A Picture Tells a Thousand Words – About You! User Interest Profiling from User Generated Visual Content

Quanzeng You[a], Sumit Bhatia[b], Jiebo Luo[a,*]

[a]Department of Computer Science, University of Rochester, Rochester NY, United States
[b]IBM Almaden Research Center, San Jose CA, United States

Abstract

Inference of online social network users' attributes and interests has been an active research topic. Accurate identification of users' attributes and interests is crucial for improving the performance of personalization and recommender systems. Most of the existing works have focused on textual content generated by the users and have successfully used it for predicting users' interests and other identifying attributes. However, little attention has been paid to user generated visual content (images) that is becoming increasingly popular and pervasive in recent times. We posit that images posted by users on online social networks are a reflection of topics they are interested in and propose an approach to infer user attributes from images posted by them. We analyze the content of individual images and then aggregate the image-level knowledge to infer user-level interest distribution. We employ image-level similarity to propagate the label information between images, as well as utilize the image category information derived from the user created organization structure to further propagate the category-level knowledge for all images. A large scale social network dataset of 1.5+ million images created from Pinterest is used for evaluation and the experimental results demonstrate the effectiveness of our proposed approach.

Keywords: User Profiling, Social Multimedia, Multimedia Analysis

*Corresponding author
Email addresses: qyou@cs.rochester.edu (Quanzeng You),
sumit.bhatia@us.ibm.com (Sumit Bhatia), jluo@cs.rochester.edu (Jiebo Luo)

1. Introduction

Online Social Networks (OSNs) such as Facebook, Twitter, Pinterest, Instagram, etc. have become a part and parcel of modern lifestyle. A study by Pew Research centre[1] reveals that three out of every four adult internet users use at least one social networking site. Such large scale adoption of OSNs and active participation of users have led to research efforts studying relationship between users' digital behavior and their demographic attributes (such as age, interests, and preferences) that are of particular interest to social science, psychology, and marketing. A large scale study of about 58,000 Facebook users performed by Kosinski et al. [1] reveals that digital records of human activity can be used to accurately predict a range of personal attributes such as age, gender, sexual orientation, political orientation, etc. Likewise, there have been numerous works that study variations in language used in social media with age, gender, personality, etc. [2, 3, 4]. While most of the popular OSNs studied in literature are mostly text based, some of them (e.g., Facebook, Twitter) also allow people to post images and videos. Recently, OSNs such as Instagram and Pinterest that are majorly image based have gained popularity with almost 20 billion photos already been shared on Instagram and an average of 60 million photos being shared daily[2].

The most appealing aspect of image based OSNs is that visual content is *universal* in nature and thus, not restricted by the *barriers of language*. Users from different cultural backgrounds, nationalizes, and speaking different languages can easily use the same *visual language* to express their feelings. Hence, analyzing the content of user posted images is an appealing idea with diverse applications. Some recent research efforts also provide support for the hypothesis that images posted by users on OSNs may prove to be useful for learning various personal and social attributes of users. Lovato et al. [5] proposed a method to learn users' latent visual preferences by extracting aesthetics and visual features from images favorited by users on Flickr. The learned models can be used to predict images likely to be favored by the user on Flickr with reasonable accuracy. Cristani et al. [6] infer personalities of users by extracting visual patterns and features from images marked as favorites by users on Flickr. Can et al. [7] utilize the visual cues of tweeted images in addition to textual and structure-based features to predict the retweet count of the posted image. Motivated by these works, we investigate *if the images posted by users on online social networks can be used to predict their*

---

[1]http://www.pewinternet.org/2013/12/30/social-media-update-2013/
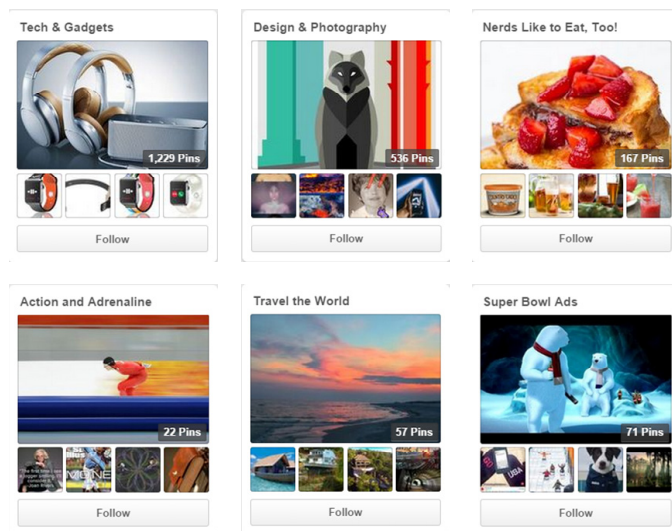[2]http://instagram.com/press/

