AISHWARYA AGARWAL

+91-8077729866 | agarwal.aishwarya2013@gmail.com | aisagarw.github.io

INTERESTS

Computer Vision, Representation Learning in Resource-constrained Scenarios, Customization of Text-to-image Generation models (Stable Diffusion), Self-supervised Learning, Machine Learning

EDUCATION

- **Interdisciplinary Dual Degree Programme, Indian Institute of Technology Bombay** July 2017 - June 2022
  Bachelor of Technology in Electrical Engineering
  Master of Technology in Centre for Machine Intelligence and Data Science
  Cumulative GPA: **9.58/10.00**

PUBLICATIONS

[1] KJ Joseph, Prateksha Udhayanan, Tripti Shukla, **Aishwarya Agarwal**, Srikrishna Karanam, Koustava Goswami, Balaji Vasan Srinivasan. **Iterative Multi-granular Image Editing using Diffusion Models**. In *Winter Conference on Applications of Computer Vision* (WACV 2024). [LINK]

[2] **Aishwarya Agarwal**, Srikrishna Karanam, KJ Joseph, Apoorv Saxena, Koustava Goswami, Balaji Vasan Srinivasan. **A-STAR: Test-time Attention Segregation and Retention for Text-to-image Synthesis**. In *International Conference on Computer Vision* (ICCV 2023). [LINK]

[3] **Aishwarya Agarwal**, Anuj Srivastava, Inderjeet Nair, Swasti Shreya Mishra, Vineeth Dorna, Sharmila Reddy Nangi, Balaji Vasan Srinivasan. **SketchBuddy: Context-Aware Sketch Enrichment and Enhancement**. In *ACM Multimedia Systems Conference* (ACM MMSys 2023). [LINK]

[4] **Aishwarya Agarwal**, Srikrishna Karanam, Balaji Vasan Srinivasan, Biplab Banerjee. **Contrastive Learning of Semantic Concepts for Open-set Cross-domain Retrieval**. In *Winter Conference on Applications of Computer Vision* (WACV 2023). [LINK]

[5] **Aishwarya Agarwal**, Biplab Banerjee, Fabio Cuzzolin, Subhasis Chaudhuri. **Semantics-Driven Generative Replay for Few-Shot Class Incremental Learning**. In *ACM International Conference on Multimedia* (ACM MM 2022). [LINK]

[6] Hrituraj Singh, Anshul Nasery, Denil Mehta, **Aishwarya Agarwal**, Jatin Lamba, Balaji Vasan Srinivasan **Mimoqa: Multimodal input multimodal output question answering** In *North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (NAACL-HLT 2021). [LINK]

MANUSCRIPTS UNDER REVIEW

[1] **Aishwarya Agarwal**, Srikrishna Karanam, Tripti Shukla, Balaji Vasan Srinivasan. **An Image is Worth Multiple Words: Multi-attribute Inversion for Constrained Text-to-Image Synthesis**.

[2] **Aishwarya Agarwal**, Srikrishna Karanam, Balaji Vasan Srinivasan. **Learning with Difference Attention for Visually Grounded Self-supervised Representations**.

[3] **Aishwarya Agarwal**, Srikrishna Karanam, Balaji Vasan Srinivasan. **AlignIT: Enhancing Prompt Alignment in Customization of Text-to-Image Models**.

[4] Abhikhya Tripathy, **Aishwarya Agarwal**, Srikrishna Karanam, Balaji Vasan Srinivasan. **Improving Text Generation in Diffusion Models via Inference-Time Self-Attention Manipulation**.

PATENTS

[1] **Aishwarya Agarwal**, Srikrishna Karanam, Tripti Shukla, Balaji Vasan Srinivasan. **An Image is Worth Multiple Words: Multi-attribute Inversion for Constrained Text-to-Image Synthesis** [Accepted for Filing]

[2] **Aishwarya Agarwal**, Srikrishna Karanam, Balaji Vasan Srinivasan. **Visual Grounding of Self-supervised Representations for Machine Learning Models utilising Difference Attention** [Filed]

[3] **Aishwarya Agarwal**, Srikrishna Karanam, KJ Joseph, Apoorv Saxena, Koustava Goswami, Balaji Vasan Srinivasan. **Text-to-image Synthesis utilizing Diffusion Models with test-time Attention Segregation and Retention Optimization** [Filed]

[4] Hrituraj Singh, Jatin Lamba, Denil Pareshbhai Mehta, Balaji Vasan Srinivasan, Anshul Nasery, **Aishwarya Agarwal**. **Modality adaptive information retrieval** [Filed]

