

Fatemeh Ghezloo

Ph.D. Candidate at University of Washington

✉ fghezloo@uw.edu |  [fghezloo](https://github.com/fghezloo) |  fghezloo.github.io

Education

University of Washington

Sep. 2019 - June 2024 (Expected)

Ph.D. in Computer Science and Engineering

Multi-modal Generative AI, Computer Vision, Language Models, Machine Learning

Amirkabir University of Technology

Sep. 2014 - Oct. 2018

B.Sc. in Computer Engineering

Experiences

Graduate Research Assistant

September 2019 - Present

University of Washington

Seattle, WA, USA

- Developing a multi-agent system leveraging large multimodal models and Llama2 as tools to navigate medical images, collecting evidence towards a diagnosis.
- Developing a text-conditioned latent diffusion model for optimize personalization and inpainting.
- Curated Quilt-Instruct, largest image-text instruction-following dataset with visual grounding for histopathology, utilizing GPT-4 (107K QA pairs).
- Developed Quilt-LLaVA, an instruction-tuned vision-language AI-based agent capable of complex reasoning for histopathology by training a LLaVA model using Quilt-Instruct.
- Curated Quilt-1M, the largest vision-language dataset for histopathology containing more than one million image-text pairs from YouTube videos utilizing LLMs (GPT3.5), ASR models (Whisper) and histopathology classifiers.
- Developed QuiltNet, a state-of-the-art model utilizing CLIP for zero-shot and few-shot classification and cross-modal image-text retrieval for histopathology.
- Developed a novel region of interest (ROI) detection pipeline by leveraging a U-Net neural network architecture to assist pathologists in their diagnosis.
- Performed statistical analysis and applied classical machine learning methods to investigate the correlation between pathologists' viewing behaviors and diagnoses accuracy.

Computer Vision and AI Research Intern

June - September 2022

Zippin

San Francisco, CA, USA

- Attained 89% accuracy in automating the cart verification process of a checkout-free store platform serving millions of shoppers, significantly minimizing reliance on manual audits.
- Developed a multimodal model by fusing video and image features from MoViNets and EfficientNet.
- Improved model architecture by integrating squeeze and excitation and Gaussian hard attention maps.
- Curated a large-scale video dataset from Zippin stores' cameras containing 5M datapoints.

Publications

Ghezloo F*, Ikezogwo WO*, Seyfioglu MS*, Krishna R, Shapiro L. "Quilt-LLaVA: Visual Instruction Tuning by Extracting Localized Narratives from Open-Source Histopathology Videos." **CVPR (2024)**

Ghezloo F, Chang OH, Knezevich S, Reisch LM, Shapiro LG, Elmore JG. "Robust ROI Detection in Whole Slide Images guided by Pathologists' Viewing Patterns." Submitted to Journal of Digital Imaging (2024)

Ghezloo F*, Ikezogwo WO*, Seyfioglu MS*, Geva D, Mohammed F, Anand P, Krishna R, Shapiro L. "Quilt-1M: One Million Image-Text Pairs for Histopathology." **NeurIPS (2023) Oral. (2.15% oral acceptance rate)**

Nofallah S, Wu W, Liu K, **Ghezloo F**, Elmore JG, Shapiro LG. "Automated analysis of whole slide digital skin biopsy images." Journal of Pathology Informatics (2022)

Ghezloo F, Wang PC, Kerr KF, Brunyé TT, Drew T, Chang OH, Reisch LM, Shapiro LG, Elmore JG. "An analysis of pathologists' viewing processes as they diagnose whole slide digital images." Journal of Pathology Informatics (2022)

Kamkar S, **Ghezloo F**, Moghaddam HA, Borji A, Lashgari R. "Multiple-target tracking in human and machine vision." PLoS computational biology (2020)

Skills

Programming skills: Python, C, C++, Java, MATLAB, SQL, HTML/CSS

Machine Learning Tools: PyTorch, TensorFlow, OpenCV, scikit-learn, NumPy, pandas, matplotlib

Developer Tools: Jupyter Notebooks, Git, Google Cloud Platform, VS Code, Azure

AI/CV Methods: Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Transformers and Attention Mechanism, Latent Diffusion, GAN, Autoencoders (AEs and VAEs), Large Language Models (LLMs), Large Multimodal Models (LMMs), Vision-Language Models (VLMs), Multiple Instance Learning (MIL).

Projects

Data Augmentation Using CycleGAN

- Trained CycleGAN, an image-to-image translation model, to augment and balance the training set of the FER2013, a facial expression recognition dataset.
- Improved the accuracy of facial expression recognition for 'disgust' and 'anger' expressions by 10% by effectively leveraging the augmented data.

AmbigQA: A Baseline Model for Ambiguous Question Answering

- Reimplemented and adjusted the AmbigQA model to be compatible with limited processing resources, achieving an F1 score (39.58) close to the original paper's (39.7).
- Conducted several ablation studies on hyper-parameters impact on the inference time.
- Evaluated the model's performance on an additional QA dataset.

Interactive Data Visualization webpage

- Designed and built a webpage for an interactive visualization tool to investigate how pathologists view a skin biopsy image.
- Developed a MIL model to classify skin cancer and visualize attention maps using Grad-CAM.

Real-Time Facial Expression Recognition

- Designed and developed a deep convolutional neural network for facial expression recognition.
 - Enabled real-time face detection and facial expression recognition on computer's camera input.
-

Teaching

| | |
|--|------------|
| CSE455 Computer Vision | 2021, 2024 |
| CSE344 Introduction to Data Management | 2020, 2023 |
| CSE473 Introduction to Artificial Intelligence | 2020 |
| CSEP501/CSE401 Compiler Construction | 2019, 2020 |

Awards

| | |
|---|------|
| Microsoft Accelerate Foundation Models Research Program | 2023 |
| NeurIPS 2023 Scholar Award | 2023 |
| University of Washington Graduate Student Conference Presentation Award | 2023 |