# PENGYU(BEN) YUAN

### **SUMMARY**

• I am a Ph.D. candidate skilled in Machine Learning and Deep Learning with applications in Medical and Seismic Image Analysis. I have published 16 research papers in conference (MICCAI, SEG) and journal: (IEEE-TMI, Geophysics) etc. I have a wealth of practical problem handling experience due to my 3 internships during my Ph.D. and published 2 patents. My career interests are bringing the advanced and mysterious AI techniques to solve the real-world problems.

### **SOFTWARE SKILLS**

- Computer programming: Python, MATLAB, C/C++, CUDA, Shell
- Machine Learning: Pytorch, TensorFlow, Keras, Theano, SciKit-Learn
- Numerical Analysis: MATLAB, NumPy, SciPy, Pandas

#### **WORK EXPERIENCE**

• Dataminr New York, NY

*May* 2021 - *Aug.* 2021

Research Intern, advised by Dr. Svebor Karaman and Dr. Mahdi Abavisani

- Studied and compared different SOTA methods on few shot object detection (FSOD) problem.
- Proposed a meta-learning based FSOD approach using dual spatial feature enhancement and sparse representation learning which improves 0.15 average precision (AP) on COCO benchmark.

• Sensia (Schlumberger) Houston, TX

Jun. 2019 - Aug. 2019

Data Scientist Intern, advised by Dr. Nam Nguyen and Dr. Jonathan Chong

- Worked on ESP (Electric Submersible Pump) close loop control for artificial lift. Used reinforcement learning to automatically respond to the alerts from Auto Events Detector (AED) and constructed a well simulator to collect training data for the model.
- o Published patent: Event driven control schemas for artificial lift.

• Schlumberger Houston, TX

Jun. 2018 – Aug. 2018

Data Scientist Intern, advised by Dr. Nam Nguyen and Dr. Jonathan Chong

- Applied deep neural network model to identifying critical events of ESP (Electric Submersible Pump) under the well.
- Established an automated complete local optimization workflow for target wells.
- Published patent: System and method for managing wellsite event detection.

• HULA Lab at University of Houston Houston, TX

Aug. 2017 - Present

Research Assistant, advised by Dr. Hien Van Nguyen and Dr. Jiefu Chen

- Developed algorithms in meta-learning, few-shot learning, Bayesian learning and applied to classification/detection/segmentation problems in different areas (e.g. medical image and seismic image).
- Published 16 research papers, featured on MICCAI/SEG conferences and IEEE TMI journal.

# ACADEMIC PROJECTS

# • Unsupervised Seismic Data Deblending and Interpolation

Sep. 2020 - Jul. 2021

- Proposed the blend-trace network to modify receptive field for unsupervised seismic data reconstruction task.
- Proposed two regularization techniques (automatic spectrum suppression and mix training) for anti-aliasing seismic interpolation.
- o Submitted patent: Method for Reconstructing at Least One Trace in a Seismic Image.

## • Brain Cell Type Classification with Few Training Samples

Sep. 2019 – Nov. 2021

- Proposed the AGILE framework which combines the strengths of data-augmentation, active learning, and meta-learning and can outperform the meta-learning baseline when the number of training tasks is limited.
- Introduce the optimal transport for data augmentation for new task and use the Bayesian Neural Network to select the most informative samples.
- With only 1% of training data (8 cells), the model can achieve 90% accuracy on the classification of unseen cell types. With 3% as training samples, it can achieve the upper bound.

### • First Break Picking on Seismic Shot Gather Image

*Jul.* 2018 – Mar. 2020

- Established a segmentation plus picking workflow for first arrival picking problem. Used U-net based deep convolutional networks for segmentation and proposed different picking strategies. The CNN plus RNN model shows the least absolute error on both synthetic and field data.
- Applied modified model-agnostic meta-learning (MAML) model for fast adaption.

#### • Lung Nodule Detection and Cancer Screening

Oct. 2017 - Present

- Established an incidental lung nodule dataset (Regular-dose CT images) from Houston Methodist Research Institute.
- Developed the meta-learning lung cancer classification model that can be adapted from public dataset (LUNA16) to lung screening dataset from another institute and incidental lung nodule dataset, which achieved clinically relevant performance (0.925 and 0.891 AUC-ROC respectively) with 30 new labeled samples.

