Image Caption Summary

Guided Open Vocabulary Image Captioning with Constrained Beam Search

- Problems:
  - previous models cannot be applied to out-of-domain cases
- Contribution
  - extend image caption models to out-of-domain cases
  - developed constrained beam search to inject constraints in output sequence

Paying More Attention to Saliency: Image Captioning with Saliency and Context Attention

- Problems:
  - when people look at an image, they tend to pay attention to salient regions
- Contribution
  - proved that teaching model to pay attention to salient regions during caption generation will improve caption quality

Towards Diverse and Natural Image Descriptions via a Conditional GAN

- Problems:
  - models training with MLE objective tend to generate repeated and boring captions since it encourage the model to use n-grams appeared in training samples
  - previous works focus on fidelity and ignore other qualities like natural and diversity
  - conventional evaluation metrics tend to favor safe but restrictive way. Under these metrics, sentences that contain matched n-grams would get substantially higher scores than those using variant expressions
- Contributions:
  - extend GAN to caption generation

Fast, Diverse and Accurate Image Captioning Guided By Part-of-Speech

- Problems:
  - beam search is computational expensively and is likely to generate generic captions
  - GAN and VAE can generate diverse captions but not accurate, and its latent variables don't have exposed semantics to explicitly show the relationship between latent variables and generated captions

- Contribution:
    - proposed to generate captions based on an image and a sequence of POS