

Badri N. Patro

Delta Lab, CC-202

IIT Kanpur-208016

☎ (+91) 9076237295

FAX patrobadi.iitb@gmail.com

✉ badri@iitk.ac.in

📄 <https://badripatro.github.io/>



Research Interests: Computer Vision, Natural language Processing & Machine Learning.

Education

- 2021–Present **Postdoctoral Researcher**, *KU Leuven, Belgium*, Working in Computer Vision and Natural language Processing..
- 2020–2021 **Postdoctoral Researcher**, *Google Research India, Bangalore*, Working in Computer Vision and Natural language Processing..
- 2015–2020 **Doctor of Philosophy in Electrical Engineering**, *Indian Institute of Technology Kanpur, India*, Specialized in Signal Processing, Communications & Networks..
- 2009–2012 **Masters of Technology in Electrical Engineering**, *Indian Institute of Technology Bombay, India*, Specialized in Communication & Signal Processing.
- 2003–2007 **Bachelor of Technology in Electronic & Tele Communication Engineering**, *National Institute of Science and Technology, Brahmapur, Orissa*.

Doctoral dissertation award

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

Journal Publication

- 2020 **Badri N. Patro**, Mayank Lunayach, Vinay P. Namboodiri, "Uncertainty-CAM: Visual Explanation using Uncertainty based Class Activation Maps", **IEEE Transactions on Image Processing (TIP)**, 2020.
- 2020 **Badri N. Patro**, Anupriy, Vinay P. Namboodiri, "Probabilistic framework for solving Visual Dialog", **Pattern Recognition**, Pages:107586, 2020.
- 2020 **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".
- 2021 **Badri N. Patro**, Vinod K. Kurmi, Sandeep Kumar, Vinay P. Namboodiri, "MUMC: Minimizing Uncertainty of Mixture of Cues", **Image and Vision Computing**, 2020, (Under Revision).
- 2021 **Badri N. Patro**, Anupriy, Vinay P. Namboodiri, "Adversarial Explanation : A Two-Player Game to obtain Attention for VQA", **Pattern Recognition**, 2021 (Under Revision).

Conference Publication

- 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.
- 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.
- 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.
- 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-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.
- 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.
- 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.
- 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.
- 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.
- 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.

- 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.

Workshop Publication

- 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)

Work Experience

- 2019 **Microsoft India (R&D) Pvt. Ltd.**, *Data Scientist Intern*, Bing-Vision, Hyderabad.
(May–July)
 - Design & developed 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 & developed 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. **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.
 - Design 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 .
- Design 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, also controlled (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 Image-based Question Answering module, question generation module, and Visual Dialog using Controllability, Uncertainty & Interpretability.
- 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 explaining answer prediction in VQA and VQG.
- In the interpretability method, we propose an adversarial explanation method to improve attention in VQA. Also, we robust explanation method to analysis predicted answers in VQA.

Masters Thesis

- Title **Real-Time Video 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 segmentation of Binary Large Objects based on Neighborhoods pixel.
- The segmented objects are tracked using the Center of 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 Roshan — 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 Course Projects

- 2019 VQuAD: Video Question Answering Diagnostic Dataset.(Vision & Language)
- 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

- 2020 **Robust Explanations for Visual Question Answering.**
○ <https://github.com/Delta-Lab-IITK/CCM-WACV>
- 2019 **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, AAAI, ACL, NAACL, TPAMI, TIP, PR, BMVC, WACV, ICVGIP, NCVPRIPG)
 - 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.
-