

ОТЗЫВ ЧЛЕНА ДИССЕРТАЦИОННОГО СОВЕТА
на диссертацию Самарина Алексея Владимировича
«Комбинированные нейросетевые модели для классификации специфичных изображений», представленную на соискание ученой степени кандидата физико-математических наук

Актуальность темы диссертации

Математические методы классификации изображений начали развиваться еще в середине прошлого века, но максимального прогресса достигли в последнее десятилетие в связи с развитием глубоких нейронных сетей. Несмотря на высокую эффективность современных подходов к распознаванию типовых изображений, существует ряд задач, для которых качество, обеспечиваемое универсальными нейросетевыми архитектурами, оказывается недостаточным. Одним из таких классов является рассмотренная в диссертации задача классификации специфичных изображений, которые содержат семантически значимую текстовую информацию и получены при вариативных условиях съемки: различные ракурсы и положение объектов, непостоянные условия освещения, наличие на изображениях различных артефактов – бликов, затенений и т.п. Таким образом, диссертационная работа Самарина Алексея Владимировича, направленная на повышение точности решения таких практически важных задач за счет комбинирования различных нейросетевых архитектур в единый ансамбль, является, безусловно, актуальной.

Научная новизна полученных результатов

На мой взгляд, в диссертации получены следующие результаты, обладающие, **научной новизной**.

- Предложена новая нейросетевая ансамблевая архитектура VCA (VGG Combined with Autoencoders), в которой характерные признаки входного изображения формируются с помощью как традиционной сверточной нейронной сети, так и с применением специально обученного автокодировщика.

- Разработана новая нейросетевая ансамблевая архитектура CCIT (Combined Classifier of Images with Text), состоящая из двух параллельных компонент – сверточной нейронной сети и компоненты детектирования и распознавания текста (OCR, Optical Character Recognition). Для последней

компоненты представлено три модификации на основе идеи многоагентного взаимодействия, отличающихся способом извлечения дескрипторов для областей изображения, содержащих текст.

Степень обоснованности и достоверности научных положений, выводов и рекомендаций, сформулированных в диссертации

Достоверность научных результатов диссертационной работы подтверждается адекватной постановкой задачи, формальными построениями предложенных комбинированных нейросетевых архитектур, а также результатами практических экспериментов и их апробацией.

Разработанные соискателем методы и алгоритмы основываются на корректном применении современных подходов к анализу изображений, а также нейросетевых архитектур. Выбор инструментария для решения поставленной задачи представляется логичным и обоснованным. В диссертации проведен достаточно подробный обзор современных методов классификации изображений. Цитируемые утверждения подтверждены ссылками на современные источники. Научные положения, выводы и рекомендации являются **обоснованными**.

Теоретическая и практическая значимость полученных результатов

Теоретическая значимость результатов диссертации заключается в создании новых ансамблевых нейросетевых архитектур на основе предобработки изображения с помощью автокодировщика и извлечения информации из детектированных на изображении областей с текстовыми данными. Для всех разработанных модификаций дескрипторов текстовых областей доказаны утверждения об их линейной вычислительной сложности относительно размера изображений.

Практическая значимость работы состоит в том, что предложенные архитектуры могут применяться для повышения качества систем распознавания сложных изображений, полученных в нестандартных условиях. Отдельно стоит отметить, что часть результатов диссертационной работы внедрены в систему восстановления доступа и систему поиска изображений социальной сети “ВКонтакте”.

Замечания по диссертации

1. Широко известно, что ансамблевые методы позволяют добиться более высокой точности классификации за счет повышения вычислительной сложности. Автору в экспериментах следовало исследовать не только значения метрик точности (F-мера, Precision/Recall), но и среднее время классификации изображений.
2. Из текста диссертации не понятно, почему в разработанной модели VCA применяются достаточно старые и очень затратные по памяти (сотни мегабайт) архитектуры VGG, хотя даже из таблицы 2 видно, что часть модели с предварительной обработкой изображения с помощью автокодировщиков (AE) для более современных архитектур оказывается более точной.
3. Имеются замечания по тексту диссертации: присутствуют опечатки, например, «архитекутры» (название раздела 1.1), «данной диссертационной» (стр. 7), Некоторые математические обозначения не пояснены., например, n в формуле L_{cr} в разделе 2.4 (здесь n – число, а в разделе 2.3 n – это отображение тензора в набор чисел), I_{in} в разделе 3.2 (в чем отличие от I_{so} ?), r там же в формуле $I_{out}[x,y]$, q в формуле резкости и т.п.
4. В разделе 2.3 не объяснена процедура создания обучающего множества для автокодировщика. Например, почему для множества тензоров P выполняется последовательное применение вертикального и горизонтального оператора Собеля, хотя обычно они применяются параллельно с последующей агрегацией результатов.
5. Стоило продемонстрировать, возможно ли применять архитектуру SSIT (рис.3.3) не с предлагаемыми дескрипторами типа А, В и С, а с признаками, извлеченными специализированными нейросетевыми моделями.

Заключение


Считаю, что отмеченные недостатки не носят принципиального характера.

Диссертация Самарина Алексея Владимировича представляет собой законченное научно-квалификационное исследование, решающее актуальную задачу классификации специфичных изображений с использованием комбинированных нейросетевых архитектур. Судя по всему, данное исследование выполнено автором самостоятельно, на высоком научном уровне. Текст диссертации написан грамотным научным языком. Результаты работы представляются вполне достоверными. Основные научные результаты опубликованы в журналах, индексируемых в международной базе научного цитирования Scopus.

Таким образом представленная к защите диссертационная работа заслуживает положительной оценки, соответствует основным требованиям, установленным Приказом от 19.11.2021 № 11181/1 «О порядке присуждения ученых степеней в Санкт-Петербургском государственном университете», а сам соискатель Самарин Алексей Владимирович заслуживает присуждения учёной степени кандидата физико-математических наук по специальности 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей. Нарушения пунктов 9 и 11 указанного Порядка в диссертации не обнаружены.

Доктор технических наук, доцент, профессор кафедры информационных систем и технологий, ведущий научный сотрудник лаборатории алгоритмов и технологий анализа сетевых структур Нижегородского филиала федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский университет «Высшая школа экономики»

Савченко Андрей Владимирович


(подпись)

14.02.2024

(дата)

Адрес: 603155, г. Нижний Новгород, ул. Большая Печерская, д. 25/12

Телефон: +7 (831) 416-95-40

Email: avsavchenko@hse.ru

URL: <https://www.hse.ru/staff/avsavchenko>



Подпись
Заведующая Начальником
отдела кадров
Н.А. Ермолиной