WORK EXPERIENCE

- **Research Associate - Adobe Inc. (Big Data Intelligence Lab)** Jul 2022 - ongoing
  Group: Collaborative Creativity *Bangalore, India*
  - Engaged in several projects involving Multimodal Design Understanding, Design Auto-generation, Customizing Text-to-image generation models (diffusion models), Brandifying Designs

- **Research Intern - Adobe Inc. (Big Data Intelligence Lab)** May 2021 - Aug 2021
  Topic: Context-Aware Sketch Enrichment and Enhancement *Sharmila Nangi, Balaji Vasan Srinivasan*
  - Worked on **Context-Aware Scene Enrichment** and **Enhancement** to aid the designers in their ideation phase
  - Proposed a novel **context-aware recommendation framework** for quicker illustrations in **sketch domain**
  - Explored deep learning architectures such as **RetinaNet** and **FRCNN** for object detection, **VAE** and **ResNet** for sketch representation, **Transformers** (Masked Region Predictor) to capture context, and **U-Net** for saliency map generation
  - Extended the proposed frameworks to **UI/UX designs** domain by giving recommendations for parallely enhancing **UI sketches** to a **lo-fi prototype** thereby allowing rapid iterations between the sketch domain and the prototype domain
  - Published in **ACM MMSys 2023**

- **Research Intern - Adobe Inc. (Big Data Intelligence Lab)** Apr 2020 - Jul 2020
  Topic: Multimodal Question Answering *Hrituraj Singh, Balaji Vasan Srinivasan*
  - Proposed a novel task - MIMOQA - **M**ultimodal **I**nput **M**ultimodal **O**utput **Q**uestion **A**nswering on documents
  - Developed a novel **transformer** based **multimodal question-answering** framework, **MExBERT**, that incorporates a **joint textual** and **visual attention** towards producing an output across **multiple modalities**
  - Curated a **multimodal dataset** for the proposed problem from publicly available unimodal datasets like MS-MARCO
  - Developed a **Flask** based web application to demonstrate the working of our **information retrieval** system
  - Published in **NAACL - HLT 2021** and filed a **patent**

KEY PROJECTS

- **Brandifying Generative AI Content Creation** April 2023 - Present
  Adobe Research
  - Developed a method for constraining text-to-image generation models (Stable Diffusion) on multiple attributes (e.g., color palette, style, layout and object semantics) of user-provided reference images
  - Proposed approach outperforms current sota methods and allows synthesizing novel generations while conditioning on different attributes in a disentangled fashion

- **Towards Improved Image-text Alignment in Diffusion Models** Dec 2022 - March 2023
  Adobe Research
  - Identified two key issues by analyzing the **cross attention representations** of the existing text-to-image diffusion models, **attention overlap** and **attention decay**, that lead to semantically inaccurate generations
  - Proposed two new loss functions called **attention segregation** loss and **attention retention** loss to explicitly address the identified issues during the **inference-time denoising process** without requiring any model retraining

- **Learning Unified Representations for Retrieval Workflows** July 2022 - Nov 2022
  Adobe Research
  - Propose a new gradient-based visual attention computation mechanism called **Visual Difference Attention** to highlight all the image salient regions SSL models attend to and hence evaluating the extend of visually groundedness of SSL models
  - Casted our proposed **Visual Difference Attention (VDA)** concept above as a fully differentiable operation, leading to a new learning objective called the **Differentiable Difference Attention (DiDA)** loss for training SSL models
  - Benchmarked the quantitative impact of the DiDA loss with multiple downstream tasks: instance segmentation, detection, and classification and reported a new state of the art with average AP improvements of **1.3**, **1.1**, and **0.8** respectively

- **Towards out-of-distribution Generalization** Jan 2021 - June 2022
  Prof. Biplab Banerjee *Dept. of CSRE, IIT Bombay*
  - **Stage I: Few-Shot Class Incremental Learning (FSCIL)**
    - Proposed a novel generative replay-based FSCIL framework, called ***Few-Shot Incremental Learning GAN* (FSIL-GAN)** which is the first attempt towards using the paradigm of **data-free memory replay** for FSCIL
    - Demonstrated that the traditional generative-replay paradigm of CIL is insufficient to solve FSCIL given less data
    - Ensured **feature discriminability** across different classes through a novel **semantic projection loss**
    - Tackled **mode-collapse** problem through a novel **anti-collapse regularizer** which penalizes the generation of similar type of synthesized data for different noise vectors
  - **Stage II: Open-set Cross-domain Retrieval** *(Collaboration with Adobe Research)*

