Presentation of a paper

September 15, 2021

1 Unknown vocabulary

Multimodal Recurrent Neural Network,

2 Paper goal

- To describe data with non-fixed categories (current SOTA has fixed label but too restrictive)

- Get large descriptions rather than one sentenced-ones

- Make a model that is free of assumptions

3 Current challenges

Many labeled datasets available but several mentions of the same categories from one dataset to another (hard to go and change labels to have a coherent dataset)

4 Solutions offered in the paper

- Make weak labels from multiple datasets to associate labels from multiple datasets

- Multimodal RNN with input an image, output text description

5 Achievements

- SOTA performances on image-ranking experiments.

- Outperforms retrieval baselines on both fullframes and region-level experoments (beats)