

# Hiroki NAGANUMA

PH.D. CANDIDATE IN COMPUTER SCIENCE

Mila - Quebec AI Institute, 6666 Rue Saint-Urbain, Montréal, QC H2S 3H, Canada

☎ (+1) 514-559-1008 | ✉ naganuma.hiroki@mila.quebec | 🏠 hiroki11x.github.io | 📧 hiroki11x | 🌐 hiroki11x | 🎓 Hiroki Naganuma

“ I am a Ph.D. candidate in Computer Science at the Université de Montréal and Mila-Quebec AI Institute. My research interests center around large-scale parallelization in machine learning and understanding the training dynamics of deep neural networks. My foundational experience in distributed and parallel computing was honed Prior to my doctoral studies, I contributed to the field of distributed and parallel computing at the Tokyo Institute of Technology. More recently, my research has pivoted towards addressing the challenges in practical scenarios of large-scale networks, aiming to enhance system robustness and reliability. ”

## Education

### Université de Montréal

Montreal, Canada

PH.D. STUDENT → CANDIDATE IN COMPUTER SCIENCE

Supervisor: Prof. Ioannis Mitliagkas

09/20 - 11/25 (Expected)

- Mila - Quebec Artificial Intelligence Institute
- Topic: Uncertainty, Out-of-Distribution and Optimization on Deep Neural Networks

### Tokyo Institute of Technology

Tokyo, Japan

M.Sc. IN COMPUTER SCIENCE

Supervisor: Prof. Rio Yokota

04/17 - 03/19

- **Completed Master Course with Summa Cum Laude (Graduate First on the List, GPA:4.0/4.0)**
- Exchange Student at Georgia Tech
- Thesis: Smoothing Objective Function in Stochastic Optimization for Large Scale Deep Learning

### Tokyo Institute of Technology

Tokyo, Japan

B.Sc. IN COMPUTER SCIENCE

Supervisor: Prof. Rio Yokota

04/13 - 03/17

- **Completed Bachelor Course with Dean's Award, School of Computing**
- Thesis: Hierarchical Low Rank Approximation for Convolutional Neural Networks

### Chuo University Suginami High School

Tokyo, Japan

HIGH SCHOOL DIPLOMA

04/10 - 03/13

- **Tokyo Metropolitan Governor Award (Graduate First on the List, Grade 4.9/5.0)**
- Thesis: Real-time Social Media Analytics in the Earthquake

## Skills

<b>Programming</b>	Python (5 years), C++ (7 years), <b>HPC skills (CUDA, MPI, NCCL, OpenMP / 5 years)</b>
<b>Framework</b>	PyTorch (5 years), Horovod (5 years), TensorFlow (2 years)
<b>DevOps / Tools</b>	Git (7 years), Slurm (7 years), Wandb (4 years), Docker, Virtual Box, Jenkins, CircleCI
<b>Languages</b>	Japanese (Native), English (Proficient), French (Beginner), Mandarin (Beginner)
<b>Personal</b>	Individualization(Teamwork), Maximizer(Leadership), Achiever, Analytical / based on Clifton Strengthsfinder

## Research Experience

### Mila - Quebec AI Institute

Montreal, Canada

PHD STUDENT → CANDIDATE

Supervisor: Prof. Ioannis Mitliagkas

09/20 - 12/25 (Expected)

- Uncertainty, Out-of-Distribution and Optimization on Deep Neural Networks

### Google DeepMind

Mountain View, U.S.

STUDENT RESEARCHER

Supervisor: Dr. George E. Dahl

09/24 - 02/25

- Machine Learning Foundations

### Microsoft Research

Redmond, U.S.

RESEARCH INTERN

Supervisor: Dr. Philipp Witte and Dr. Russell Hewett

06/24 - 09/24

- Large Language Models Foundations

### ZOZO Research

Tokyo, Japan

RESEARCH FELLOW

Supervisor: Dr. Masanari Kimura

10/22,23 - 05/23,24

- Information Geometry and Calibration

## LinkedIn - AI Foundation Team

RESEARCH INTERN

Supervisor: Dr. Chengming Jiang and Dr. Mingzhou Zhou

Sunnyvale, U.S.

06/22,23 - 09/22,23

- Calibration of DNN Model Prediction for Time Series Data

## LinkedIn - AI Foundation Team

RESEARCH INTERN

Supervisor: Mr. Aman Gupta and Dr. S. Sathiya Keerthi

Sunnyvale, U.S.

06/21 - 09/21

- Mitigating Trade-off between Calibration and AUC

## DENSO IT Laboratory

RESEARCH ASSISTANT

Supervisor: Prof. Ikuro Sato

Tokyo, Japan

05/21 - 04/23

- Generalization and Optimization in Deep Learning

## Cyber Agent - AI Lab

RESIDENCY PROGRAM

Supervisor: Mr. Masahiro Nomura

Tokyo, Japan

08/20 - 12/20

- Efficient Hyperparameter Optimization for Deep Learning

## Japan Society for the Promotion of Science

RESEARCH FELLOW

Supervisor: Prof. Rio Yokota

Tokyo, Japan

04/20 - 08/20

- Developing Efficient Algorithms based on Smoothing Loss Function for Large Batch Training

## RIKEN AIP - Deep Learning Theory Team

RESEARCH INTERN

Supervisor: Prof. Taiji Suzuki

Tokyo, Japan

09/18 - 10/18

- Large Scale Distributed Deep Learning using Natural Gradient Approximation Method
- Generalization Performance Improvement Method on Large Batch Learning using Averaging by Noise Injection

