curriculum vitæ of

Hiroki Naganuma

OUT-OF-DISTRIBUTION GENERALIZATION · OPTIMIZATION · LARGE LANGAGE MODEL

⌂ hiroki11x.github.io   ✉ naganuma.hiroki@mila.quebec   ☐ +1 514-559-1008   in hiroki11x   🎓 google scholar

📍 Mila - Quebec AI Institute, 6666 Saint-Urbain Street, #200 Montreal, QC, Canada, H2S 3H1

Education

Université de Montréal and Mila (Quebec AI Institute)   Montréal, QC, Canada
09/20 - present
Ph.D. Candidate in Computer Science
Supervisor: **Ioannis Mitliagkas**.
Topic: Uncertainty Calibration, Out-of-Distribution, and Optimization in Deep Neural Networks.

Tokyo Institute of Technology   Tokyo, Japan
04/17 - 03/19
M.Eng. in Computer Science **(Completed Master Course with Summa Cum Laude)**
Supervisor: **Rio Yokota**.
Thesis title: Smoothing Objective Function in Stochastic Optimization for Large Scale Deep Learning.
Exchange student at Georgia Institute of Technology

Tokyo Institute of Technology   Tokyo, Japan
04/13 - 03/17
B.Eng. in Computer Science **(Completed Bachelor Course with Dean's Award, School of Computing)**
Supervisor: **Rio Yokota**.
Thesis title: Hierarchical Low Rank Approximation for Convolutional Neural Networks.

Affiliations

Meta (AI and Systems Co-Design Team)   Menlo Park, CA, United States
05/25 - 08/25
Research Intern
Supervisor: **Hao-Jun Michael Shi** and **Parameswaran Raman**
Topic: Second Order Optimization for LLMs.

University of Tsukuba (Center for Cybernics Research)   Tsukuba, Japan (Remote)
09/24 - present
Visiting Scholar
Supervisor: **Yoshiaki Sankai**.
Topic: Multimodal Model for Medical Application.

Past Affiliations

Google DeepMind (Foundational Research Team)   Mountain View, CA, United States
09/24 - 02/25
Student Researcher
Supervisor: **George E. Dahl**
Topic: Machine Learning Foundations.

Microsoft Research (Phi Model Team)   Redmond, WA, United States
06/24 - 08/24
Research Intern
Supervisor: **Philipp Witte** and **Russell Hewett**.
Topic: Large Language Models Foundations.

LinkedIn (Foundation AI Technologies Team)   Sunnyvale, CA, United States
06/21,22,23 - 09/21,22,23
Research Intern
Supervisor: **Chengming Jiang**, **Mingzhou Zhou**, **Aman Gupta**, **Haichao Wei** and **S. Sathiya Keerthi**.
Topic: Mitigating Trade-off between Calibration and AUC.

Japan Society for the Promotion of Science   Tokyo, Japan
04/20 - 08/20
Research Fellow
Supervisor: **Rio Yokota**.
Topic: Developing Efficient Algorithms based on Smoothing Loss Function for Large Batch Training.

RIKEN AIP (Deep Learning Theory Team)   Tokyo, Japan
09/18 - 10/18
Research Intern
Supervisor: **Taiji Suzuki**.
Topic: Large Scale Distributed Deep Learning using Natural Gradient Approximation Method

04/17 - 03/18	**IBM Research (HPC and Large Scale Optimization Team)**	Tokyo, Japan

Student Fellow

Supervisor: **Taro Sekiyama** and **Kiyokuni Kawachiya**.

Topic: Optimization and Parallelization of Deep Learning Framework

Selected Honors

2025	Invited to GRF Workshop	JANE STREET GRADUATE RESEARCH FELLOWSHIP
2025	**Best Reviewer Award**	**AISTATS 2025**
2024	**Top Reviewer Award**	**NEURIPS 2024**
2024	Poster Award	JSPS SAN FRANCISCO ANNUAL MEETING
2024	JHPCN Emerging Research Award	JHPCN SYMPOSIUM 2024
2022	Minister of Internal Affairs and Communications Award	JAPAN CAMPUS GRANDPRIX
2018	**National Winner, Japan Representative (top 2%)**	**JAMES DYSON AWARD**
2018	Grand Prize (1st in 12 teams)	ART HACK DAY 2018
2017	Best 32, Japan Representative	MICROSOFT IMAGINE CUP 2017 WORLD FINALS
2017	3rd Prize (3rd in 52 teams)	STANFORD HEALTH HACKATHON 2017
2016	First Prize (1st in 89 teams)	JPHACKS-2016 JAPAN'S LARGEST STUDENT HACKATHON
2016	First Prize	MASHUPAWARD-2016 JAPAN'S LARGEST HACKATHON STUDENT DIVISION
2019	**Complete Master Course with Summa Cum Laude**	**TOKYO INSTITUTE OF TECHNOLOGY**
2017	**Dean's award, School of Computing**	**TOKYO INSTITUTE OF TECHNOLOGY**
2013	Tokyo Metropolitan Governor Award (Graduate First on the List)	TOKYO METROPOLITAN

