ISSN 2278-3091
Volume 9, No.5, September - October 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse70952020.pdf
https://doi.org/10.30534/ijatcse/2020/70952020

A Survey on Content Based Image Retrieval Using Convolutional Neural Networks

R.Sathya[1,2], B.Saleena[1]
[1]School of Computer Science and Engineering, VIT, Chennai
[2] Department of Computer Science and Engineering, SRMIST, Chennai

ABSTRACT

A smart image retrieval technique has been an increasing demand by the advancements in the field of computer networks and mobile computing. Generally, uploaded images have textual information associated with the owner's tags. When an input image is given as a search query, its text content is processed to identify the semantics of the image, and the output will be based on the labels stored in it. Traditional hashing techniques are most commonly used to provide high quality search results for labeled images. In this case, hashing codes are processed based on the semantic preserving information obtained from the features of input images using supervised learning methods. But the delivery of high quality pictures without human interaction, using automated annotation among very large scale image databases, is still a continuous research process. This paper focuses on comparing various methods for reducing the semantic gap between low-dimensional and high-dimensional features. It also focuses on content based image retrieval technique (CBIR), with an unsupervised learning method using convolutional Neural Networks (CNN).

Key words: CBIR, Visual hashing, unsupervised learning, Convolutional Neural Network, Deep Learning

1. INTRODUCTION

Image retrieval system [1] [2] is a process by which relevant images are searched and retrieved from a vast amount of data sets by feeding images as input queries. When the number of images [3] [4] [5] are increased, the complexity tends to increase. Conventional method of allowing people to give physical input in terms of catchphrases of pictures, in an enormous database, is time-consuming and might not catch each watchword that portrays the picture. These existing systems are categorized depending upon the features of images using low level (colors, grey scales, tint, etc.,) and middle level (area, contours, highlights, etc.,) qualities [6] [7]. In general, an image retrieval algorithm compares images according to descriptions from users' query images [8] [9] followed by similarity metrics [10]. There has been tremendous growth in the domain of image retrieval systems [11] [12], and many frameworks have been proposed in last few years. One of the frameworks called Content Based Image Retrieval System (CBIR) [6] [13] [14], is used in visual contexts to search for similar images among large scale image databases like Flickr [15] [16], Youtube, Vimeo, Wiki [17], etc. Another framework is based on region-based image retrieval system that utilizes region codebooks [18] for image retrieval.

This paper is mainly focused on CBIR, one of the applications in computer vision strategies [19], to provide solution for the tedious process of searching digital pictures from an enormous database. "Content based" implies that a query does processing and analysis of original contents, semantic, and features of images, instead of metadata such as descriptions, tags [20], and keywords associated with pictures. Here, the indexed images depend on visual context, such as hue, shape-form, and texture.

Constructiveness of CBIR system solely dependent on group of visual features selected [21] and also on feature of user's cognizance. Semantic-aided visual hashing (SAVH) [6] [22] was used to enhance the semantics embedded in the corresponding texts along with the pictures. In the first step of transformation, both visual features and textual features of pictures are retrieved and then image pixels are transformed into mathematical vectors. In the next step, a text enhanced visual graph is built with a hyper graph map, and then latent semantic elements are identified with text information associated with it.

The main problem with existing approaches lies in seeking a semantic gap among low level and high level concepts of social media images. The most important searching approach in CBIR based systems is semantic search, as humans can easily interpret images [23] [24] with visuals, but computer system may not correlate the images. As of now, no universally acceptable algorithm has been developed for semantic image retrieval system [25].

The perspectives of human beings are more accurate when compared to computer systems for content based search of images [26]. Hence, continuous research work relative to the achievement of building real-time image retrieval systems is still under progress. Earlier, hashing techniques [27] were used for construction of a similar set of binary codes from image collections. With the help of hashing functions, mapping of high dimensional visual data to low dimensional hamming (binary) spaces can be performed.

Later on, Supervised deep hashing technique [28] [29] incorporated binary hashing codes [30] from the annotated text of data in a large scale information retrieval [31]. The semantic text was based on attributes and relations among relevant text in the same cluster. The outcome of this became more effective and efficient in terms of performance and accuracy. But this approach was purely based on supervised methods.

Artificial Neural Network (ANN) was introduced to calculate the Hamming distances between binary vectors. Compared with locality sensitive hashing (LSH) [32], learning binary codes have become more efficient for mapping similar images. This system did not use training data, and thereby higher accuracy was achieved. But this was suitable only for small scale datasets under supervised learning [29]. Deep convolutional neural network (CNN) [28] [32] was introduced to overcome these issues. Rich middlelevel labels are used for image classification object detection and semantic segmentation. Usually, a huge dataset like ImageNet [34] [35] is used by deep CNN architecture [33]. Here, pre-trained data [2] has been transferred and fine-tuned to achieve better output together with improvement in its efficiency. However, scalable deep hashing [36] has been adapted for large-scale data learning and retrieval process.

This paper is organized in a way that Section 2 describes various image retrieval methods; Section 3 compares methods and algorithms associated with image retrieval process, and section 4 lists out various performance evaluation criteria. The main motivation behind this survey is discussed in section 5.

2. BENCHMARK IN IMAGE RETRIEVAL METHODS

Image retrieval system can be implemented in many different ways. The following sections deal with the most commonly used approaches for retrieval process.

