

ОТЗЫВ

официального оппонента Феоктистова Александра Геннадьевича
на диссертационную работу Валугеовой Марии Васильевны
«Разработка методов и алгоритмов построения цифровых устройств
интеллектуального анализа визуальных данных»,
представленную на соискание ученой степени кандидата технических наук по
специальности 2.3.5 – Математическое и программное обеспечение вычислительных
систем, комплексов и компьютерных сетей

Актуальность темы. Диссертационная работа Валугеовой Марии Васильевны посвящена решению задачи проектирования аппаратных ускорителей системы интеллектуального анализа изображений для повышения их производительности. Значительная требовательность искусственных нейронных сетей, особенно сверточных нейронных сетей (СНС), к вычислительным ресурсам требует качественного улучшения реализаций глубоких нейронных сетей и их технических характеристик. Одним из перспективных направлений решения поставленной задачи является аппаратная реализация компонентов нейронной сети с использованием вычислений в конечных кольцах и полях. Результаты решения поставленной задачи позволят значительно расширить создание и использование приложений для распознавания изображений, анализа речи и создания робототехнических систем.

Несмотря на то, что существуют аппаратные реализации простых нейронных сетей, в том числе осуществляющих и процесс обучения нейронной сети, разработка микроэлектронных устройств на основе архитектуры глубоких нейронных сетей с аппаратной реализацией процесса обучения является малоисследованной областью. Таким образом, тема диссертационной работы является актуальной.

Научная новизна и достоверность полученных результатов. На защиту выносятся следующие основные результаты исследования:

– математическая модель СНС с реализацией вычислений в системе остаточных классов;

– методы проектирования сверточных слоев СНС, функции активации ReLU и слоя выборки максимального элемента из окрестности, а также их программная реализация на языке описания аппаратуры VHDL,

– алгоритм проектирования аппаратной реализации СНС,

– программный комплекс для аппаратной реализации СНС для распознавания образов с использованием модулярных вычислений на FPGA.

Все результаты являются новыми. Достоверность и обоснованность применения предложенных математических моделей и методов, в частности, элементов теории модулярных вычислений в системе остаточных классов, а также разработанных алгоритмов и программ подтверждается соответствием теоретических и практических результатов, полученных в рамках имитационного моделирования и экспериментальных расчетов.

Практическая значимость. В рамках диссертации разработан программный комплекс для аппаратной реализации СНС для распознавания образов с использованием модулярных вычислений на FPGA. Имеется акт о внедрении данного программного комплекса в рамках программной платформы для обеспечения безопасности промышленных предприятий. Диссертационное исследование также поддержано грантами Совета по грантам Президента РФ, его результаты использованы при выполнении ряда научно-исследовательских проектов РФФИ и Министерства науки и высшего образования РФ.

Апробация работы. Основные результаты диссертации широко представлены на международных и всероссийских конференциях и отражены в 33 публикациях, среди которых 12 статей в научных изданиях, входящих в перечень ВАК РФ и международные системы цитирования Web of Science и Scopus. Получено 6 свидетельств о государственной регистрации программ для ЭВМ.

Структура и содержание диссертации. Диссертационная работа состоит из введения, четырех глав, заключения, списка литературных источников и двух приложений. Список используемой литературы содержит 114 источников, объем работы составляет 154 страницы. Автореферат диссертации в целом адекватно отражает структуру и содержание работы.

Во введении обоснована актуальность темы диссертации, сформулированы цель и задачи работы, выбраны объект и предмет исследования, показаны научная новизна, практическая ценность полученных результатов, приведены основные положения, выносимые на защиту.

Первая глава посвящена обсуждению математической модели СНС и сравнительному анализу методов и средств улучшения ее технических характеристик. В частности, в ней предложено использование системы остаточных классов с модулями специального вида для улучшения характеристик аппаратной реализации СНС.

Во второй главе рассмотрена задача увеличения производительности цифровых фильтров, которые применяются для реализации операции свертки в СНС. Предложен метод для проектирования устройства, реализующего операцию свертки на основе метода Винограда и системы остаточных классов с модулями специального вида. Рассмотрены случаи масок фильтров размерности 2×2 , 3×3 и 5×5 , для которых разработаны цифровые фильтры двумерной фильтрации на основе предложенного метода. Результаты анализа параметров устройств показали, что цифровые фильтры, разработанные на основе предлагаемого метода, имеют большую производительность по сравнению с рассмотренными известными аналогами.

