Ankit Jha, Assistant Professor

Dept.of CSE, The LNM Institute of Information Technology, Jaipur, India

✉ ankitjha16@gmail.com

⊙ Jaipur, India

Ankit Jha currently serves as a Assistant Professor in the Department of Computer Sceince and Engineering at the LNM Institute of Information Technology, Jaipur, Rajasthan, India. Previous to this, he worked as the Post-doctoral Researcher (Postdoc) at the National Institute for Research in Digital Science and Technology (INRIA), University Grenoble Alpes, Grenoble, France, focusing on multimodal learning (modalities like infrared and optical images) for object detection and classification. Previously, he held positions as a Research Scholar and Teaching Assistant at the Center of Studies in Resources Engineering (CSRE), Indian Institute of Technology Bombay (IITB), India. Ankit holds an M.Tech in Modelling and Simulation from Defence Institute of Advanced Technology (DIAT-DU), Pune, India and a B.Tech in Electronics and Communication Engineering from Rajasthan Technical University (RTU), Kota, India. Under the guidance of Dr. Biplab Banerjee, Associate Professor at CSRE, IIT Bombay, Ankit's research focuses on multimodal and multi-domain learning, particularly on remote sensing datasets. Along with this, his research discusses the interpretability of deeep learning models on diverse computer vision tasks. His PhD thesis titled "Learning Across Tasks and Domains in Remote Sensing" reflects his dedication to advancing knowledge in this specialized field.

Ankit's impactful contributions have been showcased at distinguished conferences such as the Computer Vision and Pattern Recognition (CVPR 2024), Computer Vision and Pattern Recognition Workshop (CVPRw 2020, 2023), British Machine Vision Virtual Conference (BMVC 2020, 2023), Winter Conference on Application of Computer Vision (WACV 2021, 2023, 2024), and Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2021, 2023) where his insights have earned recognition, including Best Paper Awards at the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD) 2023 and ICVGIP 2023. He was honored with the Best Poster Award at Product Life Cycle, Modelling, Simulation, and Synthesis (PLMSS) 2017 for his outstanding work on evaluating neural networks for handwriting and fingerprint recognition.

His groundbreaking work on multidomain learning under few-shot premises has been published in the IEEE Transactions on Geoscience and Remote Sensing (TGRS) journal. Ankit is actively involved in enhancing foundational visual language models such as CLIP through prompt learning, contributing to advancements in image recognition, semantic segmentation, and image retrieval tasks. Ankit Jha's relentless dedication, pioneering research, and profound contributions underscore his unwavering commitment to advancing the field of computer vision and making meaningful contributions to the broader scientific community. He also served as the system administrator for the Machine Learning and Visual Computing (MLVC) Lab at CSRE, IIT Bombay during his Ph D.

Education

Post Doctoral Researcher, National Institute for Research in Digital Science and Technology (INRIA), University Grenoble Alpes, Grenoble, France   Feb,24-Jun,24
Research Area: Multimodal Learning for Object Detection
Supervisor: Dr. Jocelyn Chanussot, Professor

Doctor of Philosophy (Ph. D.), Center of Studies in Resource Engineering (CSRE), Indian Institute of Technology Bombay (IITB), Mumbai-India   Jan,19-Apr,24
Final grade: 8.14
Subject of the dissertation: Learning Across Tasks and Domains in Remote Sensing
Supervisor: Dr. Biplab Banerjee, Associate Professor

Master of Technology (M.Tech.), Dept. of Applied Mathematics, Defence Institute of Advanced Technology (DIAT), Pune-India   Aug,16-May,18
Final grade: 8.24
Subject of the dissertation: Fingerprint and Signature Recognition Using Convolutional Neural Network
Supervisor: Dr. Debasish Pradhan, Associate Professor

Bachelor of Technology (B.Tech.), Dept. of Electronics and Communication Engineering, M.L.V. Textile and Engineering College (MLVTEC), Bhilwara-India   Aug,11-May,15
Final grade: 70.95%

Senior Secondary, Tagore Public School, Jaipur, Central Board of Secondary Education, India   May,10
Final grade: 63.20
Discipline: Physics, Chemistry, Mathematics (PCM)