2.1 Hashing and Ranking Based Models

Hashing models are generally used in machine learning, computer vision, and image retrieval systems, for retrieval of appropriate images from relevant data sets. Hashing enables mapping of high-level image features into compact low-level binary codes. To bridge the semantic gap, deep semantic hashing was used for a better understanding of binary representations of the images [30], and also semantic information was collected in parallel. The main benefit of this system lies in its reduction of computational cost. There were two ideas behind this model: data-dependent and data-independent. A similarity matrix for attaining semantic similarities between two sets of images was constructed with their corresponding coordinates. The next level was a feature representation between images and hash functions. Another modification to this approach was a deep semantic ranking based method [37] for preserving multilevel similarities between single and multi-labeled images [38] [39].

Yet another method, namely Locality sensitive hashing (LSH) [7], produced hash functions based on the similarity group of the same context with high probability. Generally, Euclidean distance [40] is used for finding feature-level similarities when the distance is small, the similarity of images will be high in this case. From this, top ranked images [41] are clustered together.

This process involves determination of semantic relevance [42] between targeted images and query images [43] [44]. Once it becomes zero, the targeted image is neglected and the process is repeated. A set of candidate images are obtained at the end of semantic relevance checking. Generally, Oxford and Holidays [40] datasets are used for comparison purposes. However, a large sparse matrix is required by these algorithms to coordinate similarity between data points of the training dataset.

2.2 Relevance Feedback

The basic idea behind relevant feedback [45] [46] is fetching counterpart examples from users for enhancing the performance of the system. Images can be retrieved based on user’s feedback [42]. The Query result obtained by the users is not the optimum one in most cases. Generally, user’s seeks better results. So, relevance is modified based on user’s interest. The retrieved results are stored in the database for further querying. A different table is kept for recognition of an arrangement of importance indicated by client’s question to store the feedback. As soon as it satisfies client’s interests, the resultant set is fed as input for the next level of iterations. This process continues for further queries.

No pre-defined algorithm is used for this purpose due to vast variations in user's taste and culture. Top ranked queries [42] are stored at the first level for faster access to images. Weights are assigned based on the priority of results and then adjusted based on the relevance feedback. Notable algorithms implemented in the relevance feedback are, Gaussian Estimator, genetic algorithm, BM25 algorithm, Bayesian feedback algorithm, images algorithm, Kohonen's learning vector quantization (LVQ) calculation, and tree organized self-sorting out guide (TS-SOM).

For improving performance of content based medical image retrieval system, Rajalakshmi et al, [47], incorporated a relevance feedback procedure using diverse density algorithm. The appropriate texture features have been extracted based on intensity of histogram and run-length features while retrieving brain tumor images. The drawback behind this system is that time complexity is more as it is purely based on a supervised model.

2.3 Social Re-Ranking

The boon of Internet has triggered an increase in number of users on social media. At the same time, they tend to share a huge number of images on social sites. Websites such as Flickr [16] enable clients in appending labels to the pictures, which provides an enhancement to the website or webpage picture recovery [5]. Tag-based picture seek is a significant technique for identification of images given by social clients in social-networking websites.

Primary point is to re-rank pictures by their visual words with semantic data obtained from social information. The underlying outcomes incorporate pictures conveyed by various social clients. Generally, every client contributes few pictures. To begin with, these pictures between clients are dealt with by re-positioning. Users having higher contributions are given higher query rank. At that point, intra-client positioning has been performed on the client's picture set positioned and the most applicable picture from every client's picture set alone is chosen. These extracted pictures intertwine with the end outcome. A rearranged list structure dependent on the social image dataset has been developed for accelerating the searching procedure. Test results on Flickr [16] dataset demonstrate viability and proficiency of social re-ranking technique. These methods can be adapted only for small datasets since the speed will be reduced for large scale datasets.

2.4 Supervised and Unsupervised Learning

To categorize and classify digital images, supervised learning [7] is commonly used. It is mainly dependent on labeled datasets. When there is a huge amount of images, it becomes difficult to manage labeling processes and also in some cases there will be no pre-labeled information. So images need to be classified based on weak semantic correlation, moreover, images within the same classification have to be assigned to relevant midlevel class. Partially labeled pictures are grouped and assigned to the corresponding midlevel classes on the metrics of visual semantics [6]. Thus, newly annotated pictures are taken for learning purposes and this process is continuously repeated until convergence has been achieved. Supervised learning methods may not be suitable for large datasets due to space complexity.

The key idea behind unsupervised CBIR [6] [14] is extracting semantics automatically from noisy associated text and facilitates performance of visual hashing [6]. Hash code learning is framed from a unified unsupervised framework. These hash codes are obtained by preserving visual similarities between images simultaneously [27] from the assistance of relevant text. The foremost assistance models having high order semantic relations construct topic hypergraph, the second one correlates images and then latent shared elements are fetched from collective matrix factorization. Hence, unsupervised is preferred for a huge amount of dataset for fast retrieval.

2.5. Deep Convolutional Neural Networks

Convolutional neural networks (CNN) [28] are based on ANN and used to enable categorization and clustering of images by their similarities and also to perform object recognition within similar images. Query image is converted into grayscale, resized, and then evaluated using trained models. The region of query images is classified based on corresponding labels. These particulars are further used for matching with annotated indices.

As of now, CNN can be applied directly to text analytics and image data analytics with graph convolutional networks [48]. They have been implemented in the form of drones, security mechanics, medical diagnoses, self-driving cars, robotics, and treatments for visually impaired. One of the substances of CNN named Deep learning [49] has become a great potential model for extraction of the features in various domains [50] that include image classification, object detection, and face recognition. Most commonly used deep learning algorithms [51] are deep filter pairing neural network (FPNN) and deep metric learning (DML).

