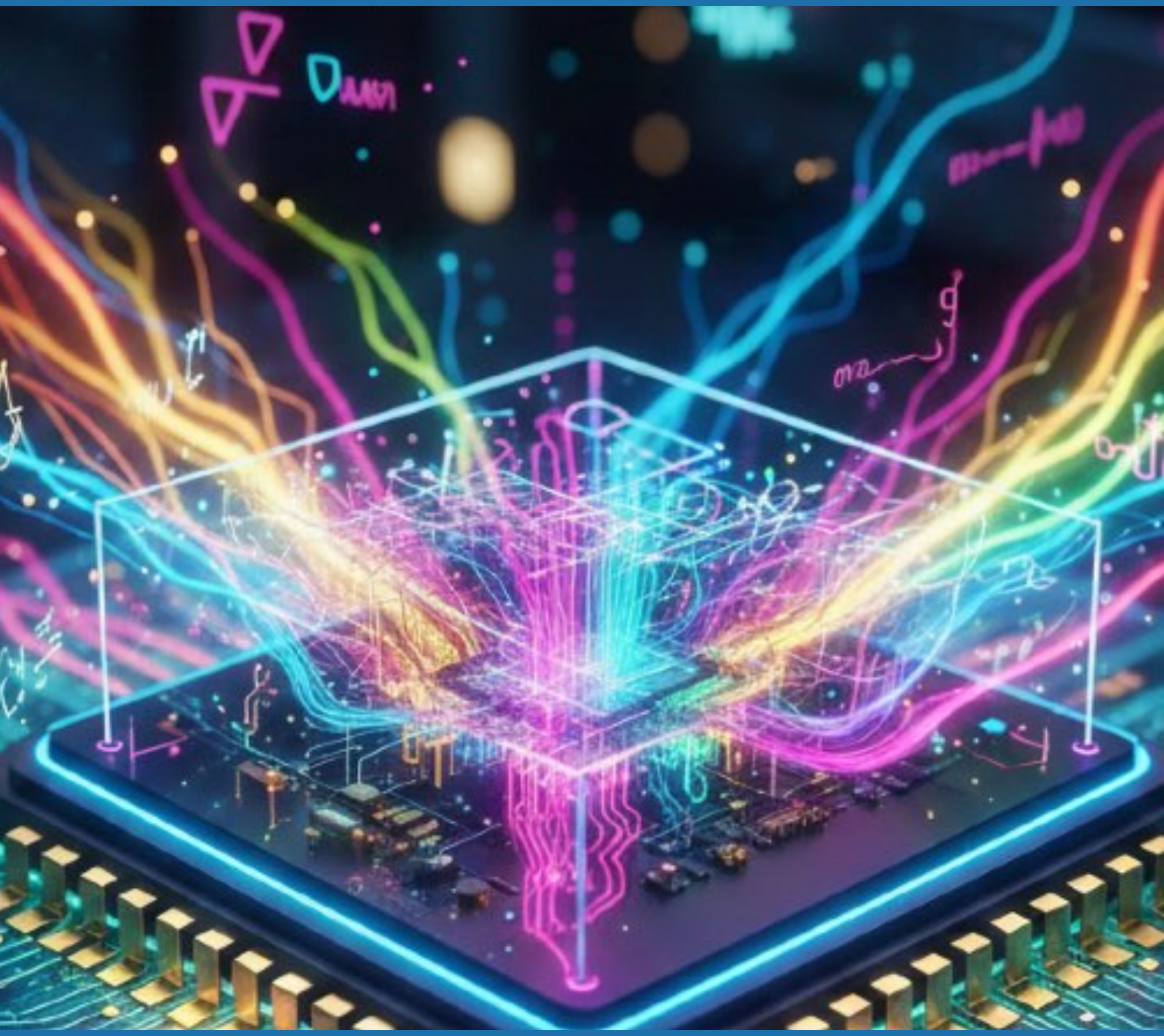
KLEPSYDRA

TECHNOLOGIES

THE Most Efficient Software for Edge Computing



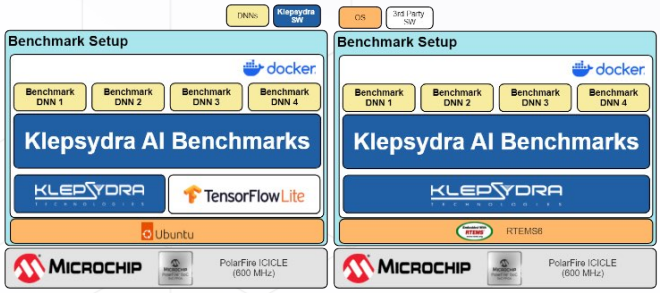
Unleash the Full Power of Your Edge Processor

Klepsydra AI Performance Benchmark Report

Version: 18 **Target**: PolarFire RISC-V. **Date**: Dec. 2024

Technical setup

The benchmark application runs in a Docker container on Ubuntu 22.04 on a PolarFire Icicle RISC-V, with the image including Klepsydra AI and TensorFlow Lite 2.4.4 (compiled with NEON extensions). For RTEMS*, the AI benchmarks are deployed directly on the target binary.



