

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.932

Царикович
Антон Александрович

Программное средство стилизации фотографий под работы известных
художников на базе нейронной сети

АВТОРЕФЕРАТ

на соискание магистра информатики и вычислительной техники
по специальности 1-40 81 02 Технологии виртуализации и облачных
вычислений

Научный руководитель
Фролов Игорь Иванович
к.т.н., доцент

Минск 2019

КРАТКОЕ ВВЕДЕНИЕ

Развитие области нейронных сетей, цифровой обработки изображений позволило автоматизировать и ускорить многие рутинные задачи человека. Доступность довольно мощного оборудования, а также возможность выполнения кода на дискретных графических адаптерах повысило скорость развития и популярность вышеописанных направлений.

Каждая картина, рисунок известного художника имеет свой уникальный и неповторимый стиль. Воссоздать такой стиль на другом изображении требует большого труда и времени мастера. Такая работа стоила больших средств.

В связи с развитием сверточных нейронных сетей стало возможным выделение содержания и стиля изображений. Суть этого подхода состоит в том, что при помощи признаков, которые можно получить при помощи сверточной нейронной сети можно выделить содержательную и стилевую часть изображения. Это позволяет создать новое изображение, которое будет содержать стиль и содержание двух разных изображений.

В данной работе рассматривается использование технологий нейронных сетей и цифровой обработки изображений в контексте редактирования и изменения входного изображения.

Решение такой проблемы впервые было предложено совсем недавно, в 2015 году, в работе «Image Style Transfer Using Convolutional Neural Networks». Сейчас можно найти определенное количество вариаций данной идеи и алгоритма. Однако задачей данных алгоритмов является одно: перенос стиля одного изображения на другое.

В связи с этим можно сделать вывод, что данное направление использования машинного обучения находится еще на стадии развития, хотя, в вышеупомянутой работе был заложен фундамент для решения данной задачи.

Разрабатываемое решение будет актуально как для рынка приложений, так и в плане исследования и разработки новых алгоритмов обработки изображений и машинного обучения. Разработка и внедрение новых алгоритмов систем технического зрения и нейронных сетей в данный момент находится на пике своей популярности. Следовательно данная тема магистерской диссертации и разрабатываемое приложение будут актуальны с точки зрения рынка и технологий еще долгое время.

Целью данной работы является разработка приложения, которое будет выполнять стилизацию изображений под работы известных художников.

Задачами данного исследования являются:

- исследовать и применить существующие алгоритмы технического зрения, необходимые для разработки приложения;
- исследовать и применить существующие архитектуры нейронных сетей в проекте;
- разработать системы алгоритмов обработки изображений;
- разработать приложение.

Результатом данной работы является приложение, которое будет выполнять стилизацию любой фотографии пользователя под одну из предложенных работ известных художников. В качестве дополнительного функционала будет возможность выполнить стилизацию под предзагруженное пользователем изображение.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Данная работа связана с исследованиями в области обработки изображений и машинного обучения. Основным направлением исследований в данной работе является тема переноса стиля между изображениями.

Конечной целью работы является разработка приложения, которое будет выполнять стилизацию изображений под работы известных художников.

В связи с вышесказанным можно сформировать следующие задачи исследования:

1. Исследовать и применить существующие алгоритмы технического зрения, необходимые для разработки приложения.

2. Исследовать и применить необходимые существующие архитектуры нейронных сетей в проекте.

3. Разработать системы алгоритмов обработки изображений.

4. Разработать приложение.

5. Провести экспериментальные исследования над разработанным алгоритмом, с целью улучшения качества и скорости работы приложения по переносу стиля между изображениями.

Объектом исследования является перенос стиля между изображениями.

Предметами исследования являются нейронные сети, алгоритмы обработки изображений, оптимизация сверточных нейронных сетей.

Основной гипотезой, положенной в основу диссертационной работы, является возможность фотореалистичного переноса стиля между изображениями. Особенностью переноса стиля между изображениями является использование актуальных технологий, алгоритмов и архитектур нейронных сетей. Данные решения появились совсем недавно и постоянно улучшаются с точки зрения как скорости работы, так и качества генерируемого изображения. Данная разработка, описанная в магистерской диссертации призвана улучшить хотя бы по одному параметру готовые решения.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

На данный момент машинное обучение и его использование для обработки изображений является важным направлением исследований как для науки, так и рынка технологий. В последние лет пять очень часто можно слышать о покупках крупными корпорациями стартапов, направления

которых связано с обработкой изображений, видео на основе нейронных сетей. Из вышесказанного можно сделать вывод: направление данного научного исследования очень востребовано на текущем рынке интернет-гигантов.

В проводимом исследовании полностью описаны и обоснованы настройки входов, слоев нейронной сети. Обоснована оптимизация с использованием L-BFGS алгоритма.