Selected Grants and Fellowships

04/25 - 03/26	Computational Resource Support (56,000 GPU hours)	UOSAKA AI RESEARCH PROJECTS (FY2025)
11/24 - 08/25	Research Grant (Approx 3,300USD)	THE ANRI FELLOWSHIP
07/24 - 03/25	Computational Resource Support (4,800 GPU hours)	JHPCN JOINT RESEARCH PROJECT (EX24402)
07/24 - 03/25	Computational Resource Support (4,800 GPU hours)	TSUBAME ENCOURAGEMENT PROGRAM (FY2024)
09/23	Computational Resource Support (2,048GPUs x 24H)	ABCI GRAND CHALLENGE PROGRAM (AIST)
07/23 - 03/24	Computational Resource Support (12,000 GPU hours)	JHPCN JOINT RESEARCH PROJECT (EX23401)
07/23 - 03/24	Computational Resource Support (12,000 GPU hours)	TSUBAME ENCOURAGEMENT PROGRAM(FY2023)
09/22 - 08/25	Scholarships for AI Field Scholar	CYBERDYNE AI FELLOWSHIP
09/22 - 08/24	Prestigious Scholarship	SHIGETA EDUCATION FOUNDATION SCHOLARSHIP
07/22 - 03/23	Computational Resource Support (12,000 GPU hours)	JHPCN JOINT RESEARCH PROJECT (EX22401)
07/22 - 03/23	Computational Resource Support (12,000 GPU hours)	TSUBAME ENCOURAGEMENT PROGRAM (FY2022)
09/22 - 08/25	**Japanese Government Fellowship**	**JAPAN STUDENT SERVICES ORGANIZATION**
09/21	Computational Resource Support (2,048GPUs x 24hours)	ABCI GRAND CHALLENGE PROGRAM (AIST)
05/21	Computational Resource Support (512GPUs x 168hours)	GRAND-CHALLENGE PROGRAM ON TSUBAME3.0
09/20 - 08/25	Tuition Fee Reduction Scholarship (Approx 68,300 USD)	UNIVERSITÉ DE MONTRÉAL
04/20 - 03/22 (decline)	Fellowship for Young Scientists (Approx 31,400 USD)	JAPAN SOCIETY FOR THE PROMOTION OF SCIENCE
04/20 - 03/22 (decline)	Research Funding (Approx 13,000 USD)	JSPS KAKENHI (20J13997)
12/19	Travel Grant for NeurIPS2019	JAPANESE SOCIETY FOR ARTIFICIAL INTELLIGENCE
04/19 - 03/22	Tokyo Tech Tsubame Scholarship	TOKYO INSTITUTE OF TECHNOLOGY
07/17 - 08/22	**Full Scholarship - Support of Tuition/Living Expenses/Research Funding**	**MASASON FOUNDATION** [1]

---

[1]A prestigious and highly competitive scholarship founded by SoftBank CEO Masayoshi Son. Awarded to <100 individuals under 25 globally for outstanding achievement.

04/17 - 03/19	Partial Support of Tuition/Living Expenses	JAPAN STUDENT SERVICES ORGANIZATION
02/19	Travel Grant for PPoPP2019	JAPAN GOVERNMENT MEXT
09/17 - 03/18	Scholarship **(Approx 2,800 USD)**	SHINANO-IKUEIKAI
11/16	Travel Grant for Undergraduates Program	SUPERCOMPUTING CONFERENCE (SC16)
03/13 (decline)	Special Scholarship for Top Score Students **(Acceptance Rate < 0.06% =(70/109,934))**	MEIJI UNIVERSITY

Peer-Reviewed Journal Publications

Authors who equally contributed to a publication are marked with a [†].

1. **Hiroki Naganuma**[†], Ryuichiro Hataya[†], Kotaro Yoshida, Ioannis Mitliagkas, "An Empirical Study of Pre-trained Model Selection for Out-of-Distribution Generalization and Calibration." Transactions on Machine Learning Research **(TMLR 2025)** (04/25)

2. Masanari Kimura[†], **Hiroki Naganuma**[†]. "Geometric Insights into Focal Loss: Reducing Curvature for Enhanced Model Calibration." **Pattern Recognition Letters** (02/25)

3. Kotaro Yoshida[†], **Hiroki Naganuma**[†]. "Towards Understanding Variants of Invariant Risk Minimization from the Perspective of Calibration." Transactions on Machine Learning Research **(TMLR 2024)** (06/24)

4. **Hiroki Naganuma**, Kartik Ahuja, Shiro Takagi, Tetsuya Motokawa, Rio Yokota, Kohta Ishikawa, Ikuro Sato, Ioannis Mitliagkas. "Empirical Study on Optimizer Selection for Out-of-Distribution Generalizations." Transactions on Machine Learning Research **(TMLR 2023)** (06/23)

