Valikhujaev Yakhyokhuja

website: https://yakhyo.github.io
linkedin: https://www.linkedin.com/in/y-valikhujaev

Email : yakhyo9696@gmail.com
Mobile : (+82) 010-9363-7088

ML Software Engineer

Passionate ML/DL/CV/NLP Engineer with ≈ 4 years of industrial and 2+ years of academic experience delivering impactful solutions across various industries. Proficient in supervised, self-supervised, and transfer learning, with in-depth experience in OCR, object detection, segmentation, tracking, video recognition, and action classification. Skilled in developing and deploying machine learning models on AWS and GCP, building and optimizing pipelines, containerization, and collaborating with cross-functional teams to drive business growth.

Skills Summary

- **Programming Languages:** Python, C/C++, Java.
- **Database Management:** MySQL, PostgreSQL, PySpark.
- **ML:** Numpy, Scikit-learn, PyTorch, PyTorch Lightning, Tensorflow, Keras, HuggingFace, Transformers, RAG, LangChain.
- **MLOps:** Docker-compose, Dockerization, Kubeflow, MLFlow, Flask, Fast API, gRCP, TorchServe, Trition, TensorRT.
- **Development Tools:** Git/Github, Docker, CI & CD.
- **Cloud:** AWS EC2, GCP.
- **Main Competencies:** Object Detection, Object Tracking, OCR, Clustering, Re-Identification, Medical Imaging, Image Restoration & Enhancement, DeepFakes, Generative Models, Vision-Language Models, Large-Language Models, Natural Language Processing, Building End-to-End Pipelines, Deployment Pipelines, GCP Deployment.

Work Experience

- **ML Engineer** — Seoul, South Korea
  Aria Studios Co. Ltd — *March 2024 - Present*
  - **In-cabin Companion**: Developed an in-cabin assistant utilizing large language models (LLMs) to enhance the driving experience. The system integrates multiple models for user interaction, including face identification, gender and age estimation, and emotion recognition. These features facilitate personalized conversations with the driver. For speech-to-text functionality, the **whisper** model was employed. Deployed all services using Fast API on GCP and local machine.
  - **LG Ground 220**: Designed and implemented the AI backend for MusicStudio and DJingStudio, enabling features such as lyrics generation from user input, cover image creation, and music generation using OpenAI, StabilityAI, and MixAudio APIs. Leveraged **LangChain** and concurrent programming to optimize API calls and reduce latency. Can be seen here
  - **Data collection tool for Virtual Character**: Developed a user interface that simulates a virtual character with predefined characteristics. The data collection team interacts with this virtual character to gather additional data for fine-tuning a large language model (LLM). The tool is deployed on Google Cloud Platform (GCP) using FastAPI, with PostgreSQL used to store the collected data.
  - **GPT Fine-tuning**: Fine-tuned the GPT-3.5-turbo model on conversation data of a virtual character to build a custom API for the virtual assistant. Performed several data augmentations using the OpenAI API for paraphrasing conversations.
  - **LLMs Deployment**: Deployed a lightweight Phi-3 model for emotion detection from text on GCP using FastAPI and prepared a deployment container. The main objective was to detect users' emotions while interacting with a virtual character and responding accordingly.
  - **VLMs Deployment**: Developed an "eye" for a virtual character to see and understand its surroundings, enabling it to interact with users. Deployed the Phi-3-Vision model on GCP for this purpose.
  - **Face Parsing**: To improve face swapping performance, employed a Face Parsing model to segment the face. Enhanced model performance by modifying the feature extractor (backbone). Implementation can be found here, 1k downloads.
  - **Image enhancement & Face restoration & DeepFake** : Worked on image enhancement and face restoration to improve DeepFakes. Created a DeepFake video for KBS election process coverage. Can be seen here
- **ML Engineer** — Seoul, South Korea
  Pyler Co. Ltd — *July 2022 - September 2023*
  - **Video-based Visual Content Moderation**: Build a Video Moderation Pipeline which helps to flag inappropriate video contents using video recognition models, achieving over a 10% improvement of model accuracy.

