

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**ВОССТАНОВЛЕНИЕ ОТСУТСТВУЮЩЕЙ ОБЛАСТИ
ИЗОБРАЖЕНИЯ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 441 группы

направления 02.03.03 Математическое обеспечение и администрирование
информационных систем

факультета компьютерных наук и информационных технологий

Чуписа Виктора Михайловича

Научный руководитель:

доцент

Е. В. Кудрина

Заведующий кафедрой

доцент, к. ф.-м. н.

М. В. Огнева

Саратов 2025

ВВЕДЕНИЕ

Актуальность темы.

В последнее время нейронные сети стали крайне популярны. Они используются в крупных корпорациях и проектах, люди сталкиваются с ними в повседневной жизни каждый день, так что можно сказать, что разработка нейронных сетей — это бурно развивающаяся и очень перспективная область. Нейронные сети применяются для большого спектра вещей: от распознавания лиц и объектов до управления машинами без вмешательства человека. Также нейронные сети применяются для восстановления изображений. Допустим, человек захочет дорисовать отсутствующую или поврежденную часть изображения. Чаще всего это старые портретные фотографии. Или часть фотографии получилась неудачной и он хочет исправить ее без использования механических средств, таких как фотошоп или других графических редакторов, требующих большого опыта работы и знания этого редактора. Приложения для генерации части фотографий или картинок пользуются большой популярностью в сфере создания медиа контента: новостные редакции, ведение соцсетей, различного рода медиашоу. Эта задача приносит пользу ряду приложений, включая удаление объектов, восстановление фотографий и т.д.

Восстановление изображений — это метод реконструкции поврежденной фотографии или изображения путем заполнения отсутствующих пикселей, чтобы восстановленное изображение выглядело реалистично и соответствовало не поврежденному изображению. Повреждения могут быть получены из разных источников, например, в случае с физическими фотографиями фотография может быть порвана, поцарапана или исписана. В случае с цифровыми изображениями восстановление также можно использовать для удаления водяных знаков или удаления мешающих объектов.

Чтобы восстановить изображение, необходимо предоставить поврежденное или замаскированное входное изображение. Отсутствующие

пиксели затем должны быть сгенерированы, в то время как остальные пиксели являются действительными и не изменяются, чтобы их можно было использовать в выходном изображении. Эти пиксели могут использоваться как руководство для алгоритма закрашивания, чтобы помочь заполнить отсутствующие пиксели.

В последние годы внимание в основном обращается на нейронные сети. С ними необходимо обучить полную модель, что может занять много времени и быть трудоемким для обучения на базовом компьютере из-за существенных требований к аппаратному обеспечению.

Эти методы учатся извлекать семантическую информацию из неповрежденных частей изображения, которую они используют для осмысленной дорисовки недостающих частей. Главное преимущество в том, что нейронные сети могут генерировать новые объекты, которые не присутствовали в поврежденных входных данных.

В настоящее время общепринято, что архитектуры моделей на основе глубокого обучения создаются путем объединения различных подходов, которые были получены в результате обучения и тонкой настройки, пока не были достигнуты удовлетворительные результаты.

Первые модели глубокого обучения использовали CNN в архитектуре кодера-декодера. Однако их недостатком является то, что они могут оставлять артефакты на выходном изображении. Эта проблема была значительно уменьшена введением GAN с моделью, предложенной D. Pathak, которые улучшили согласованность между фоном и отсутствующими частями. Однако этот метод был ограничен зарисовкой только квадратных отверстий, размещенных в середине изображения.

Еще одно улучшение было введено S. Iizuka. Их сеть использует два отдельных дискриминатора, локальный и глобальный, где целью локального дискриминатора является различение согласованности текстуры между реальными и поддельными образцами только в закрашенной области. Глобальный дискриминатор учитывает все выходное изображение и нацелен

на глобальную согласованность. Они также заменили полностью связанные слои каналов на расширенные свертки, чтобы увеличить рецептивное поле выходных нейронов.

