Predicting Viewer's Sentimental Response to Images in Social Media

Yibei Huang[1], Yijing Feng[1]

Graphen, Inc[1]

yibeih@graphen.ai, yijing@graphen.ai

Abstract

We propose a framework to analyze visual sentiment of images in social media, and predict viewer's affective response. Given an image, we predict the potential viewer affect concepts (VACs) evoked by publisher affect concepts (PACs) based on the mined publisher-viewer correlation. Furthermore, according to the suggested VACs, plausible comments for the image are chosen from a pool of relevant comments. Instead of metadata, PACs are extracted from images using DeepSentiBank, a deep convolutional neural network based detector, and further augmented with the result from object detection.

Introduction

Images are important medium for interactions in social media due to their effectiveness in information exchange and emotion expression. Massive research has been conducted in visual sentiment analysis[1], yet, the prior methods seldom look into the interactions between the image publishers and viewers.

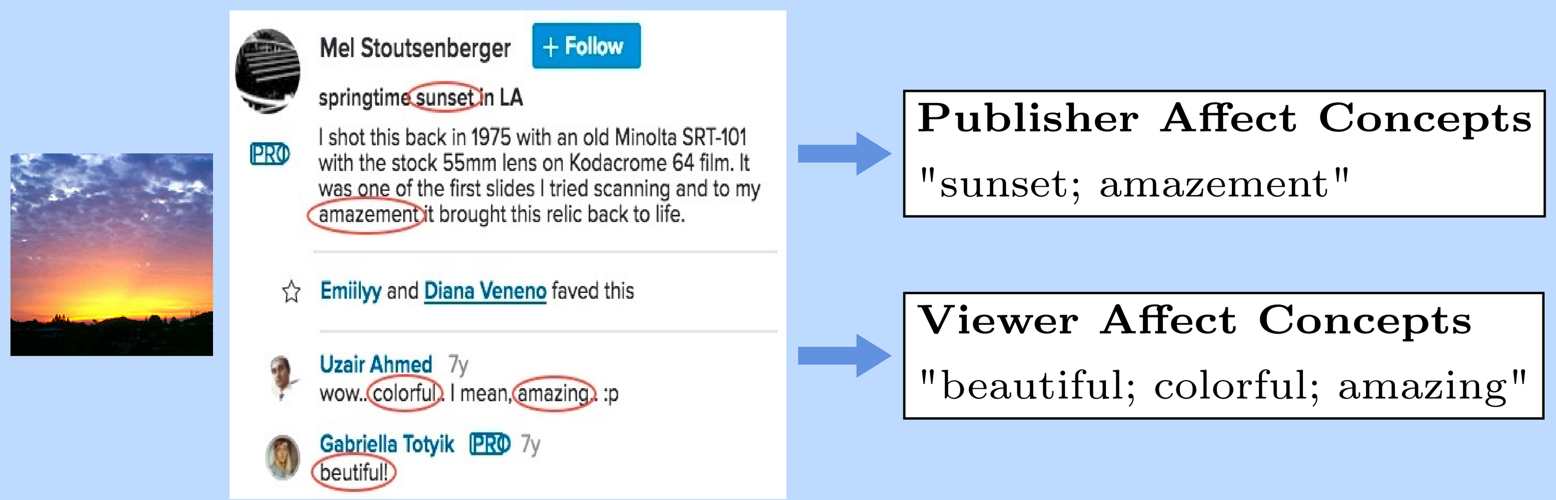


Fig.1 Image with Comments

The semantic content of an image is presented in the form of Adjective-Noun Pair (ANP). As shown in Figure 1, an image with an ANP "amazing sunset" is likely to trigger relevant comments like "beautiful", "amazing" and "colorful". Based on the mid-level understanding of how publisher affect concepts (PACs) evoke certain viewer affect concepts (VACs) provide, we present a framework to predict viewers' response. We adopt the the architecture of the basic framework proposed by [2], but specifically target publisher-viewer affective concepts correlation.

