Badri N. Patro

Ferns, Bangalore
India-560035
✆ (+91) 9076237295
FAX patrobadri.iitb@gmail.com
✉ badri.patro@microsoft.com
🖥 https://badripatro.github.io/

Research Interests: Transformers and Large Language Models(LLMs) in Computer Vision, Natural Language Processing, and Speech Processing. Responsible AI (Explainable AI, Fairness, Privacy in ML-based model, Transformers, and LLMs). Multimodal Representation Learning.

Education

2015–2020 **Doctor of Philosophy in Electrical Engineering**, *Indian Institute of Technology Kanpur, India*, Specialized in Signal Processing, Communications & Networks.
Supervisor: Prof. Vinay P. Namboodiri

2009–2012 **Masters of Technology in Electrical Engineering**, *Indian Institute of Technology Bombay, India*, Specialized in Communication & Signal Processing.
Supervisor: Prof. V. Rajbabu

2003–2007 **Bachelor of Technology in Electronic & Tele Communication Engineering**, *National Institute of Science and Technology*, Brahmapur, Orissa.
Supervisor: Prof. Rakesh Roshan

Employment

A-8 **Microsoft, India**, *Senior Research Scientist*, (June 20222 – Present) .
Advisor: Dr. Vijay Agneeswaran

A-7 **KU Leuven, Belgium**, *Postdoctoral Researcher*, (March 2021 – June 20222) .
Advisor: Prof. Luc. Van Gool

A-6 **IIIT Hyderabad, India** , *Postdoctoral Researcher*, (Jan 2021 – March 2021).
Advisor: Prof. C. V. Jawahar

A-5 **Google Research**, *Postdoctoral Researcher*, (Feb 2020–Dec 2020).
Advisor: Dr.Gaurav Aggarwal

A-4 **Microsoft (R&D)**, *Data Scientist Intern*, (May 2019 – July 2019).
Advisor: Dr. Mithun Dasgupta

A-3 **Samsung R&D Institute, Delhi**, *Lead Engineer*, (July 2013 – Dec 2015).

A-2 **Harman International**, *Associate Software Engineer*, (July 2012 – July 2013).

A-1 **Larsen & Toubro EmSyS**, *Assistant Software Engineer*, (Aug 2007 – Aug 2009).
(A tends for Affiliation)

Doctoral dissertation award

2020 **Awarded the best Doctoral Dissertation Award** , *by the "Indian Unit for Pattern Recognition and Artificial Intelligence" (IUPRAI)*, at ICVGIP, IIT Jodhpur, 2020.
.

Patents

U.S. Patent Badri N. Patro, and Vijay Srinivas Agneeswaran, "Scattering Vision Transformer." U.S. Patent, filed by Microsoft, May 17, 2023.

U.S. Patent Pranya Lohiya*, Badri N. Patro*, and Naveen Panwar*,. "SPASE: Spatial Saliency Explanation for Time Series Models." U.S. Patent, filed by Microsoft, Jan 17, 2023.

Journal Publication

TIP-21 **Badri N. Patro**, Mayank Lunayach, Vinay P. Namboodiri,"Uncertainty-CAM: Visual Explanation using Uncertainty based Class Activation Maps", **IEEE Transactions on Image Processing (TIP)**, 2021.

PR-21 **Badri N. Patro**, Anupriy, Vinay P. Namboodiri,"Adversarial Explanation: A Two-Player Game to obtain Attention for VQA", **Pattern Recognition**, 2021 (Under Revision).

IVC-21 **Badri N. Patro**, Vinod K. Kurmi, Sandeep Kumar, Vinay P. Namboodiri,"MUMC: Minimizing Uncertainty of Mixture of Cues", **Image and Vision Computing**,issn: 0262-8856, doi: https://doi.org/10.1016/j.imavis.2021.104280, 2021.

