

**АНАЛИЗ ЭФФЕКТИВНОСТИ ИСПОЛЬЗОВАНИЯ  
ГРАФИЧЕСКИХ ПРОЦЕССОРОВ ДЛЯ ОБУЧЕНИЯ  
НЕЙРОННОЙ СЕТИ РАСПОЗНАВАНИЯ РУКОПИСНЫХ  
СИМВОЛОВ.**

*Путенкова Валентина Юрьевна*

*Магистр*

*Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия*

*E-mail: misha156-8@yandex.ru*

В настоящее время особенно остро стоит задача машинного распознавания образов. Это связано, во-первых, с ежедневным увеличением объема графического контента в сети Интернет, а так же способа его систематизации и анализа. Во-вторых, с проблемой автоматизации многих рутинных операций. Таких, как распознавание текстов рукописных документов, банковских чеков, анкет, почтовых индексов и многих других.

Нейронные сети успешно применяются для решения задач классификации и распознавания образов, в частности, графически представленных, символов. Так как в силу нелинейности своей природы нейронные сети могут воспроизводить достаточно сложные зависимости. При этом уровень знаний, необходимый для успешного применения нейронных сетей, гораздо скромнее, чем, например, при использовании традиционных статистических методов.

В процессе обучения нейронной сети выполняется большое число однотипные векторно-матричных операций над, как правило, достаточно большим объемом данных. В связи с этим возникает потребность в использовании средств параллельной разработки программ для ускорения расчетов.

Целью данной работы является сравнительный анализ эффективности использования графических процессоров для обучения нейронной сети задаче распознавания образов. Данная работа проводилась на кафедре Суперкомпьютеров и Квантовой Информатики в рамках научной работы под руководством к.ф.м.н. Д.Ю. Буряка.

Данное исследование проводилось на персональном компьютере со следующими характеристиками:

Процессор: Intel(R) Core i7-3610QM; CPU 2,3 GHz; Озу 6 Гб, 64-разрядная ОС Microsoft Windows 7 Home Premium; среда разработки Microsoft Visual Studio 2010, C++. Графическая карта: GeForce GT 640M LE; CUDA Driver Version 5.5; 2 multiprocessors 1505MHz, 48 CUDA cores 1.5GHz; пакет NVIDIA Nsight HUD 3.1.

Для решения задачи была выбрана нейронная сеть прямого распространения типа "персептрон" и пять наиболее известных [3,4] методов обучения: метод обратного распространения ошибки, метод суммарного градиента с постоянным шагом, метод сопряженных градиентов [2], метод наискорейшего спуска и стохастическо-диагональный метод Левенберга-Марквардта [5]. Все пять методов были реализованы в одной программе с возможностью переключения между режимами обучения. Для каждого из методов была реализована последовательная и параллельная версии обучения, причем модуль, отвечающий за реализацию параллельной версии, исполняет все векторно-матричные вычисления на графической карте с поддержкой технологии CUDA [1].

По результатам исследования были сделаны следующие выводы:

1. Использование графического процессора позволило повысить скорость обучения в полтора раза по сравнению с параллельной версией программы.

2. Вычисления на графическом процессоре только части векторно-матричных операций недостаточно для получения приемлемого ускорения программы. Так как при этом большая часть вычислительного времени уходит на передачу данных с графической карты на центральный процессор и обратно.

### Литература

1. Боресков А. В., Харламов А. А. Основы работы с технологией CUDA. М.: ДМК Пресс, 2010.
2. Гилл Ф., Мюррей У., Райт М. Практическая оптимизация. М.: Мир, 1985.
3. Уоссермен Ф. Нейрокомпьютерная техника. Теория и практика. М.: Мир, 1992.
4. Хайкин С. Нейронные сети: полный курс. М.: Вильямс, 2006.
5. Yann LeCun Gradient based learning applied to document recognition // In Proceedings of the IEEE, nov., 1998.