Masahiro Hiramori

Greater Tokyo Area (Kanagawa), Japan contact@mshr-h.com mshr-h.com mshr-h
masahiro-hiramori

Summary

AI Compiler / Edge Inference engineer with **8+ years** of experience spanning **AI compilers, runtime systems, and FPGA/ASIC-based acceleration** for industrial and manufacturing workloads. Currently Head Engineer in **Mitsubishi Electric's Open-Source & InnerSource Program Office (OSPO)**, leading company-wide technical direction for open-source adoption across AI compilers, LLM deployment, and edge inference.

PyTorch Ambassador and **Apache TVM Committer**, with a proven record of bridging **PyTorch ↔ TVM** via modern export/import workflows and maintaining production-grade CI for compiler toolchains. Hands-on experience deploying **vision models and domain-specific LLMs** on edge devices including **NVIDIA Jetson Orin Nano, Rockchip RK3588, Hailo-8, and Renesas RZ/V2H**, and building runtime stacks for FPGA-based accelerators.

Experience

PyTorch Community, PyTorch Ambassador — Oct 2025 – present, 7 months

- Serve as a PyTorch Ambassador representing industrial and edge AI compiler use cases.
- Promote best practices for PyTorch deployment, compiler integration, and community contributions through talks, articles, and mentoring.

Apache Software Foundation, Apache TVM Committer — Sept 2024 – present, 1 year 8 months

- Committer and core contributor to Apache TVM, focusing on PyTorch frontend integration, Relax, and MetaSchedule for inference workloads.
- Maintain CI pipelines and review contributions related to PyTorch integration and model import/export workflows.

Mitsubishi Electric Corporation, Head Engineer, Open-Source & InnerSource Program Office — Japan, Apr 2025 – present, 1 year 1 month

- Lead company-wide strategy for adopting open-source AI frameworks and tools, defining technical direction and governance for AI compilers, LLMs, and edge inference.
- Drive deployment of a manufacturing domain-specific LLM to edge devices (NVIDIA Jetson Orin Nano, Rockchip RK3588), coordinating cross-functional teams to validate feasibility and align with production requirements.
- Represent the company in external OSS communities (PyTorch, Apache TVM), aligning internal R&D roadmap with upstream projects and contributing back improvements.
- Provide technical oversight and mentoring on AI/ML infrastructure, edge deployment, and OSS contribution workflows.

Mitsubishi Electric Corporation, Researcher, Information Technology R&D Center — Japan
Apr 2017 – Mar 2025
8 years

- Developed a C++ user-mode driver and runtime stack for a proprietary FPGA-based neural network accelerator, achieving 100+ fps object detection on Xilinx UltraScale+ MPSoC in real-time scenarios.
- Implemented image anomaly detection pipelines on edge AI chips such as Hailo-8 and Renesas RZ/V2H for visual inspection in manufacturing use cases, optimizing models and runtime for real-time performance on resource-constrained devices.
- Implemented deployment of a manufacturing domain-specific LLM on edge devices (NVIDIA Jetson Orin Nano, Rockchip RK3588), including model optimization and runtime integration for on-device inference.
- Prototyped end-to-end software systems for edge AI, including model conversion, compilation, and deployment pipelines using TVM and related tooling.
- Contributed to open-source projects such as Apache TVM, Hummingbird, and ONNX-MLIR, focusing on PyTorch frontend development, CI/CD infrastructure, and integration with industrial edge workloads.

Education

M.S. **Osaka Institute of Technology**, Information Science and Technology — Japan
Apr 2015 – Mar 2017

B.A. **Osaka Institute of Technology**, Information Science and Technology — Japan
Apr 2011 – Mar 2015

Skills

Languages: C, C++, Python

AI Frameworks & Compilers: PyTorch, Apache TVM (Relax / MetaSchedule), ONNX / ONNX-MLIR, Model optimization and compilation pipelines for edge devices, Integration of torch.export.ExportedProgram with Apache TVM

Hardware & Edge Inference: AI accelerators: Hailo-8, Edge SoCs: NVIDIA Jetson Orin Series, Rockchip RK3588, Renesas RZ/V2H, Performance tuning and deployment of object detection, anomaly detection, and domain-specific LLMs on edge devices