Neuro-20 **Badri N. Patro**, Dev Chauhan, Vinod K. Kurmi, Vinay P. Namboodiri,"Revisiting Paraphrase Question Generator using Pairwise Discriminator", **Neurocomputing** , issn: 0925-2312, 2020, doi: "https://doi.org/10.1016/j.neucom.2020.08.022".

PR-20 **Badri N. Patro**, Anupriy, Vinay P. Namboodiri,"Probabilistic framework for solving Visual Dialog", **Pattern Recognition**, Pages:107586, 2020.

Conference Publication

CVPR-18 **Badri N. Patro**, Vinay P. Namboodiri,"Differential Attention for Visual Question Answering", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, Utah, USA, 2018.

EMNLP-18 **Badri N. Patro**, Sandeep Kumar, Vinod K. Kurmi, Vinay P. Namboodiri,"Multimodal Differential Network for Visual Question Generation", Conference on Empirical Methods in Natural Language Processing (EMNLP), Belgium, 2018.

ICCV-19 **Badri N. Patro**, Mayank Lunayach, Shivansh Patel, Vinay P. Namboodiri,"U-CAM: Visual Explanation using Uncertainty based Class Activation Maps", International Conference on Computer Vision (ICCV), Seoul, South Korea, 2019.

AAAI-20 **Badri N. Patro**, Anupriy, Vinay P. Namboodiri,"Explanation vs Attention: A Two-Player Game to obtain Attention for VQA", Association for the Advancement of Artificial Intelligence (AAAI), Hilton, New York, USA 2020.

COLING-18 **Badri N. Patro***, Vinod K. Kurmi*, Sandeep Kumar*, Vinay P. Namboodiri,"Learning Semantic Sentence Embeddings using Pair-wise Discriminator", Proceedings of 27th International Conference on Computational Linguistics (COLING 2018), Santa Fe, New Mexico, USA, 2018.

WACV-20 **Badri N. Patro**, Shivansh Patel, Vinay P. Namboodiri,"Robust Explanations for Visual Question Answering", Winter Conference on Applications of Computer Vision (WACV '20), Colorado, USA, 2020.

WACV-20 **Badri N. Patro**, Vinod K. Kurmi, Sandeep Kumar, Vinay P. Namboodiri,"Deep Bayesian Network for Visual Question Generation", Winter Conference on Applications of Computer Vision (WACV '20), Colorado, USA, 2020.

WACV-21 **Badri N. Patro***, G.S. Kasturi*, Ansh Jain*, Vinay P. Namboodiri,"Self Supervision for Attention Networks", Winter Conference on Applications of Computer Vision (WACV '21), 2021.

WACV-21 **Badri N. Patro***, Mayank Lunayach, *, Deepankar Srivastav, Sarvesh, Hunar Singh, Vinay P. Namboodiri,"Multimodal Humor Dataset: Predicting Laughter tracks for Sitcoms", Winter Conference on Applications of Computer Vision (WACV '21), 2021.

WACV-21 Vinod K. Kurmi, **Badri N. Patro**, Vinay P. Namboodiri,"Do not Forget to Attend to Uncertainty while Mitigating Catastrophic Forgetting", Winter Conference on Applications of Computer Vision (WACV '21), 2021.

WACV-23 Jha, Abhishek*, **Badri N. Patro***, Luc Van Gool, and Tinne Tuytelaars. "Barlow constrained optimization for visual question answering." In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision(WACV '23), pp. 1084-1093. 2023.

BMVC-23 Avideep Mukherjee, **Badri N. Patro**, Vinay P. Namboodiri. "Attentive Contractive Flow: Improved Contractive Flows with Lipschitz-constrained Self-Attention." In Proceedings of the British Machine Vision Conference(BMVC '23), Aberdeen, UK. 2023.

ICASSP-21 Vinod K. Kurmi, Vipul Bajaj, **Badri N. Patro**, Venkatesh K Subramanian, Vinay P. Namboodiri," Collaborative Learning to Generate Audio-Video Jointly", IEEE International Conference on Acoustics, Speech, and Signal Processing.(ICASSP), 2021.