## IBM Research - Tokyo

STUDENT FELLOW

Supervisor: Dr. Taro Sekiyama and Dr. Kiyokuni Kawachiya

Tokyo, Japan

04/17 - 03/18

- Optimization and Parallelization of Deep Learning Framework
- Development of Communication Library for Distributed Deep Learning

## Tokyo Institute of Technology

RESEARCH ASSISTANT

Supervisor: Prof. Rio Yokota

Tokyo, Japan

04/17 - 03/20

- Compression of DNN Model for Edge Device Inference using Low Rank Approximation
- Accelerating Convolutional Neural Networks using Low Precision Arithmetic
- Smoothing of Objective Function for Large Scale Parallel Deep Learning
- An Empirical Analysis of Layer-wise Learning Rate Schedule

## Work Experience

---

### Mercari Inc.

SOFTWARE ENGINEER AND CONSULTING INTERNSHIP

SanFrancisco, U.S.

02/17 - 03/17

- Implementation of Service Prototype and Improvement Proposal of Mercari and CtoC service

### SORACOM Inc.

SOFTWARE ENGINEER INTERNSHIP

Tokyo, Japan

08/16 - 04/17

- Developed Internal Cloud Computing System using Apache Spark and Mesos

### CyberAgent Inc.

SOFTWARE ENGINEER INTERNSHIP

Tokyo, Japan

08/15 - 12/15

- Developed an Android Application in Internet TV Service AbemaTV Fresh

### Nagase Brothers Inc.

MATHEMATICAL ASSISTANT

Tokyo, Japan

01/14 - 12/16

- Since I got the perfect score in the national center test of maths 200/200, I had the opportunity to help create the practice math exam for the university entrance exams in Japan.

## Major Honors & Awards

---

### ACADEMIC AWARDS

2019 **First Place [Graduate First on the List]**, Complete Master Course with Summa Cum Laude

Tokyo, Japan

2017 **Dean's award**, Dean's Award, School of Computing at Tokyo Institute of Technology

Tokyo, Japan

2016 **First Selection for Japanese**, SC16 HPC for Undergraduates Program

Salt Lake City, U.S

2013 **First Place [Graduate First on the List]**, Tokyo Metropolitan Governor Award

Tokyo, Japan

## OTHERS

2022	<b>Minister of Internal Affairs and Communications Award</b> , Japan Campus Grandprix	Tokyo, Japan
2018	<b>National Winner, Japan Representative [Acceptance Rate 2%]</b> , James Dyson Award 2018	Online
2018	<b>Grand Prize [1st in 12 teams]</b> , Art Hack Day 2018	Tokyo, Japan
2017	<b>Best 32, Japan Representative</b> , Microsoft Imagine Cup 2017 World Finals	Seattle, U.S
2017	<b>3rd Prize [the 3rd in 52 teams]</b> , Stanford Health Hackathon 2017	San Francisco, U.S
2017	<b>First Prize</b> , Tokyo Institute of Technology Engineering Design Competition	Tokyo, Japan
2017	<b>First Prize [Acceptance Rate 3.5%]</b> , 2025 Japan World Expo Committee Creative Competition Digital Creative Department	Osaka, Japan
2016	<b>First Prize [1st in 89 teams]</b> , JPHACKS-2016 Japan's Largest Student Hackathon	Tokyo, Japan
2016	<b>First Prize</b> , MashupAward-2016 Japan's Largest Hackathon Student Division	Tokyo, Japan

## Research Interests

---

- OUT-OF-DISTRIBUTION GENERALIZATION AND CALIBRATION
- OPTIMIZATION FOR DEEP LEARNING
- INFORMATION GEOMETRY
- FOUNDATION MODEL
- HIGH PERFORMANCE COMPUTING

## Publication

---

### PREPRINT / PAPERS UNDER REVIEW

<b>"Augmenting NER Datasets with LLMs: Towards Automated and Refined Annotation"</b> , Yuji Naraki*, Ryosuke Yamaki*, Yoshikazu Ikeda, Takafumi Horie, <u>Hiroki Naganuma</u> (* denotes equal contribution), ARXIV PREPRINT (UNDER REVIEW)	03/24
<b>"Towards Understanding Variants of Invariant Risk Minimization from the Perspective of Calibration"</b> , Kotaro Yoshida*, <u>Hiroki Naganuma</u> * (* denotes equal contribution), ARXIV PREPRINT (UNDER REVIEW)	02/24
<b>"No Wrong Turns: The Simple Geometry Of Neural Networks Optimization Paths"</b> , Charles Guille-Escuret*, <u>Hiroki Naganuma</u> *, Kilian Fatras, Ioannis Mitliagkas (* denotes equal contribution), ARXIV PREPRINT (UNDER REVIEW)	05/23
<b>"Takeuchi's Information Criteria as Generalization Measures for DNNs Close to NTK Regime"</b> , <u>Hiroki Naganuma</u> , Taiji Suzuki, Rio Yokota, Masahiro Nomura, Kohta Ishikawa, Ikuro Sato, PREPRINT (OPENREVIEW)	09/21

### JOURNAL

<b>"Empirical Study on Optimizer Selection for Out-of-Distribution Generalizations"</b> , <u>Hiroki Naganuma</u> , Kartik Ahuja, Shiro Takagi, Tetsuya Motokawa, Rio Yokota, Kohta Ishikawa, Ikuro Sato, Ioannis Mitliagkas, TRANSACTIONS ON MACHINE LEARNING RESEARCH (TMLR)	06/23
--	-------

