FUNDAMENTAL VISUAL CONCEPT LEARNING FROM CORRELATED IMAGES AND TEXT ANNOTATION

Mr. B.S Murthy [1], G.Sowmya [2],

[1]**Assistant professor , MCA DEPT,** Dantuluri Narayana Raju College**, Bhimavaram, Andharapradesh**
Email: suryanarayanamurthy.b@gmail.com
[2]**PG Student of MCA, Dantuluri** Narayana Raju College**, Bhimavaram, Andharapradesh**
Email: sowmyagonaboina03@gmail.com

ABSTRACT

Heterogeneous web media consists of many visual concepts, such as objects, scenes and activities, that cannot be semantically decomposed. The task of learning fundamental visual concepts (FVCs) plays an important role in automatically understanding the elements that compose all visual media, as well as in applications of retrieval, annotation, etc. In this paper, we formulate the problem of FVC learning and propose an approach to this problem called neighboring concept distributing (NCD). Our approach models all data using a concept graph, which considers the visual patches in images as nodes and generates the inter-image edges between visual patches in different images and the intra-image edges between visual patches in the same image. The NCD approach distributes semantic information from images to visual patches based on measurements over the concept graph, including fitness, distinctiveness, smoothness and sparseness, without any pre-trained concept detectors or classifiers. We analyze the learn ability of the proposed approach and find that, under some conditions, all concepts can be correctly learned with an arbitrarily high probability as the size of the data increases .We demonstrate the performance of the NCD approach using three public datasets. Experimental results show that our approach outperforms state-of-the-art approaches when learning visual concepts from correlated media.

1 INTRODUCTION

The worldwide web is full of images, videos, audio, and text, which are not only growing rapidly in terms of quantity but are also becoming increasingly rich in terms of content. Heterogeneous web data usually coexist in multimedia documents and use similar semantics to describe the same subject from different perspectives. The various modalities of documents may be complementary in terms of expressing the semantics of content.

For example, an image can vividly inspire imagination but incompletely describe a concept. In contrast, while text can accurately describe the details of a concept, it is not intuitive enough. Currently, many real-world Internet applications involve multi-modal data, such as cross media retrieval, image tagging, multimedia searching and multimedia caption generation. Common to these applications, the relations between different modalities need to be considered and learned at the document level granularity under the supervision of labeled data.

2 RELEATED WORK

Literature survey is the most important stepin software development process. Before developing the tool, it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, ten next steps are to determine which operating system and language used for developing the tool. Once the programmers start building thetool, the programmers need lot of external support.

This support obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into for developing the proposed system.

3 IMPLEMENTATION STUDY

Existing System:

Jiang and Tan directly built a matrix transformation between visual and textual features through a set of predefined domain-specific information categories to discover the underlying associations between images and text. Du et al. presented a mixture of local linear mappings for modeling the complex semantic correlation between images and text. To maximize the correlation between two modalities, canonical correlation analysis (CCA), which linearly projects the inputs of two modalities into a subspace where the corresponding examples are generally close together, has been widely used in cross-media modeling.

- In the existing work, the media streaming is very slow with effective less techniques.
- The system is less effective due to lack of high visual concepts.

PROPOSED SYSTEM & ALOGIRTHAM

The system proposes an approach named neighboring concept distributing (NCD) to address this task. In this work, visual patches and text descriptions are represented based on deep networks and one-hot vectors, respectively. The NCD approach models all data using a concept graph, which considers the visual patches as nodes and generates inter image edges between visual patches belonging to different images and intra-image edges between visual patches in the same image.

4.1 Advantages:

- The system is more effective in image annotation since it is using Fundamental visual concept.
- The system is very effective since it is allowing heterogeneous media.

IMPLEMENTATION

4.1 MODULES

ADMIN WEB SERVER

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 Add images, View all images with its annotation, View all images ranking and its annotation, View all image details with annotation, View all image with its annotation by clicking on the images, List Users & authorize, View all images with reviews and ratings

END USER

In this module, there are n numbers of users are present. User should register with group option before doing some operations. After registration successful he has to wait for admin to authorize him and after admin authorized him. He can login by using authorized user name and password. Login successful he will do some operations like View Own Details, Search for images based on contents and annotation keyword, View my search History search images based key points of annotation and review the image.

