YIXIONG CHEN

Baltimore U.S.A.
+1 (410) 310 1390 ◇ ychen646@jh.edu

EDUCATION

Johns Hopkins University (JHU) Baltimore, United States
Ph.D. in Computer Science *Aug. 2023 - May 2027 (Expected)*

- Working on Vision Language models and Medical Imaging under the supervision of **Alan Yuille**.

Johns Hopkins University (JHU) Baltimore, United States
MSE in Computer Science (Overall GPA: 4.00 / 4.00) *Aug. 2023 - May 2025*

Chinese University of Hong Kong (CUHK) Shenzhen, China
Ph.D. in Data Science (Overall GPA: 3.90/4.00) *Sep. 2021 - Jul. 2023 (did not complete)*

- Excellent Paper on Science and Technology of Shenzhen (2023)
- Excellent Paper on Artificial Intelligence of Shenzhen (2022)

Fudan University Shanghai, China
B.Sc. in Data Science (Major GPA: 3.42/4.00, Overall GPA: 3.31/4.00) *Sep. 2016 - Jun. 2021*

- Highest Impact Award of IEEE Computer Society Biometrics Workshop (2024)
- Second Class Scholarship for Outstanding Graduates (2021)
- Huawei Cloud Scholarship (2018)

PROFESSIONAL EXPERIENCE

Waymo Research Mountain View, United States
Research Intern, AI Foundation Team *May. 2025 - Sep. 2025*

· Designed and implemented a **LLM-Trajectory-Evaluator** based on Gemini to judge the reasonableness and correctness of behavior in autonomous driving.
· Annotated a **reasoning** dataset for scene understanding and trajectory analysis as step-by-step thinking of trajectory evaluation.
· Authored a technical report with comprehensive experiments, in preparation for **CVPR 2026**.

Google Research Mountain View, United States
Student Researcher, Health AI Team *Jun. 2024 - Nov. 2024*

· Designed and implemented a **Vision Language model** based on CoCa 2B with a novel local attention mechanism for multi-image fusion, resulting in a 4% improvement in temporal classification accuracy.
· Curated a longitudinal chest X-ray image-report dataset with 1.5M samples, incorporating data from MIMIC-CXR, Chest Imagenome, and MS-CXR-T to track disease progression across multiple images.
· Authored a technical report with comprehensive experiments, published at **MICCAI 2025**.

Johns Hopkins University Baltimore, United States
Research Intern, CCVL (Alan Yuille's Lab) *Jul. 2022 - Mar. 2023*

· Analyzed the training dynamics of MLPs, CNNs, and ViTs across 8 datasets, discovering a fundamental learning preference termed *layer convergence bias*. The findings were published at **ICLR 2023**.
· Investigated radiomic features of liver tumors in CT slides and proposed a morphological algorithm for 3D tumor synthesis, which successfully passed a Visual Turing Test conducted by experienced clinicians. The results were published at the **NeurIPS Workshop 2022** and **CVPR 2023**.

Shenzhen Research Institute of Big Data (SRIBD) Shenzhen, China
Research Assistant, Medical Image Analysis Team *Jul. 2020 - Jul. 2023*

- Designed and refined **self-supervised pre-training** algorithms for ultrasound videos using contrastive learning. The results were published at **MICCAI 2021** and **TMI 2022**.
- Created layer-wise **fine-tuning algorithms** for pre-trained medical models using meta-learning techniques. The results were published at **MICCAI 2023**.
- Developed a label selection and correction algorithm inspired by KNN representation space to improve learning with **noisy data**. The results were published at **ICASSP 2024**.

Points Technology Shanghai, China
Machine Learning Intern *Mar. 2020 - Jun. 2020*

- Designed and implemented the SecureBoost algorithm with Numpy, which is an improvement of XGBoost under the horizontal federated learning standard.
- Optimized the SecureBoost to reduce its time complexity from $O(n^2)$ to $O(nlogn)$.
- Participated in the formulation of the standard *"Promotion Committee of China Communication Standardization Association: Federated learning technology tool for data circulation"*.

RESEARCH PROJECTS

Vision Language Reasoning for Autonomous Driving Evaluation Waymo, Summer 2025

- Constructed a Human-Labeled Trajectory Evaluation Dataset, comprising reasoning text descriptions and preference annotations to evaluate the VLMs' performance.
- Implemented a **scene language engine** to translate the driving scenes to natural language, enabling LLMs to understand them and do reasoning.
- Verified that **multi-modal** inputs and the CoT scoring of the trajectories are helpful for evaluation. The proposed method reaches 90% human understanding of the scene and 72.6% preference alignment.

