Abhishek Kumar

Gurgaon, Haryana | +91 9962645223 | LinkedIn | GitHub | Portfolio | Email

WORK EXPERIENCE

IAD (SemaFor SRI-UB team) **Buffalo, NY**
Machine Learning Research Engineer *Sept 2022 - Present*

- Developed a system for face morph generation using spherical interpolation on generated images (Diffusion).
- Generated high-resolution morphed datasets and annotated facial sections using LabelImg for morph detection.
- Designed a linear model over CLIP features to analyze generated images and predict aesthetic scores.
- Implemented JPEG compression analysis, achieving a 94% accuracy in classifying manipulated images.
- Created a module for localizing manipulated regions in images using noise and edge-level features.
- Utilized PyTorch parallel training to reduce training time by 37% for a 1.3 million parameter model.
- Trained YOLO-v8 model with 0.69 F1 score on a custom dataset to detect small objects in manipulated regions.
- Containerized the application using Docker and established an end-to-end CI/CD pipeline using Github Actions.
- Developed a Python tool for controlled entity replacements in long texts using SpaCy and NLTK.
- Detected inconsistencies in news articles with entity mismatch and contradiction detection algorithms.
- Built a retrieval-based chatbot on a self-curated Reddit dataset, leveraging Solr for text indexing and searching.
- Implemented embedding based query matching for ranking and co-reference resolution to maintain contexts.

Infosys **Bhubaneswar, OD**
Software Engineer - Data *Feb 2018 – Aug 2020*

- Developed employee task delegation system to manage production workflow data using Python and React.
- Built REST API to parse delegation data from JSON files received in hourly batches with Flask and Python.
- Designed scripts to extract, parse and transfer millions of parts data across 12 pricing interfaces using PLSQL.
- Reduced batch transfer failure rate by 20% through automation and monitoring of batch jobs using RPA.
- Developed pricing page displaying available dependent and independent part prices from the database in Java.
- Managed deployment and maintenance pipelines for system availability and reliability using Jenkins and Git.

SKILLS

Languages, Database: Python, C++, Java, SQL, MySQL, MongoDB, Solr, Elasticsearch, Redis

ML Libraries: PyTorch, TensorFlow, Keras, OpenCV, NumPy, Pandas, Matplotlib, Scikit-Learn, NLTK, SpaCy

Cloud, DevOps : AWS, GCP, Git, GitHub Actions, Docker, Kubernetes, Airflow, Spark

EDUCATION

University at Buffalo **Buffalo, NY**
Masters in Computer Science and Engineering, Thesis *Sept 2021 - Aug 2023*

- 3.72/4.0, Machine Learning, Computer Vision, Deep Learning, Distributed Systems, Algorithm Design etc.

PROJECTS

Ear Hair-Cell Detection and Counting, (Python, Template Matching, Non-max Suppression, Clustering)

- Detected and extracted damaged inner and outer ear hair cells for measuring deafness in animals.
- Utilized supervised and unsupervised machine learning algorithms for successful cell detection.

Player Re-Identification, (PyTorch, C++, transfer learning, ResNet, OpenPose, Bilinear Pooling, Triplet Loss)

- Improved player re-identification score in broadcast soccer videos by 6%. Paper
- Utilized dual branch network with appearance and body part features extracted by ResNet and OpenPose.

Temporal Action Spotting, (PyTorch, Transformers, NumPy, Boundary Regression, Action Classification)

- Achieved 5th rank in SoccerNet competition among 50+ participants for spotting player actions in videos.
- Classified 17 action classes and identified temporal boundaries with 52% mAP using a Transformer model with multiscale RGB features.