

ОТЗЫВ

официального оппонента на диссертацию Калиновского Ильи Андреевича «Метод нейросетевого детектирования лиц в видеопотоке сверхвысокого разрешения», представленную на соискание учёной степени кандидата технических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Цель диссертационной работы заключается в разработке эффективного высокоскоростного метода детектирования лиц в видеопотоке сверхвысокой чёткости (3840×2160 пикселей, 25 – 30 кадров/с) на маломощных вычислительных устройствах и обеспечивающего при этом высокие показатели полноты и точности.

Задача обнаружения лиц в видеопотоке стандарта *4K Ultra HD* ранее практически не рассматривалась. Автор представил максимально полный обзор современных алгоритмов обнаружения лиц, начиная с классического метода Виолы-Джонса и заканчивая самыми актуальными нейросетевыми подходами. Особое внимание уделяется изучению границ применимости предложенной системы детектирования лиц по углам поворота лица, а также по освещённости.

Предложенный детектор лиц протестирован на ряде актуальных наборов тестовых данных (*FDDDB, AFW, IJB-A*), содержащих как статические изображения, так и видеоданные. Производится сравнение с множеством других современных аналогов, в том числе с теми алгоритмами, которые широко используются в биометрии и промышленном видеонаблюдении. При этом вычисляются и графически иллюстрируются несколько метрик качества (точность, полнота, *F*-мера).

Предложенный детектор состоит из каскада компактных свёрточных нейросетей, что позволяет существенно, более чем на 99%, сузить количество анализируемых подобластей исходного изображения. Автор приложил много усилий для оптимизации детектора, что выгодно отличает его диссертационную работу, представленную на соискание учёной степени кандидата технических наук, от академических статей по обнаружению лиц. Рассмотрено несколько способов оптимизации - при помощи аппаратных средств платформы *Intel* (компиляторная

оптимизация, использование векторных функций), а также при помощи представления операции свёртки, традиционно отнимающей в свёрточных нейросетях более 90% вычислительного времени, в виде полиномиальных преобразований. Кроме того, автором предпринята оптимизация каскадного детектора с точки зрения эффективности доступа к памяти в рамках парадигм *CUDA* и *OpenCL*. Проведено профилирование детектора на нескольких аппаратных архитектурах для видеокадров разного разрешения (от *VGA* до *4K Ultra HD*).

Диссертационная работа богато проиллюстрирована: приводятся подробные чертежи архитектур свёрточных нейросетей, блок-схемы и псевдокод алгоритмов, примеры предварительной обработки изображений; показаны результаты кластеризации обучающей и валидационной выборки. Диссертация может служить полноценным учебным пособием для исследователей, интересующихся детектированием объектов.

К основным научным и практическим результатам работы, на мой взгляд, относятся:

- новый метод детектирования лиц, отличающийся от известных методов использованием в качестве классификатора каскада свёрточных нейронных сетей с компактными архитектурами, который обладает по сравнению с традиционными каскадными классификаторами лучшими свойствами, такими как значительно более низкая вероятность ложного срабатывания первой стадии каскада и высокая эффективность выполнения на процессорах с массивно-параллельной архитектурой;
- способ повышения точности каскада компактных свёрточных нейронных сетей, заключающийся в эквализации гистограммы и зеркальном отражении участков изображения, выделенных на первой стадии, перед их классификацией последующими стадиями каскада, позволяющий существенно уменьшить количество ложных срабатываний детектора практически без потери полноты обнаружения лиц;
- быстродействующие алгоритмы вычисления двумерной свёртки, позволяющие выполнять фильтрацию изображений одновременно несколькими ядрами свёртки и превосходящие по производительности соответствующие функции из ряда специализированных библиотек;

– оригинальный способ вычисления каскада классификаторов, заключающийся в асинхронном параллельном выполнении его стадий одновременно на *CPU* и *GPU* в гетерогенной вычислительной среде.

Все главы диссертации взаимосвязаны, материал изложен логично, вывод математических соотношений проведен последовательно и корректно, сделанные выводы обоснованы. Оформление иллюстративного материала и приложений способствует пониманию основных положений. Работа представляет собой самостоятельный и законченный научный труд.

Достоверность полученных выводов определяется правильным применением математического аппарата, моделированием и внедрением. Результаты работы опубликованы в 10 научных статьях, в том числе в 3 статьях в журналах и научных сборниках, рекомендованных ВАК, 5 публикациях в сборниках материалов международных и всероссийских научных конференций. Получены два свидетельства о государственной регистрации программ для ЭВМ.

