Hagai Raja Sinulingga

hagairaja.github.io | linkedin.com/in/hagai-raja-sinulingga

Education

Sejong University - Seoul, South Korea

Feb 2021 - Feb 2023

E-mail: hagairaja@gmail.com

Mobile: +44 7470813874

Master of Science, specialist in applied deep learning on computer vision

Institut Teknologi Bandung (ITB) - Bandung, Indonesia

Aug 2016 - Jul 2020

Bachelor of Science, majored in Computer Science

Selected Experiences

Lead AI Engineer (Remote Full-time)

Los Angeles, USA

MrScraper - Web Scraper AI Agent

Mar 2025 - Present

- Designed, developed, deployed, and maintained ScrapeGPT, an AI-powered browser agent serving as the core retail service of MrScraper
- Led browser and reverse engineering R&D by setting direction, managing talents, and actively contributing to complex, high-impact development tasks

Associate Research Engineer (Full-time)

Seoul, South Korea

TelePIX Co., Ltd. - Remote Sensing & Analytics Startup

Jan 2023 - Aug 2024

- Design ontology and data ingestion pipeline to load multi-modal information into a knowledge graph for South Korea's national defense
- Led team on multi-sensors building age prediction challenge by AI4EO from ESA and won #2
- Implemented distributed training pipeline to fine-tune EVA-02 (transformer-based foundation model for vision) leveraging 8 in-house H100 maximizing the usage of 640GB of GPU power

Graduate Researcher and Teaching Assistant (Full-time)

Seoul, South Korea

Imaging and Intelligent Systems Laboratory, Sejong University

Feb 2021 - Dec 2022

- Researched on efficient data collection for object detection and produced one published paper
- Built a welding anomaly detection app for our lab's client with explainable AI using Grad-CAM and SHAP values
- Prepared lecture material, including demos, tasks, and study cases for Machine Learning course attended by 50+ undergrad students about traditional ML algorithms & AI lifecycle

Data Engineer (Remote Part-time)

Bandung, Indonesia

Riset.ai - Smart City Autonomous Tram

Dec 2021 - Dec 2022

- Developed data fusion pipeline of 12 sensors: LIDAR, RADAR, Camera, IMU, and GPS with total size of data ~4TB orchestrated in the cloud using Apache Airflow
- Organized data to comply with standard multi-modal data structure: NuScenes and KITTI

Selected Projects

VRepo - Viral Content Repository built for influencers

- Spearheaded prototype in one week: social media content crawler, VLM to extract key features from a viral content, web-based app using NodeJS for user to search through the repository
- Utilized Change Scene Detection model to minimize inference time and data stored

AudioAI - Desktop app to analyze your recordings

- Developed key features: audio recordings and media player, transcribe and diarization using Whisper, and LLM integration to develop automated report generation and QA with RAG
- Engineered prompt to generate clinical report from a long-recording telephone interview, saving doctor's time by 40-50%

Multi-modal Building Construction Year Prediction - Map Your City Challenge Winner by ESA Φ-lab

• Developed deep learning model to predict construction year of buildings in european cities based on Sentinel-2, Orthophoto, and some Street View images as input