Qi Rao

Contact

EMAIL: raoqi1219@gmail.com
PHONE: +86 18986561219

Experience

Nov 2017 - Oct 2018 — Research Intern at SenseTime,Beijing

Worked as full-time intern in autopilot detection group, under supervision of Dr.Xingyu Zeng. Responsible for the GOD-Pro's development and maintaining, which is a toolkit for ConvNet models deploying including 2d detection, landmarks, face direction,etc. Also concentrated on research about object detection and instance segmentation. Spent two months in Japan(Tokyo and Kyoto) to schedule the data collection route and conduct several times real-car tests for GOD-Pro.

Jun 2017 - Sept 2017 — Research Intern at Seetatech,Beijing

Worked as full-time intern under supervision of Dr.Jie Zhang and Prof.Shiguang Shan. Help to remove waterwave on document photos with deep convolutional networks. Efficiently improve face detector's performance. Projects finished by myself went into production system.

Jan 2016 - Jun 2017 — Research Intern at ICT, Chinese Academic of Science

Multimedia Group

Under supervision of Prof.Ke Gao. Researched in visual tracking field, in particular human tracking, with emphasis on speed accelerating and performance improving. Familiar with basic deep learning based computer vision technology and related tools.

Education

Sept 2019 or later - — Ph.D candidate
University of Technology Sydney
Under supervision of Prof.Yi Yang. I am in gap year due to the VISA delay.

Aug 2013 - Jun 2017 — Bachelor of Science(Engineering) with Honours, E-Commerce
Electronic Engineering and Computer Science
Queen Mary University of London
with Honours - Second Class(Upper Division)

Aug 2013 - Jun 2017 — Bachelor of Engineering, E-Commerce Engineering
Beijing University of posts and telecommunications
Overall GPA:85+/100, ranking: 4/172

Projects

GOD-Pro — *General Object Detection toolkit for Autopilot Driving*

Developed and maintained a toolkit which contains a standard video processor with control functionality and several algorithm modules with their associated functionality. Each algorithm module is integrated with the latest deep ConvNet model within group and corresponding interface with pre/post-processing. These core algorithms are for object perception,like 2d detection(tracking), 3d detection, face/body direction, vehicle classification, traffic light status, dense estimation,etc. Controler collects the outputs of all modules and visualize in real-time. This project has been merged into SenseTime's autopilot driving system and undergone several real-car tests in highway or city road situation in Japan.

VideoCaptioner — *Generating captions for videos*

Worked with Guang Li(Phd student in UTS) as UTS-CAI team in TRECVID 2018. Built an encoder-decoder framework for the task of video caption. For encoder, we used multi-feature fusion training(c3d, resnet101 and i3d), sequence level loss to improve the performance. For decoder, we used LSTM to generate the sentences and used beam search for performance boost. We tried reinforcement method to directly improve the score like BLEU4,METEOR,CIDEr but failed once change the dataset, seemed directly improving the score lead to bad generalization ability. We also tried convolutional encoding and 'Attention is all you need', which worked well in machine translation task, in video caption after TRECVID, and indeed improved the performance. The result finally ranked 2nd in VTT task. (Ranking: CMU, UTS(ours), UPC, NTU, kslab, ku-ispl; And for Runtype=N, we ranked 1st.)

WaveNet | *Deep fully-convolutional network for waterwave removal*

Designed a deep fully-convolutional network for the specific task of removing waterwave on document photoes, whose architecture combined the FCN and ResNet-101. Experiments evaluated with different loss function design, including pixel-wise loss , perceptual loss and GAN. This work improved the performance of face detection and finally went into production system.

MDNet+ | *An improved tracking framework for human target*

Reimplemented MDNet with Caffe. Accelerated the online tracking speed. Improved human tracking representation ability with supplementary training on self-built dataset.

PoseDetector | *Human pose detection and reminding system*

Obtained the skeleton information by 3d camera, implemented self-refinement algorithm to judge human pose in real-time.

Olrenting | *A generic E-commerce online shopping framework*