Peer-Reviewed Conference Publications

5. Kotaro Yoshida, Yuji Naraki, Takafumi Horie, Ryosuke Yamaki, Ryotaro Shimizu, Yuki Saito, Julian McAuley, **Hiroki Naganuma**. "Mastering Task Arithmetic: τJp as a Key Indicator for Weight Disentanglement." The 13th International Conference on Learning Representations **(ICLR 2025)** (Singapore) (04/25)

6. Charles Guille-Escuret[†], **Hiroki Naganuma**[†], Kilian Fatras, Ioannis Mitliagkas. "No Wrong Turns: The Simple Geometry Of Neural Networks Optimization Paths." International Conference on Machine Learning **(ICML 2024)** (Vienna, Austria) (07/24)

7. Keigo Tada, **Hiroki Naganuma**. "How Image Corruption and Perturbation Affect Out-Of-Distribution Generalization and Calibration." International Joint Conference on Neural Networks (IJCNN 2023) (Gold Coast, Australia) (06/23) **[Travel Grant]**

8. **Hiroki Naganuma**, Hideaki Iiduka. "Conjugate Gradient Method for Generative Adversarial Networks." The 26th International Conference on Artificial Intelligence and Statistics **(AISTATS 2023)** (Valencia, Spain) (05/23)

9. Kilian Fatras, **Hiroki Naganuma**, Ioannis Mitliagkas. "Optimal Transport Meets Noisy Label Robust Loss and MixUp Regularization for Domain Adaptation." The Conference on Lifelong Learning Agents **(CoLLAs 2022)** (Montreal, Canada) (08/22)

10. **Hiroki Naganuma**, Rio Yokota. "Accelerating Convolutional Neural Networks Using Low Precision Arithmetic." The International Conference on High Performance Computing in Asia-Pacific Region (HPC Asia 2018) (Tokyo, Japan) (01/18)

11. Kazuki Osawa, Akira Sekiya, **Hiroki Naganuma**, and Rio Yokota. "Accelerating Matrix Multiplication in Deep Learning by using Low-Rank Approximation." The 2017 International Conference on High Performance Computing & Simulation (HPCS 2017) (Genoa, Italy) (07/17)

Preprint / Papers Under Review

12. **Hiroki Naganuma**[†], Xinzhi Zhang[†], Man-Chung Yue, Ioannis Mitliagkas, Russell J. Hewett, Philipp Andre Witte, Yin Tat Lee. "Pseudo-Asynchronous Local SGD: Robust and Efficient Data-Parallel Training." (Under Review) (04/25)

13. **Hiroki Naganuma**, "When Does Alignment Help? A Comparative Study of DCCA and Fusion-Based Approaches for Multi-modal Chest X-ray Analysis" (Preprint) (04/25)

14. **Hiroki Naganuma**[†], Kotaro Yoshida[†], Laura Gomezjurado Gonzalez[†], Takafumi Horie, Yuji Naraki, Ryotaro Shimizu, . "On Fairness of Task Arithmetic: The Role of Task Vectors." (Under Review) (04/25)

15. Yuji Naraki[†], Ryosuke Yamaki[†], Yoshikazu Ikeda, Takafumi Horie, Kotaro Yoshida, Ryotaro Shimizu, **Hiroki Naganuma**. "Augmenting NER Datasets with LLMs: Towards Automated and Refined Annotation." (Arxiv Preprint) (03/24)

16. **Hiroki Naganuma**, Taiji Suzuki, Rio Yokota, Masahiro Nomura, Kohta Ishikawa, Ikuro Sato. "Takeuchi's Information Criteria as Generalization Measures for DNNs Close to NTK Regime." (Preprint) (09/21)

International Workshop and Symposium Publications

17. Kotaro Yoshida, Yuji Naraki, Takafumi Horie, Ryosuke Yamaki, Ryotaro Shimizu, Yuki Saito, Julian McAuley, **Hiroki Naganuma**. "Mastering Task Arithmetic: τJp as a Key Indicator for Weight Disentanglement." NeurIPS2024 Workshop on Fine-Tuning in Modern Machine Learning: Principles and Scalability (Vancouver, Canada) (12/24)

18. **Hiroki Naganuma**[†], Xinzhi Zhang[†], Man-Chung Yue, Ioannis Mitliagkas, Russell J. Hewett, Philipp Andre Witte, Yin Tat Lee. "Pseudo-Asynchronous Local SGD: Robust and Efficient Data-Parallel Training." NeurIPS2024 Workshop on Optimization (Vancouver, Canada) (12/24)

19. **Hiroki Naganuma**[†], Junhyung Lyle Kim[†], Anastasios Kyrillidis, Ioannis Mitliagkas. "Smoothness-Adaptive Sharpness Aware Minimization for Finding Flatter Minima." ICLR2024 Workshop on Practical ML for Low Resource Settings, 2024 (Vienna, Austria) (05/24)

