Prateksha Udhayanan

pratekshau@gmail.com | +91-7760689583 | GitHub | LinkedIn | Website

Research Interests

My research interests primarily revolve around **computer vision**, particularly in the field of **controllable synthesis of visual content** such as **images, videos, and graphic designs**. I am especially keen on utilizing **multi-modal constraints** to effectively guide the generation.

Education

- **Integrated Master of Technology in Computer Science (BTech+MTech)** — Bangalore, India
  International Institute of Information Technology, Bangalore; CGPA: 3.76/4.00 — *2017 – 2022*
  - **Dean's Merit List:** Part of the Dean's Merit List - 2018, 2020, 2021, 2022 for academic excellence.
  - **Teaching assistant:** ESS 201 Programming II (C++, Java) (August 2020 - December 2020)
    ESS 112 Programming in Python (November 2020 - March 2021)
    AI 511 Machine Learning (September 2021 - December 2021)

Publications and Preprints

- [AAAI 2024] Koustava Goswami, Srikrishna Karanam, Joseph K J, *Prateksha Udhayanan* and Balaji Vasan Srinivasan. **Contextual Prompt Learning for Vision-Language Understanding** - Accepted at AAAI 2024 [LINK]
- [WACV 2024] KJ Joseph, *Prateksha Udhayanan*, Tripti Shukla, Aishwarya Agarwal, Srikrishna Karanam, Koustava Goswami, and Balaji Vasan Srinivasan. **Iterative Multi-granular Image Editing using Diffusion Models** - Accepted at WACV 2024 [LINK]
- [WACV 2023] *Prateksha Udhayanan*, Suryateja BV, Parth Laturia, Dev Chauhan, Darshan Khandelwal, Stefano Petrangeli, and Balaji Vasan Srinivasan. **Recipe2Video: Synthesizing Personalized Videos from Recipe Texts**. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 2268-2277). 2023 [LINK]
- [SN Computer Science] Chinchu Thomas, *Prateksha Udhayanan*, Ayush Yadav, and Dinesh Babu Jayagopi. **Multimodal Unsupervised Domain Adaptation for Predicting Speaker Characteristics from Video** - Accepted at SN Computer Science journal
- [AI-ML Systems 2021] Arjun Verma, *Prateksha Udhayanan*, Rahul Murali Shankar, Nikhil KN, and Sujit Kumar Chakrabarti. **Source-Code Similarity Measurement: Syntax Tree Fingerprinting for Automated Evaluation**, The First International Conference on AI-ML-Systems (pp. 1-7). 2021) [LINK]
- [Physica A: Statistical Mechanics and its Applications] *Prateksha Udhayanan*, Swasti S Mishra, and Shrisha Rao. **Firm dynamics and employee performance management in duopoly markets** Physica A: Statistical Mechanics and its Applications. (583, p.126298). 2021 [LINK]
- *Prateksha Udhayanan*, Srikrishna Karanam, and Balaji Vasan Srinivasan. **Learning with Multi-modal Gradient Attention for Explainable Composed Image Retrieval** - [arXiv:2308.16649]

Patents

- *Prateksha Udhayanan*, Srikrishna Karanam, and Balaji Vasan Srinivasan. **Text-Conditioned Visual Attention for Multimodal Machine Learning Models** (US Patent App. 18/351,211)
- Koustava Goswami, Srikrishna Karanam, Joseph KJ, *Prateksha Udhayanan*, and Balaji Vasan Srinivasan. **Generating Text Prompts for Digital Images Utilizing Vision-Language Models and Contextual Prompt Learning** (US Patent App. 18/342,954)
- Suryateja BV, *Prateksha Udhayanan*, Dev Chauhan, Parth Laturia, Darshan Khandelwal, Stefano Petrangeli, and Balaji Vasan Srinivasan. **Auto-Generating Video to Illustrate a Procedural Document** (US Patent App. 17/661,614)

Experience

- **Adobe Research, Bangalore, India**
  Research Associate — *July 2022 – Present*
  - Part of the Collaborative Creativity team, working on a range of projects focused on computer vision and multi-modal content understanding. Submitted papers, filed patents and developed research technologies that were successfully integrated into products.
  - Key project themes: image retrieval, editing, and generation, automated creation and editing of graphic designs, layout and color optimizations for graphic designs, and cinemagraph generation.
- **Slice, India - Fintech Startup** — Mentors: Saksham Agarwal, Khushal Bokadey
  Software Engineering Intern — *January 2022 - June 2022*
  - Contributed to designing and building a Microservice for credit-based non-card transactions in a three-member team.
  - Implemented code in Spring Boot for Paytm and bank transfers; optimized the algorithm for computing transfer charges.
  - Developed High-Level Design (HLD) and Low-Level Design (LLD) considering database requirements and schema design.

