

Observing the natural world with Flickr

Jingya Wang

Mohammed Korayem
School of Informatics and Computing
Indiana University
Bloomington, Indiana USA

David J. Crandall

{wang203,mkorayem,djcran}@indiana.edu

Abstract

The billions of public photos on online social media sites contain a vast amount of latent visual information about the world. In this paper, we study the feasibility of observing the state of the natural world by recognizing specific types of scenes and objects in large-scale social image collections. More specifically, we study whether we can recreate satellite maps of snowfall by automatically recognizing snowy scenes in geo-tagged, timestamped images from Flickr. Snow recognition turns out to be a surprisingly difficult and under-studied problem, so we test a variety of modern scene recognition techniques on this problem and introduce a large-scale, realistic dataset of images with ground truth annotations. As an additional proof-of-concept, we test the ability of recognition algorithms to detect a particular species of flower, the California Poppy, which could be used to give biologists a new source of data on its geospatial distribution over time.

1. Introduction

Digital cameras and camera-enabled smartphones are now ubiquitous, with a large fraction of the population taking photos regularly and sharing them online. These millions of people taking pictures form a massive social sensor network that is (in aggregate) observing and capturing the visual world across time and space. Modern phones and cameras record metadata like geo-tags and time-stamps in addition to the images themselves, giving (noisy) calibration information about how this ad-hoc sensor network is arranged. Social media sites like Flickr and Facebook thus contain a large amount of latent visual information about the world and how it is changing over time.

For instance, many (if not most) outdoor images contain some information about the state of the natural world, such as the weather conditions and the presence or absence of plants and animals (Figure 1). The billions of images on social media sites could be analyzed to recognize these nat-



Figure 1. Many Flickr images contain evidence about the state of the natural world, including that there is snow on the ground at a particular place and time, that a particular species of bird or animal is present, and that particular species of plants are flowering.

ural objects and phenomena, creating a new source of data to biologists and ecologists. Where are marigolds blooming today, and how is this geospatial distribution different from a year ago? Are honeybees less populous this year than last year? Which day do leaves reach their peak color in each county of the northeastern U.S.? These questions can be addressed to some extent by traditional data collection techniques like satellite instruments, aerial surveys, or longitudinal manual surveys of small patches of land, but none of these techniques allows scientists to collect fine-grained data at continental scales: satellites can monitor huge areas of land but cannot detect fine-grained features like blooming flowers, while manual surveys can collect high-quality and fine-grained data only in a small plot of land. Large-scale analysis of photos on social media sites could provide an entirely new source of data at a fraction of the cost of launching a satellite or hiring teams of biologist observers.

The idea of using crowd-sourced data for science and other purposes is of course not new. Citizen science projects have trained groups of volunteers to recognize and report

natural phenomena (like bee counts [1], bird sightings [9], and snowfall [17]) near their homes. Data mining work has shown that social networking sites like Twitter can monitor political opinions [7, 15], predict financial markets [4], track the spread of disease [10], detect earthquakes [28], and monitor weather conditions [14]. However, the vast majority of this work has used textual data from micro-blogging sites like Twitter; very few papers have tried to do this with images, despite the fact that images offer evidence that is richer, less ambiguous, and much more difficult to fabricate. This is of course because it is much easier to scan for keywords in Twitter feeds than to automatically recognize semantic content in huge collections of images.

In this paper, we test the feasibility of observing the natural world by recognizing specific types of scenes and objects in large-scale image collections from social media. We consider a well-defined but nevertheless interesting problem: deciding whether there was snow on the ground at a particular place and on a particular day, given the set of publicly-available Flickr photos that were geo-tagged and time-stamped at that place and time. This builds on our earlier work in Zhang *et al* [34] which considered a similar problem, but used only tag information (essentially scanning for photos that had the tag “snow” with some very simple image processing to remove obvious outliers). Here, we explicitly test whether large-scale recognition of image content itself could be used to do this task. Of course, snow cover can already be monitored through satellites and weather stations (although neither of these sources is perfect: weather stations are sparse in rural areas and satellites typically cannot observe snow cover through clouds [12]), so this is not a transformative application in and of itself. Instead, this application is interesting precisely because fine-grained ground truth is available, so that accuracy of crowd-sourced observations of the natural world can be tested at large scale, potentially paving the way to observe other natural phenomena for which there are no other sources of data.

