

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

УДК 004.932.4

Фисько

Дмитрий Владимирович

Условные генеративные нейросетевые модели
для генерации изображений

АВТОРЕФЕРАТ

диссертации на соискание степени магистра
по специальности 1-40 80 04 – Информатика и технологии программирования

Научный руководитель

Сиротко С.И.

к.т.н., доцент

Минск 2021

ВВЕДЕНИЕ

Одним из основных способов получения конкурентного преимущества при создании технологического продукта является изменение бизнес-модели через внедрение технологической новации. Технологические новации позволяют найти новые более эффективные способы удовлетворения существующих потребностей пользователей. Одной из таких новаций является применение условных генеративных методов. Условные генеративные методы позволяют сгенерировать по входным условиям данные из некоторого интересующего распределения. Причём эти данные могут быть самые разнообразные: изображения, 3D модели, видео, аудио, текст. Разнообразие доменных областей, к которым могут быть применены методы условной генерации является одним из основополагающих факторов, обуславливающих актуальность данного направления в современных научных исследованиях.

В данной работе рассматривается применение методов условной генерации изображений для условной генерации изображений из распределения произведений изобразительных искусств. Проводится анализ существующих методов условной генерации, делается попытка улучшения существующих методов. Затрагиваются технологические аспекты обучения больших нейронных сетей.

Обучение математических моделей условной генерации изображений является нетривиальной задачей. Большинство методов использует состязательные нейронные сети для обучения генерирующей модели. Использование состязательных нейронных сетей делает процесс обучения крайне нестабильным. Требуется на ранних стадиях обучения понимать, что дальнейшее обучение нейронной сети не приведёт к результату. Нейронные сети способные к генерации изображений высокого качества являются трудно вычислимыми, поэтому требуются значительные вычислительные ресурсы. Каждая итерация по проверке эксперимента с новой идеей может занимать несколько суток. Для решения перечисленных проблем в ходе реализации магистерской работы создаётся инфраструктура для обучения больших нейросетевых моделей и анализа результатов экспериментов.

Диссертация проверена в системе «Антиплагиат». Процент оригинальности соответствует норме, установленной кафедрой информатики. Цитирования обозначены ссылками на публикации, указанные в «Списке использованных источников».

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка математических моделей и программного обеспечения для решения задачи условной генерации изображений на основе сегментационных карт для применения в приложении по созданию произведения изобразительного искусства на основе визуальных набросков пользователя с использованием персональных компьютеров общего назначения.

Для достижения поставленной цели необходимо решить следующие задачи:

- 1) собрать и разметить обучающую выборку данных;
- 2) изучить методы условной генерации изображений;
- 3) реализовать инфраструктуру для обучения генеративных математических моделей;
- 4) реализовать несколько вариаций генеративных моделей на основе изученных методов;
- 5) обучить созданные генеративные модели;
- 6) провести экспериментальные исследования обученных генеративных моделей;
- 7) создать приложение для демонстрации работы обученной генеративной модели.

Объектом исследования являются условные генеративные математические модели.

Предметом исследования является программное обеспечение для обучения и запуска нейросетевых математических моделей для решения задачи условной генерации изображения на основе сегментационных карт.

Основной гипотезой, положенной в основу диссертационной работы, является возможность использования компьютеров общего назначения для создания произведений изобразительного искусства на основе набросков пользователя. Возможно собрать обучающую выборку для обучения с учителем на основе общедоступного пользовательского контента социальной сети. Возможно обучить генеративную математическую модель с приемлемым уровнем качества пригодную по требуемым вычислительным ресурсам для эксплуатации на персональных компьютерах общего назначения.

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

На основе проведённого анализа конкурентной ситуации на рынке приложений для рисования, было определено, что на данный момент не существует продукта, который позволил бы создавать качественные произведения искусства без затраты значительных усилий. Целью данной научной работы ставится разработка технологии позволяющей создать новый продукт с качественно новым предложением для большого высоко конкурентного рынка.

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя С. И. Сиротко, заключается в корректировке формулирования результатов исследования.

Апробация результатов диссертации

Основные положения диссертационной работы обсуждались на IX международной научной конференции «Информационные технологии и системы» (Минск, Беларусь, 2019); Международной научно–технической конференции «Актуальные вопросы современной науки и технологий» (Петрозаводск, Россия, 2021).

Публикации

По теме диссертации опубликовано 2 печатные работы в сборниках трудов и материалов международных конференций.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, четырех глав, заключения, списка использованных источников, списка публикаций автора и приложений. В первой главе представлен анализ предметной области, выявлены основные существующие проблемы в рамках тематики исследования, показаны направления их решения. Вторая глава содержит базовую теорию по нейросетевым генеративным методам. В третьей главе осуществляется обзор современных методов для условной генерации изображений. В четвертой главе предложен алгоритм формирования обучающей выборки для условной генерации изображений, предложена архитектура системы для обучения условных генеративных нейросетевых моделей, проведён анализ методов условной генерации изображений, предложена новая нейросетевая архитектура для условной генерации изображений.

