International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

SEMANTIC CROSS MEDIA HASHING TO CAPTURE TEXT SIMILARITY AT THE SEMANTIC LEVEL

AAREPU LASKHMAN[1], MOHAMMED INAYATHULLA[2], N.GRACE NAOMI[3], VISWANADHA LAKSHMI MOUNICA[4], RITU NEELAM[5], LINGALA RAJA SHEKAR REDDY[6]

[1] Assistant Professor, Department of CSE, Malla Reddy Engineering College (Autonomous), Hyderabad, India

[2] Assistant Professor, Department of CSE, Malla Reddy Engineering College (Autonomous), Hyderabad, India

[3] Assistant Professor, Department of CSE, Malla Reddy Engineering College (Autonomous), Hyderabad, India

[4,5,6] UG Student, Department of CSE, Malla Reddy Engineering College (Autonomous), Hyderabad, India

Abstract:

Hashing methods have proven to be useful for a variety of tasks and have attracted extensive attention in recent years. Various hashing approaches have been proposed to capture similarities between textual, visual, and cross-media information. However, most of the existing works use a bag-of-words method to represent textual information. Since words with different forms may have similar meaning, semantic level text similarities cannot be well processed in these methods. To address these challenges, in this paper, we propose a novel method called semantic cross-media hashing (SCMH), which uses continuous word representations to capture the textual similarity at the semantic level and use a deep belief network (DBN) to construct the correlation between different modalities. To demonstrate the effectiveness of the proposed method, we evaluate the proposed method on three commonly used cross-media data sets are used in this work. Experimental results show that the proposed method achieves significantly better performance than state-of-the-art approaches. Moreover, the efficiency of the proposed method is comparable to or better than that of some other hashing method

I.INTRODUCTION

WITH the rapid expansion of the World Wide Web, digital information has become much easier to access, modify, and duplicate. Hence, hashing based similarity calculation or approximate nearest neighbour searching methods have been proposed and received considerable attention in recent years. Various applications such as information retrieval, near duplicate detection, and data mining are performed by hashing based methods. Due to the rapid expansion of mobile networks and social media sites, information input through multiple channels has also attracted increasing attention. Images and videos are associated with tags and captions. According to research published on eMarketer, about 75 percent of the content posted by

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Facebook users contains photos.1 The relevant data from different modalities usually have semantic correlations. Therefore, it is desirable to support the retrieval of information through different modalities. For example, images can be used to find semantically relevant textual information. On the other side, images without (or with little) textual descriptions are highly needed to be retrieved with textual query.Along with the increasing requirements, in recent years, cross-media search tasks have received considerable attention [1], [2], [3], [4], [5], [6], [7]. Since each modality having different representation methods and correlational structures, a variety of methods studied the problem from the aspect of learning correlations between different modalities. Existing methods proposed to use Canonical Correlation Analysis (CCA) [8], manifolds learning [9], dual-wing harmoniums [10], deep autoencoder [11], and deep Boltzmann machine [12] to approach the task. Due to the efficiency of hashing-based methods, there also exists a rich line of work focusing the problem of mapping multi-modal high-dimensional data to low-dimensional hash codes, such as Latent semantic sparse hashing (LSSH) [13], discriminative coupled dictionary hashing (DCDH) [14], Cross-view Hashing (CVH) [15], and so on.

Most of the existing works use a bag-of-words to model textual information. The semantic level similarities between words or documents are rarely considered. Let us consider the following examples:

S1. The company announces new operating system.

S2. The company releases new operating system.

S3. The company delays new operating system.

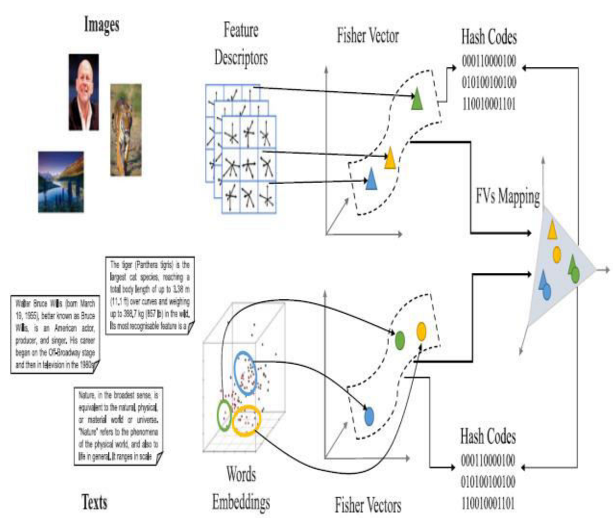
From these examples, we can observe that although only one word differs between the three sentences, sentence S3 should not be considered as the near duplicate sentence of S1 and sentence S2. The meaning expressed by S3 is much different with S1 and S2's. Since existing methods are usually based on lexical level similarities, this kind of issue cannot be well addressed by these methods.

