Songyao Jiang

360 Huntington Ave, Boston MA, USA • Tel: +1(734) 546-0695
jiangsongyao@gmail.com • GitHub • Homepage • Google Scholar • LinkedIn

EDUCATION

Northeastern University — **Boston, MA**
Ph.D. in Computer Engineering — **06/2016 – 05/2022**
- Concentration: Computer Vision and Pattern Recognition, Machine Learning.

University of Michigan — **Ann Arbor, MI**
Master of Science in Electrical Engineering: Systems — **09/2013 – 05/2015**
- Coursework: Linear Algebra, Machine Learning, Image Processing, Embedded System Programming, *etc*.

Hong Kong Polytechnic University — **Hong Kong**
Bachelor of Engineering in Electrical Engineering, — **09/2009 – 06/2013**
- Coursework: Programming, Computer Architecture, Operating Systems, Analog and Digital Circuits, *etc*.

RESEARCH INTERESTS

Multi-Modal LLMs, Human Face and Gesture Analysis, Video Classifications, Human Detection and Pose Estimation, Generative Models, Skeleton-base Action Recognition, Sign Language Recognition.

EXPERIENCE

Amazon.com, Inc. — **Boston, MA**
Applied Scientist at AGI Foundations — **11/2023 – present**
- Worked on developing Amazon Nova multi-modal LLMs.

Applied Scientist at Lab126 — **06/2022 – 11/2023**
- Worked on computer vision and machine learning in Amazon Devices AI team.

Northeastern University — **Boston, MA**
Graduate Research Assistant in SMILE Lab — **06/2016 – 05/2022**
- Advisor: Prof. Yun (Raymond) Fu
- Research topics: computer vision: pose estimation, sign language recognition, generative models, *etc*.

Graduate Research Assistant in Power Electronics Research Group — **09/2015 – 06/2016**
- Advisor: Prof. Bradley Lehman
- Research topic: machine learning based photovoltaic power prediction.

AInnovation Labs, Inc. — **Boston, MA**
Founding Member and Computer Vision Engineer Intern. — **02/2022 – 05/2022**
- Developed key machine learning algorithms in the core products, including real-time AI color calibration system, virtual makeup addon, removal and recommendation system, and face detection and alignment system.

Giaran, Inc. (Acquired by Shiseido Americas) — **Boston, MA**
Founding Member and Computer Vision Engineer Intern. — **01/2017 – 09/2017**
- Developed key machine learning algorithms in the core products, including real-time AI color calibration system, virtual makeup addon, removal and recommendation system, and face detection and alignment system.
- Our startup was then acquired by Shiseido Americas.

Teld New Energy — **Qingdao, Shandong, China**
Research Engineer in Electric Vehicle Research Team — **05/2015 – 08/2015**
- Research topic: grouped smart mass charging system for electric vehicles (EV).
- Developed a smart charging algorithm for massively grouped EV charging based on SVM and dynamic programming to mitigated charging load and surge on power system, optimized the use of renewable energy.

Nagoya University — **Nagoya, Aichi, Japan**

Research Assistant in Suzuoki Lab — **05/2014 – 08/2014**

- Advisor: Prof. Takeyoshi Kato
- Research topic: mathematical modelling of renewable energy.

CHALLENGES

CVPR 2021 Challenge on Agriculture-Vision Pattern Recognition — **04/2021 – 06/2021**

- Team leader and first contributor. Ranked the 4th place in supervised track. [GitHub][Leaderboard]
- Developed a multi-modal and self-constructing GCN for multi-label agricultural pattern recognition given RGB and infra-red aerial agriculture images.

CVPR 2021 Challenge on Signer-Independent Isolated Sign Language Recognition — **12/2020 – 04/2021**

- Team leader and first contributor. **1st place winner** in both RGB and RGB+D tracks. [GitHub][Leaderboard]
- Proposed a novel spatio-temporal GCN with attention mechanism to learn dynamics in whole-body skeleton graph as well as fusing with RGB, optical flow and depth HHA video modalities via a unified skeleton-aware multi-modal framework to recognize sign language glosses from input RGB+D videos.

