

SHIKHAR SHARMA

shikhar.sharma@microsoft.com

Senior Research SDE, Microsoft Research

<http://www.shikharsharma.com/>

RESEARCH INTERESTS

- Deep Learning
- Adversarial Learning
- Machine Learning
- Computer Vision
- Recurrent Neural Networks
- Task-oriented Dialogue

RESEARCH STATEMENT SUMMARY

My long-term research interest lies in advancing the capabilities of current artificial intelligence systems. I am highly interested in neural network architectures augmented with memory and attention mechanisms.

My recent research deals with image synthesis and visually-grounded multi-modal dialogue systems, which stand to gain significant advantage from using memory and attention similar to how we as humans do. Prior to this, I have worked on visual attention models for action recognition and video description. I have broad interests in deep learning research in natural language processing and computer vision, which I believe are important research areas on the path to true artificial intelligence.

EDUCATION

University of Toronto

M.Sc. in Computer Science with the Deep Learning and Machine Learning group

Aug'14 - Feb'16

GPA: 3.83/4

- Thesis title: *Action Recognition and Video Description using Visual Attention*
- Thesis supervisor: *Prof. Ruslan Salakhutdinov*

Indian Institute of Technology (IIT), Kanpur

B.Tech. in Computer Science

Jul'10 - Jun'14

GPA: 9.3/10

- Thesis title: *Speech Recognition using Deep Belief Networks and Hidden Markov Models*
- Thesis supervisor: *Prof. Harish Karnick*

PUBLICATIONS

[†]equal contribution

Theses

- [Shikhar Sharma](#). 2016. "Action Recognition and Video Description using Visual Attention." Masters Thesis, University of Toronto

Journal papers

- Ashesh Jain, [Shikhar Sharma](#), Thorsten Joachims, Ashutosh Saxena. 2015. "Learning preferences for manipulation tasks from online coactive feedback." *The International Journal of Robotics Research (IJRR)*, 34, 1296-1313

Conference and Workshop papers

- Alaaeldin El-Nouby, [Shikhar Sharma](#), Hannes Schulz, Devon Hjelm, Layla El Asri, Samira Ebrahimi Kahou, Yoshua Bengio, Graham W. Taylor. 2019. "Tell, Draw, and Repeat: Generating and Modifying Images Based on Continual Linguistic Instruction." *International Conference on Computer Vision (ICCV)*
- Mehdi Fatemi, [Shikhar Sharma](#), Harm van Seijen, Samira Ebrahimi Kahou. 2019. "Dead-ends and Secure Exploration in Reinforcement Learning." *International Conference on Machine Learning (ICML)*
- Dat Tien Nguyen, [Shikhar Sharma](#), Hannes Schulz, Layla El Asri. 2019. "From FiLM to Video: Multi-turn Question Answering with Multi-modal Context." *Association for the Advancement of Artificial Intelligence (AAAI): Dialog System Technology Challenge (DSTC) Workshop*