- **Detection-based Visual Content Moderation**: Utilized segmentation and detection techniques to precisely detect unsuitable content for the brand safety. Implemented latest state-of-the-art models in terms of real-time speed and efficiency, improved the model precision and recall by around a 15% by conducting active learning techniques. Built end-to-end pipeline on Kubeflow for training & deployment.
- **Classification-based Visual Content Moderation**: Leveraging multi-label and multi-head classification techniques improved the precision by approximately a 20% while using self-supervised and supervised training approaches. This novel approach showcases the adaptability and efficacy of the model for hard samples. Became standard method for visual content moderation and enhanced performance on challenging samples.
- **Model Assisted Labeling**: Created a pipeline for doing inference on existing labeled data and doing inference on unlabeled data. Improved labeling quality and speed using **Active Learning**.

- **AI Research Engineer** — Seoul, South Korea
  D-Meta Co. Ltd — *November 2020 - July 2022*
  - **Slab text Recognition**: Developed and designed text detection and recognition model to efficiently recognize handwritten texts on slab metals using Spatial Transformer Networks and Sequential modeling. Built a whole pipeline from data pre-processing to training and evaluation of the model. Achieved over a 90% accuracy, by integrating state-of-the-art detection and recognition models for scene text images.
  - **Automatic Number Plate Recognition**: Designed and developed ANPR model to accurately detect and recognize number plates. Leveraging active learning and synthetic image generation techniques improved the precision and recall by around a 15%.
  - **Car Damage Detection**: Built lightweight damage detection model and deployed it on Android device using torchscript. Improved the precision of the model by around a 10% by tuning the model parameters.

Research Experience

- **Research Assistant** — Seongnam, South Korea
  AI and SC Lab — *Sep 2018 - Nov 2020*
  - **Computer Vision based Fire and Smoke Detection**: Designed and implementation of the dilated CNN architecture for improved feature extraction and recognition in images/videos. Carefully tuning and optimizing the model, achieved a high level of accuracy in fire and smoke detection, reducing false positives and having 1.5x faster inference speed compared to the fastest counterpart.
  - **Model Optimization for Edge Devices**: Improved the FPS on Edge device (Raspberry PI 2) by using hyper-parameter tuning and quantization for detection model.

Please visit https://www.github.com/yakhyo to see more implementations of different ML models.

Education

- **Gachon University** — Seongnam, South Korea
  MSc in Computer Engineering; advised by Prof. Young Im Cho; CGPA: 4.0/4.5 — *Sep 2018 - Feb 2021*
- **Tashkent University of Information Technologies** — Tashkent, Uzbekistan
  BSc in Computer Engineering; CGPA(%): 85/100 — *Sep 2014 - June 2018*

Publications

- **Valikhujaev Y**, Abdusalomov A, Cho YI. Automatic Fire and Smoke Detection Method for Surveillance Systems Based on Dilated CNNs. Atmosphere, **IF 2.9**. 2020; 11(11):1241. https://doi.org/10.3390/atmos11111241.
- Muksimova SH, **Valikhujaev Y**, Cho YI. Automatic Fire and Smoke Detection System for Open Street CCTV Systems in Smart City Platforms. Korean Society of Information Scientists and Engineers, 412-414 pages, Domestic Conference.

Honors

Best paper award from Fire Investigation Society of Korea (FISK); (Domestic Conference, 2020)

Best presentation award from ISIS2019 & ICBAKE2019; (Domestic Conference, 2019)

Languages

English: Full Professional Proficiency (C1 Advanced, IELTS band 7);

Korean: Limited Working Proficiency (B1 Pre-Intermediate);

Russian: Limited Working Proficiency;

Uzbek: Native Proficiency;

Last Updated: September 19, 2024