Позже была выпущена улучшенная версия предыдущей работы. J. Yu использовали фреймворк грубой-тонкой закрашки с модулем контекстного внимания (CAM) в своей модели. Эта архитектура может изучать представления признаков фоновых и передних участков, чтобы заполнить отсутствующую область.

Однако использование стандартных сверточных слоев приводит к нежелательным артефактам на выходном изображении. Чтобы избежать этих проблем, G. Liu представили новый режим, который вместо этого использует частичные свертки. Свертка здесь маскируется и нормализуется, поэтому она учитывает только действительные пиксели. Им также удалось автоматически сгенерировать и обновить маску для следующего слоя, что, как они утверждают, превосходит другие модели с нерегулярными масками.

Постепенно архитектуры стали довольно сложными, что означает, что также требования к вычислительным ресурсам увеличивались. С целью сделать обучение модели дорисовки менее требовательным к вычислениям, Z. Yan представили сетевую архитектуру под названием PEPSI. Она состоит из общего кодера и двух параллельных декодеров. Задача кодирующей сети — научиться извлекать признаки из неповрежденного пути изображения и восстанавливать признаки отсутствующих частей. Грубый декодер пути реконструирует необработанное изображение, сгенерированное из карт признаков, полученных кодером, в то время как декодер восстанавливает карту признаков, которая была закодирована CAM. Последний также генерирует более приближенные изображения, поскольку он декодирует реконструированные признаки. Та же команда также предложила Diet-PEPSI, еще более легкую архитектуру, которая уменьшает количество весов, необходимых за счет использования адаптивных к скорости расширенных

сверточных слоев вместо типичных расширенных сверточных слоев, при этом сохраняя исходное качество.

Обычно проблемой является то, что модели могут давать слишком сглаженные или размытые результаты. Эта проблема рассматривается К. Назери и др. с их EdgeConnect моделью. Они предложили двухэтапную архитектуру состязательной сети, которая состоит из двух генераторов, которые разделяют дорисовку на две более мелкие задачи. Сначала генератор ребер перерисовывает контуры изображения и создает соответствующую двоичную карту. За ним следует сеть завершения изображения, которая использует контурное изображение от первого генератора для дорисовки недостающих областей.

Цель бакалаврской работы - разработать нейросеть для восстановления отсутствующей области изображения (на примере портретных фотографий людей).

Поставленная цель определила **следующие задачи**:

1. Рассмотреть историю возникновения и развития нейронных сетей.
2. Изучить теорию и основные понятия нейронных сетей.
3. Освоить инструменты работы с нейронными сетями.
4. Сформировать дата-сет для работы с портретными фотографиями людей.
5. Обучить нейросеть восстанавливать отсутствующую область прямоугольной формы на портретных фотографиях людей.
6. Продемонстрировать работу разработанной нейронной сети

Методологические основы восстановления отсутствующей области изображения представлены в классических исследованиях в разработки нейросетей и их применения для задач обработки изображений: ключевыми являются работы М. Mohammad по сверточным нейросетям [1], Ir. Purwono по генеративно-состязательным нейросетям [2], а также материалы по автокодировщикам и их видам, изложенные D. Ibomoiye [3] и Kamal

Berahmand [4]. Технологические подходы и методы — на публикациях Maximilian Seitzer [5], Luyi Yang [6] и Laura Meneghetti [7].

Практическая значимость бакалаврской работы заключается в разработке лёгкой и эффективной программы для восстановления отсутствующей области изображения. Полученное решение позволяет восстанавливать любую область различной формы и размера изображения, в частности портретных фотографий людей.

Структура и объём работы. Бакалаврская работа состоит из введения, 2 разделов, заключения, списка использованных источников и 6 приложений. Общий объем работы – 84 страницы, из них 59 страниц – основное содержание, включая 7 рисунков, список использованных источников информации – 27 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Программный комплекс молекулярного моделирования» посвящен комплексному описанию теории нейронных сетей. В рамках этого раздела даётся общее представление о структуре нейронных сетей, а также нейронных сетях, которые используются для задач обработки изображений и генерации части изображения.

