Dishant Padalia

(413)-472-0475 dpadalia@umass.edu dishant-padalia dishant26 dishantpadalia.netlify.app

Education

University of Massachusetts Amherst **Sep 2023 - May 2025**
Master of Science in Computer Science GPA: 4.0/4.0
Relevant Courses: Machine Learning, Neural Networks, Advanced NLP, Systems for Data Science

K. J. Somaiya College of Engineering **Aug 2019 - May 2023**
Bachelor of Technology in Electronics & Telecommunication Engineering CGPA: 9.18/10
Relevant Courses: Algorithms, Data Structures, Artificial Intelligence, Big Data Analytics

Technical Skills

Languages: Python (NumPy, Pandas, SpaCy, Keras, NLTK, OpenCV), C++, Java, R, SQL, MATLAB
Frameworks: TensorFlow, PyTorch, Scikit-Learn, Flask, Django, Flutter, Spark, Hadoop, FastAPI
Areas of Interest: Machine Learning, Software Engineering, Large Language Models (LLMs), Data Science

Experience

Meta **Feb 2024 – Present**
Graduate Student Researcher

- Evaluated biases in **text-to-image generative models** like **Stable Diffusion** and **DALLE**, focusing on decoupling object and background stereotypes to enhance realism and diversity in generated images.
- **Designed metrics** such as **Realism, Coverage** and **CLIP Score** to evaluate the object and regional biases and stereotype differences between generated and real images.
- Worked on methods to assess intersectional biases in AI-generated images, with a particular focus on examining the impact of demographic associations on geographic disparities in object representation.

Indian Institute of Technology (IIT) Bombay **Apr 2022 – Jun 2023**
Machine Learning Engineer

- Designed a multimodal **Vision Transformer** model for multilingual handwritten document recognition, leveraging **PyTorch** and **HuggingFace**, resulting in a **40%** improvement in the recognition rate.
- **Engineered** a synthetic dataset generation algorithm, enhancing **domain adaptation & model generalization**, boosting performance by **50%** across six benchmark OCR datasets.
- **Integrated and fine-tuned** LayoutLM and TableNet models to optimize document understanding and table detection, achieving a **30% increase in layout and text recognition efficiency**.
- Streamlined a **semi-supervised data programming** pipeline to programmatically label training data using user-defined labeling functions and rules, eliminating manual labeling.
- **Authored** a paper introducing a **Priority-Based Suppression technique** on a combination of layout models for digitization of multilingual scanned documents.

Vasundharaa Geo Technologies Pvt. Ltd. **Jun 2021 – Dec 2021**
Machine Learning Intern

- **Headed** a state government project focused on automating **license plate detection** of non-helmeted motorcycle riders, which resulted in a **25% reduction in traffic law violations** at over 10 signal junctions in Pune.
- Implemented a **three-stage system** incorporating the **YOLO model** for motorcycle, helmet, and license plate detection, achieving a **Mean Average Precision score of 99%.**
- **Integrated** the model into the hardware system using **TensorFlow Lite** and **Google Cloud**, yielding a **70% reduction in real-time response time**.

Projects

Smart Vision Glasses for the Blind and Visually Impaired | *Python, TensorFlow, Flutter, C++, Arduino*

- Designed user-friendly and affordable smart glasses for the visually impaired, featuring navigation assistance, currency recognition, and text reading through a **Flutter-based application**.
- Implemented navigation and object avoidance using **Google Maps Navigation API** for routing, **MiDaS** for depth mapping, and **SSD MobileNet** for object detection, reducing the **Object Hit Rate by 75%**.
- Engineered a real-time text reading module using **Google Cloud Vision OCR** and **Text-to-Speech AI**, providing an seamless audio-based reading experience.
- Created a currency recognition feature with the **EfficientDet-Lite4** model, trained for detection and classification of banknotes, facilitating easier financial transactions for users.

Semantic Segmentation for Anomaly Detection in Dental X-Rays | *PyTorch, HuggingFace, OpenCV*

- Built a novel soft **attention-based U-Net architecture** for teeth **segmentation** from dental X-rays, achieving a **Dice Coefficient of 93%** and an **Intersection Over Union (IoU) of 87%**.
- Evaluated seven **augmentation techniques**, including rotation, skewing, and color manipulation, leading to a **35% increase** in the Dice Coefficient on real-world clinical dental X-rays.
- **Presented and published** a research **paper** at the IEEE ICAST 2022 Conference.

Audio and Signal Analysis for Dysarthria Detection | *TensorFlow, Librosa, MATLAB*

- Designed a **hybrid CNN-LSTM model** for detecting Dysarthria, a speech disorder, from patient audio recordings and achieved an accuracy of **99%**.
- Employed **Librosa** and **MATLAB** to extract features from audio signals, utilizing methods like Pre-emphasis, Fast Fourier Transform, and Mel-Frequency Cepstral Coefficients.
- **Trained and benchmarked** multiple models, including CNN-LSTM, CNN-SVM, Random Forest, Decision Tree, and XGBoost, resulting in a research **paper** presented at an IEEE Conference.

Publications

- Vora, K., **Padalia, D.**, Mehta, D. and Sharma, D., 2022, December. Hybrid CNN-LSTM network to detect Dysarthria using Mel-Frequency Cepstral Coefficients. In 2022 5th International Conference on Advances in Science and Technology (ICAST) (pp. 615-621). IEEE.
- **Padalia, D.**, Vora, K. and Sharma, D., 2022, December. An Attention U-Net for Semantic Segmentation of Dental Panoramic X-ray images. In 2022 5th International Conference on Advances in Science and Technology (ICAST) (pp. 491-496). IEEE.
- **Padalia, D.**, Vora, K., Mehta, D. and Mehendale, N., 2022. EEF-Net: an enhanced efficientnet for breast tumor classification in mammograms. Available at SSRN 4220435.
- Mehta, D., **Padalia, D.**, Vora, K. and Mehendale, N., 2022, December. MRI image denoising using U-Net and Image Processing Techniques. In 2022 5th International Conference on Advances in Science and Technology (ICAST) (pp. 306-313). IEEE.
- Vora, K. and **Padalia, D.**, 2022. An Ensemble of Convolutional Neural Networks to Detect Foliar Diseases in Apple Plants. arXiv e-prints, pp.arXiv-2210.
- **Padalia, D.**, Mazumdar, A. and Singh, B., 2022. A CNN-LSTM Combination Network for Cataract Detection using Eye Fundus Images. arXiv preprint arXiv:2210.16093.
- **Padalia, D.**, Mehta, D., Metha, K., Bhat, A., Gupta, P. and Mehendale, N., 2022. DeepPET-3D: A Deep Learning based 3D-CNN model for diagnosis of Alzheimer's Disease using 18-FDG-PET. Available at SSRN.
- **Padalia, D.**, 2022. Detection and Number Plate Recognition of Non-Helmeted Motorcyclists using YOLO.

Achievements and Extracurriculars

- **Instructed** a course on Udemy, ***'A Hands-On Introduction to OCR'***, with over **800** enrollments.
- **Won 1st place** and ***'Best Open Innovation'*** award among 400+ teams at Electrothon Hackathon '21.