

Bill Cai

Machine Learning Engineer

Work Experience

2019
|
Present

Machine Learning Engineer

Data Science and AI Division, Government Technology Agency

Singapore

- Govtech Singapore is the technology arm of the Singapore government. The Video Analytics team works on developing and deploying computer vision and video understanding models for social good.
- Lead for AI modelling efforts in few-shot object detection, video activity recognition, and captioning models. Deployed and implemented using REST APIs in Python/Typescript, with K8s backends for infrastructure abstraction.
- Main cloud architect and DevSecOps lead for petabyte-scale cloud-native computer vision platform and ML pipelines for image and video analytics. Scaled ML infra engineering squad from 1-3 people, supporting a larger data science and analytics team of 20, with petabyte-level of data-intensive products that save >1mil man-hours annually.
- Tech lead for crowd analytics project for 200+ cameras deployed on AWS. Designed and implemented full Terraform infrastructure-as-code, serverless architecture using AWS Lambda, API Gateway, and cloud-native solutions including AWS Rekognition, S3, ECR, API Gateway.
- Practical experience with deploying, updating, and maintaining a secure and compliance-ready cloud-native system. Designed a fully security-compliant system, and main representative for successfully completed 3rd party security assessments for IM8 and security risk assessment.

2018
|
2019

Data Scientist, Computer Vision and Deep Learning

One Concern

Menlo Park, California

- One Concern is a benevolent AI company that provides trusted insights that positively impact our communities. Our mission is to drive deep social impact through benevolent intelligence to save lives and livelihoods.
- Lead in-house inference of key features from unstructured data, such as satellite images and street-level imagery. Extensive use of Keras, Tensorflow, PyTorch to build deep learning tools. Wrote and built Docker images, with deployment in Kubernetes.
- Backend engineering for a city-scale and real-time platform for infrastructure resilience. In charge of resilience and infrastructure recovery estimations, using combinatorial and graph optimisation techniques. Launched a new power and water utility modelling engineering effort that grew from a 2 person team into a 10+ person new product team, while leading algorithmic and ML engineering technical functionalities. In charge of key technical challenges including vectorising bottleneck computations to 100x in Python to increase computational output.
- Customization of open-source Javascript/HTML/CSS image annotation libraries with integration to Amazon Mechanical Turk.

2017
|
2018

Graduate Researcher, Computer Vision

MIT Senseable City Lab

Cambridge, Massachusetts

- Data mining and labelling, deep learning and computer vision models training, and large-scale deployment to quantify urban canopy cover and parking utilization on large city-wide scales
- Sensor-fusion of lidar and camera data for obstacle detection in autonomous marine vehicle applications in Amsterdam and Boston/Cambridge
- Implemented state-of-the-art CNN architectures for classification, semantic segmentation and instance segmentation, including residual network, Mask-RCNN, PSPNet. Utilized gradient class activation (Grad-CAM) maps to understand learned features
- Extensive use of ROS, including Google Cartographer for SLAM, Velodyne lidar, IMU, USB cameras, for sponsored project by SNCF in Paris

2017
|
2017

Summer Associate, Product Analytics

Thumbtack

San Francisco, California

- Built live dashboards with Python, R, SQL, Javascript/HTML/CSS to track key metrics
- Modeled two-sided matching and dynamic marketplaces in Python. Our [engineering blog post](#) that explains more!
- Analyzed A/B test results, including using quasi-experimental methods, to understand impact of product feature changes on customer behavior
- Worked closely with product managers, engineers and designers to shape product decisions

Contact Info

Website and Projects:

billcai.com

LinkedIn Profile:

[linkedin.com/in/billcai77](https://www.linkedin.com/in/billcai77)

Email Address:

billcai@alum.mit.edu

Skills

Building large-scale and operational ML systems

Deep learning and ML frameworks: PyTorch, Tensorflow, sklearn, JAX

Python, Typescript, Julia, MATLAB, R, Stata, SQL

AWS Solutions Architect Pro, AWS DevOps Eng Pro

Docker, Kubernetes, Terraform

Interests

Coding, Programming

ML Research (140 citations,

h-index of 5 on [Google Scholar](#))

Economic Theory (PhD classes in Market/Auction Design, Computational Macro)