Общий объем работы составляет 71 страниц, из которых основного текста – 50 страниц, 35 рисунков на 27 страницах, 5 таблиц на 4 страницах, список использованных источников из 40 наименований на 4 страницах и 3 приложения на 8 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В **первой главе** проведён анализ конкурентной ситуации на рынке мобильных приложений для рисования. По итогам анализа бизнес-моделей приложений для рисования, сформировано ценностное предложение нового приложения для рисования, выделяющееся на фоне конкурентов. Новое приложение позволит создавать произведение искусства по визуальным наброскам пользователя.

Рисование является очень сложным и затратным по времени процессом. Создание качественного произведения требует многих часов работы и высоких профессиональных навыков. На текущий момент существующие приложения не закрывают потребности пользователей в полной мере. Одни приложения позволяют создавать уникальные произведения искусства, но требуют многих часов, затраченных на работу. Другие приложения позволяют использовать шаблоны для создания произведений искусства буквально за минуты, но использование шаблонов нивелирует уникальность и ценность создаваемого произведения. На данный момент нет решения, которое бы позволило создавать уникальные произведения искусства за считанные минуты.

Из анализа ценностного предложения видно, что на данный момент не существует продукта, который позволил бы создавать качественные произведения искусства без затрачивания значительных усилий, поэтому главное предполагаемое изменение в бизнес-модели для нового продукта по сравнению с другими продуктами, заключается в добавлении новой технологии позволяющей упростить процесс создания качественного произведения искусства. Пользователю достаточно нарисовать несколько линий и приложение само сгенерирует для пользователя произведение искусства по наброскам.

Новый продукт позволит пользователям создавать уникальные произведения искусства на много проще и делиться своим творчеством с сообществом.

Во **второй главе** проведён обзор базовой теории по генеративным методам. Рассмотрены метод PixelRNN, метод Variable Autoencoder и метод GAN.

Суть генеративных методов сводится к максимизации правдоподобия m независимо выбранных элементов из обучающей выборки: $\prod_{i=1}^m p(x_i | \theta)$.

Свойства метода PixelRNN:

- 1) можно вычислить правдоподобность сгенерированного изображения;
- 2) относительно качественные генерируемые изображения;
- 3) последовательная генерация медленная.

Свойства метода VAE:

- 1) можно использовать энкодер для получения по изображению латентного вектора;
- 2) генерация изображения в один проход;
- 3) более худшее качество сгенерированного изображения по сравнению с PixelRNN;
- 4) сгенерированные изображения размыты и более худшего качества по сравнению с лучшими решениями основанными на GAN.

При обучении GAN используется подход с двумя нейронными сетями, которые противостоят друг другу. Первая нейронная сеть, так называемый дискриминатор, пытается научиться различать, настоящее изображение было подано на вход или сгенерированное. Вторая нейронная сеть, так называемый генератор, пытается обучиться генерации элементов достаточного качества, для того чтобы дискриминатор не смог бы отличить их от реальных данных из обучающей выборки.

Свойства GAN:

- 1) выучивается гладкое и компактное пространство, обладающее свойством, того, что прообразы похожих объектов в пространстве располагаются рядом. Для любой точки из априорного шума существует прообраз;
- 2) генератор обучается создавать близкие к реальным изображения ни разу не увидев настоящих изображений из обучающей выборки. Обучение происходит на основе градиентов, в которых закодирован сигнал, почему дискриминатор считает изображение не настоящим;
- 3) на обученном генераторе вычислительно просто производить генерацию изображений, достаточно одного шага вычислений нейронной сети.

В **третьей главе** проведён обзор современных методов условной генерации изображений на основе датасета ADE20K.

Методы условной генерации изображений прошли через серию значительных эволюционных улучшений. Первый успех в генерации изображений был достигнут с помощью состязательной нейронной сети. Данная

нейронная сеть генерировала изображения относительно хорошего качества, но при её обучении не удалось преодолеть барьер в генерации изображения с разрешением 256×256 из-за крайней неустойчивости процесса обучения.

По причине нестабильного процесса обучения состязательных нейронных сетей, в последующих работах были высказаны предложения заменить использование архитектуры состязательных нейронных сетей на более стабильную архитектуру, в которой функция потерь задана явно. В качестве такой функции потерь было предложено использовать L1 loss между слоями активаций предобученной нейронной сети на основе архитектуры VGG, так называемой perceptual loss.

Авторы следующих статей показали, что возможно обучить состязательную нейронную сеть для генерации изображений высокого разрешения, при этом улучшив метрики качества предыдущих работ. Ранее возникающая проблема с нестабильным процессом обучения была решена через добавление perceptual loss и её аналога на основе слоёв активаций дискриминатора.

Стабилизировав обучения состязательных нейронных сетей, исследователи сосредоточились на поисках наиболее эффективных архитектур подсетей. Исследователи заметили, что слои нормализации в значительной мере фильтруют информацию, переданную на вход генерирующей подсети. Чтобы решить эту проблему, был предложен подход с заменой обычной нормализации на условную нормализации, состоящей из двух стадий. На первой стадии выполняется обычная нормализация, на второй стадии выполняется денормализация с помощью параметров, вычисленных на основе исходных данных. Данный слой условной нормализации позволил в значительной мере упростить архитектуру генерирующей подсети.

