Privacy-Aware Tag Recommendation for Accurate Image Privacy Prediction

ASHWINI TONGE, Kansas State University, USA
CORNELIA CARAGEA, University of Illinois at Chicago, USA

Online images' tags are very important for indexing, sharing, and searching of images, as well as surfacing images with private or sensitive content, which needs to be protected. Social media sites such as Flickr generate these metadata from user-contributed tags. However, as the tags are at the sole discretion of users, these tags tend to be noisy and incomplete. In this paper, we present a privacy-aware approach to automatic image tagging, which aims at improving the quality of user annotations, while also preserving the images' original privacy sharing patterns. Precisely, we recommend potential tags for each target image by mining privacy-aware tags from the most similar images of the target image, which are obtained from a large collection. Experimental results show that, although the user-input tags comprise noise, our privacy-aware approach is able to predict accurate tags that can improve the performance of a downstream application on image privacy prediction, and outperforms an existing privacy-oblivious approach to image tagging. The results also show that, even for images that do not have any user tags, our proposed approach can recommend accurate tags. Crowd-sourcing the predicted tags exhibits the quality of our privacy-aware recommended tags. Our code, features, and the dataset used in experiments are available at: https://github.com/ashwinitonge/privacy-aware-tag-rec.git.

CCS Concepts: • **Security and privacy → Software and application security**; • **Social network security and privacy**;

Additional Key Words and Phrases: Social networks, image analysis, image privacy prediction, deep learning, tag recommendation, privacy-aware tags.

ACM Reference Format:
Ashwini Tonge and Cornelia Caragea. 2018. Privacy-Aware Tag Recommendation for Accurate Image Privacy Prediction. *ACM Trans. Intell. Syst. Technol.* 9, 4, Article 39 (March 2018), 27 pages. https://doi.org/0000001.0000001

1 INTRODUCTION

Images are constantly shared on social networking sites such as Facebook, Flickr, and Instagram. For instance, it is common to take photos at cocktail parties and upload them on social networking sites without much hesitation for self-promotion and personal sharing. However, when privacy settings are used inappropriately, these photos can potentially reveal a user's personal and social habits, resulting in unwanted disclosure and privacy violations [Ahern et al. 2007; Spyromitros-Xioufis et al. 2016; Squicciarini et al. 2014, 2017a; Zerr et al. 2012]. For example, malicious attackers can take advantage of these accidental leaks to launch context-aware or even impersonation attacks. Personal data can be harvested through social media without users' consent if the privacy settings of social media are not managed properly, which could lead to online privacy risks [Bullguard

Authors' addresses: Ashwini Tonge, Kansas State University, KS, USA, atonge@ksu.edu; Cornelia Caragea, University of Illinois at Chicago, IL, USA, cornelia@uic.edu.

2157-6904/2018/3-ART39 $15.00
https://doi.org/0000001.0000001

(a) *Private*: Stylish, Elegant Corporate, Style, Street Fashion, Girl, Woman Skirt, Top, Bag, Pretty

(b) *Public*: Parisi, Sabrina News, Celebrity Famous, Girl Woman, Hollywood

Fig. 1. Anecdotal evidence for visually similar images with privacy-aware user tags.

2018]. A study carried out by the Pew Research center reports that 11% of the users of social networking sites regret the content they posted [Madden 2012]. Thus, several works have been developed in recent years in an attempt to provide appropriate privacy settings for online images [Spyromitros-Xioufis et al. 2016; Squicciarini et al. 2014, 2017b; Tonge and Caragea 2016, 2018; Tonge et al. 2018b; Tran et al. 2016; Zerr et al. 2012; Zhong et al. 2017].

Prior works on privacy prediction [Squicciarini et al. 2014, 2017b; Tonge and Caragea 2016, 2018; Tonge et al. 2018b; Zerr et al. 2012] found that the tags associated with images are indicative of their sensitive content. Tags are also important for image-related applications such as indexing, sharing, searching, content detection and social discovery [Bischoff et al. 2008; Gao et al. 2011; Hollenstein and Purves 2010; Tang et al. 2009]. Yet, the tags are at the sole discretion of users, and hence, they tend to be noisy and incomplete [Sundaram et al. 2012]. Despite that many approaches to automatic image tagging have been developed [Chen et al. 2013; Feng et al. 2004; Guillaumin et al. 2009; Liu et al. 2009; Makadia et al. 2008; Yavlinsky et al. 2005], these approaches do not consider the privacy aspect of an image while making the annotations (or tagging) and could not be sufficient for identifying images' private content.

We posit that visually similar images can possess very different sets of tags if these images have different privacy orientations. For example, Figure 1 shows anecdotal evidence obtained from a Flickr dataset in which visually similar images of private and public classes display different sets of user tags. The picture of a woman that belongs to the private class in Figure 1(a) contains tags such as "Elegant," "Corporate," "Style," and "Pretty," whereas the picture of a woman that belongs to the public class in Figure 1(b) contains tags such as "Celebrity," "Famous," "News," and "Hollywood." An image is considered to be private if it belongs to the private sphere (e.g., portraits, family, friends, home) or contains information that can not be shared with everybody on the Web (e.g., private documents), whereas the remaining images are considered to be public [Zerr et al. 2012]. Figure 1 shows that the images' tags are correlated to each image's privacy patterns [Klemperer et al. 2012; Squicciarini et al. 2017b, 2011]. These tags are very useful when access to the visual content of images is not allowed due to users reluctance to share the actual images for visual content analysis (which could reveal a user's identity through the face and friends, etc.). In such cases, privacy-aware tags can become good indicators of the privacy settings and can help improve the privacy prediction methods to reduce privacy breaches.

To this end, we ask the following questions: *Can we develop an automated approach to recommend accurate image tags that can also take into account the sharing needs of the users for images in*

questions? Can this method make precise tag recommendations for newly uploaded images that have an incomplete set of user tags or no tags at all? Can these recommended tags help improve the image privacy prediction performance? We address these questions with our research agenda. In particular, we draw ideas from the collaborative filtering line of research and explore its applicability to privacy-aware image tagging. Collaborative filtering is widely used to make recommendations for unknown items to users and relies on the assumption that similar users express similar interests or preferences on similar items [Su and Khoshgoftaar 2009]. Hence, we explore tag recommendation to images based on images' similar neighbors.

Contributions and Organization. We present a privacy-aware approach to automatic image tagging, originally introduced in our prior work [Tonge et al. 2018a]. Our approach aims at improving the quality of user annotations (or user tags), while also preserving the images' original privacy sharing patterns. Precisely, we recommend potential tags for each target image by mining privacy-aware tags from the most similar images of the target image, which we obtain from a large collection of images.

In this extended version of the paper, we augment our study by providing extensive experiments to validate the proposed approach:

- We study our privacy-aware recommended tags obtained by the proposed privacy-aware weighting scheme in an ablation experiment for privacy prediction. In this experiment, we compare various privacy-aware and privacy-oblivious weighting schemes and observe how the privacy prediction performance varies for these weighting schemes. We also experiment with various parameter values to estimate the best parameter setting.
- We compare the performance of privacy prediction using tags recommended by the proposed approach against the tags recommended by a prior state-of-the-art image annotation method. Our objective in this experiment is to verify whether the recommended tags by the proposed approach can capture better privacy characteristics than the prior state-of-the-art annotation.
- We investigate tag recommendation in a binary image privacy prediction task and show that the predicted tags can exhibit relevant cues for specific privacy settings (*public* or *private*) that can be used to improve the image privacy prediction performance.
- Our results show that we achieve a better privacy prediction performance when we add the recommended privacy-aware tags to the original user tags and predicted deep tags of images as compared to prior approaches of image privacy prediction.
- We also evaluate the recommended tags by employing crowd-sourcing to identify relevancy of the suggested tags to images. The results show that, although the user-input tags comprise noise or even some images do not have any tags at all, our approach is able to recommend accurate tags. In addition, we evaluate both privacy-aware and privacy-oblivious recommended tags and show that the privacy-aware recommended tags describe an image's content more accurately as compared to the privacy-oblivious tags.

The rest of the paper is organized as follows. We summarize prior works in Section 2. In Section 3, we describe the proposed algorithm. Section 4 provides details about the dataset that we use to evaluate the proposed approach. In Section 5, we describe the experimental setting and results. We finish our analysis in Section 6, where we provide a brief discussion of our main findings, future directions and conclude the paper.

2 RELATED WORK

In this section, we briefly review the related work on three lines of research: 1) automatic image annotation, 2) tag recommendation using collaborative filtering, and 3) online image privacy.

2.1 Automatic Image Annotation

Numerous approaches to automatic image annotation (or tagging) have been proposed in the literature to improve the search and retrieval of images based on text queries. We classify these methods into following categories:

Generative methods. The generative methods try to maximize the generative likelihood of the image features and tags [Feng and Lapata 2008, 2010; Ghoshal et al. 2005; Lavrenko et al. 2004; Lienhart et al. 2009; Peng et al. 2009; Putthividhya et al. 2010; Wang et al. 2009a; Yu and Ip 2006; Zhao et al. 2009]. For example, Lavrenko et al. [2004] learned joint probabilistic models of image content features and tags. These models compute the conditional likelihood of words given image content features that can be used to infer the most likely tags for an image. Later, Feng and Lapata [2010] used LDA to infer topics that capture co-occurrences of visual features and words.

Discriminative methods. Discriminative methods perceive image annotation as a multi-label classification problem. In these works, the authors typically treat the image tagging as a classification task and train classifiers (e.g., Support Vector Machines) for each tag using image's textual and/or visual features [Ciocca et al. 2011; Dimitrovski et al. 2011; Grangier and Bengio 2008; Murthy et al. 2014]. The graph-based learning (semi-supervised) methods are also used for image annotation in which the model is the graph of the entire data. The label correlation is incorporated in the graph as graph weights [Feng and Bhanu 2016; Liu et al. 2009; Wang and Hu 2010; Wang et al. 2009b, 2011] or as an additional constraint [Bao et al. 2012; Zha et al. 2009]. In addition to the graph-based learning methods, some studies exploit the local label correlations [Huang and Zhou 2012], underlying correlations among labels using a multi-label dictionary learning [Jing et al. 2016] and handle the missing tags issues [Wu et al. 2015a].

Tag completion methods. The tag completion methods automatically annotate images by identifying the missing tags and correcting the noisy tags. The entire dataset is represented as an initial matrix with each row as an image and each column as a tag. The tag completion methods recover this initial matrix by identifying correct associations between images and labels. The tag completion-based annotation is achieved by matrix completion [Qin et al. 2015; Wu et al. 2013], linear sparse reconstruction [Lin et al. 2014, 2013], subspace clustering with matrix completion [Hou and Zhang 2015], and low-rank matrix factorization [Li et al. 2016, 2014].

Deep learning methods. The deep-learning based image annotation adopt image features and semantic tag relationships extracted using deep networks [Gong et al. 2013; Hu et al. 2016b; Jin and Nakayama 2016; Niu et al. 2018; Wang et al. 2016, 2017; Wu et al. 2015b; Yang et al. 2015]. For example, Wang et al. [2017] proposed a multitask voting automatic image annotation CNN, which contains shallow layers and regards each category as a label directly, using the raw images as inputs for large scale image annotation.

Nearest neighbors methods. The nearest neighbor model-based image annotation methods assume that visually similar images are more likely to share common labels [Bakliwal and Jawahar 2015; Chen et al. 2013; Kalayeh et al. 2014; Lin et al. 2012; Makadia et al. 2008; Tian and Shen 2014; Wu et al. 2009, 2011]. For a given target image, these methods first obtain a set of similar images and then the tags of the target image are derived based on the tags of the similar images. For example, Guillaumin et al. [2009] proposed the "TagProp" model, which integrates a weighted nearest neighbor based method and metric learning capabilities into a discriminative framework.

Furthermore, Cheng et al. [2018] discussed advantages and disadvantages of these methods in details. For instance, the generative models may not be able to capture the intricate relationship

between image features and labels, which is imperative to identify the privacy-aware tags. Additionally, the multi-label classification based discriminative approaches cannot be extended to a large number of image tags since a binary classifiers has to be trained for each tag, which is not feasible for the online images that contain diverse sets of tags. On the other hand, the tag completion models suffer from a major weakness, that is, the transformation of the tag completion process to an optimization problem. The process of optimizing the objective function may be time-consuming and computationally very complex, and cannot guarantee global optimization. Moreover, despite that deep learning based methods have shown significant improvements in the performance of image annotation, there are still a few shortcomings with these methods. The main drawback is that although RNN + CNN solve issues pertaining to label quantity prediction and label dependencies for large-scale image annotation, still a better solution to rank labels is needed as RNN requires an ordered sequential list as input, which is mostly not present in the online images. Another drawback is that the increase in the depth and breadth of the deep networks can cause the decrease in the efficiency of annotation methods. The nearest neighbor based methods are clear and intuitive, and many of them have been proven to be quite successful for tag prediction due to their high flexibility. However, improvements are still needed because of some inherent shortages. For instance, the performance of these methods is highly sensitive to the retrieval performance. Thus, an efficient way to identify appropriate neighbors for unlabeled images is highly sought.