We initially expected snow detection to be an easy problem, in which just looking for large white regions would work reasonably well. However, amongst the hundreds of papers on object and scene classification in the literature, we were surprised to find very few that have explicitly considered detecting snow. A few papers on scene classification include snow-related categories [20, 21, 33], while a few older papers on natural materials detection [6, 22] consider it along with other categories. We test a variety of recognition techniques on this problem, using a new realistic dataset of several thousand images from Flickr with labeled ground truth. We find that snow detection in consumer images is surprisingly difficult, and we hope this paper and our dataset will help spark interest in this somewhat overlooked vision problem.

Finally we present some preliminary results of applying

our recognition techniques on geo-tagged and time-stamped images from Flickr, in order to estimate the geo-spatial distribution of snow. We compare these maps to ground-truth from satellite data. We also consider an ecology application where reliable data does not exist and Flickr image analysis could be potentially quite valuable: estimating the geo-temporal flowering distribution of the California Poppy.

2. Related work

Crowd-sourcing from social media. Several recent studies have shown the power of social media for observing the world itself, as a special case of ‘social sensing’ [2]. This work includes using Twitter data to measure collective emotional state [11] (which, in turn, has found to be predictive of stock moves [4]), predicting product adoption rates and political election outcomes [15], and collecting data about earthquakes and other natural disasters [28]. Particularly striking examples include Ginsberg *et al* [10], who show that geo-temporal properties of web search queries can predict the spread of flu, and Sadilek *et al* [27] who show that Twitter feeds can predict when a given person will fall ill.

The specific application we consider here is inferring information about the state of the natural world from social media. Existing work has analyzed textual content, including text tags and Twitter feeds, in order to do this. Hyvarinen and Saltikoff [14] search for images on Flickr to validate meteorological satellite observations, albeit by hand. Zhang *et al* [34] take a large collection of geo-tagged and time-stamped Flickr photos and search for snow-related tags to produce estimates of geo-temporal snowfall distributions, and evaluate them against satellite snow maps. Singh *et al* [29] visualize geospatial distributions of photos tagged “snow” as an example of their Social Pixels framework, but they study the database theory needed to perform this analysis and do not consider the prediction problem.

Few papers have used actual image content analysis as we do here. Leung and Newsam [19] use scene analysis in geo-tagged photos to infer land cover and land use types. Murdock *et al* [23] analyze geo-referenced stationary webcam feeds to estimate cloud cover on a day-by-day basis, and then use these estimates to recreate satellite cloud cover maps. Webcams offer a complimentary data source to the social media images we consider here: on one hand, analyzing webcam data is made easier by the fact that the camera is stationary and offers dense temporal resolution; on the other hand, their observations are restricted to where public webcams exist, whereas photos on social media sites offer a potentially much denser spatial sampling of the world.

We note that these applications are related to citizen science projects where volunteers across a wide geographic area send in observations [1, 9, 17]. These projects often use social media, but require observations to be made explicitly, whereas in our work we “passively” analyze social media

feeds generated by untrained and unwitting individuals.

Detecting snow in images. We know of only a handful of papers that have explicitly considered snow detection in images. Perhaps the most relevant is the 2003 work of Singhal *et al* [22, 30] which studies this in the context of detecting “materials” like water, grass, sky, etc. They calculate local color and texture features at each pixel, and then compute a probability distribution over the materials at each pixel using a neural network. They partition the image into segments by thresholding these belief values, and assign a label to each segment with a probabilistic framework that considers both the beliefs and simple contextual information like relative location. They find that sky and grass are relatively easy to classify, while snow and water are most difficult. Follow-up work [5, 6] applied more modern techniques like support vector machines. Barnum *et al* [3] detect falling snow and rain, a complementary problem to the one we study here of detecting fallen snow.

Papers in the scene recognition literature have considered snowy scenes amongst their scene categories; for instance, Li *et al* [20, 21] mention snow as one possible component of their scene parsing framework, but do not present experimental results. The SUN database of Xiao *et al* [33] includes several snow-related classes like “snow-field,” “ski slope,” “ice shelf,” and “mountain snowy,” but other categories like “residential neighborhood” sometimes have snow and sometimes do not, such that detecting these scenes alone is not sufficient for our purposes.