20. Haruka Kumagai, Ryosuke Yamaki, **Hiroki Naganuma**. "Story-to-Images Translation: Leveraging Diffusion Models and Large Language Models for Sequence Image Generation." ACMMM2023 Workshop on User-centric Narrative Summarization of Long Videos (NarSUM) (Ottawa, Canada) (10/23) **[Travel Grant]**

21. **Hiroki Naganuma**, Ryuichiro Hataya. "An Empirical Investigation of Pre-trained Model Selection for Out-of-Distribution Generalization and Calibration." ICCV2023 Workshop on Uncertainty Quantification for Computer Vision (UNCV) (Paris, France) (10/23)

22. **Hiroki Naganuma**[†], Masanari Kimura[†]. "On the Interplay of Curvature, Calibration and Out-of-Distribution Generalization: Insights from SAM and Focal Loss Analyses." ICCV2023 Workshop on Uncertainty Quantification for Computer Vision (UNCV) (Paris, France) (10/23)

23. **Hiroki Naganuma**[†], Masanari Kimura[†]. "Necessary and Sufficient Hypothesis of Curvature: Understanding Connection Between Out-of-Distribution Generalization and Calibration." ICLR 2023 Workshop on Domain Generalization (Kigali Rwanda) (05/23)

24. **Hiroki Naganuma**, Kartik Ahuja, Ioannis Mitliagkas, Shiro Takagi, Tetsuya Motokawa, Rio Yokota, Kohta Ishikawa, Ikuro Sato. "Empirical Study on Optimizer Selection for Out-of-Distribution Generalization." NeurIPS2022 Workshop on Distribution Shifts (DistShift) (New Orleans, US) (12/22)

25. **Hiroki Naganuma**. "Towards Understanding the Relationship of Batch Size and Iterations in Deep Learning." The Machine Learning Summer School (MLSS2020 in Tübingen) (Tübingen, Germany (Online)) (06/20)

26. **Hiroki Naganuma**, Rio Yokota. "On Empirical Analysis of Layer-wise Learning Rate Schedule." The 11th Asian Conference on Machine Learning (ACML 2019) Workshop (Nagoya, Japan) (11/19)

27. **Hiroki Naganuma**. "Investigation of Second-Order Optimization in Large Mini-batch Training." International HPC Summer School, 2019 (IHPCSS2019) (Kobe, Japan) (07/19)

28. **Hiroki Naganuma**, Rio Yokota. "A Performance Improvement Approach for Second-Order Optimization in Large Mini-batch Training." 2nd High Performance Machine Learning Workshop held in conjunction with CCGrid2019 (HPML2019) (Larnaca, Cyprus) (05/19)

29. **Hiroki Naganuma**, Rio Yokota. "Noise Injection Leads to Better Generalization in Large Mini-Batch Training." Tokyo Institute of Technology and Stony Brook University Joint Science and Technology Meeting (Tokyo, Japan) (05/19)

30. Minatsu Sugimoto, Eitaro Yamatsuta, **Hiroki Naganuma**, Riku Arakawa, Yudai Ushio, Masayuki Teramoto. "Design of smart pacifier detecting dehydration symptoms in baby's and sharing parents behavior." Tsukuba Global Science Week 2018 (TGSW2018) (Ibaraki, Japan) (09/18) **[Travel Grant]**

Domestic Journal and Conference

31. Kotaro Yoshida[†], Takafumi Horie[†], Yuji Naraki, Ryotaro Shimizu, Ryosuke Yamaki, **Hiroki Naganuma**. "Evaluating the Effectiveness of Model Linearization in Task Analogies." The 39th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI2025) (Osaka, Japan) (05/25)

32. Kotaro Yoshida[†], **Hiroki Naganuma**[†]. "Towards Understanding Variants of Invariant Risk Minimization from the Perspective of Calibration." The 1st Symposium of Young Researcher Association for Machine Learning (YAML2024) (Shizuoka, Japan) (09/24) **[Travel Grant]**

33. Kotaro Yoshida[†], Yuji Naraki[†], Ryosuke Yamaki[†], Takafumi Horie, Ryotaro Shimizu, Yuki Saito, **Hiroki Naganuma**. "A Closer Look at Task Analogies: Insights from Function and Parameter Space." The 19th Symposium of Young Researcher Association for NLP Studies (YANS2024) (Osaka, Japan) (09/24) **[Travel Grant]**

34. Yuji Naraki[†], Ryosuke Yamaki[†], Yoshikazu Ikeda, Takafumi Horie, **Hiroki Naganuma**. "Toward Cost-Effective Named Entity Recognition: Navigating the Complexities of Automated Annotation through Large Language Models." The 30th Annual Meeting of the Association for Natural Language Processing (NLP2024) (Kobe, Japan) (03/24)