In contrast to previous annotation mechanisms, we take advantage of both nearest neighbors and deep learning based approaches to provide privacy-aware image annotations. We consider nearest neighbor based approaches as our strong baselines.

2.2 Tag Recommendation using Collaborative Filtering

Our tag recommendation approach draws ideas from collaborative filtering, and hence, here we briefly review the most relevant works on tag recommendation using collaborative filtering. Xu et al. [2006] designed a collaborative filtering approach to suggest high-quality tags for Web objects, according to several criteria (coverage, popularity, effort, uniformity). The authors employed a co-occurring strategy and considered that if two tags frequently co-occur when describing a specific object, they should also co-occur in the recommended set of tags. A similar approach was presented later by Sigurbjörnsson and van Zwol [2008], who recommended tags for Flickr images. They used knowledge from the Flickr community and applied it in a co-occurring strategy. Specifically, given a user-input tag, they considered the tags co-occurring with it as good candidates for recommendation. Peng et al. [2010] designed a novel technique for collaborative filtering in social tagging systems, in which all the interactions among users, items and tags are leveraged. They generated joint item-tag recommendations for users, where the tags represent topics from an item (i.e., a web resource) in which the user may be interested. Seitlinger et al. [2013] used a model of human category learning (i.e., ALCOVE) for social tags recommendation. The model uses semantic information regarding a user-specific bookmark (e.g., Wikipedia categories or LDA topics). Tags are predicted to a user by applying the semantic information to a connectionist network with three layers, which simulates the user's categorization and the bookmark formalization.

Recently, several works have been proposed to recommend tags for visual content types [Gong and Zhang 2016; Liu et al. 2014; Nguyen et al. 2017; Seah et al. 2018; Toderici et al. 2010; Zhang et al. 2017]. For example, Liu et al. [2014] explored locations to recommend tags to images. [Toderici et al. 2010] proposed a system to automatically recommend tags to YouTube videos based on their audio-visual content. Gong and Zhang [2016] adopted CNNs to recommend hashtags for microblogs. Zhang et al. [2017] proposed a co-attention network incorporating textual and visual information to recommend hashtags for multimodal tweets. Nguyen et al. [2017] presented a personalized content-aware image tag recommendation approach that combines both historical tagging information and

image-based features in a factorization model. Seah et al. [2018] concurrently generated ranked lists of comments and tags of a social image based on their joint relevance to the visual features, user comments, and user tags.

In contrast to these works, we recommend privacy-aware tags for images shared online.

2.3 Online Image Privacy

The rapid increase in images uploaded on the Web intrigued researchers to focus on establishing adequate privacy models to help protect users' sensitive information. Researchers also provided public awareness of privacy risks associated with images shared online [Henne et al. 2013; Xu et al. 2015]. Along this line, several works were carried out to study users' privacy concerns in social networks, privacy decisions about sharing resources, and the risk associated with them [Ghazinour et al. 2013; Gross and Acquisti 2005; Ilia et al. 2015; Krishnamurthy and Wills 2008; Parra-Arnau et al. 2014; Parra-Arnau et al. 2012; Simpson 2008; Song et al. 2018]. Additionally, several works on privacy analysis examined privacy decisions and considerations in mobile and online photo sharing [Ahern et al. 2007; Besmer and Lipford 2009; Gross and Acquisti 2005; Jones and O'Neill 2011]. For example, Ahern et al. [2007] studied the effectiveness of location information and tags in predicting privacy settings of images. They also conducted a study to verify whether the visual features are relevant to an image's privacy and found that content is one of the discriminatory factors affecting image privacy, especially for images depicting people. This supports the core idea underlying our work: that tags depicting private categories obtained from image content are pivotal for identifying the sensitive content from the search results. For example, tags such as "wedding," "bride," "people" describing a wedding event (private category) represent the private class that particular categories of image content are pivotal for identifying the sensitive content from the search results in establishing users' images sharing decisions.

Automated image privacy approaches are explored along following lines of research:

Social group based approaches. Several works emerged to provide the automated privacy decisions for images shared online based on the social groups or circles [Adu-Oppong et al. 2008; Bonneau et al. 2009a,b; Christin et al. 2013; Danezis 2009; Fang and LeFevre 2010; Joshi and Zhang 2009; Kepez and Yolum 2016; Klemperer et al. 2012; Mannan and van Oorschot 2008; Pesce et al. 2012; Petkos et al. 2015; Squicciarini et al. 2012, 2015, 2009; Watson et al. 2015; Yuan et al. 2017; Zerr et al. 2012]. These social group based approaches mostly consider the user trustworthiness, but ignore the image content sensitiveness, and thus, they may not necessarily provide appropriate privacy settings for online images as the privacy preferences might change according to sensitiveness of the image content.

Visual-based approaches. Several works use visual features derived from the images' content and show that they are informative for predicting images' privacy settings [Buschek et al. 2015; Chandra et al. 2018; Dufaux and Ebrahimi 2008; Hu et al. 2016a; Kuang et al. 2017; Nakashima et al. 2011, 2012, 2016; Orekondy et al. 2018; Shamma and Uddin 2014; Squicciarini et al. 2014, 2017a; Tonge and Caragea 2016, 2018; Tran et al. 2016; von Zezschwitz et al. 2016; Wu et al. 2018; Yu et al. 2017, 2018; Yuan et al. 2018; Zerr et al. 2012; Zhang et al. 2005]. Given the recent success of CNNs, several researchers [Kuang et al. 2017; Tonge and Caragea 2016, 2018; Tran et al. 2016; Yu et al. 2017, 2018] showed promising privacy prediction results compared with visual features such as SIFT and GIST. Using CNNs, some works also started to explore personalized privacy prediction models [Orekondy et al. 2017; Spyromitros-Xioufis et al. 2016; Zhong et al. 2017]. In this context, it is worth mentioning that CNNs were also used in another body of privacy related work such as multi-party privacy conflict detection [Zhong et al. 2018] and automatic redaction of sensitive image content [Orekondy et al. 2018].

Tag-based approaches. Previous work in the context of tag-based access control policies and privacy prediction for images [Apostolova and Demner-Fushman 2009; De Choudhury et al. 2009; Klemperer et al. 2012; Kurtan and Yolum 2018; Mannan and van Oorschot 2008; Pesce et al. 2012; Ra et al. 2013; Squicciarini et al. 2012, 2015, 2017b; Vyas et al. 2009; Yeung et al. 2009; Zerr et al. 2012] showed initial success in tying user tags with access control rules. For example, Squicciarini et al. [2012, 2017b], Zerr et al. [2012], and Vyas et al. [2009] explored learning models for image privacy prediction using user tags and found that user tags are informative for predicting images' privacy. However, the scarcity of tags for many online images [Sundaram et al. 2012] and the workload associated with user-defined tags preclude an accurate analysis of images' sensitivity based on this dimension. Recently, in our prior work [Tonge and Caragea 2016, 2018; Tonge et al. 2018b], we showed that the images' tags that are automatically obtained from the visual content of images using Convolutional Neural Networks (CNNs) can improve the performance of image privacy prediction. Yet, since the CNNs are trained on ImageNet (1.2M+ images labeled with 1000 object categories) [Russakovsky et al. 2014] and Places2 (which contains 365 scene classes with 2.5 million images) [Zhou et al. 2016], these tags depict objects or scenes given in the image and fail to capture the privacy characteristics (or orientation) of the image while generating the tags.

To this end, drawing ideas from collaborative filtering, we recommend privacy-aware tags for online images that have the potential to improve the set of user tags for online image sharing.

3 PRIVACY-AWARE IMAGE TAG RECOMMENDATION

Our approach to recommending privacy-aware tags for newly posted images in online content sharing sites is inspired from collaborating filtering (CF) [Shi et al. 2014]. Particularly, in user-item CF, items are recommended to users by finding the most similar users to the target user (from the user-item matrix) and recommending items to the target user based on the items that the similar users purchased/seen. The large amounts of images posted on the Web in recent years facilitate the study of potential relationships between images and tags. Our approach leverages these ideas to exchange tags between similar images. The analogy with conventional CF methods is that images correspond to users and tags correspond to items (i.e., in our setting, we deal with an image-tag matrix). Specifically, we aim to recommend tags for a target image by transferring privacy-aware tags from its most similar images, which are obtained from a large collection. We base our models on the assumption that *privacy-aware similar images possess similar tags.*

Algorithm 1 describes the process in detail. Specifically, the nearest neighbors of a target image are found by comparing rows in the image-tag matrix. Recommendations are made for the target image based on the neighboring images' tags (as a privacy-aware weighted sum of occurrences of tags). A common problem in CF is the *cold start* problem [Su and Khoshgoftaar 2009]. In our case, this refers to images that have very few tags or no tags at all, and hence, there is not enough information available to find accurate nearest neighbors for a target image. However, in our domain, images can be represented using two different views or feature types: (1) image content; and (2) image tags. We take advantage of both of these views (as shown in Algorithm 1).

The input of the algorithm is a dataset $\mathcal{D} = \{I_1, \cdots, I_n\}$ of images and their associated sets of tags, $\{T_1, \cdots, T_n\}$, respectively; a target image I and its set of tags T, which could possibly be empty; $pr(I)$ the privacy label of I, which could be *private* or *public*; k the number of nearest neighbors of I from $\mathcal{D}$; and r the number of tags to be recommended. The output of the algorithm is a ranked list of r tags, which are recommended for the target image. The algorithm starts by checking if the set of tags T corresponding to the target image I is empty (Alg. 1, line 5). If $T \neq \phi$, the similarities between I and all images in $\mathcal{D} \setminus \{I\}$ are computed based on images' tags (Alg. 1, lines 13-18). The top k most similar images to I are returned (Alg. 1, lines 20-21) and the candidate set that represents the union of the sets of tags extracted from these k similar images is ranked inside the subroutine

Algorithm 1 Tag Recommendation

1: **Input**: A dataset $\mathcal{D} = \{I_1, \cdots, I_n\}$ of images and their sets of tags $\{T_1, \cdots, T_n\}$; a target image I and its set of tags T; $pr(I)$ the privacy label of the target image I (could be private or public); k the number of nearest neighbors of I from $\mathcal{D}$; r the number of tags to be recommended.
2: **Output**: A set R of recommended tags for I.
3: $R \leftarrow \phi$; // the set of recommended tags, initially empty.
4: $S \leftarrow \phi$;
5: **if** $T = \phi$ **then** // if the set of tags is empty.
6: $\quad \mathbf{x} \leftarrow$ `ImageContentEncoding(I)`; // deep features for I
7: $\quad$ **for all** $I_j \in \mathcal{D}$ **do**
8: $\quad\quad \mathbf{x}_j \leftarrow$ `ImageContentEncoding(I_j)`; // deep features for I_j
9: $\quad\quad s_j \leftarrow$ `similarity(x, x_j)`; // compute the visual content similarity between I and I_j
10: $\quad\quad S \leftarrow S \cup (I_j, s_j)$; // store I_j and its similarity with I
11: $\quad$ **end for**
12: **else**
13: $\quad \mathbf{x} \leftarrow$ `ImageTagEncoding(I)`; // get tags' features of I
14: $\quad$ **for all** $I_j \in \mathcal{D}$ **do**
15: $\quad\quad \mathbf{x}_j \leftarrow$ `ImageTagEncoding(I_j)`; // get tags' features of I_j
16: $\quad\quad s_j \leftarrow$ `similarity(x, x_j)`; // compute the tags similarity between I and I_j
17: $\quad\quad S \leftarrow S \cup (I_j, s_j)$; // store I_j and its similarity with I
18: $\quad$ **end for**
19: **end if**
20: $S.similarities.sort()$; // sort the images in decreasing order of their similarity scores
21: $S \leftarrow$ top k (I_j, s_j) entries; // get k images with the highest similarities with I, and their similarities
22: $W \leftarrow$ `TagRanking(S, pr(I))`; // rank the tags from S images
23: $R \leftarrow r$ tags with the highest scores from W;
24: **return** R

for tag ranking (line 22). The tag ranking subroutine is described in Algorithm 2. The most highly ranked r tags from the candidate set are returned as recommended tags for the target image I (Alg. 1, line 23-24). For the cold start setting, if the initial tag set is empty, i.e., $T = \phi$ for image I, Algorithm 1 recommends r tags from the k most similar images in $\mathcal{D}$, where, this time, the similarity is computed based on image content features (not tags) (Alg. 1, lines 5-12).

