TESTING AND ANALYZING THE PERFORMANCE OF SINGLE-BOARD COMPUTERS FOR AI TASKS

D. A. Paulenka, E. V. Snezhko, V. A. Kovalev

The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus
E-mail: dmitri.pavlenko@gmail.com, eduard.snezhko@gmail.com, vassili.kovalev@gmail.com

Performance tests were conducted on the Raspberry Pi 4 Model B (Pi4) and Cool Pi 4 Model B (CoolPi) single board computers, as well as the Google Coral USB Accelerator with Google Edge TPU (Coral). Their performance for AI tasks was analyzed. CoolPi handles classification using MobileNet v3 well, but is 4 times slower than similar computations on Coral. Classification on Pi4 is 22 times slower than Coral. CoolPi outperforms Pi4 by a factor of 2 to 3 in data copying and compression tasks and almost 6 times in neural network computations. Coral gives a significant acceleration of AI tasks, but there are features related to the novelty of the technology.

Keywords: single-board computers; neural processors; neural network accelerators; AI accelerators; performance testing; benchmarking.

ТЕСТИРОВАНИЕ И АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ ОДНОПЛАТНЫХ КОМПЬЮТЕРОВ ДЛЯ РЕШЕНИЯ ЗАДАЧ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Д. А. Павленко, Э. В. Снежко, В. А. Ковалев

Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь
E-mail: dmitri.pavlenko@gmail.com, eduard.snezhko@gmail.com, vassili.kovalev@gmail.com

Проведены тесты производительности одноплатных компьютеров Raspberry Pi 4 Model B (Pi4) и CoolPi 4 Model B (CoolPi), а также USB-ускорителя Google Coral с TPU Google Edge (Coral). Проанализирована их производительность в задачах искусственного интеллекта. CoolPi хорошо справляется с классификацией с помощью MobileNet v3, но в 4 раза медленнее, чем аналогичные вычисления на Coral. Классификация на Pi4 выполняется в 22 раза медленнее, чем на Coral. CoolPi превосходит Pi4 в 2-3 раза в задачах копирования и сжатия данных и почти в 6 раз в нейросетевых вычислениях. Coral дает значительное ускорение выполнения задач ИИ, но есть и особенности, связанные с новизной технологии.

Ключевые слова: одноплатные компьютеры; нейронные процессоры; ускорители нейронных сетей; ИИ-ускорители; тестирование производительности; бенчмаркинг.

INTRODUCTION

Modern image processing methods and algorithms must utilize neural network-based artificial intelligence (AI) techniques to be innovative and competitive. The rapidly growing market for devices that use specialized microprocessors for efficient neural network training and/or operation requires new ways to practically implement neural network methods on these devices.

Two single-board computers (SBC) and one neural processor were selected and purchased for the performance tests:

- SBC Raspberry Pi 4 Model B (hereinafter referred to as Pi4) [1];
- SBC Cool Pi 4 Model B (hereinafter referred to as CoolPi) [2];
- Google Coral USB Accelerator with Google Edge TPU (hereinafter referred to as Coral) [3], which is not a standalone computer.

PERFORMANCE TESTING AND ANALYSIS

Figures 1-3 show the results of the performance tests for the three small devices (Pi4, CoolPi, Coral) and the Desktop personal computer (PC) with chipset Intel Lynx Point Z87 and CPU Intel i7-4770.

Desktop tests are conducted for illustrative and comparative purposes. It was not possible to find free cross-platform software for Windows and Linux, so the Docker containers used for Windows were "phoronix/pts" [4] and "ubuntu" [5].

In Figure 1, the smaller the bar on the chart, the better. For clarity, Figure 1 shows the time scale on a logarithmic scale.

