the present study, we make full use of ResNet (He et al., 2016) as the first method to extract image features. To make our method more robust, we also employ Autoencoder (Bank et al., 2023), a classic deep learning model with advantages in data denoising and dimensionality reduction. It is worth mentioning that when Autoencoder is combined with local sensitive hashing (LSH), an image query indexing technology for dimensionality reduction, the model performs well in terms of computational time while remaining relatively accurate. Considering the characteristics of trademark images and the practical application value of deep learning models, a combination of the ResNet50 network and Autoencoder with LSH would be applicable. To further improve the accuracy of trademark image retrieval, we perform image super-resolution reconstruction (Dong et al., 2015) on the trademark images and conduct trademark image retrieval based on sparse coding (Arora et al., 2015; Yang et al., 2009). This strategy not only improves the accuracy of retrieval but also has significant effects on challenging trademark images.

The main contributions of this paper are as follows:

1. A combination of the ResNet50 network and Autoencoder with LSH is employed for trademark image retrieval, to ensure relatively high accuracy without requiring extensive computational resources.
2. Using image super-resolution-based sparse coding, we achieve more accurate trademark image retrieval, which demonstrates consistently excellent experimental results even for challenging trademark images.
3. Our method can be effectively applied to detect popular companies and serve an important function in the field of economics.

2. METHODS

In this section, we will first introduce the ResNet50 network and Autoencoder with LSH in order. Then, we will introduce an image super-resolution-based sparse coding model. During the model introduction, we will focus on providing sufficient details and implementation processes.

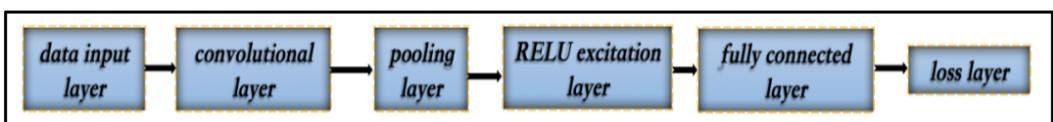
2.1 ResNet50

Convolutional neural network (CNN) architecture is widely used in image processing. It processes the input data layer by layer and automatically learns the network parameters. Loss is calculated in each processing layer and propagated backwards to optimize the network parameters.

Figure 1 shows the main layers in the CNN. As the core block, the convolutional layer extracts features through several convolutional kernels. The pooling layer is a form of nonlinear downsampling. The subsequent rectified linear unit excitation layer removes negative values (by setting them to zero), which boosts the training speed of the neural network without notable damage to the accuracy of generalization. In the fully connected layer, each neuron is connected to every neuron in the adjacent layers. Finally, the loss layer defines how the training penalizes the deviation of the predicted outputs from the true labels.

A common concern “vanishing gradients,” which occur as the number of layers increases. To address this problem, ResNet utilizes shortcut connections as a solution to guarantee that more layers

Figure 1. The main layers of CNN



will not lead to worse accuracy. In this way, it generates better accuracy by allowing the addition of more layers. Two basic blocks in ResNet are the identity block and the convolution block. We chose the well-trained ResNet50 due to its relatively small parameter size. Its structure is illustrated in Table 1. Additionally, we present various similarity measures in Table 2. In this paper, we select the most suitable similarity measures for different methods to solve practical problems. Considering that 2048-dimensional eigenvectors are used to represent images in this process, the common cosine distance is used to measure similarity.

2.2 Autoencoder With LSH

Autoencoding (Bank et al., 2023) is a type of data compression algorithm that can be used to learn efficient data coding (z) of an input image (x) in an unsupervised manner. It is data-specific and

Table 1. Structure of ResNet50

Layer	Output size	Convolutional kernel
Conv1	112×112	7×7,64, stride 2
		3×3, max_pool, stride 2
Conv2_x	56×56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3_x	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Conv4_x	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Conv5_x	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
Output	1×1	

Table 2. Common Distance Calculation Formulas

Euclidean Distance	Standardized Euclidean Distance	Manhattan Distance	Correlation Distance	Cosine Distance
$\sum_{k=1}^n (p_k - q_k)^2$	$\sum_{k=1}^n \left(\frac{p_k - q_k}{s_k} \right)^2$	$\sum_{k=1}^n p_k - q_k $	$\frac{\sum_{k=1}^n (p_k - \bar{p})(q_k - \bar{q})}{\sqrt{\sum_{k=1}^n (p_k - \bar{p})^2} \sqrt{\sum_{k=1}^n (q_k - \bar{q})^2}}$	$\frac{\sum_{k=1}^n p_k \cdot q_k}{\sqrt{\sum_{k=1}^n (p_k)^2} \sqrt{\sum_{k=1}^n (q_k)^2}}$

lossy. We use the autoregressive autoencoder to extract the compressed feature representation of trademark images.