PUBLICATIONS

- B. Sun, Y. Zhang, **S. Jiang**, and Y. Fu, "Hybrid Pixel-Unshuffled Network for Lightweight Image Super-Resolution," *AAAI*, 2023. [Preprint][GitHub][Demo]
- **S. Jiang**, B. Sun, L. Wang, Y. Bai, K. Li, and Y. Fu, "Sign Language Recognition via Skeleton-aware Multi-modal Ensemble," *Under Review*, 2022. [Preprint][GitHub]
- **S. Jiang**, B. Sun, L. Wang, Y. Bai, K. Li, and Y. Fu, "Skeleton Aware Multi-modal Sign Language Recognition," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2021. [Paper][GitHub]
- **S. Jiang**, Z. Tao, and Y. Fu, "Geometrically Editable Face Image Translation with Adversarial Networks," *IEEE Transactions on Image Processing (TIP)*, vol. 30, pp. *2771-2783*, 2021. [Paper]
- **S. Jiang**, H. Liu, Y. Wu, and Y. Fu, "Spatially Constrained GAN for Face and Fashion Synthesis," in *16th IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 2021. [Paper][GitHub][Award][Web]
- Y. Yin, J. P. Robinson, **S. Jiang**, and Y. Fu, "SuperFront: From Low-resolution to High-resolution Frontal Face Synthesis," in *Proceedings of ACM Multimedia (ACMMM)*, 2021. [Paper][GitHub]
- Y. Yin, **S. Jiang**, J. P. Robinson, and Y. Fu, "Dual-attention GAN for Large-pose Face Frontalization," in *15th IEEE International Conference on Automatic Face & Gesture Recognition (FG),* 2020. [Paper][GitHub]
- S. Sarkar, W. Kang, **S. Jiang**, K. Li, S. Ray, E. Luther, A. R. Ivanov, Y. Fu, and T. Konry, "Machine Learning-aided Quantification of Antibody-based Cancer Immunotherapy by Natural Killer Cells in Microfluidic Droplets," Lab on a Chip, 20(13), pp. *2317-2327*, 2020. [Paper]
- Z. Hong, T. Sun, **S. Jiang**, K. Li, Y. Fu, H. Xu, J. Zhang, Y. Liu, Q. Ye, and H. Cang, "Harnessing Deep Learning to Overcome Photo-toxicity for Live-cell Imaging," *Under Review*, 2020.
- **S. Jiang**, Z. Tao, and Y. Fu, "Segmentation Guided Image-to-Image Translation with Adversarial Networks," in *14th IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 2019. [Paper][GitHub]
- T. Alashkar, **S. Jiang**, and Y. Fu, "Rule-Based Facial Makeup Recommendation System," in *12th IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 2017. [Paper]
- T. Alashkar, **S. Jiang**, S. Wang, and Y. Fu, "Examples-Rules Guided Deep Neural Network for Makeup Recommendation," in *Proceedings of AAAI Conference on Artificial Intelligence (AAAI)*, 2017. [Paper]
- **S. Jiang** and T. Kato, "Dynamic Modelling of Combined Cycle Power Plant for Load Frequency Control with Large Penetration of Renewable Energy," in *7th JUACEP Workshop*. 2014.

OTHER PROJECTS

Light-weight and Video-based Multi-person 2D Pose Estimation with Tracking 02/2020 – 12/2020

- Developed a novel model that utilizes temporal information of human body movement between adjacent video frames via a temporal-aware deep neural network. Refined the pose estimation results in real-time scenarios and handled difficult occlusion cases.
- Compressed parameter size and reduced computational cost by replacing normal CNNs with our proposed novel low-rank pointwise residual modules.
- Improved performance by introducing a multi-scale heatmap fusion and supervision module.
- Collected and labeled yoga data to improve the performance of extreme poses during exercises. [Example]
- Deployed on mobile devices using CoreML (iOS) and TensorFlow Lite (Android). [Demo]
- Won GapFund360 Award and filed two patent applications (Status: Published). [Patent1][Patent2]

Face Recognition and Verification in Low-light Condition Using Transfer Learning 05/2019 – 11/2019

- In low-light condition, we utilized mid-range and long-range infra-red (IR) wavelengths to obtain the portrait images of the target persons for face recognition and verification.
- Developed a semi-supervised metric learning method and an unsupervised adversarial method to transfer the knowledge from visible spectrum to IR spectrum.
- Achieved much higher recognition rates (domain adaptation setting) and verification rate (transfer learning).