- Alaaeldin El-Nouby, [Shikhar Sharma](#), Hannes Schulz, Devon Hjelm, Layla El Asri, Samira Ebrahimi Kahou, Yoshua Bengio, Graham W. Taylor. 2018. “Keep Drawing It: Iterative language-based image generation and editing.” Neural Information Processing Systems ([NIPS](#)): Visually Grounded Interaction and Language (ViGIL) Workshop
- [Shikhar Sharma](#), Dendi Suhubdy, Vincent Michalski, Samira Ebrahimi Kahou, Yoshua Bengio. 2018. “Chat-Painter: Improving Text to Image Generation using Dialogue.” International Conference on Learning Representations ([ICLR](#)) Workshop
- Hannes Schulz[†], Jeremie Zumer[†], Layla El Asri, [Shikhar Sharma](#). 2017. “A Frame Tracking Model for Memory-Enhanced Dialogue Systems.” Proceedings of the 2nd Workshop on Representation Learning for NLP ([RepL4NLP](#))
- Layla El Asri[†], Hannes Schulz[†], [Shikhar Sharma](#)[†], Jeremie Zumer[†], Justin Harris, Emery Fine, Rahul Mehrotra, Kaheer Suleman. 2017. “Frames: A Corpus for Adding Memory to Goal-Oriented Dialogue Systems.” Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue ([SIGDIAL](#))
- [Shikhar Sharma](#), Jing He, Kaheer Suleman, Hannes Schulz, Philip Bachman. 2017. “Natural Language Generation in Dialogue using Lexicalized and Delexicalized Data.” International Conference on Learning Representations ([ICLR](#)) Workshop
- [Shikhar Sharma](#), Ryan Kiros, Ruslan Salakhutdinov. 2016. “Action Recognition using Visual Attention.” International Conference on Learning Representations ([ICLR](#)) Workshop
- Ashesh Jain, [Shikhar Sharma](#), Ashutosh Saxena. 2016. “Beyond geometric path planning: Learning context-driven trajectory preferences via sub-optimal feedback.” Robotics Research: The 16th International Symposium ([ISRR](#)). Springer International Publishing, 319-338
- [Shikhar Sharma](#), Ryan Kiros, Ruslan Salakhutdinov. 2015. “Action Recognition using Visual Attention.” Neural Information Processing Systems ([NIPS](#)): Time Series Workshop

Preprints

- [Shikhar Sharma](#), Layla El Asri, Hannes Schulz, Jeremie Zumer. 2017. “Relevance of Unsupervised Metrics in Task-Oriented Dialogue for Evaluating Natural Language Generation.” [arXiv:1706.09799](#) [cs.CL]

PATENTS

- Simultaneous dialogue state management using frame tracking. Justin Harris, Layla El Asri, Emery Fine, Rahul Mehrotra, Hannes Schulz, [Shikhar Sharma](#), Jeremie Zumer. U.S. Patent No. US 10,431,202 B2. Granted: October 01, 2019. Expires: June 20, 2037.
- Natural language generation in a spoken dialogue system. [Shikhar Sharma](#), Jing He, Kaheer Suleman, Philip Bachman, Hannes Schulz. U.S. Patent No. US 10,242,667 B2. Granted: March 26, 2019. Expires: June 02, 2037.

RESEARCH EXPERIENCE

Microsoft Research
Senior Research SDE
Research SDE II

Montreal, Quebec, Canada
Sep'19 - Present
Feb'17 - Aug'19

After Maluuba was acquired by Microsoft, I have done research on evaluation of natural language generation systems for task-oriented dialogue using unsupervised metrics. I have also worked on end-to-end trainable dialogue systems. I am currently working on systems that generate images based on text inputs (a caption or a dialogue) and spatial layouts. Apart from these, I have also worked on projects to integrate state-of-the-art research into Microsoft products.

Maluuba Research*Research Scientist**Montreal, Quebec, Canada**Mar'16 - Feb'17*

I have worked on end-to-end trainable models for task-oriented dialogue systems with the dialogue group on Maluuba's Frames dataset. Prior to this, I worked on improving the natural language generation component of task-oriented dialogue systems over the then state-of-the-art by using lexicalized slot values aligned with delexicalized slots.

University of Toronto*Graduate Research Assistant**Toronto, Ontario, Canada**Sep'14 - Feb'16*

Working together with Prof. Ruslan Salakhutdinov and Jamie Ryan Kiros, my masters thesis focused on soft-attention based recurrent neural networks for action recognition and video description generation.

Cornell University*Research Intern**Ithaca, New York, USA**May'13 - Jul'13*

I worked with Prof. Ashutosh Saxena and Ashesh Jain on learning preferences over trajectories on robots such as the Baxter. Our approach required a non-expert user for training and the preferences we learned were governed by objects and human interactions in the environment. Our work was featured on TechCrunch and various other media websites.

