Ting-Wei Wu

: waynewu6250.github.io
: waynewu@gatech.edu | : Github: waynewu6250 | : (404)-376-7640 | : ting-wei-wu

Research Interests in *deep learning, natural language processing, vision & text understanding.*

Education

Georgia Institute of Technology — Atlanta, GA
Ph.D. in Machine Learning; GPA: 3.88 — *Aug. 2019 - 2023 (Expected)*
Advised by Dr. Biing-Hwang Juang

University of California, Berkeley — Berkeley, CA
M.Eng. in Bioengineering; GPA: 3.83 — *Aug. 2017 – May. 2018*
Advised by Dr. Aaron Streets

National Taiwan University — Taipei, Taiwan
B.S. & M.S. in Electrical Engineering; GPA: B.S. 3.82 (3.93/4.3); M.S. 4.00 (4.22/4.3) — *Sep. 2012 – Jul. 2017*

Skills

Languages: **Python**, **C & C++**, **MATLAB**, **SQL**, Verilog, VBA, Java
Technologies: Pytorch, Tensorflow, Data Science, Docker, GPU, Flask, AWS, Linux OS, Git, AutoCAD, Photoshop, COMSOL
Coursera Certificates: Advanced ML (DL, Bayesian, RL, NLP), Deep Learning Specialization, Machine Learning
Courses: Math in ML, Probabilistic Graphical Model, Computational Data Analysis, Convex Optimization, Deep Learning, Deep Learning for texts, Digital Image Processing, Data Structures and Programming, Computer Vision, Operating Systems

Professional Experience

- **Applied Scientist Intern, Amazon** — Seattle, WA
  Alexa Speech - Language Modeling team **(pytorch, transformers, lightgbm, libsvm)** — *May. 2021 - now*
  - ASR second pass rescoring: Adopted LambdaMART with listwise loss to rescore N-best hypotheses with miscellaneous linguistic/non-linguistic signals. Introduced WER reduction of 7.70 % compared with 1-best of Alexa production model.
  - Feature engineering: Proposed a new ensemble approach of ASR BERT-based confidence sub-models to digest group information and other customer-related signals like query rewrites and error detection signals. Preprint for **ICASSP'22**.

- **Machine Learning Research Intern, VMware** — Palo Alto, CA
  Storage IO & Performance Engineering team **(pytorch, transformers, nltk, mysql, gensim)** — *May. - Nov. 2020*
  - Causality Extraction: Devised a new two-stream attention BiLSTM-CRF model on causality inference and paragraph-level pairing, enhancing F1 benchmark score to 0.74. Published in two internal conferences **VML'20, RADIO'21**.
  - Knowledge Graph: Established an end-to-end nlp pipeline for inter/intra sentence causality retrieval to extract 2300 useful causal relations out of 20000 problem requests within seconds to construct knowledge network for troubleshoot diagnosis.
  - Few-shot Learning: Introduced new ProtoNet-based method for few labeled data in casual tagging, boosting F1 by 11%.

- **Graduate Researcher, Speech & Spoken Language Processing Lab** — Atlanta, GA
  Task-oriented dialog understanding, advised by **Dr. Biing H, Juang** (IEEE Fellow, NAE member) — *Aug. 2019 - now*
  - Multi-intent: Developed a zero-shot multi-intent label-aware BERT framework for dialog turns. Published in **EMNLP'21**.
  - Context/Act-aware: Devised a new multi-turn dialog context-ware hierarchical BERT mechanism for multi-task training and exploited dialog act signals in state tracking. Two papers published in **INTERSPEECH'21**.
  - Machine Introspection: Adapted continual learning on task-oriented dialog domain adaptation taks and reinforced downstream dialog tasks on MultiWOZ/DSTC8 dialogue datasets by inducing knowledge and response scoring mechanism.

- **Graduate Teaching Assistant, Signal Processing & Electronic Device Innovation** — Atlanta & Taipei
  Intro to Signal Processing (ECE 2026) & Device Innovation at NTU **(matlab, SP-First)** — *2016-2017, 2019*
  - Matlab Sessions: Supervised and lectured in lab & recitation sessions with matlab programming on digital signal processing. Provided signal processing (FT, DTFT, DFT, z-transform) and matlab skills for undergraduate teaching.
  - Entrepreneurship Development: Led discussions & project management for 3 courses in Device Innovation and meetings with 100+ industry experts and students to develop new business models and research technology transfer, incubating 2 AI startups.

- **Research & Development Intern, Getac Technology Corp.** — Taipei, Taiwan
  Intelligent Baby Monitoring System **(python, raspberry pi, gpiozero)** — *Jul. 2014-Sep. 2014*
  - Project Management: Led 10+ person multi-disciplinary project team with senior engineers to design a wireless monitoring device using a raspberry pi controller for observing infant behavior.
  - Raspberry Pi: Controlled GPIO ports with python and designed chip layouts and exterior design using SolidWorks, AutoCAD.