An advanced method for image retrieval is focused on fused deep neural network [52]. It closes semantic gap between low-dimensional and high-dimensional features with the help of Lenet-L, AlexNet, and LeNet-5[53]. Accuracy of the output is more compared with traditional retrieval methods. Deep neural network [54] is categorized into three different architectures namely; deep architecture generation, identification of architecture, and mixed deep architecture. CNN is the best case for the identification of architecture. It learns different image features from different classes. CNN has three layers; convolution layer, pooling layer, and fully connected layer [55]. It is a subcategory of deep neural network, where the network is configured with weights and sigmoidal functions as the activation functions. The main purpose of pooling layer is to give a spatial resolution of input images in addition to reduce computational complexity. Fig-1 shows image retrieval steps through different layers.

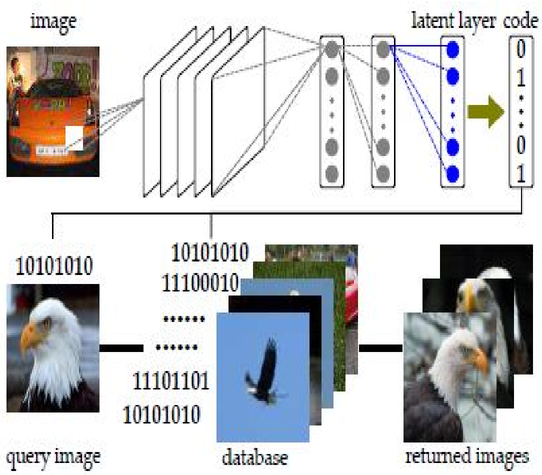

Figure 1: Images passed into hidden layers

CNN is used to encode strong invariance and capture the semantics present in the images because it contains multiple convolutional layers. To provide high-precision embedded images, a deep convolutional neural network is incorporated with different levels of invariance at various scales.

A fully unsupervised retraining model [6] exploits geometrical structure of data and enhances the performance of deep CNN. Further, a CNN pre-trained model is used to produce low dimensional image representations, and improves memory requirements which further enhance retrieval performance. Feature representations have been extracted from activation functions of last convolutional layers using max pooling operations. Weights of fully convolutional CNN [36] models are optimized for the given dataset. This is achieved by minimizing Euclidean distance between images and its “n” nearest representations. Deep CNNs [40] have become the most successful Deep Learning architectures for visual information analysis.

CNN comprises of several convolutional layers along with sub sampling having non-linear neural activation functions. It is then followed by fully connected layers. Input images are fed into neural network as a collection of 3D tensor data where height and weight are similar in dimension to the picture and depth which are equal to number of colored channels (RGB images). 3D filters along with convolutional functions are applied in each layers and output is transferred to neurons of next layer for non-linear transformation with relevant activation functions. Multiple convolution layers and sub sampling process have been performed. The structure of deep architecture normally changes into fully connected layers and also as a single dimensional signal.

The primary approach is to incorporate a deep CNN system to extract feature representations from a pre-trained model by sending images to the input layer and brining out activation function values either from fully connected layers [51] to capture high level semantic information or through convolutional layers utilizing spatial information, using either sum pooling [28] or max-pooling [40] techniques. Recent research focuses on model retraining approaches and a combination of CNN descriptors.

3. COMPARISON OF VARIOUS IMAGE RETRIEVAL METHODS

As discussed, the learning methods based on hashing have the drawback of less-feature ability, high dimensional feature, and low precision of images. So, a deep convolutional neural network [56] [57] [80] [81] is used for overcoming the flaws and for training the images with the high-dimensional features. These visual features are given as input for the hidden layers of the network where hash functions are executed. The output of these hash functions must be error-free and should reduce classification errors.

Hash codes generated are taken as the output from the hash layer and converted into hamming spaces. Profound convolutional highlights obtained are viable for visual undertakings, like, scene grouping, space adjustment, and fine-grained acknowledgment. The limits of profound portrayals are explored in [58], where middle-level portrayals [43] of a prepared CNN are exchanged and two adjustment layers are appended to the highest point of profound highlights for taking in another assignment. It describes the possibility of accomplishment of exchange learning with just the constrained measure of preparing information. This is not like in [58], where the adjustment is just activated in extra layers for arrangement, but it tweaks the whole system for area particular undertakings of protest discovery and division using region based Convolutional Network (R-CNN) [36].

A comparative study of various methods of the content-based image retrieval system is illustrated in Table-1. The different algorithms with its working principles are also discussed.

