

KAZUKI MINEMURA

Email: kazuki.minemura@gmail.com, Mobile: +81-8023704256

PROFESSIONAL SUMMARY

Experienced software optimization engineer with a broad technical background spanning high-performance computing (HPC), AI inference, computer vision, and OS kernel development. Spearheaded end-to-end initiatives, including debugging and performance tuning of compilers and AI libraries, designing computer vision systems, automating DevOps workflows, and contributing to open-source projects. Delivered technical training and customer support, while actively engaging in internal and external technical communication. Committed to team development through mentoring and collaborative leadership.

TECHNICAL SKILLS

Programming Languages:

Python, C/C++, SYCL, CUDA, FORTRAN, MATLAB, TypeScript

AI Frameworks:

OpenVINO, Pytorch, Tensorflow, vllm, oneDNN

Development Tools:

Jira, HSDDES, Jenkins, Ansible, Docker, Git, Vim, GDB, Intel VTune, Intel Analyzer

Web Frameworks:

Angular, Tornado

Data Modeling/Management:

JSON, XML MongoDB

Typesetting:

TeX, LaTeX

Operating Systems:

Ubuntu, Rokey, Windows, Yocto BSP

PROFESSIONAL EXPERIENCE

Software Technical Consultant Engineer

Intel Kabushiki Kaisha, Tokyo, Japan

July 2022 – Present

Responsibilities:

- Delivered optimization support for toolchains across diverse high-performance computing (HPC) environments.
- Worked across a wide technical scope, including compilers, runtimes, profilers, AI libraries, and OS kernel-level components.

KAZUKI MINEMURA

Email: kazuki.minemura@gmail.com, Mobile: +81-8023704256

- Led end-to-end efforts from root cause analysis and performance tuning to the development of training content and hands-on technical enablement for internal and external stakeholders.
- Contributed to cross-functional collaboration and technical knowledge sharing within the engineering community.

Key Achievements:

- Compiler Optimization & Bug Resolution: Diagnosed issues in specific compiler optimization phases; proposed and implemented patches and long-term workarounds, improving developer productivity and tool reliability.
- Deep Learning Framework Enhancement (oneDNN): Identified and fixed critical source-level bugs affecting inference accuracy and speed in the oneDNN library.
- Performance Tuning with Intel VTune: Conducted hotspot analysis and multithreading optimization for customer applications; led performance tuning workshops and internal training sessions.
- SYCL Enablement and Benchmarking: Led the migration of legacy codebases to SYCL, optimizing GPU workloads and improving performance benchmarks on Intel platforms.
- Technical Training Development: Created and delivered in-depth training programs on compiler internals, oneDNN, VTune, and SYCL for engineers and partners. Additionally, successfully introduced and conducted VTune training sessions for 14 client companies, consistently receiving positive feedback and high satisfaction ratings.
- vLLM Port to Intel XPU: Successfully ported CUDA-dependent modules in the vLLM inference engine to SYCL, enabling execution on Intel GPUs (XPU) and contributing enhancements to the open-source ecosystem.
- Linux Kernel Debugging for Woven Planet: Investigated and patched performance regressions and kernel panics on specific CPU models, stabilizing critical workloads and contributing to infrastructure reliability.

Success Factors:

- Full-Stack Technical Expertise: Demonstrated deep understanding across the software stack—from OS kernel-level development to AI libraries and GPU performance optimization.
- End-to-End Support Capabilities: Delivered comprehensive solutions covering debugging, performance tuning, and long-term maintainability.
- Effective Knowledge Sharing: Skilled in designing and delivering technical training and documentation to empower internal teams and external partners.
- Open-Source Impact: Actively contributed to the open-source community, including successful porting of vLLM to Intel XPU, accelerating ecosystem adoption.

Computer Vision Engineer

Intel Microelectronics (M), Penang, Malaysia

Jan 2019 – Jun 2022

Responsibilities:

- Designed and deployed a POC of multi-view image grading system for standardized science exams administered by China's Ministry of Education.
- Developed computer vision and deep learning-based inference frameworks, validation pipelines, and DevOps environments.
- Contributed to technical education and talent recruitment through internal training and external engagements.

Key Achievements:

KAZUKI MINEMURA

Email: kazuki.minemura@gmail.com, Mobile: +81-8023704256

- AI-Powered Grading System: Built an image classification and object detection pipeline using OpenVINO to support multi-angle exam paper analysis.
- Robust QA Framework: Developed a scalable validation system encompassing over 6,000 test cases to ensure model accuracy and reliability across hardware platforms.
- DevOps Automation: Implemented a CI/CD infrastructure using Ansible, Docker, and Jenkins to streamline deployment and reduce testing cycle times.
- Team Leadership & Mentorship: Led cross-functional teams in model validation and benchmarking; mentored junior engineers and promoted engineering best practices.
- Client Technical Enablement: Provided hands-on support and consulting to client-side engineers, resulting in improved customer satisfaction and system adoption.

