

Weilin Cong

✉ wxc272@psu.edu Ⓜ CongWeilin ☎ 814-308-2517

I am a Ph.D. student at the Department of Computer Science and Engineering, Pennsylvania State University. My research interest focuses on Computer Vision, Nature Language Processing and the combination of both. My latest research targets on understanding the latent interactions between objects in images, and this work has been developed as an website application to benefit users without experience in Computer Science area, see www.website.com.

Experience

- | | |
|---|--------------------------|
| Pennsylvania State University | State College, PA |
| ○ <i>Graduate Assistant</i> | Jan 2018–Present |
| - CMPSC 431W Database Management System Teaching Assistant. | |
| - CMPSC 497 Data Mining Teaching Assistant. | |
| China Mobile Research Institute | Beijing, China |
| ○ <i>Software Engineer Intern</i> | Sep 2016 - June 2017 |
| - Mobile face recognition application development. | |
| - Face detection algorithm development | |

Education

- | | |
|--|--------------------------|
| Pennsylvania State University | State College, PA |
| ○ <i>Ph.D. in Computer Science Engineering</i> | Aug 2017 - Present |
| Beijing Institute of Technology | Beijing, China |
| ○ <i>B.S. in Computer Science</i> | Sep 2013 - June 2017 |

Research Projects...

- **Scene Graph Generation** Jan 2018 - Present



- In this work, we design an end-to-end trainable network that generates scene graph from an image. We propose Concatenate-Relation layer that can capture the sequential order of subject and object involved in each relationship. We further propose a novel Label Feature Refinement layer that can capture label contextual information using Conditional Random Fields.
- The experiments quantitatively demonstrate that our model can outperform existing approaches on CLEVR, VRD and Visual Genome datasets. For instance, compared to current state-of-the-art models, Recall@100 of our model are raised respectively from 24.99% to 49.95%, from 41.92% to 50.47%, and from 54.69% to 54.77% on the CLEVR, VRD and Visual Genome.
- Research has been developed as a website application www.website.com with user friendly UI, to benefit users without experience in computer science area.

o Image Caption

Sep 2018 - Present

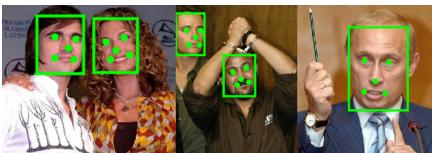


A person wear jacket and helmet is riding motorbike on the street.

- The performance of existing image caption models are limited by monolithic vector representations, which largely ignore the structural information and semantic meaning in visual scene. In this work, we design an image caption model that generate text-based descriptions according to both scene graph and the visual appearance of an image. We propose an efficient Predicate Lite layer that extract predicate feature by automatically selecting desired high level feature freely from an image according to the entity involved, which can help our model get rid of the limitation of scene graph scales (40 objects is the upper-bound for existing works), and speed up at least 35% (reduce from 0.34 sec/iter to 0.22 sec/iter) without hurting the performance. The predicted scene graphs are treated as additional information to help caption models generate meaningful text descriptions.
- The latest features will be added to website application www.website.com in the near future.

o Face Recognition

Sep 2016 – June 2017



- Face recognition contains three steps: face detection, alignment and identification. In this work, we design an coarse-to-fine joint cascade network that can solve face detection and alignment tasks simultaneously within 15 ms/image. We further proposed a Hard Example Mining method for joint cascade network training that can accelerate convergence and lead to a lower loss value.
- Caffe implementation <https://github.com/CongWeiLin/mtcnn-caffe> (more than 330 stars and 200 forks on github) has been released to benefit researchers interested in face recognition.
- Details can be found at <https://arxiv.org/pdf/1707.09364.pdf>

Technical and Personal skills

- o **Related Fields:** Machine Learning, Data Structures and Algorithms, Computer Vision, Data Mining
- o **Programming Languages:** Python, C, C++, Matlab
- o **Machine Learning Tools:** Tensorflow (current preference), Caffe