2

Figure 1: Example pinboards from one typical Pinterest user.

fine-grained interests or preferences about different topics. To understand this better, Figure 1 shows several randomly selected pinboards (collection of images, as they are called in Pinterest) for a typical Pinterest user as an example. We observe that different *pins* (a pin corresponds to an image in Pinterest) and pinboards are indicative of the user's interests in different topics such as sports, art and food. We posit that the visual content of the images posted by a user in an OSN is a reflection of her interests and preferences. Therefore, an analysis of such posted images can be used to create an interest profile of the user by analyzing the content of individual images posted by the user and then aggregating the image-level knowledge to infer user-level preference distribution at a fine-grained level.

1.1. Problem Formulation and Overview of Proposed Approach

Problem Statement: Given a set $\mathcal{I}$ of images posted by the user u on an OSN, and a set $\mathcal{C}$ of interest categories, output a probability distribution over categories in $\mathcal{C}$ as the interest distribution for the user. In order to solve this problem, we first need to understand the relationships between different categories (topics) and underlying characteristics (or features) of user posted images in the training phase. These learned relationships can then be used to predict distribution over different interest categories by analyzing images posted by a new user. Even though state-of-the-art machine learning algorithms, especially developments in deep learning, have achieved significant results for individual image classification [8], we believe

3

that incorporating OSN and user specific information can provide further performance gains. Different image based OSNs offer users capability to group together similar images in the form of albums/pinboards, etc. In Pinterest, users create pinboards for a given topic and collect similar images in the pinboard. Given this human created group information, it is reasonable to assume that strong correlations exist between objects belonging to the same curated group or categorization. For example, a user may have two pinboards belonging to the *Sports* category, one for images related to soccer and one for images related to basketball. Therefore, even though all images in the two pinboards will share some common characteristics, images within each pinboard will share some additional similarities. Motivated by these observations, we employ image level and group level label propagation in order to build more accurate learning models. We employ image-level similarity to propagate the label information between images and category correlations are employed to further propagate the category-level knowledge for all images.

2. Related Work

2.1. User Profiling from Online Social Networks

It is discovered that in social network people's relationship follows the rule *birds of a feather flock together* [9]. Similarly, people in online social network also exhibit such kind of patterns. Online social network users connected with other users may be due to very different reasons. For instance, they may try to rebuild their real world social networks on the online social network. However, most of the time, people hope to show part of themselves to the rest of the world. In this case, the content generated by online social network users may help us to infer their characteristics. Therefore, we are able to build more accurate and more targeted systems for prediction and recommendation.

There are many related works on profiling different online users' attributes. Location is quite important for advertisement targeting, local customer personalization and many novel location based applications. In Cheng et al. [10], the authors proposed an algorithm to estimate a city level location for Twitter users. More recently, Li et al. [11] proposed a unified discriminative influence model to estimate the home location of online social users. They unified both the social network and user-centric signals into a probabilistic framework. In this way, they are able to more accurately estimate the home locations. Meanwhile, since most social network users tend to have multiple social network accounts, where they exhibit different behaviors on these platforms. Reza and Huan [12] proposed MOBIUS for finding a mapping of social network users across different

4

social platforms according to people's behavioral patterns. The work in Mislove et al. [13] also proposed an approach trying to inferring user profiles by employing the social network graph. In a recent work, Kosinski et al. [1] revealed the power of using Facebook to predict private traits and attributes for Facebook volunteer users. Their results indicated that simple human social network behaviors are able to predict a wide range of human attributes, including sexual orientation, ethnic origin, political views, religion, intelligence and so on. Li et al. [14] defined *discriminative correlation* between attributes and social connections, where they tried to infer different attributes from different circles of relations.