Table 1 : Comparison of various image retrieval methods

Method	Approach/Algorithm	Discussion
Analysis of CNN for image classification [59]	CIFAR-100 dataset. Comparisons with three datasets, AlexNet, Google LeNet & ResNet	It is based on Machine intelligence for real-time object categorization problems. But the hardware requirements may not get integrate for trained desktop works
Uses visual words for Speeding up robust features (SURF) and incorporates fast retina keypoint (FREAK) feature descriptor [60]	SURF–FREAK, combination of both descriptors to resolve the issues	Performance of this system can be enhanced through the addition of spatial information factors
CBIR depend upon weighted average calculated over triangular histograms using SVM [61]	Based on SVM classification	Not suitable for very large datasets
Image retrieval system depends upon rectangular spatial histograms of visual word [62]	Weighted average of triangular histograms (WATH) of visual words	Mitigates with problem of overfitting, but storage is large as all the histograms of visual words are stored
Connected Components Objects Features For CBIR [63]	An algorithm with the region of interest for image retrieval and also for extraction of texture feature using Gray Level Co-occurrence Matrix	Results in highest Average Precision with 78% accuracy. But shape features are not classified.
Search for E-Commerce based on Deep Learning techniques [64] [86] and large Scale Visual Recommendation and	Street2Shop dataset	A visual recommendation engine is a complex system in the world of e-commerce for any e-retailer.
Neuron Importance Score Propagation [58]	Local response normalization (LRN)	Retrieved from several datasets with multiple CNN models and a significant acceleration are achieved by compressed images with negligible accuracy loss.
Joint Ranking and Regression for Image Enhancement [65]	KNN approach	Quantitative experiments produce a low RMSE with the comparison of the MIT-Adobe data of ground-truth parameters
Based on Multi-modal [84] and multi task feature extraction framework [66]	Large Margin Multi-modal Multi-task Feature Extraction for Image Classification is adopted for handling multi-modal feature extraction	Appropriate for regulated element determination. Not Suitable for unsupervised hashing
Maximum of Multiple Queries & Average of Multiple Queries [67]	Edge histogram descriptor, Local binary patterns, Texture directionality histogram	Suitable for Single group multi queries, Avg –MQ method & EHD feature extraction algorithm produces better results. Not suitable for a large query set
Unsupervised Visual Hashing [6]	To outline high-dimensional similarity images into paired codes of low dimensional Hamming spaces	Suitable for fast query response and low storage consumption. Semantic visual hashing (SAVH) is not suitable for a large data set
Bag-of-visual-words approach. Early and late fusion strategies [68]	Root SIFT feature descriptor	Suitable for object image dataset with a single object per image. It provides better performance than the single query approach. The drawback is that it is suitable only for single semantic images.
Deep Convolutional Neural Networks [7]	Administered semantics-protecting profound hashing (SSDH), develops hash works as an idle layer in a profound system	Appropriate for picture recovery and characterization for a solitary model in a basic and simple manner. It is normally versatile to vast scale sought. Not appropriate for unsupervised display

Query replacement method [69]	Color and Edge directivity descriptor	Utilizing a modest number of question pictures, the high recovery accuracy rate is achieved, yet a high calculation cost amid the run time.
Multi-feature image retrieval method [56]	Content features extraction are more reliable compared to existing algorithms, like DBN, generates huge data set to learn features	It works for classification for efficient content extraction. Can be extended for real-time data extraction
Image classification using CNN [70]	ILSVRC-2012 and ILSVRC-2010 dataset is used	For ILSVRC-2010, error rate is 37.50% top 1 and 17.00% top 5. For ILSVRC-201215, error rate is 3% top 5
Content-based image retrieval using CNN [71]	Paris 6k, INRIA Holidays, UK Bench Oxford 5k	Improved query results for supervised learning. Need to improve the performance with an unsupervised method
Generating descriptions of image regions [72]	MSCOCO, Flickr8K and Flickr30K	Appreciable results found for supervised learning methods
ConvNets, DeepID [73]	Face verification using LFW (Labeled Face in the Wild)	97.45% accuracy for the retrieved images. Should be generalized for all kinds of images
Deep collaborative embedding [74]	NUS-WIDE And MIRFlickr for Social image understanding	Performance of CBIR is 0.512 on MIRFlickr and 0.632 on NUSWID with k = 1000

4. PERFORMANCE EVALUATION CRITERIA

Unsupervised CBIR has been expanded in many different ways and directions due to the volume of uploaded images being cumbersome. This paper discussed the key contributions and techniques used in image retrieval and annotation of images. Some of the primary challenges for adoption of image retrieval systems are reduction in memory consumptions, avoiding duplications [75], and removal of noise. Annotation of the images mainly depends on the perspective, culture, location, and taste of the users. Enormous training has helped machines for learning, understanding, indexing, and annotation of images, over the past decades.

An intelligent system can be used in the manipulation of images like human beings. At present, automated annotation [76] [77] [83] tools like Google Scholar's Search tools are playing a vital role. Text-based search engines, GoogleTM, Yahoo!r, and Bing have been a success, but retrieval of exact images based on the user's requirements is still a difficult task. Search content of query images may be entirely different from the output of traditional image retrieval systems [74]. Research is still in progress, to overcome the issue of combing visual and printed information units into a typical space with their corresponding semantic similarities.

Another major issue is the location of images. Images can be stored with the exact location from which query photo was taken. Of late, the location can be easily achieved with the help of digital cameras, smart mobiles, and satellite images. This helps to figure out robot navigation, mobile landmark recognition, and real-time camera pose tracking. Direct matching methods provide a better localization performance compared with retrieval methods [78] [82], but lead to more memory space. So, inverted files and geometry feature descriptions are used. In this case, the volume of storage space is smaller.

In some of the critical applications like multi-camera tracking [79] [85] and forensic search, a deep ranking algorithm is used. It reduces the cost of disorders from the gallery set like VIPeR, CUHR-01, and CAUIAR4REID [38] [42]. It helps to build the relationship between the input image pair and their semantic scores with the help of joint representations. Various performance evaluation criteria are used for CBIR systems. Some of the commonly used evaluation criteria for performance calculation have been listed below.

4.1 Precision and Recall:

The most commonly used performance evaluation criteria for CBIR research are Precision (P) and recall (R). Precision (P) is always expressed by the ratio of relevant images obtained from the total number of images retrieved (N_{TR}):

$$P = \frac{tp}{N_{TR}} = \frac{tp}{tp + fp}$$

Here, tp is the relevant images retrieved and fp is false positive, meant that the images misclassified as relevant images.