Secondary, Maheshwari Public School, Jaipur, Rajasthan Board of Secondary Education, Jun,08
Rajasthan-India
Final grade: 79.33

Publications

ICCV 2025, FedMVP: Federated Multi-modal Visual Prompt Tuning for Vision-Language   Oct,25
Models
We propose FedMVP, a federated multimodal prompt tuning framework for CLIP that aligns visual and textual
attributes via a cross-attention−based PromptFormer, addressing the overfitting and limited adaptability of tex-
tual prompt tuning. Trained with CLIP similarity and consistency losses, FedMVP achieves strong in-distribution
accuracy while substantially improving generalization to unseen classes and domains across 20 datasets.

CVPR 2025, OSLoPrompt: Bridging Low-Supervision Challenges and Open-Set Domain   Jun,25
Generalization in CLIP
We introduce LSOSDG, a new paradigm combining low-shot learning with open-set domain generalization, where
existing CLIP-based prompt methods fall short in low-data settings and fine-grained open-set detection. To over-
come this, we propose OSLoPrompt, a prompt-learning framework that integrates domain-agnostic and visually
guided prompts via cross-attention, and leverages pseudo-open samples to improve unknown class detection
with fine-grained awareness.

TMLR 2025, Meta-Learning to Teach Semantic Prompts for Open Domain Generalization in   Mar,25
Vision-Language Models
Open-Domain Generalization (ODG) faces challenges from both domain and category shifts. Existing CNN and
CLIP-based methods struggle to generalize and distinguish unknown classes. We propose MetaPrompt, combin-
ing CLIP's semantic power with Meta-Learning to robustly classify known and unknown classes across domains
using novel, attention-guided prompts and contrastive loss for improved unknown class awareness.

WACV 2025, Foundation Models and Adaptive Feature Selection: A Synergistic Approach to   Jan,25
Video Question Answering
We propose Local-Global Question Aware Video Embedding (LGQAVE), a framework that unifies multi-modal
knowledge for VideoQA. It uses cross-attention to select question-relevant frames, models object dynamics with
miniGPT-grounded graphs refined by a question-aware transformer, and fuses local-global embeddings for an-
swer generation.

BMVC 2024, COSMo: CLIP Talks on Open-Set Multi-Target Domain Adaptation   Jul,24
We introduce COSMo, a novel framework for Open-Set Multi-Target Domain Adaptation (OSMTDA) using source
domain-guided prompt learning. Leveraging frozen CLIP encoders with minimal trainable parameters, COSMo
uses a domain-specific bias network and separate prompts for known and unknown classes to handle both
domain and class shifts. It's the first to address OSMTDA, reflecting real-world challenges, though it assumes
consistent target domain classes, which may vary in practice.

ECCV 2024, Elevating All Zero-Shot Sketch-Based Image Retrieval Through Multimodal   May,24
Prompt Learning
SpLIP, a CLIP-based framework for Zero-Shot (ZS), Generalized ZS (GZS), and Fine-Grained ZS (FG-ZS) sketch-
based image retrieval (SBIR). SpLIP introduces: (1) a multi-modal prompt learning strategy for improved cross-
modal knowledge sharing, (2) dynamic margin adjustment using CLIP's textual embeddings for sketch-photo
triplet loss, and (3) a conditional jigsaw task to refine sketch-photo associations. SpLIP consistently excels across
benchmarks, with plans to extend into broader vision-language domains.

CVPR 2024, Unknown Prompt, the only Lacuna: Unveiling CLIP's Potential for Open Domain   Feb,24
Generalization
We present ODG-CLIP, leveraging CLIP's semantic capabilities. Our framework introduces three key innovations.
Firstly, we redefine ODG as a multi-class classification task, encompassing known and novel categories. We
design a unique prompt for detecting unknown classes and train it using a stable diffusion model to generate
proxy images. Secondly, we develop a novel mechanism for learning domain-tailored classification weights,
ensuring both precision and simplicity. Finally, we enhance CLIP's visual embeddings with class-discriminative
knowledge derived from the prompt space. We introduce a novel objective to maintain the continuity of this
semantic information across domains, particularly for shared classes.