Recognizing flowers. There are a number of papers on detecting and recognizing flowers in images, although none have specifically considered the California Poppies we study here. Most work on flower classification uses datasets with close-up images of nearly-centered flowers, not the cluttered images typical of Flickr. We use the work of Nilsback and Zisserman [24] as the starting point for our experiments. They perform a binary segmentation step to separate flower from background, represent the foreground with vocabularies of color, shape, and texture features, and then perform recognition using nearest neighbors. Later work [25] uses additional features like SIFT and HOG, combines them using a multiple kernel framework, and extended the dataset to 103 classes. Other work considering flower recognition includes the semi-automated method of Zou and Nagy [35] and the work of Kanan and Cottrell [16], which uses flower recognition as an application of a recognition technique based on modeling visual attention.

3. Snow detection

As noted above, we are aware of very little work that has considered the problem of detecting snow in images: the most relevant work [30] considers snow in the context of natural materials classification, but is over 10 years old,

uses a small and biased dataset, and does not report classification results. Recent work on scene understanding [33] sometimes includes snow-related scenes, but none of this work applies directly to our problem because snow can appear across a range of different scene types. Snow is really an object, not a type of scene, but we are not aware of any work on recognizing snow in the object detection literature.

We thus begin by assembling a large-scale realistic image dataset, and test a variety of modern classification techniques on the problem of snowy scene detection. We use a labeled subset of this dataset to train classifiers and to test their performance, and then apply these classifiers to the problem of generating satellite-like snowfall maps using image analysis on geo-tagged, time-stamped Flickr photos.

3.1. Dataset

We collected a large realistic dataset of Flickr images. A subtle but important issue is how to sample these photos. The distribution of geo-tagged Flickr photos is highly non-uniform, with high peaks in population centers and tourist locations. Sampling uniformly at random from Flickr photos produces a dataset that mirrors this highly non-uniform distribution, biasing it towards cities and away from rural areas. Since our eventual goal is to reproduce continental-scale satellite maps, rural areas are very important. An alternative is biased sampling that attempts to select more uniformly over the globe, but has the disadvantage that it no longer reflects the distribution of Flickr photos. Other important considerations include how to find a variety of snowy and non-snowy images, including relatively difficult images that may include wintry scenes with ice but not snow, and how to prevent highly-active Flickr users from disproportionately affecting the datasets.

We strike a compromise on these issues by combining together datasets sampled in different ways. We begin with a collection of about 100 million Flickr photos geo-tagged within North America and collected using the public API (by repeatedly querying at different times and geo-spatial areas, similar to [13]). From this set, we considered only photos taken before January 1, 2009 (so that we could use later years for creating a separate test set), and selected: (1) all photos tagged *snow*, *snowfall*, *snowstorm*, or *snowy* in English and 10 other common languages (about 500,000 images); (2) all photos tagged *winter* in English and about 10 other languages (about 500,000 images); (3) a random sample of 500,000 images. This yielded about 1.4 million images after removing duplicates. We further sampled from this set in two ways. First, we selected up to 20 random photos from each user, or all photos if a user had less than 20 photos, giving about 258,000 images. Second, we sampled up to 100 random photos from each $0.1^\circ \times 0.1^\circ$ latitude-longitude bin of the earth (roughly $10\text{km} \times 10\text{km}$ at the mid latitudes), yielding about 300,000 images. The combination

of these two datasets has about 425,000 images after removing duplicates, creating a diverse and realistic selection of images. We partitioned this dataset into test and training sets on a per-user basis, so that all of any given user’s photos are in one set or the other (to reduce the potential for duplicate images appearing in both training and test).

We then presented a subset of these images to humans and collected annotations for each image. We asked people to label the images into one of four categories: (1) contains obvious snow near the camera; (2) contains a trace amount of snow near the camera; (3) contains obvious snow but far away from the camera (e.g. on a mountain peak); and (4) does not contain snow. For our application of reconstructing snowfall maps, we consider (1) and (2) to be positive classes and (3) and (4) to be negative, since snowfall in the distance does not give evidence of snow at the image’s geo-tagged location. In total we labeled 10,000 images.