Benchmark Deep Neural Networks

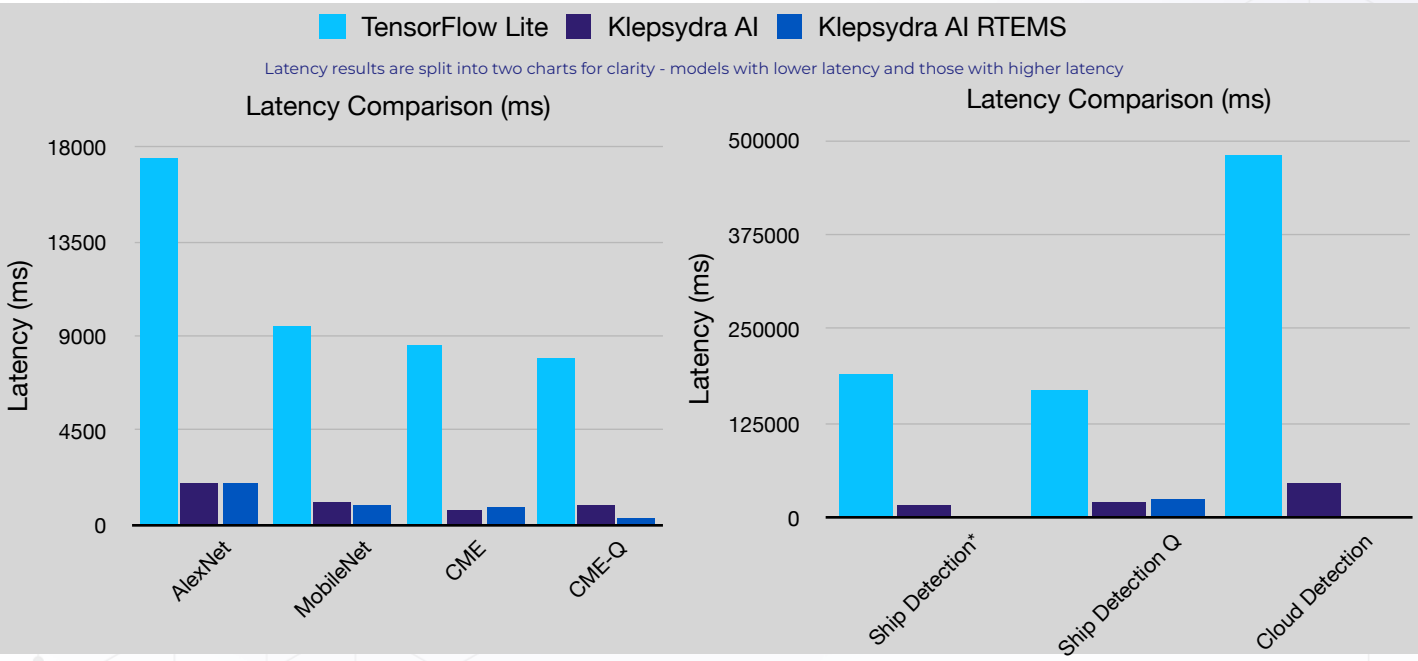
Several networks were tested as part of this campaign. These networks come from different sources as specified:

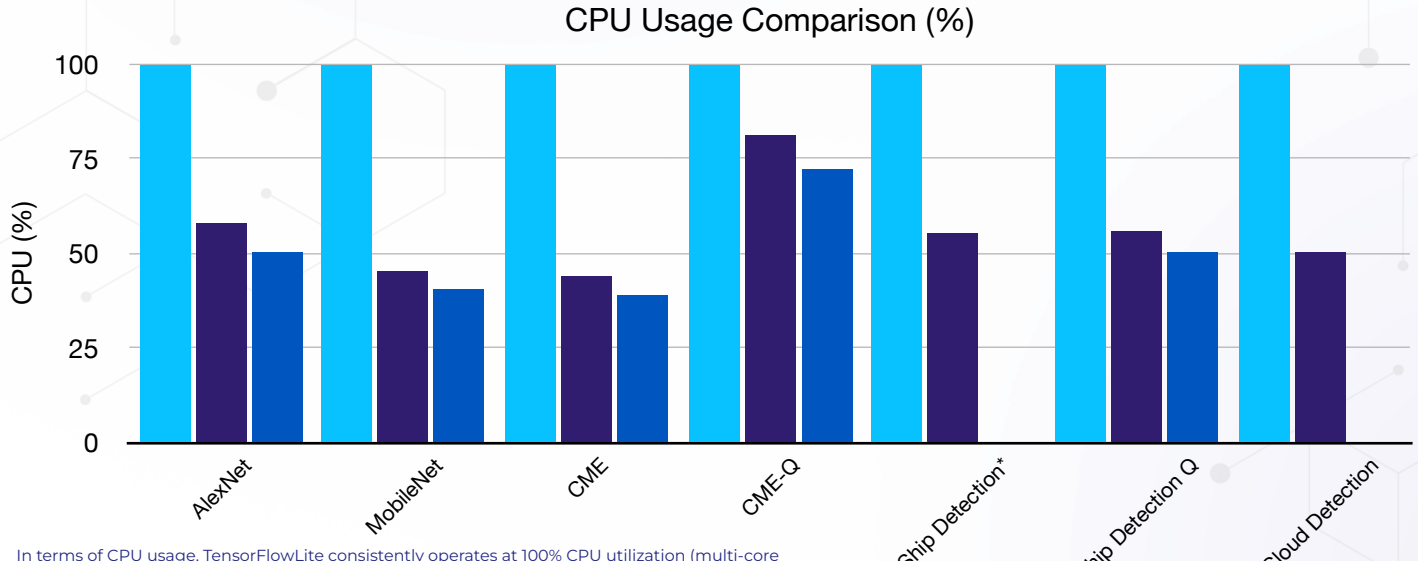
- AlexNet (open source)
- ESA Coronal Mass Ejection Detection (CME [1])
- ESA CME Quantised (ESA OBPMark-ML[1])
- YoloX / Ship Detection* (ESA OBPMark-ML[1])
- YoloX / Ship Detection Quantised (ESA OBPMark-ML[1])
- Cloud Detection (ESA OBPMark-ML[1])

[1]: https://zenodo.org/records/5638577
[2]: https://arxiv.org/abs/2309.11645
* Benchmarks for RTEMS are partially complete, with a full benchmark report expected in a new document release in Q2 2025.

The performance results are shown for latency, i.e., the time required to execute the AI algorithm for a given input data, and CPU consumption, i.e., how much CPU is used for executing the AI algorithm.

The results show that Klepsydra AI outperforms TensorFlowLite in terms of latency and CPU consumption for the RISC-V architecture.





In terms of CPU usage, TensorFlowLite consistently operates at 100% CPU utilization (multi-core optimized for latency), while Klepsydra requires significantly less CPU during inference.

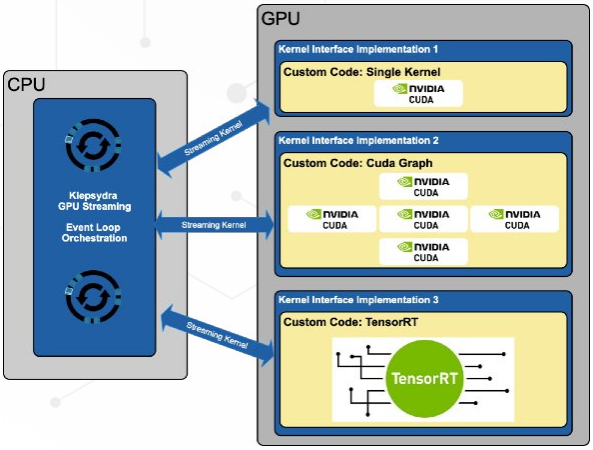
In terms of CPU usage, TensorFlowLite consistently operates at 100% CPU utilization (multi-core optimized for latency), while Klepsydra requires significantly less CPU during inference.