Highlighted Projects

- **Machine Translation Quality Estimation** *Natural Language Processing*
  Deep Learning (CS7643) Research Project advised by Facebook AI *(pytorch, transformers, polyglot)* *Jan.-May. 2020*
  - Transfer Learning: Exploited predictor-estimator model with new transformer structure to pretrain large quality-labeled translation corpus in common languages and adapt to other scarce language QE data with semi-supervised self training.
  - Ensembling: Ensembled different predictions from fine-tuning an estimator and various pretained predictors in several languages like English, Chinese, German to Estonian, Nepali with XGBoost Model, where Pearson score beats baseline 0.11 by 2 times.

- **StackBoxer: Chatroom with bilingual AI chatbots - *https://chatbox.cc*** *Natural Language Processing*
  Full-stack online plaform for functional and pretrained chatbots *(Pytorch, Django, Docker, PostgreDB)* *Jan.-Mar. 2019*
  - StackBot: Modeled intent/tag identifier from tfidf features and Starspace embeddings for matching Stackoverflow queries.
  - Movie Bot, ChickBot, YourFbBot: Established a customized 2-layer seq2seq model with attention mechanism and self-designed reward mechanism with policy gradient reinforcement set up in Django+Docker+nginx backend environment.

- **DeepEyeNet: Image Captioning with keyword-driven report generation** *Image Captioning & NLP*
  Research Project collaborated with Gatech, UvA, KAUST *(keras, tensorflow, pandas)* *Feb.-Nov. 2019*
  - Dataset Preparation and evaluation: Prepared 18,854 images annotated by experienced ophthalmologists and designed a new evaluator for the caption generator jointly in adversarial training. Work published in **WACV'21**.
  - Transformer: Devised a new contextual transformer decoder with semantic attention of technical keywords and retinal images for medical image captioning. Introduced BLEU-avg and CIDEr 74% and 87% increase over baseline models and improved abnormality interpretability. Published in **ICMR'21, ICIP'21**. Work submitted to **WACV'22, AAAI'22**.

- **Novelty Intervention in Hunter-Gatherer Game of Polycraft** *Reinforcement Learning*
  Research project in DQN funded by DARPA *(pytorch, gym, socket)* *Apr.-Nov. 2020*
  - Target-DQN: Designed a vision-based DQN agent to perform tasks of navigation and localization in Polycraft simulation.
  - Novelty: Introduced adaptation mechanism to new environments with novelty intervention and measure success of actions.

- **Cellspectra: Unsupervised cell image segmentation** *Computer Vision*
  Graduate Lab Researcher advised by Dr. Peng Qiu at Gatech *(keras, tensorflow, MATLAB)* *Jan.-May. 2020*
  - Bacterial segmentation: Developed CNN-based unsupervised object segmentation modules for cell counting and tracking.
  - Raman spectra clustering: Exploited deep embedding clustering on raman vectors from 1-d autoencoder for segmentation.

- **PillNet: A medicine pill recognition search tool in the mobile device** *Computer Vision*
  Entrepreneurship Startup Team with Ministry of Science and Technology in Taiwan *(tensorflow, opencv, c++)* *Apr.-Jul. 2019*
  - SSD-MobileNet: Developed a pharmaceutical pill identification module in real-time mobile camera to identify pill location with single shot detection model in tensorflow trained with FDA pill image database.
  - Pill Recognition: Trained siamese network by minimizing triplet loss to recognize pills and retreive relevance information.

- **Integrated Cell-sorting Sensor System** *Data Science*
  UC Berkeley Streets Lab and NTU CMOS Biotechnology Lab Graduate Researcher *(python, sklearn, R)* *2014-2016, 2018*
  - Platform: Devised new flow cytometry approach to collect impedance data and classify cell properties with frequency analysis.
  - ML Data Analysis: Utilized clustering methods (Naive-Bayes, GMM, K-means, NN) and MATLAB to extract impedance data for library creation. Published work in **MicroTAS'17, IEEE NEMS'17, IMCS'16.**
  - Chip Design: Expedited high-throughput droplet grabbing hydrogel beads with parameters by ML optimization.

Other cs-related projects:
Comics generation from wGAN, Chinese lyrics generation by charRNN, Fire event data management with selenium, pandas, SQL, Malaria cell prediction, Kaggle Sales prediction competition, Circuit Fraig and Simulation with C++.