For each tag in the candidate set, we compute its score as the privacy-aware sum of similarities between the target image and its similar images that contain that tag (Alg. 2, lines 6-12). This weighting method was employed based on the assumption that a "good" tag is very likely to be exchanged between similar images. Specifically, the weight (or score) of a tag t, w_t, is computed as:

$$w_t = \sum_{j \in \mathcal{S}} c_{jt} \cdot s_j \cdot P(t|pr(I)) \tag{1}$$

where $\mathcal{S}$ represents the neighborhood of I, i.e., its k most similar images from $\mathcal{D}$, c_{jt} is an indicator variable, which is 1 if tag t belongs to the tag set T_j of image I_j from $\mathcal{S}$ and 0 otherwise, and s_j is the similarity between image I_j and I. The probability $P(t|pr(I))$ is the likelihood of the tag t belonging to one of the privacy classes (i.e., public or private) corresponding to the privacy of the target image I. For instance, if I is of private class, then $P(t|pr(I))$ gives the probability of tag t belonging to the set of private images. In experiments, the likelihood is calculated based on the dataset $\mathcal{D}$. We wish to obtain privacy-aware tags, i.e., tags weighted by their likelihood of occurrence in private

Algorithm 2 Tag Ranking

```
1: function TagRanking(S, pr(I))
2:     W ← ϕ; // the set of tags and their scores, initially empty.
3:     for all I_j ∈ S do
4:         T_j ← I_j.tags // get the set of tags of image I_j.
5:         s_j ← I_j.similarity // similarity of target image and I_j.
6:         for all t ∈ T_j do
7:             w_t ← W.scoreOf(t) // w_t stores the score of t
8:             if w_t = null then // if tag t is not in W already
9:                 W ← W ∪ (t, 0) // add t to W
10:            end if
11:            w_t ← w_t + s_j · P(t|pr(I)) //score of t weighted by privacy
12:        end for
13:    end for
14:    W.scores.sort() // sort the scores in W in the decreasing order.
15:    return W.
16: end function
```

or public classes, without missing out on the high-quality tags. Thus, we consider privacy-aware similarity that relies on the privacy likelihood of the tag instead of considering a privacy-enforced similarity. Here, we define privacy-enforced similarity as a similarity that considers privacy as an additional parameter in the image similarity, i.e., tags could be exchanged between images of the same privacy class (either public or private). A similarity weighted with privacy likelihood favors tags with a given privacy setting as opposed to the privacy-enforced similarity that would enforce tags of the same privacy settings as the target image. For example, using privacy-enforced similarity, for Figure 1(b) (given its public nature), tags such as "Women," "Girl" (inclined to private class) would not be recommended. Conversely, privacy-aware weights can help obtain tags that are descriptive of an image content and help in identifying appropriate sharing needs of the image as it considers both image's content and the privacy aspect of the image.

Figure 2 shows the illustration of the privacy-aware tag recommendation algorithm through an example. We consider a newly uploaded target image I on the Web that is of public class and has an incomplete set of user-input tags. For this illustration, we use visual content features to compute the similarity between the target image I and the images from the collection $\mathcal{D}$ (shown in Figure 2 with a blue cylinder). Note that the images' tags can be used to compute the similarity as well (as discussed in Alg. 1). The top $k = 5$ similar images are shown in the figure where the similarity decreases from left to right (the most similar image is labeled as (1)). Using these similar images, we obtain the set of candidate tags for which we compute privacy-aware weights. The candidate tags and their privacy-aware weight calculation is shown in Table 1. For illustration purposes, we use $s_j = 1$ in Eq. 1, instead of the actual similarity between the target image and images in $\mathcal{D}$ (where $0 \leq s_j \leq 1$). As we can see from Table 1, the tag "Cute" occurred in all five similar images, once in each image (see the column labeled as "Count"). The tag "Cute" is highlighted in blue color in Figure 2. Given that the target image is annotated as public, the tag "Cute" is weighted by the privacy likelihood $P(Cute|public)$, which is 0.3 (see Table 1). Recall that $P(t|private)$ and $P(t|public)$ are calculated from $\mathcal{D}$. Thus, the privacy-aware weighted sum of occurrences is given as 1.5. Table 1 shows the calculations for privacy-aware weighted sum of tag occurrences. Likewise, final weights are calculated for all candidate tags and top $r = 3$ tags are recommended for the target image (the recommended tags are shown in bold font in Figure 2 and Table 1). Note that since we consider

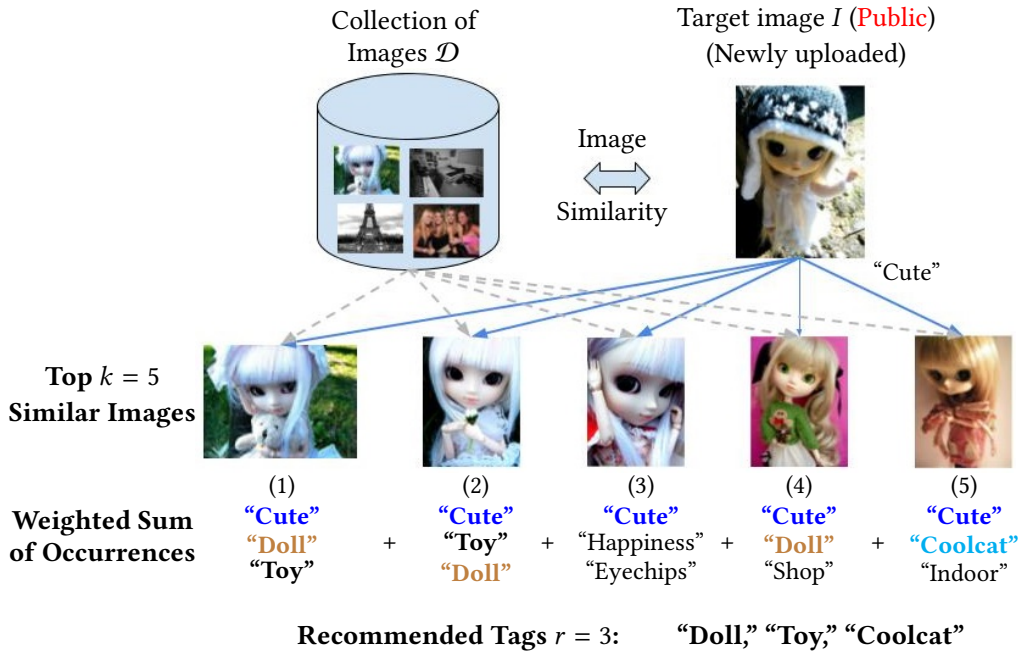


Fig. 2. Illustration of the privacy-aware tag recommendation algorithm using an example: 1) A newly uploaded image on the Web that has an incomplete set of user-input tags, i.e., {"Cute"}, is considered as the target image I. 2) We can use images' tags or content features to compute the similarity between the target image I and the images from the collection $\mathcal{D}$. For this example, we use visual content features to compute the similarity. 3) Top $r = 3$ tags { "Doll," "Toy," "Coolcat"} are recommended using top $k = 5$ similar images, through our privacy-aware tag recommendation approach. Note that the recommended tags "Doll" and "Toy" are appropriate tags for the target image I and can help correctly characterize its privacy class as *public*.

Candidate Tags	Count	$P(t\|pr = private)$	$P(t\|pr = public)$	$w_t, s_j = 1$
Doll	**3**	**0.1**	**0.9**	**3 × 0.9 = 2.7**
Toy	**2**	**0.15**	**0.85**	**2 × 0.85 = 1.7**
Cute	**5**	**0.7**	**0.3**	**5 × 0.3 = 1.5**
Coolcat	**1**	**0.0**	**1.0**	**1 × 1.0 = 1.0**
Shop	1	0.0	1.0	1 × 1.0 = 1.0
Eyechips	1	0.3	0.7	1 × 0.7 = 0.7
Indoor	1	0.6	0.4	1 × 0.4 = 0.4
Happiness	1	0.6	0.4	1 × 0.4 = 0.4

Table 1. Privacy-aware weighted sum of tag occurrences ($k = 5$) given that the target image is public. Bold words indicate the top $r = 3$ tags. Since the tag "Cute" appears already in the original set of user tags, we add the next important tag from the ranked list, i.e., "Coolcat." The tags with same weights are selected randomly.

privacy-likelihood of the tag instead of privacy-enforced similarity with the target image, the tag "Cute" describing the image content is recommended to the target image even though the tag "Cute"

Dataset	#Total Images	#Avg. Tags	#min. Tags	#max. Tags	#Private Images	#Public Images
$\mathcal{D}$	8000	9.73	1	71	2000	6000
$\mathcal{I}^E$ (Images I for Evaluation)	4189	18.65	11	78	1047	3142

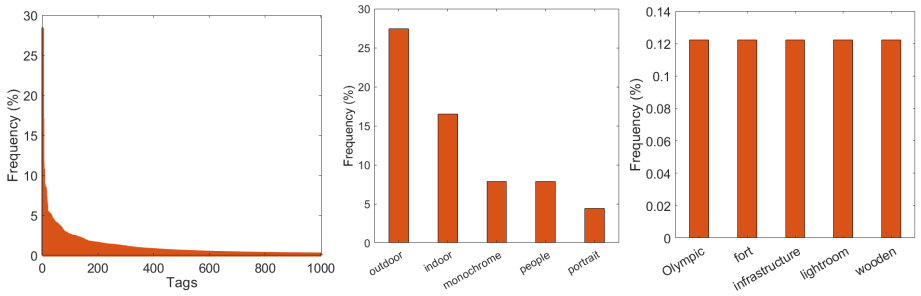
Table 2. Datasets summary.

has privacy-related ("private") connotations. However, since the tag "Cute" appears already in the original set of user tags, we do not add it to our set of recommended tags (to avoid over-counting), and add the next tag from the ranked list. We select the next tag with highest weight, i.e., "Coolcat" (shown in bold font in Table 1). The tags with the same weights are selected randomly.

4 DATASET

We explore the effectiveness of the privacy-aware recommended tags for: (1) their ability to predict the private or sensitive content of online images; and (2) their relevancy to the images' content. Hence, we evaluate our recommendation algorithm on Flickr images sampled from the PicAlert dataset, made available by Zerr et al. [2012]. The PicAlert dataset contains both user-input tags and privacy labels. PicAlert is comprised of Flickr images on various subjects, which are manually labeled as *private* or *public*. The dataset contains photos uploaded in Flickr during the period from January to April 2010. The images have been labeled by six teams providing a total of 81 users of ages between 10 and 59 years. The guideline to select the label is given as: private images belong to the private sphere (like self-portraits, family, friends, someone's home) or contain information that one would not share with everyone else (such as private documents). The remaining images are labeled as public. Each image was shown to at least two different users. In the event of disagreement, the photos were presented to additional users.

We split the PicAlert dataset into two subsets. The first subset corresponds to the dataset $\mathcal{D}$ from Alg. 1 and is a collection of 8,000 images, labeled as private or public, that are used to recommend tags for the target images. We refer to this subset as $\mathcal{D}$. The second subset corresponds to target images that we use for evaluation and consists of 4,189 images from PicAlert, also labeled as private or public. We refer to this subset as $\mathcal{I}^E$ or images I for evaluation. The ratio of public to private images in both the subsets $\mathcal{D}$ and $\mathcal{I}^E$ is 3 : 1. Table 2 shows a summary (number of total images, the average number of tags per image, the minimum number of tags per image, the maximum number of tags per image, number of private and public images) of these datasets. For each image I in $\mathcal{I}^E$, we randomly split its set of tags into two subsets (i.e., *visible* and *hidden*). The motivation behind using random split is that newly uploaded images might have an incomplete and/or noisy set of user-input tags [Sundaram et al. 2012] and we desire to know if the proposed algorithm can overcome these challenges. The *visible* subset is denoted by T in Alg. 1 and is used to compute the similarity between the *visible* subset of the target image I in $\mathcal{I}^E$ with the original set of tags of images in $\mathcal{D}$. The *hidden* subset is considered as gold standard for the evaluation of recommended tags. To calculate the precise similarity between two images using tags, we want to have at least five tags in the set of visible tags. Hence, we consider images with a number of user tags greater than 10 for the dataset $\mathcal{I}^E$ (see Table 2, #minimum tags). In case less than 10 tags are available for an image, we can use the image content similarity. We filter out stop words, numbers, URLs, words with length less than 3 characters, and words with document frequency less than 2. After preprocessing, the size of the vocabulary is reduced to $\approx 19,000$. Note that for similarity computation (cosine in our experiments), we used the stemmed version of tags and synonyms obtained from WordNet [Fellbaum 1998]. We also plot the frequency of top 1,000 tags normalized by the dataset size in Figure 3(a). The plot shows that top 200 tags befall in 3% – 30% of the data with very few tags occurring in around 20% of the dataset. Note that most of the tags occur below 3% showing the



(a) Tag frequency of top 1000 tags (b) Top 5 frequent tags (c) Less 5 frequent tags (from 1000 tags)

Fig. 3. Tag frequency (%) in the PicAlert dataset. The frequencies are normalized by the size of the dataset.

variation in the images' subjects and the complexity of our dataset. We also show the top 5 most frequent and less frequent tags with their frequency in Figures 3(b) and 3(c), respectively. Note that the frequencies of tags are normalized by the dataset size.

5 EXPERIMENTS AND RESULTS

In this section, we evaluate the tags obtained by the privacy-aware tag recommendation algorithm for images in $\mathcal{I}^E$, by transferring tags from their most similar images from $\mathcal{D}$ in several settings. That is, the quality of recommended tags is determined by: *(1) whether these tags hint to specific image privacy settings*; and *(2) whether these tags are good enough to describe the content of an image.* Hence, we adopt two evaluation mechanisms: (1) we examine the performance of models trained on the recommended tags combined with the original tags (when available) for image privacy prediction to determine their ability in building more robust models for identifying private or sensitive content for online image sharing; and (2) we compare the recommended tags against the ground-truth, i.e., the *hidden* set of tags, and also evaluate their quality through crowd-sourcing. We provide details of these evaluation types below.