The encoder transforms the eigenvector x of an input image into a condensed feature representation z by function $f(x)$. In this case, z has a smaller dimension than x . The decoder analyses the condensed representation z by function $g(z)$ to reconstruct a copy of the input image as x' . Using stochastic gradient descent, the parameters of the encoder and decoder functions can be optimized to minimize the reconstruction loss (deviation between x and x'). Figure 2 is an illustration of the main process to build the autoregressive autoencoder and the core structure of an autoencoder.

Traditional methods take a long time to retrieve images from a large dataset because they use images' statistical characteristics. Many high-dimensional indexing technologies are used to shorten the retrieval time. Common technologies include tree-based indexing, cluster-based indexing, inverted file indexing, and hash-based indexing. For images, LSH, a kind of hash-based indexing method, has proven to be suitable for our experiment. It is used to build a query index after eigenvectors are extracted through an autoencoder. A proper hash function that maps high-dimensional features to low-dimensional features is needed. It maintains the maximum possible similarity by mapping similar eigenvectors to the same collision bucket; thus, points that are close together in the original feature space are still close to each other in the generated feature space.

The step-by-step instructions for combining Autoencoder with LSH are shown in Figure 3. First, training images are selected and normalized, and we train the encoder on them. Second, we use the encoder to predict the eigenvector of 100-dimensional images. A LSH class helps to build 32-bit-long binary LSH indexes. Ultimately, based on similarity, specific images in the dataset are retrieved using a dictionary that maps binary hashes to image IDs.

The 100-dimensional vectors are then transformed into a 32-bit-long binary hash code in this process, and Euclidean distance is used for comparison and retrieval. In particular, in the execution of the above models, alternative extended queries can be applied to obtain more realistic experimental results.

2.3 Image Super-Resolution-Based Sparse Coding

The above models can achieve relatively reliable image retrieval for trademark images and have relatively fast computation speeds. However, if we need to further improve the retrieval accuracy, especially for challenging trademark images, we need to design a more sophisticated model. Specifically, image super-resolution reconstruction is performed on the trademark image, and then sparse coding is applied on this basis to calculate the similarity for image retrieval.

SRCNN (Dong et al., 2015) is an image super-resolution reconstruction model based on deep learning. It trains a deep neural network to map low-resolution images to high-resolution images and

