Deep Neural Networks In Fully Connected CRF For Image Labeling With Social Network Metadata
Chengjiang Long  Roddy Collins  Erin Swears  Anthony Hoogs
Kitware Inc. (1712 Route 9 Suite 300, Clifton Park, NY 12065)
{chengjiang.long, roddy.collins, eran.swears, anthony.hoogs}@kitware.com

Introduction

Observation: Social multimedia dataset contains (1) images, (2) text information like title, description, comments, and (3) other meta information like user information, image gallery, uploader-defined groups, and links between shared contents.

Intuition: We hypothesize that using social media context jointly with pixel information should improve the state-of-the-art in image labeling.

Goal: We seek to understand the relative contribution of pixels, text and other information in predicting image labels.

Competing Algorithms


MIR-9K Dataset

A subset of the MIRFLICKR dataset, contains 6000 + 3182 images with 24 categories. It involves a set of 3,213 users, a collection of 34,942 words and 17,687 image groups.

Effectiveness of metatdata

Compare with the state-of-the-art approaches

Effectiveness of the text-level CNN

Experiments

We propose a novel deep fully connected CRF based framework with a joint end-to-end CNN-RNN formulation for image labeling which combines the strengths of both CNNs and RNNs.

Conclusion and future work

Our future work includes investigating more effective meta information, and improving the efficiency of the current DCRF framework to handle more complicated real-world application problems.