3.2. Snow classification

Snow is a somewhat unique visual phenomenon, and we claim that detecting it in images is a unique recognition task. In some cases, snow can be detected by coarse scene recognition: ski slopes or snowy landscapes are distinctive scenes. But snow can appear in any kind of outdoor scene, and is thus like an object. However, unlike most objects that have some distinctive features, snow is simply a white, near-textureless material. (In fact, our informal observation is that humans detect snow not by recognizing its appearance, but by noticing that other expected features of a scene are occluded; in this sense, detecting snow is less about the features that are seen and more about the features that are *not* seen. We leave this as an observation to inspire future work.) We tested a variety of off-the-shelf visual features for classifying whether an image contains fallen snow. We used Support Vector Machines for classification, choosing kernels based on the feature type. Intuitively, color is a very important feature for detecting snow, and thus we focused on features that use color to some degree. Our features are:

Color histograms. We begin with perhaps the simplest of color features. We build joint histograms in CIELAB space, with 4 bins on the lightness dimension and 14 bins along each of the two color dimensions, for a total of 784 bins. We experimented with other quantizations and found that this arrangement worked best. We encode the histogram as a 784 dimensional feature and use an SVM with a chi-squared distance (as in [33]).

Tiny images. We subsample images to 16×16 pixels, giving 256 pixels per RGB color plane and yielding a 768 dimensional feature vector. Drastically reducing the image dimensions yields a feature that is less sensitive to exact alignment and more computationally feasible [31].

Spatial Moments. Tiny images capture coarse color and

spatial scene layout information, but much information is discarded during subsampling. As an alternative approach, we convert the image to LUV color space, divide it into 49 blocks using a 7×7 grid, and then compute the mean and variance of each block in each color channel. Intuitively, this is a low-resolution image and a very simple texture feature, respectively. We also compute maximum, minimum, and median value within each cell, so that the final feature vector has 735 dimensions.

Color Local Binary Pattern (LBP) with pyramid pooling. LBP represents each 9×9 pixel neighborhood as an 8-bit binary number by thresholding the 8 outer pixels by the value at the center. We build 256-bin histograms over these LBP values, both on the grayscale image and on each RGB color channel [18]. We compute these histograms in each cell of a three-level spatial pyramid, with 1 bin at the lowest level, 4 bins in a 2×2 grid at the second level, and 16 bins in a 4×4 grid at the third level. This yields a $(1 + 4 + 16) \times 4 \times 256 = 21504$ dimensional feature vector for each image.

GIST. We also apply GIST features, which capture coarse texture and scene layout by applying a Gabor filter bank followed by down-sampling [26]. Our variant produces a 1536-dimensional vector and operates on color planes. Scaling images to have square aspect ratios before computing GIST improved classification results significantly [8].

We experimented with a number of other features, and found that they did not work well; local features like SIFT and HOG in particular perform poorly, again because snow does not have distinctive local visual appearance.

3.3. Results

We tested these approaches to detecting snow on our dataset of 10,000 hand-labeled images. We split this set into a training set of 8,000 images and a test set of 2,000 images, sampled to have an equal proportion of snow and non-snow images (so that the accuracy of a random baseline is 50%). Table 1 presents the results. We observe that all of the features perform significantly better than a random baseline. Gist, Color Histograms and Tiny Image all give very similar accuracies, within a half percentage point of 74%. Spatial Moments and LBP features perform slightly better at 76.2%

Feature	Kernel	Accuracy
Random Baseline	—	50.0%
Gist	RBF	73.7%
Color	χ^2	74.1%
Tiny	RBF	74.3%
Spatial Color Moments	RBF	76.2%
Spatial pyramid LBP	RBF	77.0%
All features	linear	80.5%

Table 1. Performance of different features for snow detection, all using SVMs for classification.

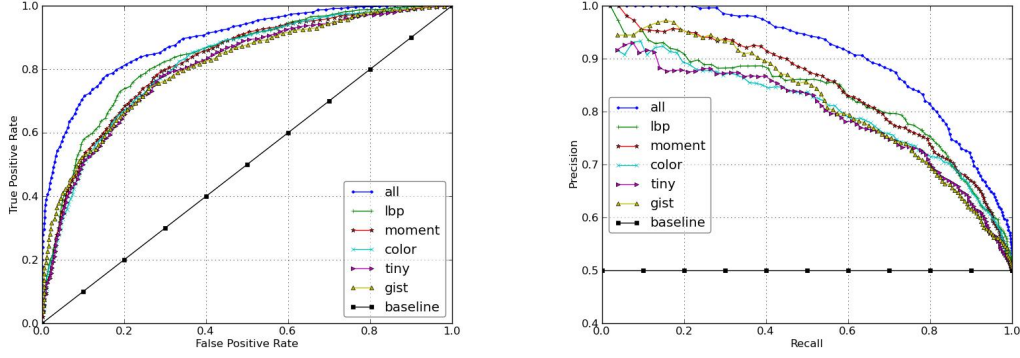


Figure 2. Snow classification results for different features and combinations, in terms of (*left*): ROC curves for the task of classifying snow vs. non-snow images; and (*right*): Precision-Recall curves for the task of retrieving snow images.

and 77.0%. We also tested a combination of all features by learning a second-level linear SVM on the output of the five SVMs; this combination performed significantly better than any single feature, at 80.5%.

