

Влияние информационной плотности входных данных на эффективность графических нейронных сетей

Н.А. Бондарева^{1,А}

Институт прикладной математики им. М.В. Келдыша РАН

¹ ORCID: 0000-0002-7586-903X, nicibond9991@gmail.com

Аннотация

Работа предлагает краткий обзор генеративных нейросетей и рассматривает роль информации при обучении генеративных нейросетей. В цифровой среде каждый объект окружён обширным информационным полем, включающим в себя неупорядоченную информацию и совокупность упоминаний о нём. Плотность информационного поля объекта определяет способность таких технологий, как искусственный интеллект, воссоздавать его образ на основе собранных данных. Чем больше данных доступно, тем точнее и полнее можно воссоздать цифровой образ. В работе рассматривается ряд проблем, возникающих при использовании сетей text-to-image и возможные методы их решения. Статья рассматривает различные аспекты роли личных данных и возможных этических и социальных последствий в эпоху генеративных технологий, а также перспективы и риски дальнейшего развития генеративных нейросетей в специализированных областях, таких как медицина и производство. Стремительное развитие нейросетевых технологий может оказать значительное влияние на образование и социальные явления.

Ключевые слова: Машинное обучение, компьютерное зрение и распознавание образов, нейронная сеть, компьютерная графика, плотность информационного поля, Text-to-image.

1. Введение

В настоящее время область нейросетевых технологий развивается стремительно, с каждым днем обретая все больше умений и возможностей. Особенно популярными становятся генеративные диффузионные модели, в которых берут свою основу большинство текстовых нейросети, способных собирать, анализировать и генерировать текстовую информацию по запросу, и графических которые могут обрабатывать медиа-контент разными способами, от анимирования фотографий до автоматического создания изображений и видео по текстовому запросу. Результаты их работы применяются во многих отраслях, от медиа до медицины, но быстрый прогресс вызывает и социальные изменения. Раньше перспективной альтернативой диффузионным моделям считались генеративно-состязательные нейросети (GAN), но они оказались менее эффективными для генерации изображений по тексту и нестабильными в обучении [1].

Модели диффузии - это итеративные алгоритмы, превращающие случайный шум в изображение. Примером служит модель DDPM (Denoising Diffusion Probabilistic Model) [2], обучаемая на тысячах изображений, к которым последовательно добавляется шум. Модель учится удалять этот шум, улучшая качество изображения. Если применять обученную модель к случайному шуму, она сможет создать новое изображение, постепенно очищая его от шума. На рисунке показан пример, где пользователь задает схематичный рисунок. Изображение зашумляется, а затем модель восстанавливает его с высокой точностью через процесс обратной диффузии.

Графические нейросети интерпретируют лингвистическую структуру запроса, обрабатывают его и формируют реалистичные визуальные образы. Они управляют несколькими объектами, их атрибутами и пространственными отношениями, устанавливая правильные связи между характеристиками объектов. Основой для таких задач служат диффузионные нейросетевые модели [2,3], которые появились в 2015 году, но получили популярность после работы [2]. Сегодня они достигают впечатляющих результатов в генерации и изменении изображений, таких как генерация изображений, музыки и видео по текстовому запросу (text-to-image), восстановление деталей (inpainting), удаление объектов и увеличение разрешения (super-resolution).

Модели text-to-image используют лингвистическую конструкцию (текстовый запрос), направляющую процесс обработки. Языковые модели, обученные на парах изображений и текстов, понимают содержание обоих типов данных. Например, модель CLIP (Contrastive Language — Image Pre-training) от OpenAI преобразует изображения и тексты в общее латентное векторное пространство, где вектора представляют собой набор значений. В таком пространстве можно находить ближайшие изображения к текстовому запросу, просто оперируя векторами. Модель Latent Diffusion [4], представленная в 2021 году, обучает генерации изображений из направленного шума, используя латентное пространство для текстов и изображений. Такие же принципы применяются в моделях Stable Diffusion, Imagen и других крупных нейросетях для преобразования текста в изображения. Основные принципы их работы описаны в [5-8].

