

# Li Yu

+18148529160 | [liyupsu@gmail.com](mailto:liyupsu@gmail.com) | <https://liyu10000.github.io>

## Education

---

<b>Pennsylvania State University, US</b>	August 2019 – June 2021 (expected)
• Master of Science in <i>Informatics</i>	GPA: 3.95/4.0
<b>Emory University, US</b>	August 2014 – August 2017
• Master of Science in <i>Chemistry</i>	GPA: 3.2/4.0
<b>Peking University, China</b>	September 2009 – July 2014
• Bachelor of Science in <i>Chemistry</i>	GPA: 3.5/4.0
• Bachelor of Science in <i>Mathematics and Applied Mathematics (double major)</i>	GPA: 3.5/4.0

## Projects

---

<b>Extent of Corrosion Image Analysis</b>	January 2020 – Present
• Work as a research assistant to collaborate with teams from PPG Industries, Inc. on the ongoing project of ship corrosion evaluation.	
• Developed a pipeline to evaluate extent of corruptions over ship surface based on images.	
<b>Explainable Image Representation Learning</b>	January 2020 – Present
• Hand crafted image features and applied discriminant analysis & convex optimization to learn explanation vectors for images.	
• Built positive gradient distance for a few-shot image classification task and achieved similar results as to Resnet.	
<b>3D Room Layout Estimation</b>	June 2020 – July 2020
• Worked in a team of three on the project of learning 3D room layout from a single panoramic image.	
• Extended current SOTA HorizonNet by incorporating an hourglass architecture to learn better layout detection. Pushed 3D intersection-over-union (IoU) from 84% to 87%.	
<b>Continuous Affect Synthesis</b>	January 2020 – June 2020
• Designed VA-StarGAN to synthesize continuous facial emotion expressions based on valence and arousal (VA) intensities.	
• Proposed to utilize basic emotions to guide the regression training on VA values.	
• Trained on in-the-wild datasets with a half million images and generated continuous, photo-realistic facial expressions that outperformed SOTA methods.	
<b>Facial Expression Recognition from Videos</b>	January 2020 – June 2020
• Worked in a team of three on the project of recognizing facial expressions from in-the-wild videos.	
• Exploited teacher-student training scheme to be able to utilize a combination of labeled and unlabeled datasets. Trained a noisy student network iteratively to enhance robustness.	
• Isolated key regions of the face and applied multi-level attention mechanism to further boost performance.	
<b>Early Identification of Students at Risk</b>	September 2019 – December 2019
• Mined students' learning statistics on online platforms such as demographics, assessment results, and daily activities into a clean format for identifying their risk of dropping or failing a course.	

- Proposed an ensemble learning framework, including eight base machine learning classification algorithms, and achieved 94.5% prediction accuracy.

## Experience

---

### Teaching Assistant

September 2019 – December 2019

- Assisted professor in a senior-level undergraduate course: Applied Data Science.
- Responsible for designing and grading assignments, class participation, helping students with semester long data science projects on real-world problems.

### Data Scientist at Pacific Online, China

May 2019 – July 2019

- Matched remarks and reviews of online products with over 100,000 entries of brand/product names via word segmentation and named entity recognition.
- Performed sentiment analysis and utilized Reddit ranking algorithms to assign scores at brand/product/attributes levels that achieved 90% precision.

### Data Scientist at Stimage Tech, China

March 2018 – April 2019

- Led a team of six for development of medical image-based analysis systems: Thinprep Cytologic Test (TCT) and Tumor Mutation Burden prediction (TMB). Deployed in Southern China Hospital.
- Modified classical YoloV3 model to contain four yolo layers to capture small objects and added one more classification branch to allow for weakly supervised training.
- Utilized SVM and XGBoost to summarize disease cells detection/classification results to achieve 99.9% accuracy on cell-level classification and 85% accuracy on disease level identification.

## Publications

---

**Yu, L.**, Davaasuren D., Rao S., Kumar V. (2020, August). StarGAN-EgVA: Emotion Guided Continuous Affect Synthesis. In *ACM Multimedia Conference HuMA Workshop*.

Kumar V., Rao S., **Yu L.** (2020, August). Noisy Student Training using Body Language Dataset Improves Facial Expression Recognition. In *European Conference on Computer Vision BEEU Workshop*.

**Yu, L.**, Cai, T. (2020, July). Ensemble Learning for Early Identification of Students at Risk from Online Learning Platforms. In *International Conference on Data Science*.

## Technical Skills

---

Languages: C/C++, Python, Matlab

Tools: TensorFlow, PyTorch, Keras