ACMMM-20 Dasgupta, Riddhiman and Tom, Francis and Kumar, Sudhir and Das Gupta, Mithun and Kumar, Yokesh and **Patro, Badri N.** and Namboodiri, Vinay,"Visually Precise Query", Proceedings of the 28th ACM International Conference on Multimedia (MM '20), Seattle, USA, 2020.

ICACCI-14 **Badri N. Patro**, "Design and implementation of novel image segmentation and BLOB detection algorithm for real-time video surveillance using DaVinci processor", International Conference on Advances in Computing, Communications and Informatics ICACCI, pp. 1909-1915, India, Sept 2014.

Under Submission

Under Submisssion   Patro, Badri N., Vinay P. Namboodiri, and Vijay Srinivas Agneeswaran. "SpectFormer: Frequency and Attention is what you need in a Vision Transformer." arXiv preprint arXiv:2304.06446 (2023).

Under Submisssion   Patro, Badri N., and Vijay Srinivas Agneeswaran. "A Critical Analysis of Explanation Techniques for Vision Transformers" 2023.

Under Submisssion   Naveen Panwar, Patro, Badri N., Krishna C. S. and Vijay Srinivas Agneeswaran. "BM25 Triumph RAG: A Critical Examination of Retrieval and Chunking in RAG" 2023.

Under Submisssion	Patro, Badri N., and Vijay Agneeswaran. "Efficiency 360: Efficient Vision Transformers." arXiv preprint arXiv:2302.08374 (2023).
Under Submisssion	Patro, Badri N., and Vijay Srinivas Agneeswaran. "Scattering Vision Transformer: Spectral Mixing is What Matters in Transformers." arXiv preprint arXiv:2304.06446 (2023).
Under Submisssion	Deepankar Srivastav, Patro, Badri Narayana, and Vinay P. Namboodiri. "Look Deeper Count Richer: Depth Based Graph Relation Network for Vqa." Available at SSRN 4063413.

Workshop Publication

WACV-22	Kumar, Sumit, **Badri N. Patro**, and Vinay P. Namboodiri."Auto QA: The Question Is Not Only What, but Also Where" In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 272-281. 2022.
WACV-22	Gupta, Vivek, **Badri N. Patro**, Hemant Parihar, and Vinay P. Namboodiri. "VQuAD: Video Question Answering Diagnostic Dataset." In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 282-291. 2022.
ICCVW-19	**Badri N. Patro**, Shivansh Patel, Vinay P. Namboodiri,"Granular Multimodal Attention Networks for Visual Dialog", ICCV Workshop (ISV), Seoul, South Korea, 2019. (8 page paper,**Oral**)
ICCVW-19	Soumik Dasgupta, **Badri N. Patro**, Vinay P. Namboodiri,"Dynamic Attention Networks for Task Oriented Grounding", ICCV Workshop (ISV), Seoul, South Korea, 2019. (8 page paper,**Oral**)
ICCVW-19	**Badri N. Patro**, Sandeep Kumar, Vinod K. Kurmi, Vinay P. Namboodiri,"Multimodal Differential Network for Visual Question Generation", ICCV Workshop (CLVL), Seoul, South Korea, 2019. (4 page paper, **Spotlight**)
ICCVW-19	**Badri N. Patro***, Vinod K. Kurmi*, Sandeep Kumar*, Vinay P. Namboodiri,"Learning Semantic Sentence Embeddings using Pair-wise Discriminator", ICCV Workshop (CLVL), Seoul, South Korea, 2019. (4 page paper, **Spotlight**)
ICCVW-19	**Badri N. Patro**, Mayank Lunayach, Shivansh Patel, Vinay P. Namboodiri,"U-CAM: Visual Explanation using Uncertainty based Class Activation Maps", ICCV Workshop (LINGIR), Seoul, South Korea, 2019. (2 page paper)
ICCVW-19	**Badri N. Patro**, Vinay P. Namboodiri,"Differential Attention for Visual Question Answering", ICCV Workshop (LINGIR), Seoul, South Korea, 2019. (2 page paper)