IMAGE ANNOTATION

In general, the task of the image annotation is to predict multiple textual labels that describe the content or visual appearance of an unseen image. In addition, a few studies focused on predicting the labels of the regions in images.

5 RESULTS AND DISCUSSION

2 SCREEN SHOTS

5.2.1 Home page with project concept

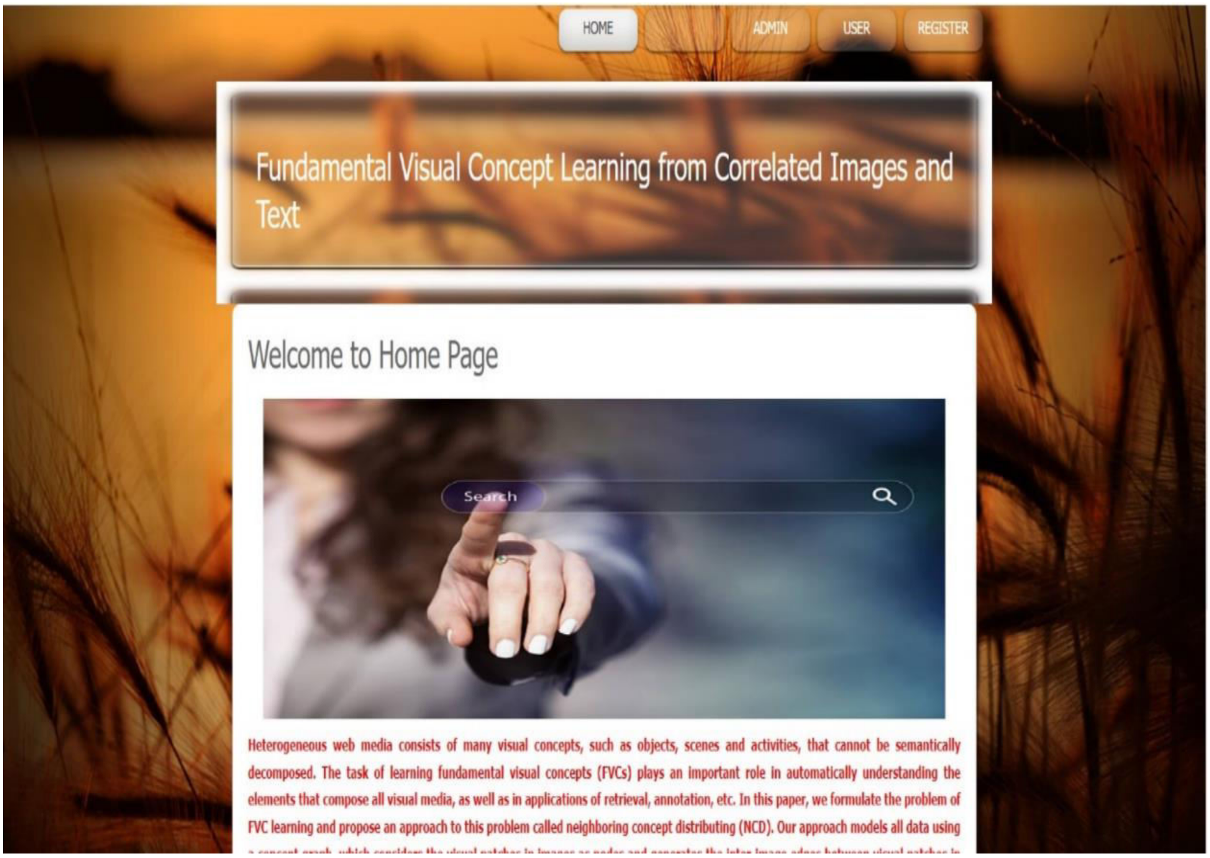

FIG:5.2.1 Home page with project concept- SCREEN SHORT

5.2.2 User Registration page

FIG:5.2.2 User registration page-SCREEN SHOT

5.2.3 Admin main page

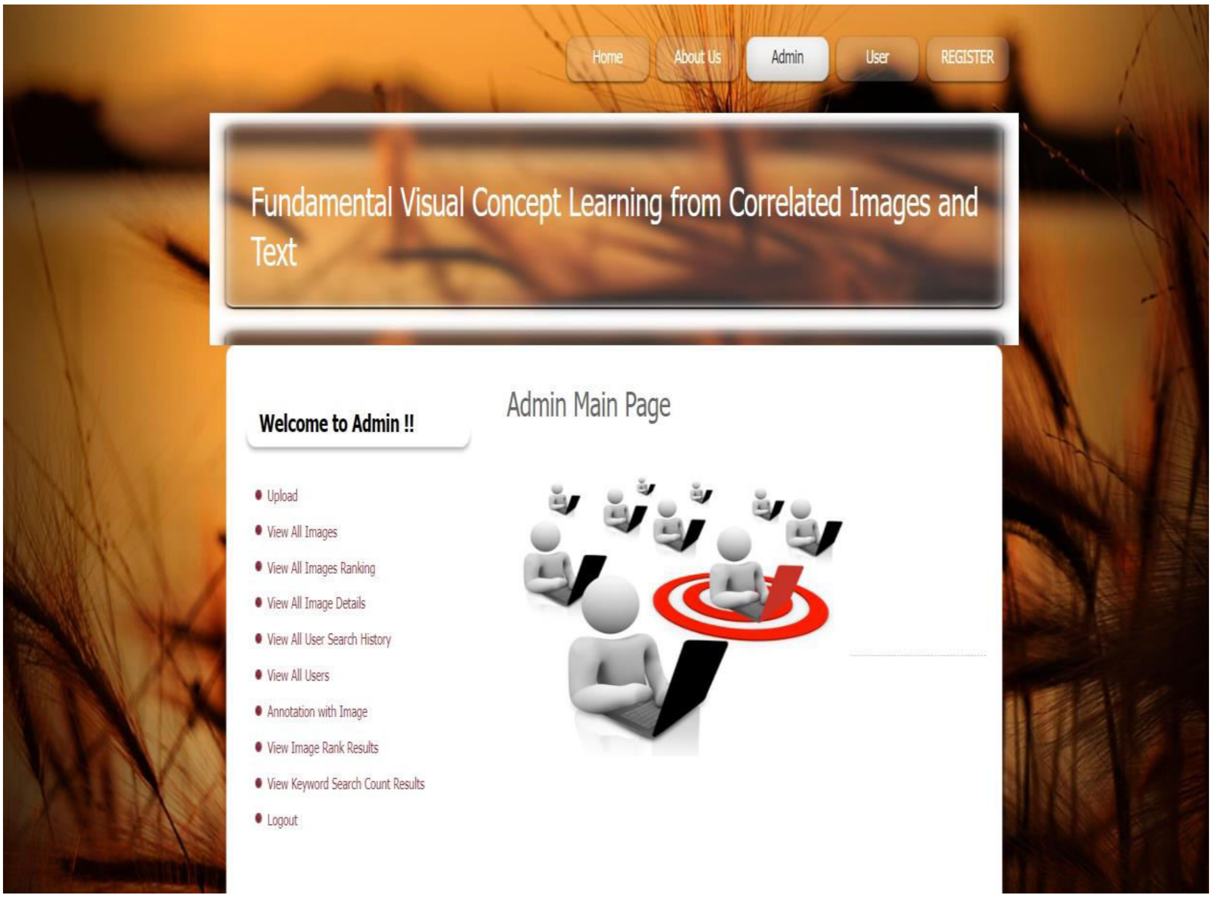

FIG: 5.2.3 Admin main page – SCREEN SHOT

5.3.4 Amin can view all user details

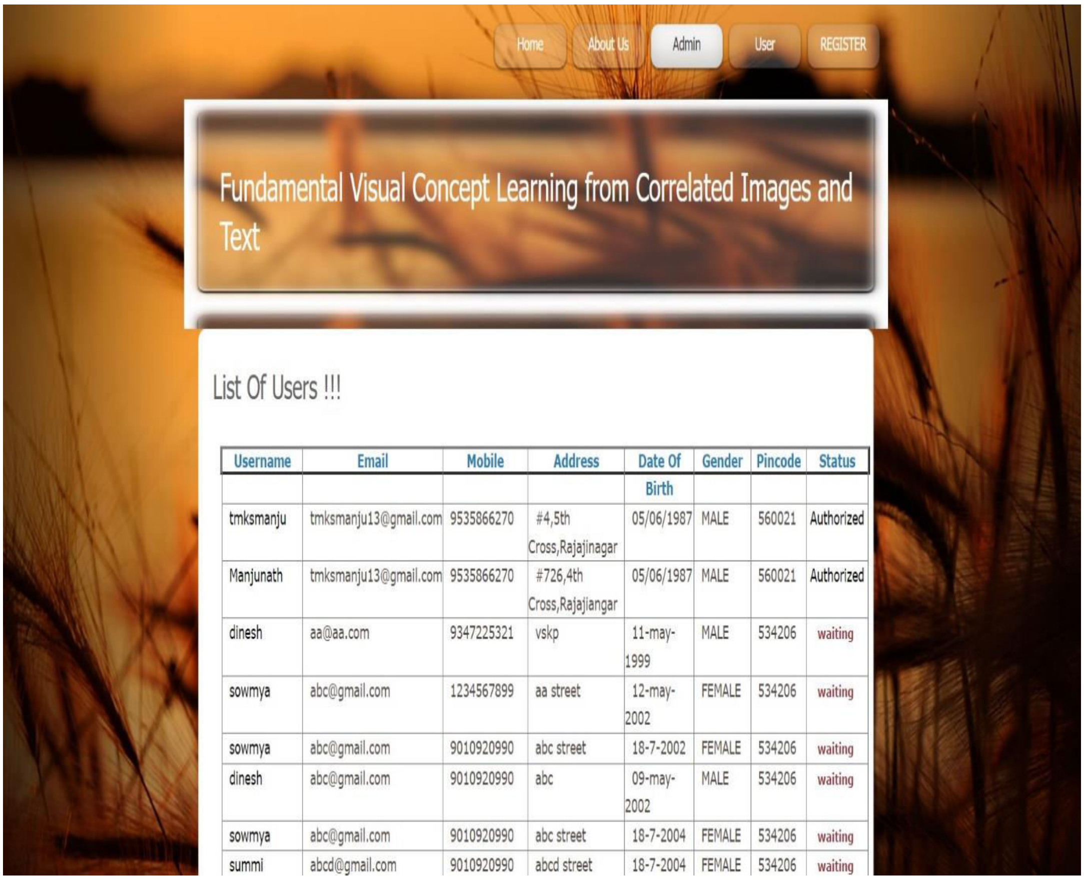

FIG:5.2.4 admin can view all user details-SCREEN SHOT

5.2.5 User main page

