Ian Huang

Palo Alto | (718)552-6811 | LinkedIn: ian--huang | Email: ianhuang@stanford.edu

SUMMARY

I'm passionate about using AI to extend our ability to tell stories. I believe that everyone should have the ability to create, and that AI has an increasingly important role to play. To that end, I'm interested in 3D content generation through interfaces intuitive to humans (like natural language and haptics), and the application of deep learning and geometric computation for that purpose.

RESEARCH EXPERIENCE

Geometric Computation Group, Stanford September 2020 - Present

CS Ph.D.

Working on developing a system that uses natural language for 3D content generation, advised by Professor Leonidas Guibas. Work in progress!

Stanford Vision & Learning Lab (SVL) June 2019-January 2020

Visiting Research Intern

The world needs an AI that can reason about unfamiliar environments, objects, and entities. I designed and developed a model that deduces the taxonomic relationship between seen and unseen objects by observing human interactions with those classes in video data. I introduced object-functionality projection, a novel technique to embed the functionality of objects as a representation of their interactions with others in a latent space. It demonstrated significant improvements upon baselines and prior works (at least a 27% relative increase – or 6% absolute – in cumulative accuracy, and under some conditions, a 4x increase in accuracy), and ablation studies validated the need of object-functionality projection. To the best of our knowledge, this is the first system for unknown object discovery in videos, and moreover the first to explicitly use object functionality for that purpose.

- Developed a novel formulation of object functionality that entangles multiple object representations and human-object interactions (HOIs).
- Designed a new neural network architecture for zero-shot prediction of an unknown object class's closest super-class.
- Proposed *object-functionality projection*, a key component of the model architecture that learns robust representations of object-object interactions that encode for the function of different noun classes.
- Demonstrated significant improvements upon comparisons against prior works in static visual data with at least 27% relative (or 6% absolute) increase in cumulative accuracy, and under some conditions, a 4x increase.
- Performed ablation studies to validate the need of *object-functional projection*.
- Built tools using Pytorch for extracting hand mesh, pose, and position for modeling human-object interaction.
- Iteratively designed, implemented, and tested various model architectures using Pytorch.
- Submitted research paper, *Functionality-Driven Object Discovery in Videos*, as first author, for CVPR2020.

Columbia Vision + Graphics Center January 2019-May 2019

Researcher

Catastrophic forgetting is the biggest bottleneck in online deep learning, and is critical to lifelong learning in robots. I initiated a novel idea of using adversarial examples, a flaw of neural networks, to fix catastrophic forgetting, another flaw of neural networks. I introduced a new way of combating catastrophic forgetting in an online learning setting by optimizing input images in an adversarial way without requiring the need for a specific type of architecture or training loss function. We demonstrated that adversarial perturbations to the data is capable of combating catastrophic forgetting.

- Introduced the concept of *gradient alignment*, a heuristic by which input images of unknown classes can be optimized iteratively to minimally disrupt the retention of prior knowledge upon backpropagation.
- Demonstrated the correlation between *gradient alignment* of an image and the extent of catastrophic forgetting caused in MNIST images.
- Maintained semantic verisimilitude of adversarial examples using regularization.

- Implemented and tested numerous gradient alignment-optimizing pipelines to reduce the extent of catastrophic forgetting for MNIST and CIFAR100.
- Designed and ran experiments proving the effectiveness of the system to reduce catastrophic forgetting while retaining the ability to learn from new samples that are adversarially modified but semantically coherent.
- Applied real analysis to ground our empirical findings and methodology in mathematical guarantees.
- Submitted research paper, *Combating Catastrophic Forgetting Without the Model,* as first author, to AISTATS2020.

Columbia Robotics Lab September 2018-May 2019

Researcher

While robotics allows for human-robot interaction, a challenge remains to maximally use this interaction for learning purposes. I built an RGB-D semantic segmentation system for visual input for robots to undertake an autonomous navigation task in a simulated environment using human-in-the-loop reinforcement learning. On another project, I also built dimensionality reduction systems using the autoencoder architecture to discover human hand synergies so that we can control an under-actuated anthropomorphic hand remotely by controlling our own.