Industrial Conference and Journal Publication

Journal Publication

MSJAR-23	Badri N. Patro, Vijay Srinivas Agneeswaran, "SpectFormer: Frequency and Attention is what you need in a Vision Transformer.", Microsoft Journal of Applied Research(MSJAR), July 2023.
MSJAR-23	Pranya Lohiya, Suhas Ranganath, Naveen Panwar, Vini Dixit, Badri N. Patro, "Detection and Remediation of Profanity in Natural Language and Generative Models in Microsoft Azure Customer Support System.", Microsoft Journal of Applied Research(MSJAR), July 2023.

MSJAR-23   Pranya Lohiya*, Badri N. Patro*, and Naveen Panwar*, "SPACE: Spatial Saliency Explanation for Time Series Models", Microsoft Journal of Applied Research(MSJAR), vol. 17, page:250-257, Jan 2023.

Conference

MLAD-23   **Badri N. Patro**, Vijay Srinivas Agneeswaran, "SpectFormer: Frequency and Attention is what you need in a Vision Transformer.", Machine Learning, AI & Data Science Conference (MLAD), Redmond, USA, 2023.

MLAD-23   **Badri N. Patro**, Vijay Srinivas Agneeswaran, "A framework for Model agnostic and Model Dependent explanation method for Vision Transformer.", Machine Learning, AI & Data Science Conference (MLAD), Redmond, USA, 2023.

MLAD-22   Pranya Lohiya*, Badri N. Patro*, and Naveen Panwar*,."SPASE: Spatial Saliency Explanation for Time Series Models.", Machine Learning, AI & Data Science Conference (MLAD), Redmond, USA, 2022.

Industrial and Academic Work Experience

2022–Present
(June–today)   **Microsoft, India**, *Senior Research Scientist*, working on Transformer and Large Language models (LLMs) for Vision, Language, and Speech, under the supervision of Dr. Vijay. Agneeswaran.
- Project Verbatim Rating Classifier(VRC): Deployed VRC using ML-based models, Transformer based models, and LLM based Models in the production pipeline
- Project Commercial Licensing Bot(CLB): Deployed (CLB) module using Large Language Model (LLM) Based Generative Question Answering and provided endpoints for customers.
- Project Fine-grained Image Classifier: Implemented Fine-grained Image Classifiers on customer data to classify the images into different classes.
- Project E5 Security Revenue Risk Model: Implemented Inference for E5 Security Revenue Risk Model and provided an explanation for the time series prediction.
- Project SpectFormer: AML Production Repo for SpectFormer Model
- Project Climax-2 (Research): Implemented and fine-tuned transformer model for Climate data with Prof. Max Welling (Vice President and Distinguished Scientist, MSR-Netherland)
- Project Globally Rain Fall Prediction (Research): Implemented deep models for Global Rain Fall Prediction using Indian Rain Fall data.

2021–22
(Mar–June)   **KU Leuven,Belgium**, *Postdoctoral Researcher*, working with the Multimedia Representation and Interaction group, under the supervision of Prof. Luc. Van Gool.
- Representation learning in Multimodal tasks.
- Fairness/Robustness in Multimodal tasks.
- Explainable AI task.

2021-2021
(Jan–March)   **IIIT Hyderabad India (R&D)**, *Postdoc Researcher* , worked with Bhasha NLP group, under the supervision of Prof. C.V. Jawahar.
- Modeled a multimodal transformer for Multi-Lingual Machine translation task.
- We use self-Supervision techniques using sentence paraphrasing.

2020–2020
(Feb–Dec)   **Google Research India,**, *Postdoctoral Researcher*, worked with Vision Research Group on Representation and Interpretability task, under the supervision of Dr. Gaurav Aggarwal.
- Perturbation-based explanation techniques for Vision tasks.
- Domain Generalisation task.