Vision Language Understanding for Longitudinal Chest X-Ray Google, Summer 2024

- Analyzed chest X-ray datasets and proposed a data pipeline to decompose reports into symptom descriptions and progress evaluations for vision-language alignment and progression prediction.
- Developed a **regional cross-attention** mechanism to improve the vision encoder's ability to perform local comparisons, enabling recognition of nuanced differences across five conditions in image pairs.
- Implemented a **multi-stage fine-tuning** strategy to synchronize pre-trained and randomly initialized parameters, resulting in an 8.6% accuracy improvement over previous SOTA in temporal classification.

Medical Image Segmentation Evaluation Tools JHU, Spring 2024

- Developed **Quality Sentinel**, a novel **text-embedding-conditioned** tool for evaluating medical image segmentation quality without reference labels.
- Constructed a large-scale dataset of CT scans paired with pseudo labels of varying quality and corresponding ground truth DSC, totaling **4M image-mask pairs**.
- Performed extensive experiments demonstrating Quality Sentinel's effectiveness in diagnosing dataset quality, supporting human-in-the-loop annotation, and enhancing semi-supervised segmentation.

Human-centric Text-to-Image Generation JHU, Autumn 2023

- Constructed a large-scale human image dataset with three subsets: human-in-the-scene images, close-up face images, and close-up hand images, totaling 1,000,000 images.
- Proposed a novel **Mix of Experts (MoE)** mechanism, fusing two low-rank adaptation experts (LoRA) to enhance the generation of hands and faces, respectively.
- Evaluated the model on two novel benchmarks based on COCO and DiffusionDB human prompts, achieving **SOTA** performance among all open-source text-to-image diffusion models.

Text-to-Image Generation Benchmarking CUHK, Spring 2023

- Introduced a novel **hierarchical Chain of Thought** approach to prompt MiniGPT-4 for explainable evaluation of images generated by text-to-image diffusion models.
- Developed a large-scale **human preference dataset**, encompassing 800k human preference choices on 430k pairs of images from diverse sources.
- Benchmarked the performance of 15 models using COCO Caption and DrawBench prompts, evaluating their fidelity, alignment, and aesthetics in image generation from text.

Liver Tumor Synthesis JHU, Autumn 2022

- Developed a **morphological liver tumor synthesis** pipeline incorporating location selection, texture/shape generation, and image warping to generate 49,000 annotated synthetic tumor examples.
- Integrated clinical knowledge into the algorithm, using a randomized hyper-parameter tuning scheme to create diverse synthetic tumors capable of deceiving doctors.
- Applied nnUNet and Swin UNETR for tumor **segmentation** with a hybrid of synthetic and real tumors, achieving **SOTA** on the LiTS dataset.

Foundational Layer Properties of DNNs CUHK and JHU, Summer 2022

- Developed a measurement for transferability based on optimization paths, identifying the **layer-wise transferability** trend from ImageNet to 12 downstream image classification tasks.
- Defined a **layer-wise convergence rate** metric, analyzing learning speeds across different layers in a DNN, and found that shallower layers tend to learn faster than deeper layers.

Transfer Learning for Medical Image Analysis CUHK, 2020 - 2022

- Developed a **meta-learning**-based fine-tuning method to automatically determine the transfer strength for different DNN layers, achieving over 2% performance improvement.
- Implemented a **meta-contrastive learning** framework for ultrasound videos, utilizing semantic clustering of 3,000 videos, resulting in a more than 10% performance gain compared to ImageNet.

Person Re-identification with Changing Clothes Fudan University, Spring 2019

- Proposed a person re-id benchmark including 23,000+ pedestrians, each with 4 different clothes.
- Collected, organized, and annotated a video dataset for pedestrians using GTA-V virtual engine.
- Tested 10 existing SOTA person re-id algorithm frameworks on our benchmark.

PUBLICATIONS

Click to view Google Scholar

1. **Chen Y**, Xiao W, Bassi PRAS, Zhou X, Er S, Hamamci IE, Zhou Z, Yuille A. Are Vision Language Models Ready for Clinical Diagnosis? A 3D Medical Benchmark for Tumor-centric Visual Question Answering. **Arxiv 2025 preprint.**
2. **Chen Y**, Xu S, Sellergren A, Matias Y, Hassidim A, Shetty S, Golden D, Yuille A, Yang L. Coca-cxr: CoCa-CXR: Contrastive Captioners Learn Strong Temporal Structures for Chest X-ray Vision-Language Understanding. **MICCAI 2025 .**
3. Zhu J, **Chen Y**, Ding M, Luo P, Wang L, Wang J. MoLE: Human-centric Text-to-image Diffusion via Mixture of Low-rank Experts. **NeurIPS 2024.**
4. Li W, Qu C, Chen X, Bassi P, Shi Y, Lai Y, Yu Q, Xue H, **Chen Y**, et al. AbdomenAtlas: A large-scale, detailed-annotated, & multi-center dataset for efficient transfer learning and open algorithmic benchmarking. **Medical Image Analysis 2024.**

