Ankit Jha, PhD Scholar

Machine Learning and Visual Computing Lab, CSRE, IIT Bombay

ankitjha16@gmail.com

Mumbai, India

Ankit Jha has been a full-time Research Scholar and Teaching Assistant (TA) at the Center of Studies in Resources Engineering (CSRE), Indian Institute of Technology Bombay (IITB), India, since January 2019. He holds M.Tech in Modelling and Simulation and B.Tech in Electronics and Communication Engineering. He is currently working under the supervision of Dr. Biplab Banerjee, Associate Professor, CSRE, IIT Bombay.

Ankit's research primarily focuses on multimodal and multi-domain learning, with a special emphasis on remote sensing datasets. His previous work has included tackling Multi-Task Learning (MTL) problems and Self-Distillation (SD) for visual scene inferences on indoor, urban, and remote sensing areas. Additionally, he has worked on domain adaptation, domain generalization, and few-shot learning in MTL.

Over the past few years, Ankit has published papers in well-known conferences, such as Computer Vision and Pattern Recognition Workshop (CVPRw 2020), British Machine Vision Virtual Conference (BMVC 2020), and Winter Conference on Application of Computer Vision (WACV, 2023). He is actively seeking research opportunities to expand his knowledge and is making significant contributions to the field.

Education

Doctor of Philosophy (Ph. D.), Center of Studies in Resource Engineering (CSRE), Indian Institute of Technology Bombay (IITB) , Mumbai-India — Jan,19-Present
Final grade: 8.14
Subject of the dissertation: Multi-task Learning for Remote Sensing Images
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

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%

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

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'

Publications

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

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.

Under-review StyLIP: Multi-Scale Style-Conditioned Prompt Learning for CLIP-based Domain Generalization (under-review) — Feb,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.

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

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 distillation and cross-task interactions — Sep,21

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.

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

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 Model For Scene Understanding — Jun,20

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 Understanding — Oct,20

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 Neural Network-A Comprehensive Review — Dec,17

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), New Delhi, India.

Contribution to the Community

Conferences:

British Machine Vision Virtual Conference (BMVC)	2020
Winter Conference on Applications of Computer Vision (WACV)	2021
British Machine Vision Virtual Conference (BMVC)	2021
Winter Conference on Applications of Computer Vision (WACV)	2022
British Machine Vision Virtual Conference (BMVC)	2023
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)	2023
CVF Computer Vision and Pattern Recognition Conference (CVPR)	2023
International Conference on Computer Vision (ICCV)	2023
IEEE Conference on Artificial Intelligence (CAI)	2023

Journals and Chapters:
Pattern Recognition, Elsevier
Geoscience and Remote Sensing Letters (GRSL), IEEE
International Journal of Computer Vision (IJCV), Springer
Chapter Reviewer: Advancements in Knowledge Distillation: Towards New Horizons of Intelligent Systems, Springer-Verlag

Experience

System Administration, Machine Learning and Visual Computing (MLVC) Lab, Indian Institute of Technology Bombay (IITB), Mumbai, India — July,2019-Present
Research Talk and Training, Advanced Data Processing Research Institute (ADRIN), Indian Space Research Organisation (ISRO), Hyderabad, India — Feb,23
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
Teaching Assistant, Machine Learning for Remote Sensing-I — Jan,21 - May,21
Research Session, FDP Optimization & Deep Learning, National Institute of Technology Patna, India — Jan,21
Teaching Assistant, Introduction to Machine Learning Minor — Jan,21 - May,21
Teaching Assistant, Machine Learning for Remote Sensing-II — Aug,20 - Dec,20
Lecture and Training, Online course on Artificial Intelligence and Machine Learning by DIAT, Pune, India — Nov,20
Teaching Assistant, Machine Learning for Remote Sensing-I — Jan,20 - May,20
Teaching Assistant, Deep Learning for Remote Sensing, Continuing Education Program (CEP), Indian Institute of Technology Bombay, India — Nov,19
Research Expert, UpGrad Education Pvt. Ltd., Mumbai, India — Nov,19-Present
Research Fellow, Indian Institute of Information Technology, Sri City, Andhra Pradesh, India — Aug,18-Nov,18

Achievements

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

Languages

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

Extra-curricular

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: ankitjha16@iitb.ac.in, ankitjha16@gmail.com
DoB: 16 March 1994
Mobile: +91 -{9284549081, 9764194921}

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." 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." 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 2023.
7. Shirsha Bose, Enrico Fini, Ankit Jha, Mainak Singha, Biplab Banerjee, and Elisa Ricci, "StyLIP: Multi-Scale Style-Conditioned Prompt Learning for CLIP-based Domain Generalization", (Under-review).
8. 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.
9. 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).
10. 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.