Figure 2. The main process to build an autoregressive autoencoder

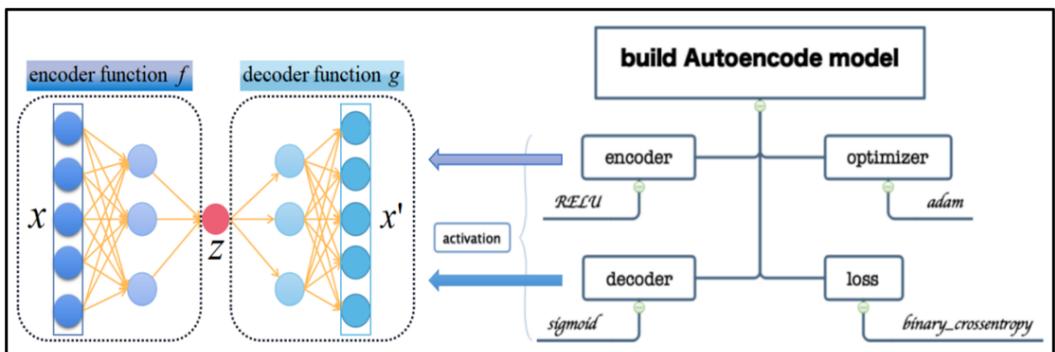
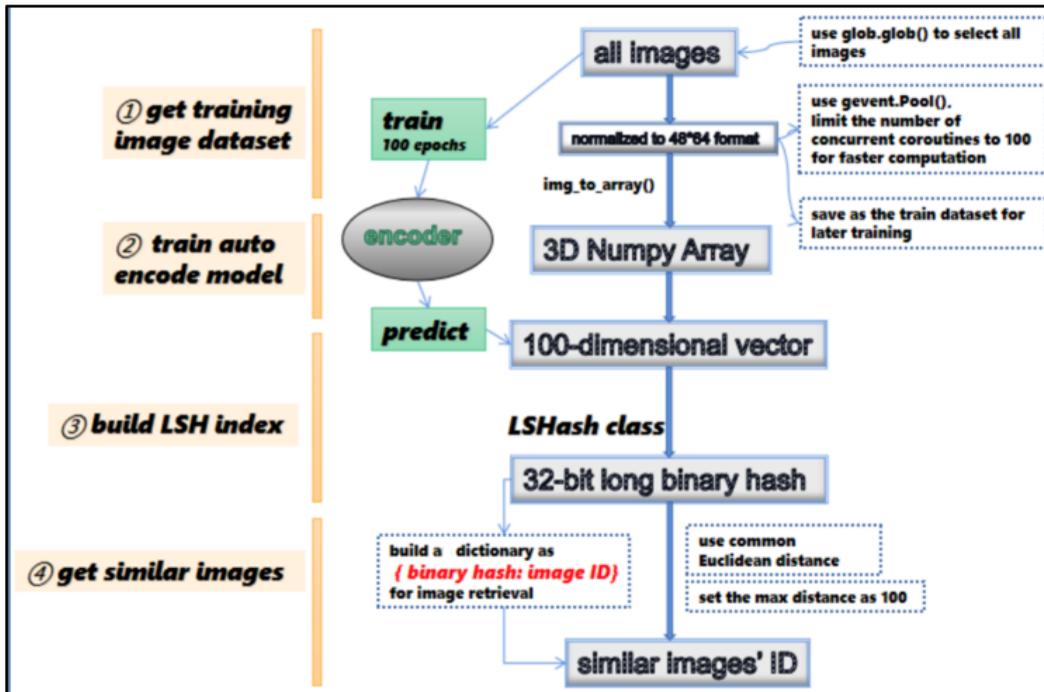


Figure 3. The step-by-step instructions for combining autoencoder and LSH



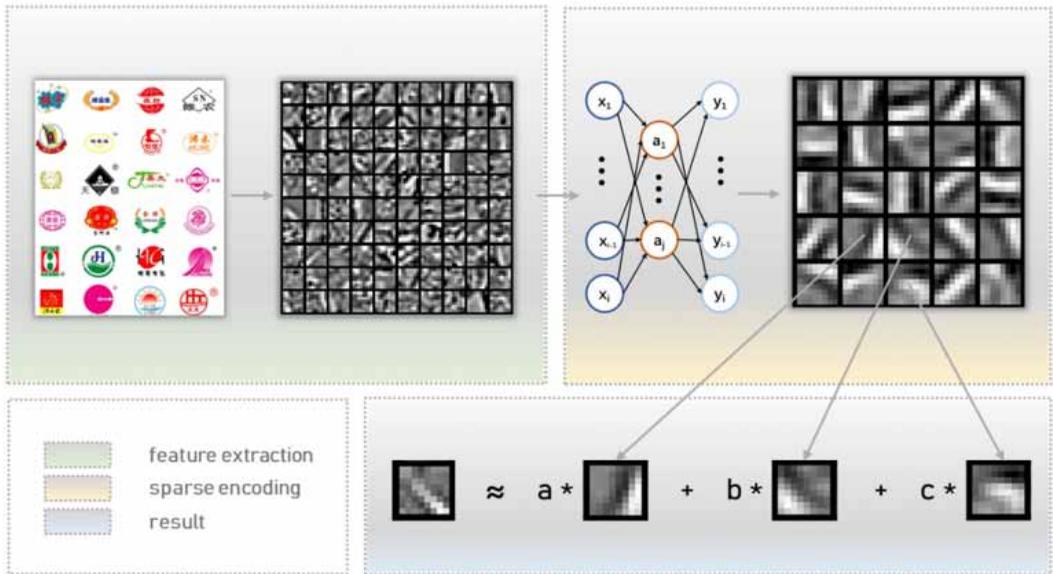
realizes the super-resolution reconstruction of images. The SRCNN is composed of three parts. Initially, during patch extraction, it divides patches of the image and extracts the features by convolution. Each image patch is mapped to a low-resolution dictionary, and a group of vectors is obtained to express the image after preprocessing. Then, in nonlinear mapping, the low-resolution features are mapped to high-resolution features to find the high-resolution features corresponding to image patches. The high-dimensional vector of the image patch obtained in the first step is mapped to another high-dimensional vector, and the high-resolution image patch is expressed through this high-dimensional vector for the final reconstruction. Finally, in the resolution process, the network reconstructs images based on high-resolution features. The final high-resolution image patches are aggregated to form the final high-resolution image. The reconstruction part is a linear operation and is implemented by a linear function. Using the above model, high-quality super-resolution reconstruction of trademark images can be achieved.