### INTERNATIONAL CONFERENCE

<b>"How Image Corruption and Perturbation Affect Out-Of-Distribution Generalization and Calibration"</b> , Keigo Tada, <u>Hiroki Naganuma</u> , INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS (IJCNN2023)	Gold Coast, Australia 06/23
<b>"Conjugate Gradient Method for Generative Adversarial Networks"</b> , <u>Hiroki Naganuma</u> , Hideaki Iiduka, THE 26TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND STATISTICS (AISTATS2023)	Valencia, Spain 05/23
<b>"Optimal Transport Meets Noisy Label Robust Loss and MixUp Regularization for Domain Adaptation"</b> , Kilian Fatras, <u>Hiroki Naganuma</u> , Ioannis Mitliagkas, THE CONFERENCE ON LIFELONG LEARNING AGENTS (COLLAs 2022)	Montreal, Canada 08/22
<b>"Accelerating Convolutional Neural Networks Using Low Precision Arithmetic"</b> , <u>Hiroki Naganuma</u> , Rio Yokota, THE INTERNATIONAL CONFERENCE ON HIGH PERFORMANCE COMPUTING IN ASIA-PACIFIC REGION (HPC Asia 2018)	Tokyo, Japan 01/18

**”Accelerating Matrix Multiplication in Deep Learning by using Low-Rank Approximation”**, Kazuki Osawa, Akira Sekiya, Hiroki Naganuma, and Rio Yokota, *Genoa, Italy*  
THE 2017 INTERNATIONAL CONFERENCE ON HIGH PERFORMANCE COMPUTING & SIMULATION (HPCS 2017) 07/17

## INTERNATIONAL WORKSHOP AND SYMPOSIUM

**”Smoothness-Adaptive Sharpness Aware Minimization for Finding Flatter Minima”**, Hiroki Naganuma\*, Junhyung Lyle Kim\*, Anastasios Kyriklidis, Ioannis Mitliagkas (\* denotes equal contribution), *Vienna Austria*  
ICLR2024 WORKSHOP ON PRACTICAL ML FOR LOW RESOURCE SETTINGS, 2024 05/24

**”Story-to-Images Translation: Leveraging Diffusion Models and Large Language Models for Sequence Image Generation”**, Haruka Kumagai, Ryosuke Yamaki, Hiroki Naganuma, *Ottawa, Canada*  
ACMMM2023 WORKSHOP ON USER-CENTRIC NARRATIVE SUMMARIZATION OF LONG VIDEOS (NARSUM) 10/23

**”An Empirical Investigation of Pre-trained Model Selection for Out-of-Distribution Generalization and Calibration”**, Hiroki Naganuma, Ryuichiro Hataya, *Paris, France*  
ICCV2023 WORKSHOP ON UNCERTAINTY QUANTIFICATION FOR COMPUTER VISION (UNCV) 10/23

**”On the Interplay of Curvature, Calibration and Out-of-Distribution Generalization: Insights from SAM and Focal Loss Analyses”**, Hiroki Naganuma\*, Masanari Kimura\* (\* denotes equal contribution), *Paris, France*  
ICCV2023 WORKSHOP ON UNCERTAINTY QUANTIFICATION FOR COMPUTER VISION (UNCV) 10/23

**”Necessary and Sufficient Hypothesis of Curvature: Understanding Connection Between Out-of-Distribution Generalization and Calibration”**, Hiroki Naganuma\*, Masanari Kimura\* (\* denotes equal contribution), *Kigali Rwanda*  
ICLR 2023 WORKSHOP ON DOMAIN GENERALIZATION 05/23

**”Empirical Study on Optimizer Selection for Out-of-Distribution Generalization”**, Hiroki Naganuma, Kartik Ahuja, Ioannis Mitliagkas, Shiro Takagi, Tetsuya Motokawa, Rio Yokota, Kohta Ishikawa, Ikuro Sato, *New Orleans, US*  
NEURIPS2022 WORKSHOP ON DISTRIBUTION SHIFTS (DISTRSHIFT) 12/22

**”Towards Understanding the Relationship of Batch Size and Iterations in Deep Learning”**, Hiroki Naganuma, *Tübingen, Germany (Online)*  
THE MACHINE LEARNING SUMMER SCHOOL (MLSS2020 IN TÜBINGEN) 06/20

**”On Empirical Analysis of Layer-wise Learning Rate Schedule”**, Hiroki Naganuma, Rio Yokota, *Nagoya, Japan*  
THE 11TH ASIAN CONFERENCE ON MACHINE LEARNING (ACML 2019) WORKSHOP 11/19

**”Investigation of Second-Order Optimization in Large Mini-batch Training”**, Hiroki Naganuma, *Kobe, Japan*  
INTERNATIONAL HPC SUMMER SCHOOL 2019 (IHPCSS2019) 07/19

**”A Performance Improvement Approach for Second-Order Optimization in Large Mini-batch Training”**, Hiroki Naganuma, Rio Yokota, *Larnaca, Cyprus*  
2ND HIGH PERFORMANCE MACHINE LEARNING WORKSHOP HELD IN CONJUNCTION WITHCCGRID2019 (HPML2019) 05/19

