

Open Source AI: Easy Inference on the Edge

Chris Paterson – Renesas Electronics Europe

What is demonstrated

We have ported six popular Open Source AI frameworks to the Renesas RZ/G1 and RZ/G2 series of embedded SoCs:

- Arm NN SDK v19.08
- Caffe2 v0.8.1
- ONNX Runtime v0.4.0
- OpenCV v4.1.1
- TensorFlow v1.10.0
- TensorFlow Lite v1.10.0

The demo allows users to switch between the various AI frameworks on the fly and test object recognition with a range of pre-trained models.

Inference is done by the Arm Cortex + NEON cores.

What was improved

Most AI frameworks are designed to run in the cloud on large servers. More and more use cases are arising where inference needs to be run on embedded platforms – on the ‘edge’.

Many technologies aren’t easily cross-compiled though, so we have created OpenEmbedded/Yocto layers to do the hard work. These layers can easily be added on top of a BSP, for example meta-rzg2 from Renesas.

Sample applications have been included so that users can start testing/benchmarking their models straight away.

All code has been published to GitHub.

We plan to continue enhancing meta-renesas-ai with the most popular leading edge AI frameworks coming from the OS community.

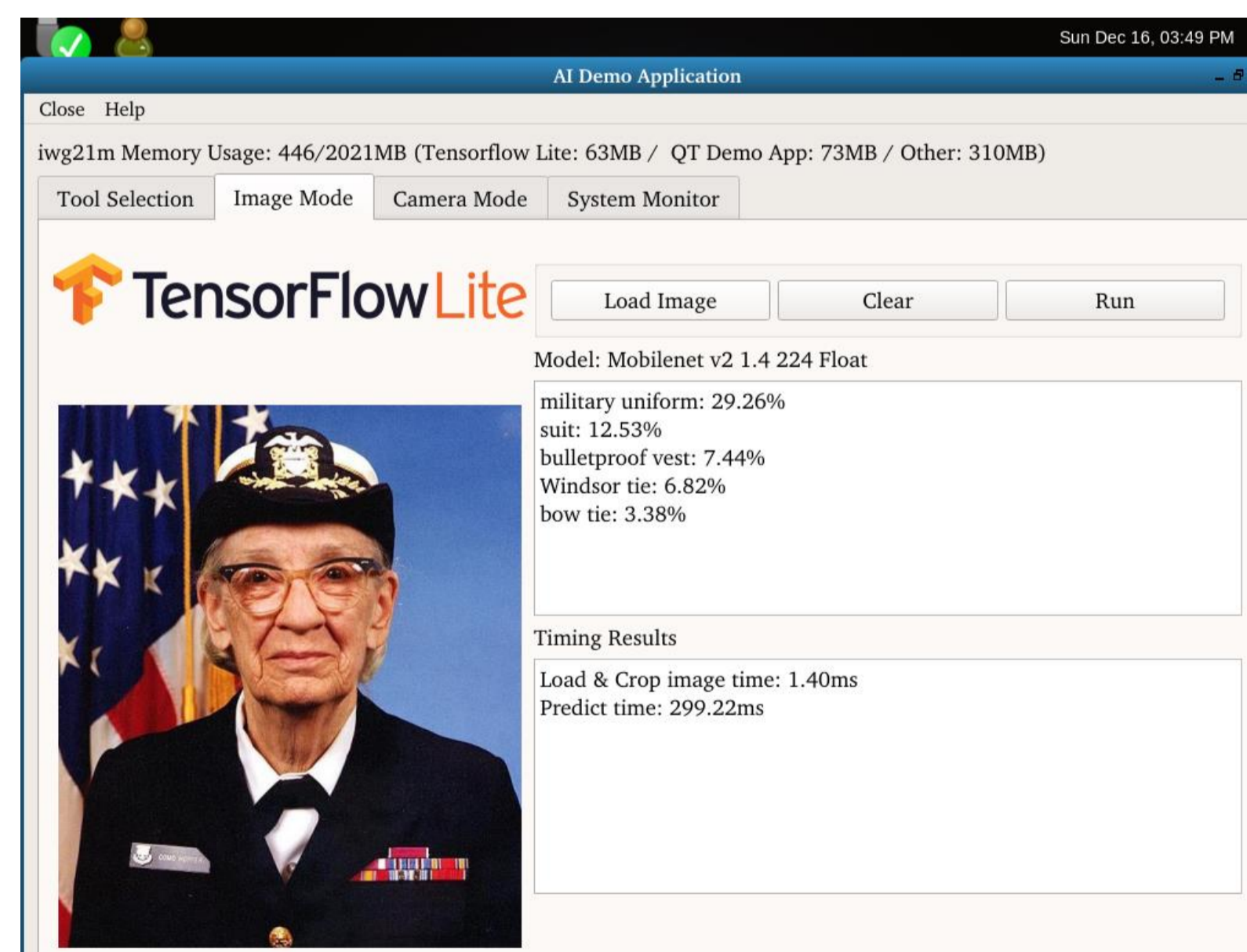
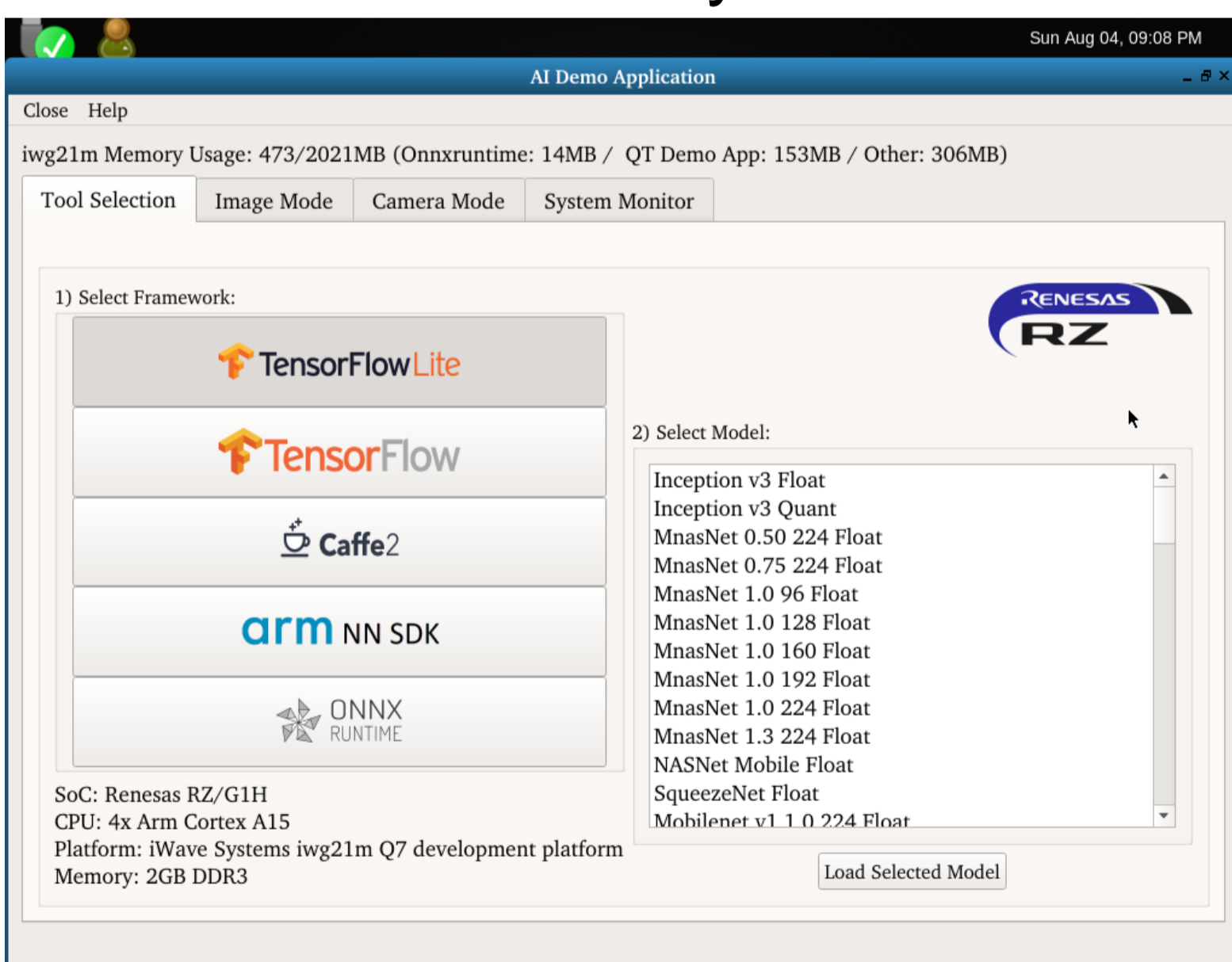
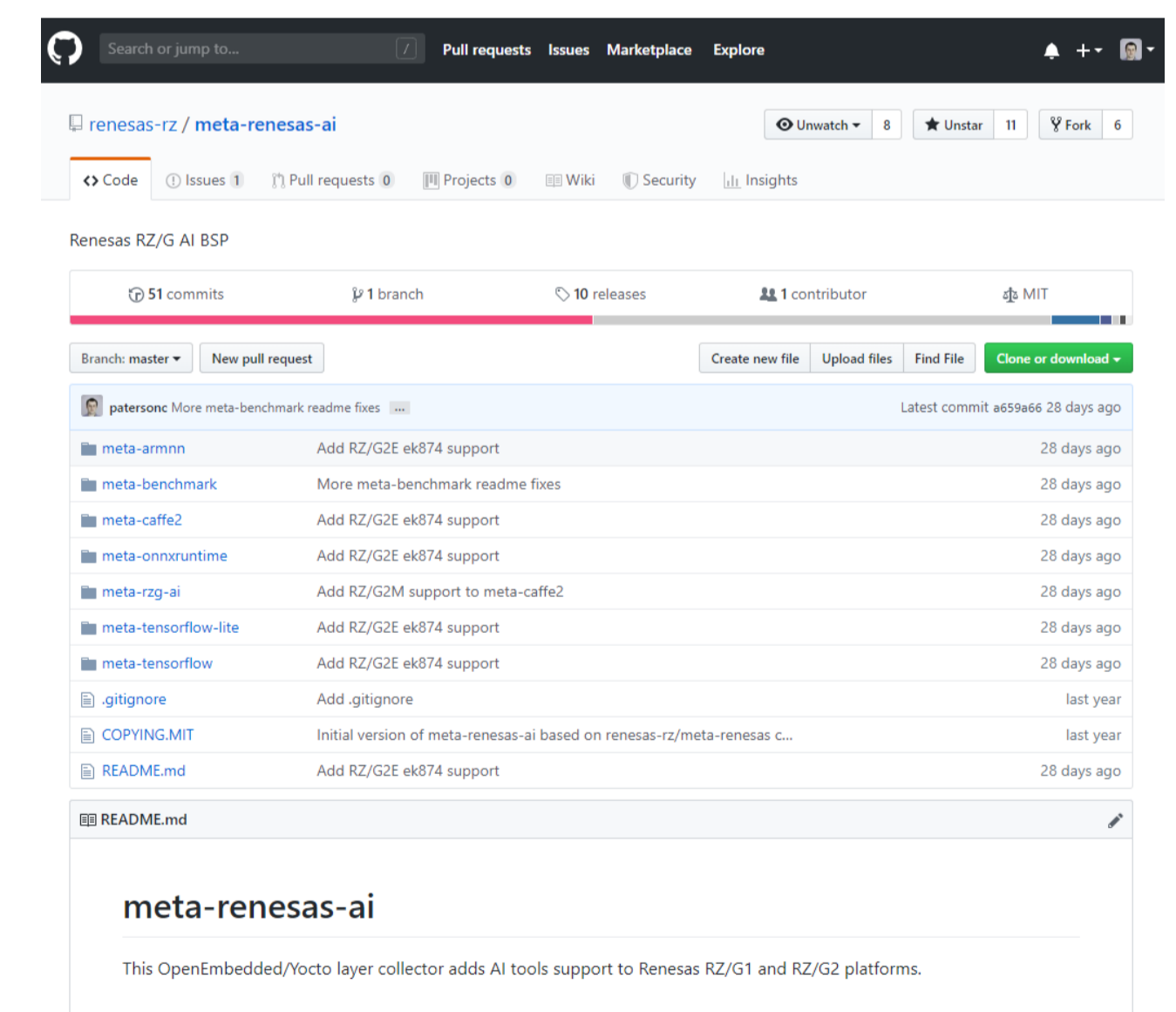
Feedback & contributions welcome!

Source code or detail technical information availability

RZ/G AI BSP: <https://github.com/renesas-rz/meta-renesas-ai>

RZ/G2 BSP: <https://github.com/renesas-rz/meta-rzg2>

HiHope-g2m: r8a774a1-hihopec-rzg2m-ex.dts (mainline & CIP SLTS v4.19)



Hardware Information

HiHope-g2m: <http://hihope.org/product/detail/rzg2>

- Renesas RZ/G2M SoC
- 2x Arm Cortex-A57 and 4x Arm Cortex-A53
- 4GB LPDDR4 SDRAM