- Proposed a new approach, **SCNNet**, for generalized **unseen-class** and **unseen-domain** image retrieval
- Implemented an end-to-end data-driven strategy for learning an **orthogonal set of local visual concepts** that can be used to generate local feature representations for any unseen-class image
- Used natural language (class labels) in a novel contrastive distance learning framework to ensure features of **same-class-different-domain** images are mapped close in the feature space

- **OCR & Machine Translation** Feb 2021 - April 2021
  Prof. Pushpak Bhattacharyya *Dept. of CSE, IIT Bombay*
  * Finetuned a **pretrained model mt5** for Hindi to English translation on IIT-B English-Hindi parallel corpus
  * Extended **transformer** architecture to include **layerwise coordination** between encoder and decoder layers
  * Performed comparative analysis between encoders like **RNN**, **LSTM** and **transformer** based on **BLEU** score

- **Multi-domain Learning — Visual Decathlon Challenge** Jan 2021 - April 2021
  Prof. Amit Sethi *Dept. of EE, IIT Bombay*
  * **Overview:** Simultaneously solving ten image classification tasks representative of very different visual domains
  * Performed extensive literature review on **feature transferability** and **sharing** between deep learning models
  * Worked towards developing modularization techniques with an aim to achieve reduction in number of trainable parameters while matching the accuracies of fully fine-tuned CNNs on new domains
  * Proposed a novel architecture by extending **parallel residual adapters** for multi-domain few shot learning

- **Flow-Based Image Abstraction** Oct 2020 - Dec 2020
  Prof. Suyash Awate *Dept. of CSE, IIT Bombay*
  * Implemented **non-photorealistic rendering** technique by using **Edge tangent flow**, Flow-Based Difference of **Gaussians Filter** and Flow-Based **Bilateral Filter** iteratively for **line drawing** and **region smoothing**
  * Inspected improvement over non-flow based algorithms like Canny Edge Detector and Mean Shift Segmentation

TEACHING EXPERIENCE

- Undergraduate Teaching Assistant for *Programming for Data Science (DS203)* course at IIT Bombay 2021
- Undergraduate Teaching Assistant for *Introduction to Machine Learning (DS303)* course at IIT Bombay 2022

MISCELLANEOUS PROJECTS

- **Invention Factory**: Designed and prototyped a **mechanical washing device** for washing and spin drying reusable menstrual pads, received recognition in national media like **India Today**, **Outlook** and **The Better India** for the invention
- **Inter IIT Tech Challenge by Bridgei2i**: Worked on automated identification, summarization, and entity-based sentiment analysis of articles and tweets by fine-tuning models for specific tasks e.g., distill-BART for extreme summarisation of articles
- **D.R.D.O. SASE's UAV Fleet Challenge**: Contributed to the implementation of algorithms to **detect camouflaged targets** among a clutter of objects and subsequently relaying the location of identified targets to a map
- **Team Rakshak**: Part of the initiative to develop robust **Unmanned Aerial Vehicles** at IIT Bombay

RELEVANT COURSES

- **CS and ML Courses:**
  Deep Learning for Natural Language Processing, Information Retrieval & Mining for Hypertext & the Web, Decision Analysis and Game Theory, Advanced Topics in Machine Learning, Fundamentals of Digital Image Processing, Introducing to Machine Learning, Programming for Data Science, Network Security, Data Structures and Algorithms
- **Math Courses:**
  A First Course in Optimization, Markov Chains and Queuing Systems, Probability and Random Processes, Data Analysis and Interpretation, Complex Analysis, Differential Equations I & II, Linear Algebra, Calculus

SERVICES AND EXTRA-CURRICULAR ACTIVITIES

- Reviewer at CVPR 2024, ACMMM 2023, Pattern Recognition 2022 & 2023
- Secured overall **second position** in the 9th **Inter IIT Tech Meet** organised by IIT Guwahati
- Formed and led a 24*7 helpline team called **Covid Help Moradabad** to provide medical and emotional assistance to covid patients and families, and co-ordinated with the doctors in city to ensure availability and timely supply of oxygen cylinders
- Pursued Hindustani Vocals in **NSO (National Sports Organization)** under Prof. Veena Sawale
- Secured **third position** in **Hindustani Vocal Music Festival** organized by DPS Varanasi
- Actively involved in events of **State** and **National level Spic Macay Convention Society**