**”Noise Injection Leads to Better Generalization in Large Mini-Batch Training”**, Hiroki Naganuma, Rio Yokota, *Tokyo, Japan*  
TOKYO INSTITUTE OF TECHNOLOGY AND STONY BROOK UNIVERSITY JOINT SCIENCE AND TECHNOLOGY MEETING 05/19

**”Design of smart pacifier detecting dehydration symptoms in baby’s and sharing parents behavior”**, Minatsu Sugimoto, Eitaro Yamatsuta, Hiroki Naganuma, Riku Arakawa, Yudai Ushio, Masayuki Teramoto, [Excellent Poster Award] *Ibaraki, Japan*  
TSUKUBA GLOBAL SCIENCE WEEK 2018 (TGSW2018) 09/18

## DOMESTIC JOURNAL

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

## DOMESTIC CONFERENCE, SYMPOSIUM, AND WORKSHOP

**”Toward Cost-Effective Named Entity Recognition: Navigating the Complexities of Automated Annotation through Large Language Models”**, Yuji Naraki\*, Ryosuke Yamaki\*, Yoshikazu Ikeda, Takafumi Horie, Hiroki Naganuma, *Kobe, Japan*  
THE 30TH ANNUAL MEETING OF THE ASSOCIATION FOR NATURAL LANGUAGE PROCESSING (NLP2024) 03/24

**”Averaging Multi-domain Models for Continual Reinforcement Learning to Mitigate Catastrophic Forgetting”**, Kaisei Takahashi, Hiroki Naganuma, *Yokohama, Japan (Online)*  
THE 86TH NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2024) 03/24

<p><b>”Towards Understanding Variant Methods of Invariant Risk Minimization from the Perspective of Calibration”</b>, Kotaro Yoshida*, Hiroki Naganuma* (* denotes equal contribution), <b>[Student Encouragement Award]</b>  THE 86TH NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2024)</p>	<p>Yokohama, Japan  (Online)  03/24</p>
<p><b>”Averaging Multi-domain Models for Robust Deep Reinforcement Learning under Distribution Shifts”</b>, Kaisei Takahashi, Hiroki Naganuma, <b>[FIT Encouragement Award]</b>  FORUM FOR INFORMATION AND TECHNOLOGY 2023 (FIT2023)</p>	<p>Osaka, Japan  (Online)  09/23</p>
<p><b>”Uncertainty Calibration in Deep Neural Networks through Invariant Risk Minimization”</b>, Kotaro Yoshida, Keigo Tada*, Hiroki Naganuma* (* denotes equal contribution),  FORUM FOR INFORMATION AND TECHNOLOGY 2023 (FIT2023)</p>	<p>Osaka, Japan  (Online)  09/23</p>
<p><b>”Diffusion Model Incorporating Context Information of Narratives”</b>, Haruka Kumagai, Ryosuke Yamaki, Hiroki Naganuma,  MEETING ON IMAGE RECOGNITION AND UNDERSTANDING 2023 (MIRU2023)</p>	<p>Shizuoka, Japan  07/23</p>
<p><b>”Measuring the Effect of Image Corruption and Perturbation through the Lens of Calibration and Out-of-Distribution Generalization”</b>, Keigo Tada, Hiroki Naganuma, <b>[Student Encouragement Award]</b>  THE 85TH NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2023)</p>	<p>Tokyo, Japan  (Online)  03/23</p>
<p><b>”A Weight Averaging Method Mitigates the Effect of Distribution Shift in Deep Reinforcement Learning”</b>, Kaisei Takahashi, Hiroki Naganuma,  THE 85TH NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2023)</p>	<p>Tokyo, Japan  (Online)  03/23</p>
<p><b>”Analysis of Script Data with Large Language Model: Exploring and Evaluating Product Placement Insertion Points”</b>, Ryosuke Yamaki, Yuji Naraki, Hiroki Naganuma, <b>[Hakuhodo DY Media Partners Inc. Award]</b>  THE 29TH ANNUAL MEETING OF THE ASSOCIATION FOR NATURAL LANGUAGE PROCESSING (NLP2023)</p>	<p>Okinawa, Japan  03/23</p>
<p><b>”Effect of Distribution Shift on Out-of-Distribution Generalization and Uncertainty”</b>, Keigo Tada*, Hiroki Naganuma* (* denotes equal contribution),  FORUM FOR INFORMATION AND TECHNOLOGY 2022 (FIT2022)</p>	<p>Kanagawa, Japan  (Online)  09/22</p>
<p><b>”A Preliminary Study of Explore-Exploit Tradeoff in Reinforcement Learning with Stochastic Weight Averaging”</b>, Kaisei Takahashi*, Hiroki Naganuma* (* denotes equal contribution),  FORUM FOR INFORMATION AND TECHNOLOGY 2022 (FIT2022)</p>	<p>Kanagawa, Japan  (Online)  09/22</p>
<p><b>”Scheduling of Damping in Natural Gradient Method”</b>, Hiroki Naganuma, Gaku Fujimori, Mari Takeuchi, Jumpei Nagase,  THE 36TH ANNUAL CONFERENCE OF THE JAPANESE SOCIETY FOR ARTIFICIAL INTELLIGENCE (JSAI2022)</p>	<p>Kyoto, Japan  (Online)  06/22</p>
<p><b>”The Effect of Damping on Positive Definiteness of Fisher Information Matrix in Natural Gradient Descent and Escaping Saddle Point”</b>, Gaku Fujimori, Jumpei Nagase*, Hiroki Naganuma* (* denotes equal contribution),  THE 84TH NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2022)</p>	<p>Ehime, Japan  (Online)  03/22</p>
<p><b>”Effects of Adaptive Damping in Second-Order Optimization of DNN Models with Skip-Connection”</b>, Gaku Fujimori, Jumpei Nagase, Hiroki Naganuma,  JAPAN SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATIC (JSIAM2021)</p>	<p>Saitama, Japan  (Online)  09/21</p>
<p><b>”Trainability of Neural Networks with Skip-Connection by Preconditioned Optimizer”</b>, Jumpei Nagase, Hiroki Naganuma,  FORUM FOR INFORMATION AND TECHNOLOGY 2021 (FIT2021)</p>	<p>Online  08/21</p>
<p><b>”Relationship between Consistency Loss Decay and Generalization in Semi-Supervised Learning”</b>, Yuji Naraki, Tetsuya Motokawa, Hiroki Naganuma,  THE 83RD NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2021)</p>	<p>Online  03/21</p>
<p><b>”A Study of the Effect on Hessian Spectrum according to the Approximation Method in Deep Neural Networks”</b>, Tetsuya Motokawa, Hiroki Naganuma, Tatsuro Ide,  THE 83RD NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2021)</p>	<p>Online  03/21</p>
<p><b>”An Analysis of Generalization Measures in Training Process of DNN, and Application to Hyperparameter Optimization”</b>, Hiroki Naganuma, Masahiro Nomura, Rio Yokota,  THE 23RD INFORMATION-BASED INDUCTION SCIENCES WORKSHOP (IBIS2020)</p>	<p>Tsukuba, Japan  (Online)  11/20</p>
<p><b>”An Experimental Analysis of Hyperparameters Effect to Generalization in Stochastic Weight Averaging”</b>, Takahiro Shohata, Hiroki Naganuma, Rio Yokota,  THE 23RD INFORMATION-BASED INDUCTION SCIENCES WORKSHOP (IBIS2020)</p>	<p>Tsukuba, Japan  (Online)  11/20</p>
<p><b>”Towards Understanding the Effect of Model Averaging in Semi-Supervised Learning from the Perspective of Hessian Spectrum Analysis”</b>, Yuji Naraki, Tetsuya Motokawa, Hiroki Naganuma,  THE 23RD INFORMATION-BASED INDUCTION SCIENCES WORKSHOP (IBIS2020)</p>	<p>Tsukuba, Japan  (Online)  11/20</p>