In short text segments (e.g., microblogs, captions, and tags), the similarities between words are especially important for retrieval. For example: journey versus travel, coast versus shore. According to human-assigned similarity judgments [16], more than 90 percent of subjects thought that these pairs of words had similar meanings. Fig. 1 illustrates a set of images retrieved from Flickr using different queries. From these examples, we can see that images may express similar concepts, even though there is little overlap in terms of annotated tags. Since users rarely annotate a single image using multiple words with similar meaning, semantic level textual similarities should be incorporated into the crossmedia retrieval.

Motivated by the success of continuous space word representations (also called word embeddings) in a variety of tasks, in this work we propose to incorporate word embeddings to meet these challenges. Words

in a text are embedded in a continuous space, which can be viewed as a Bag-of-Embedded-Words (BoEW). Since the number of words in a text is dynamic, in [17], we proposed a method
to aggregate it into a fixed length Fisher Vector (FV), using a Fisher kernel framework [18]. However, the proposed method only focus on textual information. Another challenge in this task is how to determine the correlation between multi-modal representations. Since we propose the use of a Fisher kernel framework to represent the textual information, we also use it to aggregate the SIFT descriptors [19] of images. Through the Fisher kernel framework, both textual and visual information is mapped to points in the gradient space of a Riemannian manifold. However, the relationships that exist between FVs of different modalities are usually highly non-linear. Hence, to construct the correlation between textual and visual modalities, we introduce a DBN based method to model the mapping function, which is used to convert abstract representations of different modalities from one to another.
The main contributions of this work are summarized as
follows.
_ We propose to incorporate continuous word representations to handle semantic textual similarities and adopted for cross-media retrieval.
_ Inspired by the advantages of DBN in handling highly non-linear relationships and noisy data, we introduce a novel DBN based method to construct the correlation between different modalities.
_ A variety of experiments on three cross-media commonly used benchmarks demonstrate the effectiveness of the proposed method. The experimental results show that the proposed method can significantly outperform the state-of-the-art methods.

II.SYSTEM ARCHITECTURE:



III.EXISTING SYSTEM

Along with the increasing requirements, in recent years, cross-media search tasks have received considerable attention. Since, each modality having different representation methods and correlational structures, a variety of methods studied the problem from the aspect of learning correlations between different modalities. Existing methods proposed to use Canonical Correlation Analysis (CCA), manifolds learning, dual-wing harmoniums, deep autoencoder, and deep Boltzmann machine to approach the

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

task. Due to the efficiency of hashing-based methods, there also exists a rich line of work focusing the problem of mapping multi-modal high-dimensional data to low-dimensional hash codes, such as Latent semantic sparse hashing (LSSH), discriminative coupled dictionary hashing (DCDH), Cross-view Hashing (CVH), and so on.

Disadvantages of Existing System:

1. Most of the existing works use a bag-of-words to model textual information. The semantic level similarities between words or documents are rarely considered.
2. Existing works focused only on textual information.
3. Another challenge in this task is how to determine the correlation between multi-modal representations.

IV.PROPOSED SYSTEM

We propose a novel hashing method, called semantic cross-media hashing (SCMH), to perform the near-duplicate detection and cross media retrieval task. We propose to use a set of word embeddings to represent textual information. Fisher kernel framework is incorporated to represent both textual and visual information with fixed length vectors. For mapping the Fisher vectors of different modalities, a deep belief network is proposed to perform the task. We evaluate the proposed method SCMH on three commonly used data sets. SCMH achieves better results than state-of-the-art methods with different the lengths of hash codes.

Advantages of Proposed System:

1. We introduce a novel DBN based method to construct the correlation between different modalities.
2. The proposed method can significantly outperform the state-of-the-art methods.