Recall (R) is defined as the ratio of relevant images extracted from the number of relevant images reposted in the database

$$R = \frac{tp}{N_{RI}} = \frac{tp}{tp + fn}$$

Here, tp and N_{RI} are the relevant images retrieved and number of relevant images from the corresponding databases respectively. NRI is retrieved from tp + fn, where fn is false negative, meaning that the images that belong to the relevant class but misclassified as they belong to some other classes.

4.2. Average Precision:

Considering single query *k*, *A*verage precision (AP) is commonly measured after getting mean over precision values of each relevant images:

$$AP = \frac{\sum_{k=1}^{N_{RI}}(P(k) X\, R(k))}{N_{RI}}$$

***4.3. Mean Average Precision**:*

*M*ean average precision (MAP) for a set of queries *S* is equal to the mean of Average Precision values for each query (q). It is represented by

$$MAP = \frac{\sum_{q=1}^{S} AP(q)}{S}$$

Here, *S* is the number of queries.

5. DISCUSSION

In recent years, a lot of research work is being done in the domain of automated vehicles and highway transport systems. The main objective is to make driving safe and accident-free. Vehicles usually lose balance if they face larger potholes or humps. While the motorists slow down the acceleration speed for avoiding the effect of a pothole, there may be enormous chances of collisions and accidents [87] with vehicles following them. The highest priority to get the solutions to avoid these dangerous situations has to be done.

This survey paper aims to propose a novel pothole detection and avoidance mechanism, to assist drivers by indicating humps and potholes over the roads and distributing prior warning messages. The proposed idea is to develop an automatic pothole detection method using deep convolutional neural networks, which is focused on detecting potholes and communicating information to the nearby vehicles. By sharing information, the probability of possible accidents or collisions can be reduced.

These perspectives may be extended further for designing vehicles that are capable of reducing acceleration speed automatically whenever they detect the humps or speed breakers and any other irregularities on the roads. The following approaches may be incorporated.

1. The speed of the vehicles can be controlled or reduced automatically based on sensor values.

2. Balanced Speed control of other vehicles can be done if the details are shared via Google maps.

6. CONCLUSION

A detailed survey of the research related to CBIR with automated annotations has been presented. With the vast development, unsolved problem of semantic relevance can be a solution with the help of deep convolutional neural networks and deep hashing techniques.

Deep learning has become the most acceptable method for solving the memory utilization of visual codes. The reduction of semantic gap in CBIR for the long scale image sets is described. In particular, the frameworks which are relevant to deep learning along with their applications in coordination with CBIR were discussed. The motivation behind automated pothole detection mechanism has been proposed based on deep convolutional neural network.

REFERENCES

[1] Wengang Zhou, Houqiang Li, and Qi Tian,.**Recent advance in content-based image retrieval**: A literature survey.CoRR/abs/1706.06064, http://arxiv.org/abs/1706.06064, 2017.
[2] Mehmood Z., Anwar S., and Altaf M, "**A Survey on Different Image Retrieval Techniques**", *International Research Journal of Engineering and Technology (IRJET),* volume 5, Issue 10, 2018
[3] Chutel, Pushpa M. and Sakhare, Apeksha, **Reverse Image Search Engine using Compact Composite Descriptor**, *International Journal of Advance Research in Computer Science and Management Studies,* Vol. 2 No. 1, pp. 564-570, 2014.
[4] Bo Luo, Xiaogang Wang, Xiaoou Tang, **A World Wide Web Based Image Search Engine Using Text and Image Content Features**. *Proceedings of the SPIE, Volume 5018,* p. 123-130, 2003
[5] M. L. Kherfi, D. Ziou, A. Bernardi, **Image Retrieval from the World Wide Web: Issues, Techniques**, and System.,

ACM Transaction on Computing Surveys, Vol. 36, No. 1, pp. 35–67, 2004

[6] Lei Zhu, JialieShen, Liang Xie, "**Unsupervised Visual Hashing with Semantic Assistant for Content-based Image Retrieval**', *IEEE Transactions On Knowledge And Data Engineering, 2016.*

[7] Huei-Fang Yang, Kevin Lin, and Chu-Song Chen, "**Supervised Learning of Semantics-Preserving Hash via Deep Convolutional Neural Networks**", *IEEE Transactions On Pattern Analysis And Machine Intelligence, 2017.*

[8] X. Li, T. Uricchio, L. Ballan, M. Bertini, C. G. Snoek, and A. D. Bimbo, "**Socializing the semantic gap: A comparative survey on image tag assignment, refinement, and retrieval**" *ACM Comput-ing Surveys (CSUR)*, vol. 49, no. 1, p. 14, 2016.

[9] Flickner M, et al.,**Query by image and video content: the QBIC system**, *IEEE Computer Magazine, Vol. 28 No .9, pp. 23-32,* 1995

[10] A. Gionis, P. Indyk, and R. Motwani, "**Similarity search in high dimensions via hashing**" in *Proceedings of the 25th International Conference on Very Large Data Bases*, pp. 518–529, 1999

[11] R. Datta, D. Joshi, J. Li, and J. Z. Wang,**Image Retrieval: Ideas, Influences, and Trends of the New Age**. *ACM Comput. Survey, 40 (2):1–60*, 2008

[12] Y. Rui, T.S. Huang and S.F. Chang, "**Image Retrieval: Current Techniques, Promising Directions and Open Issues**", *Journal of Visual Comm. and Image Representation, vol. 10, no. 4, pp. 39-62, Apr. 1999.*