- Developed and tested deep learning models using Tensorflow and Pytorch for semantic segmentation with RGB-D data, enabling the robot to identify and locate target destinations in simulation and real life.
- Led the investigation on using eye-tracking data as a method to guide learning in human-in-the-loop training schemas.
- Used eye-tracking to embed higher level reasoning in our system to enhance few-shot learning and lifelong learning.
- Designed and proposed a data collection pipeline to gain joint angle and sEMG data spanning over 5 different grasp types on a total of 20 different objects.
- Built a deep learning system to map real-time human hand joint data to robotic hands using sEMG sensors by employing deep autoencoder architectures
- Pretrained autoencoder on data gathered by the aforementioned pipeline using a loss function that incentivizes the preservation of the grasp type information in the latent space.

Columbia Data Science Institute Wu Lab September 2017-May 2018

Researcher

The explainability of deep learning systems remains a bottleneck. I worked on a system to verify the internal logic of recurrent neural networks. I designed various experiments that sought to validate the identification of important neuron groups. We introduced a novel method called Deep Neural Inspection (DNI) to make sense of internal neuron-level behaviors within recurrent neural networks to dissect and validate their logic for their decisions.

- Formulated the problem and the necessary experiments to evaluate our system's efficacy.
- Used Keras and Tensorflow to build, train and test machine learning models and design neural network architectures.
- Spearheaded the branch of the project that investigated the conversion of continuous activation spaces of neurons into more interpretable finite state machines so as to be able to validate the logic of the network.
- Used Principal Component Analysis (PCA) and T-SNE to visualize and analyze the high-dimensional results of tests.
- Co-authored research papers accepted to SIGMOD2019 and SysML2018.

INDUSTRY EXPERIENCE

LinkedIn May 2018-September 2018

Machine Learning & Relevance Engineer Intern

What do the contents of an image tell us about the relevance of advertisements to different users? My project was to develop usable semantic embeddings of images to better help LinkedIn's ad ranking algorithms send more relevant advertisements to users. My work showed that visual information – something that had never been used for this purpose before at LinkedIn – can improve the relevance of advertisements.

- Designed deep learning architectures for image embedding extraction.

- Implemented, trained, and tested deep learning models using Tensorflow and monitored training process using Tensorboard.
- Qualitatively evaluated and visualized higher-dimensional embeddings in three-dimensional space using T-SNE.
- Used Python to implement a parallelized web scraping script for data aggregation, image processing, and data cleaning.

Intel Corporation June 2017-September 2017

Data Science Intern at the Artificial Intelligence Product Group

Computers can indicate the location of certain objects in images, but oftentimes not well. My work enabled the state-of-the-art detection model to predict posed bounding boxes that more tightly enclose the object of interest. This sparked a partnership with Red Cross, who is currently using my work on satellite images to estimate the population of remote areas that are stricken by disaster, thus serving as a better estimate for the amount of aid needed.

- Implemented Jarvis March and Rotating Calipers, as well as various self-developed algorithms, in Python to preprocess the data for 18,000 satellite images in the Spacenet dataset.
- Implemented and tested the modified architecture in Python (total >5500 lines) using Intel Nervana's deep learning library (Neon) to detect the position, dimensions, and angle (i.e. pose) of buildings in satellite images.
- Created and ran unit tests for a dozen components of the system and documentation to facilitate further work.
- Vectorized operations using NumPy to optimize computational speed of some components up to 300 times.
- Trained and tested architecture, optimized for hyperparameters and benchmarked components.
- Participated in discussions of potential partnerships and applications of this project with Red Cross in aid estimation and the Transportation Security Administration (TSA) in object identification.

Zhenfund December 2016-January 2017

Information Technology & Engineering Intern

- Designed architecture of a computer-aided deal-sourcing system that would save deal sourcing time by 50%.
- Reduced projected developmental cost to 85% by streamlining the design of Zhenfund's current system.
- Visited and evaluated 3 startups according to Zhenfund's investment philosophy

TEACHING EXPERIENCE

Columbia University January 2019-May 2019, January 2020-May 2020

Teaching Assistant, Graduate-Level Machine Learning (COMS 4771)

As one of the most difficult CS classes offered at Columbia, this graduate-level machine learning class teaches more than 200 students semesterly on topics from linear regression to Vapnik–Chervonenkis dimensionality and introductory computational learning theory.