V.IMPLEMENTATION:

- **Admin**

  In this module, the Admin has to login by using valid user name and password. After login successful he can do some operations such as upload images, view uploaded images, view the searching history, view all image ranking and view all users, search images and logout.

  Search History

  This is controlled by admin; the admin can view the search history details. If he clicks on search history button, it will show the list of searched user details with their tags such as user name, user searched for image name, time and date.

  Rank of images

  In user's module, the admin can view the list of ranking images. If admin click on list of ranking images, then the server will give response with their tags image and rank of image.

  Upload Images

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

In this module, the admin can upload n number of images with hash code and the Hash code will be generated based on the image name and contents. If the image is same and then the hash code will be similar for all corresponding images. Admin want to upload new image then he has enter some fields like image name, image color, image description, image type, image usage, browse the image file and upload. After uploading successfully he will get a response from the server. Initially new uploaded image rank is zero. After viewing that image rank will re-rank.

- **User**

In this module, there are n numbers of users are present. User should register before doing some operations. And register user details are stored in user module. After registration successful he has to login by using authorized user name and password. Login successful he will do some operations like view my details, search images based on Hashing code of the group of images, request secrete key and logout. The user click on my details link then the server will give response to the user with all details such as user name, phone no, address, e mail ID and location. An end user can search the images based on **Images based on keyword or hash code** and gets the details like image name, image color, image usage and image type. And server will give response to the user, then that image rank will be increased.

VI.CONCLUSION:

In this work, we propose a novel hashing method, SCMH, to perform the near-duplicate detection and cross media retrieval task. We propose to use a set of word embeddings to represent textual information. Fisher kernel framework is incorporated to represent both textual and visual information with fixed length vectors. For mapping the Fisher vectors of different modalities, a deep belief network is proposed to perform the task. We evaluate the proposed method SCMH on three commonly used data sets. SCMH achieves better results than state-of-the-art methods with different the lengths of hash codes. In NUS-WIDE data set, the relative improvements of SCMH over LSSH, which achieves the best results in these datasets, are 10.0 and 18.5 percent on the Text ! Image and Image ! Text tasks respectively. Experimental results demonstrate the effectiveness of the proposed method on the cross-media retrieval task.

VII.REFERENCES:

[1] G. Ding, Y. Guo, and J. Zhou, “Collective matrix factorization hashing for multimodal data,” in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2014, pp. 2083–2090.

[2] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin, “Iterative quantization: A procrustean approach to learning binary

codes for large-scale image retrieval," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 12, pp. 2916–2929, Dec. 2013.

[3] Y. Pan, T. Yao, T. Mei, H. Li, C.-W. Ngo, and Y. Rui, "Clickthrough- based cross-view learning for image search," in Proc. 37th Int.ACMSIGIR Conf. Res. Develop. Inf. Retrieval, 2014, pp. 717–726.

[4] J. Song, Y. Yang, Y. Yang, Z. Huang, and H. T. Shen, "Inter-media hashing for large-scale retrieval from heterogeneous data sources," in Proc. Int. Conf. Manage. Data, 2013, pp. 785–796.

[5] D. Zhai, H. Chang, Y. Zhen, X. Liu, X. Chen, and W. Gao, "Parametric local multimodal hashing for cross-view similarity search," in Proc. 23rd Int. Joint Conf. Artif. Intell., 2013, pp. 2754–2760.

[6] D. Zhang and W.-J. Li, "Large-scale supervised multimodal hashing with semantic correlation maximization," in Proc. 28th AAAI Conf. Artif.Intell., 2014, pp. 2177–2183.

[7] Y. Zhuang, Y. Yang, F. Wu, and Y. Pan, "Manifold learning based cross-media retrieval: A solution to media object complementary nature," J. VLSI Signal Process. Syst. Signal, Image Video Technol., vol. 46, pp. 153–164, 2007.

[8] F. Wu, H. Zhang, and Y. Zhuang, "Learning semantic correlations for cross-media retrieval," in Proc. IEEE Int. Conf. Image Process., 2006, pp. 1465–1468.

