Chun Yang

Los Angeles, CA ✉ chunyang.dev@gmail.com ☐ 213-284-1509 in/tom-yang-profile/ https://cy3021561.github.io/personal-website/

SUMMARY

Software engineer with 3+ years of experience, working on machine learning projects. Skilled in deep learning, generative AI, NLP, LLM, computer vision, and predictive modeling with proficiency in deriving insights from big data and creating data-driven models.

SKILLS

Full Stack: HTML, CSS, Java, JavaScript, Bootstrap, Angular, SwiftUI, Node.js, EJS, jQuery, Flask, SQL, Git

ML : Python, MLOps, Generative AI, Prompt Engineering, NLP, LLM, Pandas, PyTorch, TensorFlow, HuggingFace, LangChain, GCP, Kubernetes

3D Programming: C++, Unreal Engine

WORK EXPERIENCE

Software Engineer
Universal Global Scientific Industrial   **November 2018 - May 2022, Taipei, TW**

· Engineered AI-based defect detection systems, utilizing YOLO for targeted object detection and VGG for image classification within quality control processes. Achieved an 80% reduction in false alarms, significantly enhancing operational efficiency.
· Conducted real-time analysis of component usage from 200+ machines, fanout to 6 national sites and 50 production lines.
· Collaborated with international teams in Portland, Mexico, and China to develop over 20 software modules for various machines.

EDUCATION

Master of Science in Computer Science
University of Southern California · Los Angeles, CA · 2024 · 3.95 / 4.0

Master of Science in Computer Science
National Chiao Tung University · Hsinchu, TW · 2018 · 4.03 / 4.3

PROJECTS

Deepfake video detection using multimodal middle fusion
· Developed the 2D3MF model, a novel deepfake detection framework utilizing multi-modal middle fusion to synergistically analyze audio and visual data for detecting inconsistencies in manipulated media.
· Demonstrated superior performance on multiple benchmarks, achieving state-of-the-art results with over 95% accuracy on leading deepfake datasets such as DFDC and FaceForensics++. Rigorously tested the model across various conditions to ensure robustness and reliability.

Martian Terrain Image Classification: Enhancing Planetary Diagnostic Tools with CNN and Transfer Learning
· Built a CNN-based classifier project for identifying frost in Martian images, involving image augmentation and data preprocessing of over 119,920 tiles.
· Employed transfer learning with models such as EfficientNetB0 and ResNet50, achieving a notable 87.9% F1 score. Conducted a comprehensive comparative analysis of model performances.

Sentiment Analysis in Mental Health: Detecting Hidden Signs of Anxiety and Depression with NLP
· Drove the advancement of the NLP-MentalHealth-Insight project, leveraging advanced NLP methods to uncover concealed indicators of anxiety and depression within the text, building upon the research foundation set by Wolohan et al. (2018).
· Engineered and benchmarked a suite of machine learning algorithms, such as Perceptron, SVM, and BERT, to analyze sentiment and classify emotions in mental health-related text from online sources, including Reddit.

PUBLICATIONS

IF-NET: An Illumination-invariant Feature Network
IEEE International Conference on Robotics and Automation (ICRA 2020)
· Developed IF-Net, a novel approach in descriptor learning, focusing on overcoming the challenges of illumination variations in computer vision applications, leading to unprecedented improvements in image stitching, retrieval, and visual localization.
· Implemented advanced techniques including dataset scheduling, separation training, ROI loss, and hard-positive mining, achieving state-of-the-art results on patch-matching benchmarks and top localization accuracy in real-world conditions with significant illumination changes.

CERTIFICATIONS

Machine Learning Engineering for Production (MLOps)
DeepLearning.AI
· Design an ML production system end-to-end: project scoping, data needs, modeling strategies, and deployment requirements.
· Establish a model baseline, address concept drift, and prototype how to develop, deploy, and continuously improve a productionized ML application.
· Build data pipelines by gathering, cleaning, and validating datasets. Establish data lifecycle by using data lineage and provenance metadata tools.