

# Ruth Catherine Fong

St. John's College, St. Giles • Oxford OX1 3JP, United Kingdom

Email: [ruthfong@robots.ox.ac.uk](mailto:ruthfong@robots.ox.ac.uk) • Phone: +44-7490-568259 • Web: [www.ruthfong.com](http://www.ruthfong.com)

[RF]

## Education

---

**University of Oxford** – D.Phil. in Engineering Science, advised by Dr. Andrea Vedaldi **Oct. 2016-present**  
Conducting research in explainable AI. Thesis title: "Understanding Deep Neural Networks."

**University of Oxford** – M.Sc. in Neuroscience, *distinction* **Sept. 2015-Sept. 2016**  
Completed two computational neuroscience research rotations with Professor Rafal Bogacz as well as Professor Andy King, Dr. Ben Willmore, and Dr. Nicol Harper in the Oxford Auditory Neuroscience group.

**Harvard University** – A.B. in Computer Science, *magna cum laude with Highest Honors* **Sept. 2011-May 2015**  
Conducted thesis research with Professors David Cox and Walter Scheirer. Relevant coursework in machine learning, computer vision, computational linguistics, discrete math, cryptography, differential privacy, algorithms, programming languages, neuroscience, and biology. GPA: 3.8.

## Refereed Journal Articles

---

**Fong, R.**, Scheirer, W., and Cox, D. "Using Human Brain Activity to Guide Machine Learning." *Scientific Reports* (2018).

## Refereed Conference Papers

---

**Fong, R.\***, Patrick M.\*, and Vedaldi, A. "Understanding Deep Networks via Extremal Perturbations and Smooth Masks." *Proceedings of the IEEE International Conference on Computer Vision* (2019).

**Fong, R.** and Vedaldi, A. "Net2Vec: Quantifying and Explaining how Concepts are Encoded by Filters in Deep Neural Networks." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018).

**Fong, R.** and Vedaldi, A. "Interpretable Explanations of Black Boxes by Meaningful Perturbation." *Proceedings of the IEEE International Conference on Computer Vision* (2017).

\* denotes equal contribution.

## Refereed Workshop Papers

---

**Fong, R.** and Vedaldi, A. "Occlusions for Effective Data Augmentation in Image Classification." *ICCV Workshop on Interpreting and Explaining Visual Artificial Intelligence Models* (2019).

## Book Chapters

---

**Fong, R.** and Vedaldi, A. "Explanations for Attributing Deep Neural Network Predictions." Samek, W., Montavon, G., Vedaldi, A., Hansen, L., Müller, K.-R., editors. *Interpretable AI: Interpreting, Explaining, and Visualizing Deep Learning*. Springer LNCS. (2019).

## Research Projects

---

**"Interactive Similarity Overlays."\*** **Spring 2019-present**  
Introduced a novel, interactive visualization technique for comparing the CNN representation of image patches.

**"Occlusions for Effective Data Augmentation in Image Classification."\*** **Spring 2019-present**  
Developed a simple paradigm for leveraging stochastic and saliency-based perturbations as occlusions for effective data augmentation in image classification.

**"Understanding Deep Neural Networks via Perturbation Analysis."** **Fall 2016-present**  
Developed a theoretical framework for learning "explanations" of black box functions like CNNs as well as saliency methods for identifying "where" a computer vision algorithm is looking (Fong and Vedaldi, 2017; Fong and Vedaldi, Springer 2019; Fong et al., ICCV 2019).

**"Net2Vec: Quantifying and Explaining how Concepts are Encoded by Filters in DNNs."** **Summer 2017-2018**  
Investigated how human-interpretable visual concepts (i.e., textures, objects, etc.) are encoded across hidden units of a convolutional neural network (CNN) layer as well as across CNN layers (Fong and Vedaldi, 2018).



**“Modeling Blind Single Channel Sound Separation Using Predictive Neural Networks.” Summer 2016**

*M.Sc. Dissertation #2.* Developed an unsupervised learning paradigm for sound separation using fully connected and recurrent neural networks to predict the future from past cochleagram data.

**“Optimizing Deep Brain Stimulation to Dampen Tremor.” Spring 2016**

*M.Sc. Dissertation #1.* Developed a computational oscillator model that modeled the tremor-dampening effects of phasic deep brain stimulation and analyzed it on experimental data.

**“Leveraging Human Brain Activity to Improve Object Classification.” Spring 2014-Summer 2015**

*A.B. Thesis.* Introduced a biologically-informed machine learning paradigm for object classification that biases models to better match the learned, internal representations of the visual cortex (Fong et al., 2018).

**“Applying Differential Privacy to Complex Genomics Data” with Louis Li. Fall 2014**

Generalized private data-release mechanisms for genomics data linked to K possible disease categories.

\* Working towards publication.