- Setup infrastructure pipelines, automated build, test, and deployment workflows across different environments.
- **Internal Reconciliation Service** - Led the development of an internal reconciliation service, designed to compare records maintained by Slice with transaction details provided by external entities.

- **Adobe Research, Bangalore, India** — Mentor: Dr. Balaji Vasan Srinivasan
  *Research Intern [**Published in WACV 2023**]* — *May 2021 – August 2021*
  - Built a novel end-to-end deep-learning based architecture to convert instructional documents into multimodal videos.
  - Proposed re-ranking methods to retrieve and select the optimal combination of multimodal visuals aligned with constraints.
  - Used Viterbi algorithm-based optimization for multi-modal frame selection, resulting in a coherent video composition.
  - Established evaluation metrics based on cognitive models of procedural text understanding.
- **Anveshan** — Mentor: Ayush Goyal
  Software Engineering Intern at Anveshan — *March 2020 – April 2020*
  - Collaborated closely with the startup's tech team to develop an application for operational supply chain management, to help eliminate middlemen from the supply chain and promote fair trade for farmers.
  - Co-led the design and development of the back-end for an Inventory Management app using Django.
  - Implemented Optical Character Recognition techniques to enable automated product quality check - verifying MRP and expiry date.
- **Siemens Technology and Services Pvt. Ltd. Bangalore** — Mentors: Varghese Alex, Vinay Sudhakaran
  Research Intern — *May 2020 – July 2020*
  - Implemented an LSTM-based hierarchical model in TensorFlow for Anomaly Detection in procedural videos.
  - The model is trained to learn the sequential structure of procedural activities from text domain and transfer it to visual domain which enables visual appearance learning and future prediction.
- **Multimodal Perception Lab (MPL) - IIITB** — Mentor: Dr. Dinesh Babu J
  Research Intern at MPL — *June 2019 – September 2019*
  - Indian Sign Language Synthesis - implemented 3D CNN models to estimate 3D hand joint coordinates from depth images.
  - Developed a pipeline to map the predicted coordinates to a virtual avatar for gesture generation.
  - Reduced the gesture generation time for each sign from 20 minutes to less than 2 minutes by automating the process.
- **Rails Girls Summer of Code Scholar 2018** — Mentors: Julia Nguyen, Camille Villa
  Software Engineering Intern at "if-me" organization — *July 2018 – Sept 2018*
  - Designed and developed a full stack feature for blocking and reporting users on the if-me web app using Ruby on Rails.
  - Improved the UI and built several accessible and modular ReactJS components and wrote tests in Enzyme and Jest.

Selected Projects

- **Design Co-pilot**
  Adobe Research | Being integrated to product
  - Implemented novel methodologies to generate graphic design variants based on color, layout and content.
  - Designed and built an end-to-end pipeline to edit and transform a design based on user prompt.
  - Implemented transformer-based masked field prediction approach for layout and color optimization in design.
  - Built on stable diffusion models to develop a latent-space iteration and manipulation technique for iterative image editing in design workflows.
- **Generating Cinemagraphs from Static Images**
  Adobe Research
  - Currently working on a method to animate periodic motions involving large displacements using a single image as input.
  - Helped in product integration of automatic animation for fluid elements such as water, smoke, and fire, from a single image.
- **Composed Image Retrieval**
  Adobe Research
  - Developed a novel multi-modal gradient attention computation mechanism for composed image retrieval.
  - Proposed an end-to-end learning scheme for vision-language siamese transformers with our attention-based loss function.
  - Improved the performance by $\approx 3\%$ on the *Recall*@1 metric and by $\approx 20\%$ on the $Recall_{subset}$@1 metric.
- **Firm Dynamics and Employee Performance Management in Duopoly Markets**
  Supervisor - Dr. Shrisha Rao | [Published in Physica A: Statistical Mechanics and its Applications 2021]
  - Modelled and studied the effects of different performance management strategies in a firm using agent-based simulations.
  - Studied the effects of Peter Principle on firms competing in Stackelberg games and Cournot games.
- **Multimodal Unsupervised Domain Adaptation for Predicting Speaker Characteristics from Videos**
  Supervisor - Dr. Dinesh Babu Jayagopi | [Accepted in SN Computer Science journal]
  - Built a multimodal unsupervised domain adaptation framework for predicting speaker characteristics from videos.
  - Used MAG-BERT to infuse multimodal data and trained the model in an adversarial setting with knowledge distillation.
  - Curated a dataset of videos obtained from informal online learning platforms and annotated it with speaker characteristics.
- **Syntax Tree Fingerprinting for Source-Code Similarity Measurement**
  Supervisor - Dr. Sujit Kumar Chakrabarti | [Published in AI-ML Systems 2021]
  - Developed a winnowing-based AST fingerprinting technique to compute structural similarity between two programs.
  - Built a dataset of 7 problems with 100 samples each. Developed evaluation rubrics for these problems to test the system.