Figure 2 shows classification performance in terms of an ROC curve, as well as a precision-recall curve in which the task is to retrieval photos containing snow. The precision-recall curve shows that at about 20% recall, precision is very near to 100%, while even at 50% recall, precision is close to 90%. This is a nice feature because in many applications, it may not be necessarily to correct classify all images, but instead to find some images that most likely contain a subject of interest. To give a sense for the difficulty and failure modes of our dataset, we show a random sample of correct and incorrect classification results in Figure 4.

Reconstructing satellite snow maps. Finally, we tested whether this automated photo classification run on large-scale collections of geo-tagged, time-stamped social images could be used to approximate snow maps generated by satellites. An advantage of considering the snow recognition task is that ground truth, in the form of daily snow cover maps, is publicly available from NASA and others [34]. This is thus a somewhat artificial task because very good datasets already exist for snow cover, but we use this problem here as a test case of the more general idea of using Flickr to observe nature. (Nevertheless, satellites are also limited because they require the ground to be visible, and thus are not effective when there is cloud cover.)

To test this idea, we downloaded public, geo-tagged, time-stamped Flickr photos taken in North America on three days: March 3, April 6, and December 21 2009 (4422, 5606, and 9906 photos respectively). We ran our combined classifiers on these images. We discretized the image geo-tags into 1 degree by 1 degree bins, and interpreted each snowy image as evidence of snow in that bin and each non-snowy image as evidence against snow in that bin. We combined this evidence together using the simple Bayesian approach proposed by [34]. Figure 3 shows the resulting map

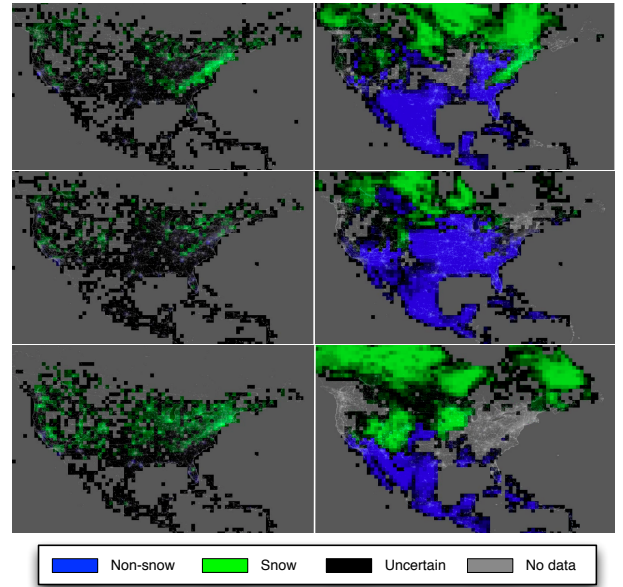


Figure 3. Snow cover maps generated by our Flickr analysis (left), compared with satellite maps (right), on three days in 2009: March 3 (top), April 6 (middle), and December 21 (bottom). Green indicates snow, blue indicates no snow, and gray indicates uncertainty (caused by too few photos in Flickr analysis, or by cloud cover in satellite maps).

produced by our automated Flickr analysis, and compares it to the corresponding snow cover map produced by NASA’s MODIS instrument [12]. We note that the Flickr map is much sparser than the satellite map, especially in sparsely populated areas like northern Canada and the western U.S. On the other hand, the Flickr maps give some observations even when the satellite maps are missing data due to clouds.

