Image Tag Recommendation by Exploiting Social-Network Metadata

Amin Salehi, Avinash Reddy Kaitha

Abstract—With the unprecedented surge of images shared on the social media platforms, users are overloaded with many choices which can result in poor decisions. Social tagging alleviates this problem by enabling users to categorize and also retrieve images by using social tags. However, users might find it difficult to tag all the images, so they may leave some images untagged. To facilitate the process of tagging images for users, many image tag recommender systems have been proposed. A majority of these systems face two challenges. First, they do not capture the most representative features reflecting the visual content of images. Second, they neglect to efficiently exploit social-network metadata such as social interactions between users. Recently, deep learning has achieved a great success in image-related tasks (e.g., image classification) thanks to its capability to capture salient features. Moreover, according to social influence theory, users interacting with each other can change one another's attitude and behavior. Therefore, in this project, we propose a framework which exploits deep learning to capture highly representative features for the visual content of images and social-network metadata in order to recommend the most relevant tags to users.

Index Terms—Image Tag Recommendation, Social Tags, Social-Network Metadata, Social Interactions.

✦

1 INTRODUCTION

THe amount of multimedia data, especially images, shared on the social media platforms is enormously growing in recent years. Therefore, users are facing with the challenge of finding the relevant images from huge number of available ones. Fortunately, many social media platforms enable users to tag their images. User-generated tags provide insight to the image categories important to users. Therefore, users can utilize such categories to find the relevant images.

Users might not tag some images because (1) they might not come up with relevant tags for such images, and (2) they might find this process tedious. Many tag recommendation systems have been developed in order to assist users to tag such images.

Although a large body of work focuses mostly on the image visual content in order to recommend tags to users, they do not extract the most representative features. At present, many efforts have demonstrated that deep learning performs very well in the context of extracting the most relevant features for images. Therefore, deep learning can enhance tag recommender systems to recommend tags to users for their images according to the correspondence between tags and visual content.

According to social influence theory [1], [2], social influence can be defined as a change in an individual's attitudes and behavior resulting from social interactions with other individuals. Tagging behavior can be impacted by social influence. In other words, users interacting with each other tend to use the same tags for similar images.

2 RELATED WORK

A large body of work has been devoted to image tag recommendation [3], [4], [5], [6], [7], [8], [9], [10]. For instance, [9] utilizes the co-occurrence of social tags in order to detect concepts; a concept is defined as a set of highly co-occurred tags. Next, they recommend tags to users based on these concepts by using retrieval techniques. Another work [8] utilizes the social tags assigned to users' favorite images as well as users' friendships to recommend tags. However, these methods does not leverage the visual content of images. Several efforts [3], [5], [6] focus on exploiting visual similarity between images, using simple similarity measures, in order to recommend tags. Another work [4] learns tag relevance by aggregating votes from visually similar neighboring images. Moreover, [10] incorporates not only image similarity but also friendships, geotags, user groups, and historical tags into a tensor decomposition framework to suggest tags. Furthermore, [7] propose a convolutional neural network which first models visual content of images and the context which images is captured separately. Then, their proposed CNN merges these two models by using a concatenation layer.

The difference between our proposed framework compared with the previously proposed methods is that we aim to exploit social-network metadata as well as the features extracted from visual content of images using deep learning to recommend tags to users.

3 PROBLEM STATEMENT

Given visual content of images, user profiles, social interactions and image groups, the problem of image tag recommendation can be defined as utilizing these various sources of information in order to recommend the most relevant tags for a given image to a specific user. To accomplish this task, we aim to propose a model based framework exploiting convolutional neural networks as well as other state of the art models useful for such a task.