- **Composite Hybrid Visualization Tool for Unbalanced Bipartite Graphs**
  Supervisor - Dr. Jaya Sreevalsan Nair | [Project Report]
  - Designed and built a visualization tool in D3.js that integrates matrix representation and radial graph layout.
  - Implemented algorithms for edge-rendering and node-ordering: force-directed edge bundling and Barycentric ordering.
- **A Java-Based Tool for Storing OWL Ontologies in Object Oriented Databases**
  Supervisor - Dr. Chandrashekar Ramanathan | [GitHub]
  - Implemented various static mapping and dynamic mapping rules to persist OWL ontologies in ObjectDB databases.
  - Built an inferencing engine based on forward chaining algorithm for inverse, transitive, and symmetric properties.
  - Integrated a querying tool that converts input queries from SPARQL to JPQL and retrieves results from the OODB.

Achievements

- **Winner of CodeHers Coding Challenge 2021:** One of the **top 15** winners out of 50,000+ women participants.
- **Rails Girls Summer of Code Scholar 2018:** One of the 7 full-time teams selected out of 200+ international teams.

Miscellaneous Projects

- **Unsupervised Visual Representation Learning:** Conducted a detailed study and analysis of unsupervised visual representation learning methods like MoCo, SimCLR and PIRL. Implemented a BOW-based model for visual words in PyTorch and conducted experiments over different datasets.
- **Malaria Parasite Detection from Blood Cell Images:** Compared the performance of random forest classifier to a convolutional neural network to predict if a given cell image is parasite infected or not.
- **Predicting Binary and Graded Lexical Entailment on Cross-Lingual word-pairs:** Combined the ideas of LEAR and CLEAR to transform the input word vector space into a semantically specialised cross-lingual vector space that emphasises on lexical entailment relation.
- **Fitness App to Count Reps:** Built a fitness web app that incorporates pose estimation models to automatically count the number of reps as the user exercises. Developed the app using NextJS and Firebase. Jenkins, Docker and Ansible were used for CI/CD and ELK stack was used for monitoring.
- **Employee Attrition Analysis:** Conducted data preprocessing and feature engineering on Kaggle's HR Analytics dataset. Implemented XGBoost to predict employee attrition.
- **Data Processing Toolkit:** Implemented a modified version of "Minimum Degree Ordering Algorithm" in C++ which was integrated in the project to build a small-scale tabular data processing toolkit.
- **Web Scraping Project:** Built a Python based scraper which uses the BeautifulSoup library and extracts information from college's learning management system. Configured as a cronjob that runs every half an hour to check for new updates and notifies the user via email.

Relevant Courses

- **Artificial Intelligence and Machine Learning:** Machine Learning, Math for Machine Learning, Natural Language Processing, Advanced Visual Recognition, Artificial Intelligence
- **Computer Science:** Programming Languages, Data Structures and Algorithms, Computer Networks, Signals and Systems, Database Systems, Automata Theory, Software Engineering, Operating Systems, Data Modeling, Discrete Mathematics, Data Visualization, Foundations of Cryptography, Computing on Private Data
- **Mathematics:** Real Analysis, Linear Algebra, Abstract Algebra, Calculus, Probability and Statistics

Programming Skills

- **Programming Languages:** Python, C, C++, Java, HTML, CSS, Django, Ruby on Rails
- **Libraries and Software:** PyTorch, TensorFlow, Keras, Scikit-Learn, Pandas, NumPy, OpenCV, Git, MySQL, LaTeX

Extra-curricular Activities

- **Online Asian Machine Learning School(OA^{M}LS), a part of Asian Conference on Machine Learning (ACML) 2021:** Was selected to participate and present a poster on my work - Firm Dynamics and Employee Performance Management in Duopoly Markets.
- **Founding Member, Lean In Chapter IIIT-B:** Mentored girls by conducting lectures and workshops on machine learning and open source. Members have been selected in GSoC, MLH fellowship, GSSoC and various other prestigious programs.
- **Curator, TEDxIIITBangalore 2019:** Managed the speaker selection process and worked with the speakers to curate their talks.
- **Bharatnatyam:** Trained at Khechara Academy for 12 years and secured 91.75% in the Junior Grade Exam.
- **Core Member, Dance Club-IIITB:** Choreographed several dance performances for Infin8, Sangam, the institute's foundation day, and many more.
- **Core Member, Alumni Committee, 2018-2019:** Organized the alumni meet - Sangam 2019 and coordinated alumni talks and workshops throughout the academic year.