

Yash Goyal

CONTACT INFORMATION

Apt 12, 706 Appalachian Drive, Blacksburg, VA 24060
Phone: +1 540-449-7761
Webpage: www.computing.ece.vt.edu/~ygoyal
E-mail: ygoyal@vt.edu

EDUCATION

Virginia Tech, Blacksburg, VA, USA **August 2014 - Present**
Ph.D. in Computer Engineering
CGPA: 4.0/4.0
Advised by: Dhruv Batra
Research Interests: Computer Vision and Natural Language Processing

Indian Institute of Technology Gandhinagar, Gujarat, India **May 2014**
Bachelor of Technology (B. Tech.) in Electrical Engineering
with Honors, and Minor in Computer Science and Engineering
CPI: 9.33/10

CONFERENCE AND WORKSHOP PUBLICATIONS

Yash Goyal*, Tejas Khot*, Douglas Summers-Stay, Dhruv Batra, Devi Parikh.
Making the V in VQA Matter: Elevating the Role of Image Understanding in Visual Question Answering.
arXiv preprint arXiv:1612.00837, 2016
*equal contribution

Gordon Christie*, Ankit Laddha*, Aishwarya Agrawal, Stanislaw Antol, Yash Goyal, Kevin Kocherberger, Dhruv Batra.
Resolving Language and Vision Ambiguities Together: Joint Segmentation & Prepositional Attachment Resolution in Captioned Scenes.
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2016;
arXiv preprint arXiv:1604.02125, 2016;
Language & Vision Workshop; Scene Understanding Workshop, CVPR, 2015.
*equal contribution

Yash Goyal, Akrit Mohapatra, Devi Parikh, Dhruv Batra.
Interpreting Visual Question Answering Models.
International Conference on Machine Learning (ICML) Workshop on Visualization for Deep Learning, 2016. (**Best Student Paper**);
arXiv preprint arXiv:1608.08974, 2016;
Visual Question Answering Challenge Workshop; Scene Understanding Workshop, CVPR 2016.

Peng Zhang*, Yash Goyal*, Douglas Summers-Stay, Dhruv Batra, Devi Parikh.
Yin and Yang: Balancing and Answering Binary Visual Questions.
Conference on Computer Vision and Pattern Recognition (CVPR), 2016;
arXiv preprint arXiv:1511.05099, 2015
*equal contribution

Harsh Agrawal, Clint Solomon Mathialagan, Yash Goyal, Neelima Chavali, Prakriti Banik, Akrit Mohapatra, Ahmed Osman, Dhruv Batra.
CloudCV: Large-Scale Distributed Computer Vision as a Cloud Service.
Book Chapter, *Mobile Cloud Visual Media Computing*. Editors: Gang Hua, Xian-Sheng Hua. Springer, 2015;
arXiv preprint arXiv:1506.04130, 2015

Yash Goyal, Harsh Agrawal, Clint Solomon, Dhruv Batra.
CloudCV: Fast Addition of New Classes to a Pre-trained ConvNet.
Mid-Atlantic Computer Vision (MACV) Workshop, 2015.

Deepesh Kumar, Yash Goyal, Sunil Nair, Arvind Chauhan, Uttama Lahiri.
Design of a physiologically informed virtual reality based interactive platform for individuals with upper limb impairment.
IEEE International System on Robot and Human Interactive Communication (Ro-MAN 2014), UK.

SELECTED
PROJECTS

Elevating the Role of Image Understanding in Visual Question Answering *Virginia Tech*
Advisors: Devi Parikh, Dhruv Batra *Summer and Fall 2016*
Description: We counter the language priors present in the popular Visual Question Answering (VQA) dataset (Antol et al., ICCV 2015) and make vision (the V in VQA) matter by balancing the dataset. Our dataset is by construction more balanced than the original VQA dataset and has approximately twice the number of image-question pairs. Our complete balanced dataset will be publicly released as part of the 2nd iteration of the Visual Question Answering Challenge (VQA v2.0). This work is in submission to Conference on Computer Vision and Pattern Recognition (CVPR) 2017.

Interpreting Visual Question Answering Models *Virginia Tech*
Advisor: Dhruv Batra, Devi Parikh *Spring 2016*
Description: We experimented with two visualization methods – guided backpropagation and occlusion – to interpret deep learning models for the task of Visual Question Answering. Specifically, we find what part of the input (pixels in images or words in questions) the VQA model focuses on while answering a question about an image. This work was presented as an oral in International Conference on Machine Learning (ICML) Workshop on Visualization for Deep Learning, 2016.