2.2. Visual Content Analysis

Visual content becomes increasingly popular in all online social networks. Recent research works have indicated the possibility of using online user generated visual content to learn personal attributes. In Kosinsky et al. [1], a total of $58,000$ volunteers provided their Facebook likes as well detailed demographic profiles. Their results suggest that digital records of human activities can be used to accurately predict a range of personal attributes such as age, gender, sexual orientation, and political orientation. Meanwhile, the work in Lovato et al. [15] tried to build the connection between cognitive effects and the consumption of multimedia content. Image features, including both aesthetics and content, are employed to predict online users' personality traits. The findings may suggest new opportunities for both multimedia technologies and cognitive science. More recently, Lovato et al. [16] proposed to learn users' biometrics from their collections of favorite images. Various perceptual and content visual features are proven to be effective in terms of predicting users' preference of images. You, Bhatia and Luo [17] exploited visual features to determine the gender of online users from their posted images. Meanwhile, there are also related works on applying multimodal approaches on visual content analysis including annotations [18], web image ranking [19] and multimedia retrieval [20].

3. Proposed Approach

As discussed in Section 1, the problem of inferring user interests can be considered as an image classification problem. However, in contrast to the classical image classification problem where the objective is to maximize classification performance at individual level, we are focused more on learning the overall user-level *image category distribution*, which in turn yields users' interests distribution. In our work, we use data crawled from pinterest.com, which is one of the most
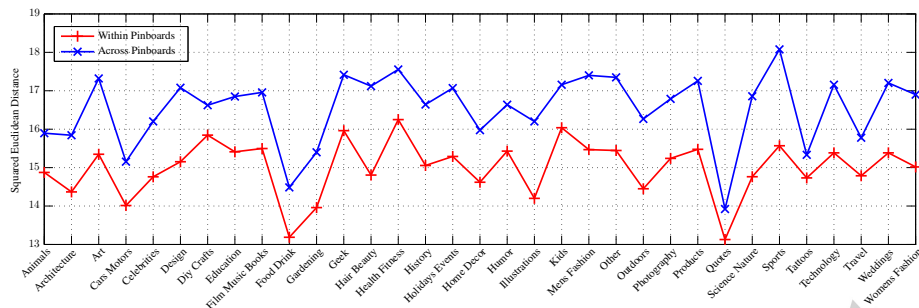
5

Figure 2: Average distance between images in different groups.

popular image based social networks. In Pinterest, users can share/save images that are known as *pins*. Users can categorize these pins into different *pinboards* such that a pinboard is collection of related pins (images). Also, while creating a pinboard the user has to choose a descriptive category label for the pinboard from a pre-defined list specified by Pinterest. There are a total of 34 available categories for users to choose from (listed in Table 1). A typical Pinterest manages/owns many different pinboards belonging to different categories and each pinboard will contain closely related images of the same category. Pinterest users mainly use pinboards to attract other users and also to organize the images of interest for themselves. Therefore, often times they will choose interesting and high-quality pins to add to their pinboards and one can use these carefully selected and well organized high quality images to infer the interests of the user.

Table 1: List of 34 categories in Pinterest.

Animals	Architecture	Art	Cars & Motorcycles	Celebrities	Design	DIY & Crafts
Education	Film, Music & Books	Food & Drink	Gardening	Geek	Hair & Beauty	Health & Fitness
History	Holidays & Events	Home Decor	Humor	Illustrations & Posters	Kids	Men's Fashion
Outdoors	Photography	Products	Quotes	Science & Nature	Sports	Tattoos
Technology	Travel	Weddings	Women's Fashion	Other		

3.1. Training Image-level Classification Model

We employ the Convolutional Neural Network (CNN) to train our image classifier. The main architecture comes from the Convolutional Neural Network proposed by Krizhevsky et al. [8], that has achieved state of the art performance in the challenging ImageNet classification problem. We train our model on top of the model of Krizhevsky et al. [8]. For a detailed description of the architecture of the model, we direct the reader to the original paper. In particular, we fine-tune our model by employing images and labels from Pinterest. For each image, we assign

6

the label of the image to be the same as the label of the pinboard it belongs to. More details on the preparation of training data for our model will be discussed in the experimental section. From the trained deep CNN model, it is possible to extract deep features for each image. These features, also known as high-level features, have shown to be more appropriate for performing various other image related tasks, such as image content analysis and semantic analysis [21]. Specifically, we extract the deep CNN features from the last fully connected layer of our trained CNN model and employ these deep features for label propagation as described in the next section. It is noteworthy that this work differs from image annotation, which deals with individual images and have been extensively studied. As will become more clear, the approach we take in the form of label propagation, is designed to exploit the strong connection between the images curated within and across collections by users. For the same reason, the typical noise in user data is suppressed in Pinterest due the same user curation process.

