

ИССЛЕДОВАНИЕ ПРОИЗВОДИТЕЛЬНОСТИ ВЫЧИСЛИТЕЛЬНОГО МОДУЛЯ NEUROMATRIX MC127.05 ПРИ ВЫПОЛНЕНИИ НА НЕМ НЕЙРОННЫХ СЕТЕЙ YOLO

С.Г. Небаба, к.т.н., старший преподаватель ИШИТР

С.А. Ткачёв, аспирант ИШИТР

Томский политехнический университет

E-mail: sat12@tpu.ru

Введение

Разработчики вычислительного модуля (ВМ) NEUROMATRIX MC127.05 позиционируют его как высокопроизводительный ускоритель при выполнении искусственных нейронных сетей, в частности, сверточных нейронных сетей (СНС), применяемых для распознавания объектов на изображениях [1]. В комплект поставки ВМ NEUROMATRIX MC127.05 входит пример СНС классической архитектуры YOLOv3 [2], предварительно обученной разработчиками (далее – предобученная СНС YOLOv3). Такая СНС позволяет решать задачи детектирования (обнаружения и локализации) объектов на изображениях. В комплект поставки ВМ также входят несколько цветных изображений из известного датасета MS COCO [3], каждое размером 416x416 пикселей.

Цель данной работы - исследовать производительность ВМ NEUROMATRIX MC127.05 при выполнении на нем СНС YOLOv3, а также провести сравнительный анализ производительности ВМ с производительностью других вычислительных устройств при вычислении этой же СНС.

Эксперименты по оценке производительности ВМ NEUROMATRIX MC127.05

К входящим в комплект поставки ВМ NEUROMATRIX MC127.05 цветным изображениям из датасета MS COCO были добавлены цветные изображения летательных аппаратов, в том числе класса «drone», и в итоге сформирована выборка из 35 изображений. Эта выборка использовалась в качестве тестовой в экспериментах с предобученной СНС YOLOv3.

Для проведения экспериментов по оценке производительности ВМ NEUROMATRIX MC127.05 нами была обучена СНС YOLOv3 с той же архитектурой, что и у предобученной СНС. Обучение такой СНС проводилось отдельно на созданных датасетах тепловизионных и цветных изображений объектов класса «drone».

В этих датасетах объемы обучающих выборок составляли 802 тепловизионных изображения и 802 цветных изображения, каждое изображение имело размер 416x416 пикселей. Также были сформированы две тестовые выборки: 122 тепловизионных изображения и 122 цветных изображения. При обучении были получены две новые модели СНС YOLOv3 в формате Darknet, отличающиеся друг от друга весовыми коэффициентами.

В первой серии экспериментов исследовалась производительность ВМ NEUROMATRIX MC127.05 при выполнении предобученной СНС. При этом на ее вход подавались цветные изображения из описанного выше тестового набора, созданного на основе входящих в состав модуля цветных изображений. Измерялось время, затрачиваемое СНС на обработку (анализ) каждого из 35 изображений. Результаты экспериментов в виде значений минимального и максимального времени обработки одного изображения, а также значение метрики производительности ВМ NEUROMATRIX MC127.05 в виде среднего значения времени обработки (анализа) одного тестового изображения приведены в таблице 1.

Таблица 1. Производительность ВМ NEUROMATRIX MC127.05 при выполнении СНС YOLOv3

Показатель	Предобученная СНС YOLOv3 из комплекта поставки вычислительного модуля	Обученная на тепловизионных изображениях модель СНС YOLOv3	Обученная на цветных изображениях модель СНС YOLOv3
Минимальное время обработки изображения, с	0,270	0,268	0,269
Максимальное время обработки изображения, с	0,271	0,270	0,271
Среднее время обработки изображения, с	0,270	0,269	0,270

Число изображений для тестирования, шт.	35	35	35
---	----	----	----

Во второй серии экспериментов также исследовалась производительность ВМ NEUROMATRIX MC127.05, но вычисления проводились для каждой из обученных и загруженных в память модуля моделей СНС. На вход первой из них подавались тестовые изображения из датасета тепловизионных изображений объектов класса «drone», а на вход второй модели СНС – цветные изображения объектов того же класса. Измерения времени обработки проводились для каждого из тестовых изображений. Результаты экспериментов представлены в таблице 1 для каждой из моделей СНС.