Нейронная сеть — это программа или модель машинного обучения, которая принимает решения аналогично человеческому мозгу, используя процессы, имитирующие совместную работу биологических нейронов для распознавания явлений, оценки вариантов и принятия решений. Входной слой считается первым слоем нейронной сети, а выходной слой — последним. Различные уровни между этими двумя слоями обычно называются скрытые слои. Каждый уровень обычно состоит из простого алгоритма, поддерживающего одну функцию активации.

Сверточная нейронная сеть (CNN, ConvNet) — это усовершенствованная версия искусственных нейронных сетей (ANN), в первую очередь разработанная для извлечения признаков из наборов данных в виде сетки. Это особенно полезно для визуальных данных, таких как изображения или видео. CNN широко используются в приложениях компьютерного зрения из-за их эффективности в обработке визуальных данных. Сверточные нейронные сети являются мощным инструментом, который постепенно стал очень популярным, в основном благодаря своей способности обрабатывать большие объемы данных.

Генеративные состязательные сети (GAN) относятся к категории генеративных моделей. Эти модели создают новые искусственные данные на основе предоставленного обучающего набора в неконтролируемом режиме; следовательно, сеть сама обнаруживает закономерности в данных. Высокий уровень результатов достигается также за счет их архитектуры, которая

состоит из генератора и дискриминатора. Оба эти компонента сделаны из отдельных моделей нейронных сетей, и обычно они состоят из сверточных нейронных сетей; однако это не является условием. Генератор является частью модели, которая пытается создать новые данные, которые настолько реалистичны, что дискриминатор решает, что они реальны.

Автокодировщики — это тип нейронной сети, которая пытается воспроизвести данные, которые ей подаются, что означает, что выход сети такой же, как и ее входные данные. Архитектура автокодировщика состоит из энкодера, представления скрытого пространства (также называемого закодированным пространством, кодированием) и декодера. После того, как сеть получает некоторые данные, она сжимает их, используя первую часть сети, энкодер, в представление скрытого пространства, которое обычно имеет меньшую размерность. Когда размерность представления скрытого пространства меньше размерности ввода, то гарантируется, что автокодировщик не может просто скопировать вводные данные, но должен изучить важные признаки.

Для разработки нового клиентского графического модуля был выбран язык программирования Python, позволяющий писать простой и доступный код. Библиотека Tensorflow позволяет реализовать архитектуру нейросети.

Таким образом, нейронные сети представляет собой популярный инструмент для работы с изображениями, в частности задачи восстановления изображений. Существует много типов нейронных сетей, которые подходят для различных задач обработки изображений. Для реализации проекта были рассмотрены технологии и выбраны инструменты.

Второй раздел «Практическая часть» посвящен описанию практической реализации нейросети для восстановления отсутствующей области изображений. Основной упор сделан на компоненты модели: класс для обработки входных изображений, классы архитектуры нейросети и классы для ее обучения. При этом внимание сосредоточено на том, как и с помощью каких средств и библиотечных инструментов классы и модули программы обеспечивают обучение и валидацию модели.

В процессе работы был собран дата-сет, написан класс для обработки тренировочных и валидационных данных, произведена настройка и оптимизация системы для обучения модели, создана нейросеть для восстановления отсутствующей области изображения, а также браузерное расширение для выделения восстанавливаемой области входных изображений, обучена нейросеть на различном техническом оборудовании и сохранены параметры модели на различных стадиях обучения.

В завершение раздела подводятся итоги проделанной работы: демонстрируется работоспособность программы, разработанный функционал и возможности программы, а также показываются различные стадии разработки и обучения нейросети.

ЗАКЛЮЧЕНИЕ

В ходе выполнения работы были достигнуты все поставленные задачи: разобрано понятие нейронных сетей, разобраны их спецификации и методы работы с ними, рассмотрен теоретический способ создания нейросети и выбран язык программирования. Также была реализована программа для восстановления отсутствующей области изображения с помощью нейросети.