4. Detecting California Poppies

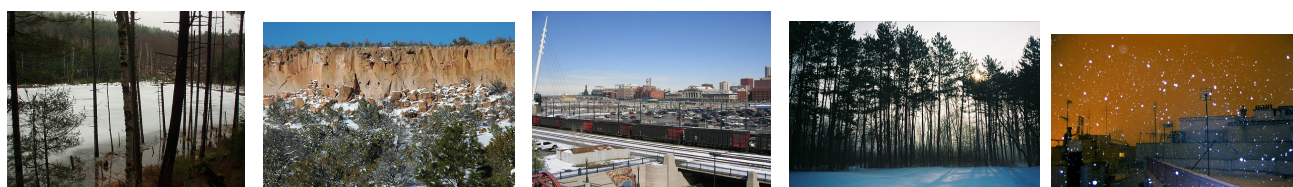
We have also studied whether we can apply computer vision analysis of Flickr photos to a problem of interest to biologists: tracking the geo-temporal distribution of flow-



(a) Random true negatives (non-snow images classified as non-snow)



(b) Random true positives (snow images classified as snow)



(c) Random false negatives (snow images classified as non-snow)



(d) Random false positives (non-snow images classified as snow)

Figure 4. Snow classification results on some random images from our dataset, including (from top) true negatives, true positives, false negatives, and false positives. These results were obtained using the combined classifier that uses all of the image features.

ering plants. Plants and animals will respond as climates change over time, and biologists would like fine-grained, continental scale information about how flowering and migratory patterns are changing. Unlike weather conditions, this data is very difficult to monitor from satellites or aircraft, so biologists currently rely mostly on traditional data collection techniques like longitudinal studies of small plots of land by expert biologists. Analyzing Flickr images could provide an alternative data source for these studies.

As a step in this direction, here we consider one particular class of flower: the California Poppy. We chose this flower both because of its distinct visual appearance (a bright orange) and because it is of interest to biologists because it grows in a relatively small area of the western U.S. and thus may be particularly sensitive to changes in climate.

4.1. Dataset

From our collection of about 100 million U.S. Flickr photos, we selected all images tagged “poppy” (about 8100

images). Some of these images are of California Poppies but most are not, since there are other species of poppies and amateur photographers often confuse them with other flowers. We took a random sample of about 2000 images and asked biology students to label them into one of four categories: (1) close-up of a California Poppy; (2) multiple California Poppies (e.g. in a photo of a field perhaps amongst other flowers); (3) no California Poppy; and (4) special cases like drawings of poppies. We discarded images from category (4) and sampled from the remaining dataset to have an equal proportion of the three classes. This gave 150 training images and 450 independent test images. Figure 5 shows a few sample images from our dataset.

4.2. Classifying poppies

We used the same features described above in Section 3.2 for classifying snow images, including tiny images, color histograms, color-aware local binary pattern with spatial pyramids, and GIST. For comparison, we also implemented

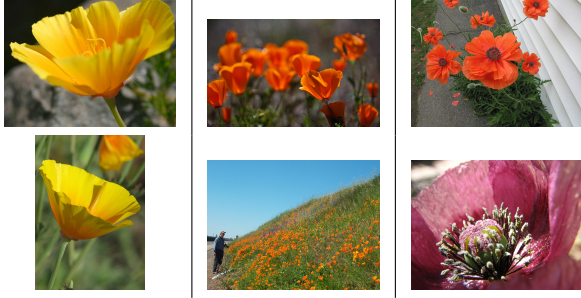


Figure 5. Some images from the three classes in our California Poppy dataset: (left): close-ups of true poppies; (center): longer-range images of true poppies; and (right): images with no poppies.

techniques based on bag-of-words vocabularies of color, texture, and shape features that have been applied to flower recognition in past work [24]. In particular:

Color vocabulary. We clustered the HSV color values from all images into a 200-word vocabulary, and then represented each image as a histogram over these visual words.

Shape vocabulary. We use SIFT features to represent local image “shape.” We extracted SIFT features densely (on 25×25 pixel regions, at strides of 20 pixels), and again built a vocabulary using k -means clustering with $k = 200$.

Texture vocabulary. We used MR8 features [32] to capture local texture information. MR8 applies a filter bank of 8 filters (4 Gaussians and 4 Laplacians of Gaussians) at different scales and orientations, and then characterizes local texture in terms of the maximal filter responses. We again cluster these into a vocabulary with size 200.

We also define a combined feature which incorporates all three of the above features. This feature computes the histogram for each of the three filters, and then concatenates these together after normalizing each vector.

4.3. Results

Table 2 shows the performance of the different features on the problem of classifying close-up California Poppy photos versus photos of fields and non-poppies. We observe that the vocabulary-based features work significantly better in combination than separately, yielding a combined accuracy of 65.0% versus the 33.3% baseline. The LBP and Gist features perform better, with the best performance achieved by the combination of features (72.1%). Figure 6 shows ROC and Precision-Recall curves.