Single-Image Robust Automatic White Balance Under Mixed Light 09/2016 – 01/2017

- Developed a mixed-light automatic white balance algorithm using iterative neutral color pixels voting scheme and chromatic analysis as additional constraints and solve least square using matting Laplacian matrix.
- Estimated faithful skin color under mixed light with guidance from facial landmarks for neutral color voting.
- Deployed using OpenCV/native C++ and also on Universal Windows Platform (UWP) apps using C#. [Report]

Facial Attributes Classification, Makeup Recommendation and Addon Systems 02/2016 – 08/2016

- Collected a facial attribute and makeup dataset (e.g., skin color, face, and eye shapes). Developed a facial attribute classification system using pretrained deep features and multi-class SVM.
- The predicted classes of facial attributes were then used to recommend makeup styles for users using a learned knowledge-based system learned from YouTube makeup videos. A makeup add-on system is developed to virtually visualize the recommended makeup.
- Awarded NSF I-Corps Grant. Used in our startup company "Giaran, Inc." [Patent]

Machine-Learning Based Snow Effect and Photovoltaic Power Output Prediction 12/2015 – 03/2016

- Predicted the snow effects on photovoltaic (PV) power output during winter when PV panels experienced snowfalls. A fully-connected neural regression and clustering model was trained on historical weather and power data of solar farms to predict the snow effect on the PV power output.

PATENTS

- Y. Fu, **S. Jiang**, B. Sun, "Light-Weight Pose Estimation Network with Multi-Scale Heatmap Fusion," Granted. *US Patent 12,205,317.* [Patent]
- Y. Fu, **S. Jiang**, "Segmentation Guided Image Generation with Adversarial Networks," Granted. *US Patent 10,825,219.* [Patent]
- Y. Fu, **S. Jiang**, "Video 2D Multi-person Pose Estimation using Multi-frame Refinement and Optimization," Published. *WIPO Patent App. No.: WO 2020/232069*. [Patent]
- Y. Fu, S. Wang, S. Lee, **S. Jiang**, B. Sun, H. Mao, K. H. E. Cheung, "Systems and Methods for Virtual Facial Makeup Removal and Simulation, Fast Facial Detection and Landmark Tracking, Reduction in Input Video Lag and … ," Published. *US Patent App. No: 16/584,310.* [Patent]

ACADEMIC SERVICE

Conference PC Member and Reviewer

- International Conference on Computer Vision (ICCV)
- International Joint Conferences on Artificial Intelligence (IJCAI)
- IEEE International Conference on Automatic Face & Gesture Recognition (FG)
- IEEE International Conference on Data Mining (ICDM)
- IEEE International Conference on Multimedia Information Processing and Retrieval (MIPR)

Journal Reviewer

- IEEE Transactions on Image Processing (TIP)
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- IEEE Transactions on Multimedia (TMM)
- Journal of Visual Communication and Image Representation (JVCI)
- The Vision Computer (TVCJ)
- IET Image Processing
- Journal of Electronic Imaging (JEI)

Workshop Reviewer

- IEEE International Workshop on Analysis and Modeling of Faces and Gestures Workshops (AMFG)

HONORS & AWARDS

- NSF I-Corps Grant 2022
- NVIDIA CCS Best Student Paper Award 2021
- Champion of the CVPR 2021 Challenge on Sign Language Recognition (both RGB & RGBD tracks) 2021
- 4th Rank in CVPR 2021 Challenge on Agriculture-Vision (supervised track) 2021
- PhD Network Travel Grant, Northeastern University, USA 2019
- GapFund360 Award, Northeastern University, USA 2018
- NSF I-Corps Grant, National Science Foundation 2016
- JASSO Scholarship, Nagoya University, Japan 2014
- Outstanding Scholarship, Hong Kong Polytechnic University 2010, 2011, 2012, 2013

SKILLS

Languages:	English (full professional), Chinese (native), Cantonese, Japanese (basic).
Deep Learning Frameworks:	PyTorch (proficient), TensorFlow, CoreML (good knowledge).
Programming Languages:	Python, C/C++, C#, Java, HTML, JavaScript.
Others:	OpenCV, MATLAB, AWS E2 S3, Google Colab, Slurm, UWP, Git, etc.