3.2. Image and Group Level Label Propagation for Prediction

During the prediction stage, we try to solve a more general problem, where we are given a collection of images without group information. However, we want to predict the users' interests from these unorganized collection of images. We propose to use label propagation for image labels. The work in Yu et al. [22]) also tried to use collection information for label propagation. However, differently from their work, where collection information is employed to extract sparsity constraint to the original label propagation algorithm, we propose to impose the additional group-level similarity to further propagate the image labels for the same user.

Meanwhile, we observe that for most of the categories in Table 1, the average distance between images in the same pinboard have closer distance than images in the same category. Figure 2 shows the average distance between images in the same pinboards and the average distance between images in different pinboards but in the same categories. The distance is represented by the squared Euclidean distance between deep features from last layer of Convolutional Neural Network. Intuitively, this can be explained by the fact that most of the categories are quite diverse in that they can contain quite different sub-categories. On the other hand, users generally create pinboards to collect similar images into the same group. These images are more similar in that they are more likely to be in the same sub-category of the category label chosen by the user. Hence, the motivation for implementing label propagation at user level.

7

Let n be the number of categories, we define matrix $G \in R^{n \times n}$ to be the affinity matrix between these n categories (G is normalized such that the columns of G sum 1). The intuitive idea is that in general, there exists some correlation between a person's interests. If one user likes sports, then he is likely to be interested in Health & Fitness. This kind of information may help us to predict users' interest more accurately, especially for users with very few images in their potential interest categories. Therefore, during the label propagation, we also consider the propagation of category-level information.

To incorporate the group level information into our model, we define the following iterative process for image label propagation.

$$Y^{t+1} = (1 - \Lambda)WY^tG + \Lambda Y^0 \tag{1}$$

where Y^0 is the initial prediction of image labels from the trained CNN model, W is the normalized similarity matrix between images and Λ is a diagonal matrix. Following Yu et al. [22]), Λ is defined as

$$\Lambda_{i,i} = \max_j \frac{Y_{i,j}^0}{\sum_k Y_{i,k}^0}. \tag{2}$$

We can consider the above process as two stages. In the first stage $(1 - \Lambda)WY^t$, we use the similarity between the images to propagate the labels at image level. In the second stage, the group relationship matrix G is employed to further propagate the group relationship to all images, i.e. we multiply matrix G with the results of $(1 - \Lambda)WY^t$. Next, we present the convergence analysis of the proposed label propagation framework.

3.2.1. Convergence Analysis

From Eqn. (1), we have the following formula for Y^{t+1}.

$$Y^{t+1} = ((1 - \Lambda)W)^{t+1}Y^0G^{t+1} + \sum_{i=0}^{t} ((1 - \Lambda)W)^i \Lambda Y^0 G^i \tag{3}$$

Since we have $0 \le \lambda_{i,j}, w_{i,j}, G_{i,j} < 1$ for all i, j, therefore

$$\lim_{t \to \infty} ((1 - \Lambda)W)^{t+1} Y^0 G^{t+1} = 0. \tag{4}$$

---

[3]In our implementation, we use Jaccard index to calculate G, see the experimental section for detailed description.

Algorithm 1 User Profiling by Group Constraint Label Propagation

---

Require: $X = \{x_1, x_2, \ldots, x_N\}$ a collection of images

     M: Fine-tuned Imagenet CNN model

     G: Correlation matrix between the labels

1: Predict the categories $Y^0 \in R^{N \times K}$ (K is the number of categories) of X using the trained CNN model M

2: Extract deep features from M for all X.

3: Calculate the similarity matrix W' between X using Gaussian kernel function $W'(i, j) = \exp(-\frac{\|x_i - x_j\|^2}{2\delta^2})$, where x_i and x_j are the deep features for image i and j respectively.

4: Normalize W' to get $W = D^{-1}W'$, where D is diagonal matrix with $D_{ii} = \sum_j W'_{i,j}$.

5: Calculate the diagonal matrix Λ according to Eqn.(2).

6: Calculate the affinity matrix between G between different categories[3].

7: Initialize, iteration index $t = 0$

8: **repeat**

9:   Employ Eqn.(1) to update Y^{t+1} according to Y^t

10: **until** Convergence or t reaches the maximum iteration number

11: Normalize rows of $Y^t \in R^{N \times K}$ to get $Y'^t \in R^{N \times K}$.

12: **return** Y'^t.

