

ОТЗЫВ

научного руководителя на диссертационную работу

Самарина Алексея Владимировича на тему

**«Комбинированные нейросетевые модели для классификации
специфичных изображений»,**

представленную на соискание ученой степени кандидата физико-математических наук по специальности 2.3.5. Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

На сегодняшний день в компьютерном зрении особую актуальность имеют задачи классификации изображений, содержащих визуальное представление текстовой информации. Сложность таких задач связана с неспособностью современных механизмов оптического распознавания текста устойчиво распознавать текст, представленный на изображении при варьировании шрифтов, стилей оформления, условий съемки и прочих условий.

Кроме этого, имеются и другие факторы, осложняющие классификацию различных типов изображений. Например, современные графические редакторы предоставляют пользователям широкий спектр возможностей по модификации исходного изображения. Обработка ситуаций, связанных с различными искусственными апостериорными модификациями исходных изображений, усложняет задачу классификации изображений с текстом.

Следует отметить, что на данный момент не существует универсальных методов и инструментов для решения задач классификации изображений с элементами текста в сложных условиях и, видимо, соответствующие ограничения носят принципиальный характер. Однако, некоторые частные задачи этого класса, имеющие практическое значение, могут быть успешно решены.

Диссертационная работа Алексея Владимировича Самарина посвящена созданию математических моделей для классификации изображений в сложных условиях, что обуславливает несомненную актуальность выбранной темы.

В рамках данной работы предложены новые комбинированные нейросетевые архитектуры, а также на их основе реализован ряд классификаторов для специфичных видов изображений. Для извлечения дополнительной семантики из фрагментов изображений с текстом разработан ряд специальных дескрипторов, которые повышают эффективность комбинированных нейросетевых архитектур.

Теоретическая ценность диссертационной работы заключается в создании математических моделей нейросетевых архитектур, в рамках которых удалось формализовать (учесть) ряд условий, влияющих на распознавание и качество классификации изображений с текстом. Это сделано впервые и составляет несомненную **научную новизну** работы.

Разработанные в рамках диссертационной работы классификаторы изображений интегрированы в систему автоматизированного восстановления доступа к учётным записям и ряд внутренних алгоритмов обработки изображений социальной сети «ВКонтакте», что подтверждено соответствующими актами внедрения и составляет **практическую ценность** данной работы.

В диссертации представлены следующие основные результаты, перечисленные ниже.

- Рассмотрены проблемы обработки и распознавания изображений, содержащих фрагменты с текстом. Особое внимание уделено двум типам подходов к классификации изображений таких. Первая группа подходов явно распознаёт текст на изображении и выполняет последующий анализ текстовой информации, что актуально в случае, когда имеется непосредственная зависимость результата классификации от семантики текста в изображении. Вторая группа неявно использует текстовые фрагменты изображения, не выполняя их оптического распознавания.

Подобные подходы актуальны при решении задач классификации изображений, где семантика запечатленного текста не имеет первостепенного значения при выборе того или иного класса.

- Представлена архитектура комбинированной нейросетевой модели классификации изображений, не выполняющая явное распознавание текстовой информации. Предложенная архитектура была использована для решения задачи определения пригодности изображения документа для подтверждения личности при условии варьирования условий съемки.
- Предложена архитектура комбинированного классификатора с явным распознаванием текстовой информации. Для улучшения качества финальных результатов на основе этой архитектуры был реализован ряд дополнительных моделей. Эффективность данных подходов изучена при решении задачи классификации фотографий фасадов коммерческих строений по типу оказываемых услуг.
- Выполнено экспериментальное исследование показавшее, что комбинированные нейросетевые классификаторы, реализующие предложенные в диссертации архитектуры, демонстрируют более качественные результаты по сравнению с общими методами классификации изображений.

Следует особенно подчеркнуть глубокий профессионализм Алексея Владимировича в данной тематике исследований, что подтверждается, в частности, значительным количеством публикаций (в научных изданиях, включенных в перечень ВАК РФ, индексируемых РИНЦ и международными наукометрическими базами данных Scopus и WoS), личным знакомством со многими ведущими специалистами в мире по этой тематике, а также представлением результатов исследований на ведущих конференциях в области, в частности, на 25-ой международной конференции “International Conference on Pattern Recognition (ICPR)” (10-15 января 2021 г., Милан, Италия), на 8-ой

международной конференции “Image Mining. Theory and Applications (IMTA)” (21–25 августа 2022 г., Монреаль Квебек, Канада).

Достоверность и обоснованность результатов подтверждается использованием строгих математических методов, экспериментами и апробацией основных положений диссертации на международных конференциях, а также путём интеграции в промышленные системы.

Диссертация Алексея Владимировича на тему «Комбинированные нейросетевые модели для классификации специфичных изображений» соответствует требованиям, установленным Приказом от 01.09.2016 №6821/1 «О порядке присуждения научных степеней в Санкт-Петербургском государственном университете», а сам соискатель, Алексей Владимирович Самарин, заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 2.3.5. Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Научный руководитель диссертации
Доктор технических наук,
кандидат физико-математических наук,
доцент кафедры системного программирования СПбГУ
Кознов Д.В.

25 мая 2023 год

Подпись руки	<i>Кознова</i>
УДОСТОВЕРЯЮ	
Специалист по кадровой работе	<i>Л.И. Нестерова</i>
« 25 » 05	2023 г.

Нестерова Л.И.

