

Wilka Torrico De Carvalho, *Aspiring Brain Scientist*

CONTACT INFORMATION

Website: wcarvalho.github.io Github: github.com/wcarvalho
E-mail: wcarvalh@umich.edu Google Scholar

RESEARCH INTERESTS

My long-term goal is to develop cognitive theories of learning that help us understand how humans infer, reason with, and exploit the rich structure present in realistic visual scenes to enable sophisticated behavioral policies. Towards this end, I am currently studying how object-centric representation learning and reinforcement learning can bring us closer to human-level artificial intelligence.

EDUCATION

University of Michigan—Ann Arbor, Ann Arbor, Michigan USA Sep 2018 - Present
School of Engineering, Ph.D. in Computer Science
Advisors: [Honglak Lee](#), [Satinder Singh](#), [Richard Lewis](#)

University of Southern California, Los Angeles, California USA Aug 2015 - May 2017
Viterbi School of Engineering, M.S. in Computer Science
Advisor: [Yan Liu](#)

Stony Brook University, Stony Brook, New York USA Aug 2011 - May 2015
College of Arts and Sciences, B.S. in Physics
Advisor: [Axel Drees](#)

Brooklyn Technical High School, Brooklyn, New York USA May 2007 - May 2011
Diploma in Applied Physics

HONORS & AWARDS

1/200 chosen internationally for Heidelberg Laureate Forum 2018
GEM National Fellowship sponsored by IBM, Adobe 2017, 2018
University of Michigan Rackham Merit Fellowship 2017
ICLR, Neurips Travel Award 2017, 2019
NSF Graduate Research Fellowship (Neuroscience) 2015
Provost Award for Academic Excellence ($\sim 0.5\%$ of graduates chosen) 2015
Researcher of the Month (1 monthly at Stony Brook University) 2014
HHMI Minority Undergraduate Research Fellowship 2014
 $\Sigma\Pi\Sigma$ Physics Honor Society (sponsored by [Alfred Goldhaber](#)) 2013
NSF Louis Stokes Alliance for Minority Participation Scholar 2011
Deans List 2011-2015

CONFERENCE PUBLICATIONS

Sanjay Purushotham*, **Wilka Carvalho***, Tanachat Nilanon, Yan Liu. “*Variational Recurrent Adversarial Domain Adaptation*.” In International Conference on Learning Representations (ICLR), 2017

Wilka Carvalho. “*Modeling a Detection of internally reflected Cherenkov light (DIRC) Particle Detector for High-Multiplicity Collisions*.” In State University of New York Undergraduate Research Conference (SURC), 2015

WORKSHOP PUBLICATIONS

Wilka Carvalho, Anthony Liang, Kimin Lee, Sungryull Sohn, Richard L. Lewis, Satinder Singh, Honglak Lee. “*Reinforcement Learning for Sparse-Reward Object-Interaction Tasks in First-person Simulated 3D Environments*.” In ICML Object-Oriented Learning Workshop, NeurIPS DeepRL Workshop, 2020

Chris Hoang, Sungryull Sohn, Jongwook Choi, **Wilka Carvalho**, Honglak Lee. “*Successor Landmarks for Efficient Exploration and Long-Horizon Navigation.*” In NeurIPS DeepRL Workshop, 2020

Wilka Carvalho, Kimin Lee, Anthony Liang, Ryan Krueger, Richard L. Lewis, Satinder Singh, Honglak Lee. “*Efficiently Learning to Perform Household Task with Object-oriented Exploration.*” In NeurIPS Black In AI Workshop (Neurips BAI), 2019 **(Oral)**

Bryant Chen*, **Wilka Carvalho***, Benjamin Edwards, Taesung Lee, Ian Molloy, Heiko Ludwig. “*Detecting Backdoor Attacks on Deep Neural Networks by Activation Clustering.*” In Artificial Intelligence Safety Workshop at Association for the Advancement of Artificial Intelligence (AAAI AASW), 2018 **(Best Paper)**

Sanjay Purushotham*, **Wilka Carvalho***, Yan Liu. “*Variational Adversarial Deep Domain Adaptation for Health Care Time Series Analysis.*” In NeurIPS Workshop on Machine Learning for Healthcare (Neurips ML4HC), 2016 **(Spotlight)**