Сравнение производительности ВМ NEUROMATRIX MC127.05 с другими устройствами

С целью получения результатов для сравнительного анализа производительности ВМ NEUROMATRIX MC127.05 были проведены исследования производительности графического процессора NVIDIA Tesla K80 (видеокарта со средними вычислительными ресурсами 2014 года выпуска) и ноутбука с процессором Intel Core i5-8265U CPU (тактовая частота 1,80 ГГц) при выполнении двух обученных моделей СНС YOLOv3. Обученные модели СНС загружались в эти устройства в формате Darknet. Результаты третьей серии экспериментов по исследованию производительности этих двух устройств при выполнении на них двух обученных моделей СНС YOLOv3 приведены в таблице 2. Они получены при использовании тех же тестовых изображений, что применялись во второй серии экспериментов. Данные по пиковой производительности устройств в таблицу 2 взяты из документации, а результаты для ВМ перенесены в нее из таблицы 1.

Таблица 2. Результаты исследования производительности вычислительных устройств

Показатель	NEUROMATRIX MC127.05	NVIDIA Tesla K80	Intel Core i5-8265U
Пиковая производительность в формате одинарной точности, GFlop/s	512	8740	422
Среднее время обработки изображения, с	0,270	0,073	1,224
Потребляемая мощность, Вт	30	300	15

Заключение

Сравнительный анализ полученных результатов по оценке производительности трех устройств при выполнении на них СНС YOLOv3 позволяет сделать следующие выводы.

1. ВМ NEUROMATRIX MC127.05 уступает по скорости обработки изображений при выполнении СНС YOLOv3 графическому процессору NVIDIA Tesla K80 в среднем в 4,2 раза.
2. По сравнению с ноутбуком с процессором Intel Core i5-8265U CPU 1,80 ГГц без графического ускорителя скорость обработки изображений при выполнении СНС YOLOv3 на ВМ NEUROMATRIX MC127.05 выше в 4,5 раза.
3. Скорость обработки изображений на ВМ NEUROMATRIX MC127.05 при выполнении СНС YOLOv3 с одной архитектурой, но с разными весовыми коэффициентами и с разными типами входных изображений практически не меняется.
4. Потребляемая мощность NEUROMATRIX MC127.05 в 10 раз меньше, чем у NVIDIA Tesla K80.

Исходя из сравнения заявленной пиковой производительности, потребляемой мощности и полученных в экспериментах значений скорости обработки изображений, можно заключить, что ВМ NEUROMATRIX MC127.05 лучше адаптирован для выполнения СНС, по крайней мере, с архитектурой YOLOv3, поскольку разница в скорости обработки изображений по сравнению с графическим процессором NVIDIA Tesla K80 не так велика, как это можно было ожидать при сравнении пиковой производительности этих устройств. Вместе с тем, узкая специализация этого ВМ и средние показатели производительности ограничивают сферу его применения в задачах компьютерного зрения.

Список использованных источников

1. Модуль NEUROMATRIX MC127.05 [Электронный ресурс]. – URL: <https://www.neuromatrix.ru/directions/multimedia/12705> (дата обращения 27.02.2022).
2. YOLO: Real-Time Object Detection [Электронный ресурс]. – URL: <https://pjreddie.com/darknet/yolo> (дата обращения 27.02.2022).
3. MS COCO Dataset [Электронный ресурс]. – URL: <https://cocodataset.org/#home> (дата обращения 27.02.2022).