**Formula 02**

|  |  |
| --- | --- |
| VIETNAM NATIONAL UNIVERSITY, HANOI  **INFORMATION TECHNOLOGY INSTITUTE** | **VIETNAM SOCIALIST REPUBLIC**  **Independence – Freedom – Happiness** |

**RECRUITMENT INTERN  
Embedded AI engineer**

**Title: Optimizing Machine Learning for Edge Devices with RISC-V Architecture and Apache TVM.**

* **Job Description**

Traditional AI models are often too resource-intensive for edge devices, resulting in low performance and high power consumption. This project focuses on optimizing AI solutions to run efficiently on the RISC-V architecture, a flexible and open-source platform. The goal is to make AI models lightweight and efficient enough to be integrated directly into compact, power-efficient hardware.

The core task is to implement and optimize AI models using techniques like quantization and pruning. Crucially, this project will leverage the Apache TVM deep learning compiler, a powerful tool that automatically generates optimized code. It takes full advantage of RISC-V’s unique features, such as its parallel processing capabilities, thereby boosting performance and significantly reducing power consumption.

The project’s outcome will be a complete workflow, from training models to deploying them on RISC-V hardware. By using TVM, the project will deliver Ultra-Low-Power, high-performance AI solutions, enabling smart applications like image recognition or anomaly detection to run effectively on edge devices. This opens up new possibilities for IoT and automation where energy efficiency is critical.

**Objective:**

The primary objective of this research is to optimize Machine Learning based on Apache TVM on RISC-V architectures. This project will leverage RISC-V's flexibility and efficiency to create a complete image processing system for UAV.

1. Model Optimization & TVM Integration: Optimize the AI model with techniques like quantization and pruning, then integrate it with the Apache TVM deep learning compiler.
2. RISC-V Hardware Deployment: optimized code that fully leverages RISC-V's parallel processing capabilities, building a complete and efficient image processing system.
3. Performance & Power Evaluation: Evaluate the frame processing rate and minimize power consumption to ensure the system is suitable for practical, Ultra-Low-Power applications.

**Keywords**: FPGA, RISC-V Architecture, Parallel Processing, Ultra Low power, Apache TVM, Edge AI, High-Level Synthesis, Hardware Acceleration.

* **Project Supervision**

The project is supervised by Prof. Tran Xuan Tu and Dr. Bui Duy Hieu. Prof. Xuan-Tu Tran and Dr. Duy-Hieu Bui have been working on hardware design and accelerations for FPGA since 2010. They have my publications and patents specifically on hardware design for image processing and security. This work is of interest in designing HW accelerators for edge devices.

* **Candidate Profile**

Education: GPA >= 3.21/4.0.

**Research Experience**:

* Having experience with FPGA, embedded system design & implementation is a plus.

**Technical Skills:**

* Programming: Python, C++/C++ and/or basic Verilog.
* Basic skills in neuron network training and inference.
* Experience with ARM programming, electronic circuit design is a plus.

**Language:**

* English: Proficient in communication and reading.
* **Bibliography**

1. Y. -R. Chen et al., "Experiments and optimizations for TVM on RISC-V Architectures with P Extension," 2020 International Symposium on VLSI Design, Automation and Test (VLSI-DAT), Hsinchu, Taiwan, 2020, pp. 1-4, doi: 10.1109/VLSI-DAT49148.2020.9196477.
2. Jenq-Kuen Lee1 et al., “Supporting TVM on RISC-V Architectures with SIMD Computations,” TVM Conference, 2019. URL: https://tvmconf.org/2019/slides/2019/E08-Jenq-Kuen-Lee.pdf
3. Jenq-Kuen Lee et al., “Supporting TVM on RISC-V Architectures,” TVM Confernce, 2018. URL: <https://tvmconf.org/2018/slides/07-Jenq-Kuen-Lee.pdf>
4. Frank K. Gürkaynak, “PULP and AI Acceleration,” Taichip Winter School – Frankfurt a.d. Oder – February 2025. URL: <https://pulp-platform.org/docs/taichip2025/kgf_taichip_winterschool25.pdf>

* **Internship location:** Building E3, VNU Information Technology Institute, 144 Xuan Thuy, Cau Giay, Hanoi.
* **How to apply:** Interested candidates, please email your CV and academic record to hieubd [at] vnu.edu.vn. Email subject: [INTERN] - [Position name] - [Fullname].