2019–2019 (May–July)	**Microsoft India (R&D) Pvt. Ltd.**, *Data Scientist Intern*, worked with Bing-Vision Group, under the supervision of Dr. Mithun Dasgupta.

- Modeled a Multimodal Transformer for active tag prediction.
- Created Multimodal Active Tag Prediction dataset.

2013–2015	**Samsung R&D Institute, Delhi**, *Lead Engineer*, Audio Processing and Multimedia on Tizen D2TV, Delhi, India.

- Design & develop audio processing modules for Visual Impaired people in IPTV. The Module has Dolby digital audio switching feature, audio codec switching, sub-surround sound(5.1, 2.1), language changing in live & PVOD channels of IPTV on Tizen os.
- Designed subtitle, Teletext, and caption modules for MPEG-2 TS & rendered in D2TV.
- Designed and implemented Forward Error Correction algorithm using RTP and RTSP on GStreamer multimedia framework in MVPD Architecture(Streaming, Player, FFmpeg Demuxer, GStreamer).
- Worked on audio and language modules for IPTV at Samsung Electronics, **South Korea**.

2012–2013	**Harman International Limited**, *Associate Software Engineer*, Multimedia and Audio Processing, Pune, India.

- Designed and implemented audio post-processing algorithms (Parametric Equalizer, Doppler Effect, DRC, and SRC) for car audio acoustic system using OMAP3530 processor.

2007–2009	**Larsen & Toubro EmSyS Ltd, Mysore**, *Assistant Software Engineer*, DC to DC and AC to DC Power Converter Designer .

- Designed and developed an end-to-end hardware module for universal input AC-DC Power Converter using Flyback Topology.
- Design and developed an end-to-end hardware module for DC-DC converter using Active Clamp Technology. Also, work on power controller (PD, PID) all the dc-dc modules by generating PWM signal using CPLD.

Teaching Experience

2019	**Tutor**, *ESC201: Introduction to Electronics*, Autumn,IIT Kanpur .
2019	**Teaching Assistant**, *MSO201A: Probability and Statistics*, Winter,IIT Kanpur .
2018	**Tutor**, *ESC201: Introduction to Electronics*, Autumn,IIT Kanpur .
2018	**Teaching Assistant**, *Research Lab Development Committee*, Summer,IIT Kanpur.
2018	**Teaching Assistant**, *Department Post-Graduation Committee*, Winter,IIT Kanpur.
2017	**Teaching Assistant**, *ESC201: Introduction to Electronics*, Autumn, IIT Kanpur.
2017	**Teaching Assistant**, *ESC201: Introduction to Electronics*, Summer,IIT Kanpur .
2017	**Teaching Assistant**, *EE301A: Digital Signal processing*, Winter,IIT Kanpur.
2016	**Teaching Assistant**, *EE601A: Image Signal processing*, Autumn,IIT Kanpur.
2012	**Research Assistant**, *Texas Instrument-Digital Signal processing Lab*, IIT Bombay.

PhD Thesis Work

Title	**Towards Understanding vision and language systems: Controllability, Uncertainty & Interpretability for VQA and VQG** .
	Supervisors -:Prof. Vinay P. Namboodiri. (IIT Kanpur)

Description.
- Understanding vision and language system based on Visual Question Answering task, Visual question generation task, and Visual Dialog using Controllability, Uncertainty & Interpretability techniques.
- In the controllability method, we propose an exemplar base deep network for VQA and VQG. Also, we use similarity kind of exemplar concept in the loss function for paraphrase question generation.
- In the uncertainty method, we propose an uncertainty-based method to improve attention and explain answer prediction in VQA and VQG.
- In the interpretability method, we propose an adversarial explanation method to improve attention in VQA. Also, we robust the explanation method to analyze predicted answers in VQA.