Academic Service/Talks

ICLR 2020, NeurIPS 2020,

NeurIPS 2021, ICML 2021

Climate Change Workshop

Program Committee

Reviewer for CVPR, NeurIPS,

IEEE Internet of Things Journal

IEEE BigData Congress 2018

Presentation

MIT CCE Student Seminar 2018

Education

● M.S. in Computational Science and Engineering

Center for Computational Science and Engineering, Massachusetts Institute of Technology
Cambridge, Massachusetts

- GPA: 5.00/5.00, thesis on applying computer vision and deep learning for large-scale quantification of urban and city dynamics (advised by [Carlo Ratti](#))
- Selected Coursework: Advances in Computer Vision, Statistical Learning Theory and Applications, Numerical Methods in Partial Differential Equations, Optimization Methods

● B.A. in Economics

University of Chicago
Chicago, Illinois

- GPA: 3.87/4.00, Graduated with Phi Beta Kappa (highest honors) and Dean's List from 2014-2017
- Selected Coursework: Planning, Learning and Estimation for AI, Machine Learning, Market Design, Inequality: A Perspective from Macroeconomics, Honors Econometrics, Applied Statistical Modeling

Research, Journal and Conference Publications

- Dec 2020 ● DAMSL: Domain Agnostic Meta Score-based Learning
[CVPR 2021 Workshop on Learning from Limited and Imperfect Data](#)
John Cai, **Bill Yang Cai**, Shengmei Shen
- Oct 2019 ● Quantifying Urban Canopy Cover with Deep Convolutional Neural Networks
[Published in NeurIPS Workshop on Climate Change AI](#)
Bill Yang Cai, Xiaojiang Li, Carlo Ratti
- Dec 2018 ● Quantifying Legibility in Indoor Spaces Using Deep Convolutional Neural Networks: A Case Study in Train Stations
[Published in Building and Environment](#)
Wang Zhoutong, Liang Qianhui, **Bill Yang Cai**, Louis Charron, Fabio Duarte, Carlo Ratti
- Dec 2018 ● Deep Learning Architect: Classification for Architectural Design through the Eye of Artificial Intelligence
[Published in Computational Urban Planning and Management for Smart Cities](#)
Yuji Yoshimura, **Bill Yang Cai**, Wang Zhoutong, Carlo Ratti
- Aug 2018 ● Deep Learning Based Video System for Accurate and Real-Time Parking Measurement
[Published in IEEE Internet of Things Journal](#)
[Special Issue on Enabling a Smart City: Internet of Things Meets AI](#)
Bill Yang Cai, Ricardo Alvarez, Michelle Sit, Fabio Duarte, Carlo Ratti
- Apr 2018 ● Treepedia 2.0: Applying Deep Learning for Large-scale Quantification of Urban Tree Cover
[Published in IEEE BigData Congress 2018, arXiv preprint](#)
Bill Yang Cai, Xiaojiang Li, Ian Seiferling, Carlo Ratti
- Feb 2018 ● Using Street-level Images and Deep Learning for Urban Landscape Analysis
[Published in Landscape Architecture Frontiers](#)
Xiaojiang Li, **Bill Yang Cai**, Carlo Ratti

Research Projects

- Sep 2017 | Sep 2018 ● [Treepedia](#)
Role: Computer Vision and Deep Learning Lead
Treepedia is a project by the MIT Senseable City Lab in partnership with the World Economic Forum to measure canopy cover and green spaces in cities globally. This project has inspired planners and policymakers to design greener cities, and has been featured on the [Wall Street Journal](#), [Time](#), [Wired](#), and [Forbes](#). I developed the Tensorflow-based deep learning algorithms used to detect and quantify canopy cover from Google Street View images, from data annotation to model training/hyperparameter tuning to eventual [trained model/library](#).
- Jan 2018 | Sep 2018 ● [Roboat](#)
Role: Computer Vision Engineer
Roboat is a 5 year research project and collaboration between the Amsterdam Institute for Advanced Metropolitan Solutions and MIT to develop the world's first fleet of urban autonomous floating vessels. Our project has been featured on [CNBC](#), [The Verge](#), [Reuters](#), [Quartz](#), and [Fortune](#). I tuned and deployed Tensorflow-based instance segmentation and object detection models on boat-based GPUs/mini PCs to detect live obstacles, and provide estimated obstacle locations and types to the ROS-based motion planner.