It follows that

$$\lim_{t \to \infty} Y^t = \sum_{i=0}^{t-1} \left((1 - \Lambda) W\right)^i \Lambda Y^0 G^i. \tag{5}$$

According to the theorem that *the product of two converging sequences converges* [23], we define two different matrix series $A^t = \left((1 - \Lambda) W\right)^t \Lambda Y^0$ and $B^t = G^t$. For matrix series A^t, it converges [22]. For matrix series B^t, it also converges since $\rho(B) < 1$. Therefore, we conclude that the product of the two matrix series A^t and B^t also converges. More importantly, we have $\sum_i A_i = (I - \Lambda) \Lambda Y^0$ and $\sum_i B_i = (I - Q)^{-1}$. All the elements of A_i and B_i are non-negative, which leads to the fact that $\sum_i A_i B_i <_p (\sum_i A_i)(\sum_i B_i)$. We use $<_p$ to represent the element-wise comparison between two matrix. In other words, we have a upper bound for Y^t, together with the fact that $A_i B_i \geq 0$, we conclude that Y^t converges.

Meanwhile, since we implement label propagation in each user's collection of images, thus the main computational and storage complexity is in the order of

$O(n_i^2)$, where n_i is the number of images for user i.

Algorithm 1 summarizes the main steps using the proposed group constrained label propagation framework. Note that in step 3, we use a Gaussian kernel function to calculate the similarity between two images using the deep features. However, one can employ other techniques such as locally linear embedding (LLE) [24] to calculate the similarity between different instances.

Eventually, we obtain users' interests distribution by aggregating the label distribution of the collections of their images. In other words, we simply sum up the label distribution of individual images and then normalize it to produce the users' interest prediction results. Equ.(6) shows the detail of aggregating the j-th interest for user u_i from his collection of images I_{u_i}.

$$Y(u_i)_j = \frac{\sum_{i \in I_{u_i}} p_i(j)}{\sum_k \sum_{i \in I_{u_i}} p_i(k)} \tag{6}$$

Another advantage of the proposed approach is that we can also update user profiling in an *online* approach. In particular, OSN users may continuously post messages. Our model can also update users' profiles in an online way. First, we predict newly added images and then employ the proposed label propagation on all the available images. In this way, we are able to more accurately estimate users' current interests.

4. Experiments and Evaluations

To evaluate the proposed algorithm, we crawl data from Pinterest according to a randomly chosen user list consisting of 748 users. Table 2 gives the statistics of our crawled dataset. This dataset is used to train our own deep convolutional neural network. We use 80% of these images as the training data and the remaining 20% as validation data. We also downloaded another 77 users' data as our testing data. Since each pinboard has one category, we use the category label as the label for all the images in that pinboard.

Table 2: Statistics of our dataset.

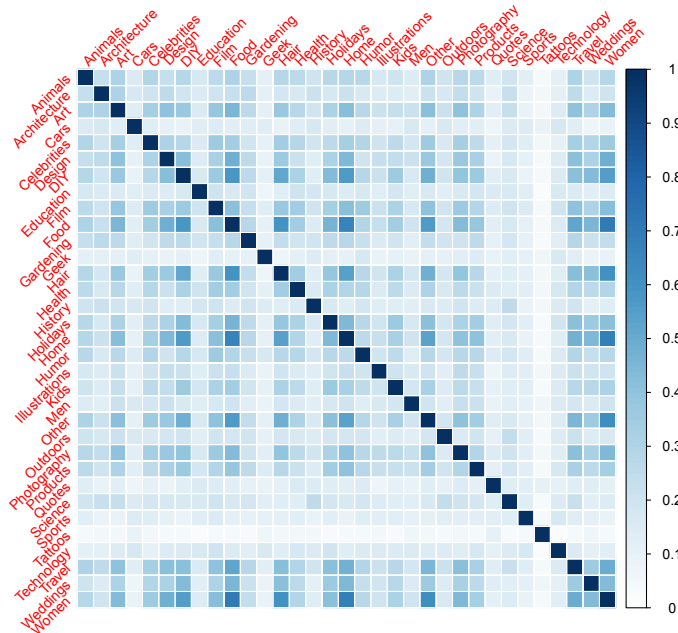|	Training and Validating	Testing
Num of Users	748	77
Num of Pinboards	30,213	1,126
Num of Pins	1,586,947	66,050

10

Figure 3: Jaccard similarity coefficients between different categories in Table 1.