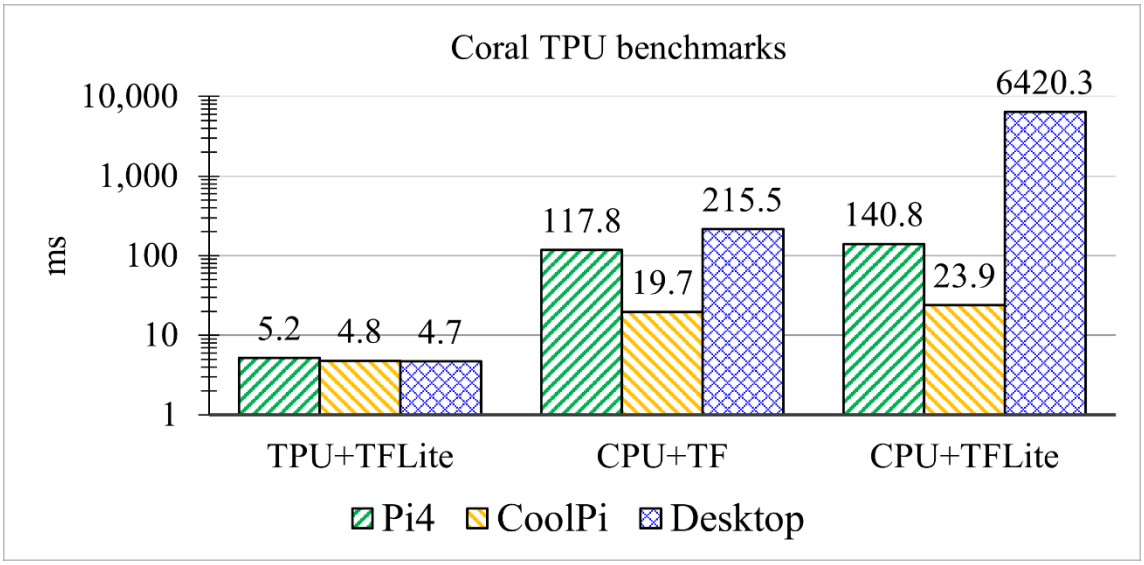

Fig. 1. Performance testing of the Coral TPU

"TPU+TFLite" contains performance test results for TPU Coral and TensorFlow Lite (TFLite) library. "CPU+TF" contains results for CPU and TensorFlow (TF) library. "CPU+TFLite" contains results for CPU and TensorFlow Lite (TFLite) library. TPU can only work with the TFLite library, so the fourth combination "TPU+TF" is not applicable.

Performance tests for Figure 1 were conducted on the classification task. The MobileNet v3 model pre-trained on ImageNet images was used for classification. The model files can be found in the Coral repository [6]:

- file "tf2_mobilenet_v3_edgetpu_1.0_224_ptq_edgetpu.tflite" (TPU);
- file "tf2_mobilenet_v3_edgetpu_1.0_224_ptq.tflite" (CPU).

The performance tests software can be found in the files "tf_lite_benchmarks.py" and "tf_lite.py" [7]. The log file with the test results is located in the GitHub repository file "data/coral-tpu-benchmark-results.txt" [7].

Figure 1 shows that the speed of calculations on Coral TPU ("TPU+TFLite") differs insignificantly for different devices and significantly outperforms CPU calculations. A comparison of the "CPU+TF" and "CPU+TFLite" showed that TensorFlow was slightly faster than TensorFlow Lite, except for the Desktop.

Perhaps, for Desktop PC, the significant 30-fold speed difference of "CPU+TF" (215.5 ms) and "CPU+TFLite" (6,420.3 ms) is due to some internal optimizations of the TF and TFLite libraries.

In Figure 2, the higher the bar of the chart, the better.

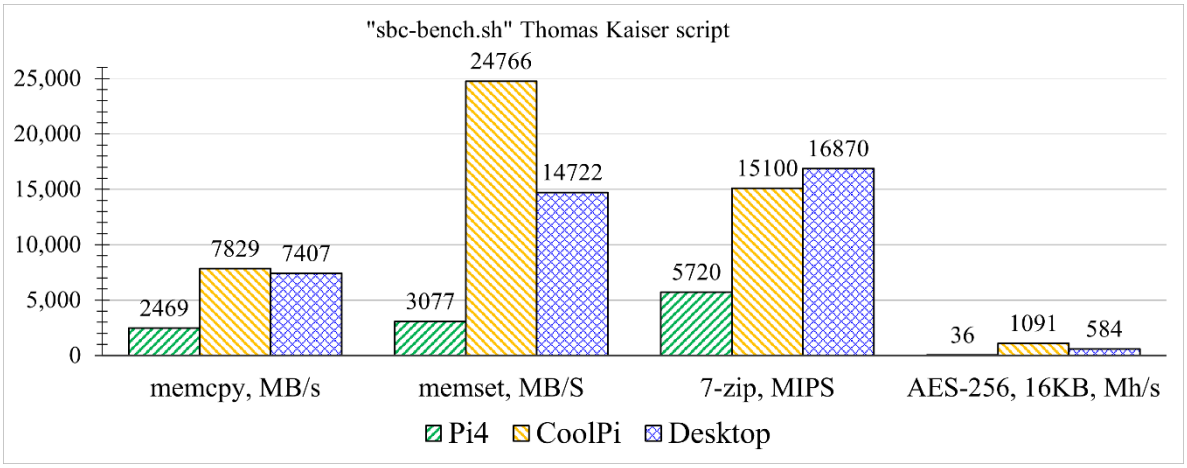

Fig. 2. Performance testing with the "sbc-bench.sh" script

In Figure 2, performance testing is performed using Thomas Kaiser software «sbc-bench.sh» [8], which is designed specifically for SBC performance tests.

"memcpy" is the speed of copying data from one block of memory to another using the "memcpy" function of the C/C++ programming language.

