

This checklist must be submitted as a PDF as part of your submission.

Name of Certifying Engineer(s): Parker Zhang

Email of Certifying Engineer(s): quic_xiaopeng@quicinc.com

Name(s) of System Under Test: Low Power AI Subsystem on Next Generation Snapdragon Mobile Platform Hardware Development Kit (HDK)

Division (check one):

- ☐ Open
- ☒ Closed

Category (check one):

- ☐ Available
- ☒ Preview
- ☐ Research, Development, and Internal (RDI)

Benchmark(s) (check all that apply):

- ☐ Visual Wake Words
- ☐ Keyword Spotting
- ☒ Anomaly Detection
- ☐ Image Classification

Please fill in the following table adding lines as necessary:

System Under Test Name	Benchmark	Accuracy/AUC
eAI on Snapdragon® 8 Gen 2 HDK	Visual Wake Words	
eAI on Snapdragon® 8 Gen 2 HDK	Keyword Spotting	
eAI on Snapdragon® 8 Gen 2 HDK	Anomaly Detection	77.4% / 0.86
eAI on Snapdragon® 8 Gen 2 HDK	Image Classification	

For each SUT, is the benchmark Accuracy/AUC target met? (Not a requirement for the Open division) (check all that apply):

- ☐ Yes (Visual Wake Words ... 80% Accuracy)
- ☐ Yes (Keyword Spotting ... 90% Accuracy)
- ☒ Yes (Anomaly Detection ... 0.85 AUC)
- ☐ Yes (Image Classification ... 85% Accuracy)
- ☐ No, for some combination of benchmark, scenario and SUT

For each SUT and benchmark, did the submission run on the whole validation set in accuracy mode? (check one):

- ☒ Yes

This checklist must be submitted as a PDF as part of your submission.

☐ No

For each SUT and benchmark, does the submission use the EEMBC Runner? (check one)

✓ Yes

☐ No

For each SUT and benchmark, is the same code run in accuracy and performance modes? (check one)

✓ Yes

☐ No

Are the weights calibrated using data outside of the official calibration set? (check one)

☐ Yes

✓ No

What numerics does the submission use? (check all that apply)

☐ INT4

✓ INT8

✓ INT16

☐ UINT8

☐ UINT16

☐ FP11

☐ FP16

☐ BF16

☐ FP32

☐ Other, please specify:

What backend does the submission use? (check all that apply)

✓ Vendor backend, please name: Qualcomm AI Stack

☐ TF-Lite Micro

☐ Micro TVM

☐ Other, please specify:

Which of the following caching techniques does the submission use? (check all that apply, ideally none):

☐ Caching Inputs between iterations

☐ Caching responses between iterations

☐ Caching intermediate computations between iterations

Which of the following techniques does the submission use? (check all that apply, ideally none if submitting to the closed division.)

☐ Quantization aware training

This checklist must be submitted as a PDF as part of your submission.

- ☐ Wholesale weight replacement
- ☐ Weight supplements
- ☐ Discarding non-zero weight elements
- ☐ Pruning
- ☐ Modifying weights during the timed portion of an inference run
- ☐ Hard coding the total number of queries
- ☒ None of the above

Is the submission congruent with all relevant MLPerf rules?

- ☒ Yes
- ☐ No

If the answer to the above question is no, please explain:

For each SUT, have you filled out the JSON system description file?

- ☒ Yes
- ☐ No

For each SUT, does the submission accurately reflect the real-world performance of the SUT?

- ☒ Yes
- ☐ No

Does your submission include the following: (check all that apply)

- ☒ System description file
- ☒ Code that implements the benchmarks
- ☐ Code/scripts that train the model(s) (Open Division)
- ☐ Metadata that describes each system-implementation combination tested
- ☐ Scripts that set up and execute each system implementation tested
- ☒ Result logs for each system implementation tested
- ☒ This Checklist