Практическая значимость результатов исследований заключается в том, что разработанный детектор фронтальных лиц превосходит все ранее предложенные алгоритмы по скорости обработки кадров как на *CPU*, так и на *GPU*, и позволяет анализировать видеопоток с разрешением вплоть до *4K Ultra HD* в режиме реального времени на маломощных вычислительных устройствах.

У меня также имеются замечания по диссертационной работе:

– в представленном в диссертационной работе детекторе лиц присутствуют свертки 3×3 , 4×4 , 5×6 , даже 7×8 . В работе автор применяет полиномиальные преобразования для ускорения вычисления этих свёрток. Для сверточных ядер, превосходящих по размерам 3×3 элиза, эти формулы довольно сложны. Неясно, насколько необходимо применение "неквдратных" свёрток (5×6 , 7×8), требующих больше всего вычислений. Представляется, что можно существенно уменьшить количество вычислений, оставив в детекторе лишь свёртки малого размера - 3×3 и 2×2 , для которых полиномиальное представление (известное также под именем формул Винограда) просто выписывается и хорошо реализуется на *GPU* (A. Lavin, S. Gray, "Fast algorithms for convolutional neural networks", arXiv:1509.09308v2). Возможно, при применении свёрток меньшего размера для обеспечения требуемой точности детектирования пришлось бы увеличить глубину нейросети, но при этом можно было рассмотреть сети с остаточной связью типа ResNet;

– учитывая проведенную автором обширную оптимизацию вычисления сверток на *CPU* и *GPU*, было бы интересно реализовать предложенный каскадный нейросетевой детектор лиц с представлением параметров сети с фиксированной запятой. Эксперименты в области портирования нейросетей на встраиваемые процессоры, получившие развитие в последние два года, показывают, что при переходе от плавающей к фиксированной запятой вполне реально существенно (от 2 до 10 раз) ускорить вычисления, понизить требования к памяти (а, следовательно, снова увеличить скорость за счёт локальности данных в кэше), при этом не сильно жертвуя точностью детекции.

Отмеченные недостатки, которые, скорее, носят характер пожеланий, не влияют на общую положительную оценку диссертации, которая является законченным научно-исследовательским трудом, выполненным лично автором на высоком научном уровне.

Полученные результаты являются достоверными и обоснованными, имеют не только важное научное значение, но и доведены до уровня *практического использования*:

- алгоритмы и программные средства, разработанные в диссертационной работе, использовались при выполнении работ по гранту РФФИ № 12-08-00296 «Создание комплексных технологий распознавания объектов на изображениях ...» (2012–2014 гг.);
- в учебном процессе Института кибернетики ТПУ;
- в проекте «Разработка системы биометрической видеоаналитики ...», выполняемом по гранту «Фонда содействия развитию малых форм предприятий в научно-технической сфере» в рамках программы «У.М.Н.И.К.» («Участник молодёжного научно-инновационного конкурса») (2016–2017 гг.).

Все работы по реализации и внедрению проводились под руководством или при непосредственном участии автора.

Диссертация Калиновского И.А. является законченной научной работой, которую характеризует новизна результатов, достоверность основных положений и выводов, практическая направленность, хороший стиль изложения. По каждой главе и работе в целом сделаны четкие выводы. Автореферат соответствует основному содержанию диссертации.

Изложенные выше выводы об актуальности темы, о научной новизне исследований и содержащихся в диссертации результатах, их теоретической и практической значимости позволяют сделать общее заключение о том, что диссертационная работа Калиновского Ильи Андреевича на тему «Метод нейросетевого детектирования лиц в видеопотоке сверхвысокого разрешения», представленная на соискание учёной степени кандидата технических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей, удовлетворяет требованиям Положения о присуждении ученых степеней, предъявляемым к кандидатским диссертациям. Автор диссертации Калиновский И.А. заслуживает присуждения ему учёной степени кандидата технических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Официальный оппонент,
директор Управления перспективных мобильных технологий,
ООО «Исследовательский центр Самсунг»,
доктор физико-математических наук
(01.04.06 - акустика), профессор


08.12.2016

М.Н. Рычагов

Подпись Рычагова Михаила Николаевича заверяю,
начальник административного управления





С.В. Навасардян

ООО «Исследовательский центр Самсунг»

Адрес: 127018, г. Москва, ул. Двинцев, дом 12, корпус 1, офис 1500

Тел.: +7 (495) 797-25-00; cv.ru@samsung.com

<http://www.samsung.com/ru/aboutsamsung/samsungelectronics/careers/srr.html>