"memset" is the memory block initialization speed using the C/C++ "memset" function. "7-zip" is the speed of data compression using "7-zip" software. "AES-256" is the data encryption speed using "AES-256" algorithm. Performance tests have shown that only the Pi4 has no throttling, perhaps, due to the poor cooling system of CoolPi and Desktop PC.

The log file with the test results is located in the GitHub repository file "data/Thomas-Kaiser-sbc-bench-results.txt" [7].

In Figure 3, the higher the bar of the chart, the better.

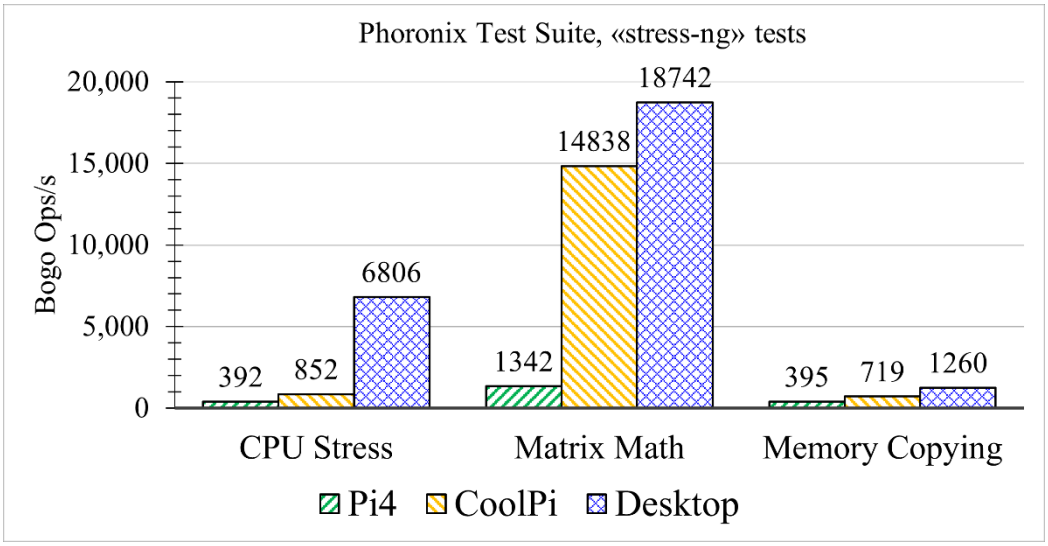

Fig. 3. Performance testing with Phoronix Test Suite, "stress-ng" test

In the Phoronix Test Suite software [9], performance is measured in Bogo Ops/s (bogus operations per second), which is a way of measuring the speed at which a computer executes instructions in the Linux kernel.

Phoronix Test Suite software [9] has hundreds of different performance tests, from which the "stress-ng" test was selected, and performance tests have been carried out: "CPU Stress", "Matrix Math" and "Memory Copying".

The log file with the test results is located in the GitHub repository file "data/phoronix-stress-ng-results.txt" [7].

CONCLUSIONS

Coral TPU gives a significant acceleration of AI tasks, but there are features related to the novelty of the technology that will be investigated in future works. CoolPi is 4 times slower than Coral. Pi4 is 22 times slower than Coral. CoolPi outperforms Pi4 by a factor of 2 to 3 in data copying and compression tasks and almost 6 times in AI tasks.

ACKNOWLEDGEMENTS

This work was carried out with the financial support of the Complex-SG space program.

REFERENCES

1. Raspberry Pi 4. [Electronic resource]. URL: https://www.raspberrypi.com/products/raspberry-pi-4-model-b (date of access: 27.03.2024).
2. Cool Pi 4 Model B – A much faster alternative to Raspberry Pi 4 SBC [Electronic resource]. URL: https://www.cnx-software.com/2022/12/04/cool-pi-4-model-b-powerful-raspberry-pi-4-alternative (date of access: 27.03.2024).
3. USB Accelerator [Electronic resource]. URL: https://coral.ai/products/accelerator (date of access: 27.03.2024).
4. phoronix/pts [Electronic resource]. URL: https://hub.docker.com/r/phoronix/pts (date of access: 27.03.2024).
5. ubuntu [Electronic resource]. URL: https://hub.docker.com/_/ubuntu (date of access: 27.03.2024).
6. google-coral/test_data [Electronic resource]. URL: https://github.com/google-coral/test_data. (date of access: 27.03.2024).
7. foobar167/junkyard/coral_tpu [Electronic resource]. URL: https://github.com/foobar167/junkyard/tree/master/coral_tpu (date of access: 27.03.2024).
8. ThomasKaiser/sbc-bench [Electronic resource]. URL: https://github.com/ThomasKaiser/sbc-bench (date of access: 27.03.2024).
9. Phoronix Test Suite [Electronic resource]. URL: https://www.phoronix-test-suite.com. (date of access: 27.03.2024).