

УТВЕРЖДАЮ

Директор Федерального государственного
учреждения "Федеральный исследовательский
центр Институт прикладной математики им.
М.В. Келдыша Российской академии наук"
член. корр. РАН

А.И. Аптекарев
10 мая 2023 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

Федерального государственного учреждения «Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук» на диссертационную работу **ШЕШКУСА Александра Владимировича** на тему «Использование преобразования Хафа в качестве слоя нейронной сети», представленную на соискание ученой степени кандидата технических наук по специальности 1.2.2 — «Математическое моделирование, численные методы и комплексы программ» (технические науки)

Актуальность темы диссертации

Диссертация Шешкуса Александра Владимировича посвящена важной и актуальной проблеме — созданию современных автоматизированных систем подготовки и поддержки принятия решений.

Искусственные нейронные сети — быстро развивающийся инструмент, применяемый для решения широкого спектра задач, в том числе задач машинного зрения и распознавания образов. Относясь к классу так называемых обучаемых машин, нейросетевые модели являются универсальными аппроксиматорами и позволяют на практике реализовать идею «обучения на примерах», при которых инстанцирование модели осуществляется путем стохастической оптимизации в пространстве ее параметров. Известно, что решение такой оптимизационной задачи можно значительно облегчить за счет выбора правильной архитектуры нейронной сети. Так, в задачах обработки изображений можно исходить из предположений о пространственной ограниченности и небольшом размере рецептивного поля нейронов, а также

перестановочности операций обработки со сдвигом изображения. В частности, поэтому сверточные нейросетевые архитектуры, реализующие эти ограничения, существенно эффективнее в обработке изображений, чем классические полносвязные.

В то же время, в задачах машинного зрения, таких как определение собственного положения и измерение углов и расстояний по изображению сцены, рецептивное поле нейронов последнего слоя нейросети должно включать существенную площадь изображения. Это делает актуальной разработку новых нейросетевых архитектур с большим рецептивным полем и хорошими показателями обучаемости. Решению этой задачи посвящена диссертационная работа Шешкуса А.В., в которой исследованы нейросетевые архитектуры, основанные на включении необучаемых слоев прямого и транспонированного интегрального преобразования Хафа между сверточными слоями сети. Данный подход позволяет обеспечить требуемую конфигурацию рецептивных полей нейронов в слоях, идущих после Хаф-слоя, но, в то же время, сохранить хорошую обучаемость при решении прикладных задач машинного зрения.

Научная новизна

Автором получены новые научные результаты, определяющие научную новизну с обоснованными отличительными признаками, а именно:

1. В работе впервые проведено исследование трансформации рецептивного поля нейронной сети при добавлении в нее слоев прямого и транспонированного быстрого преобразования Хафа. В работе показано, что для того, чтобы нейронная сеть была способна решать задачи компьютерного зрения, ее рецептивное поле должно обладать достаточным размером.
2. В работе предложена новая архитектура сетей со слоями прямого преобразования Хафа, а также метод ее использования для решения задачи поиска внешней точки схода.

3. В работе впервые предложена архитектура нейронной сети с использованием транспонированного преобразования Хафа для решения задач компьютерного зрения и показано, что такая архитектура способна решать задачи поиска внутренних точек схода и семантической сегментации четырехугольных объектов с превосходным качеством, при этом обладая на два порядка меньшим количеством обучаемых параметров.
4. Выявлен новый эффект при применении известной технологии переноса знаний для адаптации к домену (domain adaptation): перенос знаний успешно происходит, но для высокоуровневых, а не для низкоуровневых признаков.

Общая характеристика работы

Диссертация Шешкуса А.В. состоит из введения, четырех глав, заключения и списка литературы из 95 пунктов. Общий объем диссертации составляет 123 страницы.

Во введении Шешкус А.В. обосновывает актуальность и научную новизну диссертационной работы, ставит цель и задачи диссертационного исследования и приводит положения, выносимые на защиту.

В первой главе на основе анализа развития нейросетевых технологий и инженерных алгоритмов для решения задач машинного зрения дает обоснование перспективности идеи совмещения интегрального оператора Радона и сверточной нейронной сети в единую сущность, для чего исследуется поведение нейронных сетей, включающих в свою архитектуру слои быстрого преобразования Хафа в явном виде, в практических задачах машинного зрения.

Во второй главе Шешкус А.В. описывает теоретические основы, на которых базируется проведенное в диссертации исследование. Подробно описывается трансформация рецептивного поля при применении преобразования Хафа к картам признаков и алгоритм быстрого преобразования Хафа, основанный на аппроксимации пикселизации прямых диадическими паттернами. Автор доказывает, что транспонированное быстрое

преобразование Хафа можно вычислять тем же самым алгоритмом, если правильным образом произвести отражения входных и выходных изображений. Это позволяет не только распространять градиент алгоритмом с хорошей асимптотической сложностью, но и строить сети, которые используют транспонированный оператор быстрого преобразования Хафа в явном виде.

В третьей главе Шешкус А.В. дает описание разработанных нейросетевых архитектур, использующих слои прямого и транспонированного быстрого преобразования Хафа, и исследует их свойства и способы применения. Важным аспектом становится соотнесение координат на карте признаков, обрабатываемых нейросетью, с координатами в трехмерном пространстве анализируемой сцены.

Четвертая глава посвящена практическим применениям предложенных в главе 3 архитектур и описанию проведенного внедрения их в индустриальное программное обеспечение компании “Смарт Энджинс Сервис”.

