

УДК 004.93

ПРОГРАММНО-АППАРАТНАЯ ПОДДЕРЖКА РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ СВЕРТОЧНОЙ НЕЙРОСЕТИ

Холошня В.Д. (Университет ИТМО)

Научный руководитель – к.т.н. Болдырева Е.А.

(Университет ИТМО)

В работе рассматривается вопрос программно-аппаратной поддержки распознавания изображений с помощью сверточной нейронной сети и использования данных средств в сфере образования. В настоящее время существует большое количество программных продуктов для работы с изображениями, однако все они требуют фундаментальной теоретической и практической предварительной подготовки и используют аппаратную часть в качестве “черного ящика”. В ходе исследования выдвинута гипотеза: для того, чтобы снизить порог вхождения и улучшить качество образования в сфере искусственного интеллекта, необходимо раскрыть для обучаемых аппаратную архитектуру, на которой выполняется нейронная сеть, и дополнить программную часть системы пользовательским интерфейсом. Это позволит сместить фокус с программной составляющей нейронной сети на программно-аппаратную и продемонстрировать студентам полную картину функционирования систем распознавания изображений. Разрабатываемая система представляет из себя программно-аппаратный комплекс для распознавания изображений с помощью сверточной нейронной сети YOLOv3 с аппаратной поддержкой в виде USB-ключа Kneron. Программное обеспечение представляет из себя систему, состоящую из двух серверов и веб-приложения. Разрабатываемый комплекс позволит облегчить обучение основам искусственного интеллекта, а также уменьшит порог вхождения в данную область.

Введение. Анализ существующих на рынке программных продуктов для изучения искусственного интеллекта показал, что эти системы отлично справляются с тем, чтобы продемонстрировать выполнение алгоритмов машинного обучения с точки зрения теории внутреннего устройства нейронных сетей. Они реализуют упрощенный вариант применения полученных знаний на практике, а аппаратная поддержка сети остается для студента “черным ящиком”.

Программно-аппаратные комплексы, в свою очередь, позволяют в полной мере изучать работу алгоритмов на практике за счет возможности выполнять необходимые алгоритмы на вычислительном устройстве локально, однако таких систем мало. Существующие программно-аппаратные системы, такие как Intel Neural Compute Stick, являются профессиональными инструментами, созданными, в основном, для использования во встраиваемых системах. Они не предназначены для использования в сфере образования.

Таким образом, на данный момент не найдено доступной системы, адаптированной под использование в качестве системы для изучения работы нейронных сетей. В данной исследовательской работе предложен прототип программно-аппаратного комплекса для распознавания изображений с помощью нейросетей. Особенность комплекса - прозрачность программной и аппаратной частей для студента, который занимается разработкой и обучением алгоритмов распознавания изображений. Такой комплекс вместе с методическим сопровождением позволит уменьшить порог вхождения в сферу искусственного интеллекта и повысить качество обучения.

Основная часть. В качестве аппаратной части разрабатываемой системы были выбраны USB-ключи. Данные доступные устройства позволяют обеспечить универсальность аппаратной части системы, поскольку их можно использовать совместно с любыми учебными компьютерами, имеющими соответствующий способ подключения.

Из возможных производителей USB-ключей были выделены два: Intel и Kneron. Intel предоставляет USB-ключ Neural Compute Stick. Kneron предоставляет USB-ключи и чипы. Данные продукты поддерживают работу с нейронными сетями и включают свое программное обеспечение, однако Kneron также предоставляет дополнительные возможности. Разработчики Kneron предоставляют средства для эмуляции работы чипов с помощью Docker, что упрощает разработку. Помимо этого программное обеспечение Kneron включают в себя заранее созданные и оптимизированные реализации нейронных сетей и алгоритмов машинного обучения, а также библиотеки для взаимодействия с ними, что позволяет ускорить разработку системы в целом. Учитывая данные преимущества, для данного проекта могут быть выбраны чипы Kneron KL520 и KL720.

Помимо предоставляемого разработчиками программного обеспечения, ведется работа над прошивками, которые изменяют вид принимаемых и возвращаемых данных в необходимом для остальной системы формате. Данные прошивки позволяют работать с различными нейронными сетями, а также проводят нужные виды постобработки.

Программная часть представляет из себя систему из двух серверов и веб-приложения. Один из серверов размещается на каждом USB-ключе и обеспечивает доступ к низкоуровневому API. Данный сервер позволяет взаимодействовать непосредственно с аппаратной частью системы.

Второй сервер устанавливается на главном учебном компьютере и позволяет сохранять образцы и результаты исследования между практическими занятиями, а также реализует систему авторизации пользователей.

Клиентское веб-приложение представляет из себя интерактивный пользовательский интерфейс для проведения необходимых работ при изучении искусственного интеллекта. Помимо интерфейса приложение отвечает за взаимодействие и связь между пользователем, главным сервером и сервером на USB-ключе.

В качестве алгоритма для распознавания объектов на изображении была выбрана сверточная нейронная сеть YOLOv3. В настоящее время, данная нейронная сеть является одной из самых точных. По умолчанию она предоставляет возможность распознавания 80-ти различных классов объектов на скорости до 45-ти кадров в секунду. Данная скорость обеспечивается внутренним устройством сети и, чтобы распознать объекты, достаточно пропустить изображение через сеть всего один раз. Данная особенность позволяет демонстрировать работу нейронной сети в реальном времени.

Выводы. В данной работе рассмотрен разрабатываемый программно-аппаратный комплекс для изучения распознавания объектов на изображении с помощью сверточной нейронной сети YOLOv3. В качестве аппаратной поддержки системы предпочтительны USB-ключи Kneron с чипами KL520 и KL720.

Разрабатываемая система позволяет в полной мере изучить устройство нейронной сети, так как система предоставляет возможность локального проведения вычислений и выполнения алгоритмов. Данные возможности позволяют облегчить, ускорить и улучшить качество обучения в сфере искусственного интеллекта.

Холошня В.Д. (автор)

Подпись

Болдырева Е.А. (научный руководитель)

Подпись