Image Privacy Prediction. Similar to prior works on privacy prediction [Squicciarini et al. 2014, 2017a; Tonge and Caragea 2016, 2018; Zerr et al. 2012], we aim at identifying generic privacy patterns using the recommended tags to verify if these tags are indicative of the privacy classes. For this, we split $\mathcal{I}^E$ into two subsets *Train* and *Test* to determine if the recommended tags are able to enhance the training set and learn better privacy characteristics. From $\mathcal{I}^E$, we randomly sample 3,689 images for *Train* and 500 for *Test*. We use *Train* to train Support Vector Machine (SVM) classifiers based on the recommended tags and use *Test* to test these classifiers. We provide the privacy class of images in *Train* as input to Alg. 1 and generate privacy-aware recommended tags for these images by exchanging tags from similar images in $\mathcal{D}$. The similarity between images is computed between the *visible* set of a target image in *Train* and all available tags from an image in $\mathcal{D}$. We train SVM classifiers on these recommended tags of *Train* and evaluate them on the visible tags of the images in *Test*. Note that we do not recommend tags to images in *Test* as we assume that we do not know the privacy class of these images. We use the Weka implementation of SVM classifiers and choose the hyper-parameters that give best performance on *Train* using 10-fold cross-validation. For hyper-parameters, we experimented with $C \in \{0.001, 0.01, 1.0, \cdots, 10\}$, kernels: Polynomial and RBF, the γ parameter in RBF, and the degree d of a polynomial.

Tag relevance. To evaluate the relevancy of the recommended tags, we randomly sample 500 images from $\mathcal{I}^E$. We denote this subset as *DRel*. We recommend privacy-aware tags for images in *DRel* by exchanging tags from similar images in $\mathcal{D}$. The similarity between images is computed between the *visible* set of a target image in *DRel* and all available tags from an image in $\mathcal{D}$. The *hidden* subset is considered as gold standard for evaluation, and contrasted with the predicted tag set. We also conduct a crowd-sourcing experiment to determine whether the recommended tags of the *DRel* dataset are relevant to the image's content.

For all the experiments, we generate five random splits of visible and hidden subsets of tags and report performance (Accuracy, F1-measure, Precision, Recall) averaged over these five splits. We use a Boolean representation of tags, i.e., 1 if a tag is present for an image and 0 otherwise, since tags generally appear only once per image.

5.1 Evaluation of Privacy-Aware Recommended Tags by Privacy Prediction

The performance of privacy-aware recommended tags for image privacy prediction. We first evaluate our privacy-aware recommended tags obtained by the proposed weighting scheme in an ablation experiment for image privacy prediction. Specifically, we compare the performance of SVM classifiers trained only on recommended tags, where the recommended tags are obtained in several settings: (1) by our privacy-aware scoring mechanism, denoted as **p-Weights**, that ranks candidate tags using a privacy-aware weighted sum of tag occurrences (see Eq. 1); (2) recommending privacy-aware tags from the candidate set of tags based on their frequency in the k similar images, without considering images' similarity s_j (see Eq. 1 with $s_j = 1$), denoted as **p-Freq** (privacy-aware); (3) recommending tags by weighted sum of occurrences without considering the privacy likelihood (i.e., Eq. 1 without the term $P(t|pr(I))$), denoted as **Weights** (not privacy-aware); and (4) recommending tags based on their frequency in the k similar images, without considering images' similarity s_j and the privacy likelihood $P(t|pr(I))$, denoted as **Freq** (not privacy-aware). We also compare **p-Weights** with a random approach that recommends r tags randomly from the vocabulary of tags, denoted as **Random** (not privacy-aware).

To compare these methods, we study Algorithm 1 in the setting where each image in *Train* has a seed set of tags associated with it, i.e., $T \neq \phi$ (lines 13-18). The similarity between images is thus computed between the *visible* set of a target image in *Train* and all available tags from an image in $\mathcal{D}$. The similarity between two sets of tags is given as the cosine similarity of the corresponding bag-of-words vectors. We experiment with various numbers of similar images $k = 2, \cdots, 10$, in steps of 1, and recommended tags $r = 5, \cdots, 20$, in steps of 5.

Figure 4 shows the average F1-measure achieved by SVM classifiers using the four ranking strategies for different values of k (number of similar images) and r (number of recommended tags), and the Random naive approach. The SVMs are trained on the recommended tags of the *Train* dataset and evaluated on the visible tags of the *Test* dataset. We can see from the figure that recommended tags obtained using **p-Weights** can learn better privacy characteristics than **Random**, **Weights**, **Freq** (not privacy-aware) and perform comparable to **p-Freq** (privacy-aware), for values of $r = \{10, 15, 20\}$ regardless of the value of k. We also notice that the **p-Weights** scoring mechanism achieves the best performance for $r = 5$ and $k = 4$, outperforming all the other models including **p-Freq**, which shows that all scoring components (s_j and $P(t|pr(I))$) play a role in the overall performance. It is also interesting to mention that **Weights** (not privacy-aware) scoring mechanism consistently performs better than **Freq** (not privacy-aware) scoring method.

In the previous experiment, we only used recommended tags to compare various scoring schemes. Next, we wish to identify how recommended tags perform when we add them to the visible set of images in *Train* for privacy prediction. In what follows, since the results are generally similar for

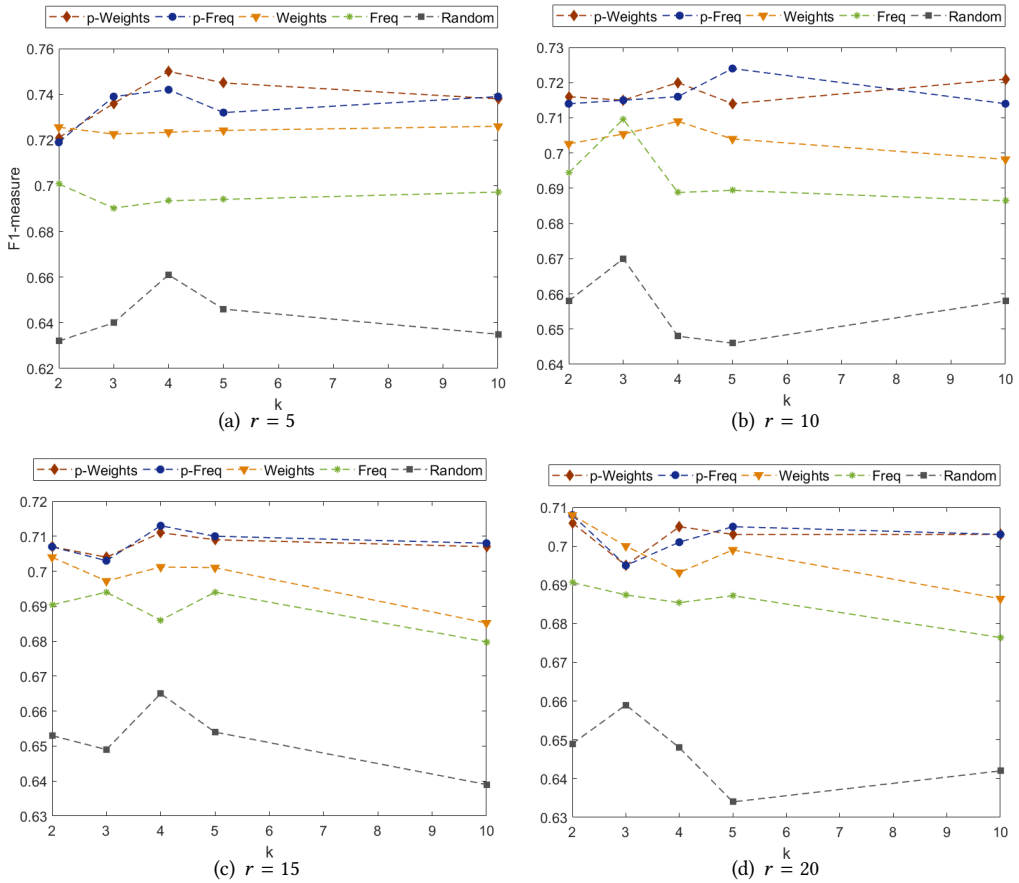


Fig. 4. F1-measure obtained for various parameter values, k and r of Alg. 1. p-Weights and p-Freq are privacy-aware scoring mechanism whereas Weights, Freq and Random are privacy-oblivious scoring mechanisms.

$k = 4, 5, 10$ (see Figure 4), we use these values to augment the set of tags for *Train* with recommended tags by **p-Weights**.

The performance of privacy-aware recommended tags for image privacy prediction when added to the visible tags. Table 3 shows the performance (Accuracy, F1-measure, Precision, Recall) obtained by the SVM classifiers trained on the combination of recommended tags (rt) and visible tags (vt) (as we increase r from 5 to 20) for the images in *Train* and evaluated on the fixed set of visible tags of the images in *Test* (for consistency). The results show that the performance of privacy prediction improves when we add recommended tags to the set of visible tags for images in *Train*. Specifically, we get the best performance when we use $k = 10$ and $r = 5$ with F1-measure of 0.772, whereas the SVM trained on only visible tags achieves 0.743 F1-measure, yielding an improvement of 3% in overall performance. We notice that generally, the performance increases with the decreasing value of r (best performance is given by $r = 5$) and increasing value of k (best performance is given by $k = 10$). This can be justified by the fact that given the diverse nature of the data and the large vocabulary, a large r may introduce noise in the results. Similarly, a high value of k leads to higher number of similar images from which we get a set of good candidate tags.

Features	Acc. %	F1	Precision	Recall
vt	74.83	0.743	0.739	0.748
	$k = 4$			
vt & rt ($r = 5$)	**77.84**	**0.766**	**0.755**	**0.778**
vt & rt ($r = 10$)	77.47	0.763	0.752	0.776
vt & rt ($r = 15$)	77.31	0.757	0.744	0.771
vt & rt ($r = 20$)	76.83	0.754	0.741	0.769
	$k = 5$			
vt & rt ($r = 5$)	**77.96**	**0.769**	**0.758**	**0.781**
vt & rt ($r = 10$)	77.80	0.766	0.755	0.778
vt & rt ($r = 15$)	77.60	0.764	0.752	0.776
vt & rt ($r = 20$)	77.27	0.760	0.747	0.773
	$k = 10$			
vt & rt ($r = 5$)	**78.20**	**0.772**	**0.762**	**0.783**
vt & rt ($r = 10$)	77.80	0.765	0.754	0.777
vt & rt ($r = 15$)	77.92	0.767	0.758	0.778
vt & rt ($r = 20$)	77.43	0.758	0.745	0.771

Table 3. Performance for privacy prediction after adding recommended tags. "vt" denotes a set of visible tags and "rt" denotes a set of recommended tags, e.g., {"cute", "toy", "doll"}. "r" is the number of tags recommended.

The previous experiments used image tags to find the neighborhood of an image. However, not all images on social networking sites have user tags associated with them [Sundaram et al. 2012], and this gives rise to the cold start problem for collaborative filtering. Next, we discuss how we overcome the problem. In the following experiments, we use the privacy-aware weighting scheme **p-Weights**, $k = 10$ and $r = 5$.

5.2 Solution to the Cold Start Problem

Cold start is a challenging problem particularly in many collaborative filtering approaches, where the absence of items (i.e., tags, in our case) that are used to bootstrap the algorithms may theoretically hinder the recommendations to be produced. Hence, we evaluate our approach **p-Weights** for image tag recommendation in the setting where we assume that each image in *Train* has no tags associated with it, i.e., $T = \phi$. This involves recommending tags from visually similar images (lines 5-12 of Alg. 1). The similarity between two images is given as the cosine similarity of the corresponding feature vectors. We consider two types of image features extracted from a deep convolutional neural network (CNN): 1) *deep visual feature* $pool_5$, and 2) *deep tags*. The choice of the features is motivated by their performance for privacy prediction in prior works [Tonge and Caragea 2016, 2018; Tran et al. 2016].

We extract the deep visual features and deep image tags using GoogLeNet architecture [Szegedy et al. 2014], which implements a 22 layer deep network with the Inception architecture. The architecture is a combination of all layers with their output filter bank concatenated to form input for the next stage. We extract visual features $pool_5$ from the layer named as "$pool_5$/drop_7x7_s1" (dropout layer). For deep tags, we use the probability distribution over 1,000 object categories for the input image obtained by applying the softmax function over the last fully-connected layer of the CNN. We consider the top k objects of highest probabilities as *deep tags*. The GoogLeNet network is pre-trained on a subset of the ImageNet dataset [Russakovsky et al. 2014], which is distributed with the CAFFE open-source framework for CNN [Donahue et al. 2013].

Features	Acc.%	F1	Precision	Recall
rt-pool$_5$	**75.74**	0.743	0.729	**0.757**
rt-DT	74.19	0.731	0.725	0.742
vt	74.83	0.743	0.739	0.748
DT	68.54	0.645	0.619	0.685

Table 4. Visual content-based similarity ($k = 10$, $r = 5$).