35. Kaisei Takahashi, **Hiroki Naganuma**. "Averaging Multi-domain Models for Continual Reinforcement Learning to Mitigate Catastrophic Forgetting." The 86th National Convention of Information Processing Society of Japan (IPSJ2024) (Yokohama, Japan (Online)) (03/24)

36. Kotaro Yoshida[†], **Hiroki Naganuma**[†]. "Towards Understanding Variant Methods of Invariant Risk Minimization from the Perspective of Calibration." The 86th National Convention of Information Processing Society of Japan (IPSJ2024) (Yokohama, Japan (Online)) (03/24) **[Student Encouragement Award]**

37. Kaisei Takahashi, **Hiroki Naganuma**. "Averaging Multi-domain Models for Robust Deep Reinforcement Learning under Distribution Shifts." Forum for Information and Technology 2023 (FIT2023) (Osaka, Japan (Online)) (09/23) **[FIT Encouragement Award]**

38. Kotaro Yoshida, Keigo Tada[†], **Hiroki Naganuma**[†]. "Uncertainty Calibration in Deep Neural Networks through Invariant Risk Minimization." Forum for Information and Technology 2023 (FIT2023) (Osaka, Japan (Online)) (09/23)

39. **Hiroki Naganuma**. "Comprehensive Evaluation of the Next Generation Optimizers on Out of Distribution Generalization." The 15th JHPCN symposium (Tokyo, Japan (Online)) (07/23)

40. Haruka Kumagai, Ryosuke Yamaki, **Hiroki Naganuma**. "Diffusion Model Incorporating Context Information of Narratives." Meeting on Image Recognition and Understanding 2023 (MIRU2023) (Shizuoka, Japan) (07/23)

41. Keigo Tada, **Hiroki Naganuma**. "Measuring the Effect of Image Corruption and Perturbation through the Lens of Calibration and Out-of-Distribution Generalization." The 85th National Convention of Information Processing Society of Japan (IPSJ2023) (Tokyo, Japan (Online)) (03/23) **[Student Encouragement Award]**

42. Kaisei Takahashi, **Hiroki Naganuma**. "A Weight Averaging Method Mitigates the Effect of Distribution Shift in Deep Reinforcement Learning." The 85th National Convention of Information Processing Society of Japan (IPSJ2023) (Tokyo, Japan (Online)) (03/23)

43. Ryosuke Yamaki, Yuji Naraki, **Hiroki Naganuma**. "Analysis of Script Data with Large Language Model: Exploring and Evaluating Product Placement Insertion Points." The 29th Annual Meeting of the Association for Natural Language Processing (NLP2023) (Okinawa, Japan) (03/23) **[Hakuhodo DY Media Partners Inc. Award]**

44. Keigo Tada[†], **Hiroki Naganuma**[†]. "Effect of Distribution Shift on Out-of-Distribution Generalization and Uncertainty." Forum for Information and Technology 2022 (FIT2022) (Kanagawa, Japan (Online)) (09/22)

45. Kaisei Takahashi[†], **Hiroki Naganuma**[†]. "A Preliminary Study of Explore-Exploit Tradeoff in Reinforcement Learning with Stochastic Weight Averaging." Forum for Information and Technology 2022 (FIT2022) (Kanagawa, Japan (Online)) (09/22)

46. **Hiroki Naganuma**, Gaku Fujimori, Mari Takeuchi, Jumpei Nagase. "Scheduling of Damping in Natural Gradient Method." The 36th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI2022) (Kyoto, Japan (Online)) (06/22)

47. Gaku Fujimori, Jumpei Nagase[†], **Hiroki Naganuma**[†]. "The Effect of Damping on Positive Definiteness of Fisher Information Matrix in Natural Gradient Descent and Escaping Saddle Point." The 84th National Convention of Information Processing Society of Japan (IPSJ2022) (Ehime, Japan (Online)) (03/22)

48. Gaku Fujimori, Jumpei Nagase, **Hiroki Naganuma**. "Effects of Adaptive Damping in Second-Order Optimization of DNN Models with Skip-Connection." Japan Society for Industrial and Applied Mathematic (JSIAM2021) (Saitama, Japan (Online)) (09/21)

49. Jumpei Nagase, **Hiroki Naganuma**. "Trainability of Neural Networks with Skip-Connection by Preconditioned Optimizer." Forum for Information and Technology 2021 (FIT2021) (Online) (08/21)

50. Yuji Naraki, Tetsuya Motokawa, **Hiroki Naganuma**. "Relationship between Consistency Loss Decay and Generalization in Semi-Supervised Learning." The 83rd National Convention of Information Processing Society of Japan (IPSJ2021) (Online) (03/21)

51. Tetsuya Motokawa, **Hiroki Naganuma**, Tatsuro Ide. "A Study of the Effect on Hessian Spectrum according to the Approximation Method in Deep Neural Networks." The 83rd National Convention of Information Processing Society of Japan (IPSJ2021) (Online) (03/21)