5. **Chen Y**, Zhou Z, Yuille A. Quality Sentinel: Estimating Label Quality and Errors in Medical Segmentation Datasets. **Arxiv 2024 preprint.**
6. Jiang H, **Chen Y**, Ding C, Liu L, Han X, Zhang X. Leveraging Noisy Labels of Nearest Neighbors for Label Correction and Sample Selection. **ICASSP 2024.**
7. **Chen Y**, Li J, Jiang H, et al. MetaLR: Meta-tuning of Learning Rates for Transfer Learning in Medical Imaging. **MICCAI 2023.**
8. Wu X, Hao Y, Sun K, **Chen Y**, Zhu F, Zhao R, Li H. Human Preference Score v2: A Solid Benchmark for Evaluating Human Preferences of Text-to-Image Synthesis. **Arxiv 2023 preprint.**
9. **Chen Y**, Liu L, Ding C. X-IQE: eXplainable Image Quality Evaluation for Text-to-Image Generation with Visual Large Language Models. **Arxiv 2023 preprint.**
10. Hu Q, **Chen Y**, Yuille A, Zhou Z. Label-Free Liver Tumor Segmentation. **CVPR 2023.**
11. **Chen Y**, Yuille A, Zhou Z. Which Layer is Learning Faster? A Systematic Exploration of Layer-wise Convergence Rate for Deep Neural Networks. **ICLR 2023.**
12. **Chen Y**, Li J, Liu L, Ding C. Rethinking Two Consensus of the Transferability in Deep Learning. **Arxiv 2023 preprint.**
13. **Chen Y**, Zhang C, Liu L, Ding C. Generating and Weighting Semantically Consistent Sample Pairs for Ultrasound Contrastive Learning. **IEEE Transactions on Medical Imaging.**
14. Hu Q, Xiao J, **Chen Y**, Sun S, Chen JN, Yuille A, Zhou Z. Synthetic Tumors Make AI Segment Tumors Better. **NeurIPS 2022 Workshop.**
15. Zhang C, **Chen Y**, Liu L, et al. HiCo: Hierarchical Contrastive Learning for Ultrasound Video Model Pretraining. **ACCV 2022.**
16. **Chen Y**, Zhang C, Liu L, et al. Uscl: Pretraining deep ultrasound image diagnosis model through video contrastive representation learning. **MICCAI 2021 (oral).**
17. Wan F, Wu Y, Qian X, **Chen Y**, Fu Y. When person re-identification meets changing clothes. **CVPR 2020 Workshop.**

PRESENTATIONS

Learning to Predict Multi-time Point Progressions (Poster) Daejon, Korea
The 28th Int. Conf. on Medical Image Computing and Computer Assisted Intervention *Sep. 2025*

Meta-Learning-Rate for Medical Transfer Learning (Poster) Vancouver, Canada
The 26th Int. Conf. on Medical Image Computing and Computer Assisted Intervention *Oct. 2023*

Layer-wise Convergence Rate of DNNs (Poster) Kigali, Rwanda
The 11th International Conference on Learning Representations *May. 2023*

Meta-Reweighting for Contrastive Learning (Poster) Shenzhen, China
The 2nd Doctoral & Postdoctoral Academic Forum at SRIBD *Aug. 2022*

Ultrasound Contrastive Learning (Oral) Online
The 24th Int. Conf. on Medical Image Computing and Computer Assisted Intervention *Sept. 2021*

ACDEMIC ACTIVITIES

Reviewer: NeurIPS, MICCAI, TASE, MedIA, IJCV, ICLR 2025

Reviewer: NeurIPS, ICLR, MICCAI, TASE 2024

Reviewer: MICCAI, ICML, TASE 2023

Teaching Assistant: AI in Medical Imaging 2022 Spring

Teaching Assistant: Statistical Inference 2021 Fall

RELEVANT COURSES

Core Courses	**Other Courses**
Machine Learning	Probabilistic Models in the Visual Cortex
Machine Learning: Deep Learning	Vision as Bayesian Inference
Intro Algorithms	Software Defined Networks
Computer Vision	Artificial Intelligence

SKILLS

Programming	Python, Bash, Matlab, R, JavaScript
Tech Stack	Linux, Git, Slurm, PyTorch, JAX, Numpy, Pandas, OpenCV, LaTeX
Language	English (TOEFL 107), Chinese (mother tongue)