2. Роль информации при обучении генеративных нейросетей

Генеративные нейросети, такие как Generative Adversarial Networks (GAN), Variational Autoencoders (VAE), и трансформеры (например, GPT), представляют собой сложные системы, способные создавать новые данные, которые невозможно отличить от реальных. Роль информации в этих моделях многогранна и критична для их успешного функционирования и обучения. Основываясь на принципах машинного обучения, данные становятся фундаментом, на котором строится обучающая модель, позволяя ей обобщать, интерполировать и экстраполировать закономерности.

Отдельной задачей помимо обучения нейронной сети становится формирование обучающих датасетов [9]. Датасет является основным источником информации, на котором нейросеть строит свою модель. Модель учится на датасете, анализируя его, выявляя паттерны и закономерности, которые она впоследствии использует для генерации. Его качество и разнообразие напрямую влияют на способности генерируемой модели: чем лучше подготовлены данные, тем меньше времени потребуется на отладку модели, ее тренировки, поиск и устранение неточностей распознавания.

Среди основных критериев качественного датасета можно выделить следующие:

1. Полнота данных. Это означает достаточность объема, глубины и широты наборов данных. Набор данных должен содержать достаточное количество параметров или признаков, чтобы не осталось неохваченных пограничных случаев. Неполнота приводит либо к невозможности анализа, либо к необходимости отталкиваться от некоторых предположений или допущений относительно пропущенной информации.

2. Точность. Данные должны быть максимально приближены к реальным условиям, в которых будет работать модель нейросети.

3. Правильность и корректность. Этот пункт подразумевает соответствие данных реальности и корректность их интерпретации, а также соответствия формата и аннотаций данных в датасете с теми, в которых работает фреймворк и архитектура модели нейросети.

4. Однородность (uniformity) — значения всех атрибутов должны быть сравнимыми для всех данных. Неравномерности или наличие выбросов в наборах данных отрицательно влияют на качество данных обучения.

5. Наличие разделенных между собой датасетов для обучения, валидации и тестирования.

Большие языковые модели (LLM) произвели революцию в обработке естественного языка в таких задачах, как понимание прочитанного, рассуждение и генерация языка. Они представляют собой мощные инструменты в области обработки естественного языка (NLP). Такие модели обучаются на основе огромных наборов текстовых данных, и в результате оказываются способны улавливать сложные паттерны и нюансы языка.

Основной принцип работы LLM заключается в использовании архитектуры трансформеров, которая была предложена в работе [10]. Трансформеры позволяют выявлять и анализировать зависимости между словами в предложении, независимо от их позиции. Это улучшает качество генерации текста и понимания контекста.

LLM активно используются в генеративных нейросетях многих типов, — в частности в таких, как «text-to-image», «text-to-video», и «text-to-3D» — а также при разнообразной обработке текста и текстовых запросов от автоматического перевода до генерации полноценного связного текста или кода. Модели обучаются на многих тысячах гигабайт текстовых данных, в которые входят книги, статьи, веб-сайты и прочие различные текстовые ресурсы. Объем информации, содержащейся в модели LLM, можно охарактеризовать как чрезвычайно высокий. Например, модель GPT-3 от OpenAI имеет 175 миллиардов параметров, а её предшественница GPT-2 - всего 1,5 миллиарда параметров.

Модель графической нейросети Kandinsky обучалась на платформе SberCloud ML Spase в течение двух месяцев, потратив 20 352 GPU-V100 дней. Для её обучения использовался датасет из 60 млн пар изображений и текстовых описаний, который впоследствии был сокращён до 28 млн. В обучении были задействованы такие известные датасеты, как ConceptualCaptions [11] (датасет, содержащий более 3 миллионов изображений с подписями на естественном языке, необработанные описания для которых собираются из Интернета) и YFCC100m [12] (крупнейшая общедоступная и свободно используемая коллекция мультимедиа, содержащая метаданные около 99,2 миллионов фотографий), переведённые на русский язык. Первый этап обучения составил 250 тыс. итераций.

