

Haiyi Mao

Email : HAM112@pitt.edu

Mobile : +1-617-818-8720

EDUCATION

- **University of Pittsburgh** Pittsburgh, PA
PhD program in Intelligence Systems Sept. 2019 – Present
- **Northeastern University** Boston, MA
Master of Science in Computer Science Sept. 2014 – Dec. 2016
- **Xidian Univerisity** Xi'an, China
Bachelor of Engineering in Computer Science and Technology Sept. 2008 – Jul. 2012

RESEARCH INTEREST

Biomedical Informatics: Clinical Data Analysis, Computational Biology

Machine Learning: Deep Learning, Multi-view Data Analysis, Bayesian learning, Interpretable Deep learning, Representation Learning, Graph Model, Feature Selection

INDUSTRIAL EXPERIENCE

- **Shiseido Americas Corp.(Giaran acquired by Shiseido Amdericas)** Cambridge, MA
Machine Learning Engineer Nov 2017 - Aug, 2019
 - **2D & 3D Face Landmarks Detection based on Deep Learning:** Designed a real-time face landmark detection algorithm based on deep learning (**Pytorch**) with Lucas-Kanade loss to mitigate landmarks drifting for videos
 - **Optimizing the deep learning model for CPU real time inference:** Designed a light-weight deep learning landmark detection algorithm for fast inference on CPU. Speed was enhanced from **1 FPS to 40 FPS**
 - **Deep learning model in Python conversion to JavaScript for website application:** Created a unique tool to translate the deep learning model from Python to Javascript in order to integrate with web application
 - **C++ facial landmarks detection algorithm conversion to Javascript :** Developed a tool to convert C++ code and OpenCV into Javascript without sacrificing the speed of execution (**Tensorflow.js** drops 15 FPS but ours only drops 3 FPS). This breakthrough saved 1 millions for Shiseido Corp.
- **Giaran Inc.** Boston, MA
Machine Learning Engineer(one of the core engineers before acquisition) Aug 2017 - Nov 2017
 - **Face landmarks detection:** Developed the facial landmarks detection algorithm in **C++** with **OpenCV**
 - **Algorithm performance optimization:** Optimized algorithm performance with **multi-threading** with **OpenGL** to enable the real-time inference
 - **Machine learning model size reduction:** Reduced the model size by sparse matrix to improve the mobile and web application performance
- **Philips Health** Cambridge, MA
Research Intern Jan 2017 - Oct 2017
 - **Machine Learning in Health Care:** Designed and implemented a clinical early warning algorithm based on machine learning methods for electronic health records (**EHR**) databases. The algorithm took real-time spo2, heart rate, blood rate, etc as inputs to generate alerts for patients in general ward need to be delivered to ICU
 - **Data Process:** Developed a package based on **Pyspark** to do parallel pre-processing and outlier detection for real time **EHR** from Philips devices
 - **Distributed Computation:** Embedded the early warning algorithm into **Pyspark** parallel machine learning pipeline. Improved the time efficiency 200 times (1 PB testing data set) compared with the original version.
 - **Parallel Deep Learning:** Applied **Pyspark** on deep learning model in order to do parallel inference with multi-GPUs

RESEARCH PROJECTS

- **Analyze Mice Neuron Data** Boston, MA
Northeastern University, Boston University Advised by Raymond Fu, Xue Han *May 2015 - Sept 2015*
 - **Analyze data:** Created clustering analysis pipeline on the collected data from monkeys' brains
- **Deep Transfer Boosting** Boston, MA
Northeastern University Advised by Yun Fu *May 2015 - Sept 2015*
 - **Model Creation:** Proposed a deep boost algorithm with the *Rademacher* complexities loss to avoid over-fitting in transfer learning scenario
 - **Theorem Proof:** Proved two theorems to show the algorithm can converge well with two different distributions data
- **Multi-view Feature Selection** Boston, MA
Northeastern University Advised by Yun Fu *Jan 2016 - March 2016*
 - **Model implementation:** Implemented a feature selection algorithm for multi-view data by aligning the the different views into same space by k-means. This algorithm achieved good performance with only $O(k)$ time complexity
- **Super Resolution** Boston, MA
Northeastern University Advised by Yun Fu *Feb 2016 - May 2016*
 - **Partially pixelized images:** Proposed a brand new problem for super resolution from partially pixelized images.
 - **GAN for super resolution:** Designed a **GAN**(Generative Adversarial Network)(one of the first works about GAN in super resolution) in **auto-encoder** structure to solve super resolution problem
 - **Special Deconvolutional Layers:** Added two layers of special deconvolutional layer to reconstruct pixelized images
- **Video Summarization with Text Information** Boston, MA
Northeastern University Advised by Ehsan Elhamifar *Jun 2016 - Dec 2016*
 - **Feature Extraction and NLP features:** Extracted video spatio-temporal features by **3D convolutional network** and texts descriptions features
 - **Subset Selection Algorithm:** Implemented the subset selection algorithm by using **multi-view** features (3D deep features and NLP features)

PUBLICATIONS

- **Partial Pixelated Image Super Resolution with Convolutional Neural Network**
Haiyi Mao, Jun Li, Yue Wu, Yun Fu
ACM Multi-Media ACMMM 2016, rank 5%, **best paper nominated**
- **Robust Multi-View Feature Selection**
Hongfu Liu, Haiyi Mao, Yun Fu
IEEE International Conference of Data Mining ICDM 2016, 8.26% acceptance rate)
- **Video Summarization with Side Information**
Ehsan Elhamifar , Haiyi Mao
(under review)
- **Deep Transfer Boosting**
ShuhuiJiang, Haiyi Mao, Zhengming Ding, Yun Fu
IEEE Transactions on Neural Networks and Learning Systems
- Yun Fu, Shuyang Wang, Sumin Lee, Songyao Jiang, Bin Sun, Haiyi Mao, Kai Ho Edgar Cheung. 2018
Systems and Methods for Virtual Facial Makeup Removal and Simulation, Fast Facial Detection and Landmark Tracking, Reduction in Input Video Lag and Shaking, and a Method for Recommending Makeup US PATENT US20190014884A1

PROGRAMMING SKILLS

- **Languages:** Python, C++, Java, Javascript, SQL, MATLAB
- **System:** Linux, MacOS, Windows, Git
- **Deep Learning:** Tensorflow, Pytorch, MXNet, Caffe, Keras, Torch
- **Data Base:** MySQL, MongoDB, Elastic Search
- **Packages:** OpenCV, NumPy, SciPy, Scikit-learn
- **Parallel Computation:** Pyspark, OpenGL

TEACHING & RESEARCH ASSISTANCE

- **Research Assistance Advised by Ehsan Elhamifar** Fall, 2016
- **Teaching Assistance For Machine Learning** Fall, 2016
- **Research Assistance Advised by Raymond Yun Fu** Spring, 2016
- **Teaching Assistance For Math In Computer Science** Fall, 2015