Masters Thesis

Title **Real-Time Video and Image Processing for Object Tracking using DaVinci Processor**.
Supervisors -:Prof. V. Rajbabu. (IIT Bombay)

Description.
- Developed a novel segmentation-based BLOB detection algorithm for target tracking using DaVinci Multimedia DSP Processor(DM6437) for the fixed surveillance camera.
- Developed a novel object detection-based algorithm to solve the target tracking problem. This algorithm is based on the segmentation of Binary Large Objects based on Neighborhoods pixels.
- The segmented objects are tracked using the Center of the mass-based tracking algorithm.
- The algorithm is demonstrated for multiple ball tracking using DaVinci Multimedia DSP Processor(DM6437) for a fixed surveillance camera.

Bachelor Project

Title **Bit Error Rate Analysis of Multi-Carrier CDMA**.
Supervisors -: Rakesh Roshaon — Electronics & Telecommunications, NIST.

Description.
In this project, we have derived bit error rate for multi-path frequency selective fading with Gaussian noise for Multi Code-Multi Carrier-CDMA.

Technical Projects

2019 **Auto QA** , *The Question Is Not Only What, but Also Where*.
- We investigate on localization-based question answering task in the context of autonomous driving where this functionality is important.
- We provide a new dataset, Auto-QA. Our new dataset is built over the Argoverse dataset and provides a truly multi-modal setting with seven views per frame and point-cloud LIDAR data being available for answering a localization-based question.

2019 **VQuAD**, *Video Question Answering Diagnostic Dataset*.
- We investigate the task of Video based Question Answering.
- We provide a diagnostic dataset that can be used to evaluate the extent of reasoning abilities of various methods for solving this task.

2019 **MHD**, *Multimodal Humor Detection Dataset*.
- We Provide the Multimodal Humor Dataset (MHD), having textual dialogues with the corresponding video counterparts.
- We present a large scale manually annotated dataset for a comprehensive multimodal understanding of visual humor.

2018 **GAN**, *Granular Attention Network for Visual Dialog*.
- We investigate various attention models proposed in the past. However, the scale at which attention needs to be applied has not been well examined.
- We provide a new method Granular Multi-modal Attention, where we aim to particularly address the question of the right granularity at which one needs to attend while solving the Visual Dialog task.

2018 **DAN**, *Dynamic Attention Network for Task-Oriented Grounding*.
- We provide a novel Dynamic Attention Network architecture for the efficient multi-modal fusion of text and visual representations which can generate a robust definition of state for the policy learner.
- Our model assumes no prior knowledge from visual and textual domains and is an end to end-trainable.
- We show that Dynamic Attention helps in achieving grounding and also aids in the policy learning objective.

Course Projects

2016 Visual Question Answering.(Computer Vision)

2016 Object Recognition and Localization.(Selected Topics of Image Processing)

2016 Direction of Arrival Based Spatial Co-variance Model For Blind Source Separation. (Speech Signal Processing)

2016 Robust Video Stabilization Based on Particle Filter Tracking of Projected Camera Motion. (Video Processing)

2011 Run length encoding, Barrel Shifter, floating point adder & Bus behavior design Projects using VHDL and Verilog. (VLSI Design Lab)

2010 SENSE: Sensitive Encoding technique for Fast MRI using Back Projection. (Medical Image Processing)

2010 An Semi-Autonomous, External Command Reading White line Follower Robot. (Embedded System-Robotics)

2010 Adaptive Beamforming using microphone array for hands-free Telephony with the help of generalized side lobe technique. (Adaptive Signal Processing)

2010 Detection of Duplicate Forgery in Handwritten Signature using Statistical DWT & EDM.(Wavelet Transform)

2009 Frequency Code(LFM) and Phase code(Barker code) Pulse Compression Techniques in Mono Pulse Radar.(Digital Signal Processing)

Industrial Workshops

2017 Summer school on advance computer vision using Deep learning (DL for vision and language, DL for videos, object detection, semantic segmentation, Domain Adaption, and advances in 3D (IIITH).