Однако возможности нейросетей не ограничиваются обработкой естественного языка и генерацией изображений. Для специализированных задач — таких, как производство или медицина, возникают потребности в специальных нейросетях, обученных на профессионально ориентированных датасетах. Такие нейросети должны быть способны воспринимать узкопрофильные жаргоны и научно-технические термины. Это необходимо, чтобы предотвратить возможные двусмысленные трактовки. Учитывая огромное количество накопленных данных и наличие специализированных архивов для множества областей, создание графических нейросетей с конкретной направленностью — лишь вопрос времени. Потенциальное их применение открывает широкие возможности для анализа и сопоставления различных типов данных, а также для их визуализации в доступной и наглядной форме.

Такие нейросети могут оказаться полезными и в учебно-методических пособиях. Например, нейросеть способна отразить типичное состояние органа или тканей при наличии определенного набора симптомов, упомянутых в запросе. Если в тексте описания указана какая-либо патология, визуализация может помочь выделить её особенности, что способствует принятию правильного диагностического решения.

В области производства огромные базы данных технических чертежей и стандартных формализованных и подробно описанных 3D моделей дают потенциальную возможность для обучения специализированных нейросетей, ориентированных на строгую формализацию запроса и ограниченную тематику. Такие специализированные нейросети могут быть использованы для генерации CAD-файлов с помощью текстовых подсказок, как показано на рисунке 1, для восстановления 3D модели по чертежам и для оценки и сопоставления деревьев 3D моделей с целью выявления идентичных мо-

делей, несмотря на различия в их построении. Генерируемые модели можно импортировать в выбранную программу CAD или создавать специализированные генераторы Text-to-CAD без построения и поддержания инфраструктуры.

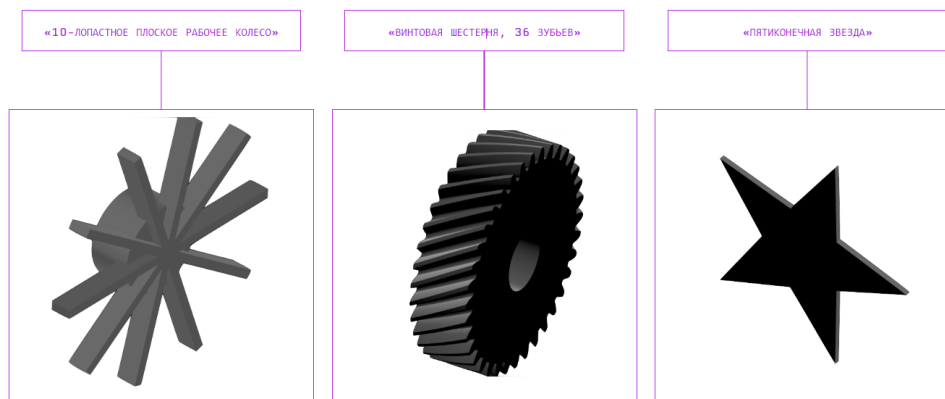


Рисунок 1 — Text-to-CAD, интерфейс для создания файлов CAD с помощью текстовых подсказок [13]

Кроме того, современные нейросети text-to-image, позволяющие создавать изображения или 3D модели по текстовому запросу, могут применяться в производственном процессе для получения предварительного визуального вида детали, который в дальнейшем может быть переработан в соответствии с видением конструктора. Пример такой нейросети на рисунке 2. Предварительный концепт, не требующий значительных затрат и предоставляемый в неограниченном количестве вариантов, может значительно сократить трудоемкость и расходы по созданию прототипов в процессе научно-исследовательских и опытно-конструкторских работ (НИР и НИОКР).