Balancing & Answering Binary Visual Questions for Abstract Scenes *Virginia Tech, ARL*
Advisors: Devi Parikh, Dhruv Batra, Douglas Summers-Stay *Summer and Fall 2015*
Description: We balance the existing VQA dataset so that VQA models are forced to understand the image to improve their performance. We propose an approach that focuses heavily on vision and answers the question by visual verification. This work was presented in Conference on Computer Vision and Pattern Recognition (CVPR) 2016.

Holistic Scene Understanding via Multiple Structured Hypotheses from Perception Modules *Virginia Tech*
Advisor: Dhruv Batra *March - May 2015*
Description: We present an approach to simultaneously perform semantic segmentation and prepositional phrase attachment resolution for captioned images. We show that our vision and language modules have complementary strengths, and that joint reasoning produces more accurate results than any module operating in isolation. This work has been accepted in Conference on Empirical Methods in Natural Language Processing (EMNLP) 2016.

Fast Addition of New Classes to a Pretrained ConvNet *Virginia Tech*
Advisor: Dhruv Batra *January - March 2015*
Description: We created a demo which enable users to efficiently add new categories to the existing trained caffeNet model with 1000 ImageNet categories in an online fashion, without the need of training the model from scratch. This work was presented as a poster at Mid-Atlantic Computer Vision (MACV) Workshop 2015.

Virtual Reality Based Stroke Rehabilitation Platform *IIT Gandhinagar*
Advisor: Uttama Lahiri *Fall 2013 and Spring 2014*
Description: We created a virtual reality based interactive exercise environment consisting of neuromuscular tasks using haptic device and joystick for stroke patients' rehabilitation in *Vizard*. The

platform investigates the patient's physiological indices such as electromyogram, heart beat rate, etc. while a participant interacts with it. This work was presented in IEEE International System on Robot and Human Interactive Communication (Ro-MAN 2014), UK .

WORK
EXPERIENCE

Visiting Scholar, Summer 2015

Army Research Lab, Adelphi, MD, USA

Supervisor: Douglas Summers-Stay

- Worked on developing visual question answering models for abstract scenes.
- Designed image features for abstract scenes for visual question answering.

Teaching Assistant, Fall 2014 and Spring 2015

Bradley Department of Electrical and Computer Engineering, Virginia Tech

- Teaching assistant for ECE 2534, Microcontroller Interfacing
- Responsibilities included helping students with lab projects and homeworks, and grading them.

Visiting Scholar, Summer 2013

Duke University, Durham, NC, USA

Supervisor: Prof. Krishnendu Chakrabarty

- Simulated a capacitive sensing circuit for detecting presence of microfluidic droplets on digital biochips on LTSPICE.
- Implemented the circuit on PCB through surface mounting and through-hole mounting.

Engineering Intern, Summer 2012

GridAnt Technologies, Ahmedabad, India

- Worked on accumulation of real-time data from Facebook for natural language processing using Facebook graph API platform.

SIGNIFICANT
ACTIVITIES

Co-organizer of the Visual Question Answering (VQA) Challenge workshop at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

Reviewer for ECCV, CVPR, ICLR.

Regular participation in Computer Vision and Machine Learning Reading Group at Virginia Tech.

Successful Participant, The Mathematical Contest in Modelling by COMAP.

HONORS AND
AWARDS

Best Student Paper, Workshop on Visualization for Deep Learning, ICML 2016.

Registration Fee Waiver Award, Deep Learning Summer School 2016, Montreal, Canada.

Dean's List for 7 semesters, IIT Gandhinagar.

S.P. Mehrotra scholarship, 2011-12, IIT Gandhinagar.

Merit-cum-Means scholarship, 2010-11, 2012-14, IIT Gandhinagar.

0.1% merit certificate in Mathematics, class XII board, CBSE.

RELEVANT
COURSEWORK

Graduate: Advanced Computer Vision, Computer Vision, Deep Learning for Perception, Intro to Machine Learning, Data Analytics, Intro to Artificial Intelligence.

Undergraduate: 3D Computer Vision, Fundamentals of Artificial Neural Networks, Intro to Python and Scientific Computing, Artificial Intelligence, Markov Chains and Queueing Models, Intro to Optimization, Probability and Random Processes, Intro to Computational Complexity Theory.

TECHNICAL SKILLS

Programming Languages: Python, MATLAB, Lua, C/C++

Libraries: Torch, Caffe, Keras, OpenCV, Vizard

Human Computation: Amazon Mechanical Turk