The novel contributions of our work will include,

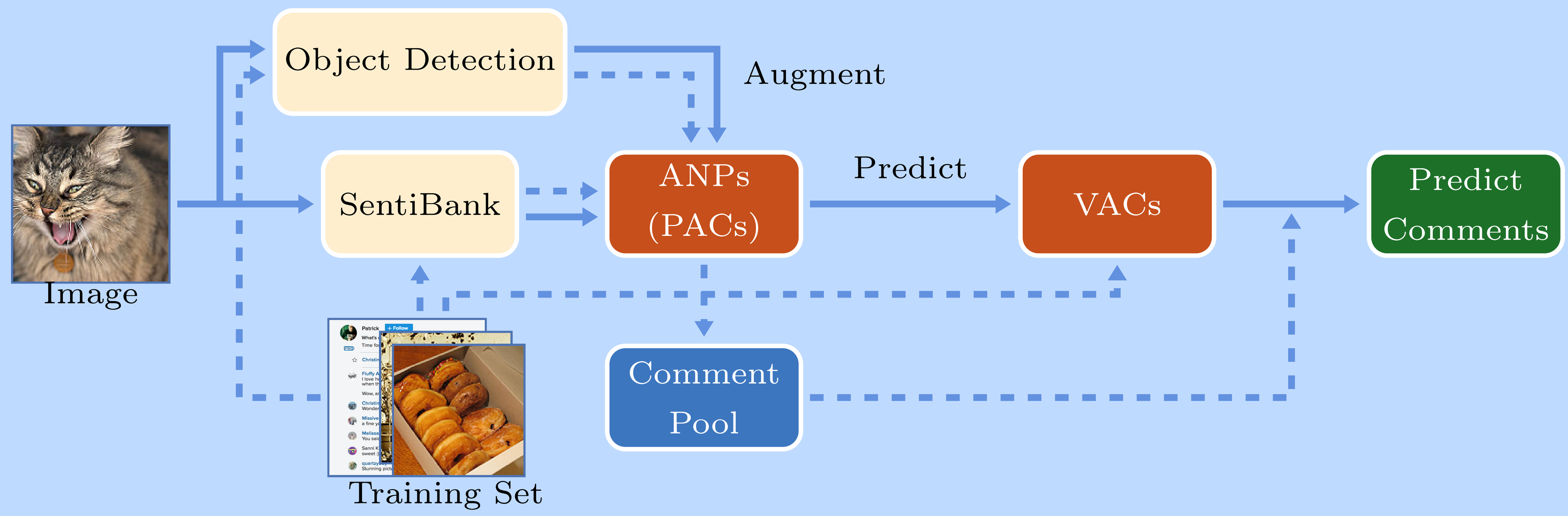
- employing an auxiliary object detection module to augment ANPs, and thereby reduce inherited error induced by false detection;
- constructing a pool of candidate comments for prediction by clustering the embedding features of images;
- utilizing hierarchical probability model to address publisher-viewer affective concepts correlation.

Dataset

The dataset contains 91162 Flickr images and their associated comments and metadata. The training corpus containing the Flickr images and corresponding metadata are downloaded from `http://www.ee.columbia.edu/ln/dvmm/vso/download/flickr_dataset.html`. The framework is trained on 80% of the data and test on 20% of the data.

System Overview

We employ DeepSentiBank, a CNN-based visual sentiment concept classifier, to extract the PACs from images. Adjectives with strong sentiment value are mined from comments associated with Flickr images to construct a pool of VACs. Given a test image, we predict its potential VACs based on the probabilistic correlation between PACs and VACs, and thereby propose comments from a comment pool which contains comments extracted from images with similar embedding features from the training set.



Method

PAC Detection

DeepSentiBank produces a probabilistic distribution over 2089 ANPs defined in Visual Sentiment Ontology[3]. The ANPs with highest probability are assigned to the image as their PACs. Furthermore, we will train a multi-object detection module[4] to remove ambiguous PACs detected by mistake.

VAC Detection

VACs are adjectives with strong sentiment value discovered from real user comments associated with Flickr Images. The sentiment value is measured by SentiWordNet. The VAC pool includes 324 adjectives with absolute sentiment value larger than 0.125, and frequently occur in the comments.

Probabilistic Correlation Model

Currently, we apply Naive Bayes probabilistic models to estimate the posterior probability of each VAC given a test image and the co-occurrence statistics from training data. The posterior probability of VACs given a test image d_i is,

$$P(v_j|d_i;\theta) \propto P(v_j|\theta)P(d_i|v_j;\theta), \tag{1}$$

where $P(v_j|\theta)$ is the frequency of VAC v_j appearing in the training set. The likelihood of a given image d_i and VAC v_j is represented by a multinomial model. Unlike the multivariate Bernoulli model used in [2], multinomial model[5] captures probability of how likely a PAC is assigned to an image. We also apply non-negative matrix factorization to smooth the conditional probability matrix and recover the missing correlation between PACs and VACs.

Our future work will focus on capturing publisher-viewer affect correlation using Latent Dirichlet allocation (LDA) model. We assume that each image contains a mixture of VACs, which represent a higher-level interpretation of the image; and the existence of PAC is attributable to the topic of the image, that is VAC.

Application

Comment Prediction

We generate a comment pool which contains comments from images with similar embedding features detected by DeepSentiBank. Specifically, we use K-Nearest Neighbor (KNN) for clustering. Comments in the pool are represented by bag-of-words as a vector C_l, indicating the occurrence of VAC in the comments. Then we measure the relevance between test image and candidate comments by taking inner product $C_l \cdot V_i$ where V_i is the posterior probability of any VAC given test image i. Candidate comments with highest relevance score are recommended to the user.

Evaluation

We evaluate the VAC prediction performance by and overlap-ratio,

$$\text{overlap} = \frac{|\{\text{ground truth}\} \bigcap \{\text{Predicted}\}|}{|\{\text{ground truth}\} \bigcup \{\text{Predicted}\}|}. \tag{2}$$

References

1. R. Balasubramanyan, B. R. Routledge, and N. A. Smith. From tweets to polls : Linking text sentiment to public opinion time series. In International Conference on Weblogs and Social Media, 2010
2. Y. Chen, T. Chen, W. H. Hsu, and H. M. Liao and S. Chang. Predicting Viewer Affective Comments Based on Image Content in Social Media. Proceedings of International Conference on Multimedia Retrieval, 233:233–233:240, 2014.
3. T. Chen, D. Borth, T. Darrell and S. Chang. DeepSentiBank: Visual Sentiment Concept Classification with Deep Convolutional Neural Networks, CoRR, 2014.
4. T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar. Focal Loss for Dense Object Detection. CoRR, 2017
5. A. McCallum and K. Nigam. A comparison of event models for Naive Bayes text classification. IN AAAI-98 Workshop on Learning for Text Categorization, 41-48, 1998.