ADVISE: Symbolism and External Knowledge for Decoding Advertisements

Keren Ye
Adriana Kovashka
Department of Computer Science, University of Pittsburgh

Introduction
- Advertisements embed references to outside knowledge, and inspire us to ask:
  - How to utilize symbolic references and knowledge to understand the meaning of an ad?
- We formulate the ad understanding task as matching an ad image to human-written statements about the ad’s message.
- We interpret an ad using symbolic region proposals and apply bottom-up attention to aggregate information.
- We use external knowledge as a constraint to regularize the model, and incorporate discovered object-symbol mappings.

Method
- **Basic image-text triplet embedding**
  - The distance between an image and its corresponding statement should be smaller than the distance between that image and any other statement, or between other images and that statement.
- **Image embedding using symbols**
  - We use Huang et al., 2017 to train a region proposal network and fine-tune on symbol box annotations of Hussain et al., 2017.
  - We use the bottom-up attention mechanism (Anderson et al., 2017) to aggregate features from different proposals.

Dataset
- We use the PITT image ads dataset (Hussain et al., CVPR 2017)

Experiments
- **Evaluate on the main ranking task**
  - Rank of the highest-ranked true matching statement
  - Recall@3: number of correct statements ranked in the Top-3

Synonyms learnt by the extra constraints

Ablation study (% improvement over basic embedding)

Association of image regions and words
- Given the query words, we use k-NN to retrieve the most related image regions from the test images

Acknowledgement

NSF Grant Nr 1566270
Google Faculty Research Award
NVIDIA hardware grant