As the essential representative detector, sparse coding (Arora et al., 2015; Yang et al., 2009) can represent trademark images fully, on the foundation of exact representation, and the similarity between different trademark images can be calculated. The essential content of sparse coding is that some patches from the complete training process are used to represent the trademark image contributing to similarity evaluation. Figure 4 illustrates the essential process of sparse coding. After the above process is completed, Euclidean distance is used to retrieve the trademark images.

3. EXPERIMENTS

To verify the effectiveness of the ResNet50 network and Autoencoder with LSH, we established a naturalistic trademark image database containing tens of thousands of representative trademark images.

Figure 4. The essential process of image representation by sparse coding



Our database consists of trademark images from sports and other industries. Based on our hypothesis, the more popular a brand is, the more its trademark will be represented in the database. Additionally, we included thousands of trademark images belonging to additional industries from the FlickrLogo-47 dataset (Universität Augsburg, n.d.). Some trademark image instances are shown in Figure 5.

First, we tested the effectiveness of ResNet50. We input one image of each of the Nike and Reebok trademarks and retrieved the most similar trademark images, partial representative results after screening are shown in Figure 6.

Similarly, we tested the effectiveness of Autoencoder with LSH. The same Nike and Reebok trademark images were input, and the most similar trademark images were retrieved; partial representative results after screening are shown in Figure 7.

Figure 5. Trademark image instances from FlickrLogo-47



Figure 6. Partial representative results after screening using ResNet50

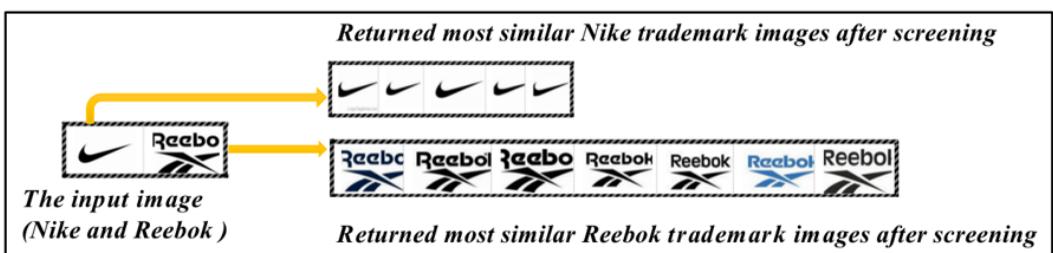


Figure 7. Partial representative results after screening using autoencoder with LSH



In the same way, we conducted multiple other experiments and carefully analysed the experimental results. Both models can quickly retrieve target trademark images with relatively high accuracy. In particular, when the two models are combined, retrieval accuracy can be further improved.

Next, we validated the image super-resolution-based sparse coding through adequate comparisons. In particular, we built a more representative database that includes not only regular trademark images but also some challenging trademark images, such as images with noise, lower resolution, and missing regions. By considering the widest possible variety of scenarios, we can verify the effectiveness of the method more thoroughly. Through comparison with some appropriately modified typical baseline methods (the improved sparse coding-based method [Sun et al., 2019], the spatial pyramid matching-based method [Lazebnik et al., 2006] and the GIST-based similarity calculation method [Hays & Efros, 2007]), we observe that our method could be able to achieve the retrieval of target trademark images most accurately as evaluated by both AP values and AUC values (Figures 8 and 9).