Table 4 shows the performance of privacy prediction obtained by the SVM models trained on the privacy-aware tags recommended from visually similar images based on pool$_5$ (rt-pool$_5$) and deep tags (rt-DT) for the images in *Train* and evaluated on the visible tags of the images in *Test*. For this experiment, we assume that we do not know the set of visible tags for images in *Train*. However, we wish to examine how would the recommended tags obtained using visual content similarity perform as compared to the visible tags and predicted deep tags (DT) of images in *Train*, as done in our prior work [Tonge and Caragea 2016, 2018]. Thus, we also show the performance of the models trained on visible tags alone (vt) and deep tags (DT) in Table 4. The results show that the models trained on the recommended tags yield similar results to the models trained on visible tags (user-input tags - if we would know them) and outperform those trained on the top k predicted deep tags (from GoogLeNet) for each image in *Train* [Tonge and Caragea 2016, 2018]. Precisely, we obtain maximum value of F1-score as 0.743 and best recall of 0.757 with recommended tags $r = 5$.

From the table, we observe that the models trained on tags recommended from visually similar images calculated based on pool$_5$ (rt-pool$_5$) outperform those trained on tags recommended from visually similar images calculated based on deep tags (rt-DT). The models trained on recommended tags obtained using pool$_5$ also outperform the models trained on the top k predicted deep tags (DT) presented in our prior works [Tonge and Caragea 2016, 2018], that are generated without any tag recommendation (i.e., the exchange of tags from similar images). This can be explained by the fact that the deep tags belong to only 1,000 object categories due to which many relevant tags can not be captured. For example, tags such as "walking" and "culture" are not present in the 1,000 object categories, but may be relevant tags for a given picture.

5.3 The Proposed Approach vs. Prior Privacy Prediction Works

We compare the performance of privacy prediction models trained on the user tags improved by the set of recommended tags with the performance obtained by following prior privacy prediction approaches. Mainly, we compare the performance obtained with the recommended tags with two types of features, viz., visual features (fc$_8$ and PCNH) and tag features (User Tags, Deep Tags, and their combination).

1. fc$_8$ [Tonge and Caragea 2016, 2018]: We consider the model trained on the features extracted from the last fully-connected layer of AlexNet, i.e., fc$_8$ as our baseline, since in our previous work we achieved a good performance using these features for privacy prediction.

2. PCNH privacy framework [Tran et al. 2016]: This framework combines features obtained from two architectures: one that extracts convolutional features (size = 24, referred as Convolutional CNN), and another that extracts object features (size = 24, referred as Object CNN). The Convolutional CNN contains two convolutional layers and three fully-connected layers of size 512, 512, 24, respectively. On the other hand, the object CNN is an extension of AlexNet architecture that appends three fully-connected layers of size 512, 512, and 24, at the end of the last fully-connected layer of AlexNet and forms a deep network of 11 layers. The two CNNs are connected at the output layer. The PCNH framework is first trained on the ImageNet dataset and then fine-tuned on a small privacy dataset.

3. Image Tags: Previous works used user tags (UT) [Squicciarini et al. 2014; Zerr et al. 2012], deep tags (DT) [Tonge and Caragea 2016, 2018] and their combination (UT+DT) [Tonge and Caragea 2016, 2018] for privacy prediction and hence, we consider models trained on these tags as other baselines. Note that we describe deep tags in details in our previous experiment where we evaluated the cold start problem.

Features	Acc.%	F1	Precision	Recall
With Recommended Tags				
RT	75.37	0.75	0.747	0.754
UT+RT	78.20	0.772	0.762	0.783
UT+DT+RT	**81.9**	**0.810**	**0.811**	**0.819**
Visual features				
fc_8 [Tonge and Caragea 2016, 2018]	81.16	0.805	0.803	0.812
PCNH [Tran et al. 2016]	77.91	0.768	0.764	0.779
Tag features				
UT [Squicciarini et al. 2014, 2017a; Zerr et al. 2012]	74.83	0.743	0.739	0.748
DT [Tonge and Caragea 2016, 2018]	68.54	0.645	0.619	0.685
UT+DT [Tonge and Caragea 2016, 2018]	78.81	0.786	0.784	0.789

Table 5. Comparison of privacy prediction performance obtained using the proposed approach and prior privacy prediction approaches.

Table 5 compares the privacy prediction performance obtained with the recommended tags (RT) with the performance obtained by the prior works. The table shows that when we add the recommended tags (RT) to the existing user tags (UT), the F1-measure improves by 3% over the user tags alone. Similarly, when we add the recommended tags (RT) to the combination of user tags and deep tags (UT + DT), we get improvement in the F1-measure of 3% over the combination of user tags and deep tags. We also observe that the model trained on the tag features with the recommended tags (UT+DT+RT) yields a better performance to the models trained on the visual features fc_8 and PCNH. For example, the UT+DT+RT achieves an F1-measure of 0.81, whereas fc_8 and PCNH obtain F1-measure of 0.805 and 0.768, respectively. Even though the tag features do not yield a great improvement over visual features (fc_8), tag features are also essential for the privacy prediction as they provide other aspects of an input image that have not been captured by the visual content. For example, consider an image containing "people with glasses in their hands." Solely using visual content, one cannot differentiate from a "birthday party" to "event launch party." User tags (generated by image owner) can contain such information, which can provide relevant cues for privacy prediction. It is interesting to mention here that, improving user tags with the set of recommended tags reduces the performance gap between the tag and visual features. Visual features and tag features can complement each other, and hence, can be combined to obtain improved privacy prediction performance in the future. Additionally, these privacy-aware tags can predict privacy of an image accurately even when access to the visual content of the image is not allowed due to users reluctance to share the actual image for visual content analysis (which could reveal a user's identity through the face and friends, etc.).

Next, we compare the privacy prediction performance of recommended tags by our approach (privacy-aware) with the tags generated by prior image annotation mechanism (not privacy-aware).

5.4 The Proposed Approach vs. Prior Image Annotation Works

In this experiment, we compare the performance of privacy prediction using tags recommended by the proposed approach "p-Weights" against the tags recommended by prior nearest neighbors based image annotation works. Particularly, we consider the modified version of "Fast image tagging" (or FastTag) [Chen et al. 2013] and image annotation by Makadia et al. [2008] as our baselines. We provide details of our baselines as follows.

1. FastTag [Chen et al. 2013]: FastTag addresses the tag sparsity problem, which motivates the choice of FastTag as our baseline. This is particularly critical to our dataset as we can see in Figure 3, very few tags occur in around 20% of the dataset. Additionally, similar to our approach, FastTag also considers images with partial tags to predict tag annotations. For FastTag, authors considered traditional image features such as Gist descriptor [Oliva and Torralba 2001], global color histograms, and bag-of-word visual features. Recently, Mayhew et al. [2016] trained nearest neighbors-based image annotation algorithms using the features derived from CNNs and achieved better performance than using traditional image features. Thus, similar to our approach, we use $pool_5$ (CNN based feature representation) as image features in the FastTag algorithm. For other parameters in FastTag, we consider the best (default) values given by the authors.

2. Makadia et al. [2008]: Similar to our work, Makadia et al. [2008] also transfers tags from the most similar images of a target image and thus, we consider it as our another baseline. The tag transfer mechanism of Makadia et al. [2008] is different from our tag scoring mechanism "p-Weights". Makadia et al. [2008] follows a three step process to transfer tags to a target image from its neighbors. First, the authors rank the tags according to their frequency in the dataset (in our case $\mathcal{D}$). Second, the highest ranking tags of the first neighbor (first similar image) are transferred to the target image. If the number of tags of the first neighbor is greater than r, then only the top r tags are transferred. Last, the tags of neighbors $2, \cdots, k$ are ranked based on two factors: 1. co-occurrence of tags in training ($\mathcal{D}$) with the tags transferred in step 2; and 2. frequency of tags of neighbors $2, \cdots, k$. The highest ranking tags are selected and the remaining tags (r - tags transferred in step 2.) are transferred to the target image. Makadia et al. [2008] also considered color and texture-based visual features (traditional image features). Even in this case, we use $pool_5$ as image features for an unbiased comparison. Similar to our approach, we compute cosine similarity between two visual feature vectors to obtain top $k = 10$ neighbors and recommend $r = 5$ tags.

For this comparison, we include the tags obtained by both the settings when the seed set $T \neq \phi$ (tag similarity, Alg. 1, lines 13-18) and $T = \phi$ (visual content similarity, Alg. 1, lines 5-12). Specifically, we compare the models for privacy prediction trained on the combination of visible tags and recommended tags by the proposed approach with the models trained on the combination of visible tags and the tags obtained by FastTag.

Table 6 shows the privacy prediction performance comparison between the models trained on the combination of visible tags (vt) and the recommended tags (rt) by Alg. 1, FastTag and Makadia et al. [2008]. From the table, we can observe that the models trained on the tags obtained by the proposed approach perform better than the models trained on the tags obtained by FastTag and Makadia et al. [2008]. Specifically, models trained on the combination of visible tags and tags recommended by visual content yield F1-measure as high as 0.752 (Table 6, #3 Visual Content Similarity), whereas the models trained on the combination of visible tags and tags obtained by FastTag and Makadia et al. [2008] get F1-measure of 0.741 and 0.730 respectively. We achieve the best performance of 0.772 (F1-measure) using the tag similarity (Table 6, #4 Tag Similarity). Note that the F1-measure obtained by models trained on the combination of visible tags and the tags obtained by FastTag (0.741) or Makadia et al. [2008] (0.730) (Table 6, #2 Prior Image Annotation Works) is even slightly worse than the F1-measure (0.743) obtained for the models trained on only visible tags (Table 6, #1

Features	Acc.%	F1	Precision	Recall
#1 Original User Tags (Visible Tags)				
vt	74.83	0.743	0.739	0.748
#2 Prior Image Annotation Works				
vt & rt by FastTag	74.55	0.741	0.738	0.745
vt & rt by Makadia et al. [2008]	74.87	0.730	0.723	0.749
#3 Visual Content Similarity ($T = \phi$)				
vt & rt ($r = 5$)	75.23	0.741	0.730	0.752
vt & rt ($r = 10$)	75.63	0.742	0.727	0.757
vt & rt ($r = 15$)	***76.71***	***0.752***	***0.737***	***0.768***
vt & rt ($r = 20$)	76.27	0.747	0.732	0.763
#4 Tag Similarity ($T \neq \phi$)				
vt & rt ($r = 5$)	**78.20**	**0.772**	**0.762**	**0.783**

Table 6. Privacy-aware Tag recommendation vs. Prior Image Annotation Works.

Original User Tags). The results show that even though we use the same set of visual features (deep features) for all the three methods (p-Weights, FastTag and Makadia et al. [2008]) to generate the tags, the tags obtained by FastTag and Makadia et al. [2008], which are privacy-oblivious, are not very helpful for identifying images' private content. Despite that FastTag performs well for general image annotation [Cheng et al. 2018], it fails to recommend privacy preserving tags on the PicAlert dataset because, unlike our approach, the impact of the privacy of an image is not considered.

5.5 Quality Assessment of Recommended Tags

In the above experiments, we compared the privacy prediction performance obtained by privacy-aware and privacy-oblivious tags. In this experiment, we determine which set of recommended tags (privacy-aware vs. privacy-oblivious) describe an image's content appropriately. Precisely, we obtain two sets of recommended tags: (1) using our privacy-aware weighting scheme, referred as privacy-aware tags (see Eq. 1), and (2) using weighting scheme without privacy likelihood, referred as privacy-oblivious tags (Eq. 1 without the term P(t|pr(I))). We compare these tags against the ground-truth (i.e., the *hidden* set of tags). For this experiment, we recommend tags (using both privacy-aware and privacy-oblivious weighting schemes) for images in *DRel*, where each image has a seed set of tags associated with it, i.e., $T \neq \phi$. For each image in *DRel*, we randomly split its set of tags into two subsets, i.e., visible and hidden, where the visible set is used for tag similarity and the hidden set is used as gold-standard set. The similarity between images is thus computed between the *visible* set of a target image in *DRel* and all available tags from an image in $\mathcal{D}$.

Table 7 shows the performance (Precision@r) obtained for $r \in \{1, \cdots, 10\}$ tags recommended for the images in *DRel* when compared against the gold-standard set of tags (those that are hidden from the original user tags). We compute Precision as the total number of *recommended* and *relevant* tags over the number of tags recommended (i.e., r). The results show that the privacy-aware tags obtain better precisions than the privacy-oblivious tags, yielding the highest precision of 0.197 ($r = 4$) using gold-standard. The gold-standard set is nothing but a subset of user annotated tags, which may not provide all the possible tags that can be associated with an image content. Hence, the gold-standard set may fail to capture highly relevant tags provided by the recommendation strategy. For example, in Figure 5, we can see that tags relevant to the image content (shown in italic) are recommended, but do not appear in the user-input tags. Specifically, even though tags such as *culture, street, walking* are consistent with the image content, these tags are not considered

Fig. 5. Image with recommended tags, r=10.

