



СПбГЭТУ «ЛЭТИ»  
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

МИНОБРНАУКИ РОССИИ  
федеральное государственное автономное образовательное учреждение высшего образования  
«Санкт-Петербургский государственный электротехнический университет  
«ЛЭТИ» им. В.И. Ульянова (Ленина)»  
(СПбГЭТУ «ЛЭТИ»)

Проректор по научной и  
инновационной деятельности



СПбГЭТУ «ЛЭТИ»

д.т.н., доцент.

А. А. Семенов

06 февраля 2024 г.

## ОТЗЫВ

ведущей организации на диссертацию Самарина Алексея Владимировича  
на тему **«Комбинированные нейросетевые модели для классификации  
специфичных изображений»**,

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

### Актуальность темы

Одной из основных задач, решаемых в области компьютерного зрения, является классификация изображений. В настоящее время за счет развития методов и технологий глубокого обучения в данной области достигнут существенный прогресс. Можно говорить о том, что современная теоретическая проработка соответствующих методов и алгоритмов позволяет в общем случае получать удовлетворительное решение для задачи классификации изображений. При этом существуют частные случаи, для которых общепринятые подходы не позволяют получить удовлетворительный результат. К этой группе относятся изображения, полученные в сложных условиях наблюдения, снимки с низким контрастом, бликами и пр. Проблему в данном контексте представляет в том числе классификация изображений, содержащих текстовую информацию.

33-06-216 от 28.02.2024

Работа А.В. Самарина посвящена созданию нейросетевых моделей для классификации изображений, полученных в сложных условиях съемки. Тема диссертационного исследования является актуальной и востребованной на современном этапе.

### **Структура и содержание работы**

Работа состоит из введения, трех глав, заключения и одного приложения.

Объем работы составляет 92 страницы, включает 37 рисунков и 5 таблиц.

Во введении обосновывается актуальность темы работы, сформулированы задачи исследования, приведены краткие сведения о достигнутых результатах, апробации, соответствия специальности.

В первой главе проведен аналитический обзор предметной области, рассмотрены существующие прототипы решений.

Вторая глава посвящена синтезу комбинированной нейросетевой архитектуры для распознавания изображений без учета семантики содержащегося в них текста. Описаны принципы построения архитектуры. Приведены обоснования выбранных решений, проведено моделирование и экспериментальное исследование полученных характеристик.

В третьей главе предложена комбинированная нейросетевая архитектура для классификации изображений с учетом текстовой информации. Детализирована постановка задачи, предложен алгоритм предварительной обработки изображений, предложены дескрипторы нового типа. Описаны проведенные эксперименты и изложены полученные результаты исследования характеристик полученного решения.

В заключении приведены основные полученные результаты и сделаны выводы о возможном направлении дальнейшей работы.

### **Основные результаты работы**

1. Предложена комбинированная нейросетевая архитектура для распознавания изображений без учета семантики содержащегося в них текста. Указанная архитектура была успешно апробирована и используется на сегодняшний день ООО «В Контакте» для определения пригодности снимков, полученных в сложных условиях съемки, для анализа при прохождении процедуры подтверждения личности. Полученный таким образом классификатор продемонстрировал высокие значения в точности классификации, а также

значительный прирост в эффективности по сравнению с прототипами. Значение Error Reduction при апробации архитектуры составило более 43% в сравнение с ближайшими аналогами.

2. Предложена комбинированная нейросетевая архитектура для классификации изображений с учетом текстовой информации. Архитектура позволяет получать точность классификации в среднем на 15% выше по сравнению с аналогами.
3. Разработан ряд новых дескрипторов для извлечения из текста характеристик стиля и деталей визуального оформления, что существенно улучшило качество распознавания изображений. Применение созданных в рамках диссертационной работы специальных дескрипторов такого рода позволило повысить значение метрик качества классификации (в среднем на 15 процентов по сравнению с использованием других векторизованных представлений в рассмотренном контексте).