2017 Summer School on Machine Learning using Deep Learning (Optimization for DL, GAN, VAE, DL for RL and game theory)(IIITH).

2016 Mysore Park Workshop on Vision, Language and AI (Video Caption, guided LSTM, GAN, Adversarial auto-encoders, reinforcement learning, deep contextual models)(VLAI 2016, Mysore).

2016 Summer School on computer vision using Deep Learning(CNN, RNN, Auto-encoder, optimization for DL, Symbolic DL & Face, Pose and Egocentric action recognition, model compression)(IIITH).

2012 Audio Engineering(Acoustics, Recording, Broadcasting Technology, Surround Sound, Microphones& Speakers) & Audio Post Processing (Harman International).

Academic Talks/Seminars

2020 Delivered presentation on "Towards Understanding Vision and Language Systems: Controllability, Uncertainty and Interpretability for VQA and VQG" at CDS, **IISC** Bangalore, India.

2020 Delivered presentation on Explanation vs Attention: A Two-Player Game to Obtain Attention for VQA' at **AAAI** in Hilton, Newyork USA.

2019 Presented poster on 'U-CAM: Visual Explanation using Uncertainty based Class Activation Maps"' at **ICCV** conference in Seoul, South Korea.

2019 Delivered oral presentation on 'Dynamic Attention Networks for Task Oriented Grounding' at **ICCV ISV** workshop in Seoul, South Korea.

2019 Delivered spotlight presentation on 'Multimodal Differential Network for Visual Question Generation' at **ICCV CLVL** workshop in Seoul, South Korea.

2019 Delivered spotlight presentation on 'Learning Semantic Sentence Embeddings using Pair-wise Discriminator"' at **ICCV CLVL** workshop in Seoul, South Korea.

2019 Delivered Poster presentation on Differential Attention for Visual Question Answering' at **ICCV ISV** workshop in Seoul, South Korea.

2018 Presented Poster on Visual Question Answering in **ICVGIP** conference at Hyderabad (India).

2018 Delivered talks on 'Computer Vision and Image Processing' at "TEQIP training session", IIT Kanpur(India)

2018 Presented poster on 'Learning Semantic Sentence Embeddings using Pair-wise Discriminator' at **COLING** conference in Santa Fe, New Mexico, USA.

2018 Presented poster on 'Multimodal Differential Network for Visual Question Generation' at **EMNLP** conference in Brussels, Belgium.

2018 Presented Poster on 'Differential Attention Network Visual Question Answering' at **CVPR** conference in Salt Lake City, Utha, USA.

2018 Delivered State-of-the-Art seminar on 'Visual Question Answering and Visual Question Generation' at Electrical Engineering **IIT Kanpur**

2017 Presented poster on Visual Question Answering in **Advance Computer Vision using Deep Learning** at IIIT Hyderabad (India).

2016   Delivered talks on '**Basics of Deep Learning Platforms (Torch, Caffe, Keras, Tensorflow**' at IIT Kanpur (India)

Industrial Presentation

2014   MPEG-2 Transport Stream Standard (ISO/IEC-13818-1)– PAT, PMT, Descriptor, Section, TS, PES and ES information (Samsung R&D)

2014   ATSC System Information Standard–A/53 part-1, A/65 and CEA-708,608 for Close Caption Decoder (Samsung R&D)

2014   DVB Service Information Standard –EN 300468 and EN-300743 Subtitle Decoder (Samsung R&D)

2013   Digital Audio Processing–Audio Representation, Compression, Microphones, and Speakers module and Audio post processing (Samsung R&D)

2013   Forward Error Correction Techniques– Uneven Length Protection(Samsung R&D)

2012   DSP algorithm and Filter Design– FIR/IIR digital filter and transform technique(DFT, DCT, DST, FFT and Wavelet) (Harman International)

Github

2023   **SpectFormer: Frequency and Attention is what you need in a Vision Transformer**.
- https://badripatro.github.io/SpectFormers/