We train the CNN on top of the ImageNet convolutional neural model [8]. We use the publicly available implementation Caffe [25] to train our model. All of our experiments are evaluated on a Linux X86_64 machine with 32G RAM and two NVIDIA GTX Titan GPUs. We finish the training of CNN after a total of $200,000$ iterations.

In the proposed group constrained label propagation, we need to know the similarity matrix G between all categories. We propose to learn the similarity matrix G from pinterest users' behaviors of building their collections of pinboards. No additional textual semantic or visual semantic information is employed to build the similarity matrix. In our experiments, we employ Jaccard index to calculate the similarity coefficient between different categories in Table 1. Specifically, the entry of G_{ij} is defined as the Jaccard index between category i and j, which is the ratio between the number of users who have pinboards of both categories i and j and the number of users who have pinboards of category i or j. Figure 3 shows the coefficients between all different categories. It shows that users prefer to choose some of the categories together, which may suggest potential category recommendations for Pinterest users.

4.1. Evaluation Criteria

To evaluate the performance of different models, we employ two different criteria. 1) **Normalized Discounted Cumulative Gain (NDCG) score** NDCG is a popular measure for ranking tasks [26]. Discounted Cumulative Gain at position n is defined as

$$DCG_n = p_1 + \sum_{i=1}^{n} \frac{p_i}{\log_2(i)}, \tag{7}$$

where p_i is the relevance score at position i. In our case, we use the predicted probability as the relevance score. NDCG is the normalized DCG, where the value is between 0 and 1. To calculate the NDCG, for a user u_i, we first get the ground truth distribution of his interest according to the category labels of his pinboards and then rank the ground truth categories in descending order of their distribution probabilities. Note that the pinboard labels of u_i may not contain all the categories and our trained classifier may give a class distribution over all the categories. Therefore, when calculating the NDCG score we add a prior to the interest distribution of each user. We first calculate the prior category distribution p_0 from all the training and validating data set according to their labels. Then, for each testing user i we first calculate the category distribution of p_i according to the ground truth labels. Next, we smooth the distribution by updating $p_i = p_i + 0.1 * p_0$ and then normalize it to get the ground truth interest distribution for user i. In this way, we are able to calculate the NDCG score for each user according to the prediction of the model. 2) **Recall@Rank K** In this metric, we firstly build the ground truth for one user u_i, we rank all of his categories in descending order according to the distribution probabilities. Then, we calculate the shared top ranked categories with the ground truth. Users' data in Pinterest may be incomplete. For example, user i may be interested in Sports and Women's Fashion. However, she only has created pinboards belonging to Women's Fashion. She has not created Sports related pinboards. However, it does not mean that she is not interested in Sports. Thus, it makes sense to evaluate different algorithms' performance using the Recall@Rank k.

4.2. Experimental Results

We report the results for three different models. First one is the results from our fine-tuned CNN model, i.e. the initial prediction for label propagation. Second model is label propagation [22] (LP) and the last one is the proposed group constrained label propagation (GLP).
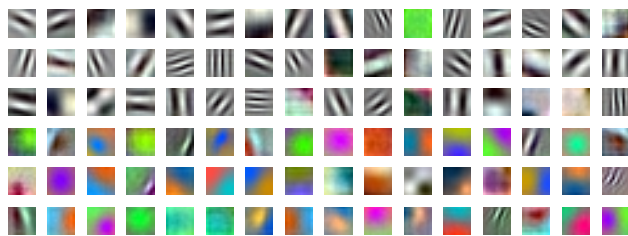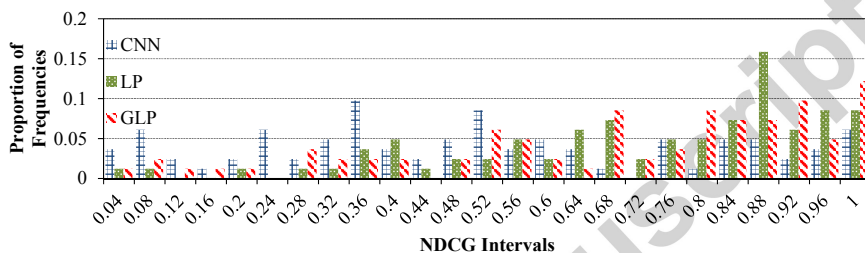
Figure 4: Filters of the first convolutional layer.

Figure 5: Distribution of NDCG scores for three different approaches.