### **Научная новизна полученных результатов**

1. В рамках диссертационного исследования решена задача классификации изображений с ограничениями на качество съёмки, а также способами обработки текстовой информации, включенной в изображения.
2. Предложена новая нейросетевая архитектура для случая, когда не выполняется без учета семантики содержащейся в них текстовой информации. Новая архитектура использует совмещение нескольких дескрипторов изображений — промежуточных представлений глубоких нейросетевых кодировщиков, работающих с исходным изображением, и результаты работы глубоких кодировщиков на изображениях, получаемых из исходных путем специальных алгоритмов предобработки
3. Предложена новая нейросетевая архитектура классификации изображений с учетом семантики содержащейся в них текстовой информации. Архитектура основывается на специальной комбинаторной схеме, использующей дескрипторы текстовых областей изображений.

### **Достоверность полученных научных положений, результатов и выводов**

Положения, выносимые на защиту, в достаточной степени подтверждены результатами моделирования и экспериментальных сравнительных исследований с аналогами.

Результаты диссертационного исследования прошли апробацию в ряде отечественных и зарубежных тематических конференций. Основные результаты

диссертации изложены в рецензируемых изданиях, входящих в перечень ВАК и индексируемых в международных базах.

### **Практическая значимость работы**

Реализованные в рамках данной диссертационной работы нейросетевые модели классификаторов интегрированы в систему восстановления доступа к учетным записям социальной сети "ВКонтакте", в составе автоматического фильтра для отклонения заявок, непригодных для восстановления доступа. Также произведена интеграция разработанных моделей классификаторов в систему поиска изображений для удаления фотографий удостоверений личности из поискового индекса в целях предотвращения распространения персональных данных пользователей. Нейросетевые классификаторы изображений с явной стадией изолированного анализа регионов изображений внедрены во внутренние алгоритмы социальной сети "ВКонтакте" и выполняют задачи для классификации рекламных объявлений и составления векторизованных семантических представлений для изображений, визуальный ряд которых содержит графические элементы текста.

### **Замечания по диссертационной работе**

Диссертация выполнена на актуальную тему, посвященную классификации изображений, полученных в сложных условиях съемки. В качестве замечаний следует отметить следующее.

1. В первой главе нет в явном виде завершающей части проведенного аналитического обзора. Не проведена постановка задач исследования, которые детально описаны только в дальнейшем тексте работы.
2. В исследовании рассматривается работа со снимками, полученными в сложных условиях съемки. При этом не конкретизированы численные оценки характеристик изображений, входящих в выборки: ракурс, разрешение, отношение сигнал/шум и прочее.
3. Не заданы ограничения на применение полученных решений. До каких пределов ухудшения качества изображения, изменения ракурса и пр. предложенные решения будут обеспечивать декларируемые характеристики.
4. В итоговых выводах не приведены интегральные численные оценки, характеризующие преимущества предложенных моделей по сравнению с аналогами. При этом необходимые данные для этих оценок были получены в ходе моделирования и экспериментов.

Указанные недостатки не влияют в значительной мере на степень обоснованности полученных результатов и не отменяют положительную оценку диссертационного исследования.

### **Заключение**

Диссертация Самарина Алексея Владимировича на тему «Комбинированные нейросетевые модели для классификации специфичных изображений» соответствует требованиям, установленным Приказом от 19.11.2021 № 11181/1 «О порядке присуждения научных степеней в Санкт-Петербургском государственном университете». Соискатель, Самарин Алексей Владимирович, заслуживает присуждения ученой 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.

Доклад по диссертационной работе заслушан и одобрен на открытом заседании кафедры телевидения и видеотехники СПбГЭТУ «ЛЭТИ» (протокол №05-23/24 от 07.02.2024 г.).



Наталия Александровна Обухова,

д.т.н., доцент,

декан факультета радиотехники и телекоммуникаций

СПбГЭТУ «ЛЭТИ»