В **четвёртой главе** описан алгоритм формирования обучающей выборки без использования ручной разметки. Описан процесс проектирования и реализации архитектуры системы для обучения условных генеративных нейросетевых моделей. Описан процесс реализации и обучения математических моделей, проведён анализ результатов обучения генерирующих нейронных сетей.

Описан алгоритм создания обучающей выборки. Для создания обучающей выборки было реализовано 4 подзадачи:

- 1) найден источник качественных изображений требуемого прикладного домена;
- 2) изображения были отфильтрованы с помощью эвристик на основе метаданных и обученного классификатора;

- 3) произведена нормализация элементов обучающей выборки;
- 4) получены сегментационные карты признаков без использования ручной разметки.

Процесс обучения нейронной сети сопряжён с необходимостью создания инфраструктуры позволяющей быстро совершать итерации изменений и отслеживать процесс обучения.

Критерии, ускоряющие процесс совершения итераций:

- 1) архитектура проекта легко модифицируема;
- 2) деплой изменений кода быстрый;
- 3) в процессе обучения задействовано наибольшее количество доступных вычислительных ресурсов.

Для отслеживания процесса обучения математических моделей необходимо:

- 1) визуализировать данные, на основе которых происходило обучения математических моделей;
- 2) визуализировать по отдельности работу каждого компонента;
- 3) визуализировать итерации процесса обучения;
- 4) персистентно сохранять результаты процесса обучения.

В данной главе был описан процесс проектирования и реализации системы для обучения нейронных сетей соответствующий перечисленным критериям.

На основе изученной литературы были выделены 3 нейросетевые архитектуры, для проведения экспериментов по обучению генерирующей нейронной сети:

- 1) архитектура на основе метода SPADE;
- 2) архитектура на основе метода OASIS;
- 3) новая архитектура на основе методов OASIS и SPADE.

На основании сравнения качества генерируемых изображений по метрике FID, архитектура SPADE показала наилучший результат по сравнению с другими рассмотренными методами. Архитектура SPADE также показала наилучший визуальный результат при визуальном сравнении результатов обучения.

Предложенный метод на основе комбинирования идей методов SPADE и OASIS показал результаты, превосходящие исходную архитектуру OASIS.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1) Проведён анализ методов условной генерации изображений. По результатам проведённого анализа были выделены и реализованы лучшие подходы для условной генерации изображений. Был проведён сравнительный анализ обученных нейросетевых моделей. Выбрана нейросетевая архитектура, наиболее подходящая для обучения условной генерации изображений применительно к сформированной обучающей выборке.

2) Предложен алгоритм формирования обучающей выборки для условной генерации изображений из распределения произведений искусств на основе источника открытых данных без необходимости ручной разметки изображений.

3) Предложена новая нейросетевая архитектура для условной генерации изображений. Проведён сравнительный анализ данной архитектуры по отношению к другим методам генерации, проведён сравнительный анализ данной архитектуры по отношению к другим методам генерации

4) Предложена архитектура системы для обучения условных генеративных нейросетевых моделей. Система позволяет гибко добавлять новые архитектуры нейросетевых моделей, построенных на основе состязательных нейронных сетей. Система позволяет отслеживать и систематизировать процесс обучения моделей. Позволяет вносить корректировки в гиперпараметры модели в ходе обучения.

Рекомендации по практическому использованию результатов

1) Результаты работы могут быть использованы для создания мобильного приложения в значительной степени упрощающего процесс рисования. Сформированная бизнес-модель мобильного приложения позволяет выделяться на фоне конкурентов и получить значительную долю рынка мобильных приложений для рисования.

2) Предложенный алгоритм формирования обучающей выборки без необходимости ручной разметки обучающей выборки может быть использован для других прикладных доменов.

3) Спроектированная архитектура системы обучения состязательных нейронных сетей для условной генерации изображений позволяет легко организовать процесс разработки новых методов на основе состязательных нейронных сетей, удовлетворяющих требуемым критериям.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1–А. Фисько, Д.В. Использование дистилляции между слоями нейронной сети для улучшения качества сегментации дорожной разметки / Фисько Д.В., Сиротко С. И. // Информационные технологии и системы 2019 (ИТС 2019): материалы международной научной конференции, Минск, 30 октября 2019 г. / Белорусский государственный университет информатики и радиоэлектроники; редкол. : Л. Ю. Шилин [и др.]. – Минск, 2019. – С. 126 – 127.

2–А. Фисько, Д.В. Обзор методов условной генерации изображений нейросетевыми моделями / Фисько Д.В. // АКТУАЛЬНЫЕ ВОПРОСЫ СОВРЕМЕННОЙ НАУКИ И ТЕХНОЛОГИЙ: сборник статей Международной научно–практической конференции (20 мая 2021 г.) / МЦНП «Новая наука»; редкол.: М.В. Посновой [и др.]. – Петрозаводск, 2021. – с. 57–62.

Библиотека БГУИР