**Експерименти з паралельним навчанням згорткових нейронних мереж**

Дослідження проводилися з такою конфігурацією:

Таблиця 1 ‒ Використані комплектуючі

|  |  |
| --- | --- |
| Процесор | AMD Athlon 64 x2 Dual Core, 2.51ГГц |
| Оперативна пам’ять | 4ГБ, DDR2 |
| Відеокарта | Nvidia GTX 950 Asus, серія Strix, 2Гб, GDDR5 |
| PCI Express | 2.0 |
| CUDA Toolkit | 8.0 |

Для експериментів було вибрано два види зображень: цитологічні та гістологічні. Зображення кожного виду були Зображення кожного виду поділені на 5 класів. Кількість цитологічних зображень 80, гістологічних ‒ 91.

Для побудови нейронної мережі використовувалася бібліотека Deeplearning4j, яка дозволяє виконувати обчислення на графічному процесорі.

Модель використаної нейромережі зображено на рисунку 1.1.

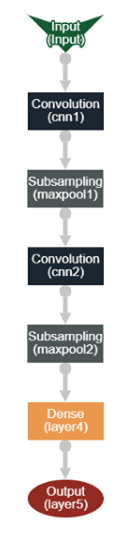


Рисунок 1.1 ‒ Модель нейромережі

Для першого експерименту використовувалися кольорові RGB зображення роздільною здатністю 128х128 пікселів.

Для кожного типу зображень було проведено 3 досліди. Після кожного досліду, в моделі мережі коригувалися деякі параметри для досягнення кращого результату, а саме: норма навчання, алгоритм оптимізації. Результати навчання мережі наведено у таблиці 1.1.

Таблиця .1 ‒ Результати навчання (128px)

|  |  |  |  |
| --- | --- | --- | --- |
| Вид зображень | Час навчання на GPU (хв) | Час навчання на CPU (хв) | Точність класифікації (%) |
| Цитологічні | 3 | 15 | 100 |
| 3 | 15 | 93 |
| 2 | 11 | 100 |
| Гістологічні | 4 | 18 | 90 |
| 3 | 16 | 93 |
| 2 | 10 | 93 |

Для другого експерименту використовувалися ті ж самі зображення, але з роздільною здатністю 224х224 пікселі. Для цих зображень також було проведено 3 досліди. Результати наведено у таблиці 1.2.

Таблиця 1.2 ‒ Результати навчання (224px)

|  |  |  |  |
| --- | --- | --- | --- |
| Вид зображень | Час навчання на GPU (хв) | Час навчання на CPU (хв) | Точність класифікації (%) |
| Цитологічні | 11 | 47 | 93 |
| 10 | 45 | 100 |
| 9 | 34 | 100 |
| Гістологічні | 9 | 36 | 90 |
| 10 | 45 | 90 |
| 10 | 45 | 90 |

Як видно з наведених таблиць, графічний процесор дає прискорення приблизно в чотири рази.

Також, для всіх наборів даних було використано batchSize = 5. Це значення вказує скільки зображень буде використовуватися одночасно при обчисленнях.

Після проведення тестів із зображеннями різної роздільної здатності (від 128 до 3200 пікселів), я дійшов висновку, що використана відеокарта не дає можливості працювати із зображеннями більшого розміру, ніж наведено в таблицях, оскільки має дуже малий обсяг відеопам’яті.

**Апаратні вимоги**

В процесі навчання нейронної мережі велику роль відіграють комплектуючі комп’ютера. Неможливо виграти в часі тільки за рахунок розпаралелювання на програмному рівні. Навіть, якщо весь код буде виконуватися паралельно на 100%, то він буде повільнішим на слабкішому процесорі. Я маю на увазі, що тут велику роль відіграє швидкість та обсяг оперативної пам’яті, швидкість шини PCI Express, тип відеопам’яті та багато інших факторів. Інколи, під час навчання нейронної мережі, не вистачає пам’яті відеокарти (це при малому розмірі зображень), тому, я б рекомендував наступні комплектуючі:

* RAM 8Гб, DDR4;
* Intel Pentium G4400 або i3-4170;
* Відеокарта Nvidia з обсягом пам’яті 4Гб;
* PCI Express 3.0.

**Попередня обробка зображень (CPU vs GPU)**

Для попередньої обробки зображень було використано бібліотеку *OpenCV*. Тестування проводилися на зображеннях з різним розміром. Над зображеннями виконувалися наступні операції: перетворення в градації сірого, фільтр Гауса, гістограм не вирівнювання. Результати наведені у таблиці 3.

Таблиця 3 ‒ Результати попередньої обробки

|  |  |  |
| --- | --- | --- |
| Розмір зображення, пікс. | Час виконання CPU, мс | Час виконання GPU, мс |
| 200x150 | 0.000004 | 0.000070 |
| 300x300 | 0.000005 | 0.000096 |
| 1200x675 | 0.000046 | 0.000120 |
| 3888x2592 | 0.000597 | 0.000145 |
| 15000x10800 | 0.019987 | 0.005338 |