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя И.И. Фролова, заключается в полном сопровождении и консультации при разработке проекта и написании текста данной магистерской диссертации.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Диссертация состоит из общей характеристики работы, введения, четырех глав, заключения, списка использованных источников, списка публикаций автора и приложений, одного графического приложения. В первой главе описаны основные программные аналоги проекта. Также описана история развития идеи, алгоритмы, существующие на данный момент. Описаны алгоритмы параллельного исследования проблемы магистерской диссертации, которые так или иначе повлияли на разрабатываемый проект. Вторая глава посвящена теоретической части разрабатываемого проекта. В ней представлены теоретические выкладки, которые необходимы для понимания читателю всех аспектов алгоритмической и теоретической части проекта. Описаны важные функции, которые являются основополагающими для разрабатываемого проекта, нейронные сети, а также алгоритмы. Третья глава посвящена разработке программного проекта. В ней описаны и обоснованы все используемые программные методы и решения, которые использовались для реализации алгоритма переноса стиля изображения. В четвертой главе представлены результаты экспериментов, обоснование их результатов и выводы, которые были получены.

Общий объем работы составляет 61 страницу, из которых основного текста – 50 страниц, список использованных источников из 31 наименования на 3 страницах, 1 лист формата А3 и 1 приложение на 8 страницах.

ЗАКЛЮЧЕНИЕ

Результатом данной магистерской работы стало приложение, позволяющее переносить стилизацию картин известных художников на любые изображения пользователей.

Поставленные цели были достигнуты и задачи выполнены в полном объеме.

В процессе рассмотрения аналогов, были найдены очень функциональные и проработанные приложения, которые и стали отправной точкой исследований данной магистерской диссертации. Также была подробно рассмотрена решаемая проблема с точки зрения зарождения и развития алгоритмов. Было уделено внимание работам, которые решают данную проблему отличным от разработанного путем. Однако, было выяснено, что вклад данных исследований в развитие решения проблемы очень велик. Рассмотренные аналоги и научные работы были досконально исследованы и изучены.

Разработанное решение основано на общеизвестных алгоритмах. Работоспособность приложения полностью соответствует поставленным задачам. Сильными сторонами разработанного решения можно назвать:

- скорость работы. Даже на видеокарте средней мощности алгоритм показывает неплохую скорость работы;

- качество результата. В сравнении с описанными алгоритмами в главе 1, которые могли работать на мобильных видеокартах в режиме реального времени, разработанный алгоритм может генерировать изображения высокого качества (вплоть до HD), при условии наличия мощного GPU. Однако такой подход исключает режим работы в реальном времени;

- гибкость. Разработанное приложение может применять стиль к абсолютно любому изображению, и, что самое важное, извлекать стиль из любого загруженного изображения. Приложение перед генерацией изображения можно настроить на сохранение содержания, либо на большую степень переноса стиля;

- простота. Приложение написано на одном из самых популярных языков в мире. Также были использованы актуальные библиотеки и фреймворки для машинного обучения, которые стабильно получают обновления в виде новых функций, а также оптимизацию. Следовательно, со временем приложение может работать еще быстрее.

Любой проект, решение имеет свои слабые стороны. Разработанное решение не является исключением:

- отсутствие предобработки и постобработки. Некоторые современные аналоги имеют такой функционал. В разработанном приложении он

отсутствует, в виду отсутствия первоначального требования такого функционала к приложению.

– отказоустойчивость. Приложение не покрыто тестами и не проходило любое другое серьезное тестирование. В будущем это может плохо сказаться на отказоустойчивости приложения.

В перспективе развития данного приложения можно выделить несколько вариантов:

1. Функциональный вариант. Добавление и улучшение существующего функционала.

2. Коммерческий вариант развития. Интеграция с существующими приложениями либо продажа самостоятельного продукта.

Функциональный подход развития может в себя включать:

1. Добавление к функционалу предобработку и постобработку. На данный момент существует огромное количество фильтров, которые могут улучшить качество результирующего изображения.

2. Изменение параметров или архитектуры нейронной сети. Сейчас архитектура VGG используется во всех проектах данного направления. Следовательно, можно поменять архитектуру на VGG19. Другой путь – полное изменение архитектуры нейронной сети, что будет являться серьезным исследованием.

3. Добавление модуля, который будет проводить предварительный анализ изображений стиля и контента, с целью рекомендаций пользователю оптимальных настроек и фильтров.

4. Разработка данного приложения как облачного сервиса.

К коммерческим вариантам развития можно отнести следующие решения, с учетом реализации вариантов функционального развития:

1. Написание мобильного клиента, который будет являться оберткой приложения для пользователя. Клиент будет выводить для пользователя всю необходимую информацию, а сама генерация изображения должна производиться на облачном сервисе. Данное решение самое популярное у аналогов разработанного проекта.

2. Интеграция данного решения в дизайнерские приложения, что позволит значительно расширить функционал последнего.

3. Интеграция решения с популярными фотохостингами как стадию предобработки изображения либо фотографии.

СПИСОК ПУБЛИКАЦИЙ АВТОРА

[1-А.] Царикович А.А., Перенос стиля изображения // Сборник тезисов 55-й юбилейной научной конференции аспирантов, магистрантов и студентов БГУИР на кафедре ЭВМ (секция "Электронные вычислительные машины", Минск, Беларусь, 2019) // Минск 2019.

[2-А.] Царикович А.А., Проблемы переноса стиля изображения // Globus / г.Москва, 2019.