Yandong Li

4328 Scorpius Street, Orlando, FL 32826

Email : lyndon.leeseu@outlook.com

Mobile : +1-321-202-0296

Education

- **University of Central Florida** — Orlando, FL
  Ph.D. in Computer Science, Advisor: Dr.s Liqiang Wang and Boqing Gong (remote) — *Aug. 2017 – Present*
- **Southeast University** — Nanjing, China
  Bachelor of Software Engineering; GPA: 3.7/4.0; Ranking: 3/109 — *Aug. 2012 – July. 2016*

Research Interests

I am a third-year Ph.D. student at UCF. My primary research areas include machine learning and computer vision. My recent works mainly investigate object detection, the distributions of adversarial examples for deep neural networks, multi-task learning, visual question answering and segmentation, action recognition, and video summarization.

Publications

- **Improving Object Detection with Selective Self-Supervised Self-Training (ECCV 2020)**
  ***Yandong Li*, Di Huang, Danfeng Qin, Liqiang Wang, Boqing Gong**
  We study how to leverage Web images to augment human-curated object detection datasets. We customize self-training for the object detection task by introducing a selective net and a self-supervised loss. We train the selective net in such a way that it not only selects correct bounding boxes out of proposals but also creates a safe zone to avoid messing up the hard negative mining used in many object detection algorithms. We report state-of-the-art results on detecting backpacks and chairs from everyday scenes, along with other challenging object classes.
- **Neural Networks Are More Data-Efficient Teachers Than Human Raters:Active Mixup for Knowledge Distillation from a Blackbox Teacher Model (CVPR 2020 Oral)**
  ***Yandong Li**, Dongdong Wang*, Liqiang Wang, Boqing Gong (* Equal Contribution)**
  We study how to train a student deep neural network for visual recognition in a data-efficient manner when there exists a blackbox teacher model. Progress on this problem can significantly reduce the dependence on large-scale human curated datasets for learning high-performing visual recognition models. Our approach blends image mixup and active learning. The former effectively augments the few available unlabeled images by a big pool of synthetic images sampled from the convex hull of the original images, and the latter actively chooses hard examples from the candidate pool for the current student network and then uses them to query labels from the teacher model. We validate our approach with extensive experiments.
- **BachGAN: High-Resolution Image Synthesis from Salient Object Layout (CVPR 2020)**
  ***Yandong Li*, Yu Cheng, Zhe Gan, Licheng Yu, Liqiang Wang, Jingjing Liu**
  We propose a new task towards more practical application for image generation - high-quality image synthesis from salient object layout. This new setting allows users to provide the layout of salient objects only (i.e., foreground bounding boxes and categories), and have the model complete the drawing with an invented background and a matching foreground. We propose Background Hallucination Generative Adversarial Network (BachGAN) for the novel task, which first selects a set of segmentation maps from a large candidate pool via a background retrieval module, then encodes these candidate layouts via a background fusion module to hallucinate a suitable background for the given objects. Experiments on Cityscapes and ADE20K datasets demonstrate the advantage of BachGAN over existing methods, measured on both visual fidelity of generated images and visual alignment between output images and input layouts.
- **AdaFilter: Adaptive Filter Fine-tuning for Deep Transfer Learning (AAAI 2020)**
  Yunhui Guo, *Yandong Li*, Liqiang Wang, Tajana Rosing
  There is an increasing number of pre-trained deep neural network models. However, it is still unclear how to effectively use these models for a new task. Fine-tuning is a popular technique for deep neural networks where a few rounds of training are applied to the parameters of a pre-trained model to adapt them to a new task. Despite its popularity, in this paper we show that fine-tuning suffers from several drawbacks. We propose an

adaptive fine-tuning approach, called AdaFilter, which selects only a part of the convolutional filters in the pre-trained model to optimize on a *per-example* basis. We use a recurrent gated network to selectively fine-tune convolutional filters based on the activations of the previous layer. We experiment with 7 public image classification datasets and the results show that AdaFilter can reduce the classification error of the standard fine-tuning by up to 5%.

- **$\mathcal{N}$ATTACK: Learning the Distributions of Adversarial Examples for an Improved Black-Box Attack on Deep Neural Networks (ICML 2019)**
  Yandong Li*, Lijun Li*, Liqiang Wang, Tong Zhang, Boqing Gong (* Equal Contribution)**
  We propose a black-box adversarial attack algorithm that can defeat both vanilla DNNs and those generated by various defense techniques developed recently. Instead of searching for an "optimal" adversarial example for a benign input to a targeted DNN, our algorithm finds a probability density distribution over a small region centered around the input, such that a sample drawn from this distribution is likely an adversarial example, without the need of accessing the DNN's internal layers or weights. Our approach is *universal* as it can successfully attack different neural networks by a single algorithm. It is also *strong*; according to the testing against 2 vanilla DNNs and 13 defended ones, it outperforms state-of-the-art black-box or white-box attack methods for most test cases.