Klepsydra AI Performance Benchmarks Report

Version: 23 **Target:** Jetson ORIN NX **Date:** Feb. 2026

Technical setup

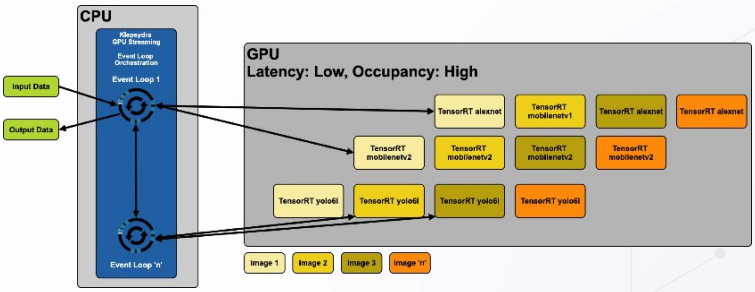
The Klepsydra GPU Connector maximizes NVIDIA TensorRT performance by enabling true parallel execution of multiple AI models on a single application. By leveraging our proprietary CUDA stream management, Klepsydra decouples hardware resources from software instances.

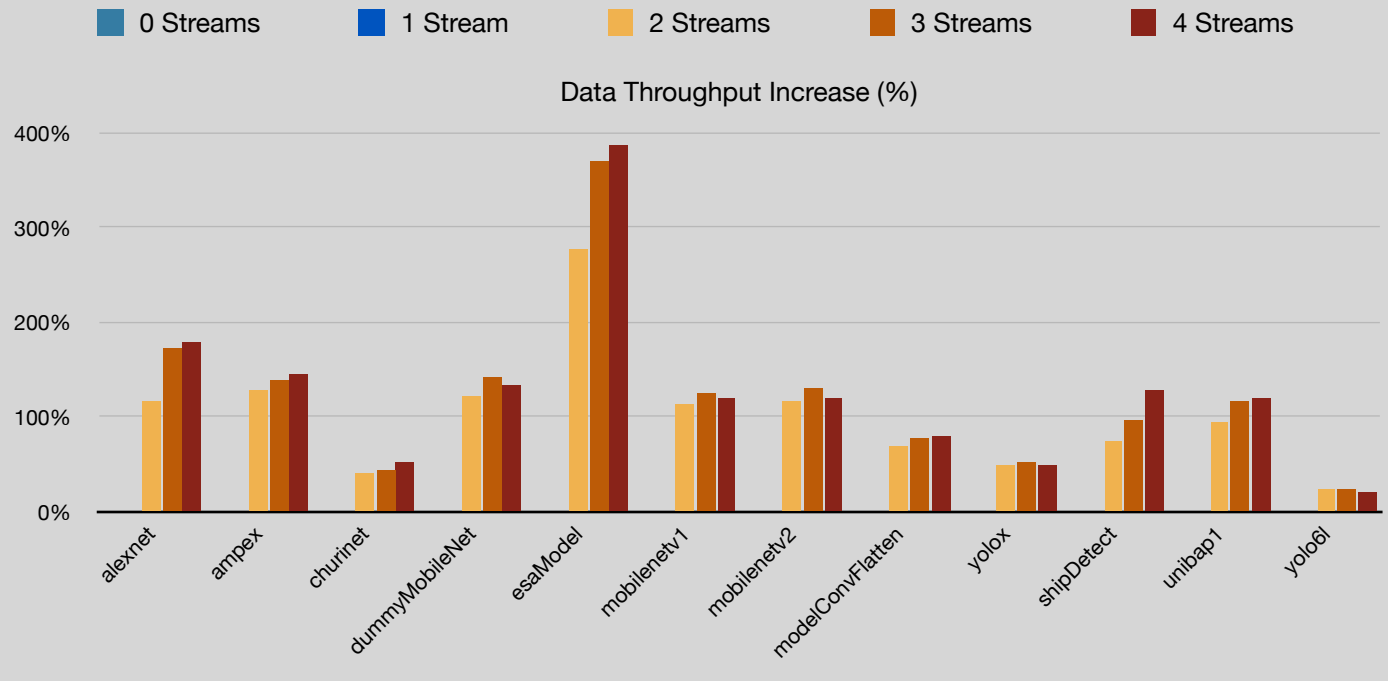


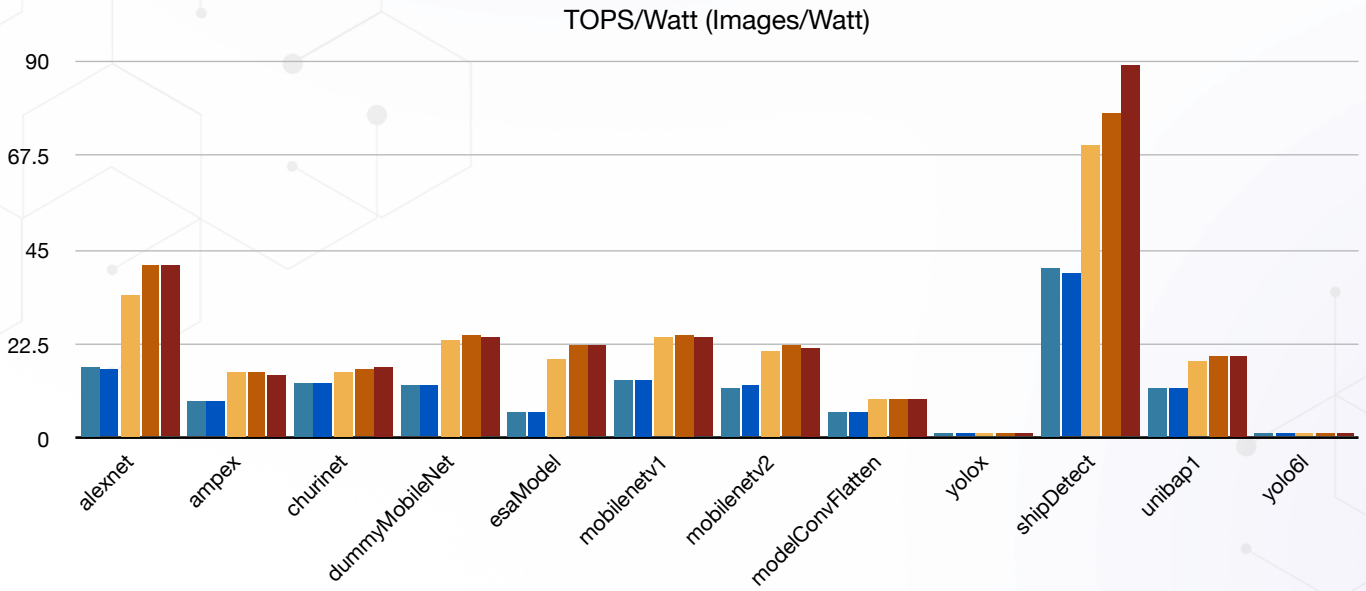
Benchmark Deep Neural Networks

Current benchmarks reflect a high-speed pipelined approach (sequential execution). Upcoming benchmarks will demonstrate the performance gains of fully independent, parallel TensorRT engine execution.