To further test the role of super-resolution reconstruction and sparse coding, we added an ablation experiment to verify these two steps. From the experimental results in Figure 10, we can see that each step plays an important role, and there is a significant decrease in retrieval accuracy if any of these steps are abandoned or replaced by another common method. Similarly, in Figure 11, the AUC values further illustrate the necessity of the two steps.

Figure 8. AP values calculated using our method and baseline methods

Note. Baseline 1: The improved sparse coding-based method. Baseline 2: The spatial pyramid matching-based method. Baseline 3: The GIST-based similarity calculation method.

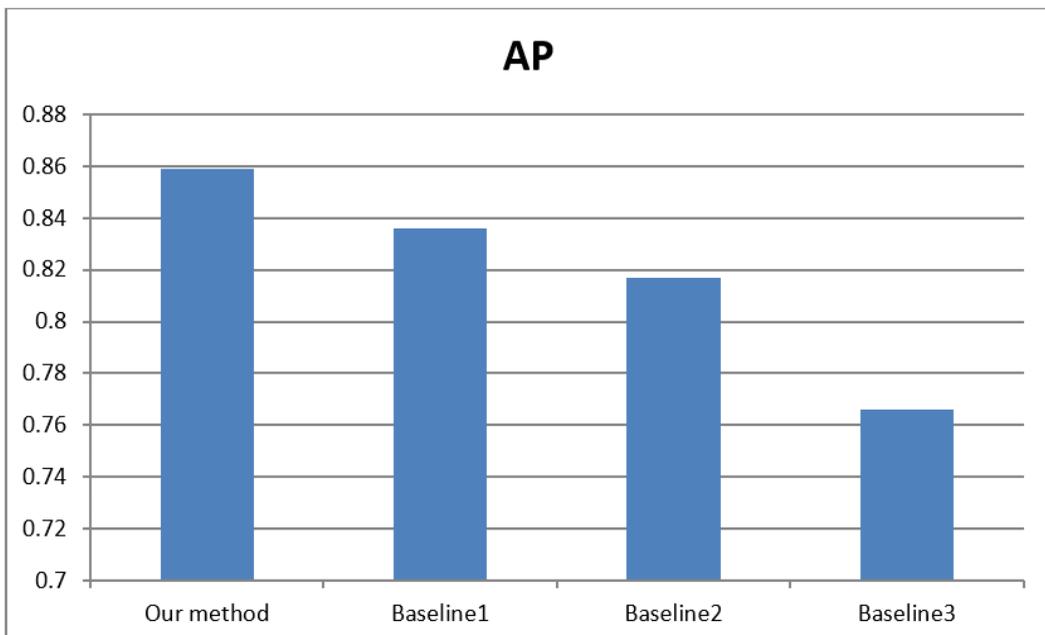


Figure 9 AUC values calculated using our method and baseline methods

Note. Baseline 1: The improved sparse coding-based method. Baseline 2: The spatial pyramid matching-based method. Baseline 3: The GIST-based similarity calculation method.