CVPR-L3D-IVU 2024, Bridging Domain Gaps in Generalized Category Discovery   Apr,24
We introduce AD-GCD, a challenging task where disparate distributions hinder clustering in the GCD setting. Our
CDAD-NET framework addresses this with targeted domain alignment and domain-centric metric learning using
contrastive methods. Additionally, we incorporate conditional image inpainting to enrich feature associations
within image embeddings.

TGRS 2023, MDFS-Net: Multi-domain Few shot Classification for Hyperspectral Images With   Oct,23
Support Set Reconstruction

We propose the concept of few-shot multi-domain learning (FS-MDL) within the realm of HSI classification. We present an intuitive meta-learning strategy for training model parameters, which are predominantly shared across domains. To retain domain-specific information, we employ separate batch-normalization parameters for each domain. In addressing overfitting, we introduce the concept of minimizing feature uncertainty and incorporate a self-supervised task for sample reconstruction.

TMLR 2023, Beyond Boundaries: A Novel Data-Augmentation Discourse for Open Domain Generalization   Oct,23

We create an open-set classifier in a discriminative, unbiased, and disentangled semantic space. Using a generative augmentation framework, we enrich data diversity by generating novel domains through style interpolation. This strategy effectively synthesizes images by disentangling style and content information. To address style bias, we represent images in relation to source domain properties, enhancing visual features. This leads to a multi-class semantic object classifier with closed and open class capabilities, and a style classifier. This joint approach disentangles the latent space, improving the semantic classifier's generalization.

ICVGIP 2023, C-SAW: Self-Supervised Prompt Learning for Image Generalization in Remote Sensing (**Best Paper Award**)   Oct,23

C-SAW, tackles the limitations of part embeddings in CLIP by incorporating a reconstruction task to enhance the latent visual space of the distorted input image. Furthermore, we generate the visual attention tokens using token generator $\mathcal{G}_{VAT}$ before the frozen text encoder in CLIP to impose desired prompt constraints on the input visual embeddings.

BMVC 2023, GoPro: Generate and Optimize Prompts in CLIP using Self-Supervised Learning   Aug,23

We improve CLIP's prompt learning with SSL and disentangled image-conditioned prompts. Our contributions include enhancing vision and text projectors on a frozen CLIP with visual-space SSL contrastive loss, image-text contrastive loss, and a novel prompt consistency loss for diverse image views. Additionally, we introduce prompt distribution learning using multi-scale visual content and style information from CLIP.

WACV 2024, Learning Class and Domain Augmentations for Single-Source Open-Domain Generalization   Aug,23

we propose a novel framework called SODG-NET that simultaneously synthesizes novel domains and generates pseudo-open samples using a learning-based objective, in contrast to the ad-hoc mixing strategies commonly found in the literature. Our approach enhances generalization by diversifying the styles of known class samples using a novel metric criterion and generates diverse pseudo-open samples to train a unified and confident multiclass classifier capable of handling both open and closed-set data.

WACV 2024, StyLIP: Multi-Scale Style-Conditioned Prompt Learning for CLIP-based Domain Generalization   Aug,23

We addresses the issue of prompt learning under the domain-shift and improve CLIP's generalization ability across domains. StyLIP, a novel approach for Domain Generalization (DG) based on a domain-agnostic prompt learning strategy. In the absence of explicit domain knowledge, we aim to disentangle the visual style and the content information extracted from the pre-trained CLIP in the prompts so they can be effortlessly adapted to novel domains during inference.

Pattern Recognition Letters, Elsevier 2023, MAML-SR: Self-Adaptive Super-Resolution Networks via Multi-scale Optimized Attention-aware Meta-Learning   Aug,23

A meta-learning approach for Super-Resolution (SR) that optimizes LR image patch patterns using multi-scale optimization. We enhance stability with second-order optimization, employ cross-scale attention and deep-supervision modules, and introduce a metric-dependent constraint for efficient gradient updates during meta-testing.

