Prateksha Udhayanan

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

Research Interests

My research interests primarily revolve around computer vision, with a strong emphasis on understanding multi-modal input forms to enhance image and video retrievals as well as generations.

Education

- **Integrated Masters of Technology in Computer Science (BTech+MTech)** — Bangalore, India
  International Institute of Information Technology, Bangalore; CGPA: 3.76/4.00 — *2017 – 2022*
  - **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

- [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]
- [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** - [Preprint]
- Koustava Goswami, Srikrishna Karanam, Joseph K J, Prateksha Udhayanan and Balaji Vasan Srinivasan. **Contextual Prompt Learning for Vision-Language Understanding** - Under review at AAAI 2024 [Preprint]

Patents

- *Prateksha Udhayanan*, Srikrishna Karanam, and Balaji Vasan Srinivasan. **Text-Conditioned Visual Attention for Multimodal Machine Learning Models** (US Patent App. 18/351,211)
- *Prateksha Udhayanan*, Dev Chauhan, Parth Laturia, Darshan Khandelwal, Suryateja BV, Stefano Petrangeli, and Balaji Vasan Srinivasan. **Auto-Generating Video to Illustrate a Procedural Document** (US Patent App. 17/661,614)
- 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)

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 techniques with multimodal inputs, image editing techniques, automated creation and editing of graphic designs, layout and color optimizations for graphic designs.
- **Slice, India - a fintech startup** — Mentors: Saksham Agarwal, Khushal Bokadey
  Software Engineering Intern — *January 2022 - June 2022*
  - Part of a three-member team responsible for designing and building a **Microservice for credit-based non-card transactions**
  - Implemented code in Spring Boot for Paytm wallet and bank transfers; optimized the computation algorithms for 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: Balaji Vasan Srinivasan
  *Research Intern [**Published in WACV 2023**]* — *May 2021 – August 2021*
  - Built a novel end-to-end deep-learning based architecture that converts instructional documents into multimodal illustrative videos.
  - Implemented modules for retrieval and selection of multimodal visuals based on novel ranking schemes, tailored to user needs.
  - Used a Viterbi-algorithm-based optimization technique for multi-modal frame selection, resulting in a coherent video composition.
  - Established evaluation metrics based on cognitive models of procedural text understanding.
- **Siemens Technology and Services Pvt. Ltd. Bangalore** — Mentor: Varghese Alex
  Research Intern — *May 2020 – July 2020*
  - Implemented a 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.

- **Anveshan** Mentor: Ayush Goyal
  Software Engineering Intern at Anveshan *March 2020 – April 2020*
  - Co-led the designing 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.
- **Multimodal Perception Lab (MPL) - IIITB** Mentor: Dr.Dinesh Babu J
  Research Intern at MPL *June 2019 – September 2019*
  - Translation system for Indian Sign Language - Estimated the coordinates of hand joints using 3D CNN models and mapped them to a human avatar. Reduced the gesture generation time for each sign from 20 minutes to less than 2 minutes.
- **Rails Girls Summer of Code Scholar'18** Mentor: Julia Nguyen, Camille Villa
  Software Engineering Intern at "if-me" organization *July 2018 – Sept 2018*
  - **One of the 7 full-time teams** selected out of 200+ international teams.
  - Designed and developed a feature for blocking and reporting users using Ruby on Rails.

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 filed prediction approach for layout and color optimization in design.
  - Developed a latent-space iteration technique to enable iterative editing of images in design workflows
- **Generating Cinemagraphs by Animating Elements from a Single Image Input**
  Adobe Research
  - Currently working on a method to animate periodic motions using a single image as input.
  - Helped integrate animation of fluid elements work into product.
- **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 that integrates our attention-based loss function.
  - Significantly improved the performance by $\approx 4\%$ on *Recall*@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]
  - Modeled 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.
- **Syntax Tree Fingerprinting for Automated Evaluation**
  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.
- **Multimodal Unsupervised Domain Adaptation for Predicting Speaker Characteristics from Video**
  Supervisor - Dr. Dinesh Babu Jayagopi
  - Built multimodal unsupervised domain adaptation framework to predict the persuasiveness and expertise of the speaker from videos.
  - Used knowledge distillation along with adversarial discriminative domain adaptationn.
  - Curated a dataset of videos obtained from informal online learning platform and annotated it with speaker characteristics
- **Unsupervised Visual Representation Learning**
  Supervisor - Dr. Dinesh Babu Jayagopi
  - 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.
- **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 different 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.
- **Fitness App to Count Reps**
  Supervisor: Dr. Thangaraju B | [GitHub] [Project Report]
  - Built a fitness web app that incorporates pose estimation models to automatically count the number of reps as the user exercises.
  - Developed using NextJS and Firebase. Jenkins, Docker and Ansible were used for CI/CD and ELK stack was used for monitoring.

Achievements

- **Winner of CodeHers Coding Challenge 2021:** One of the **top 15** winners out of 50,000+ women participants.
- **Dean's Merit List:** Part of the Dean's Merit List - 2018, 2020, 2021, 2022 for academic excellence.

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