- **Depthwise Convolution is All You Need for Learning Multiple Visual Domains (AAAI 2019)**
  Yandong Li*, Yunhui Guo*, Rogerio Feris, Liqiang Wang, Tajana Rosing (* Equal Contribution)**
  We propose a multi-domain learning architecture based on depthwise separable convolution. The proposed approach is based on the assumption that the images from different domains share cross-channel correlations but have domain-specific spatial correlations. The proposed model is compact and have minimal over-head when being applied to new domains. We have evaluated our approach on Visual Decathlon Challenge, a benchmark for testing the ability of multi-domain models. According to the results of experiments, our approach can achieve the highest score while only requiring 50% of the parameters compared with the state-of-the-art approaches.

- **How Local is the Local Diversity? Reinforcing Sequential Determinantal Point Processes with Dynamic Ground Sets for Supervised Video Summarization (ECCV 2018)**
  ***Yandong Li*, Liqiang Wang, Tianbao Yang, Boqing Gong**
  We propose a novel probabilistic model, built upon SeqDPP, to dynamically control the time span of a video segment upon which the local diversity is imposed. In particular, we enable SeqDPP to learn to automatically infer how local the local diversity is supposed to be from the input video. The resulting model is extremely involved to train by the hallmark maximum likelihood estimation (MLE), which further suffers from the exposure bias and non-differentiable evaluation metrics. To tackle these problems, we instead devise a reinforcement learning algorithm for training the proposed model. Extensive experiments verify the advantages of our model and the new learning algorithm over MLE-based methods.

- **VQS: Linking Segmentations to Questions and Answers for Supervised Attention in VQA and Question-Focused Semantic Segmentation. (ICCV 2017)**
  Chuang Gan, *Yandong Li*, Haoxiang Li, Chen Sun, Boqing Gong
  We present the preliminary work of linking the instance segmentations provided by COCO to the questions and answers (QAs) in the VQA dataset and name the collected links as VQS. They transfer human supervision among the previously separate tasks, providing more effective leverage to existing problems, which offer inspiration and references for new research problems and models. We study two applications of the VQS data in this paper, which is supervised attention for VQA and a novel question-focused semantic segmentation task.

Experience

- **Google Research (AI:D)** — Seattle, US
  Research Intern — *May 2020 - August 2020*
  - **Ranking Neural Checkpoints:** Establish a benchmark and propose a creative approach to rank the neural checkpoints without training.

- **Google Research (Mobile Vision)** — Seattle, US
  Student Researcher — *Jan 2020 - May 2020*

- **Auto Data:** Self-supervised learning

- **Google Cloud&&AI** Sunnyvale, US
  Research Intern *Sep 2019 - Dec 2019*
  - **Object Detection:** Improving object detection with selective self-supervised self-training. Note that our paper has been accepted by ECCV2020!
- **Microsoft Cloud&&AI** Seattle, US
  Research Intern *May 2019 - Aug 2019*
  - **Generative model:** High-resolution image generation. Note that our paper has been accepted by CVPR2020!
- **IBM Thomas J. Watson Research Center** New York, US
  Research Intern *May 2018 - Aug 2018*
  - **Multitask learning** We proposed a novel approach for multitask learning.
- **Baidu Institute of Deep Learning (IDL)** Beijing, China
  Research Intern *Feb 2017 - Aug 2017*
  - **Action Recognition:** Our team won the 3^{rd} place in Youtube-8M competition and 1^{st} place in ActivityNet challenge which are prestigious competitions for action recognition
- **Microsoft Reseach Asia (MSRA)** Beijing, China
  Research Intern *Aug 2015 - July 2016*
  - **Video Thumbnail Tools:** I have taken part in the development of Microsoft Cognitive Service.
  - **Hashing:** Deep cross-modal hashing won the outstanding dissertation for a Bachelor's degree.

Programming Skills

- **Languages: Python, Java, C++**
- **Familiar with Pytorch, Caffe and other deep learning development tools.**

Awards

- **Sep. 2014 National Scholarship for Undergraduate Student**
- **Jun. 2016 Outstanding Undergraduate Student**
- **Apr. 2017** 3^{rd} **place in Youtube-8M**
- **Sep. 2017** 1^{st} **place in ActivityNet challenge**

Professional services & activities

- **PC member**: NeurIPS 2020, MM 2020, ECCV 2020, ICML 2020, CVPR 2020, ICLR 2020, Bigdata 2019, MM 2019, WACV 2019, Bigdata 2018, WACV 2018
- **Reviewer**: NeurIPS 2019, CVPR 2019, ICCV 2019, IJCAI 2019, IJCAI 2018