

# Chengkun Li

[sjtulck@gmail.com](mailto:sjtulck@gmail.com) | [github.com/pipme](https://github.com/pipme) | [Google Scholar](https://scholar.google.com/citations?user=...) | [LinkedIn](https://www.linkedin.com/in/...)

**Summary:** Artificial intelligence researcher with a PhD in computer science, specializing in probabilistic machine learning, computational statistics, and computer vision. Experienced in Python, generative models, open source software development and interdisciplinary collaboration across scientific and engineering domains.

## EDUCATION

---

**University of Helsinki** | *Ph.D. in Computer Science* **Aug 2021 – Apr 2026**

- Research on computationally efficient Bayesian inference and probabilistic machine learning methods.

**Shanghai Jiao Tong University** | *B.E., M.E. in Mechanical Engineering* **Sep 2014 – Jun 2021**

- Honors program and outstanding graduate award (B.E.). Ranked top 0.1% in the national college entrance exam.

**CentraleSupélec** | *Diplôme d'ingénieur (Master's degree)* **Sep 2016 – Jun 2018**

- Coursework in mathematics and computer science. Concurrent Licence 3 in Mathematics (Paris-Sud University).

## EXPERIENCE

---

**University of Helsinki** | *Researcher* **Aug 2021 – Present**

- **PyVBMCMC:** Core developer of a Python package for efficient Bayesian inference using *active learning*, achieving  $10\text{--}100\times$  speed-ups on expensive models through *sample/evaluation efficiency*. The package has 100+ GitHub stars and is used by researchers in both science and engineering fields. Paper published in JOSS.
- **Generative Models for Scientific Inference:** Developed a hybrid framework for *reliable and efficient statistical inference* by combining deep generative models and MCMC, demonstrating  $15\text{--}60\times$  speed-ups on neuroscience and psychology models. Presented at NeurIPS workshop; published in TMLR.
- **Surrogate Modeling:** Developed surrogate modeling methods including both deep learning (normalizing flows) and classical ML approaches (sparse Gaussian processes, variational inference), achieving accurate approximation of posterior distributions. Published in a top-tier computational statistics journal and PMLR.
- **AI4S – Geophysics:** Collaborated with geophysicists to apply surrogate modeling for analyzing expensive scientific simulators, reducing costs by *an order of magnitude*. Published in Geoscientific Model Development.
- **Ensemble Method:** Developed a stacking/ensembling approach combining multiple surrogate models (Gaussian processes and mixtures) from parallel runs, improving accuracy with *negligible extra costs*. Published in TMLR.

**Shanghai Jiao Tong University** | *Master's student* **Sep 2018 – Jun 2021**

- **KeypointNet:** Built a *large-scale 3D keypoint dataset* with 80k+ keypoints and 8k+ object models, enabling object keypoint detection and semantic understanding in both 2D and 3D spaces; published in CVPR and TPAMI.
- **Object Embedding & Object Detection:** Developed a method for achieving *dense object semantic embedding* extraction from correspondence annotations; published in ECCV. Contributed to a voting-based method for 3D object oriented bounding box detection; published in CVPR.
- National champion, China RoboCup Small Size League (2019). Second prize, China Post-Graduate Mathematical Contest in Modeling (2020).

**Microsoft Research Asia** | *Research intern (remote)* **Jun 2020 – Nov 2020**

- Worked on neural rendering methods for novel view synthesis from sparse multi-view images. Explored approaches including NeRF, hypernetworks, and transformers, and developed an experimental research codebase.

**ETH Zürich (Autonomous System Lab)** | *Visiting student* **Jun 2018 – Aug 2018**

- Developed deep learning and geometric methods for line detection and description in RGB-D images, contributing preliminary work later acknowledged in a 3DV publication.

## SKILLS

---

**Programming:** Python, C++, MATLAB, Mathematica

**Languages:** English, Chinese, French

**Tools & Infrastructure:** PyTorch, JAX, Git, Linux, CI/CD, HPC (Slurm), Hydra, Weights & Biases

**Academic Service:** Reviewer for NeurIPS, AISTATS, TMLR, Nature Human Behaviour

## SELECTED PUBLICATIONS

---

Li et al. Amortized Bayesian workflow. *TMLR*, 2026.

Li et al. Fast post-process Bayesian inference with variational sparse Bayesian quadrature. *Statistics and Computing*, 2025.

Lou\*, You\*, Li\* et al. Human correspondence consensus for 3D object semantic understanding. *ECCV*, 2020.

You, Lou\*, Li\* et al. KeypointNet: A large-scale 3D keypoint dataset. *CVPR*, 2020. (\*equal contribution)