We first train our CNN model using the data in Table 2. Figure 4 shows the learned filters in the first layer of the CNN model, which is consistent with the filters learned in other image related deep learning tasks [27]. These Gabor-like filters can be employed to replicate the properties of human brain V1 area [28], which can detect low-level texture features from raw input images. Next, we employ this model to predict the labels of the testing data, which is going to be the initial label prediction for LP and GLP models.

Table 3: Mean and Standard deviation of NDCG score

Model	Mean	STD
CNN	0.692	0.179
LP	0.818	0.138
GLP	0.826	0.138

For both LP and GLP models, we choose the δ in the Gaussian kernel to be the variance of distance between each pair of deep features [29]. We set the maximum number of iterations to be 100 for both models[4]. Figure 5 shows the performance

---

[4]In our experiments, the increase of iteration number over 100 leads to similar results with the

13

of using NDCG for the three models. The results shows that both LP and GLP try to move the distribution to the right, which suggests that both models can improve the performance in terms of NDCG score for most of the users. To quantitatively analyze the results, we give the mean and the standard derivation of the NDCG scores in Table 3. Both LP and GLP improve the NDCG score from about 0.69 to over 0.80 and reduce the standard deviation. GLP shows slight advantage over LP in terms of both mean and standard deviation of NDCG score.
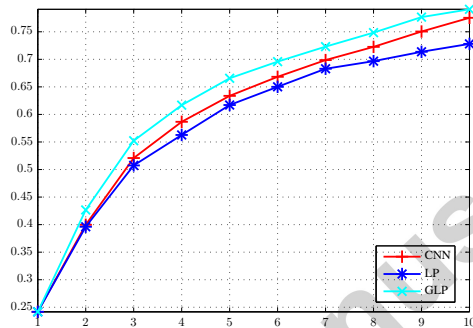

Figure 6: Recall@K for different K for the three models respectively.

The evaluation results of using Recall@K are shown in Figure 6. Differently from the NDCG score, CNN and GLP show better recall performance then LP for all different values of Ks. Meanwhile, GLP consistently outperforms the CNN model. The results suggest that without the propagation of category or group similarity, label propagation may cause the increase of irrelevant categories, which leads to poor recall.

Besides the user-level interests prediction, we are also interested in *whether or not label propagation can improve the performance of image-level classification performance*. Since we have the pinboard labels as the ground-truth labels for each testing image, we report the accuracy of the three models on the testing images. Table 4 summarizes the accuracy. Indeed, the performance of these three models is quite similar. These results may suggest that the label propagation have limited impact on the overall label distribution of individual image, i.e. the max index in the probability distribution of labels. However, they may change the probability distribution, which may benefits the overall user-level interests estimation.

---

reported results in this work.

Table 4: Accuracy on Image-level classification of different models.

Model	Accuracy
CNN	0.431
LP	0.434
GLP	0.434

5. Conclusions

We addressed the problem of user interest distribution by analyzing user generated visual content. We framed the problem as an image classification problem and trained a CNN model on training images spanning over 748 users' photo albums. Taking advantage of the human intelligence incorporated through the user curation of the organized visual content, we used image-level similarity to propagate the label information between images, as well as utilized the image category information derived from the user created organization structure to further propagate the category-level knowledge for all images. Experimental evaluation using data derived from Pinterest provided support for the effectiveness of the proposed method. In this work, our focus was on using image content for user interest prediction. We plan to extend our work to also incorporate user generated textual content in the form of comments and user profile information for improved performance.

Reference

[1] M. Kosinski, D. Stillwell, T. Graepel, Private traits and attributes are predictable from digital records of human behavior, Proceedings of the National Academy of Sciences 110 (15) (2013) 5802–5805.

[2] J. D. Burger, J. Henderson, G. Kim, G. Zarrella, Discriminating gender on twitter, in: EMNLP, 2011, pp. 1301–1309.

[3] D. Bamman, J. Eisenstein, T. Schnoebelen, Gender identity and lexical variation in social media, Journal of Sociolinguistics 18 (2) (2014) 135–160.

[4] H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, L. Dziurzynski, S. M. Ramones, M. Agrawal, A. Shah, M. Kosinski, D. Stillwell, M. E. Seligman, et al., Personality, gender, and age in the language of social media: The open-vocabulary approach, PloS one 8 (9).

[5] P. Lovato, A. Perina, D. S. Cheng, C. Segalin, N. Sebe, M. Cristani, We like it! mapping image preferences on the counting grid., in: ICIP, 2013, pp. 2892–2896.