ICCVw, OOD-CV 2023, AD-CLIP: Adapting Domains in Prompt Space Using CLIP   Aug,23

A domain-agnostic prompt learning technique for CLIP. It uses supervised contrastive learning for the source domain, entropy minimization for domain alignment in the embedding space with target domain data, and a cross-domain style mapping network for testing with only target domain samples.

ECML-PKDD2023, HAVE-Net: Hallucinated Audio-Visual Embeddings for Few-Shot Classification with Unimodal Cues (**Best Paper Award**)   Jul,23

We introduce a novel modality-agnostic multimodal framework in a few-shot context, namely, HAVE-Net, to battle against the missing modalities in the meta-testing phase. We propose a novel Conditional Multimodal Adversarial loss to pre-train our CMM-GAN classifier from the available unimodal data (either audio or visual) to hallucinate missing modality features in testing time.

Book Chapter, Springer, Self-Distillation with the New Paradigm in Multi-Task Learning   June,23

We demonstrate the setting of self-distillation under the multi-task learning (MTL) scenario. We compiled both soft and hard parameter based strategies for the self-distilled MTL networks, results showcased the effectiveness of such methods in dense-prediction tasks.

ISSCS2023, RS-MCQA: Multi-class Question Aware Visual Question Answering for Optical Remote Sensing Datasets   June,23

We tackle the problem of visual question answering (VQA) towards automated landcover information retrieval. Analysis of the remote sensing datasets involves simple classification problems to complex regression challenges of the widely diversified ground objects with varying shapes and sizes.

CVPR-EV2023, APPLeNet: Visual Attention Parameterized Prompt Learning for Few-Shot   Apr,23
Remote Sensing Image Generalization using CLIP

We study the problem of domain generalization in remote sensing using the foundational model. APPLeNet emphasizes the importance of multi-scale feature learning in RS scene classification and disentangles visual style and content primitives for domain generalization tasks.

WACV2023, GAF-Net: Improving the Performance of Remote Sensing Image Fusion using   Jan,23
Novel Global Self and Cross Attention Learning

The self-attention models fail to incorporate the global context due to the limited size of the receptive fields, cross-attention learning may generate ambiguous features as the feature extractors for all the modalities are jointly trained. We introduce the within-modality feature refinement module through global spectral-spatial attention learning using the query-key-value processing where both the global spatial and channel contexts are used to generate two channel attention masks.

ICVGIP2021, S^3DMT-Net: Improving soft sharing based multi-task CNN using task-specific   Sep,21
distillation and cross-task interactions

We tackle the problem of self-distillation for multi-task learning using the soft-sharing mechanism at the encoder end and task-specific decoders at the decoder end. It aims at distilling knowledge from deeper CNN layers into the shallow layers and generate the explainability of the multitask dense prediction model.

WACV2021, ADA-AT/DT: An Adversarial Approach for Cross-Domain and Cross-Task   Jan,21
Knowledge Transfer

We deal with the problem of cross-task and cross-domain knowledge transfer in the realm of scene understanding for autonomous vehicles. We consider the scenario where supervision is available for a pair of tasks in a source domain while it is available for only one of the tasks in the target domain. Given that, the goal is to perform inference for the task in the target, which is devoid of any training information.

BMVC2020, SD-MTCNN: Self-Distilled Multi-Task CNN   Sep,20

A multi-task network with the self-distillation mechanism. SD-MTCNN aims at distilling knowledge from deeper CNN layers into the shallow layers. It is a novel paradigm of self-distillation within the network. It guarantees improved multi-task performance from different parts of the network.

EDLCV, CVPRw2020, AdaMT-Net: An Adaptive Weight Learning Based Multi-Task Learning   Jun,20
Model For Scene Understanding

A typical U-Net based encoder-decoder architecture called AdaMT-Net, where the densely-connected deep convolutional neural network (CNN) based feature encoder is shared among the tasks while the soft-attention based task-specific decoder modules produce the desired outcomes. It is a novel adaptive weight learning strategy by carefully exploring the loss-gradients per-task over the training iterations.

