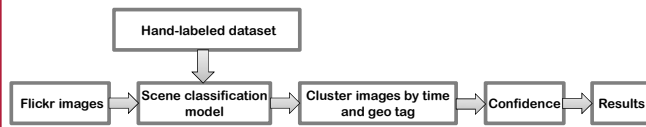


# Observing ecology phenomenon on Flickr

Stefan Lee, David Crandall

School of Informatics and Computing, Indiana University, Bloomington, IN

## 1. Overview



## 2. Method

- Confidence score** is measuring the ratio of log likelihood of being positive or not at each time and location.

$$P(scene | s^+, \vec{s}^+) = \frac{P(s^+, \vec{s}^+ | scene) P(scene)}{P(s^+, \vec{s}^+)}$$

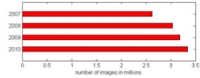
$$= \frac{\binom{m+n}{n} p^n (1-p)^m P(scene)}{P(s^+, \vec{s}^+)}$$

$$\frac{P(scene | s^+, \vec{s}^+)}{P(scene | s^-, \vec{s}^-)} = \frac{P(scene)}{P(scene)} \left( \frac{p}{q} \right)^n \left( \frac{1-p}{1-q} \right)^m$$

At a place and time of interest,  
 $m$ : the number of photos contains evidence of scene (event  $s$ )  
 $n$ : the number of photos without evidence of scene (event  $s^-$ )  
 $p$ :  $P(s | scene)$   
 $q$ :  $P(s^- | scene)$   
 $P(scene)$ : the prior probability of certain scene.

## 3. Dataset

- Very large dataset from Flickr



- Diversity of public sharing images



Random positive images in vegetation dataset



Random negative images in vegetation dataset

- Hand-labeled dataset

We build 2 hand labeled datasets with equal number of positive and negative images. **Vegetation** dataset has over 10000 images. They are taken before 2009, and are composed by images with "forest" and "summer" like tags and also random images without any tag preference. We build a positive set with images in category "Outdoor Greenery" and a negative set with images in categories "Outdoor non-Greenery" and "Indoor". **Snow** dataset also has about 10000 images. These images are tagged with "snow" and "winter" like words. They are labeled in 4 categories: (1) obvious snow; (2) contains a trace amount of snow near the camera; (3) obvious snow but far away; and (4) does not contain snow. We consider (1) and (2) to be positive classes and (3) and (4) to be negative.

## 4. Scene Classification

- Combining image features
- Snow: GIST, color histogram, tiny image, spatial color moments, spatial pyramid LBP
- Vegetation: color SIFT, color GIST
- Convolutional Neural Network

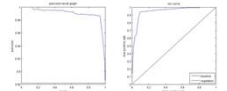
	Image features	CNN
Snow scene	80.5%	~85%
Vegetation scene	85.9%	~88%

- Sample image of snow Scene classification

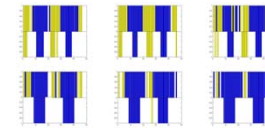


## Vegetation coverage

- Numbers



- Examples of single places over 2009-2010



Blue: place covered by vegetation  
 Yellow: place without vegetation  
 Top: prediction  
 Bottom: ground truth

- North America maps in 2010



## 5. Experiments

### Snow coverage

## 6. Conclusion and future work