[6] M. Cristani, A. Vinciarelli, C. Segalin, A. Perina, Unveiling the multimedia unconscious: Implicit cognitive processes and multimedia content analysis, in: ACM MM, ACM, 2013, pp. 213–222.

[7] E. F. Can, H. Oktay, R. Manmatha, Predicting retweet count using visual cues, in: Proceedings of the 22nd ACM international conference on Conference on information and knowledge management, ACM, 2013, pp. 1481–1484.

[8] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks., in: NIPS, Vol. 1, 2012, p. 4.

[9] M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a feather: Homophily in social networks, Annual review of sociology (2001) 415–444.

[10] Z. Cheng, J. Caverlee, K. Lee, You are where you tweet: a content-based approach to geo-locating twitter users, in: Proceedings of the 19th ACM international conference on Information and knowledge management, ACM, 2010, pp. 759–768.

[11] R. Li, S. Wang, H. Deng, R. Wang, K. C.-C. Chang, Towards social user profiling: unified and discriminative influence model for inferring home locations, in: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2012, pp. 1023–1031.

[12] R. Zafarani, H. Liu, Connecting users across social media sites: a behavioral-modeling approach, in: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2013, pp. 41–49.

[13] A. Mislove, B. Viswanath, K. P. Gummadi, P. Druschel, You are who you know: inferring user profiles in online social networks, in: Proceedings of the third ACM international conference on Web search and data mining, ACM, 2010, pp. 251–260.

[14] R. Li, C. Wang, K. C.-C. Chang, User profiling in an ego network: co-profiling attributes and relationships, in: Proceedings of the 23rd international conference on World wide web, International World Wide Web Conferences Steering Committee, 2014, pp. 819–830.

[15] P. Lovato, A. Perina, N. Sebe, O. Zandonà, A. Montagnini, M. Bicego, M. Cristani, Tell me what you like and ill tell you what you are: discriminating visual preferences on flickr data, in: Computer Vision–ACCV 2012, Springer, 2013, pp. 45–56.

[16] P. Lovato, M. Bicego, C. Segalin, A. Perina, N. Sebe, M. Cristani, Faved! biometrics: Tell me which image you like and i'll tell you who you are, Information Forensics and Security, IEEE Transactions on 9 (3) (2014) 364–374.

[17] Q. You, S. Bhatia, J. Luo, The eyes of the beholder: Gender prediction using images posted in online social networks, in: Social Multimedia Data Mining, 2014.

[18] M. Wang, X.-S. Hua, R. Hong, J. Tang, G.-J. Qi, Y. Song, Unified video annotation via multigraph learning, Circuits and Systems for Video Technology, IEEE Transactions on 19 (5) (2009) 733–746.

[19] M. Wang, H. Li, D. Tao, K. Lu, X. Wu, Multimodal graph-based reranking for web image search, Image Processing, IEEE Transactions on 21 (11) (2012) 4649–4661.

[20] Y. Gao, M. Wang, R. Ji, Z. Zha, J. Shen, k-partite graph reinforcement and its application in multimedia information retrieval, Information Sciences 194 (2012) 224–239.

[21] Y. Bengio, Learning deep architectures for ai, Foundations and trends® in Machine Learning 2 (1) (2009) 1–127.

[22] J. Yu, X. Jin, J. Han, J. Luo, Collection-based sparse label propagation and its application on social group suggestion from photos, ACM Transactions on Intelligent Systems and Technology (TIST) 2 (2) (2011) 12.

[23] M. Rosenlicht, Introduction to analysis, Courier Dover Publications, 1968.

[24] D. L. Donoho, C. Grimes, Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data, Proceedings of the National Academy of Sciences 100 (10) (2003) 5591–5596.

[25] Y. Jia, Caffe: An open source convolutional architecture for fast feature embedding, `http://caffe.berkeleyvision.org/` (2013).

[26] W. B. Croft, D. Metzler, T. Strohman, Search engines: Information retrieval in practice, Addison-Wesley Reading, 2010.

[27] M. D. Zeiler, R. Fergus, Visualizing and understanding convolutional neural networks, arXiv preprint arXiv:1311.2901.

[28] H. Lee, C. Ekanadham, A. Y. Ng, Sparse deep belief net model for visual area v2, in: Advances in neural information processing systems, 2008, pp. 873–880.

[29] U. Von Luxburg, A tutorial on spectral clustering, Statistics and computing 17 (4) (2007) 395–416.