5. Conclusion

In this paper, we propose using photo-sharing social media sites as a means of observing the state of the natural world, by automatically recognizing specific types of scenes and objects in large-scale social image collections. This work is an initial step towards a long-term goal of monitoring important ecological events and trends through online

Feature	Accuracy
Random Baseline	33.3%
Shape Vocabulary	45.0%
Texture Vocabulary	48.6%
Color Vocabulary	53.6%
Combination of color, shape and texture	65.0%
Tiny Image	58.8%
RGB histogram	61.3%
LBP	68.4%
GIST	68.8%
Spatial pyramid LBP	70.4%
Combined	72.1%

Table 2. Results for California Poppy classification.

social media. Our study shows that snowy scene recognition is not nearly as easy a problem as one might expect, when applied to realistic consumer images; our best result using modern vision techniques gives 81% accuracy. Nevertheless, as a proof-of-concept we demonstrated that this recognition accuracy still yields a reasonable map that approximates observations from satellites. We also test recognition algorithms on their ability to recognize a particular species of flower, the California Poppy. In future work, we plan to combine evidence from tags and other metadata with visual features for more accurate estimates, and to develop novel techniques for these challenging recognition problems. More generally, we hope the idea of observing nature through photo-sharing websites will help spark renewed interest in recognizing natural and ecological phenomenon in consumer images.

6. Acknowledgements

This work was supported in part by the National Science Foundation (IIS-1253549), the Lilly Endowment and by the Data to Insight Center at Indiana University. We also thank Prof. Gretchen LeBuhn and biology students at San Francisco State University for labeling the poppy images.

References

- [1] <http://www.greatsunflower.org>. 2
- [2] C. Aggarwal and T. Abdelzaher. Social Sensing. In *Managing and Mining Sensor Data*. Springer, 2013. 2
- [3] P. Barnum, S. Narasimhan, and T. Kanade. Analysis of rain and snow in frequency space. *IJCV*, 86(2-3), 2009. 3
- [4] J. Bollen, H. Mao, and X.-J. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8, 2011. 2
- [5] M. Boutell. *Exploiting context for semantic scene classification*. PhD thesis, University of Rochester Department of Computer Science, 2005. 3
- [6] M. Boutell, A. Choudhury, J. Luo, and C. M. Brown. Using semantic features for scene classification: How good do they need to be? In *ICME*, 2006. 2, 3
- [7] J. DiGrazia, K. McKelvey, J. Bollen, and F. Rojas. More tweets, more votes: Social media as a quantitative indicator of political behavior. *PLOS One*, 2013. 2

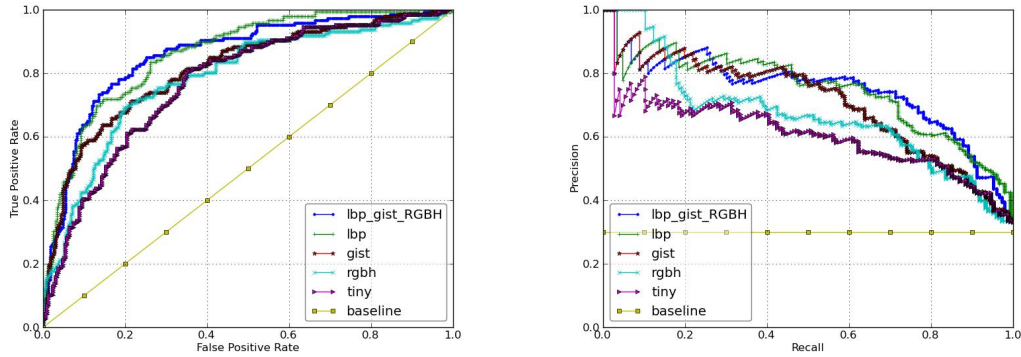


Figure 6. California Poppy classification results for different features, where the goal is to find close-up pictures of California Poppies, in terms of (left): ROC curves of classification performance and (right): Precision-Recall curves showing retrieval performance.