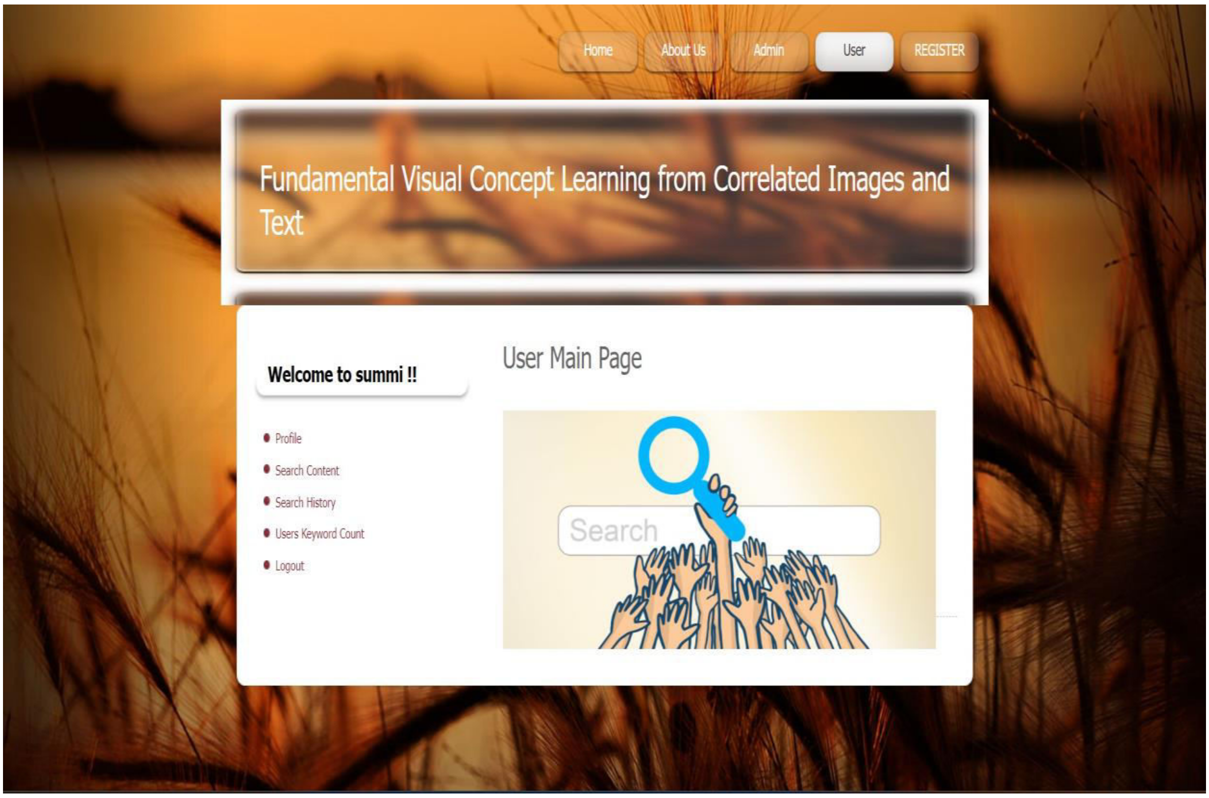

FIG:5.2.5 User main page -SCREEN SHOT

5.2.6 Admin can view all image details

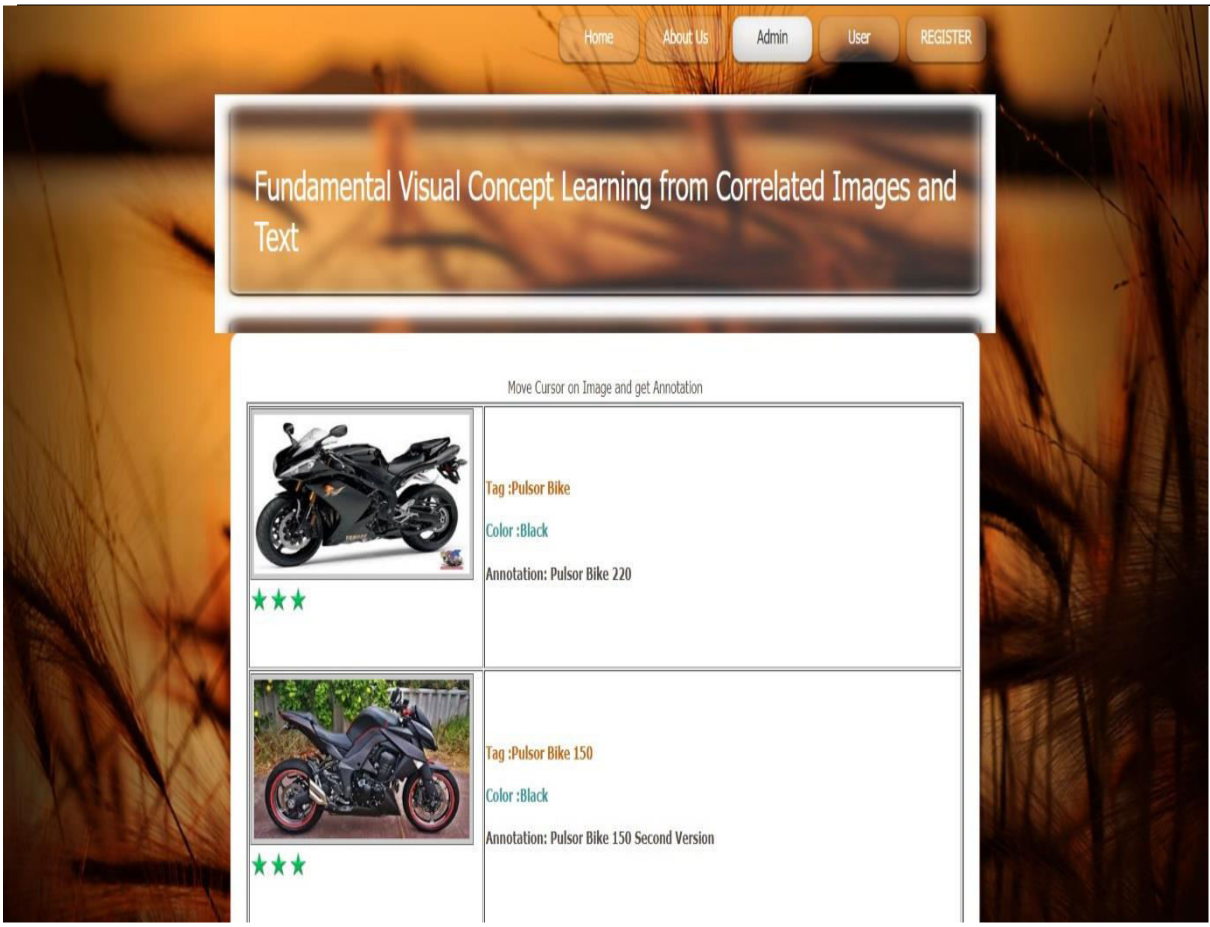

FIG:5.6 Admin can view all image details

6. CONCLUSION AND FUTURE WORK

CONCLUSION

In this paper, we formulate the problem of fundamental visual concept learning from correlated images and text and propose an approach to this problem called neighboring concept distributing. The proposed NCD approach introduces the concept graph, which consists of two kinds of edges, i.e., intra image edges and inter-image edges, to model the relations between patches belonging to the same image and different images, respectively. The approach distributes the semantic information from the images to the patches and propagates it across different