52. **Hiroki Naganuma**, Masahiro Nomura, Rio Yokota. "An Analysis of Generalization Measures in Training Process of DNN, and Application to Hyperparameter Optimization." The 23rd Information-Based Induction Sciences Workshop (IBIS2020) (Tsukuba, Japan (Online)) (11/20)

53. Takahiro Shohata, **Hiroki Naganuma**, Rio Yokota. "An Experimental Analysis of Hyperparameters Effect to Generalization in Stochastic Weight Averaging." The 23rd Information-Based Induction Sciences Workshop (IBIS2020) (Tsukuba, Japan (Online)) (11/20)

54. Yuji Naraki, Tetsuya Motokawa, **Hiroki Naganuma**. "Towards Understanding the Effect of Model Averaging in Semi-Supervised Learning from the Perspective of Hessian Spectrum Analysis." The 23rd Information-Based Induction Sciences Workshop (IBIS2020) (Tsukuba, Japan (Online)) (11/20)

55. Tatsuro Ide, Tetsuya Motokawa, **Hiroki Naganuma**. "Performance Evaluation of Approximation Methods of Calculation for Information Matrix." The 23rd Information-Based Induction Sciences Workshop (IBIS2020) (Tsukuba, Japan (Online)) (11/20)

56. **Hiroki Naganuma**, Rio Yokota. "A Preliminary Study of the LARS Effect to Large Batch Training Problem." CREST AI Projects Joint Meeting 2020 (Online) (09/20)

57. **Hiroki Naganuma**, Tatsuro Ide, Rio Yokota. "Measuring the Effects to Beneficial Batch Size and Required Iteration by LARS on Neural Network Training." The 34th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI2020) (Kumamoto, Japan (Online)) (06/20)

58. Takahiro Shohata, **Hiroki Naganuma**, Rio Yokota. "Verification of Generalization Performance Improvement using Stochastic Weight Averaging in Large Batch Training." The 82nd National Convention of Information Processing Society of Japan (IPSJ2020) (Kanazawa, Japan (Online)) (03/20)

59. **Hiroki Naganuma**, Rio Yokota. "Proposal of Initial Learning Rate Determination Method in Training with Layer-wise Adaptive Rate Scaling." The 22nd Information-based Induction Sciences Workshop (IBIS2019) (Nagoya, Japan) (11/19)

60. Keiko Yamamoto, **Hiroki Naganuma**, Toshiya Yui, Masafumi Harada, Kazuyoshi Fukutani, Osao Hori, Takayuki Fukusawa. "Astral Body: The Expression of "Sense of Life" by a Kinetic Surface Using Ferromagnetic Powder." Transactions of the Virtual Reality Society of Japan, Vol. 24, No. 3 (09/19)

61. **Hiroki Naganuma**, Rio Yokota. "A Preliminary Study of Adaptive Learning Rate for Stochastic Optimization in Convolutional Neural Network." CREST 3 AI Projects Joint Meeting 2019 (Kanagwa, Japan) (08/19)

62. **Hiroki Naganuma**, Rio Yokota. "Investigation of Smoothing in Natural Gradient Method for Large Mini-batch Training." The 3rd cross-disciplinary Workshop on Computing Systems, Infrastructures, and Programming (xSig2019) (Kanagwa, Japan) (05/19) **[Outstanding Presentation Award]**

63. **Hiroki Naganuma**, Shun Iwase, Rio Yokota. "Verification of the Reducing the Number of Iterations in Large Mini-Batch Training by Applying Mixup." The 3rd cross-disciplinary Workshop on Computing Systems, Infrastructures, and Programming (xSig2019) (Kanagwa, Japan) (05/19)

64. **Hiroki Naganuma**, Rio Yokota. "A Study on Generalization Performance Improvement Method on Large Batch Learning Using Averaging by Noise Injection." The Institute of Electronics Information and Communication Engineers General Conference 2019 (IEICE2019) (Tokyo, Japan) (03/19) **[Poster Award]**

65. **Hiroki Naganuma**, Rio Yokota. "Smoothing of Objective Function for Large Scale Parallel Deep Learning." The 81st National Convention of Information Processing Society of Japan (IPSJ2019) (Fukuoka, Japan) (03/19)

66. **Hiroki Naganuma**, Rio Yokota. "Smoothing of the Objective Function in Stochastic Optimization for Large Scale Parallel Learning." CREST-Deep Symposium (Tokyo, Japan) (11/18)

67. **Hiroki Naganuma**, Shun Iwase, Kinsho Kaku, Hikaru Nakata, Rio Yokota. "Hyperparameter Optimization of Large Scale Parallel Deep Learning using Natural Gradient Approximation Method." Forum for Information and Technology 2018 (FIT2018) (Fukuoka, Japan) (09/18)