2023   **Efficient360: Efficient Vision Transformer**.
- https://github.com/badripatro/efficient360

2023   **Barlow constrained optimization for Visual Question Answering**.
- https://github.com/abskjha/Barlow-constrained-VQA

2021   **Multimodal Humor Dataset: Predicting Laughter tracks for Sitcoms**.
- https://delta-lab-iitk.github.io/Multimodal-Humor-Dataset/

2020   **Robust Explanations for Visual Question Answering**.
- https://github.com/DelTA-Lab-IITK/CCM-WACV

2020   **Explanation vs Attention: A Two-Player Game to Obtain Attention for VQA**.
- https://delta-lab-iitk.github.io/TwoPlayer/

2019   **PDUN:Probabilistic framework for solving Visual Dialog**.
- https://delta-lab-iitk.github.io/PDUN/

2019   **U-CAM: Visual Explanation using Uncertainty based Class Activation Maps**.
- https://delta-lab-iitk.github.io/U-CAM/

2018   **Multimodal Differential Network for Visual Question Generation**.
- https://badripatro.github.io/MDN-VQG/

2018   **Learning Semantic Sentence Embeddings using Pair-wise Discriminator**.
- https://badripatro.github.io/Question-Paraphrases/

2018   **Differential Attention for Visual Question Answering**.
- https://badripatro.github.io/DVQA/

Technical skills

Deep learning	Pytorch, Torch, Tensorflow.
Language :	Lua, Python, C, C++, VHDL, Verilog, MATLAB.
Processor :	DM6437, DM6467, OMAP3530, C5510, MSP430, PIC, u8059.
Tools :	Source Insight, LaTeX, Rhapsody, Perforce, Beyond Compare.
IDE :	Code Composer Studio, Xilinx, GHDL, Icurus Verilog, Keil, Sublime.
Analyzer :	Audacity, Praat audio analysis, Eagle, Pspice.

Student Volunteer Awards

2019	Received Student Volunteer Award from ICCV 2019.
2018	Received Student Volunteer Award from EMNLP 2018.
2018	Received Student Volunteer Award from CVPR 2018.

Travel Grant Awards

2020	Received Partial Conference Travel Grant from WACV for WACV 2020.
2020	Received Partial Conference Travel Grant from IIT Kanpur India for WACV 2020.
2020	Received Partial Conference Travel Grant from Google India for AAAI 2020.
2020	Received Partial Conference Travel Grant from Microsoft India for AAAI 2020.
2019	Received Partial Conference Travel Grant from ICCV for ICCV 2019.
2019	Received Partial Conference Travel Grant from IIT Kanpur for ICCV 2019.
2018	Received Partial Conference Travel Grant from EMNLP for EMNLP 2018.
2018	Received Partial Conference Travel Grant from ACM India for EMNLP 2018.
2018	Received Partial Conference Travel Grant from Microsoft India for EMNLP 2018.
2018	Received Conference Travel Grant from IIT Kanpur for COLING 2018.
2018	Received Conference Travel Grant from Google India for CVPR 2018.
2017	Selected in Quiz competition in Deep learning summer school for Computer Vision.
2017	Selected in Quiz competition in Deep learning summer school for Machine Learning.

Professional Service

2019	Participated in Amazon Research Days at Bangalore sponsored by Amazon India.
2018	Served as a reviewer in conferences (such as CVPR, ICCV, ECCV, NeurIPS, ICLR, AAAI, ACL, EMNLP, NAACL, BMVC, WACV, ICVGIP, NCVPRIPG)
2018	Served as a reviewer in Journals (such as Transactions on Pattern Analysis and Machine Intelligence (T-PAMI), Transactions-Image Processing (T-IP), Pattern Recognition(PR))
2017	Participated in ACM-MSR Summit 2018 at IIITH sponsored by Microsoft India
2016	Participated in the Mysore Park workshop on Vision, Language and AI, sponsored by Google and Infosys.