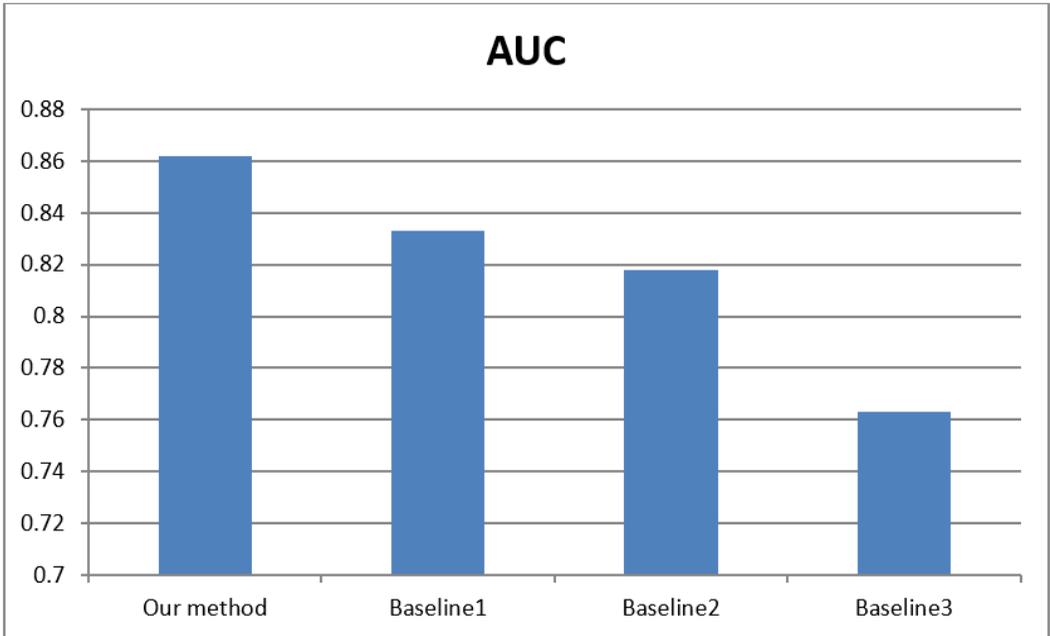


Figure 10. Comparison of AP values between our method and baseline methods

Note. Baseline method 1 is to retrieve trademark images without using the image super-resolution reconstruction step. Baseline method 2 is to retrieve trademark images using only a common encoding model.

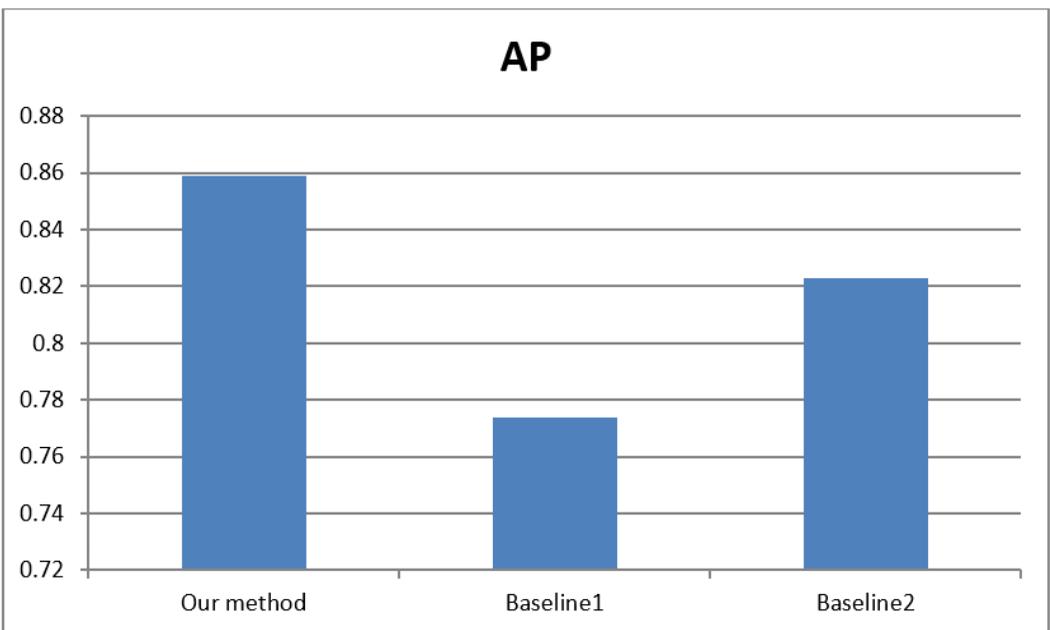
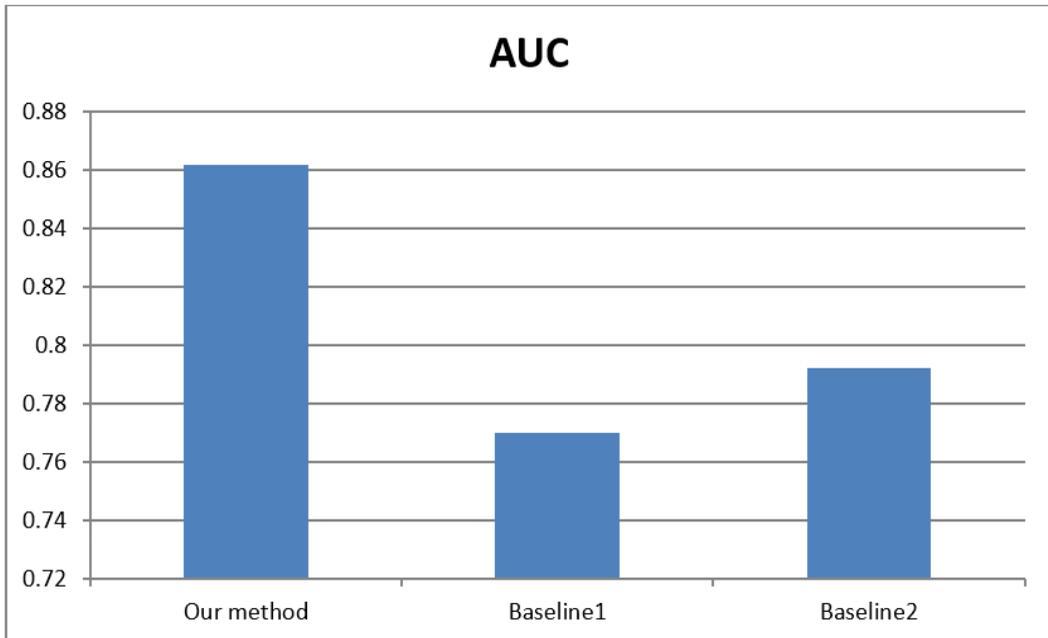