Третья глава описывает методы реализации вычислительно сложных операций в системе остаточных классов и устройства для их выполнения. Предложены архитектуры устройств определения знака числа, сравнения чисел в системе остаточных классов, а также архитектура устройства обратного преобразования в системе остаточных классов на основе китайской теоремы об остатках. Результаты теоретического анализа и аппаратного

моделирования показали, что разработанные архитектуры устройств реализации вычислительно сложных операций в системе остаточных классов обладают меньшей задержкой и требуют меньше аппаратных ресурсов. Предложены метод проектирования устройства, реализующего функцию активации ReLU, и метод проектирования устройства, реализующего операцию выбора максимального элемента из окрестности на основе Китайской теоремы об остатках с дробными величинами.

В четвертой главе предложен алгоритм проектирования аппаратной реализации СНС с вычислениями в системе остаточных классов и архитектура СНС, разработанная по предложенному алгоритму. Весовые коэффициенты сети преобразуются к целочисленному формату. Результаты эксперимента подтвердили ранее полученные теоретические результаты. Результаты аппаратного моделирования показали, что использование метода Винограда позволяет увеличить производительность устройства в 4.3-5.56 раз, а применение предлагаемого подхода организации вычислений в системе остаточных классов увеличивает производительность устройства на 2.84-12.13% по сравнению с рассмотренными аналогами.

Заключение содержит краткое изложение основных результатов работы.

Замечания по диссертации.

1. В разделе 1.3 и далее по тексту диссертации говорится о важности балансировки модулей системы остаточных классов. В то же время явная оценка этого показателя в работе не приводится.

2. На рисунке 2.2.1 (стр. 34 диссертации) приведено z -преобразование входного сигнала. Однако пояснение его назначения в тексте диссертации отсутствует.

3. Во второй главе диссертации при описании моделирования работы цифровых двумерных фильтров в процессе обработки изображений используется набор критериев, в том числе противоречивых. Например, минимизация энергопотребления зачастую не позволяет одновременно решать задачу повышения производительности. Поэтому применение формальных методов многокритериальной оценки результатов работы фильтров возможно более ярко подчеркнуло бы преимущества предложенной архитектуры фильтра.

4. Основное внимание в разделе 4.2 диссертационной работы уделено структуре разработанного комплекса программ для аппаратной реализации СНС и назначению его основных компонентов. Между тем архитектура данного комплекса в целом и схема его функционирования нуждаются в детализации.

5. В разделе 4.3 диссертации рассматриваются вопросы, связанные с обучением СНС. К сожалению, накладные расходы на обучение сети и обучающие выборки в работе не приводятся. В этой связи не совсем ясно то, как размер и состав обучающей выборки влияют на результаты работы СНС и изменение накладных расходов на ее обучение.

6. Замечания по оформлению: в тексте диссертации отсутствуют ссылки на некоторые рисунки (см., например, рисунки 1.1.2 и 2.1.2), аббревиатура ОПСС не раскрыта,

а аббревиатура ПСС введена дважды; в автореферате на стр. 18 и 19 нарушена последовательность нумерации основных результатов и выводов по работе; в текстах диссертации и автореферата встречаются несущественные опечатки.

Следует отметить, что указанные замечания не влияют на несомненную теоретическую и практическую ценность диссертационной работы.

Заключение. Диссертационная работа Валуевой Марии Васильевны выполнена по актуальному направлению исследований на высоком научном уровне. Содержание диссертации полностью соответствует поставленным целям и задачам, обладает внутренним единством и корректно отражает последовательность выполнения основных этапов исследования. Представленная работа изложена в хорошем научном стиле, ясно и доказательно. Основные положения и выводы диссертации являются вполне обоснованными. Диссертация Валуевой Марии Васильевны является самостоятельно подготовленной, завершенной научно-квалификационной работой, в которой на основании выполненных автором исследований эффективно решается актуальная задача разработки методов проектирования аппаратных ускорителей системы интеллектуального анализа изображений для повышения их производительности.

Диссертация соответствует требованиям п. 9 Положения ВАК РФ о присуждении ученых степеней, предъявляемым к кандидатским диссертациям, а ее автор, Валуева Мария Васильевна, заслуживает присуждения ей ученой степени кандидата технических наук по специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Официальный оппонент

главный научный сотрудник лаборатории параллельных
и распределенных вычислительных систем,

ФГБУН Институт динамики систем и теории управления

имени В.М. Матросова Сибирского отделения

Российской академии наук,

доктор технических наук, доцент,

Феоктистов Александр Геннадьевич

Контактные данные организации: Федеральное государственное бюджетное учреждение науки Институт динамики систем и теории управления имени В.М. Матросова Сибирского отделения Российской академии наук (ИДСТУ СО РАН), 664033, г. Иркутск, ул. Лермонтова, д. 134, e-mail: idstu@icc.ru, тел.: +7 (3952) 42-71-00

Подпись заверяю
Нач. отдела делопроизводства
и организационного обеспечения
ИДСТУ СО РАН

Г.Б. Кононенко

21.09.2023