Рисунок 2 — Пример создания САПР по текстовым подсказкам

3. Информационное поле объекта

Каждый объект в цифровой среде обладает так называемым информационным полем. Информационное поле объекта определяется как весь объем неупорядоченной информации, связанной с искомым объектом и совокупность упоминаний в цифровой среде. Иными словами, это то количество открытой и публичной информации, которое окружает искомый объект и позволяет воссоздать его образ искусственно.

Информационное поле включает в себя все упоминания объекта в цифровой среде. Это могут быть:

и обработка позволяют использовать эти данные в целях воссоздания искусственного облика, манеры речи, видео и аудио материалов.

Современные технологии синтеза речи уже способны создавать практически неотличимые от оригинала голоса медиа-персон, основываясь на многочисленных аудио-записях. С помощью методов машинного обучения и нейронных сетей можно также воссоздавать видеовersion этих людей, синтезируя изображения, которые будут с максимальным приближением к реалистичности передавать мимику, жесты и движения оригиналов.

Кроме того, все эти данные позволяют создавать подробные психологические и поведенческие профили медиа-персон и государственных деятелей. Анализируя их публичные выступления, интервью и посты в социальных сетях, можно выявить их предпочтения, убеждения и мотивации. Например, даже использование определенных слов и фраз может дать представление о стиле общения человека, его эмоциональном состоянии и даже профессиональной компетенции. Эти данные могут использоваться не только для создания точных симуляций, но и для прогнозирования поведения этих личностей в тех или иных ситуациях.

Использование таких потенциальных возможностей нейронных сетей может быть разным. В статье [14] рассматриваются примеры известных случаев вредоносного использования нейросетей, от мошенничества до манипулирования общественным мнением.

Сейчас в большинстве популярных нейросетей, особенно тех, которые предоставляют свои услуги на платной основе, разработчики предпринимают широкий ряд специальных мер, направленных на уменьшение потенциального вредоносного влияния и обеспечение безопасности общества. Разработчики вводят всё больше ограничений, касающихся использования образов известных государственных и медийных лиц в изображениях, сгенерированных нейросетями. Однако поскольку остается возможность локального пользовательского обучения отдельных моделей нейросетей, то можно предположить, что проблема будет оставаться актуальной на протяжении продолжительного времени.

Вместе с тем набирает силу и популярность тенденция создания цифровых двойников реальных людей, как живых, так и давно умерших. Идея «цифрового пантеона» не нова, но теперь, имея возможность обучить нейросетевую модель на кадрах хроники, выдержках из личной переписки и собраниях сочинений известных исторических деятелей, возникает риск возникновения новых приемов манипуляции общественным мнением и засорения информационного пространства и образовательной системы ложной или искусственно сгенерированной информацией.

Подобное явление можно назвать созданием так называемой цифровой псевдоличности, которая сможет с определенной мерой достоверности воспроизводить речь и образ мышления определенного человека. В настоящее время уже существует услуга создания такой псевдоличности для публичных людей или ведущих бизнес [15]. Для этого потребуется загрузить в базу данных свой голос, цифровые образцы внешности (фото и видео), образцы личной и деловой переписки, а также примеры текстов в различных стилях. Утверждается, что такая цифровая псевдоличность сможет имитировать стиль общения своего оригинала и вести переговоры от его имени (например, с клиентами).

В этой идее существует ряд рискованных моментов:

- 1) Необходима строгая конфиденциальность личных данных. Если разработчик позволит им «уплыть» в свободный интернет, то проконтролировать дальнейшее развитие своих потенциальных двойников пользователь уже не сможет.

- 2) Фейковые факты и высказывания. Пользователь может столкнуться с тем, что в его биографии всплывают ему самому неизвестные факты, которые на самом деле являются порождением генерации цифровой модели. Также ему могут приписать слова, которые он не говорил.

3) Манипуляции и мошенничество. Сбор и хранение биометрических и поведенческих данных в частных коммерческих организациях несут в себе риск, что данные окажутся в руках злоумышленников. В случае использования цифровой копии теми, кто не является оригиналом, легко распространить дезинформацию или совершать действия от имени другого человека.