В заключении приведено общее описание и основные результаты работы Шешкуса А.В.

Достоверность и обоснованность результатов работы

В работе используются методы цифровой обработки изображений, методы стохастической непрерывной оптимизации, методы статистического анализа. Все результаты обосновываются формально-дедуктивно или вычислительным экспериментом.

Достоверность полученных результатов подтверждается соответствием теоретических и экспериментальных результатов, продемонстрированных в работе, успешной апробацией результатов и внедрением в коммерческие системы распознавания документов.

Основные результаты по теме диссертации опубликованы в 8 печатных изданиях (1 статья в журнале, рекомендованном ВАК, 7 – в изданиях, индексируемых в Web of Science и Scopus), докладывались и обсуждались на конференциях International Conference on Machine Vision (ICMV) в 2016-м, 2017-м и 2019-м годах, International Conference on Document Analysis and

Recognition (ICDAR) в 2019-м году, International Conference on Image Processing” (ICIP) в 2020-м году.

Значимость для науки и практики

Значимость работы А.В. Шешкуса для науки заключается в выполненном исследовании подходов и методов обучения нейросетевых архитектур, включающих необучаемые слои интегральных преобразований, а именно — преобразований Хафа.

Выполненная работа имеет большую практическую значимость. Это подтверждается тем, что методы обучения нейронных сетей с использованием исследованных в работе слоев преобразования Хафа внедрены в отечественную систему обучения нейронных сетей компании ООО «Смарт Энджинс Сервис». Обученные нейросетевые модели применяются для обработки изображений и распознавания документов в информационной инфраструктуре ряда государственных структур Российской Федерации, а также коммерческих организаций

Автором получено свидетельство о регистрации программы для электронных вычислительных машин (ЭВМ) и патент США.

Рекомендации по использованию результатов исследования

Результаты диссертационной работы могут быть применены при разработке новых нейросетевых алгоритмов машинного зрения, основанных на анализе линейных признаков на изображении, а также для разработки новых нейросетевых архитектур.

Диссертационная работа и ее результаты могут быть применены в педагогическом процессе для изучения дисциплин «Обработка изображений», «Техническое зрение» и «Распознавание образов» в Московском государственном университете им. М.В. Ломоносова (МГУ), Московском физико-техническом институте (национальном исследовательском университете), Национальном исследовательском университете «Высшая школа экономики», Сколковском институте науки и технологий и т.д. Результаты диссертации могут быть использованы в институтах Российской

академии наук (Институте прикладной математики им. М.В. Келдыша РАН, Институте проблем передачи информации, Федеральном исследовательском центре «Информатика и управление») и Научно-исследовательском центре «Курчатовский институт».

Недостатки и замечания по диссертационной работе

В таблице 1 диссертационной работы приведены показатели точности различных нейросетевых архитектур при наличии и отсутствии результатов преобразования Хафа, однако не приведена методика вычисления представленных в таблице значений. Можно предположить, о чем идет речь, исходя из того, что оценивается качество классификации, однако для ясности повествования стоило бы дать развернутое описание.

В тексте диссертации не разъяснены обозначения k_i , R_i , s_i , a_i , f_i , использованные в формуле (2.1).

Матрица оператора отражения C начинает использоваться в формулах до того, как в тексте вводятся обозначения данного оператора и его матрицы.

В работе присутствует некоторая путаница с использованием терминов внутренняя, внешняя собственная и несобственная точка схода.

Высказанные замечания по работе не носят принципиального характера и не умаляют достоинств диссертационного исследования.

Аннотация адекватно отражает содержание диссертации. Основные научные результаты достаточно полно отражены в публикациях, в том числе в журналах перечня ВАК и в изданиях, индексируемых в Web of Science и Scopus.

Заключение

Диссертация Шешкуса А.В. является завершённой научно-квалификационной работой, посвященной исследованию свойств нейронных сетей, включающих слои прямого и транспонированного преобразований Хафа при решении прикладных задач машинного зрения.

Работа соответствует требованиям раздела II Постановления Правительства РФ от 24.09.2013 №842 «О порядке присуждения учёных

степеней» (ред. от 26.09.2022), предъявляемых к кандидатским диссертациям, а также пунктам 2 «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий», 3 «Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента», 7 «Качественные или аналитические методы исследования математических моделей (технические науки)» паспорта научной специальности 1.2.2 «Математическое моделирование, численные методы и комплексы программ» (технические науки), а её автор, **Шешкус Александр Владимирович**, заслуживает присуждения учёной степени кандидата технических наук по данной специальности.

Зав. отделом, главный научный сотрудник
ИПМ им. М.В. Келдыша РАН
доктор физико-математических наук

Орлов Юрий Николаевич

ведущий научный сотрудник
ИПМ им. М.В. Келдыша РАН
доктор технических наук,

Баканова Нина Борисовна

Подписи Орлова Ю.Н. и Бакановой Н.Б. заверяю
Ученый секретарь ИПМ им. М.В. Келдыша РАН

Давыдов А.А.

Сведения об организации

Полное наименование: Федерального государственного учреждения "Федеральный исследовательский центр Институт прикладной математики им. М.В. Келдыша Российской академии наук"

Почтовый индекс и адрес организации: 125047, Москва, Миусская пл., д.4

Телефон +7 499 978-13-14

Адрес электронной почты office@keldysh.ru

Сайт <https://www.keldysh.ru/>