- [8] M. Douze, H. Jegou, H. Sandhawalia, L. Amsaleg, and C. Schmid. Evaluation of gist descriptors for web-scale image search. In *ICIVR*, 2009. 4
- [9] D. Fink, T. Damoulas, and J. Dave. Adaptive spatio-temporal exploratory models: Hemisphere-wide species distributions from massively crowdsourced eBird data. In *AAAI*, 2013. 2
- [10] J. Ginsberg, M. Mohebbi, R. Patel, L. Brammer, M. Smolinski, and L. Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457:1012–1014, 2009. 2
- [11] S. A. Golder and M. Macy. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science*, 333(6051):1878–1881, Sept. 2011. 2
- [12] D. Hall, G. Riggs, and V. Salomonson. MODIS/Terra Snow Cover Daily L3 Global 0.05Deg CMG V004. National Snow and Ice Data Center, updated daily. 2, 5
- [13] J. Hays and A. A. Efros. IM2GPS: Estimating geographic information from a single image. In *CVPR*, 2008. 3
- [14] O. Hyvarinen and E. Saltikoff. Social media as a source of meteorological observations. *Monthly Weather Review*, 138(8):3175–3184, 2010. 2
- [15] X. Jin, A. Gallagher, L. Cao, J. Luo, and J. Han. The wisdom of social multimedia: Using Flickr for prediction and forecast. In *ACM Multimedia*, 2010. 2
- [16] C. Kanan and G. Cottrell. Robust classification of objects, faces, and flowers using natural image statistics. In *CVPR*, 2010. 3
- [17] J. King, A. Cabrera, and R. Kelly. The Snowtweets Project: Communicating snow depth measurements from specialists and non-specialists via mobile communication technologies and social networks. In *AGU Fall Meeting*, 2009. 2
- [18] M. Korayem, A. Mohamed, D. Crandall, and R. Yampolskiy. Solving avatar captchas automatically. In *AMLTA*, 2012. 4
- [19] D. Leung and S. Newsam. Proximate Sensing: Inferring What-Is-Where From Georeferenced Photo Collections. In *CVPR*, 2010. 2
- [20] L.-J. Li and L. Fei-Fei. What, where and who? classifying events by scene and object recognition. In *ICCV*, 2007. 2, 3
- [21] L.-J. Li, R. Socher, and L. Fei-Fei. Towards total scene understanding: classification, annotation and segmentation in an automatic framework. In *CVPR*, 2009. 2, 3
- [22] J. Luo, A. Singhal, and W. Zhu. Natural object detection in outdoor scenes based on probabilistic spatial context models. In *ICME*, 2003. 2, 3
- [23] C. Murdock, N. Jacobs, and R. Pless. Webcam2Satellite: Estimating cloud maps from webcam imagery. In *WACV*, 2013. 2
- [24] M.-E. Nilsback and A. Zisserman. A visual vocabulary for flower classification. In *CVPR*, 2006. 3, 7
- [25] M.-E. Nilsback and A. Zisserman. Automated flower classification over a large number of classes. In *ICVGIP*, 2008. 3
- [26] A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *IJCV*, 42(3):145–175, 2001. 4
- [27] A. Sadilek, H. Kautz, and V. Silenzio. Predicting Disease Transmission from Geo-Tagged Micro-Blog Data. In *AAAI*, 2012. 2
- [28] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes Twitter users: real-time event detection by social sensors. In *WWW*, 2010. 2
- [29] V. K. Singh, M. Gao, and R. Jain. Social pixels: genesis and evaluation. In *ACM Multimedia*, 2010. 2
- [30] A. Singhal, J. Luo, and W. Zhu. Probabilistic spatial context models for scene content understanding. In *CVPR*, 2003. 3
- [31] A. Torralba, R. Fergus, and W. T. Freeman. 80 million tiny images: A large data set for nonparametric object and scene recognition. *PAMI*, 30(11):1958–1970, 2008. 4
- [32] M. Varma and A. Zisserman. Classifying images of materials: Achieving viewpoint and illumination independence. In *ECCV*, pages 255–271. Springer, 2002. 7
- [33] J. Xiao, J. Hays, K. A. Ehinger, A. Oliva, and A. Torralba. SUN database: Large-scale scene recognition from abbey to zoo. In *CVPR*, 2010. 2, 3, 4
- [34] H. Zhang, M. Korayem, D. Crandall, and G. LeBuhn. Mining Photo-sharing Websites to Study Ecological Phenomena. In *WWW*, 2012. 2, 5
- [35] J. Zou and G. Nagy. Evaluation of model-based interactive flower recognition. In *ICPR*, 2004. 3