<p><b>”Performance Evaluation of Approximation Methods of Calculation for Information Matrix”</b>, <i>Tatsuro Ide, Tetsuya Motokawa, Hiroki Naganuma</i>, THE 23RD INFORMATION-BASED INDUCTION SCIENCES WORKSHOP (IBIS2020)</p>	<p><i>Tsukuba, Japan (Online) 11/20</i></p>
<p><b>”A Preliminary Study of the LARS Effect to Large Batch Training Problem”</b>, <i>Hiroki Naganuma, Rio Yokota</i>, CREST AI PROJECTS JOINT MEETING 2020</p>	<p><i>Online 09/20</i></p>
<p><b>”Measuring the Effects to Beneficial Batch Size and Required Iteration by LARS on Neural Network Training”</b>, <i>Hiroki Naganuma, Tatsuro Ide, Rio Yokota</i>, THE 34TH ANNUAL CONFERENCE OF THE JAPANESE SOCIETY FOR ARTIFICIAL INTELLIGENCE (JSAI2020)</p>	<p><i>Kumamoto, Japan (Online) 06/20</i></p>
<p><b>”Verification of Generalization Performance Improvement using Stochastic Weight Averaging in Large Batch Training”</b>, <i>Takahiro Shohata, Hiroki Naganuma, Rio Yokota</i>, THE 82ND NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2020)</p>	<p><i>Kanazawa, Japan (Online) 03/20</i></p>
<p><b>”Proposal of Initial Learning Rate Determination Method in Training with Layer-wise Adaptive Rate Scaling”</b>, <i>Hiroki Naganuma, Rio Yokota</i>, THE 22ND INFORMATION-BASED INDUCTION SCIENCES WORKSHOP (IBIS2019)</p>	<p><i>Nagoya, Japan 11/19</i></p>
<p><b>”A Preliminary Study of Adaptive Learning Rate for Stochastic Optimization in Convolutional Neural Network”</b>, <i>Hiroki Naganuma, Rio Yokota</i>, CREST 3 AI PROJECTS JOINT MEETING 2019</p>	<p><i>Kanagawa, Japan 08/19</i></p>
<p><b>”Investigation of Smoothing in Natural Gradient Method for Large Mini-batch Training”</b>, <i>Hiroki Naganuma, Rio Yokota, [Outstanding Presentation Award]</i> THE 3RD CROSS-DISCIPLINARY WORKSHOP ON COMPUTING SYSTEMS, INFRASTRUCTURES, AND PROGRAMMING (xSIG2019)</p>	<p><i>Kanagawa, Japan 05/19</i></p>
<p><b>”Verification of the Reducing the Number of Iterations in Large Mini-Batch Training by Applying Mixup”</b>, <i>Hiroki Naganuma, Shun Iwase, Rio Yokota</i>, THE 3RD CROSS-DISCIPLINARY WORKSHOP ON COMPUTING SYSTEMS, INFRASTRUCTURES, AND PROGRAMMING (xSIG2019)</p>	<p><i>Kanagawa, Japan 05/19</i></p>
<p><b>”A Study on Generalization Performance Improvement Method on Large Batch Learning Using Averaging by Noise Injection”</b>, <i>Hiroki Naganuma, Rio Yokota, [Information and Systems Society Poster Award (Acceptance Rate 6.3%)]</i> THE INSTITUTE OF ELECTRONICS INFORMATION AND COMMUNICATION ENGINEERS GENERAL CONFERENCE 2019 (IEICE2019)</p>	<p><i>Tokyo, Japan 03/19</i></p>
<p><b>”Smoothing of Objective Function for Large Scale Parallel Deep Learning”</b>, <i>Hiroki Naganuma, Rio Yokota</i>, THE 81ST NATIONAL CONVENTION OF INFORMATION PROCESSING SOCIETY OF JAPAN (IPSJ2019)</p>	<p><i>Fukuoka, Japan 03/19</i></p>
<p><b>”Smoothing of the Objective Function in Stochastic Optimization for Large Scale Parallel Learning”</b>, <i>Hiroki Naganuma, Rio Yokota</i>, CREST-DEEP SYMPOSIUM</p>	<p><i>Tokyo, Japan 11/18</i></p>
<p><b>”Hyperparameter Optimization of Large Scale Parallel Deep Learning using Natural Gradient Approximation Method”</b>, <i>Hiroki Naganuma, Shun Iwase, Kinsho Kaku, Hikaru Nakata, Rio Yokota</i>, FORUM FOR INFORMATION AND TECHNOLOGY 2018 (FIT2018)</p>	<p><i>Fukuoka, Japan 09/18</i></p>
<p><b>”Verification of speeding up using low precision arithmetic in convolutional neural network”</b>, <i>Hiroki Naganuma, Akira Sekiya, Kazuki Osawa, Hiroyuki Ootomo, Yuji Kuwamura, Rio Yokota</i>, GTC JAPAN 2017 POSTER SESSION</p>	<p><i>Tokyo, Japan 12/17</i></p>
<p><b>”Improvement of speed using low precision arithmetic in deep learning and performance evaluation of accelerator”</b>, <i>Hiroki Naganuma, Akira Sekiya, Kazuki Osawa, Hiroyuki Ootomo, Yuji Kuwamura, Rio Yokota</i>, TECHNICAL COMMITTEE ON PATTERN RECOGNITION AND MEDIA UNDERSTANDING (PRMU)</p>	<p><i>Kumamoto, Japan 10/17</i></p>
<p><b>”Accelerating Convolutional Neural Networks Using Low-Rank Tensor Decomposition”</b>, <i>Kazuki Osawa, Akira Sekiya, Hiroki Naganuma, Rio Yokota</i>, TECHNICAL COMMITTEE ON PATTERN RECOGNITION AND MEDIA UNDERSTANDING (PRMU)</p>	<p><i>Kumamoto, Japan 10/17</i></p>
<p><b>”Acceleration of Compression Model Using Half-Precision Arithmetic in Deep Learning”</b>, <i>Hiroki Naganuma, Kazuki Osawa, Akira Sekiya, Rio Yokota</i>, JAPAN SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATIC (JSIAM2017)</p>	<p><i>Tokyo, Japan 09/17</i></p>
<p><b>”Accelerating Convolutional Neural Networks Using Low-Rank Approximation”</b>, <i>Kazuki Osawa, Akira Sekiya, Hiroki Naganuma, Rio Yokota</i>, PROCEEDINGS OF THE CONFERENCE ON COMPUTATIONAL ENGINEERING AND SCIENCE VOL. 22</p>	<p><i>Saitama, Japan 05/17</i></p>
<p><b>”Acceleration of Matrix Multiplication of Deep Learning Using Low Rank Approximation”</b>, <i>Akira Sekiya, Kazuki Osawa, Hiroki Naganuma, Rio Yokota</i>, 158TH RESEARCH PRESENTATION SEMINAR IN HIGH PERFORMANCE COMPUTING</p>	<p><i>Shizuoka, Japan 05/17</i></p>