#Tags	Gold-standard		Crowd-sourcing	
(r)	Privacy-aware (PA) P@r	Privacy-oblivious (PO) P@r	Privacy-aware (PA) P@r	Privacy-oblivious (PO) P@r
1	0.162	0.182	**0.87**	0.863
2	0.186	0.182	0.85	0.84
3	0.195	0.180	0.812	0.822
4	**0.197**	0.186	0.791	0.793
5	0.190	0.184	0.77	0.77
6	0.184	0.178	0.753	0.75
7	0.174	0.169	0.742	0.738
8	0.168	0.164	0.731	0.72
9	0.162	0.158	0.72	0.71
10	0.156	0.153	0.71	0.704

Table 7. Gold-standard and User evaluation of privacy-aware and privacy-oblivious recommended tags.

for calculating the precision values since they do not appear among the tags in the *hidden* set or gold-standard set.

Crowd-sourcing can be used to address the above limitation. Hence, we employ crowd-sourcing to make use of the "wisdom of the crowd," as follows: we use two annotators from Figure Eight[1] to determine if the recommended tags are relevant to images' content. For each tag, annotators were asked to choose between: *relevant*, *irrelevant* and *not sure*. To calculate precision values, we consider a tag as *Relevant* if at least one annotator marked it as *relevant* as the tags can be subjective and one annotator can observe more in an image than the other.

Table 7 also shows the performance obtained through crowd-sourcing. We notice that the results of crowd-sourcing are higher than those obtained by relying only on *gold standard* to compute the performance. Precisely, through crowd-sourcing, the precision increased from 0.197 (gold-standard set) to 0.87 for privacy-aware tags, reassuring that the generated tags are relevant to images' content. Similarly, for privacy-oblivious tags, the precision increased from 0.182 to 0.863. The difference in the results can be justified by the fact that the user tags tend to be noisy, incomplete, and may not relate to the image content [Sundaram et al. 2012]. We observe that, for the crowd-sourcing

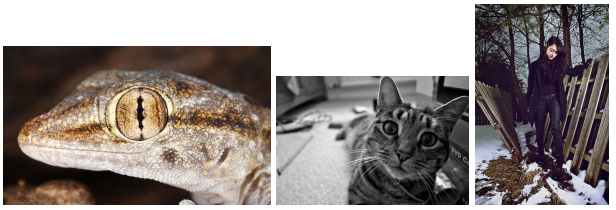
[1]https://make.figure-eight.com/

#Tags	Nouns Only		Verb Only		Adjective Only		Noun & Verb	
	PA	PO	PA	PO	PA	PO	PA	PO
(r)	P@r	P@r	P@r	P@r	P@r	P@r	P@r	P@r
1	**0.812**	0.808	**0.64**	0.626	**0.872**	0.865	**0.815**	0.805
2	0.792	0.796	0.633	0.626	0.849	0.848	0.792	0.790
3	0.778	0.776	0.638	0.626	0.848	0.846	0.780	0.780
4	0.756	0.750	0.638	0.626	0.848	0.846	0.755	0.753
5	0.741	0.735	0.638	0.626	0.848	0.846	0.742	0.736
6	0.737	0.730	0.638	0.626	0.848	0.846	0.735	0.729
7	0.735	0.728	0.638	0.626	0.848	0.846	0.734	0.726
8	0.734	0.727	0.638	0.626	0.848	0.846	0.733	0.726
9	0.734	0.727	0.638	0.626	0.848	0.846	0.732	0.725
10	0.734	0.727	0.638	0.626	0.848	0.846	0.726	0.731

Table 8. User evaluation of recommended tags that are Noun, Verb, Adjective, and Noun & Verb. Privacy-aware tags are denoted as PA and privacy-oblivious are denoted as PO.

experiment, precision obtained using privacy-aware tags is higher than the precision obtained using privacy-oblivious tags for $r = \{1, 2, 7-10\}$. Note that for r ranging from 3 to 6, the performance of privacy-aware tags is comparable to the performance of privacy-oblivious tags. One reason could be that some relevant tags have higher weights and are recommended irrespective of their privacy likelihood. Consider a private image of "people on the beach" for which "beach" (being considered as nature) would be recommended even though it has higher likelihood towards the public class.

The tags depicting objects (such as beach, furniture) or actions (such as walking) in images are more objectively identified by annotators, whereas abstract tags such as "beautiful," "pretty," etc., are more subjective. This could be another justification for the similar results that we obtain for privacy-aware and privacy-oblivious tags for values of $r = \{3-6\}$ in Table 7. To understand this, we further investigate both privacy-aware and privacy-oblivious sets of tags by obtaining part-of-speech (POS) tags for the recommended tags. The recommended tags for *DRel* contain approximately 45% of nouns, 4% of verbs, 5% of adjective POS tags, and the remaining are the proper nouns (44%). Table 8 shows the user evaluation of the recommended tags (privcay-aware and privacy-oblivious) that are nouns, verbs, adjectives, and nouns & verbs. Note that we do not consider proper nouns as solely from the visual content (without user's information), it is difficult to identify whether a particular place or a person is relevant to a target image. For example, one can recognize a "beach" from the visual content of a target image, but for some images it is difficult to know the exact location (i.e., a proper noun) of the beach (e.g., oregon coast). In the table, the privacy-aware tags are denoted as "PA", and privacy-oblivious tags are denoted as "PO". The table shows that for nouns (that depict objects and scenes in the image), privacy-aware tags obtain higher performance than the privacy-oblivious tags for almost all values of r. Similarly, for verbs (that depict actions of the objects in the image), privacy-aware tags yield higher performance than the privacy-oblivious tags. Note that images might not have more than $1-2$ verbs; thus the performance does not change after $r = 3$. Conversely, for adjectives, we observe that the performance is comparable for both sets of tags. One reason might be that the adjectives are subjective and even though the privacy-aware tags have recommended good adjective tags, those are not reflected in the performance for values of $r = \{3-6\}$ in Table 7. To illustrate this, we provide some examples in Figure 6 that contain subjective adjectives (tags). For example, for image (a), some people might identify that the shot was taken beautifully, and hence, they might consider tag "beautiful" as relevant tag for the image. On the other hand, others might find the animal scary and they might not consider the tag relevant.



(a) "beautiful" (b) "pretty" (c) "picturesque"

Fig. 6. Subjective Adjective (Tags)

6 CONCLUSIONS

We proposed an approach to recommending privacy-aware image tags that can improve the original set of user tags and, at the same time, preserve images' privacy to help reduce the private content from the search results. Our approach draws ideas from collaborative filtering (CF). Although the user-input tags are prone to noise, we were able to integrate them in our approach and recommend accurate tags. More importantly, we simulated the recommendation strategy for newly-posted images, which had no tags attached. This is a particularly challenging problem, as in many CF approaches, the absence of items (tags in our case) may theoretically hinder the recommendations to be produced, due to the lack of enough information available to find similar images to a target image. Through our experiments, we showed that we achieve better performance for image privacy prediction with recommended tags than the original set of user tags, which in turn indicates that the suggested tags comply to the images' privacy. We also show that improving user tags with a set of privacy-aware recommended tags can reduce the performance gap between the tag and visual features for privacy prediction. Visual features and tag features can complement each other, and hence, can be combined to obtain improved privacy prediction performance in the future. Last, we conducted a user evaluation to inspect the quality of our privacy-aware recommended tags. The results show that the proposed approach is able to recommend highly relevant tags.

In future work, it would be interesting to study the algorithm for multiple sharing needs of the user such as friends, family, and colleagues by considering privacy likelihood with respect to multi-class privacy settings. We plan to explore alternative ways of computing images' similarity, such as combining information from both tags and visual content. Also, another interesting direction would be to explore image-content features depicting various image subjects such as scene and location, which could lead to more accurate results.

ACKNOWLEDGMENTS

This research is supported by the NSF grant #1421970. Any opinions, findings, and conclusions expressed here are those of the authors and do not necessarily reflect the views of NSF. The computing for this project was performed on AWS. We also thank our reviewers for their feedback.

REFERENCES

Fabeah Adu-Oppong, Casey K. Gardiner, Apu Kapadia, and Patrick P. Tsang. 2008. Social Circles: Tackling Privacy in Social Networks. In *Symposium on Usable Privacy andSecurity (SOUPS)*.

Shane Ahern, Dean Eckles, Nathaniel S. Good, Simon King, Mor Naaman, and Rahul Nair. 2007. Over-exposed?: privacy patterns and considerations in online and mobile photo sharing. In *CHI '07*.

Emilia Apostolova and Dina Demner-Fushman. 2009. Towards Automatic Image Region Annotation - Image Region Textual Coreference Resolution. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Short Papers*. 41–44.

P. Bakliwal and C. V. Jawahar. 2015. Active learning based image annotation. In *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*. 1–4.

B. Bao, T. Li, and S. Yan. 2012. Hidden-Concept Driven Multilabel Image Annotation and Label Ranking. *IEEE Transactions on Multimedia* 14, 1 (Feb 2012), 199–210.

Andrew Besmer and Heather Lipford. 2009. Tagged photos: concerns, perceptions, and protections. In *CHI '09*.

Kerstin Bischoff, Claudiu S. Firan, Wolfgang Nejdl, and Raluca Paiu. 2008. Can All Tags Be Used for Search?. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM '08)*. ACM, New York, NY, USA, 193–202.

Joseph Bonneau, Jonathan Anderson, and Luke Church. 2009a. Privacy Suites: Shared Privacy for Social Networks. In *Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09)*. ACM, New York, NY, USA, Article 30.

Joseph Bonneau, Jonathan Anderson, and George Danezis. 2009b. Prying Data out of a Social Network. In *Proceedings of the 2009 International Conf. on Adv. in Social Network Analysis and Mining (ASONAM '09)*. IEEE Computer Society, 249–254.

Bullguard. 2018. Privacy violations, the dark side of social media. (2018). http://www.bullguard.com/bullguard-security-center/internet-security/social-media-dangers/privacy-violations-in-social-media.aspx.

Daniel Buschek, Moritz Bader, Emanuel von Zezschwitz, and Alexander De Luca. 2015. Automatic Privacy Classification of Personal Photos. In *Human Computer Interaction INTERACT 2015*. Vol. 9297. 428–435.

D. Keerthi Chandra, W. Chowgule, Y. Fu, and D. Lin. 2018. RIPA: Real-Time Image Privacy Alert System. In *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)*. 136–145.

Minmin Chen, Alice Zheng, and Kilian Q. Weinberger. 2013. Fast Image Tagging, In Proceedings of the 30th International Conference on Machine Learning. https://www.microsoft.com/en-us/research/publication/fast-image-tagging/

Qimin Cheng, Qian Zhang, Peng Fu, Conghuan Tu, and Sen Li. 2018. A survey and analysis on automatic image annotation. *Pattern Recognition* 79 (2018), 242 – 259.

Delphine Christin, Pablo SáNchez LóPez, Andreas Reinhardt, Matthias Hollick, and Michaela Kauer. 2013. Share with Strangers: Privacy Bubbles As User-centered Privacy Control for Mobile Content Sharing Applications. *Inf. Secur. Tech. Rep.* 17, 3 (Feb. 2013), 105–116.

Gianluigi Ciocca, Claudio Cusano, Simone Santini, and Raimondo Schettini. 2011. Halfway Through the Semantic Gap: Prosemantic Features for Image Retrieval. *Inf. Sci.* 181, 22 (Nov. 2011), 4943–4958.

George Danezis. 2009. Inferring Privacy Policies for Social Networking Services. In *Proceedings of the 2nd ACM Workshop on Security and Artificial Intelligence (AISec '09)*. ACM, New York, NY, USA, 5–10.

Munmun De Choudhury, Hari Sundaram, Yu-Ru Lin, Ajita John, and Doree Duncan Seligmann. 2009. Connecting Content to Community in Social Media via Image Content, User Tags and User Communication. In *Proceedings of the 2009 IEEE International Conference on Multimedia and Expo (ICME'09)*. IEEE Press, Piscataway, NJ, USA, 1238–1241.

Ivica Dimitrovski, Dragi Kocev, Suzana Loskovska, and Saso Dzeroski. 2011. Hierarchical annotation of medical images. *Pattern Recognition* 44, 10 (2011), 2436 – 2449. Semi-Supervised Learning for Visual Content Analysis and Understanding.

Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. 2013. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. *CoRR* (2013).

F. Dufaux and T. Ebrahimi. 2008. Scrambling for Privacy Protection in Video Surveillance Systems. *IEEE Trans. Cir. and Sys. for Video Technol.* 18, 8 (Aug. 2008), 1168–1174.

Lujun Fang and Kristen LeFevre. 2010. Privacy Wizards for Social Networking Sites. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*. ACM, New York, NY, USA, 351–360.

Christiane Fellbaum. 1998. ed. WordNet: an electronic lexical database. *MIT Press, Cambridge MA* (1998).

L. Feng and B. Bhanu. 2016. Semantic Concept Co-Occurrence Patterns for Image Annotation and Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38, 4 (April 2016), 785–799.