Benchmarks were carried out on the Jetson ORIN NX with 4 TensorRT instances for the algorithms: alextnet, mobilev2 and yolo.





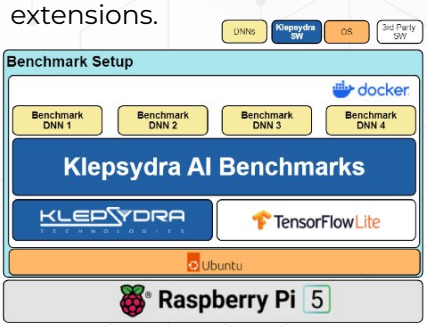


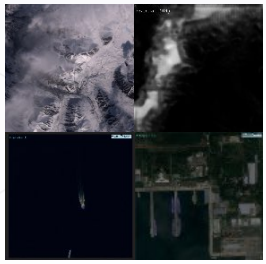
Klepsydra AI Performance Benchmarks Report

Version: 18 **Target**: Raspberry Pi5. **Date**: Dec. 2024

Technical setup

The benchmark application is run on a docker container on top of the Ubuntu 22.04 on the Raspberry Pi5 4-core CPU. The docker image contains both Klepsydra AI as well as TensorFlowLite 2.4.4 compiled with NEON extensions.





Benchmark Deep Neural Networks

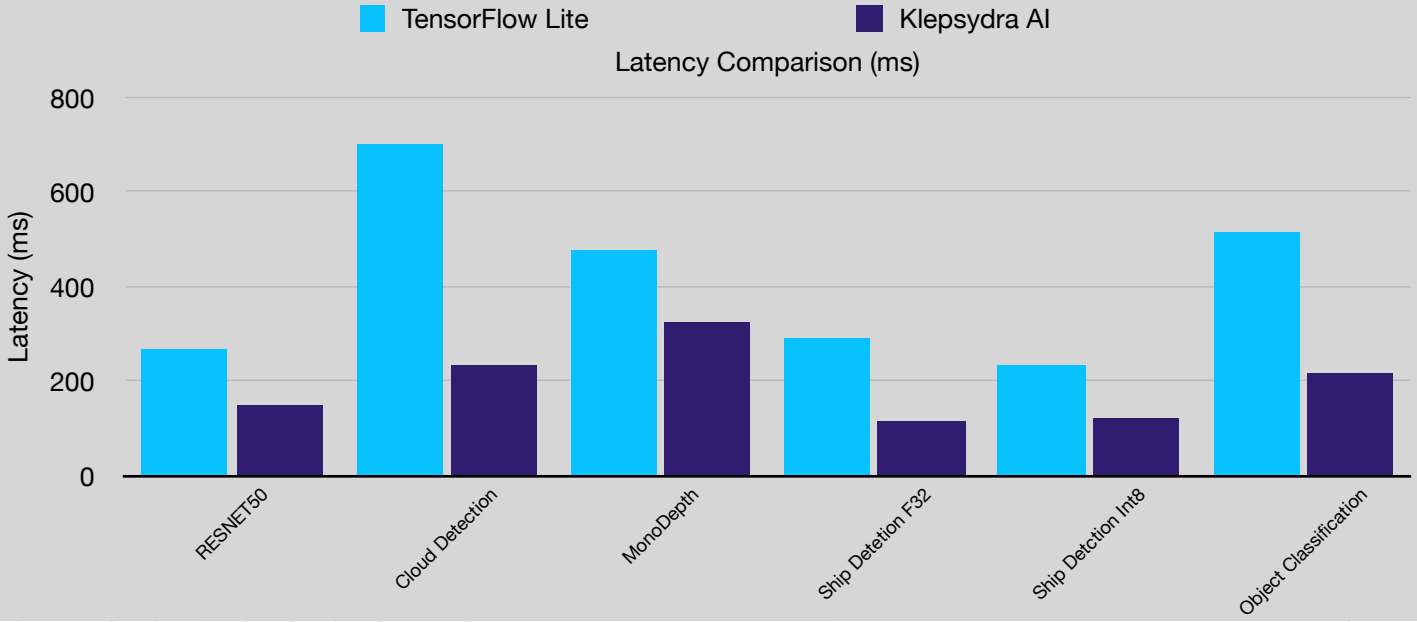
Several networks were tested as part of this campaign. These networks come from different sources as specified:

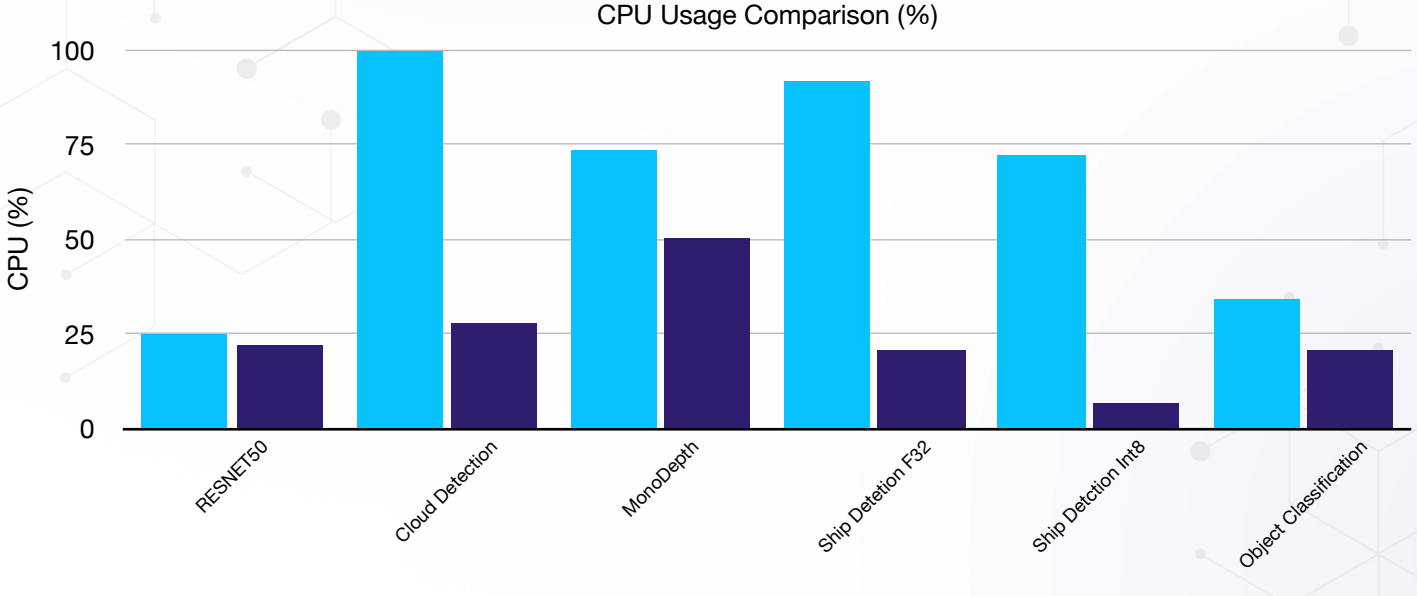
- RESNET50 (open source)
- Cloud Detection Quantised (ESA OBPMark-ML[1])
- Monodepth (open source)
- YoloX / Ship Detection (ESA OBPMark-ML[1])
- YoloX / Ship Detection Quantised (ESA OBPMark-ML[1])
- Object Classification (D'Amico paper[2])

[1]: https://zenodo.org/records/5638577
[2]: https://arxiv.org/abs/2309.11645

The performance results are shown for latency, i.e., the time required to execute the AI algorithm for a given input data, and CPU consumption, i.e., how much CPU is used for executing the AI algorithm.

The results show that Klepsydra AI outperforms TensorFlowLite in terms of latency and CPU consumption for the Raspberry Pi5.



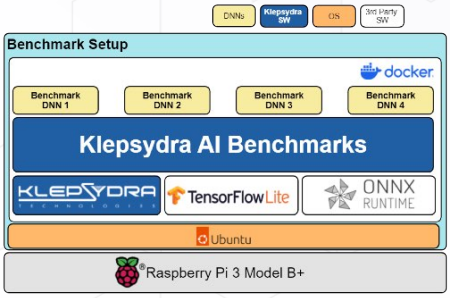


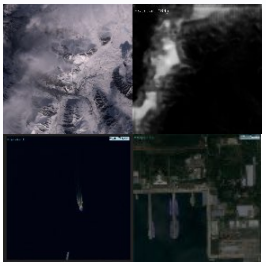
Klepsydra AI Performance Benchmarks Report

Version: 23 **Target**: Raspberry Pi3B+ **Date**: Dec. 2025

Technical setup

The benchmark application is run on a docker container on top of the Ubuntu 22.04 on the Raspberry Pi3B+ 4-core CPU. The docker image contains both Klepsydra AI as well as TensorFlowLite 2.4.4 and ONNX Runtime 1.20 compiled with NEON extensions.





Benchmark Deep Neural Networks

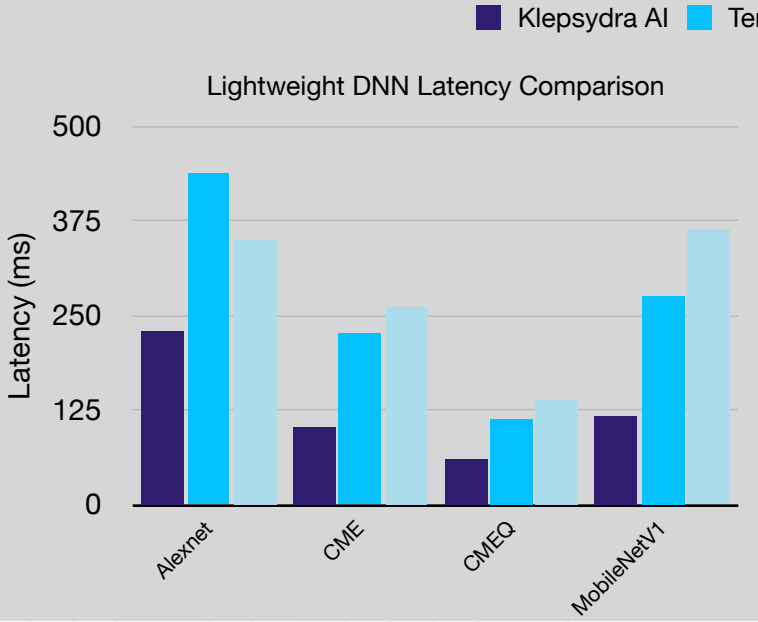
Several networks were tested as part of this campaign. These networks come from different sources as specified:

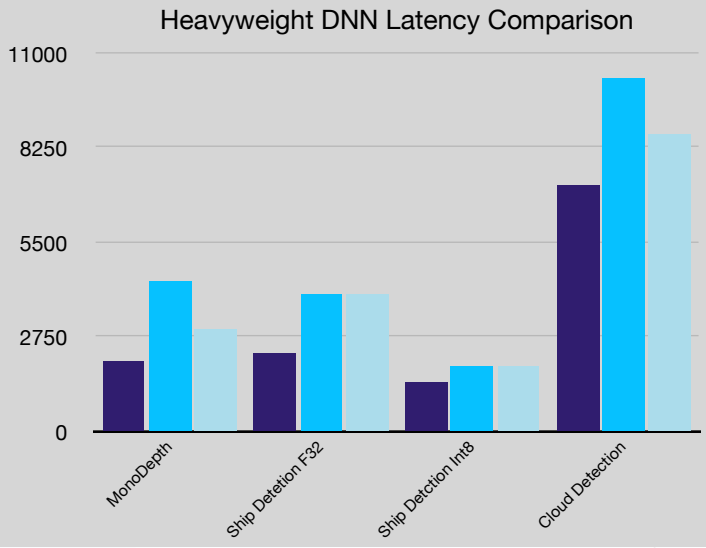
- RESNET50 (open source)
- Cloud Detection Quantised (ESA OBPMark-ML[1])
- Monodepth (open source)
- YoloX / Ship Detection (ESA OBPMark-ML[1])
- YoloX / Ship Detection Quantised (ESA OBPMark-ML[1])
- Coronal Mass Ejection Detect (ESA OBPMark-ML[1])
- MobileNetV1 (open source)

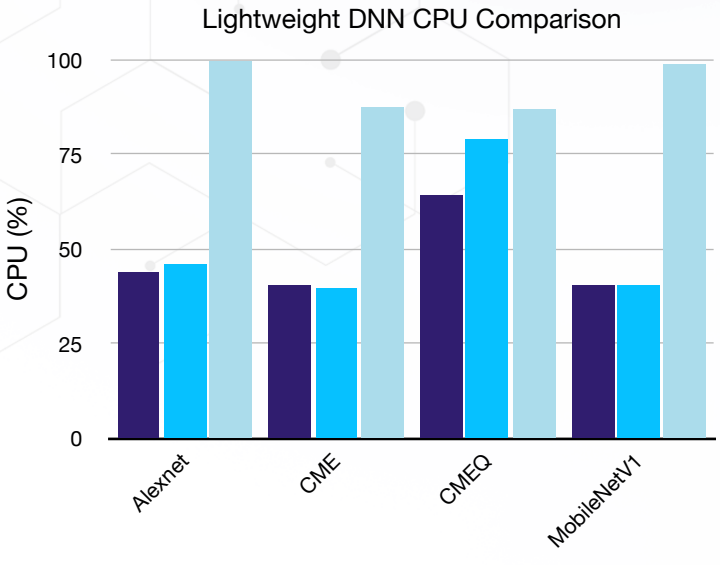
[1]: https://zenodo.org/records/5638577

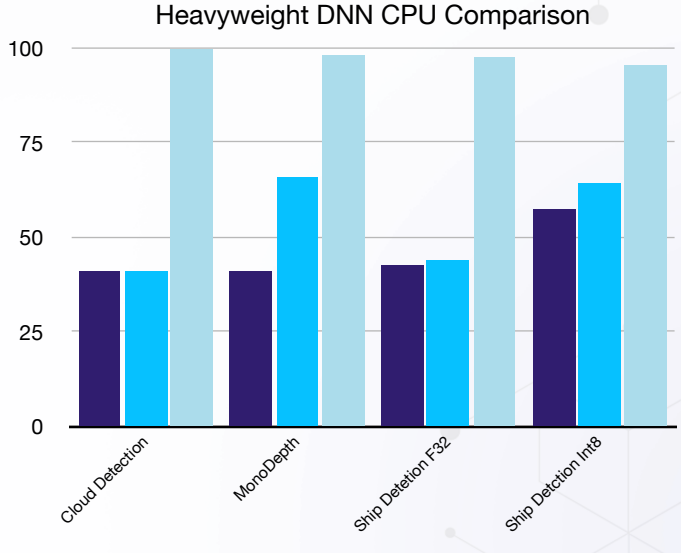
The performance results are shown for latency, i.e., the time required to execute the AI algorithm for a given input data, and CPU consumption, i.e., how much CPU is used for executing the AI algorithm.

The results show that Klepsydra AI outperforms both, TensorFlowLite and ONNX RT in terms of latency and CPU consumption for the Raspberry Pi3B+.







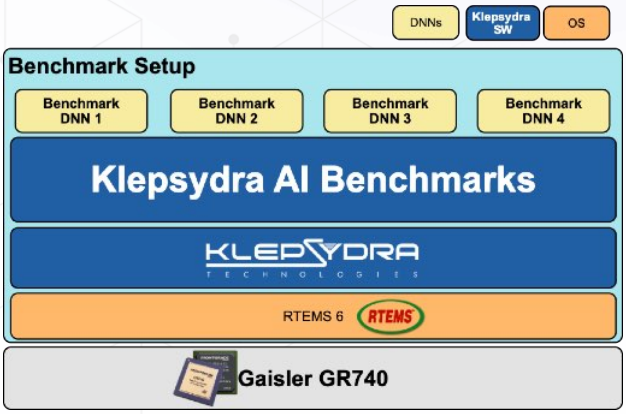


Klepsydra AI Performance Benchmarks Report

Version: 19 **Target**: Gaisler GR740. **Date**: Dec. 2025

Technical setup

A Deep Neural Network (DNN) is trained using standard tools like TensorFlow or PyTorch and exported as an ONNX file. This file is then loaded into Klepsydra AI, running on the GR740.