- Developed and wrote solution manuals for biweekly problem sets.
- Held office hours twice a week to conduct mini-lectures covering topics ranging from Support Vector Machine and linear regression to VC-dimensionality, dimensionality reduction techniques, and graphical networks.
- Organized exam review sessions.
- Graded 120 submissions bi-weekly, and handled all regrade requests.

PUBLICATIONS

- **I. Huang**, J. Ji, J. C. Niebles, *Functionality-Driven Object Discovery in Videos*. CVPR 2020. Under review.
- **I. Huang**, L. Arnett, C. Vondrick, *Combating Catastrophic Forgetting without the Model*. AISTAT 2020. Under review.
- T. Sellam, K. Lin, **I. Huang**, Y. Chen, M. Yang, C. Vondrick, E. Wu, *DeepBase: Deep Inspection of Neural Networks.* SIGMOD 2019.
- T. Sellam, K. Lin, **I. Huang**, E. Wu, C. Vondrick. *"I Like the Way You Think!"* - Inspecting the Internal Logic of Recurrent Neural Networks. SysML Conference 2018.

- **I. Huang**, M. Chen, W. Yang, Z. Qin. *An Evolutionary Model for Efficient Transportation Networks Based on Minimum Spanning Trees.* The 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, 2015.

PATENTS

- *"WhispEAR Utility Patent"* cn ZL 2013 2 0361698.0, *"WhispEAR Exterior Design Patent"* cn ZL 2013 3 0272231.4

FELLOWSHIP AND DISTINCTIONS

- Ng Teng Fong Scholarship and Fellowship Fund for my research at the Stanford Vision & Learning Lab, 2019
- Columbia University Dean's List for 6 Consecutive Semesters, Fall 2016 - Spring 2019
- 1st Place (twice) & 2nd Place Winner at the American Protégé Music Competition and performed solo violin at Carnegie Hall in 2015, 2016, 2017.

EDUCATION

Columbia University, New York, NY 2016-Present
B.A. in Applied Mathematics and Computer Science
Graduate-Level: Machine Learning, Probability Theory, Real Analysis, Abstract Algebra, Numerical Methods for Partial Differential Equations, Topology, Advanced Computer Vision, Natural Language Processing, Independent Research.
Undergraduate-Level: Analysis & Optimization, Discrete Math, Linear Algebra, Partial Differential Equations, Ordinary Differential Equations, Data Structures, Computer Science Theory, Computer Systems, Advanced C and C++ Programming, Cellular Neurobiology, Electrical Engineering, Multivariate Calculus.

Online Courses 2016-Present
Courses: Deep Learning for Natural Language Processing (Stanford), Convolutional Neural Networks for Visual Recognition (Stanford), Machine Learning (Stanford), Programming in Python (Rice University), Modern Robotics: Motion, Kinematics and Dynamics (Northwestern University), Computational Neuroscience (University of Washington), Ethical Hacking

SKILLS

- **Advanced**: Machine Learning, Deep Learning, Computer Vision, PyTorch, Python, Tensorflow, Keras, Numpy, Scipy, Scikit-Learn, Pandas, Java, C, R, MATLAB, Reinforcement Learning, OpenAI gym, Robot Operating System (ROS), Arduino, Raspberry Pi, Autodesk Fusion 360, Solidworks, Cura, 3D printing
- **Proficient:** Finite Element Analysis, C++, Mathematica, HTML, CSS, Javascript, NLTK, Gensim
- **Languages**: English, French, Spanish, Mandarin

EXTRACURRICULAR ACTIVITIES

- Member of Columbia Space Initiative 2018-Present
- Board Member, Columbia Organization of Rising Entrepreneurs (CORE) 2016-2017
- Member of Columbia Application Development Initiative Club (ADI) 2017-2018
- Vice President, Columbia University Medical Informatics Society (CU MIS) 2016-2017
- Founder, Business Development Leader, Web Developer of CollegeWalk (Startup!) 2016-2017
- Violinist in the Columbia University Orchestra 2016-2017