IIT Kanpur + Microsoft Research, Redmond*Research Intern**IIT Kanpur, India**Apr'12 - Jul'12*

I worked with Dr. Sumit Gulwani, Prof. Amey Karkare and Prof. Subhajit Roy on creating an Intelligent Tutoring System for teaching First Order Logic. We utilized the Stanford CoreNLP Parser to parse and retrieve word-dependencies and generated logical first order formulae.

AWARDS AND SCHOLASTIC ACHIEVEMENTS

2014	Department Entrance Scholarship (CS) of CAD\$10,000 at University of Toronto for 2014-2016
2013	Selected for the prestigious Cornell IIT-Internship Program, 2013
2012	IIT Kanpur Academic Excellence Award for 2011-12 for distinctive academic achievements
2011	O.P. Jindal Engg. and Mgmt. Scholarship for excellence in academics and leadership
2011	IIT Kanpur Academic Excellence Award for 2010-11 for distinctive academic achievements
2010	Awarded the CBSE Merit Scholarship for Professional Studies - AIEEE for 2010-2014
2010	All India Rank 434 in IIT-Joint Entrance Examination (amongst 0.48 million candidates)

KEY RESEARCH PROJECTS**Tell, Draw, and Repeat: Continual Text-based Image Generation***Sep'18 - Feb'19**Dialogue group**Microsoft Research, Montreal, Canada*

- Implemented a recurrent-GAN architecture for generating and modifying images based on continual text input
- Introduced Generative Neural Visual Artist (GeNeVA) task for iterative text-based image generation
- Introduced two datasets and a relational similarity metric for the GeNeVA task

Blog post <https://www.aka.ms/blog-geneva>**Project page** <https://www.aka.ms/project-geneva>**Source code** <https://github.com/Maluuba/GeNeVA>**Datasets** https://github.com/Maluuba/GeNeVA_datasets**Evaluation of Task-oriented Natural Language Generation systems***Feb'17 - May'17**Dialogue group**Microsoft Research, Montreal, Canada*

- Implemented baselines and hierarchical recurrent models to generate natural language
- Set up a pipeline for evaluating generated sentences on word-overlap-based and word embeddings-based metrics
- Performed analysis of correlation between human scores and unsupervised metrics

Blog post <https://www.aka.ms/blog-nlg-eval>
Source code <https://github.com/Maluuba/nlg-eval/>

A Dialogue System from Human-Human Dialogues

Jun'16 - Feb'17

Dialogue group

Maluuba Research, Canada

- Organized and supervised Frames data collection and annotation with data analysts: 19,986 turns in total
- Generalized the state tracking task to frame tracking, which requires adding memory
- Designed and implemented tools to parse and preprocess the dataset
- Implemented a baseline and a natural language understanding module which extracts, anonymizes slot values

Project page <https://datasets.maluuba.com/Frames/>

Action Recognition using Visual Attention

Feb'15 - Aug'15

Research Project with Prof. Ruslan Salakhutdinov

University of Toronto, Canada

- Proposed a soft attention model to perform action recognition in video datasets like HMDB-51, Hollywood2
- Used GoogLeNet to extract features and an attention mechanism to dynamically pool the features
- Performed the classification task using a Recurrent Neural Network based on Long Short-Term Memory units
- Compared our models with relevant baselines and state-of-the-art Deep Learning based models
- Presented an extensive qualitative analysis of possibly how the attention model selected glimpses

Project page <http://shikharsharma.com/projects/action-recognition-attention/>

Source code <https://github.com/kracwarlock/action-recognition-visual-attention/>

Mentions Talk by Prof. Ruslan Salakhutdinov at Deep Learning Summer School, Montreal 2016