ICIP2020, MT-UNET: A Novel U-Net Based Multi-Task Architecture For Visual Scene   Oct,20
Understanding

MT-UNET, a densely-connected encoder with task-specific decoders. It also encourages the cross-talk (CT) between the tasks by introducing cross-task skip connections at the decoder end with adaptive weight learning for the task-specific loss functions in the final cost measure.

PLMSS17, Poster Presentation - Hand Written Character Recognition Using Convolutional   Dec,17
Neural Network-A Comprehensive Review

The review comprises with the comparison with various traditional and advanced computing (deep learning) approaches on hand written characters and fingerprints.

CHAKRAVYUH MAGAZINE, Article on Primer on Satellite Communication   Aug,14

In this article, we provided a brief summary on essentials of satellite communication, which got published in Chakravyuh magazine of Scientific Analysis Group (SAG), Defence Research and Development Organization (DRDO), New Delhi, India.

Contribution to the Community

Conferences:

Winter Conference on Applications of Computer Vision (WACV)	2025
Computer Vision and Pattern Recognition Conference (CVPR)	2025
International Conference on Computer Vision (ICCV)	2025
Winter Conference on Applications of Computer Vision (WACV)	2024
Computer Vision and Pattern Recognition Conference (CVPR)	2024
Earth Vision Workshop at CVF Computer Vision and Pattern Recognition Conference (CVPRw)	2024
IEEE Conference on Artificial Intelligence (CAI)	2024

European Conference on Computer Vision (ECCV)	2024
Winter Conference on Applications of Computer Vision (WACV)	2023
Computer Vision and Pattern Recognition Conference (CVPR)	2023
International Conference on Computer Vision (ICCV)	2023
IEEE Conference on Artificial Intelligence (CAI)	2023
Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)	2023
British Machine Vision Virtual Conference (BMVC)	2023
Winter Conference on Applications of Computer Vision (WACV)	2022
European Conference on Computer Vision (ECCV)	2022
British Machine Vision Virtual Conference (BMVC)	2022
Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)	2022
Winter Conference on Applications of Computer Vision (WACV)	2021
British Machine Vision Virtual Conference (BMVC)	2021
British Machine Vision Virtual Conference (BMVC)	2020

Technical Committee:

PAN India Forum, Indian Symposium on Machine Learning (IndoML)	2025
Domain Generalization: Evolution, Breakthroughs and Future Horizon, IEEE Computer Vision and Pattern Recognition (CVPR) Workshop	2025
Earth Vision, IEEE Computer Vision and Pattern Recognition (CVPR) Workshop	2024

Journals and Chapters:

Transactions on Geoscience and Remote Sensing (TGRS)
Pattern Recognition, Elsevier
Pattern Recognition Letters, Elsevier
Geoscience and Remote Sensing Letters (GRSL), IEEE
Journal of Selected Topics in Applied Earth Observations and Remote Sensing, IEEE
International Journal of Computer Vision (IJCV), Springer
Chapter Reviewer: Advancements in Knowledge Distillation: Towards New Horizons of Intelligent Systems, Springer-Verlag

Research Talk and Training

Teaching, Madhya Pradesh and Foundation for Innovation and Technology Transfer (FITT), IIT Delhi	Apr,25
Research Training and Organizing, FDP cum Workshop on Generative AI, The LNMIIT, Jaipur, India	Mar,25
Teaching, Madhya Pradesh and Foundation for Innovation and Technology Transfer (FITT), IIT Delhi	Nov,24
Teaching, Madhya Pradesh and Foundation for Innovation and Technology Transfer (FITT), IIT Delhi	Jan,24
Research Training, ATAL - FDP, Sardar Patel Institute of Technology, Mumbai, India	Jan,24
Research Training, Introduction to SAR Image Processing and Deep Learning Techniques, Center of Studies in Resources Engineering, Indian Institute of Technology Bombay, India	Sep,23
Research Talk and Training, Advanced Data Processing Research Institute (ADRIN), Indian Space Research Organisation (ISRO), Hyderabad, India	Feb,23
Research Talk, GeoAI Roundtable, The Open Geospatial Consortium (OGC),	Mar,23
Research Session, FDP Optimization & Deep Learning, National Institute of Technology Patna, India	Jan,21
Lecture and Training, Online course on Artificial Intelligence and Machine Learning by DIAT, Pune, India	Nov,20
Teaching Assistant and Volunteer, Deep Learning for Remote Sensing, Continuing Education Program (CEP), Indian Institute of Technology Bombay, India	Nov,19