68. **Hiroki Naganuma**, Akira Sekiya, Kazuki Osawa, Hiroyuki Ootomo, Yuji Kuwamura, Rio Yokota. "Verification of speeding up using low precision arithmetic in convolutional neural network." GTC Japan 2017 Poster Session (Tokyo, Japan) (12/17)

69. **Hiroki Naganuma**, Akira Sekiya, Kazuki Osawa, Hiroyuki Ootomo, Yuji Kuwamura, Rio Yokota. "Improvement of speed using low precision arithmetic in deep learning and performance evaluation of accelerator." Technical Committee on Pattern Recognition and Media Understanding (PRMU) (Kumamoto, Japan) (10/17)

70. Kazuki Osawa, Akira Sekiya, **Hiroki Naganuma**, Rio Yokota. "Accelerating Convolutional Neural Networks Using Low-Rank Tensor Decomposition." Technical Committee on Pattern Recognition and Media Understanding (PRMU) (Kumamoto, Japan) (10/17)

71. **Hiroki Naganuma**, Kazuki Osawa, Akira Sekiya, Rio Yokota. "Acceleration of Compression Model Using Half-Precision Arithmetic in Deep Learning." Japan Society for Industrial and Applied Mathematic (JSIAM2017) (Tokyo, Japan) (09/17)

72. Kazuki Osawa, Akira Sekiya, **Hiroki Naganuma**, Rio Yokota. "Accelerating Convolutional Neural Networks Using Low-Rank Approximation." Proceedings of the Conference on Computational Engineering and Science Vol. 22 (Saitama, Japan) (05/17)

73. Akira Sekiya, Kazuki Osawa, **Hiroki Naganuma**, Rio Yokota. "Acceleration of Matrix Multiplication of Deep Learning Using Low Rank Approximation." 158th Research Presentation Seminar in High Performance Computing (Shizuoka, Japan) (05/17)

Skills

Programming: Python, C++ , HPC (CUDA, MPI, NCCL, OpenMP)
Framework: PyTorch, Jax, Horovod, LLM (DeepSpeed, Apex, GPT-NeoX, Huggingface)
DevOps/Tools: Git, Slurm, Wandb, Docker, Azure, GCP
Languages: Japanese (Native), English (Proficient), French (Beginner), Mandarin (Beginner)
Personal: Teamwork, Leadership, Achiever, Analytical / based on Clifton Strengthsfinder

Academic Service

2023 - present
Reviewer
COLM 2025
AISTATS 2024, 2025 **(Best Reviewer)**
NeurIPS 2024 **(Top Reviewer)**
ICLR 2025
Workshops: ICLR2023 Domain Generalization

Mentoring

Predoctoral Students

08/24 - present
Laura Gomezjurado Gonzalez
Undergraduate Student at Stanford University → Research Intern at Microsoft Research
Under Review

04/23 - present
Kotaro Yoshida
Undergraduate Student at Tokyo Institute of Technology → Research Intern at A*STAR Singapore
TMLR2024, ICLR2025

04/22 - 03/23
Keigo Tada
Master Student at Ritsumeikan University → SONY
IJCNN2023

Talks

03/25   Spring Seminor Series: Ottawa Japanese Researchers Association (Host: Yasutsugu Tamura)

11/24   The 25th Lecture Series: Japan Healthcare Innovation Hub (Host: Anna Nakayasu)

09/24   Seminor Series: Montreal Academy Club (Host: Norio Tomita)

08/24   The 4th Lecture Series: Seattle Japanese Research Community (Host: Masa Nakashima)

11/23   The 8th Mathematical Optimization Laboratory Seminar at Meiji University (Host: Hideaki Iiduka)

02/23   RICOS×ZOZO NeurIPS2022 Paper Reading Group (Host: Masanari Kimura)

01/22   The 54th AI Seminar ABCI Grand Challenge (Host: Yusuke Tanimura)

06/21	The 6th Mathematical Optimization Laboratory Seminar at Meiji University (Host: Hideaki Iiduka)
02/22	Montréal Machine Learning and Optimization (Host: Ioannis Mitliagkas)
08/20	Summer Seminer Series: The Japanese Graduate Student Association in the United States (JGSAU) (Host: Kai Narita)
05/20	The 78th Artificial Intelligence Seminar: AI Trend - Top Conference Report Session (NeurIPS2019) (Host: Togami Masahito)

Extracurricular Activities

08/24 - 09/24 Kids Code Club   Seattle, WA, United States
Teaching Volunteer
Volunteered as an instructor for a programming class serving junior high school students.

02/23 - present JSPS Alumni Association - Québec Chapter   Montréal, QC, Canada
Founding Member
Organized academic research exchange events to foster networking and collaboration among JSPS award recipients.

02/23 - present Montréal Japanese Academy Society   Montréal, QC, Canada
President
Organized academic research exchange events to foster networking and collaboration among scholars and students; offered support to academics and their families by coordinating social gatherings and providing tailored resources.