На практике данную разработку можно использовать для создания коммерческого приложения, предлагающего реконструкцию поврежденных портретных фотографий. В последнее время подобные нейросети появились в коммерческих графических редакторах для автоматизации процесса редактирования изображений, например, в Adobe Photoshop. Также подобная разработка может быть востребована в медиа сфере, где часто для создания контента используются не оригинальные фотографии.

В дальнейшем эту тему можно развить следующим образом: натренировать нейросеть для восстановления не только портретных фотографий людей, но и любых фотографий человека в полный рост, чтобы нейросеть была способна восстанавливать не только лица, но и все тело человека: одежду, любые части тела, окружение. Это позволит восстанавливать фотографии людей любого формата. Коммерческие приложения, специализирующиеся на данной задаче, пользуются большой популярностью в последнее время. А также же можно обучить нейросеть восстанавливать фотографии любого профиля. Например, фотографию разрушенного здания или местности. Такую нейросеть можно использовать для еще большего круга задач. Однако это требует значительно больших вычислительных мощностей и значительно большего дата-сета, содержащего изображения различной тематики.

Основные источники информации:

1. Mohammad Mustafa Taye. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. [Электронный ресурс] / Mohammad Mustafa Taye. // Multidisciplinary Digital Publishing Institute. 2023. - URL : <https://www.mdpi.com/2079-3197/11/3/52> (дата обращения: 25.12.2024). - Загл. с экрана. - Яз англ.
2. Ir. Purwono. Understanding Generative Adversarial Networks (GANs): A Review. [Электронный ресурс] / Ir. Purwono, Annastasya Nabila Elsa Wulandari, Alfian Ma'arif, Wael A. Salah. // Control Systems and Optimization Letters. 2025. - URL : https://www.researchgate.net/publication/388803152_Understanding_Generative_Adversarial_Networks_GANs_A_Review (дата обращения: 17.01.2025). - Загл. с экрана. - Яз англ.
3. Ibomoiye Domor Mienye. Deep Autoencoder Neural Networks: A Comprehensive Review and New Perspectives. [Электронный ресурс] / Ibomoiye Domor Mienye, Theo G. Swart. // Archives of Computational Methods in Engineering. 2025. - URL : https://www.researchgate.net/publication/389877168_Deep_Autoencoder_Neural_Networks_A_Comprehensive_Review_and_New_Perspectives (дата обращения: 20.01.2025). - Загл. с экрана. - Яз англ.
4. Kamal Berahmand. Autoencoders and their applications in machine learning: a survey. [Электронный ресурс] / Kamal Berahman, Fatemeh Daneshfar, Elaheh Sadat Salehi, Yuefeng Li, Yue Xu. // Artificial Intelligence Review. 2024. - URL : https://www.researchgate.net/publication/377955158_Autoencoders_and_their_applications_in_machine_learning_a_survey (дата обращения: 27.01.2025). - Загл. с экрана. - Яз англ.
5. Maximilian Seitzer. On the pitfalls of heteroscedastic uncertainty estimation with probabilistic neural networks. [Электронный ресурс] /

Maximilian Seitzer, Arash Tavakoli, Dimitrije Antic, Georg Martius. // International Conference on Learning Representations. 2022. - <https://openreview.net/forum?id=aPOpXlnV1T> (дата обращения: 13.02.2025). - Загл. с экрана. - Яз англ.

6. Luyi Yang. Theoretical Analysis of Adam Optimizer in the Presence of Gradient Skewness. [Электронный ресурс] / Luyi Yang. // International Journal of Applied Science. 2024. – URL : https://www.researchgate.net/publication/384654097_Theoretical_Analysis_of_Adam_Optimizer_in_the_Presence_of_Gradient_Skewness (дата обращения: 15.03.2025). - Загл. с экрана. - Яз англ.
7. Laura Meneghetti. A dimensionality reduction approach for convolutional neural networks. [Электронный ресурс] / Laura Meneghetti, Nicola Demo, Gianluigi Rozza. // Applied Intelligence. 2023. – URL : https://www.researchgate.net/publication/372105681_A_dimensionality_reduction_approach_for_convolutional_neural_networks (дата обращения: 28.02.2025). - Загл. с экрана. - Яз англ.