Experience

Assistant Dean of Academics Affairs, The LNM Institute of Information Technology, Jaipur, Rajasthan, India	Sep,25-Present
Assistant Professor, Dept. of Computer Science and Engineering, The LNM Institute of Information Technology, Jaipur, Rajasthan, India	Jul,24-Present
Postdoctoral Researcher, INRIA, Grenoble France	Feb,24-Jun,24
Content Creation, Madhya Pradesh and Foundation for Innovation and Technology Transfer (FITT), IIT Delhi	Jan,24

Doctoral Researcher, Indian Institute of Technology Bombay, India — Jan,19-Dec,23

System Administration, Machine Learning and Visual Computing (MLVC) Lab, Indian Institute of Technology Bombay (IITB), Mumbai, India — Jul,19-Jan,24

Inter-IIT Tech meet, Mentor for Inter-IIT Tech Meet — Feb,23

Teaching Assistant, Machine Learning for Remote Sensing-II — Jun,22-Dec,22

Teaching Assistant, Machine Learning for Remote Sensing-II — Jun,21-Dec,21

Research Internship, Defence Geoinformatics Research Establishment (DGRE), Defence Research & Development Organisation (DRDO), Chandigarh-India — Oct,21-Nov,21
Supervisor: Dr. Pinaki Roy Chowdhury, Sc.'G'

Teaching Assistant, Machine Learning for Remote Sensing-I — Jan,21 - May,21

Teaching Assistant, Introduction to Machine Learning Minor — Jan,21 - May,21

Teaching Assistant, Machine Learning for Remote Sensing-II — Aug,20 - Dec,20

Workshop in IRISS 2020, Indian Institute of Technology Gandhinagar — Feb,20

Teaching Assistant, Machine Learning for Remote Sensing-I — Jan,20 - May,20

Research Expert, UpGrad Education Pvt. Ltd., Mumbai, India — Nov,19-Dec,22

Research Fellow, Indian Institute of Information Technology, Sri City, Andhra Pradesh, India — Aug,18-Nov,18

Summer Internship, Scientific Aalysis Group (SAG), Defence Research & Development Organisation (DRDO), New Delhi-India — Jun,14-Aug,14
Supervisor: Dr. Devendra Jha, Sc.'F'

Achievements

Best Paper Award: Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP) — Dec,23

Best Paper Award: Adapting to Change: Reliable Multimodal Learning Across Domains in European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD) — Sep,23

Doctoral Consortium: Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP) — Dec,21

Doctoral's Fellowship: Indian Institute of Technology Bombay, India — Jan,19-Dec,23

Junior Research Fellowship: Joint Cipher Bureau, Defence Research and Development Organization, New Delhi, India — Aug,18

Best Poster Award: Product Lifecycle Modeling Simulation and Synthesis Conference (PLMSS) — Dec, 17

Master's Fellowship: Defence Institute of Advanced Technology, Pune, India — Aug,16-May,18

Extra-curricular

Local organiser in ICISS, 2024

Organising member for Post-Quantum Cryptography in ICISS, 2024

Volunteer in IndoML, 2023

Departmental Coordinator, 2023

Local organising committee for CODS-COMAD, 2023

Organising committee for inter-hostel volleyball tournament, 2022

Serving as the sports-secretary for Hostel-13, IITB, 2022

Sports organiser in Spandan, DIAT, Pune, 2018

Personal interest: Volleyball, Cricket, Table Tennis, Science, Technology, Photography, Music

Personal details

Parents: Mrs. Ranju Jha, Mr. Narayan Jha
E-mail: ankit.jha@lnmiit.ac.in, ankitjha16@gmail.com

DoB: 16 March 1994
Mobile: +91 - {9284549081, 9764194921}

Languages

English: Full professional proficiency, **Hindi**: Proficiency, **Maithili**: Native proficiency

Publications

1. **Ankit Jha**, Awanish Kumar, Shivam Pande, Biplab Banerjee and Subhasis Chaudhuri, "MT-UNET: A Novel U-Net Based Multi-Task Architecture For Visual Scene Understanding," 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 2191-2195, doi: 10.1109/ICIP40778.2020.9190695.