09/22 - 08/25 The Japanese Graduate Student Association in the United States   Online
Divisional Director
Organized a series of seminars on pursuing graduate schools abroad, overseeing a session on "Families and Partnerships."

09/22 - 03/23 NEDO Technology Commercialization Program / FY2022   Tokyo, Japan
Selected Attendee
My proposal was selected for a research and development commercialization program. I participated in a training program sponsored by a Japanese government-contracted organization, receiving mentoring and support for commercialization.

07/20 The Machine Learning Summer School 2020   Tübingen, Germany
Selected Attendee
Selected to attend MLSS 2020 in Tübingen, Germany, featuring top global instructors. Out of more than 1300 applicants, only 180 were admitted (acceptance rate <14%).

03/20 Top Conference Reports: NeurIPS 2019 (Supported by JSAI)   Tokyo and Osaka, Japan
Selected Reporter and Presenter
Delivered a presentation titled "Trends in Deep Learning Theory at NeurIPS 2019" at the 78th Artificial Intelligence Seminar (AI Trends and Top Conferences Debriefing Session), highlighting key insights in deep learning theory observed at NeurIPS 2019.

09/19 Leadership Training Programs powered by JT   Nagano, Japan
Selected Attendee
Chosen from among more than 1,000 applicants to participate in a leadership workshop led by prominent astronauts, management consultants, and brain scientists in Japan.

01/19 International HPC Summer School 2019   Kobe, Japan
Selected Attendee
Attended an international program offering lectures and workshops by leading HPC scientists, providing opportunities to expand knowledge, network globally, and explore cutting-edge HPC research.

01/19 Georgia Institute of Technology Leadership Program   Atlanta, GA, United States
Selected Attendee
Completed a month-long leadership training program at Georgia Institute of Technology, featuring lectures and workshops that systematically covered leadership principles. This experience honed my leadership skills through valuable insights and feedback from knowledgeable instructors.

03/18   KOBE HPC Spring School   Kobe, Japan
Selected Attendee

Participated in a workshop led by leading Japanese HPC researchers, focusing on HPC techniques for students in related fields. Learned to implement communication overlap techniques, crucial in distributed deep learning research.

11/18   Gpu Technology Conference Japan, Deep Learning Institute Workshop   Tokyo, Japan
Teaching Assistant

Served as a teaching assistant, instructing participants on NVIDIA DIGITS and guiding them in building CNN and RNN models. Also introduced basic Python/Jupyter techniques for beginners.

11/17 - 12/19   Organization for Inviting EXPO2025 to OSAKA   Osaka, Japan
Student Representative of Computer Science Major in Japan

Served as the student representative for Computer Science in Japan. Presented medical data analysis software and wearable-device-based health management applications to the BIE team in March 2018, contributing to Osaka's successful EXPO2025 bid.

05/17   TEDxUTokyo 2017   Tokyo, Japan
Visual Director

Oversaw the visual direction for the opening and finale of TEDxUTokyo 2017 using computer vision and computer graphics. Developed a real-time large-scale data analysis system and performed interactive content integrated with Facebook.

11/16   SC16 HPC for Undergraduates Program   Salt Lake City, UT, United States
Selected Attendee

Became the first Japanese undergraduate selected for the HPC for Undergraduates program at SC16, a premier high-performance computing conference. Received committee sponsorship for travel, accommodation, and conference participation expenses, and learned from leading HPC experts.

09/16   Musashino Art University Festival Projection Mapping Project 2015   Tokyo, Japan
Leader, Chief Engineer

Led the projection mapping project for the opening and finale of the Musashino Art University Art Festival 2015. Executed real-time large-scale data analysis and developed interactive Twitter-linked content to foster audience engagement.

08/16 - 03/19   Tokyo Institute of Technology Advertisement and Media Art Society (TITAMAS)   Tokyo, Japan
Vice President

Served as Vice President when TITAMAS was chosen to represent Japan at the Microsoft Imagine Cup 2017, one of the world's largest IT development competitions. Led active participation in domestic and international programming contests, achieving multiple accolades.

08/16   Supercomputing Contest 2016 Held in Tokyo Institute of Technology and U Osaka   Osaka, Japan
Teaching Assistant

Taught high school students how to use the Linux system and run scripts on the TSUBAME 2.5 supercomputer at Tokyo Institute of Technology. Also introduced basic C++/CUDA techniques for the Supercomputing Contest.

03/16   Silicon Valley Workshop 2016 Supported by Recruit Holdings   San Francisco, CA, United States
Selected Attendee, 1st Generation

Selected as a distinguished student engineer in Japan to join the 2016 Silicon Valley Workshop, sponsored by Recruit Holdings. Spent 10 days in San Francisco and Silicon Valley, attending lectures at Stanford University and touring IT companies in Palo Alto, gaining exposure to cutting-edge technology and diverse perspectives.