4 CHALLENGES

One of the challenges of this problem is data acquisition from Flickr. Although there are a lot of image datasets extracted from Flickr, there are no dataset providing social interaction between users. Therefore, we aim to collect a dataset containing social interactions by crawling Flickr. The main challenge here is that Flickr API provides directed fellowship information. Therefore, we cannot crawl a large enough connected component of user accounts by only using simple crawling methods in social networks such as Breadth First Search (BFS). Another challenges is proposing a framework which utilize all available metadata such as historical tags, image visual content, and social interactions in order to recommend the most relevant tags to a user for a given image. Since we are dealing with large amount of multimedia data as well as social-network metadata, another challenge we need to address is using a multi-thread processing framework such as Tensorflow to process this huge amount of data.

5 DATASET

We are going to use Flickr Data for this project. As mentioned in the challenges section, there are many existing datasets with Flickr Images but we would like to exploit the social-network metadata which isn't available in any of the public datasets so we would like to crawl Flickr and get all the images and other user data . The following are some of the attributes we would like to get from Flickr (not only for this project even for future ones usable for researchers): social interactions, historical tags, image visual content, community-curated image groups, favorite images for users, photo set, user profile, comments, and image attributes such as timestamp, view count, locations-geotags.

6 EVALUATION

We compare our proposed framework with the existing methods based on precision, recall and accuracy measures. To that end, we first predict top-k tags for each of the test images and compared with the ground truth. As suggested by [7], we compute the precision as the proportion of top k generated tags appearing in the user tags, recall as the proportion of the user tags appearing in the top k generated tags, and accuracy as 1 if at least one of the top k generated tags is present in the user tags and 0 otherwise.

7 TIMELINE

The tentative timeline of the project is as follows:

- Data Acquisition: 02/17/2017 - 03/10/2017
- Implementation:

  - Deep Learning: 03/10/2017 - 04/07/2017
  - Social Network Modeling: 03/10/2017 - 04/07/2017
  - Evaluation: 04/07/2017 - 04/14/2017

- Experiments: 04/14/2017 - 04/21/2017
- Final Paper: 04/21/2017

8 RESPONSIBILITIES

The tasks are divided as follows:

- Data Acquisition: Avinash
- Implementation:

  - Deep Learning: Avinash, Amin
  - Social Network Modeling: Amin
  - Evaluation: Amin

- Experiments: Avinash, Amin
- Final Paper: Avinash, Amin

REFERENCES

[1] Turner, John C. Social influence. Thomson Brooks/Cole Publishing Co, 1991.
[2] French, John RP, Bertram Raven, and D. Cartwright. The bases of social power. Classics of organization theory 7 (1959).
[3] Lindstaedt, Stefanie, et al. Automatic image annotation using visual content and folksonomies. Multimedia Tools and Applications 42.1 (2009): 97-113.
[4] Li, Xirong, Cees GM Snoek, and Marcel Worring. Learning social tag relevance by neighbor voting. IEEE Transactions on Multimedia 11.7 (2009): 1310-1322.
[5] Lee, Sihyoung, et al. Map-based image tag recommendation using a visual folksonomy. Pattern Recognition Letters 31.9 (2010): 976-982.
[6] Wu, Lei, et al. Learning to tag. Proceedings of the 18th international conference on World wide web. ACM, 2009.
[7] Rawat, Yogesh Singh, and Mohan S. Kankanhalli. ConTagNet: Exploiting User Context for Image Tag Recommendation. Proceedings of the 2016 ACM on Multimedia Conference. ACM, 2016.
[8] Haifeng, Guo, Su Shoubao, and Sun Zhoubao. Image tag recommendation based on friendships. Multimedia Tools and Applications (2016): 1-17.
[9] Sun, Aixin, Sourav S. Bhowmick, and Jun-An Chong. Social image tag recommendation by concept matching. Proceedings of the 19th ACM international conference on Multimedia. ACM, 2011.
[10] Panagopoulos, Michail, and Constantine Kotropoulos. "Image tagging using tensor decomposition." Information, Intelligence, Systems and Applications (IISA), 2015 6th International Conference on. IEEE, 2015.