#### **SELECTED PUBLICATIONS**

- 1. Yuan, P., et al. "Self-supervised learning for anti-aliasing seismic data interpolation." First International Meeting for Applied Geoscience & Energy. Society of Exploration Geophysicists, 2021.
- 2. **Yuan, P.**, et al. "A deep learning model-based lung cancer risk assessment for incidental pulmonary nodules." *AACR Annual Meeting* (2021): 2614-2614.
- 3. Yuan, P., et al. "Few Is Enough: Task-Augmented Active Meta-Learning for Brain Cell Classification." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2020.
- 4. **Yuan, P.**, et al. "Adaptive first arrival picking model with meta-learning." *SEG Technical Program Expanded Abstracts* 2020. Society of Exploration Geophysicists, 2020. 1486-1490.
- 5. **Yuan, P.**, et al. "A robust first-arrival picking workflow using convolutional and recurrent neural networks." *Geophysics* 85.5 (2020): U109-U119.
- 6. **Yuan**, **P.**, et al. "First arrival picking using U-net with Lovasz loss and nearest point picking method." *SEG Technical Program Expanded Abstracts* 2019. Society of Exploration Geophysicists, 2019. 2624-2628.
- 7. Yuan, P., et al. "Phasetime: Deep learning approach to detect nuclei in time lapse phase images." *Journal of clinical medicine* 8.8 (2019): 1159.
- 8. Vo, Hung Q., Yuan, P., He, T., Wong, S. T., & Nguyen, H. V. "Multimodal Breast Lesion Classification Using Cross-Attention Deep Networks." 2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI). IEEE, 2021.
- 9. Mobiny, A., Yuan, P., Cicalese, P. A., Moulik, S. K., Garg, N., Wu, C. C., ... & Nguyen, H. V. "Memory-Augmented Capsule Network for Adaptable Lung Nodule Classification." *IEEE Transactions on Medical Imaging* (2021).
- 10. Mobiny, A., Yuan, P., Cicalese, P. A., & Van Nguyen, H. "DECAPS: Detail-Oriented Capsule Networks." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2020.
- 11. Cicalese, P. A., Mobiny, A., **Yuan, P.**, Becker, J., Mohan, C., & Van Nguyen, H. "StyPath: Style-Transfer Data Augmentation for Robust Histology Image Classification." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2020.
- 12. Padmanabhan, M., Yuan, P., Chada, G., & Nguyen, H. V, "Physician-friendly machine learning: A case study with cardiovascular disease risk prediction." *Journal of clinical medicine* 8.7 (2019): 1050.
- 13. Mobiny, A, Cicalese, P. A., Zare, S., **Yuan**, **P.**, Abavisani, M., Wu, C. C., ... & Van Nguyen, H. "Radiologist-level covid-19 detection using ct scans with detail-oriented capsule networks." *arXiv preprint arXiv*:2004.07407 (2020).
- 14. Cicalese, P. A., Rizvi, S. A., Wang, V., Patibandla, S., **Yuan, P.**, Zare, S., ... & Van Nguyen, H. "MorphSet: Improving Renal Histopathology Case Assessment Through Learned Prognostic Vectors." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2021.

#### **ACTIVITIES**

### • Invited Tutorials & Workshops

- Presented the "Bayesian Meta-Learning" tutorial in MICCAI '21.
- Presented the "Bayesian Deep Learning on Medical Image Analysis" tutorial in MICCAI '19.
- "Data Augmentation, Labeling, and Imperfections (DALI)" workshop in MICCAI '21.
- o "Medical Image Learning with Less Labels and Imperfect Data (MIL3ID)" workshop in MICCAI '19, and '20.

#### Teaching

- o Introduction to Machine Learning and Computer Vision (Spring 2018, 2019, and 2020)
- Neural Networks and Deep Learning (Fall 2018, and 2019)
- o Principles of Internetworking (Fall 2017)

### • Professional Services

- Reviewer for several conferences including MICCAI, SEG, IMAGE, and MOBIQUITOUS.
- Reviewer for several journals including IEEE Journal on Multiscale and Multiphysics Computational Techniques (JMMCT), Computerized Medical Imaging and Graphics, and Geophysics.
- o Reviewer for a book entitled "Meta-Learning: An Overview" in Elsevier.

## **HONORS & AWARDS**

- Featured as "great innovative idea" in NSF funded Computing Community Consortium for our research on "Physician-Friendly Machine Learning Algorithms for Medical Diagnosis".
- Outstanding Student (The highest honor for undergraduates and only 10 students are awarded each year), University of Electronic Science and Technology of China, China.
- Honorable Mention in American Mathematical Contest in Modeling.

# **EDUCATION**

• University of Houston

Ph.D., Electrical and Computer Engineering GPA: 3.81/4

• University of Electronic Science and Technology of China (UESTC) B.S., Electrical Engineering GPA: 3.96/4 (Ranked 16/351)

Houston, Texas, USA
Aug. 2017 – Expected: May 2022
Chengdu, Sichuan, China
Sep.2013 – Jul. 2017