patches by considering fitness, distinctiveness, smoothness and sparseness. The learn ability analysis reveals that, under some conditions, all concepts can be learned with an arbitrarily high performance as the correlated images and text data increase. Experimental results demonstrate that the proposed NCD approach outperforms state-of-the-art methods. **7. REFRENCES**

[1] Y. Zhuang, Y. Yang, and F. Wu, “Mining semantic correlation of heterogeneous multimedia data for cross-media retrieval,” *IEEE Transactions*

on Multimedia, vol. 10, no. 2, pp. 221–229, 2008.

[2] X. Wu, Y. Qiao, X. Wang, and X. Tang, “Cross matching of music and image,” in *Proceedings of the 20th ACM International Conference on Multimedia*. ACM, 2012, pp. 837–840.

[3] X. Zhai, Y. Peng, and J. Xiao, “Effective heterogeneous similarity measure with nearest neighbors for cross-media retrieval,” in *Proceedings of the International Conference on Multimedia Modeling*. Springer, 2012, pp. 312–322.

[4] T. Zhang and J. Wang, “Collaborative quantization for cross-modal similarity search,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2036–2045.

[5] M. Wang, B. Ni, X.-S. Hua, and T.-S. Chua, “Assistive tagging: A survey of multimedia tagging with human-computer joint exploration,” *ACM Computing Surveys (CSUR)*, vol. 44, no. 4, p. 25, 2012.

[6] Y. Gao, M. Wang, Z.-J. Zha, J. Shen, X. Li, and X. Wu, “Visualtextual joint relevance learning for tag-based social image search,” *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 363–376, 2013.

[7] Y. Feng and M. Lapata, “Automatic caption generation for news images,” *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 4, pp. 797–812, 2013.

[8] S. Yang, Y. Jiang, and Z. Zhou, “Multi-instance multi-label learning with weak label,” in *Proceedings of the International Joint Conference on Artificial Intelligence*. AAAI Press, 2013, pp. 1862–1868.

[9] X. Xue, W. Zhang, J. Zhang, B. Wu, J. Fan, and Y. Lu, “Correlative multi-label multi-instance image annotation,” in *Proceedings of the IEEE International Conference on Computer Vision*. IEEE, 2011, pp. 651–658.

[10] A. T. Pham, R. Raich, X. Z. Fern, J. P´erez Arriaga *et al.*, “Multiinstance multi-label learning in the presence of novel class instances,” in *Proceedings of the International Conference on Machine Learning*, 2015, pp. 2427–2435.