Figure 11. Comparison of AUC values between our method and baseline methods

Note. Baseline method 1 is to retrieve trademark images without using the image super-resolution reconstruction step. Baseline method 2 is to retrieve trademark images using only a common encoding model.



Overall, the above experiments are generally in line with expectations. The methods of ResNet50 and Autoencoder with LSH could retrieve trademark images quickly. In particular, the combination of the two methods could further improve the effectiveness of trademark image retrieval. The image super-resolution-based sparse coding method shows superiority in terms of accuracy, especially when the database contains more challenging trademark images. In comparison with multiple baseline methods, it is the best performer in terms of accuracy and other aspects. Furthermore, the steps of image super-resolution reconstruction and sparse coding are fundamentally reasonable and play a sufficient role.

4. CONCLUSION

In this paper, ResNet50 and the Autoencoder with LSH are effectively combined to achieve satisfactory trademark image retrieval with reduced consumption of computational power. Furthermore, to enhance the precision of trademark image retrieval and address challenging trademark images comprehensively, the image super-resolution-based sparse coding method is fully utilized. Moreover, we conduct sufficient experiments on an authoritative database and compare the proposed methods with some classic methods to better verify their effectiveness. Our methods can effectively detect popular companies by retrieving trademark images, serving an important function in the field of economics.

ACKNOWLEDGMENT

This research was supported by National Nature Science Foundation of China (Grant Nos. 62072043) and Blockchain Innovation Application Pilot Project.