# Fellowship, Scholarship, and Grants-in-Aid

---

## **TSUBAME Encouragement Program for Young/Female Users (FY2024)**

COMPUTATIONAL RESOURCE SUPPORT (12,000 GPU HOURS)

07/24 - 03/25

- Topic: Curvature Regularization in Large Language Model Training for Downstream Task.

## **ABC Grand Challenge Program, National Institute of Advanced Industrial Science and Technology (AIST)**

COMPUTATIONAL RESOURCE SUPPORT (2,048GPUS X 24H)

09/23

- Topic: Large-Scale Evaluation of Training Strategies and Model Selection for Foundation Models: Impact on Out-of-Distribution Generalization and Uncertainty in Language and Image Tasks.

## **International Information Science Foundation**

TRAVEL GRANT

06/23

- Support Program for Researchers Traveling Abroad: Funding for expenses to participate in IJCNN2023

## **Exploratory Joint Research Project Support Program from JHPCN (EX23401)**

COMPUTATIONAL RESOURCE SUPPORT (12,000 GPU HOURS)

07/23 - 03/24

- Topic: Comprehensive Evaluation of the Next Generation Optimizers on Out-of-Distribution Generalization.

## **TSUBAME Encouragement Program for Young/Female Users (FY2023)**

COMPUTATIONAL RESOURCE SUPPORT (12,000 GPU HOURS)

07/23 - 03/24

- Topic: Comprehensive Evaluation of the Next Generation Optimizers on Out-of-Distribution Generalization.