Learning Trajectory Preferences for Manipulators via Iterative Improvement

May'13 - Jul'13

Research Project with Prof. Ashutosh Saxena

Cornell University, USA

- Learnt preferences over trajectories for 7 Degree of Freedom robot manipulators of Baxter Research Robot
- The algorithm requires the user to incrementally improve over the trajectory currently proposed by the robot
- Theoretical regret bounds of our algorithm match the asymptotic rates of optimal trajectory algorithms
- The robot was used to perform grocery checkout tasks

Project page <http://pr.cs.cornell.edu/coactive/>

YouTube demo <https://www.youtube.com/watch?v=uLktpkd7ojA>

Mentions Cornell University CS News (<https://goo.gl/owFPss>),
TechCrunch (<https://goo.gl/y2hWl2>), KurzweilAI (<https://goo.gl/Kyfn9U>)

TECHNICAL SKILLS

Programming Languages	Python, C, C++
Machine Learning tools	PyTorch, Keras, Theano, Tensorflow, Matlab
Web Programming	HTML, PHP, CSS, JavaScript, JQuery, MySQL

INTERNS SUPERVISED

- Tristan Sylvain. Ph.D. Candidate at Université de Montréal.
Co-supervisor(s): Devon Hjelm.
- Nouha Dziri. Ph.D. Candidate at University of Alberta.
Co-supervisor(s): Alessandro Sordoni, Geoff Gordon, Hannes Schulz.
- Alaaeldin El-Nouby. M.Sc. Candidate at University of Guelph.
- Dat Tien Nguyen. Ph.D. Candidate at University of Amsterdam.
Co-supervisor(s): Hannes Schulz, Layla El Asri.
- Eugene Vorontsov. Ph.D. Candidate at Polytechnique Montréal.
Co-supervisor(s): Samira Shabanian.

TEACHING EXPERIENCE

- 2015** Teaching Assistant for CSC411: Introduction to Machine Learning at University of Toronto
- 2015** Teaching Assistant for CSC321: Introduction to Neural Networks at University of Toronto
- 2014** Teaching Assistant for CSC309: Programming on the Web at University of Toronto

REVIEWING ACTIVITY

- 2020** ICML, IJCAI (Program Committee Member), IEEE Transactions on Multimedia (Impact Factor: 5.452)
- 2019** NAACL-HLT, Machine Vision and Applications (Impact Factor: 1.788), Montreal AI Symposium
- 2018** ACL, IEEE Transactions on Multimedia (Impact Factor: 3.509), IEEE Transactions on Circuits and Systems for Video Technology (Impact Factor: 3.599), Montreal AI Symposium, Machine Vision and Applications (Impact Factor: 1.788)

LINKS

- Website** <http://www.shikharsharma.com/>
- GitHub** <https://github.com/kracwarlock>
- LinkedIn** <https://ca.linkedin.com/in/sharmashikhar>
- Scholar** <https://scholar.google.ca/citations?user=bLA7DYwAAAAJ>

VOLUNTEER EXPERIENCE

MariHacks Feb'18
Mentor and Judge *Montreal, Quebec, Canada*

MariHacks was a day-long hackathon for CEGEP and high school students where they had to create revolutionary software projects. I was one of the mentors who helped the learners. I was also one of the judges who had to evaluate the projects on Originality, Usefulness, Polish, and Technical Difficulty.

Ladies Learning Code Sept'15
Mentor *Toronto, Ontario, Canada*

Toronto: Data Insights with Python for Beginners was a hands-on, beginner-friendly python workshop. Beginners learnt basic programming fundamentals and also how to read and manipulate large data sets to draw insights for marketing and analysis. I was one of the mentors who helped the learners.

Lighthouse Labs Feb'15
Mentor *Toronto, Ontario, Canada*

The HTML500 was a one-day event where 50 of Canada's top tech companies came together to teach 500 people how to code, for free. I was one of the mentors who helped out beginners in learning HTML and CSS and creating a web page.

REFERENCES

- Ruslan Salakhutdinov, Associate Professor, Carnegie Mellon University rsalakhu@cs.cmu.edu