REFERENCES

- Arora, S., Ge, R., Ma, T., & Moitra, A. (2015, June). Simple, efficient, and neural algorithms for sparse coding. In *Conference on learning theory* (pp. 113–149). PMLR. <https://proceedings.mlr.press/v40/Arora15.html>
- Bank, D., Koenigstein, N., & Giryas, R. (2023). Autoencoders. In L. Rokach, O. Maimon, & E. Shmueli (Eds.), *Machine learning for data science handbook*. Springer. doi:10.1007/978-3-031-24628-9_16
- Bao, Y., Cheng, X., & Zarifis, A. (2021). Exploring the Impact of Country-of-Origin Image and Purchase Intention in Cross-Border E-Commerce. [JGIM]. *Journal of Global Information Management*, 30(2), 1–20. doi:10.4018/JGIM.20220301.oa7
- Chavan, A., Shen, Z., Liu, Z., Liu, Z., Cheng, K. T., & Xing, E. P. (2022). Vision transformer slimming: Multi-dimension searching in continuous optimization space. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4931–4941). IEEE. https://openaccess.thecvf.com/content/CVPR2022/html/Chavan_Vision_Transformer_Slimming_Multi-Dimension_Searching_in_Continuous_Optimization_Space_CVPR_2022_paper.html?ref=roboflow-blog
- Del Vecchio, P., & Salvini, A. (2000). Neural Network and Fourier Descriptor macromodeling dynamic hysteresis. *IEEE Transactions on Magnetics*, 36(4), 1246–1249. doi:10.1109/20.877666
- Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2), 295–307. doi:10.1109/TPAMI.2015.2439281 PMID:26761735
- Hays, J., & Efros, A. A. (2007). Scene completion using millions of photographs. *ACM Transactions on Graphics (ToG)*, 26(3), 4-es. ACM. 10.1145/1276377.1276382
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778). IEEE. https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700–4708). IEEE. https://openaccess.thecvf.com/content_cvpr_2017/html/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.html
- Huang, J., Kumar, S. R., Mitra, M., Zhu, W. J., & Zabih, R. (1997, June). Image indexing using color correlograms. In *Proceedings of IEEE computer society conference on Computer Vision and Pattern Recognition* (pp. 762–768). IEEE., doi:10.1109/CVPR.1997.609412
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. doi:10.1145/3065386
- Lazebnik, S., Schmid, C., & Ponce, J. (2006, June). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06) (Vol. 2, pp. 2169–2178). IEEE. doi:10.1109/CVPR.2006.68
- Lee, S., Seong, H., Lee, S., & Kim, E. (2022). Correlation verification for image retrieval. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 5374–5384). The CVF. https://openaccess.thecvf.com/content/CVPR2022/html/Lee_Correlation_Verification_for_Image_Retrieval_CVPR_2022_paper.html
- Mafla, A., Dey, S., Biten, A. F., Gomez, L., & Karatzas, D. (2021). Multi-modal reasoning graph for scene-text based fine-grained image classification and retrieval. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 4023–4033). IEEE. https://openaccess.thecvf.com/content/WACV2021/html/Mafla_Multi-Modal_Reasoning_Graph_for_Scene-Text_Based_Fine-Grained_Image_Classification_and_WACV_2021_paper.html
- Pass, G., Zabih, R., & Miller, J. (1997, February). Comparing images using color coherence vectors. In *Proceedings of the fourth ACM international conference on multimedia* (pp. 65–73). Association for Computing Machinery. doi:10.1145/244130.244148

Perez, C. A., Estévez, P. A., Galdames, F. J., Schulz, D. A., Perez, J. P., Bastías, D., & Vilar, D. R. (2018, July). Trademark image retrieval using a combination of deep convolutional neural networks. In 2018 international joint conference on neural networks (IJCNN) (pp. 1–7). IEEE. doi:10.1109/IJCNN.2018.8489045

Sun, X., Nasrabadi, N. M., & Tran, T. D. (2019). Supervised deep sparse coding networks for image classification. *IEEE Transactions on Image Processing*, 29, 405–418. doi:10.1109/TIP.2019.2928121 PMID:31331886

Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, 7(1), 11–32. doi:10.1007/BF00130487

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1). doi:10.1609/aaai.v31i1.11231

Trappey, C. V., Chang, A. C., & Trappey, A. J. (2021). Building an internet-based knowledge ontology for trademark protection. [JGIM]. *Journal of Global Information Management*, 29(1), 123–144. doi:10.4018/JGIM.2021010107

Universität Augsburg. (n.d.). FlickrLogos [Data Set]. <https://www.uni-augsburg.de/en/fakultaet/fai/informatik/prof/mmc/research/datensatze/flickrlogos/>

World Intellectual Property Organization. (2018). *WIPO intellectual property statistics*. WIPO. <https://www.wipo.int/ipstats/en/>

Yang, J., Yu, K., Gong, Y., & Huang, T. (2009, June). Linear spatial pyramid matching using sparse coding for image classification. In 2009 *IEEE Conference on computer vision and pattern recognition* (pp. 1794–1801). IEEE. doi:10.1109/CVPR.2009.5206757

Zou, X., Zhu, D., Huang, J., Lu, W., Yao, X., & Lian, Z. (2022). WGAN-Based Image Denoising Algorithm. [JGIM]. *Journal of Global Information Management*, 30(9), 1–20. doi:10.4018/JGIM.300821