Yiwen **Shao**

Tel: 410-900-3948 | Email: yshao18@jhu.edu

EDUCATION

Johns Hopkins University, Baltimore, USA — Aug.2017 - May.2019
Whiting School of Engineering
- MSE in Computer Science

Southeast University, Nanjing, China — Sep.2013 - Jun.2017
School of Electronic Science and Engineering
- B.Eng in Internet of Things Engineering
- GPA 3.81/4 RANK: 1/30

WRITINGS

Conference Paper

[1] **Yiwen Shao**, Qigong Lin, "Use of Pitch Continuity for Robust Speech Activity Detection", **ICASSP 2018.**
[2] Qiguang Lin, **Yiwen Shao**, "A Novel Normalization Method for Autocorrelation Function for Pitch Detection and for Speech Activity Detection", **Interspeech 2018**.
[3] Ke Li, Nanxin Chen, **Yiwen Shao**, Daniel Povey, Sanjeev Khudanpur, "Syntax-aware Pointer Sentinel Mixture Model for Code-switched Language Modeling", **ICASSP 2019** (submitted).

Project Report

[1] **Yiwen Shao***, Chun-Chieh Chang*, "Parallelizing SGD".
[2] Dhananjay Singh*, Manya Wadhwa*, **Yiwen Shao*,** "GANerating Images from Text"

EXPERIENCES

CLSP, JHU, Baltimore, MD — Aug.2017 - Present
Research Assistant, advised by Daniel Povey

Auxiliary Augmentation

A novel label-unpreserved data-augmentation method that adds an auxiliary label of the permutation (e.g. rotation) we did on the data and trains the network to recognize it as well as its class label in a multitask learning way.
- Implemented it in both Kaldi and PyTorch, and tested it on cifar, mnist, f-mnist with ResNet, ResNetX, WideResNet, and PyramidNet, resulting in a consistent improvement.
- code available at: https://github.com/YiwenShaoStephen/Auxiliary-Augmentation

Waldo

A toolkit for text localization and instance segmentation in PyTorch using a new non-detection method. It trains the neural net with two pixel-wise outputs -- semantic output and offset output. Then we use priority queue and greed algorithm to "decode" the image. That is, gradually merging adjacent pixel(s) to instances according to the output from networks.
- Implemented basic utilities including neural network models, data preparation, train/test etc. and wrote the first example in Waldo on DSB2018 dataset that outperforms Mask-RCNN.
- code available at: https://github.com/waldo-seg/waldo (extended: https://github.com/YiwenShaoStephen/mergeNet)

Kaldi

A state-of-the-art comprehensive ASR toolkit.
- Implemented basic image processing utilities.
- Implemented feature extraction, data augmentation, slant-correction etc. for OCR**.**
- Got state-of-the-art results for image recognition on cifar10/100 with an improvement of 1%/6% absolutely.
- Implemented and tested Mix-up augmentation, rank reduction by SVD decomposition.

Code Switch

A syntax-aware gated pointer sentinel mixture model (**G-PSMM**) for code-switched language modeling.
- Contributed to the train/test scripts in PyTorch.
- Run experiments and tuned models.

In-Progress Projects

- **PyChain:** Implementing Lattice-Free MMI in PyTorch.
- **TorchSpeech:** Implementing pre-processing and data-loader for speech recognition, and integrating it with Fairseq, OpenNMT.
- **Weighted Softmax** for language model integration in attention model, working with Prof. Shinji Watanabe.

Baihu Co., Ltd, Wuxi, China Jun.2016 - Aug.2017

Speech Processing & Software Engineering Intern, advised by Qiguang Lin

- Implemented a Speech Activity Detection (SAD) system in C++/Java based on the Combo-SAD for feature extraction and the Ring Processing and SVMs for classification.
- Proposed a novel algorithm that utilizes pitch continuity to improve SAD performance
- Proposed exponential ACF that gets better results on both SAD and Pitch Detection.
- Setup the Baihu Corpus using sox and FaNt.
- Trained the GMM-MFCC SAD system using CMU Sphinx.

PROJECTS

MergeNet (extended Waldo)
- Implemented FCN, PSPNet, FPN etc. to do both semantic segmentation and instance segmentation.
- Wrote example scripts for Cityscape, COCO dataset.
- code available at: https://github.com/YiwenShaoStephen/mergeNet

GANerating Images from Text
- Used DC-GAN and WGAN to generate images from text.
- Reproduced "Attention GAN with Deep Attentional Multimodal Similarity Model" (DAMSM) to improve the results.

Parallelizing SGD
- Implemented Hogwild! and synchronized SGD in Pytorch.
- Did experiments on MNIST and analyzed the accuracy and efficiency plot of different synchronizing methods.

IMDB Movie Score Prediction
- Analyzed and visualized the relationship between movie features and its IMDB score.
- Predicted IMDB movie score with Naive Bayes, Random Forest, Decision Tree, Ada Boost. MLP, SVM, KNN, and compared the results.

Toy NLP
- Implemented HMM Part-of-Speech Tagging.
- Implemented Log-Linear ngram Language Model to do text categorization and speech recognition on switchboard
- Built Earley Parser with pruning and proposed a special pruning trick to avoid duplicate checks in "attach" operation, which achieved a significant speedup.

Path Planning & Navigation System of the SEU KidSize Robot
- Developed the upper control module of the Path Planning & Navigation System in C++ on Linux.
- Designed a robust guidelines extraction method based on the Otsu's algorithm for image thresholding.
- Rank 3rd in Path Planning in the 11st SEU RoboCup.

SKILLS

- **Languages**: Python, Shell, C/C++, Java, Perl
- **Deep Learning:** Pytorch, Kaldi, Tensorflow

RESARCH INTERESTS

- Sequence Modeling
- Automatic Speech Recognition, NLP
- Unsupervised/Semi-supervised Learning