Success Factors:

- Delivered end-to-end AI solutions, from model development and optimization to deployment and support.
- Built cross-hardware validation systems enabling scalable testing across edge and cloud environments.
- Fostered a collaborative and knowledge-sharing culture as a technical leader and mentor.

Software Validation Engineer **Jan 2018 - Dec 2018**

Responsibilities:

- Designed and implemented an automation framework that significantly accelerated AI software development workflows.
- Developed object detection models leveraging LiDAR data for robust performance in real-world scenarios.
- Promoted the company's technical branding by presenting research results at international conferences and public events.

Key Achievements:

- 5x Efficiency Improvement: Increased testing efficiency fivefold by introducing scalable automation scripts, reducing manual effort and regression cycle times.
- LiDAR-Based AI Modeling: Designed and optimized object detection algorithms based on LiDAR signals, enhancing environmental perception accuracy.
- Global Research Dissemination: Presented research findings at international conferences and contributed to the company's visibility within the AI research community.
- High-Impact Public Engagement: Delivered a keynote speech at the IPIARTI Symposium, strengthening the company's reputation in the field of AI and robotics.

Success Factors:

- Proven ability to architect automation infrastructure that balances speed, scalability, and quality.
- Deep technical expertise in LiDAR-based perception systems tailored to real-world deployment.
- Strong public speaking and communication skills for effectively sharing technical insights with external audiences.

Graduate Trainee **July 2016 – Dec 2017**

KAZUKI MINEMURA

Email: kazuki.minemura@gmail.com, Mobile: +81-8023704256

Responsibilities:

- Supported the development of LiDAR-based object detection models using computer vision and deep learning techniques.
- Conducted research on deep learning trends and contributed to technical communication across teams.
- Assisted in creating demos and materials for internal and external events.

Key Achievements:

- Model Development Support: Contributed to the design and experimentation of LiDAR object detection models, gaining hands-on experience with CV/DL pipelines.
- Team Collaboration: Facilitated knowledge sharing through weekly research meetings, improving alignment among team members.
- Technical Demonstration: Developed demos for internal events, effectively showcasing ongoing technical efforts.
- Open Source Contribution: Translated Autoware documentation into English, enhancing accessibility and encouraging broader OSS adoption.

Success Factors:

- Demonstrated the ability to quickly grasp and apply emerging technologies in real-world projects.
- Strong communication and collaboration skills that helped strengthen team synergy.
- Proven capacity to contribute effectively as a beginner-level engineer with a proactive mindset.

EDUCATION

Ph.D. in Computer Science
2017

Feb 2013 – Jan

University of Malaya, Kuala Lumpur, Malaysia

- Conducted research on "Sketch - An Investigation into Feature Extraction in Compressed Domain"
- Published 2 ISI-indexed journal articles, 1 book chapter, and 8 peer-reviewed conference papers

M.S. in Electrical and Electronics Engineering
2012 Shinshu University, Nagano, Japan

Apr 2010 - Mar

B.S. in Electrical and Electronics Engineering
2010 Shinshu University, Nagano, Japan

Apr 2006 - Mar

PUBLICATIONS (Last 6 years)

ISI Indexed Journal

- J1. Raphaël C.-W. Phan, Yin-Yin Low, KokSheik Wong, **Kazuki Minemura**, "Strengthening speech content authentication against tampering". Speech Communication. Vol 6. 2021, (IF 2.017)
- J2. **Kazuki Minemura**, KokSheik Wong, C.-W Phan, Kiyoshi Tanaka, "A novel sketch attack for H.264/AVC format-compliant encrypted video". IEEE Transactions on Circuits and Systems

KAZUKI MINEMURA

Email: kazuki.minemura@gmail.com, Mobile: +81-8023704256
for Video Technology. Jul. 2016, (IF 9.9)

- J3. **Kazuki Minemura**, KokSheik Wong, Xiaojun Qi and Kiyoshi Tanaka, "A Scrambling Framework for Block Transform Compressed Image," Multimedia Tools and Application, Feb. 2016, (IF 2.313)

Peer Reviewed Conference Paper

- C1. **Kazuki Minemura**, Hengfui Liao, Abraham Monrroy and Shinpei Kato, "LMNet: Real-time Multiclass Object Detection on CPU using 3D LiDAR", IEEE Conference on Intelligent Robot Systems (ACIRS), pp. 28-34, 2018.
- C1. Yiqi Tew, **Kazuki Minemura** and KokSheik Wong, "HEVC selective encryption using transform skip signal and sign bin", Asia-Pacific Signal and Information Processing Association (APSIPA), pp. 963-970, 2015.
- C2. Masaya Moriyama, **Kazuki Minemura** and KokSheik Wong, "Moving Object Detection in HEVC Video by Frame Sub-sampling," IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), pp. 48-52, 2015.