Honors & Awards

- **Taiwan Government Scholarship to Study Abroad**, Taiwan Ministry of Education *2021*
- **Graduate Research Assistantship**, Georgia Tech Electrical & Computer Engineering *2019*
- **Graduate Honor Fellowship**, UC Berkeley Fung Institute of Engineering *2018*
- **Member**, UC Berkeley Golden Key International Honor Society *2018*
- **Graduate Honor Fellowship**, National Taiwan University Graduate Institute of Electronics Engineering *2017*
- **Travel Award**, Taiwan Ministry of Science and Technology. IMCS conference 2016 *2016*
- **Delegate**, Taiwan Model APEC 2014 *2014*
- **School Delegate**, AIESEC Asia-Pacific Exchange and Leadership Development Seminar Symposium *2010*
- **Nominee of Representative of Taiwan**, Global Yong Leaders Conference *2010*

Selected Publications

- **Ting-Wei Wu**, Ruolin Su and Biing Juang, "A Label-Aware BERT Attention Network for Zero-Shot Multi-Intent Detection in Spoken Language Understanding." *The 2021 Conference on Empirical Methods in Natural Language Processing (**EMNLP**)*, Nov 2021.

- **Ting-Wei Wu**, Ruolin Su and Biing Juang, "A Context-Aware Hierarchical BERT Fusion Network for Multi-turn Dialog Act Detection." *The 22nd Annual Conference of the International Speech Communication Association (**Interspeech**)*, Aug 2021.

- Ruolin Su, **Ting-Wei Wu** and Biing Juang, "Act-Aware Slot-Value Predicting in Multi-Domain Dialogue State Tracking." *The 22nd Annual Conference of the International Speech Communication Association (**Interspeech**)*, Aug 2021.

- **Ting-Wei Wu**, Jia-Hong Huang, Chao-Han Yang, Elbert Liu, Hiromasa Morikawa and J. N. Tegner, "Deep Context-Encoding Network for Retinal Image Captioning." *2021 IEEE International Conference on Image Processing (**IEEE ICIP**)*, Sep 2021.

- Jia-Hong Huang, **Ting-Wei Wu** and Marcel Worring, "Contextualized Keyword Representations for Multi-modal Retinal Image Captioning." *ACM International Conference on Multimedia Retrieval (**ICMR**)*, Apr 2021.

- **Ting-Wei Wu**, Chien-Chun Hung, Chien-Chia Chen, Razvan Cheveresan, Rajesh Somasundaran, "Two-stream Self-attentive Network for Cross-sentence Causality Reasoning." *The 3rd VMware Machine Learning Conference (**VML**)*, Oct 2020. *Largest VMware General Research Conference (**RADIO**)*, May 2021.

- **Ting-Wei Wu**, Yung-An Hsieh and Yi-Chieh Liu, "Ensemble-based Transfer Learning for Low-resource Machine Translation Quality Estimation.", *arXiv:2105.07622*, May 2021.

- J. H. Huang, C. H. Yang, F. Liu, M. Tian, Y. C. Liu, **T. W. Wu**, I. H. Lin, K. Wang, H. Morikawa, H. H. Chang J. N. Tegner, "DeepOpht: Medical Report Generation for Retinal Images via Deep Models and Visual Explanation." *The 2021 IEEE Winter Conference on Applications of Computer Vision (**WACV**)*, Jan 2021.

- **Ting-Wei Wu**, I-Fan Chen, Ankur Gandhe, Yi Gu, Ivan Bulyko, "Learning to rank with BERT Confidence Model in Automatic Speech Recognition." submitted to *2022 IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP**)*, Jan 2022.

- **Ting-Wei Wu**, Jia-Hong Huang, Chao-Han Yang, Elbert Liu, Hiromasa Morikawa, J. N. Tegner, "KOTR: Keyword-oriented Transformer-based Retinal Image Captioning." submitted to *The 36th AAAI Conference on Artificial Intelligence (**AAAI**)*, Jan 2022.

- **Ting-Wei Wu**, Chia-Hong Gao, Yi-Zhan Huang, Ting-Wei Lin and Chih-Ting Lin, "Electrode Spatial Design for a New Microfluidics Impedance Cytometer," *The 21st International Conference on Miniaturized Systems for Chemistry and Life Sciences (**MicroTAS**)*, October 2017.

- **Ting-Wei Wu**, Chia-Hong Gao, Fan-En Chen and Chih-Ting Lin, "Impedance Spectroscopy for Microfluidic Particle-analyzing Device with Spatial-Coplanar Electrode Design," *The 12th Annual IEEE International Conference on Nano/Micro Engineered and Molecular Systems (**IEEE NEMS**)*, April 2017.

- **Ting-Wei Wu**, Chia-Hong Gao and Chih-Ting Lin, "A microfluidic cell counting device based on impedance sensing," *16th International Meeting on Chemical Sensors (**IMCS**)*, July 2016.