PATENTS

Bryant Chen, **Wilka Carvalho**, Heiko Ludig, Ian Molloy, Jialong Zhang, Benjamin Edwards. “*Detecting poisoning attacks on neural networks by activation clustering.*” 2020

Wilka Carvalho, Bryant Chen, Benjamin Edwards, Taesung Lee, Ian Molloy, Jialong Zhang. “*Using Gradients to Detect Backdoors in Neural Networks.*” 2018

Wilka Carvalho, Yan Liu, Tanachat Nilanon, Sanjay Purushotham. “*Effective Knowledge Transfer Among Patient Populations via Deep Learning.*” 2017

INVITED TALKS

Object Representations for Learning and Reasoning Workshop. NeurIPS. (December, 2020)

Cognitive Science Community. University of Michigan. (November, 2020)

Stony Brook Society of Physics. Stony Brook, NY. (April, 2020)

Machine Learning Lunch Seminar. University of Southern California. (April, 2017)

SYMPOSIUM PRESENTATIONS

“Variational Adversarial Deep Domain Adaptation for Healthcare Time Series.” *Southern California Machine Learning Symposium. California Institute of Technology*, Pasadena, CA, 2016. **Runnerup, Best Poster. Worth \$1000 in Amazon AWS credit.**

“Modeling a DIRC Particle Detector for High-Multiplicity Collisions.” *23rd Annual CSTEP Statewide Student Conference*, Bolton Landing, NY, 2015. **2nd Place, Physics and Math.**

“Modeling the Cognitive Process of Attributing Traits to Others.” *Summer Seminar Day. California Institute of Technology*, Pasadena, CA, 2014.

“Modeling Deep Brain Stimulation of Globus Palidus Internus.” *22nd Annual CSTEP Statewide Student Conference*, Bolton Landing, NY, 2014.

“Modeling a Detection of internally reflected Cherenkov light (DIRC) Particle Detector for High-Multiplicity Collisions.” *URECA Celebration of Undergraduate Research & Creativity. Stony Brook University*, Stony Brook, NY, 2014.

“Modeling Deep Brain Stimulation of Globus Palidus Internus.” *Poster Symposium. University of Minnesota*, Minneapolis, MN, 2013.

“Testing Theories in Fluid Dynamics.” *Global Lab Poster Symposium. Suny Oswego*, Oswego, NY, 2012.

RESEARCH
EXPERIENCE

University of Michigan–Ann Arbor, Ann Arbor, Michigan USA

AI Lab, *August 2018 – Present*

Advisors: [Honglak Lee](#), [Satinder Singh](#), [Richard Lewis](#)

Project: Data-Efficient Reinforcement Learning with Object-Oriented Representations

- Developed and implemented novel nonparametric neural network that employs contrastive learning to learn an invariant object recognition function that enables sample-efficient reinforcement learning in complex 3D domains
- Implemented reinforcement learning and computer vision baselines

Microsoft Research, Redmond, Washington USA

Medical Devices Group, *June 2018 – August 2018*

Advisor: [Sumit Basu](#)

Project: Predicting clinical measures from physiological signals measured by a wearable device

- Developed a machine learning model that can predict clinical measures from physiological signals measured by a wearable device.

IBM Research, San Jose, California USA

AI Platform Research Group, *September 2017 – December 2017*

Advisor: [Heiko Ludwig](#)

Project: Detecting Backdoor Attacks on Deep Neural Networks by Activation Clustering

- Contributed to novel research algorithm by suggesting subspace projection technique that increased our performance from 15% to 95% accuracy. Designed and implemented environment for testing algorithm.
- Developed baseline and state-of-the-art neural networks using Tensorflow.
- Built data pipeline for large image dataset.
- Research was presented in global meeting to IBM executives including CEO and selected for funding.

Visa Research, Palo Alto, California USA

Data Analytics Group, *June 2017 – August 2017*

Advisor: [Hao Yang](#)

Project: Learning latent language models for improved machine reading comprehension

- Implemented model and baselines for language generation and question answering using Tensorflow and Facebook's ParlAI NLP software.
- Performed extensive literature reviews on machine reading comprehension and generative models.

