Inseong Han

Email | LinkedIn | GitHub | Palo Alto, Stanford, CA

Data Scientist with over 3 years of relevant experience driven by a deep love for problem-solving.

Education

Master of Science in Data Science — ***July 2023 – Now***
University of San Francisco — *San Francisco, CA*

Bachelor of Science in Computer Engineering — ***Mar 2014 – Aug 2020***
Seoul National University — *Seoul, South Korea*
Thesis: Image-based synchronization of multiple videos — *Advisor: Gunhee Kim*

Professional Experience

Data Scientist, Intern — ***Sep 2021 – Now***
LexisNexis — *San Francisco, United States*

- Managed a Scala codebase for data pipelining while facilitating data analysis of 2.4 million entities with 11 million documents (+50GB) and effective communication within the team.
- Developed a semantic search engine utilizing relational database and advanced prompt-engineering techniques, notably leveraging RAG (Retriever Augmented Generation) and React prompt engineering methods.
- Enhanced search engine capabilities to understand diverse user queries in natural language, outperforming prior lexical methods.
- Collaborated with a cross-functional team of 20+ members, including the data science team lead and director

Machine Learning Engineer — ***Sep 2021 – May 2022***
Rapport Labs — *Seoul, South Korea*

- Built a recommender system for the fashion retail application with implicit feedback data from 300k users on 100k items.
- Eliminated the need for daily manual annotation of item ranks, relieving domain experts from this time-consuming task and saving hours of their valuable time.

Computer Vision Engineer — ***Aug 2019 – Sep 2021***
Bepro Company — *Seoul, South Korea*

- ***Camera calibration***
  - Dedicated to research on reducing distortions in stitched images.
  - Achieved a significant reduction of the distortion, enhancing overall product satisfaction for users.

- ***Synchronization of multiple videos taken from different perspectives***
  - Developed and implemented a synchronization module, enhancing the quality of generated stitched videos.
  - Achieved frame-level synchronization across multiple videos.
  - Saved costs by the improved synchronization process, reducing monthly expenses on cloud compute resources from $500 to $1500.

- ***Optimization of video stitching pipeline***
  - Worked on optimizing video stitching code to produce 30 FPS stitched videos through the pipeline.
  - Improved system efficiency of stitching with multiple 4K videos.
  - Obtained an additional 20 FPS throughput without additional resources.

Select Academic Projects

Implementation of Zhang's calibration method — ***Apr 2021 – Mar 2021***

- Applied Zhang's calibration method to estimate crucial camera parameters—such as intrinsic, extrinsic, and radial distortion coefficients—by processing corner points identified in chessboard images.
- Developed Python scripts for efficient homography estimation, camera parameter initialization, and optimization, enhancing the accuracy and speed of the calibration process.
- Validated the implementation's effectiveness by achieving comparable camera parameters on test images to those obtained using the OpenCV library, ensuring reliability and accuracy in calibration results for potential real-world applications.

Technical Skills

- Python, Large Language Model (LLM), LangChain, Hugging Face, PyTorch, Matplotlib, Pandas, Numpy, MongoDB, PostgreSQL, GitHub, Airflow, Docker, C++, OpenCV, Kubeflow, MLflow, GCP (Bigquery, GCS, GAR, GKE), AWS (S3, EC2), Nsight System (NVIDIA performance analysis tool), GPU programming (CUDA)