2. **Ankit Jha**, Awanish Kumar, Biplab Banerjee and Subhasis Chaudhuri, "AdaMT-Net: An Adaptive Weight Learning Based Multi-Task Learning Model For Scene Understanding," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, doi: 10.1109/CVPRW50498.2020.00361.

3. Ruchika Chavhan, **Ankit Jha**, Biplab Banerjee and Subhasis Chaudhuri, "ADA-AT/DT: An Adversarial Approach for Cross-Domain and Cross-Task Knowledge Transfer," 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 2021, pp. 3501-3510, doi: 10.1109/WACV48630.2021.00354.

4. **Ankit Jha**, Awanish Kumar, Biplab Banerjee and Vinay Namboodiri, "SD-MTCNN: Self-Distilled Multi-Task CNN." British Machine Vision Conference (BMVC), 2020.

5. **Ankit Jha**, Biplab Banerjee, and Subhasis Chaudhuri, "S^3DMT-Net: improving soft sharing based multi-task CNN using task-specific distillation and cross-task interactions." In Proceedings of the Twelfth Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP '21). Association for Computing Machinery, New York, NY, USA, Article 16, 1–9. DOI:https://doi.org/10.1145/3490035.3490274.

6. **Ankit Jha**, Shirsha Bose, and Biplab Banerjee, "GAF-Net: Improving the Performance of Remote Sensing Image Fusion using Novel Global Self and Cross Attention Learning." IEEE Winter Conference on Applications of Computer Vision (WACV), 2023.

7. Mainak Singha, **Ankit Jha**, Bhupendra Solanki, Shirsha Bose and Biplab Banerjee, "APPLeNet: Visual Attention Parameterized Prompt Learning for Few-Shot Remote Sensing Image Generalization using CLIP", IEEE, CVPR, EarthVision 2023.

8. **Ankit Jha** and Biplab Banerjee, "Self-Distillation with the New Paradigm in Multi-Task Learning", Advancements in Knowledge Distillation: Towards New Horizons of Intelligent Systems, 2023 (Book Chapter).

9. Hitul Desai, Debabrata Pal, **Ankit Jha**, Avik Hati, and Biplab Banerjee, "RS-MCQA: Multi-class Question Aware Visual Question Answering for Optical Remote Sensing Datasets", International Symposium on Signals Circuits and Systems - ISSCS 2023.

10. **Ankit Jha**, Debabrata Pal, Mainak Singha, and Biplab Banerjee, "HAVE-Net: Hallucinated Audio-Visual Embeddings for Few-Shot Classification with Unimodal Cues", The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), 2023.

11. Mainak Singha, Harsh Pal, **Ankit Jha**, and Biplab Banerjee, "AD-CLIP: Adapting Domains in Prompt Space Using CLIP", Out Of Distribution Generalization in Computer Vision, The International Conference on Computer Vision (ICCV), 2023.

12. Debabrata Pal, Shirsha Bose, Deeptej More, **Ankit Jha**, Biplab Banerjee and Yogananda Jeppu, "MAML-SR: Self-Adaptive Super-Resolution Networks via Multi-scale Optimized Attention-aware Meta-Learning", Pattern recognition Letters, Elsevier, 2023.

13. Shirsha Bose, **Ankit Jha**, Enrico Fini, Mainak Singha, Biplab Banerjee, and Elisa Ricci, "STyLIP: Multi-Scale Style-Conditioned Prompt Learning for CLIP-based Domain Generalization", IEEE Winter Conference on Applications of Computer Vision (WACV), 2024.

14. Mainak Singha, **Ankit Jha**, and Biplab Banerjee, "GOPro: Generate and Optimize Prompts in CLIP Using Self-Supervised Learning", British Machine Vision Conference (BMVC), 2023.