University of Southern California, Los Angeles, California USA

Melady Machine Learning Lab, *November 2015 – May 2017*

Advisor: [Yan Liu](#)

Samsung and NSF funded project: "Variational Adversarial Deep Domain Adaptation for Health Care Time Series Analysis"

- Implemented novel neural network that employed variational inference and adversarial training for transfer learning of multivariate time-series.
- Proposed analyses used in publications to empirically verify that our model (a) performed domain adaptation by creating domain-invariant representations and (b) transferred temporal dependencies across domains. Research led to 2 publications and a patent.
- Communicated research to general public through research feature by the USC Graduate School and to technical audience at ICLR poster presentation.

Stony Brook University, Stony Brook, New York USA

Heavy Ion Research Group, *January 2013 – August 2015*

Advisor: [Axel Drees](#)

DOE funded project: “Modeling a Detection of internally reflected Cherenkov light Particle Detector for High-Multiplicity Collisions”

- Built and maintained a simulator in C++ for the heavy ion particle detector at the Brookhaven National Laboratory.
- Contributed methods from multivariate calculus and linear algebra to particle detection algorithm. Accuracy improved from 60% to 80%.
- Designed and implemented a statistical analysis pipeline in C++ for measuring efficacy of particle detection algorithm.

Stony Brook University, Stony Brook, New York USA

Computational Neuroscience Group, *Fall 2014*

Advisor: [Giancarlo La Camera](#)

NSF LSAMP funded project: “Spectral Analysis of Rodent Neural Data”

- Performed spectral analyses on neural data to determine behavioral correlates of neural activity.

California Institute of Technology, Pasadena, California USA

Emotion and Social Cognition Laboratory, *Summer 2014*

Advisor: [Ralph Adolphs](#)

HHMI funded project: “Modeling the Cognitive Process of Attributing Traits to Others”

- Formulated a trait learning behavioral experiment to study human inference.
- Built online platform to administer psychology experiments using Javascript, PHP, and HTML.

University of Minnesota, Minneapolis, Minnesota USA

Neuromodulation Research and Technology Laboratory, *Summer 2013*

Advisor: [Matthew Johnson](#)

NIH funded project: “Modeling Deep Brain Stimulation of Globus Pallidus Internus”

- Implemented python script to build a biologically feasible computational model of neural networks
- Created template for using python to simulate deep brain stimulation with “Neuron” software

National Central University, Jhongli City, Taiwan

Turbulent Combustion Laboratory, *Summer 2012*

Advisor: [Shenqyang Shy](#)

TEACHING EXPERIENCE

Stony Brook University, Stony Brook, NY

Calculus Instructor, *Spring 2015*

Worked with two math professors to develop and teach a supplementary calculus curriculum that promoted minority representation in stem majors.

Stony Brook University, Stony Brook, NY

Educational Opportunity Program Personal Tutor, *Spring 2013 - Fall 2014*

Tutored marginalized students in introductory physics and math courses

SERVICE

Reviewer, NeurIPS Deep Reinforcement Learning Workshop, 2020

Student Volunteer, ICLR, 2017

OUTREACH

Stony Brook Society of Physics, Stony Brook, NY, 2020

Michigan Explore Graduate Studies Workshop, Ann Arbor, MI, 2019

Research and Fellowships Week NSF Panel, Los Angeles, CA, 2016
National Society of Black Engineers Grad Panel, Los Angeles, CA, 2016
Graduate School External Fellowship Boot Camp, Los Angeles, CA, 2016
Mentored marginalized high school youth through the Pullias Center for Higher Education, Los Angeles, CA, 2016
Engineering Graduate Diversity Symposium, Los Angeles, CA, 2015
Black Student Association: What it takes to go to Graduate School, Los Angeles, CA, 2015
Collegiate Science and Technology Entry Program Undergraduate Research Panel, Stony Brook, CA, 2014

SKILLS Machine Learning Software: Pytorch, TensorFlow, Theano, Keras
 Neuroscience Software: Neuron
 Languages: Python, C++, C, Java
 Systems: Unix, Linux, OSX

PRESS [This Week In Machine Learning \(TWIML\) Research Feature](#)
 [University of Michigan Research Feature](#)
 [Exploring the source of social stereotypes](#)
 [Black History Month: Why a career in science?](#)
 [Research Feature by the USC Graduate School](#)
 [2015 NSF Graduate Research Fellow Wilka Carvalho](#)
 [Biomath Learning Center Launches Modified Supplemental Instruction Program](#)
 [URECA Research of the Month: Wilka Carvalho](#)
 Student Feature by Stony Brook University

INTERESTS ● traveling ● chess ● software development ● improvisational dance ● deadpan humor