Benchmark Deep Neural Networks

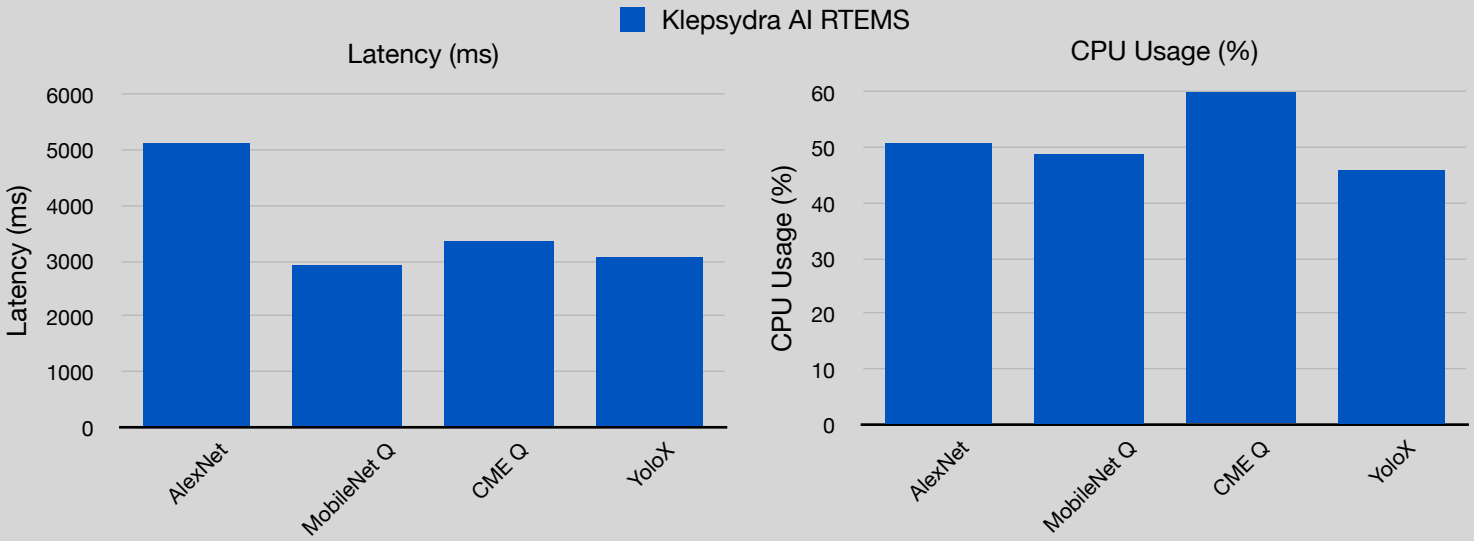
Several networks were tested as part of this campaign. These networks come from different sources as specified:

- AlexNet (open source)
- MobilNet Quantised
- ESA CME Quantised (ESA OBPMark-ML[1])
- YoloX / Ship Detection (ESA OBPMark-ML[1])

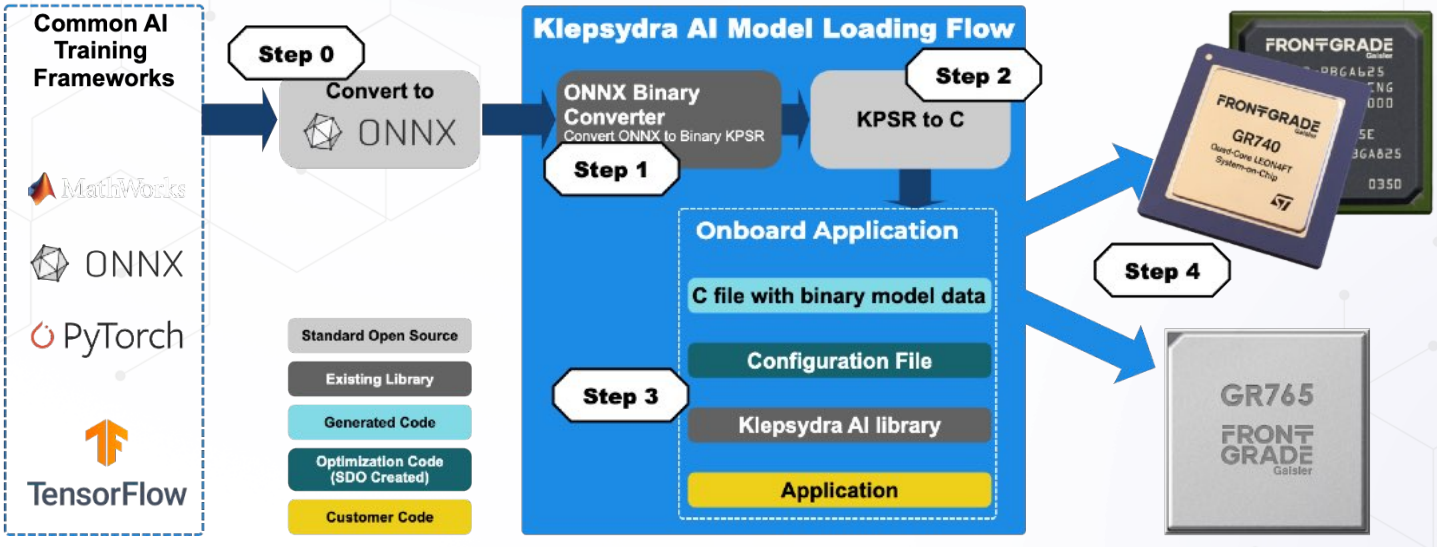
[1]: https://zenodo.org/records/5638577

The performance results are shown for latency, i.e., the time required to execute the AI algorithm for a given input data, and CPU consumption, i.e., how much CPU is used for executing the AI algorithm.

TensorFlow Lite does not run on GR740 - Klepsydra AI is currently the only AI inference framework capable of of running AI algorithms on the GR740.



Easy to use architecture for integrating AI into space systems

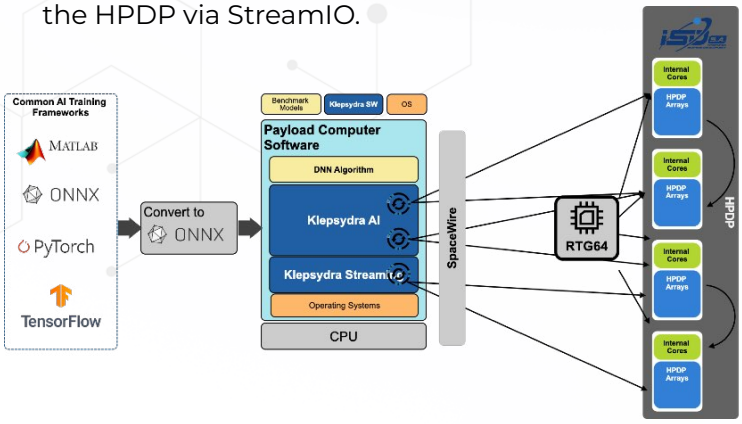


Klepsydra AI Performance Benchmarks Report

Version: 19 **Target**: HPDP **Date**: Mar. 2025

Technical setup

A Deep Neural Network (DNN) is trained using standard tools like TensorFlow or PyTorch and exported as an ONNX file. This file is then loaded into Klepsydra AI, running on the payload processor. The subset layers of the DNN that will be offloaded to the HPDP are sent to the HPDP via SpaceWire so that the FNC can load them into the XPP (code and configuration). Alternatively, this subset of layers can be loaded into the RTG4 RISC-V processor for execution in the HPDP via StreamIO.