15. Avigyan Bhattacharya, Mainak Singha, **Ankit Jha**, and Biplab Banerjee, "`C-SAW`: Self-Supervised Prompt Learning for Image Generalization in Remote Sensing", In Proceedings of the 14th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP '23).

16. **Ankit Jha**, and Biplab Banerjee, "MDFS-Net: Multi-domain Few shot Classification for Hyperspectral Images With Support Set Reconstruction", In IEEE Transactions on Geoscience and Remote Sensing (TGRS).

17. Prathmesh Bele, Valay Bundele, Avigyan Bhattacharya, **Ankit Jha**, Gemma Roig, and Biplab Banerjee. "Learning Class and Domain Augmentations for Single-Source Open-Domain Generalization", IEEE Winter Conference on Applications of Computer Vision (WACV), 2024.

18. Shirsha Bose, **Ankit Jha**, and Biplab Banerjee. "Beyond Boundaries: A Novel Data-Augmentation Discourse for Open Domain Generalization", Transactions on Machine Learning Research (TMLR).

19. Mainak Singha, **Ankit Jha**, Shirsha Bose, Moloud Abdar, and Biplab Banerjee. "Unknown Prompt, the only Lacuna: Unveiling CLIP's Potential for Open Domain Generalization", IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2024.

20. **Ankit Jha**, Mainak Singha, Avigyan Bhattacharya, and Biplab Banerjee. "RS^{3}Lip: Consistency for Remote Sensing Image Classification on Part Embeddings using Self-supervised Learning and CLIP".

21. Sai Bhargav Rongali, Sarthak Mehrotra, **Ankit Jha**, Mohamad Hassan N C, Shirsha Bose, Tanisha Gupta, Mainak Singha, Biplab Banerjee. "Bridging Domain Gaps in Generalized Category Discovery", IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshop, 3rd Workshop on Learning with Limited Labelled Data for Image and Video Understanding, 2024.

22. Mainak Singha, **Ankit Jha**, Divyam Gupta, Pranav Singla, Biplab Banerjee. "Elevating All Zero-Shot Sketch-Based Image Retrieval Through Multimodal Prompt Learning", the European Conference on Computer Vision (ECCV), 2024

23. Munish Monga, Sachin Kumar Giroh, **Ankit Jha**, Mainak Singha, Biplab Banerjee, Jocelyn Chanussot. "COSMo: CLIP Talks on Open-Set Multi-Target Domain Adaptation", British Machine Vision Conference (BMVC), 2024.

24. Sai Bhargav Rongali, Mohamad Hassan N C, **Ankit Jha**, Neha Bhargava, Saurabh Prasad, Biplab Banerjee. "Foundation Models and Adaptive Feature Selection: A Synergistic Approach to Video Question Answering", IEEE Winter Conference on Applications of Computer Vision (WACV), 2025.

25. Alessia Boccallate, **Ankit Jha**, Jocelyn Chanussot. "Leveraging Large-scale Aerial Data for Accurate Urban Rooftop Solar Potential Estimation via Multitask Learning", Journal for Solar Energy, Elsevier, 2025.

26. Mohamad Hassan N C, Divyam Gupta, Mainak Singha, Sai Bhargav Rongali, **Ankit Jha**, Muhammad Haris Khan, Biplab Banerjee. "OSLoPrompt: Bridging Low-Supervision Challenges and Open-Set Domain Generalization in CLIP", IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2025.

27. Shirsha Bose, Mainak Singha, **Ankit Jha**, Souradeep Mukhopadhyay, Biplab Banerjee. "Meta-Learning to Teach Semantic Prompts for Open Domain Generalization in Vision-Language Models", Transactions on Machine Learning Research (TMLR) 2025.

28. Shirsha Bose, Mainak Singha, **Ankit Jha**, Souradeep Mukhopadhyay, Biplab Banerjee. "FedMVP: Federated Multi-modal Visual Prompt Tuning for Vision-Language Models", The International Conference on Computer Vision (ICCV) 2025.