Shaolei Feng, Raghavan Manmatha, and Victor Lavrenko. 2004. Multiple Bernoulli Relevance Models for Image and Video Annotation.. In *CVPR (2)* (2007-01-17). 1002–1009.

Yansong Feng and Mirella Lapata. 2008. Automatic Image Annotation Using Auxiliary Text Information. In *ACL-08: HLT*.

Yansong Feng and Mirella Lapata. 2010. Topic Models for Image Annotation and Text Illustration *(HLT '10)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 831–839.

Yue Gao, Meng Wang, Huanbo Luan, Jialie Shen, Shuicheng Yan, and Dacheng Tao. 2011. Tag-based Social Image Search with Visual-text Joint Hypergraph Learning. In *Proceedings of the 19th ACM International Conference on Multimedia (MM '11)*. ACM, New York, NY, USA, 1517–1520.

Kambiz Ghazinour, Stan Matwin, and Marina Sokolova. 2013. Monitoring and Recommending Privacy Settings in Social Networks. In *Proceedings of the Joint EDBT/ICDT 2013 Workshops (EDBT '13)*. ACM, New York, NY, USA, 164–168.

Arnab Ghoshal, Pavel Ircing, and Sanjeev Khudanpur. 2005. Hidden Markov Models for Automatic Annotation and Content-based Retrieval of Images and Video. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '05)*. ACM, New York, NY, USA, 544–551.

Yunchao Gong, Yangqing Jia, Thomas Leung, Alexander Toshev, and Sergey Ioffe. 2013. Deep Convolutional Ranking for Multilabel Image Annotation. *CoRR* abs/1312.4894 (2013). arXiv:1312.4894 http://arxiv.org/abs/1312.4894

Yuyun Gong and Qi Zhang. 2016. Hashtag Recommendation Using Attention-based Convolutional Neural Network. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI'16)*. AAAI Press, 2782–2788.

D. Grangier and S. Bengio. 2008. A Discriminative Kernel-Based Approach to Rank Images from Text Queries. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 8 (Aug 2008), 1371–1384.

Ralph Gross and Alessandro Acquisti. 2005. Information Revelation and Privacy in Online Social Networks. In *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*. 71–80.

Matthieu Guillaumin, Thomas Mensink, Jakob Verbeek, and Cordelia Schmid. 2009. TagProp: Discriminative Metric Learning in Nearest Neighbor Models for Image Auto-Annotation. In *ICCV*.

Benjamin Henne, Christian Szongott, and Matthew Smith. 2013. SnapMe if You Can: Privacy Threats of Other Peoples' Geo-tagged Media and What We Can Do About It *(WiSec '13)*. 12.

Livia Hollenstein and Ross Purves. 2010. Exploring place through user-generated content: Using Flickr tags to describe city cores. *J. Spatial Information Science* 1, 1 (2010), 21–48.

Y. Hou and X. Zhang. 2015. A geometric constrained HCRF for object recognition. In *2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. 1–4.

D. Hu, F. Chen, X. Wu, and Z. Zhao. 2016a. A Framework of Privacy Decision Recommendation for Image Sharing in Online Social Networks. In *2016 IEEE First International Conference on Data Science in Cyberspace (DSC)*. 243–251.

H. Hu, G. Zhou, Z. Deng, Z. Liao, and G. Mori. 2016b. Learning Structured Inference Neural Networks with Label Relations. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2960–2968.

Sheng-Jun Huang and Zhi-Hua Zhou. 2012. Multi-label Learning by Exploiting Label Correlations Locally. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence (AAAI'12)*. AAAI Press, 949–955.

Panagiotis Ilia, Iasonas Polakis, Elias Athanasopoulos, Federico Maggi, and Sotiris Ioannidis. 2015. Face/Off: Preventing Privacy Leakage From Photos in Social Networks. In *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security (CCS '15)*. ACM, New York, NY, USA, 781–792.

Jiren Jin and H. Nakayama. 2016. Annotation order matters: Recurrent Image Annotator for arbitrary length image tagging. In *2016 23rd International Conference on Pattern Recognition (ICPR)*. 2452–2457.

X. Jing, F. Wu, Z. Li, R. Hu, and D. Zhang. 2016. Multi-Label Dictionary Learning for Image Annotation. *IEEE Transactions on Image Processing* 25, 6 (June 2016), 2712–2725.

Simon Jones and Eamonn O'Neill. 2011. Contextual dynamics of group-based sharing decisions *(CHI '11)*. 10.

James B. D. Joshi and Tao Zhang (Eds.). 2009. *The 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing, CollaborateCom 2009, Washington DC, USA, November 11-14, 2009*. ICST / IEEE.

M. M. Kalayeh, H. Idrees, and M. Shah. 2014. NMF-KNN: Image Annotation Using Weighted Multi-view Non-negative Matrix Factorization. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 184–191.

Berkant Kepez and Pinar Yolum. 2016. Learning Privacy Rules Cooperatively in Online Social Networks. In *Proceedings of the 1st International Workshop on AI for Privacy and Security (PrAISe '16)*. ACM, New York, NY, USA, Article 3, 4 pages.

Peter Klemperer, Yuan Liang, Michelle Mazurek, Manya Sleeper, Blase Ur, Lujo Bauer, Lorrie F. Cranor, Nitin Gupta, and Michael Reiter. 2012. Tag, you can see it! Using tags for access control in photo sharing. In *CHI'12*.

Balachander Krishnamurthy and Craig E. Wills. 2008. Characterizing Privacy in Online Social Networks. In *Proceedings of the First Workshop on Online Social Networks (WOSN '08)*. ACM, New York, NY, USA, 37–42.

Zhenzhong Kuang, Zongmin Li, Dan Lin, and Jianping Fan. 2017. Automatic Privacy Prediction to Accelerate Social Image Sharing. In *Third IEEE International Conference on Multimedia Big Data, BigMM*. 197–200.

Abdurrahman Kurtan and Pinar Yolum. 2018. PELTE: Privacy Estimation of Images from Tags. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '18)*. Richland, SC, 1989–1991.

Victor Lavrenko, R. Manmatha, and Jiwoon Jeon. 2004. A Model for Learning the Semantics of Pictures. In *NIPS 16*.

X. Li, B. Shen, B. Liu, and Y. Zhang. 2016. A Locality Sensitive Low-Rank Model for Image Tag Completion. *IEEE Transactions on Multimedia* 18, 3 (March 2016), 474–483.

X. Li, Y. Zhang, B. Shen, and B. Liu. 2014. Image tag completion by low-rank factorization with dual reconstruction structure preserved. In *2014 IEEE International Conference on Image Processing (ICIP)*. 3062–3066.

Rainer Lienhart, Stefan Romberg, and Eva Hörster. 2009. Multilayer pLSA for Multimodal Image Retrieval. In *Proceedings of the ACM International Conference on Image and Video Retrieval (CIVR '09)*. ACM, New York, NY, USA, Article 9, 8 pages.

Zijia Lin, Guiguang Ding, Mingqing Hu, Yunzhen Lin, and Shuzhi Sam Ge. 2014. Image tag completion via dual-view linear sparse reconstructions. *Computer Vision and Image Understanding* 124 (2014), 42 – 60.

Zijia Lin, Guiguang Ding, Mingqing Hu, Jianmin Wang, and Jiaguang Sun. 2012. Automatic Image Annotation Using Tag-related Random Search over Visual Neighbors. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM '12)*. ACM, New York, NY, USA, 1784–1788.

Z. Lin, G. Ding, M. Hu, J. Wang, and X. Ye. 2013. Image Tag Completion via Image-Specific and Tag-Specific Linear Sparse Reconstructions. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*. 1618–1625.

Jing Liu, Mingjing Li, Qingshan Liu, Hanqing Lu, and Songde Ma. 2009. Image Annotation via Graph Learning. *PR* (2009).

J. Liu, Z. Li, J. Tang, Y. Jiang, and H. Lu. 2014. Personalized Geo-Specific Tag Recommendation for Photos on Social Websites. *IEEE Transactions on Multimedia* 16, 3 (April 2014), 588–600.

Mary Madden. 2012. Privacy management on social media sites. http://www.pewinternet.org/2012/02/24/privacy-management-on-social-media-sites. (2012). [Online; accessed 14-January-2018].

Ameesh Makadia, Vladimir Pavlovic, and Sanjiv Kumar. 2008. A New Baseline for Image Annotation. In *ECCV 2008*.

Mohammad Mannan and Paul C. van Oorschot. 2008. Privacy-enhanced Sharing of Personal Content on the Web. In *Proceedings of the 17th International Conference on World Wide Web*. 487–496.

M. B. Mayhew, B. Chen, and K. S. Ni. 2016. Assessing semantic information in convolutional neural network representations of images via image annotation. In *2016 IEEE International Conference on Image Processing (ICIP)*. 2266–2270.

Venkatesh N. Murthy, Ethem F. Can, and R. Manmatha. 2014. A Hybrid Model for Automatic Image Annotation. In *Proceedings of International Conference on Multimedia Retrieval (ICMR '14)*. ACM, New York, NY, USA, 369:369–369:376.

Yuta Nakashima, Noboru Babaguchi, and Jianping Fan. 2011. Automatic generation of privacy-protected videos using background estimation. In *Proceedings of the 2011 IEEE International Conference on Multimedia and Expo, ICME*. 1–6.

Yuta Nakashima, Noboru Babaguchi, and Jianping Fan. 2012. Intended human object detection for automatically protecting privacy in mobile video surveillance. *Multimedia Syst.* 18, 2 (2012), 157–173.

Yuta Nakashima, Noboru Babaguchi, and Jianping Fan. 2016. Privacy Protection for Social Video via Background Estimation and CRF-Based Videographer's Intention Modeling. *IEICE Transactions* 99-D, 4 (2016), 1221–1233.

Hanh T. H. Nguyen, Martin Wistuba, and Lars Schmidt-Thieme. 2017. Personalized Tag Recommendation for Images Using Deep Transfer Learning. In *Machine Learning and Knowledge Discovery in Databases*. Cham, 705–720.

Y. Niu, Z. Lu, J. Wen, T. Xiang, and S. Chang. 2018. Multi-Modal Multi-Scale Deep Learning for Large-Scale Image Annotation. *IEEE Transactions on Image Processing* (2018), 1–1.

Aude Oliva and Antonio Torralba. 2001. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *IJCV* 42, 3 (May 2001), 145–175.

Tribhuvanesh Orekondy, Mario Fritz, and Bernt Schiele. 2018. Connecting Pixels to Privacy and Utility: Automatic Redaction of Private Information in Images. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.

Tribhuvanesh Orekondy, Bernt Schiele, and Mario Fritz. 2017. Towards a Visual Privacy Advisor: Understanding and Predicting Privacy Risks in Images. In *IEEE International Conference on Computer Vision, ICCV 2017*. 3706–3715.

J. Parra-Arnau, A. Perego, E. Ferrari, J. FornÃľ, and D. Rebollo-Monedero. 2014. Privacy-Preserving Enhanced Collaborative Tagging. *IEEE Transactions on Knowledge and Data Engineering* 26, 1 (Jan 2014), 180–193.

Javier Parra-Arnau, David Rebollo-Monedero, Jordi Forné, Jose L. Muñoz, and Oscar Esparza. 2012. Optimal tag suppression for privacy protection in the semantic Web. *Data Knowl. Eng.* 81-82 (2012), 46–66.

Jing Peng, Daniel Dajun Zeng, Huimin Zhao, and Fei-yue Wang. 2010. Collaborative Filtering in Social Tagging Systems Based on Joint Item-tag Recommendations *(CIKM '10)*. ACM, 809–818.

Yuxin Peng, Zhiwu Lu, and Jianguo Xiao. 2009. Semantic Concept Annotation Based on Audio PLSA Model. In *Proceedings of the 17th ACM International Conference on Multimedia (MM '09)*. ACM, New York, NY, USA, 841–844.

João Paulo Pesce, Diego Las Casas, Gustavo Rauber, and Virgílio Almeida. 2012. Privacy Attacks in Social Media Using Photo Tagging Networks: A Case Study with Facebook. In *Proceedings of the 1st Workshop on Privacy and Security in Online Social Media (PSOSM '12)*. ACM, New York, NY, USA, Article 4, 8 pages.

Georgios Petkos, Symeon Papadopoulos, and Yiannis Kompatsiaris. 2015. Social Circle Discovery in Ego-Networks by Mining the Latent Structure of User Connections and Profile Attributes. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015 (ASONAM '15)*. ACM, New York, NY, USA, 880–887.

Duangmanee Putthividhya, Hagai Attias, and Srikantan S. Nagarajan. 2010. Topic regression multi-modal Latent Dirichlet Allocation for image annotation. *IEEE Computer Society Conf. on Computer Vision and Pattern Rec.* (2010), 3408–3415.

Z. Qin, C. Li, H. Zhang, and J. Guo. 2015. Improving tag matrix completion for image annotation and retrieval. In *2015 Visual Communications and Image Processing (VCIP)*. 1–4.

Moo-Ryong Ra, Ramesh Govindan, and Antonio Ortega. 2013. P3: Toward Privacy-preserving Photo Sharing. In *Proceedings of the 10th USENIX Conference on Networked Systems Design and Implementation (nsdi'13)*. Berkeley, CA, USA, 515–528.

Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. 2014. ImageNet Large Scale Visual Recognition Challenge.. In *arXiv:1409.0575*.

Boon-Siew Seah, Aixin Sun, and Sourav S. Bhowmick. 2018. Killing Two Birds With One Stone: Concurrent Ranking of Tags and Comments of Social Images. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR 2018, Ann Arbor, MI, USA, July 08-12, 2018*. 937–940.

Paul Seitlinger, Dominik Kowald, Christoph Trattner, and Tobias Ley. 2013. Recommending tags with a model of human categorization *(CIKM '13)*. ACM, New York, NY, USA.

S. Shamma and M. Y. S. Uddin. 2014. Towards privacy-aware photo sharing using mobile phones. In *8th International Conference on Electrical and Computer Engineering*. 449–452.

Yue Shi, Martha Larson, and Alan Hanjalic. 2014. Collaborative Filtering Beyond the User-Item Matrix: A Survey of the State of the Art and Future Challenges. *ACM Comput. Surv.*, Article 3 (May 2014), 45 pages.

Börkur Sigurbjörnsson and Roelof van Zwol. 2008. Flickr Tag Recommendation Based on Collective Knowledge *(WWW '08)*. ACM, New York, NY, USA, 327–336.

Andrew Simpson. 2008. On the Need for User-defined Fine-grained Access Control Policies for Social Networking Applications. In *Proceedings of the Workshop on Security in Opportunistic and SOCial Networks (SOSOC '08)*. ACM, 1:1–1:8.

Xuemeng Song, Xiang Wang, Liqiang Nie, Xiangnan He, Zhumin Chen, and Wei Liu. 2018. A Personal Privacy Preserving Framework: I Let You Know Who Can See What. In *SIGIR*. ACM, 295–304.

Eleftherios Spyromitros-Xioufis, Symeon Papadopoulos, Adrian Popescu, and Yiannis Kompatsiaris. 2016. Personalized Privacy-aware Image Classification. In *ICMR '16*. ACM, New York, NY, USA, 71–78.

Anna Squicciarini, Cornelia Caragea, and Rahul Balakavi. 2014. Analyzing Images' Privacy for the Modern Web *(HT '14)*. ACM, New York, NY, USA, 136–147.

Anna Squicciarini, Cornelia Caragea, and Rahul Balakavi. 2017a. Toward Automated Online Photo Privacy. *ACM Transactions on the Web* 11, 1, Article 2 (April 2017).

Anna Squicciarini, D. Lin, S. Karumanchi, and N. DeSisto. 2012. Automatic social group organization and privacy management. In *8th International Conference on Collaborative Computing: Networking, Applications and Worksharing*. 89–96.

Anna Squicciarini, Dan Lin, Smitha Sundareswaran, and Joshua Wede. 2015. Privacy Policy Inference of User-Uploaded Images on Content Sharing Sites. *IEEE Trans. Knowl. Data Eng.* 27, 1 (2015), 193–206.

Anna Squicciarini, Andrea Novelli, Dan Lin, Cornelia Caragea, and Haoti Zhong. 2017b. From Tag to Protect: A Tag-Driven PolicyRecommender System for Image Sharing. In *PST '17*.

Anna Squicciarini, Mohamed Shehab, and Federica Paci. 2009. Collective Privacy Management in Social Networks. In *Proceedings of the 18th International Conference on World Wide Web (WWW '09)*. ACM, New York, NY, USA, 521–530.

Anna Squicciarini, Smitha Sundareswaran, Dan Lin, and Josh Wede. 2011. A3p: adaptive policy prediction for shared images over popular content sharing sites. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*. 261–270.

Xiaoyuan Su and Taghi M. Khoshgoftaar. 2009. A Survey of Collaborative Filtering Techniques. *Adv. in AI*, Article 4 (2009).

Hari Sundaram, Lexing Xie, Munmun De Choudhury, Yu-Ru Lin, and Apostol Natsev. 2012. Multimedia Semantics: Interactions Between Content and Community. *Proc. IEEE* 100, 9 (2012), 2737–2758.

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2014. Going Deeper with Convolutions. *CoRR* abs/1409.4842 (2014).

Jinhui Tang, Shuicheng Yan, Richang Hong, Guo-Jun Qi, and Tat-Seng Chua. 2009. Inferring Semantic Concepts from Community-contributed Images and Noisy Tags. In *Proceedings of the 17th ACM International Conference on Multimedia (MM '09)*. ACM, New York, NY, USA, 223–232.

F. Tian and X. Shen. 2014. Learning Label Set Relevance for Search Based Image Annotation. In *2014 International Conference on Virtual Reality and Visualization*. 260–265.

G. Toderici, H. Aradhye, M. Pasca, L. Sbaiz, and J. Yagnik. 2010. Finding meaning on YouTube: Tag recommendation and category discovery. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 3447–3454.

Ashwini Tonge and Cornelia Caragea. 2016. Image Privacy Prediction Using Deep Features. In *AAAI*.

Ashwini Tonge and Cornelia Caragea. 2018. On the Use of "Deep" Features for Online Image Sharing. In *Companion Proceedings of The Web Conf.* 1317–1321.

Ashwini Tonge, Cornelia Caragea, and Anna Squicciarini. 2018a. Privacy-Aware Tag Recommendation for Image Sharing. In *Proceedings of the 29th on Hypertext and Social Media (HT '18)*. ACM, New York, NY, USA, 52–56.

Ashwini Tonge, Cornelia Caragea, and Anna Squicciarini. 2018b. Uncovering Scene Context for Predicting Privacy of Online Shared Images. In *AAAI '18*.

Lam Tran, Deguang Kong, Hongxia Jin, and Ji Liu. 2016. Privacy-CNH: A Framework to Detect Photo Privacy with Convolutional Neural Network Using Hierarchical Features. In *AAAI '16*. 7.

Emanuel von Zezschwitz, Sigrid Ebbinghaus, Heinrich Hussmann, and Alexander De Luca. 2016. You Can'T Watch This!: Privacy-Respectful Photo Browsing on Smartphones. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4320–4324.

Nitya Vyas, Anna Squicciarini, Chih-Cheng Chang, and Danfeng Yao. 2009. Towards automatic privacy management in Web 2.0 with semantic analysis on annotations. In *CollaborateCom*. 1–10.

Chong Wang, David M. Blei, and Fei-Fei Li. 2009a. Simultaneous image classification and annotation. In *CVPR 2009*.

H. Wang and J. Hu. 2010. Multi-label image annotation via Maximum Consistency. In *2010 IEEE International Conference on Image Processing*. 2337–2340.

Hua Wang, Heng Huang, and C. Ding. 2009b. Image annotation using multi-label correlated Green's function. In *2009 IEEE 12th International Conference on Computer Vision*. 2029–2034.

H. Wang, H. Huang, and C. Ding. 2011. Image annotation using bi-relational graph of images and semantic labels. In *CVPR*.

J. Wang, Y. Yang, J. Mao, Z. Huang, C. Huang, and W. Xu. 2016. CNN-RNN: A Unified Framework for Multi-label Image Classification. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2285–2294.

Ronggui Wang, Yunfei Xie, Juan Yang, Lixia Xue, Min Hu, and Qingyang Zhang. 2017. Large scale automatic image annotation based on convolutional neural network. *Journal of Visual Communication and Image Representation* (2017).

Jason Watson, Heather Richter Lipford, and Andrew Besmer. 2015. Mapping User Preference to Privacy Default Settings. *ACM Trans. Comput.-Hum. Interact.* 22, 6, Article 32 (Nov. 2015), 20 pages.

Baoyuan Wu, Siwei Lyu, Bao-Gang Hu, and Qiang Ji. 2015a. Multi-label learning with missing labels for image annotation and facial action unit recognition. *Pattern Recognition* 48, 7 (2015), 2279 – 2289.

J. Wu, Yinan Yu, Chang Huang, and Kai Yu. 2015b. Deep multiple instance learning for image classification and auto-annotation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 3460–3469.

Lei Wu, Steven C.H. Hoi, Rong Jin, Jianke Zhu, and Nenghai Yu. 2009. Distance Metric Learning from Uncertain Side Information with Application to Automated Photo Tagging. In *Proceedings of the 17th ACM International Conference on Multimedia (MM '09)*. ACM, New York, NY, USA, 135–144.

Lei Wu, Rong Jin, and Anil K Jain. 2013. Tag Completion for Image Retrieval. *IEEE transactions on pattern analysis and machine intelligence* 35, 3 (March 2013).

Pengcheng Wu, Steven Chu-Hong Hoi, Peilin Zhao, and Ying He. 2011. Mining Social Images with Distance Metric Learning for Automated Image Tagging. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining (WSDM '11)*. ACM, New York, NY, USA, 197–206.

Zhenyu Wu, Zhangyang Wang, Zhaowen Wang, and Hailin Jin. 2018. Towards Privacy-Preserving Visual Recognition via Adversarial Training: A Pilot Study. In *ECCV 2018*, Vol. 11220. Springer, 627–645.

Haitao Xu, Haining Wang, and Angelos Stavrou. 2015. Privacy Risk Assessment on Online Photos.. In *RAID*. 427–447.

Zhichen Xu, Yun Fu, Jianchang Mao, and Difu Su. 2006. Towards the semantic web: Collaborative tag suggestions. In *Collaborative Web Tagging Workshop at 15th Int. WWW Conference*.

Yang Yang, Wensheng Zhang, and Yuan Xie. 2015. Image automatic annotation via multi-view deep representation. *Journal of Visual Communication and Image Representation* 33 (2015), 368 – 377.

Alexei Yavlinsky, Edward Schofield, and Stefan Rüger. 2005. *Automated Image Annotation Using Global Features and Robust Nonparametric Density Estimation*. Springer Berlin Heidelberg.

Ching-man Au Yeung, Lalana Kagal, Nicholas Gibbins, and Nigel Shadbolt. 2009. Providing access control to online photo albums based on tags and linked data. *Social Semantic Web: Where Web* 2 (2009).

F. Yu and H. H. S Ip. 2006. Automatic Semantic Annotation of Images using Spatial Hidden Markov Model. In *2006 IEEE International Conference on Multimedia and Expo*. 305–308.

Jun Yu, Zhenzhong Kuang, Zhou Yu, Dan Lin, and Jianping Fan. 2017. Privacy Setting Recommendation for Image Sharing. In *16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*. 726–730.

Jun Yu, Zhenzhong Kuang, Baopeng Zhang, Wei Zhang, Dan Lin, and Jianping Fan. 2018. Leveraging Content Sensitiveness and User Trustworthiness to Recommend Fine-Grained Privacy Settings for Social Image Sharing. *IEEE Trans. Information Forensics and Security* 13, 5 (2018), 1317–1332.

Lin Yuan, Joel Regis Theytaz, and Touradj Ebrahimi. 2017. Context-Dependent Privacy-Aware Photo Sharing based on Machine Learning. *Proc. of 32nd International Conference on ICT Systems Security and Privacy Protection (IFIP SEC)* (2017).

X. Yuan, X. Wang, C. Wang, Anna Squicciarini, and K. Ren. 2018. Towards Privacy-Preserving and Practical Image-Centric Social Discovery. *IEEE Transactions on Dependable and Secure Computing* 15, 5 (Sept 2018), 868–882.

Sergej Zerr, Stefan Siersdorfer, Jonathon Hare, and Elena Demidova. 2012. Privacy-aware image classification and search. In *ACM SIGIR*. ACM, NY, USA.

Zheng-Jun Zha, Tao Mei, Jingdong Wang, Zengfu Wang, and Xian-Sheng Hua. 2009. Graph-based semi-supervised learning with multiple labels. *Journal of Visual Communication and Image Representation* 20, 2 (2009), 97 – 103.

Qi Zhang, Jiawen Wang, Haoran Huang, Xuanjing Huang, and Yeyun Gong. 2017. Hashtag Recommendation for Multimodal Microblog Using Co-Attention Network. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*. 3420–3426.

Wei Zhang, S. S. Cheung, and Minghua Chen. 2005. Hiding privacy information in video surveillance system. In *IEEE International Conference on Image Processing 2005*, Vol. 3. II–868.

Yufeng Zhao, Yao Zhao, and Zhenfeng Zhu. 2009. TSVM-HMM: Transductive SVM based hidden Markov model for automatic image annotation. *Expert Systems with Applications* 36, 6 (2009), 9813 – 9818.

Haoti Zhong, Anna Squicciarini, and David Miller. 2018. Toward Automated Multiparty Privacy Conflict Detection. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM '18)*. 1811–1814.

Haoti Zhong, Anna Squicciarini, David Miller, and Cornelia Caragea. 2017. A Group-Based Personalized Model for Image Privacy Classification and Labeling. In *Proceedings of the 26th International Joint Conf. on Artificial Intelligence*. 3952–3958.

Bolei Zhou, Aditya Khosla, Agata Lapedriza, Antonio Torralba, and Aude Oliva. 2016. Places: An Image Database for Deep Scene Understanding. *arXiv preprint arXiv:1610.02055* (2016).