## **Shigeta Education Foundation Scholarship**

LIVING EXPENSE

09/22 - 08/24

- One of the most prestigious scholarships for degree study abroad for Japanese student

## **Exploratory Joint Research Project Support Program from JHPCN (EX22401)**

COMPUTATIONAL RESOURCE SUPPORT (12,000 GPU HOURS)

07/22 - 03/23

- Topic: Development of Cost-Efficient and Fast Convergence Optimization Algorithms for Generative Adversarial Networks.

## **TSUBAME Encouragement Program for Young/Female Users (FY2022)**

COMPUTATIONAL RESOURCE SUPPORT (12,000 GPU HOURS)

07/22 - 03/23

- Topic: Development of Cost-Efficient and Fast Convergence Optimization Algorithms for Generative Adversarial Networks.

## **Graduate Scholarship for Degree Seeking Students, Japan Student Services Organization**

PARTIAL SUPPORT OF TUITION, LIVING EXPENSES

09/22 - 08/25

- The most prestigious research funding for Ph.D. students from Japan **[Acceptance Rate < 20%]**

## **ABC Grand Challenge Program, National Institute of Advanced Industrial Science and Technology (AIST)**

COMPUTATIONAL RESOURCE SUPPORT (2,048GPUS X 24H)

09/21

- Topic: Exhaustive Study of Optimizer Selection for Image and Language Model Training for OOD Generalization.

## **Grand-challenge Program on TSUBAME3.0 Super Computer**

COMPUTATIONAL RESOURCE SUPPORT (512GPUS X 168H)

05/21

- Topic: An Exhaustive Study of Optimizer Characteristics for OOD Generalization on the ImageNet Scale

## **Université de Montréal Tuition Fee Reduction Scholarship for International Student**

84,000 CAD / APPROX 68,300 USD

09/20 - 08/25

- Tuition Fee Reduction for the international doctoral student at Université de Montréal

## **Research Fellowship for Young Scientists (Declined)**

LIVING EXPENSE

04/20 - 03/22

- The most prestigious fellowship for Ph.D. students in Japan **[Acceptance Rate < 20%]**

## **JSPS KAKENHI Grant Number 20J13997 (Declined)**

RESEARCH FUNDING

04/20 - 03/22

- The most prestigious research funding for Ph.D. students in Japan **[Acceptance Rate < 20%]**

## **NeurIPS 2019 Travel Support by The Japanese Society for Artificial Intelligence**

FULL SUPPORT OF TRAVEL EXPENSE AND REGISTRATION FEES

12/19

- Selected as an excellent student participates in machine learning research

## Tokyo Tech Tsubame Scholarship

PARTIAL SUPPORT OF TUITION, LIVING EXPENSES

04/19 - 03/22

- As a doctoral student at Tokyo Institute of Technology, my master's degree graduation with Summa Cum Laude and a research plan were highly evaluated and selected.

## Masason Foundation Scholarship supported by Masayoshi Son, CEO of SoftBank Group

FULL SUPPORT OF TUITION, LIVING EXPENSES, AND RESEARCH FUNDING

07/17 - 08/22

- Selected as one of the most talented 96 young scholars under 25 years old among the more than 1,100 applicants who have won the International Contest Winning Experience and the Math Olympics National Representative and so on from over the world **[Acceptance Rate 8.7%]**

## Japan Student Services Organization, Master's Scholarship

PARTIAL SUPPORT OF TUITION, LIVING EXPENSES

04/17 - 03/19

- Repayment Exemption for Students with Excellent Grades **[Acceptance Rate 6.8%]**

## Japan MEXT (Ministry of Education, Culture, Sports, Science and Technology) Scholarship

FULL SUPPORT OF TRAVEL EXPENSE AND REGISTRATION FEES

02/19

- Selected as an excellent Japanese master's student; International academic societies etc. participation support

## Shinano-Ikueikai Scholarship

300,000 JPY / APPROX 2,800 USD

09/17 - 03/18

- Highly evaluated the development of support devices using visual recognition for visually impaired people and supported development costs by the foundation who admit volunteer achievement

## Incentive Award of Director of School of Computing, Tokyo Institute of Technology

300,000 JPY / APPROX 2,800 USD

03/17

- Awarded by excellent achievement and grade of the undergraduate course

## SC16 HPC for Undergraduates Program

FULL SUPPORT OF TRAVEL EXPENSE AND REGISTRATION FEES

11/16

- Full support of travel expense and registration fees for attending Supercomputing Conference **[First selection for Japanese student]**

## Meiji University Special Scholarship for Top Score Students (Declined)

FULL SUPPORT OF TUITION FEES

03/13

- Full Support of Tuition FY2013-FY2017 **[Acceptance Rate < 0.06% =(70/109,934)]**

## Extracurricular Activity

---

### Montreal Japanese Academy Society

Montreal, Canada

COMMITTEE MEMBER

02/23 -

- Organized academic research exchange events to facilitate networking and collaboration among scholars and students.
- Provided support for academics and their families by coordinating social events and offering resources tailored to their needs.

### The Japanese Graduate Student Association in the United States (JGSAU)

Tokyo, Japan

DIVISIONAL DIRECTOR

09/22 - 08/25

- We organized a series of seminars on graduate schools abroad. In particular, I was responsible for a session on "Families and Partnerships."