[13] Wan, J., Wang, D., Hoi, S. C., Wu, P., Zhu, J., Zhang, Y., Li, J.,**Deep learning for content-based image retrieval: A comprehensive study**. *In: Proceedings of the ACM International Conference on Multimedia*, MM '14, Orlando, FL, USA. pp. 157-166, 2014

[14] Smeulders, A.W.M..et.al, **Content-Based Image Retrieval at the End of the Early Years**, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349-1380,2000

[15] M. J. Huiskes and M. S. Lew, "**The mirflickr retrieval evaluation**" in *ACM international conference on Multimedia information retrieval (MIR)*, pp. 39–43, 2008

[16] M. J. Huiskes and M. S. Lew, "**The mirflickr retrieval evaluation**" *in Proc. ACM Int. Multimedia Information Retrieval (MIR), pp.39–43, 2008*

[17] https://www.wikipedia.org/

[18] Muhammad HammadMemon ; Jianping Li ; Imran Memon ; Qasim Ali Arain ; Muhammad HunainMemon, **Region based localized matching image retrieval system using color-size features for image retrieval**, 14th International Computer Conference on Wavele-t Active Media Technology and Information Processing (ICCWAMTIP), 2017.

[19] Yu, Q., Liu, F., SonG, Y.-Z., Xiang, T., Hospedales, T., Loy, C. C. **Sketch me that shoe.***In: Computer Vision and Pattern Recognition*. pp. 799-807, 2016

[20] *A. Zhang et al., "**Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network**," J. Comput.-Aided Civil Infrastruct. Eng., vol. 32, no. 10, pp. 805–819, 2017*

[21] Ko-Jen Hsiao, Jeff Calder and Alfred O. Hero, ***Pareto-Depth for Multiple-Query Image Retrieval***, IEEE Transactions on image processing, 2013.

[22] A. Alzubi, A. Amira, and N. Ramzan, "**Semantic content-based image retrieval: A comprehensive study**," *Journal of Visual Com-munication and Image Representation*, vol. 32, pp. 20–54, 2015.

[23] van der Walt, S., Schonberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., **the scikit-image contributors, 6, *scikit-image: image processing in Python***. *PeerJ 2, e453*, 2014.

[24] Van den Broek, E. L., van Rikxoort, E. M., Schouten, T. E.,.***Human-Centered Object-Based Image Retrieval****.Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 492-501,*2005.

[25] Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V.,& Garcia-Rodriguez, J. "**A Review on Deep Learning Techniques Applied to Semantic Segmentation**.", IEEE conference on computer vision and pattern recognition, 2017.

[26] J. Wang, S. Kumar, and S.F. Chang, "**Semi-supervised hashing for large-scale search**.," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 34, no. 12, pp. 2393–2406, 2012.

[27] M. Wang, W. Zhou, Q. Tian, Z. Zha, and H. Li, "**Linear distance preserving pseudo-supervised and unsupervised hashing**" in *ACM International Conference on Multimedia (MM)*. ACM, pp. 1257–1266, 2016.

[28] *PengTian-qiang, Li Fang,* Henan Institute of Engineering Department of Computer Science and Engineering, Zhengzhou,China, "**Image Retrieval Based On Deep Convolutional Neural Networks And Binary Hashing Learning**", *IEEE, 2017.*

[29] H. Liu, R. Wang, S. Shan, and X. Chen, "**Deep supervised hashing for fast image retrieval**" in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2064–2072, 2016.

[30] Nadav Ben-Haim, Boris Babenko and Serge Belongie, "I**mproving Web-based Image Search via Content Based Clustering**" CVPRW '06 *Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop, New York, IEEE Press, pp. 106-112,* 2006

[31] Jian Zhang and YuxinPeng, "**SSDH: Semi-supervised Deep Hashing for Large Scale Image Retrieval**", *IEEE Transactions On Circuits And Systems For Video Technology, 2017.*

[32] M. Slaney and M. Casey, "**Locality-sensitive hashing for finding nearest neighbors**," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 128–131, 2008.

[33] A. Gordo, J. Almazan, J. Revaud, and D. Larlus, "**Deep image retrieval: Learning global representations for image search**," in *European Conference on Computer Vision (ECCV)*, 2016.

[34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "**ImageNet classification with deep convolutional neural networks**," in *Advances in neural information processing systems (NIPS)*, pp. 1097–1105, 2012.

[35] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., Fei-Fei, L.,"**ImageNet Large Scale Visual**

Recognition Challenge", *International Journal of Computer Vision (IJCV)* 115 (3), 211-252, 2015

[36] Xuchaa Lu, Li Sang, Rang Xie, Xiaakang Yang, Wenjun Zhang, Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University, Future Medianet Innovation Center, Shanghai, China, IEEE, "**Deep Hash Learning For Efficient Image Retrieval**", 2017

[37] J. Yu, D. Tao, M. Wang, and Y. Rui,"**Learning to rank using user clicks andvisual features for image retrieval**," *IEEE Transactions on Cybernetics*, vol. 45, no. 4, pp. 767–779, 2015.

[38] Hanjiang Lai, Pan Yan, XiangboShu,Yunchao Wei, Shuicheng Yan, Senior Member, IEEE, "**Instance-Aware Hashing for Multi-Label Image Retrieval**", *IEEE Transactions On Image Processing, 2016.*

[39] F. Yang, B. Matei, and L. S. Davis, "**Re-ranking by multi-feature fusion with diffusion for image retrieval**," in *IEEE Winter Confer-ence on Applications of Computer Vision (WACV).*IEEE, pp.572–579, 2015.

[40] K. Lin, H.-F. Yang, J.-H.Hsiao, and C.-S. Chen, "**Deep learning of binary hash codes for fast image retrieval**," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 27–35, 2015.