**Research Talks**

---

**Continental** – Oxford, U.K.: “Understanding Deep Networks via Extremal Perturbations” **April 2019**

**OpenAI** – San Francisco, CA: “Extremal Perturbations” **April 2019**

**Stanford Vision and Learning Lab** – Stanford, CA: “Extremal Perturbations” **April 2019**

**Mutual Benefits of Cognitive and Computer Vision (MBCC) Workshop** – Venice, Italy **October 2017**

Contributing talk on “Using Human Brain Activity to Guide Machine Learning” (MBCC is co-located at the International Conference on Computer Vision [ICCV]).

**Girls Advancing in STEM (GAINS) Network Conference** – Cambridge, MA **April 2015**

Invited technical talk on “Leveraging Human Brain Activity to Improve Object Classification” (GAINS conference was hosted by MIT’s Broad Institute).

**Student Advising**

---

**Mandela Patrick**, PhD student at University of Oxford (co-advised by Andrea Vedaldi) **2018-present**

**Kurtis David**, Masters student at University of Texas, Austin (co-advised by Qiang Liu) **2019-present**

**Research Media Coverage**

---

**New Scientist** – “MRI brains scans train machines to see the world more like us” **March 2017**

**Work Experience**

---

**Pro Unlimited @ Facebook: Research Collaborator** with Andrea Vedaldi – London, UK **March 2019-Aug. 2019**  
Developed an PyTorch library for explainable AI with a focus on attribution techniques.

**Google Research: Research Intern** with Vitto Ferrari – Zurich, Switzerland **July 2018-Nov. 2018**  
Researched leveraging the internal representation of image recognition CNNs to perform localization.

**D.E. Shaw, Co.: Quantitative Software Engineer Intern** – New York, NY **June 2014-Aug. 2014**  
Worked on a proprietary software engineering project and a quantitative financial research project.

**Apple: Software Engineer Intern** – Cupertino, CA **May 2013-Aug. 2013**  
Implemented HTML5 features and improved developer tools in the open-source WebKit project.

**Microsoft: Software Engineer Intern** – Redmond, WA **May 2012-Aug. 2012**  
Worked for the Windows 8 Device Drivers team on a diagnostic tool for Windows 8-style device apps.

**Technical Projects**

---

**CrimsonCommunity** with Annie Ryu, Styliani Pantella, and Victoria Gu. **Jan. 2013-Dec. 2014**  
Developed an SMS-based service that provides students easy access to Harvard’s mental health services.



**Health Data Collection & Visualization Project** as a 2014 Tech in the World Fellow. **Dec. 2013-Jan. 2014**  
 Partnered with Ifakara Health Institute to build centralized web platform that collects vital health statistics from throughout Tanzania and taught seminars at Dar es Salaam Institute of Technology.

**BagIt** with Eric Lu. **Dec. 2012-May 2013**  
 Developed an iOS & mobile app that allowed Harvard students to order take-away meals from dining halls.

## Leadership and Service Activities

---

**University of Oxford, Dept. of Engineering Science: Tutorial Instructor** **Fall 2019-Spring 2020**  
 Will teach sections and grade assignments for B14: Information Engineering, a third-year undergraduate course covering topics in signal processing, computer vision, machine learning, and statistics.

**NJ Governor's School in Engineering & Technology: Course Instructor** – Piscataway, NJ **Summer 2015**  
 Designed and taught "Mathematics in the World", an introductory course on discrete math, probability, and their real-world applications to gifted 3<sup>rd</sup> year high school students in New Jersey, USA.

**Harvard Women in CS: Webmaster, Board Member, & Mentor** – Cambridge, MA **2012-2015**

**Harvard Course Assistant & Teaching Fellow** – Cambridge, MA **2012, 2014**  
 Held office hours, taught sections, and authored and graded assignments for CS50: "Intro. to Comp. Sci. I" (Fall '12), CS20: "Intro. to Discrete Math" (Spring '14), CS121: "Intro. to the Theory of Computation" (Fall '14).

## Honors

---

**Open Philanthropy AI Fellow ('18)**, CVPR Outstanding Reviewer ('18), Women in Computer Vision CVPR Travel Grant ('17, '18), International Computer Vision Summer School Best Poster Award ('17), Murray Speight Research Grant ('17, '18), Women in Machine Learning NIPS Travel Grant ('16; '17 declined), **Rhodes Scholar ('15)**, **National Science Foundation (NSF) Graduate Research Fellowship ('15, declined)**, **Fulbright Scholar to Tanzania ('15, declined)**, **Hoopes Prize (for outstanding undergraduate thesis) ('15)**, Derek Bok Certificate of Distinction (for outstanding CS121 teaching evaluations) ('15), Tech in the World Fellowship ('14), Apple iOS Scholarship ('13)

## Reviewing

---

*Neural Information Processing Systems (NeurIPS), European Conference in Computer Vision (ECCV), IEEE Conference on Computer Vision & Pattern Recognition (CVPR), IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI), IEEE Signal Processing Letters, Black in AI (BAI), Women in Machine Learning (WiML), Women in Computer Vision (WiCV), Distill*

## Skills

---

Proficient in the following programming languages and deep learning frameworks: Python, MATLAB, Java, C#, C; PyTorch, TensorFlow, MatConvNet, PyCaffe. Experience with HTML, CSS, Javascript, MySQL, Objective-C.