### NEDO Technology Commercialization Program / FY2022

Tokyo, Japan

SELECTED ATTENDEE

09/22 - 03/23

- My proposal has been selected for a program of commercialization of research and development, and I am participating in a training program in which a Japanese government-contracted organization provides mentoring and support for commercialization.

### The Machine Learning Summer School 2020

Tübingen, Germany

SELECTED ATTENDEE

07/20

- I attended MLSS 2020 in Tübingen, Germany. The summer school was very mainstream and had a lot of top instructors from all over the world. There were 180 participants selected out of more than 1300 applications, with **an acceptance rate of less than 14%**.

### Top Conference Reports: NeurIPS 2019 (Supported by JSAI)

Tokyo and Osaka, Japan

SELECTED REPORTER AND PRESENTER

03/20

- At the 78th Artificial Intelligence Seminar AI Trends and Top Conferences debriefing session (NeurIPS2019), I published a slide titled "Trends in Deep Learning Theory at NeurIPS 2019." The slide summarized the key deep learning theory trends that were observed at NeurIPS 2019.

### Leadership Training Programs powered by JT

Nagano, Japan

SELECTED ATTENDEE

09/19

- I was selected out of more than 1,000 times the number of applicants to participate in a workshop to learn from prominent astronauts, management consultants, and brain scientists in Japan about leadership and their lives.



## International HPC Summer School 2019

Kobe, Japan

SELECTED ATTENDEE

01/19

- The international summer school allows talented young researchers to broaden their perspectives and acquire new knowledge through lectures and workshops offered by the leading scientists in the HPC field. I could network with other participants from many parts of the world, which was a great opportunity to learn about different research areas and build collaborations. I learned a wide range of fields related to cutting-edge HPC and networked at this school.

## Georgia Institute of Technology Leadership Program

Atlanta, US

SELECTED ATTENDEE

01/19

- I attended a month-long leadership training program at Georgia Institute of Technology. The program included lectures and workshops, and I learned about leadership systematically. This was an excellent opportunity to develop my leadership skills and learn more about what it takes to be successful. The instructors were knowledgeable and provided great feedback to help me grow as a leader.

## KOBE HPC Spring School

Kobe, Japan

SELECTED ATTENDEE

03/18

- This workshop is a program in which leading-edge Japanese HPC researchers widely teach HPC techniques to students in related fields. I learned at this school to incorporate the efficient communication overlap technique, which is actively researched in the area of HPC regarding communication in distributed deep learning.

## Gpu Technology Conference Japan, Deep Learning Institute Workshop

Tokyo, Japan

TEACHING ASSISTANT

11/18

- In this workshop, I taught participants how to use NVIDIA DIGITS and build CNN and RNN. I also introduced basic Python/Jupyter techniques for beginners.

## Organization for Inviting EXPO2025 to OSAKA

Osaka, Japan

STUDENT REPRESENTATIVE OF COMPUTER SCIENCE MAJOR IN JAPAN

11/17 - 12/19

- The BIE team visited Osaka in March 2018, and I gave them a presentation on medical data analysis software and health management applications that I developed using wearable devices as an engineer and researcher. We succeeded to attract the Osaka Expo in 2025 with these innovative applications.

## TEDxUTokyo 2017

Tokyo, Japan

VISUAL DIRECTOR

05/17

- I helped present the opening and finale at the TEDxUTokyo 2017 conference using computer vision and computer graphic techniques. I developed a real-time analysis system of large-scale data and performed interactive content linked with Facebook.

## SC16 HPC for Undergraduates Program

Salt Lake City, US

SELECTED ATTENDEE

11/16

- I am honored to have been selected as the first Japanese student for the undergraduate program at SC16, the top conference of High-Performance Computing (HPC). The committee offered me travel, accommodation, and conference participation expenses. I attended this important event and learned from some of the best minds in HPC.

## Musashino Art University Festival Projection Mapping Project 2015

Tokyo, Japan

LEADER, CHIEF ENGINEER

09/16

- Our team presented the opening and finale at the Musashino Art University Art Festival 2015 by performing real-time large-scale data analysis and developing interactive content linked with Twitter. We displayed a large monitor showing real-time tweets and an interactive map showing how people were tweeting about our performance. Our goal was to create a unique experience for festival attendees that would allow them to connect with each other and the performers in new ways.

## Tokyo Institute of Technology Advertisement and Media Art Society (TITAMAS)

Tokyo, Japan

VICE PRESIDENT

08/16 - 03/19

- We were selected as the Japan representative for Microsoft Imagine Cup 2017. This is the world's largest IT development contest, and we are proud to be a part of it. We have been actively participating in programming contests and so on., domestically and abroad, and achieved many results.

## Supercomputing Contest 2016 Held in Tokyo Institute of Technology and U Osaka

Osaka, Japan

TEACHING ASSISTANT

08/16

- I had the opportunity to teach high school students how to use the Linux system and how to run scripts on Supercomputer TSUBAME 2.5 at Tokyo Institute of Technology. I also taught basic C++/CUDA techniques for the Supercomputing Contest.

## Silicon Valley Workshop 2016 Supported by Recruit Holdings

Tokyo, Japan

SELECTED ATTENDEE, 1ST GENERATION

03/16

- I was selected as an outstanding student engineer in Japan and participated in the Silicon Valley Workshop with the support of Recruit Holdings. I stayed in San Francisco and Silicon Valley for 10 days, took some lectures at Stanford University, and toured IT-related companies around Palo Alto. Learning about cutting-edge technology and meeting people from various backgrounds was a great opportunity.