[41] Jun Yu, *Member, IEEE*, Xiaokang Yang, *Senior Member, IEEE*, FeiGao, and Dacheng Tao, *Fellow, IEEE, s* "**Deep Multimodal Distance Metric Learning Using Click Constraints for Image Ranking**", *IEEE Transactions On Cybernetics, VOL. 47, NO. 12, December 2017.*

[42] AneliaGrigorova, Francesco G. B. De Natale, Senior Member, IEEE, Charlie Dagli, and Thomas S. Huang, Life Fellow, IEEE, "**Content-Based Image Retrieval by Feature Adaptation and Relevance Feedback**", *IEEE Transactions On Multimedia, 2007.*

[43] Basura Fernando and TinneTuytelaars**mining multiple queries for image retrieval: On-the-fly learning of an object-specific mid-level representation**, IEEE International conference on computer vision, 2013.

[44] Ko-Jen Hsiao, Jeff Calder and Alfred O. Hero, ***Pareto-Depth for Multiple-Query Image Retrieval***, IEEE Transactions on image processing, 2013.

[45] Raghuwanshi, Ghanshyam, Nishchol Mishra, and Sanjeev Sharma. "**Content based image retrieval using implicit and explicit feedback with interactive genetic algorithm**" *International Journal of Computer Applications 43.16: 8-14, 2012.*

[46] S. D. MacArthur, C. E. Brodley, A. C. Kak, and L. S. Broderick. **Interactive Content-Based Image Retrieval Using Relevance Feedback.*Computer Vision and Image Understanding***, 88(2):55–75, 2002.

[47] Rajalakshmi, T., and R. I. Minu. "**Improving relevance feedback for content based medical image retrieval**" Information Communication and Embedded Systems (ICICES), 2014 *International Conference on. IEEE*, February 2014.

[48] H. Liu, R.Wang, S. Shan, and X. Chen, "**Deep supervised hashing for fast image retrieval**," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 2064–2072, 2016

[49] Y. Lecun, Y. Bengio, and G. Hinton, "**Deep learning**," *Nature*, vol. 521, pp. 436–44, 2015.

[50] Ushizima, D. M., Bale, H. A., Bethel, E. W., Ercius, P., Helms, B. A., Krishnan, H., Grinberg, L. T., Haranczyk, M., Macdowell, A. A., Odziomek, K., Parkinson, D. Y., Perciano, T., Ritchie, R. O., Yang, C.Ideal: **Images across domains, experiments, algorithms and learning**. *The Journal of The Minerals, Metals & Materials Society,* 1-10, Sep 2016.

[51] Peizhong Liu, Jing-Ming Guo, *Senior Member, IEEE*, Chi-Yi Wu, and DanlinCai, "**Fusion of Deep Learning and Compressed Domain Features for Content-Based Image Retrieval**", *IEEE 'Transactions On Image Processing, VOL. 26, NO. 12, December 2017.*

[52] F. Sabahi, *Member*, IEEE, M. Omair Ahmad, *Fellow*, IEEE, and M.N.S. Swamy, *Fellow*, IEEE, "**An Unsupervised Learning Based Method for Content-based Image Retrieval using Hopfield Neural Network**", 2017

[53] R Yu, A Li, CF Chen, JH Lai, VI Morariu, X Han, M Gao, CY Lin, LS Davis, CVPR: *NISP: **Pruning Networks using Neuron Importance Score Propagation**, 2018.*

[54] D.-H. Lee, "**Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks**," in *Workshop on Challenges in Representation Learning, ICML*, vol. 3, pp. 2–7, 2013

[55] G. Tolias, R. Sicre, and H. J´egou, "**Particular object retrieval with integral max-pooling of CNN activations**," *International Conference on Learning and Representation (ICLR)*, 2016.

[56] R. Rani Saritha · Varghese Paul · P. Ganesh Kumar, "**Content based image retrieval using deep learning process**", *Springer Science, Business Media, LLC, part of Springer Nature, 2018*.

[57] AnshumanVikram Singh. "**Content-Based Image Retrieval using Deep Learning**". *A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science, in B. Thomas Golisano College of Computing and Information Sciences Rochester Institute of Technology Rochester, New York, June 2015.*

[58] *Peng Li and PengRen, ,* "**Partial Randomness Hashing for Large-Scale Remote Sensing Image Retrieval**", *IEEE Geoscience And Remote Sensing Letters, Vol. 14, No. 3, March 2017.*

[59] NehaSharma,Vibhor Jain, Anju Mishra , "**An Analysis Of Convolutional Neural Networks For Image Classification**", *Science Direct, Elsevier, International conference on convolutional neural networks for image classification, 2018*

[60] Jabeen S, Mehmood Z, Mahmood T, Saba T, Rehman A, Mahmood MT, **"An effective content-based image retrieval technique for image visuals representation based on the bag-of-visual-words model",** *PLoS ONE 13(4): e0194526, 2018*

[61] Mehmood Z., Mahmood T., and Javid M.A., "**Content-based image retrieval and semantic automatic image annotation based on the weighted average of triangular histograms using support vector machine**", *Applied Intelligence, 48(1): p. 166–181, 2018*