[9] Y. Yang, Y.-T.Zhuang, F. Wu, and Y.-H. Pan, "Harmonizing hierarchical manifolds for multimedia document semantics understanding and cross-media retrieval," IEEE Trans. Multimedia, vol. 10, no. 3, pp. 437–446, Apr. 2008.

[10] E. P. Xing, R. Yan, and A. G. Hauptmann, "Mining associated text and images with dual-wing harmoniums," in Proc. 21st Conf. Uncertainty Artif. Intell., 2005, pp. 633–641.

[11] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal deep learning," in Proc. 28th Int. Conf. Mach. Learn., 2011, pp. 689–696.

[12] N. Srivastava and R. Salakhutdinov, "Multimodal learning with deep boltzmann machines," in Proc. Adv. Neural Inf. Process. Syst., 2012, pp. 2222–2230.

[13] J. Zhou, G. Ding, and Y. Guo, "Latent semantic sparse hashing for cross-modal similarity search," in Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2014, pp. 415–424.

[14] Z. Yu, F. Wu, Y. Yang, Q. Tian, J. Luo, and Y. Zhuang, "Discriminative coupled dictionary hashing for fast cross-media retrieval," in Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2014, pp. 395–404.

[15] S. Kumar and R. Udupa, "Learning hash functions for cross-view similarity search," in Proc. Int. Joint Conf. Artif.Intell., 2011, pp. 1360–1365.

[16] L. Finkelstein, E. Gabrilovich, Y. Matias, E. Rivlin, Z. Solan, G. Wolfman, and E. Ruppin, "Placing search in context: The concept revisited," in Proc. 10th Int. Conf. World Wide Web, 2001, pp. 406–414.

[17] Q. Zhang, J. Kang, J. Qian, and X. Huang, "Continuous word embeddings for

detecting local text reuses at the semantic level," in Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2014, pp. 797–806.

[18] T. Jaakkola, D. Haussleret al., "Exploiting generative models in discriminative classifiers," in Proc. Adv. Neural Inf. Process. Syst., 1999, pp. 487–493.

[19] D. G. Lowe, "Object recognition from local scale-invariant features," in Proc. Int. Conf. Comput. Vis., 1999, p. 1150. [20] X. Wang, Y. Liu, D. Wang, and F. Wu, "Cross-media topic mining on wikipedia," in Proc. 21st ACM Int. Conf. Multimedia, 2013, pp. 689–692.

[21] H. Zhang, J. Yuan, X. Gao, and Z. Chen, "Boosting cross-media retrieval via visual-auditory feature analysis and relevance feedback," in Proc. ACM Int. Conf. Multimedia, 2014, pp. 953–956.

[22] P. Daras, S. Manolopoulou, and A. Axenopoulos, "Search and retrieval of rich media objects supporting multiple multimodal queries," IEEE Trans. Multimedia, vol. 14, no. 3, pp. 734–746, Jun. 2012.

[23] A. Karpathy and L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Boston, MA, USA, Jun. 2015, pp. 3128–3137.

[24] A. Z. Broder, "On the resemblance and containment of documents," in Proc. SEQUENCES, 1997, p. 21.

.

[25] A. Z. Broder, "Identifying and filtering near-duplicate documents," in Proc. Combinatorial Pattern Matching, 2000, pp. 1–10.

[26] M. Theobald, J. Siddharth, and A. Paepcke, "Spotsigs: Robust and efficient near duplicate detection in large web collections," in Proc. 31st Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2008, pp. 563–570.

[27] A. Chowdhury, O. Frieder, D. Grossman, and M. C. McCabe, "Collection statistics for fast duplicate document detection," ACM Trans. Inf. Syst., vol. 20, no. 2, pp. 171–191, 2002.

[28] A. Ko»cz, A. Chowdhury, and J. Alspector, "Improved robustness of signature-based near-replica detection via lexicon randomization," in Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2004, pp. 605–610.

[29] J. Seo and W. B. Croft, "Local text reuse detection," in Proc. 31st Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2008, pp. 571–578.

[30] Q. Zhang, Y. Zhang, H. Yu, and X. Huang, "Efficient partialduplicate detection based on sequence matching," in Proc. 31st Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2010, pp. 571–578.

[31] J. W. Kim, K. S. Candan, and J. Tatemura, "Efficient overlap and content reuse detection in blogs and online news articles," in Proc. Int. Conf. World Wide Web, 2009, pp. 571–578