DevOps & Tooling: CI/CD (e.g., GitHub Actions), test automation, Reproducible ML workflows, containerization, continuous integration for compiler/tooling codebases

Open Source & Community: Large-scale OSS contribution workflows, code review, and maintainer responsibilities, Program committee and community leadership (PyTorch Conference, PyTorch Community Awards)

Languages

Japanese: Native

English: Full professional proficiency

Open Source Contributions

Apache TVM

- Introduced and implemented the ExportedProgram frontend to enable integration of torch.export.ExportedProgram with TVM.
- Actively develop the Relax PyTorch frontend and maintain CI pipelines for PyTorch-related workflows.
- Review and maintain contributions related to Relax frontend, CI pipelines, and model import/export.

PyTorch

- Implemented and maintained integration paths between PyTorch and Apache TVM to offload models for optimized inference on edge devices.

Microsoft Hummingbird
- Integrated Apache TVM as a backend for classical ML model conversion.
- Extended CI pipelines to validate conversion and execution of models targeting TVM-based runtimes.

Community Service

Program Committee Member, PyTorch Conference 2025
- Compilers & Kernels track. Reviewed proposals focused on compiler integration, deployment, and performance optimization topics.

Projects

vscode-verilog-hdl-support
- Maintainer of one of the most widely used Verilog/SystemVerilog HDL extensions for Visual Studio Code. Provides syntax highlighting, basic language support, and integration with typical hardware design workflows used in FPGA/ASIC development.

Publications

OSS コミッターが海外カンファレンスに飛び込んだ話 Feb 2026
Masahiro Hiramori
www.docswell.com/s/mshr-h/KVMQWX-lf-ai-data-japan-rug-2026-02-20 (LF AI & Data Japan Regional User Group (Japan RUG) 第 3 回 Meetup)

Bridging PyTorch and TVM: Integrating torch.export.ExportedProgram With Apache TVM 2025
Masahiro Hiramori
pytorchconference.sched.com/event/28nTQ/poster-presentations-compilers-kernels (PyTorch Conference 2025 (Poster))

製造業ドメイン特化の言語モデル Sept 2025
Masahiro Hiramori
www.mitsubishielectric.co.jp/giho/2020s/202509 (三菱電機技報, 2025 年 9 月号)

三菱電機における **AI** のコンパクト化に向けた **OSS** 活用と **OSS** コミュニティ活動 Feb 2025
Masahiro Hiramori
www.docswell.com/s/mshr-h/59VLGY-lf-ai-data-japan-rug-kickoff (LF AI & Data Japan Regional User Group (Japan RUG) Kick Off)

Performance Evaluation of Apache TVM's MetaSchedule for Image Classification Models on Edge Device 2024
Masahiro Hiramori
ieeexplore.ieee.org/document/10760375 (IEEE 13th Global Conference on Consumer Electronics (GCCE))

Caffe-SSD Inference on Edge Device Using TVM and Hybrid Script 2021
Masahiro Hiramori
tvmconf.org/2021/index.html%3Fp=1168.html (TVMCon 2021)

A Study on Fast Motion Estimation Algorithm 2016
Masahiro Hiramori
(IEEE 5th Global Conference on Consumer Electronics)

A Study on Motion Estimation Algorithm for Moving Pictures 2016
Masahiro Hiramori
(IEEE 5th Global Conference on Consumer Electronics)

Patents

Data processing execution device, data processing execution method and data processing execution program: JP6815563B1 (Tsutomu MOTOHAMA, Masahiro DEGUCHI, Masahiro HIRAMORI, Hidetoshi OKI) Link

Container migration device, container migration method, and computer readable medium: JP7527523B2 (Masahiro HIRAMORI, Takehisa MIZUGUCHI) Link

Media Coverage

三菱電機が設立した **OSPO** のメンバーにインタビュー。「インナーソース」を戦略的に使う背景とは？ Oct 2025

- Think IT
- Japanese; interview on OSPO and InnerSource strategy

OSS 活用推進に向け三菱電機が **OSPO** 設立、インナーソースで風土を醸成 July 2025

- 日経クロステック／日経コンピュータ
- Japanese; feature on OSPO establishment and OSS utilization

Awards

PyTorch Community Awards 2025 — Nominee

- Awarded by PyTorch

PyTorch Community Awards 2024 — PyTorch Innovator Nominee

- Awarded by PyTorch