[62] Mehmood Z., Anwar S., and Altaf M., "**Novel image retrieval based on rectangular spatial histograms of visual words**", *Kuwait Journal of Science*,45(1): p. 54–69, 2018.
[63] FaiqBaji, Mihas Mocanu, "**Connected components objects features for CBIR**", *IEEE, 18th International Carpathian Control Conference (ICCC), 2017*
[64] Devashish Shankar, SujayNarumanchi, Ananya H A, PramodKompalli, KrishnenduChaudhury, "**Deep Learning based Large Scale Visual Recommendation and Search for E-Commerce**", IEEE conference on Computer vision and pattern recognition, 2017.
[65] Parag S. Chandakkar and Baoxin Li, "**Joint Regression and Ranking for Image Enhancement**", in IEEE WACV, 2017.
[66] Yong Luo, YonggangWen, Dacheng Tao, JieGui, Member, and Chao Xu, "**Large Margin Multi-Modal Multi-Task Feature Extraction for Image Classification**", *IEEE Transactions On Image Processing, Vol. 25, No. 1, January 2016*
[67] Elena Stojanova, Katarina Trojacanec, IvicaDimitrovski and SuzanaLoshkovska, ***Multi-query content based medical image retrieval***, the 10th conference for Informatics and Information Technology, 2013.
[68] Fatihcalisir, MuhammetBastan, UgurGudukbay and OzgurUlusoy, ***Mobile image search using multi-query images***, IEEE conference on signal processing and communication application, 2015
[69] Vimina E R, Ramakrishnan K, NavyaNandakumar and Poulose Jacob K "**A*n efficient multi query system for content based image retrieval using query replacement*"**, 16th IEEE conference on Software engineering, artificial intelligence, networking and parallel/distributed computing, 2015.
[70] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "**ImageNet classification with deep convolutional neural networks**,"*Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2012.
[71] C. Zhang, J. Cheng, and Q. Tian, "**Semantically modeling of object and context for categorization**," *IEEE Transactions onNeural Networks and Learning Systems*, vol. 30, no. 4, pp. 1013–1024, 2018.
[72] A. Karpathy and L. Fei-Fei, "**Deep visual-semantic alignments for generating image descriptions**," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3128–3137, Boston, MA, USA, June 2015.
[73] Y. Sun, X. Wang, and X. Tang, "**Deep learning face representation from predicting 10,000 classes**," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1891–1898, Columbus, OH, USA, June 2014.
[74] Z. Li, J. Tang, and T. Mei, "**Deep collaborative embedding for social image understanding**," *IEEE Transactions on PatternAnalysis and Machine Intelligence*, vol. 41, no. 9, pp. 2070–2083, 2018.
[75] L. Paulev´e, H. J´egou, and L. Amsaleg, "**Locality sensitive hashing: A comparison of hash function types and querying mechanisms**," *Pattern Recognition Letters*, vol. 31, no. 11, pp. 1348–1358, 2010
[76] J. R. Smith and S. F. Chang.VisualSEEk: **A fully automated content-based image query system**. In *Proceedings of the ACM Multimedia*, pages 87–98, Boston, MA, November 1996.ack
[77] A. Fischer and C. Igel, "**Training restricted Boltzmann machines: An introduction**," *Pattern Recognition*, vol. 47, no. 1, pp. 25–39, 2014.
[78] Yanzhi Chen, Xi Li, Anthony Dick and Anton Van den Hengel, ***Boosting object retrieval with group queries IEEE Signal Processing Letters****, Vol. 19, No. 11*, 2012
[79] Cai-Zhi Zhu, Yu-Huii Huang and Shin'ichi Satoh, ***Multi-image aggregation for better visual object retrieval***,IEEE International conference on acoustic, speech and signal processing, 2014.
[80] J. Weston, F. Ratle, H. Mobahi, and R.Collobert, "**Deep learning via semi-supervised embedding**," in *International Conference on Machine Learning (ICML)*, pp. 1168–1175, 2008
[81] Zeng, Y., Xu, X., Fang, Y., Zhao, K..Tra_c,**Sign Recognition Using Deep Convolutional Networks and Extreme Learning Machine.*Springer International Publishing***, *Cham*, pp. 272-280, 2015
[82] Sivasanker.N. P, AshishAnand, Nishant Raj and R. Sathya, **Lane Keeping System Using Real Time Pathway Detection**, *International Journal of Future Generation Communication and Networking Vol. 13, No. 3, pp. 42 -47 , 2020*
[83] R. Sathya, Abhijit Kumar Sanu, Avinash Singh, Saurav Chaurasia, Vishal Agrawal "**Detection of Manipulated Images using Convolution Neural Network**", *International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, Vol-8 No-4, pp 1431 - 1438 April 2019.*
[84] N Poornima, B Saleena, **Multi-modal features and correlation incorporated Naive Bayes classifier for a semantic-enriched lecture video retrieval system**, *The Imaging Science Journal 66 (5), 263-277, 2018*
[85] R.Sathya, Abhishek Kumar, Ninad Mathpati, Hans Raj, Suraj Kumar **"Vehicle to Vehicle Communication and Collision Avoidance with Auto breaking System",** Journal of Network Communications and Emerging Technologies (JNCET), 2395-5317, pp 343-345, April 2018
[86] Madala Gayathri , Murala Meghana , Madala Trivedh ,D Manju, "**Suspicious Activity Detection and Tracking through Unmanned Aerial Vehicle Using Deep Learning Techniques**", International Journal of Advanced Trends in Computer Science and Engineering, 2278-3091, Vol 9 No 3, pp 2812-2816, 2020
[87] Tirumuru Ketha , S Sagar Imambi, "**Analysis of Road Accidents to Indentify Major Causes and Influencing Factors of Accidents – A Machine Learning Approach**", International Journal of Advanced Trends in Computer Science and Engineering, 2278-3091, Vol- 8 No 6, pp 3492-3497, 2019