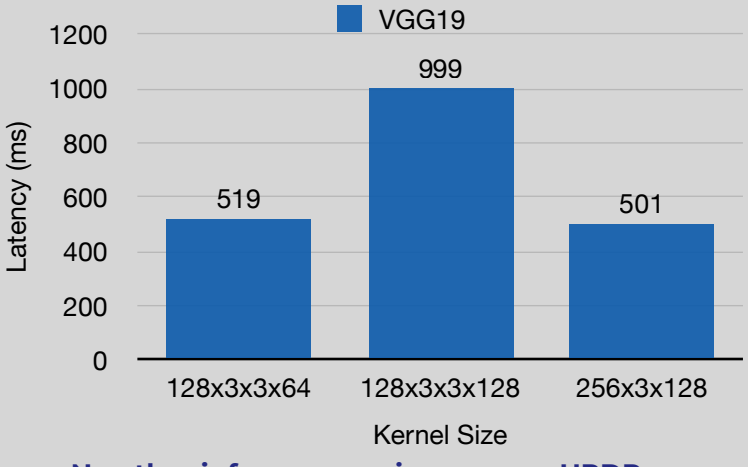
Benchmark Deep Neural Networks

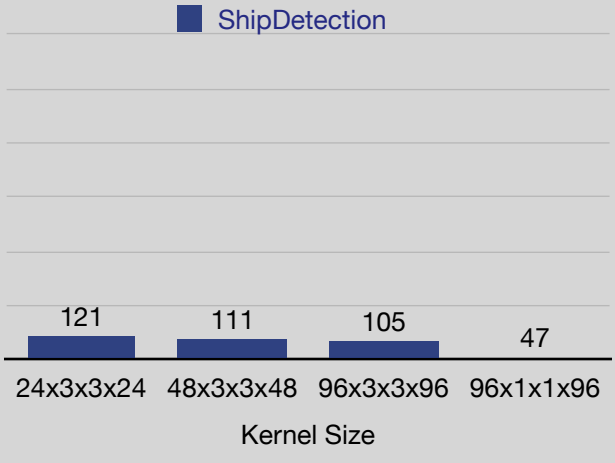
Several networks were tested as part of this campaign. These networks come from different sources as specified:

- ResNet50 Quantised (MLAB Project)
- MobileNetV2 Quantised
- YoloX / Ship Detection Quantised (ESA OBPMark-ML[1])
- VGG19 Quantised (MLAB Project)

New Capabilities

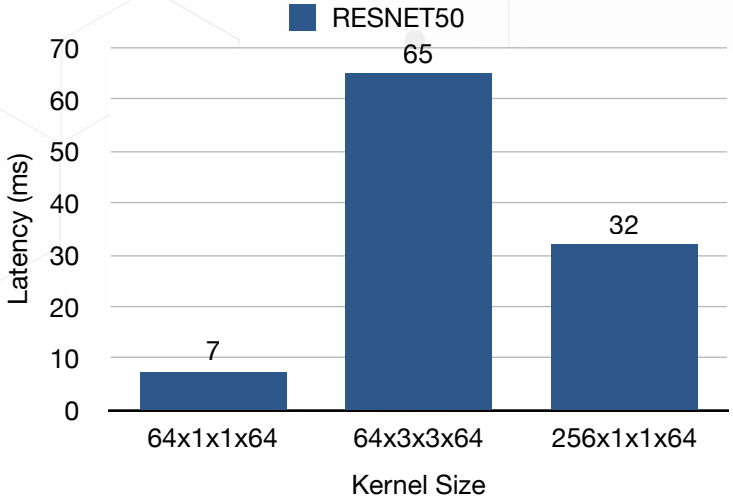
- Enable execution of AI on HPDP, a radiation hardened processor.
- Easy to use: Execute AI in 2-3 lines of code.
- No requirement for development of NML code for the HPDP: users can simply provide their existing AI models.
- Interoperability solution with Machine Learning (ML) frameworks: user can train AI models in any framework and provide the AI model file to payload computer + HPDP for execution without any development effort.

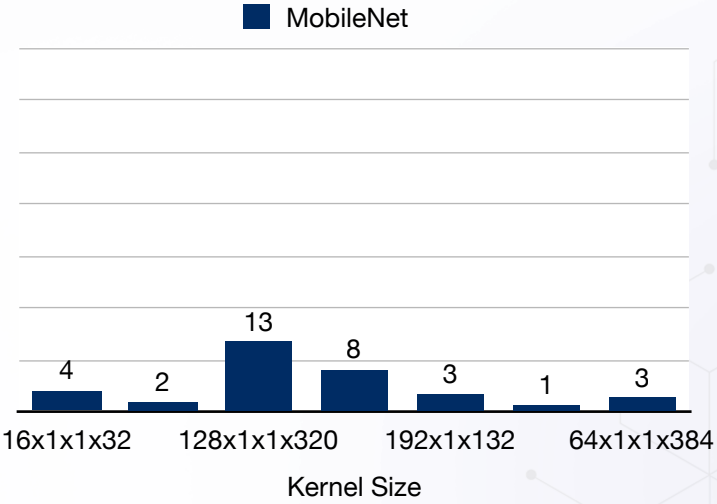




No other inference engine runs on HPDP

Klepsydra AI is the only AI inference framework capable of of running AI algorithms on the HPDP.





KLEPSYDRA

TECHNOLOGIES

Fast, secure and safe AI software for Space and Defence Systems



Software is available at
https://downloads.klepsydra.com/
Or contact
sales@klepsydra.com

Available in three formats:

- Docker images
- Debian packages
- Tarballs

A trial license covers:

- A 90-day license
- 10 hours of technical support
- A 1-hour technical workshop

Contact

Klepsydra Technologies AG
6300 Zug, Switzerland

www.klepsydra.com
sales@klepsydra.com

+Klepsydra Technologies
wa.me/41782493720