4) Юридическая ответственность. Неясность в правовом регулировании создания и использования цифровых копий может создать правовые вакуумы, которые будут использоваться для незаконных целей, так как непонятно, кто будет нести ответственность за слова и действия цифровой копии реального человека.

5) Этические аспекты. В августе 2024 года впервые распространилось видео от якобы «мертвого» человека, который продолжил существовать в цифровом виде и вести блог. Видео было признано фейком, а оригинальный человек оказался жив и объявил произошедшее арт-перформансом [16], тем не менее, в обществе это создало информационный прецедент использования лица и личных данных умершего человека для воспроизведения его цифровой псевдоличности и дальнейших манипуляций от его имени. Злоупотребление подобными действиями приведет к нарушениям прав на неприкосновенность частной жизни и личных данных.

При использовании личности известных исторических деятелей, обладающих плотным информационным полем, позволяющим создать цифровую псевдоличность, может возникнуть следующий ряд проблем:

1) Искажение исторической правды, неточности и фальсификации. Станет сложно отделить сгенерированные высказывания (особенно если они становятся крылатыми) от реальных, а также будет затруднительно установить подлинность того или иного высказывания, благодаря чему возрастет риск манипуляции общественным мнением в исторических и политических диспутах. Использование таких копий для интерпретации исторических событий может внести значительные изменения в восприятие и понимание истории, что не всегда будет соответствовать действительности.

2) Политическая и социальная манипуляция. Цифровые копии исторических деятелей могут быть использованы для пропаганды и политической манипуляции, а некорректное представление исторических личностей (особенно если они были спорными фигурами в своей исторической эпохе) может вызвать социальную напряженность.

3) Образовательные риски. У будущих поколений существует риск столкнуться с явлением мнимой достоверности, когда цифровая копия будет восприниматься как достоверное представление личности исторического деятеля, что в свою очередь приведет к поверхностному восприятию и массовым искажениям реальных фактов и культурного контекста той или иной эпохи.

Между тем, по мере наполнения интернета информацией, сгенерированной нейросетями, возникают новые проблемы и потенциальные риски от последствий такого обучения. Обучение на данных, созданных самим ИИ, может приводить к накоплению ошибок и избыточным повторениям. Это сопровождается риском появления артефактов, которые трудно заметить и исправить. Статья [17] рассматривает процесс того, как использование сгенерированного моделью контента в обучении вызывает необратимые дефекты в полученных моделях, где хвосты исходного распределения контента исчезают. Авторы называют этот эффект коллапсом модели и демонстрируют, что он может возникать в вариационных автоэнкодерах, моделях гауссовой смеси и LLM.

Специалисты всего мира высказывают разные мнения [18] на предмет будущего нейросетей и тех мер, которые можно предпринять для минимизации будущего коллапса, однако многие сходятся в том, что чистота и релевантность данных будут играть ключевое значение при обучении будущих поколений нейросетей.

4. Роль личных данных в эпоху генеративных нейросетей

В эпоху расцвета генеративных нейронных сетей особую роль приобретает та личная информация о человеке, которую чаще всего он разглашает о себе сам. Локальные модели вполне реально обучить даже на том небольшом информационном поле, которое складывается из социальных сетей и иных источников.

К числу таких данных можно отнести:

- фотографии в социальных сетях;
- видео;
- аудиозапись голоса — из видео или телефонных разговоров
- личные записи и блоги
- дневники
- образцы подписи

В наше время люди постоянно становятся объектами съемки вне зависимости от собственного желания — от уличных видеокамер, до случайного попадания на чужие видео, выложенные в интернет. Персонализировано собирать такую информацию непросто, это требует очень большого количества ресурсов, но, тем не менее, подобные материалы могут стать обучающим материалом для новых моделей нейросетей. На рисунке 4 показана аннотированная фотография из обучающего набора данных Diversity in Faces от IBM, подготовленная для обучения процессу распознавания лиц.

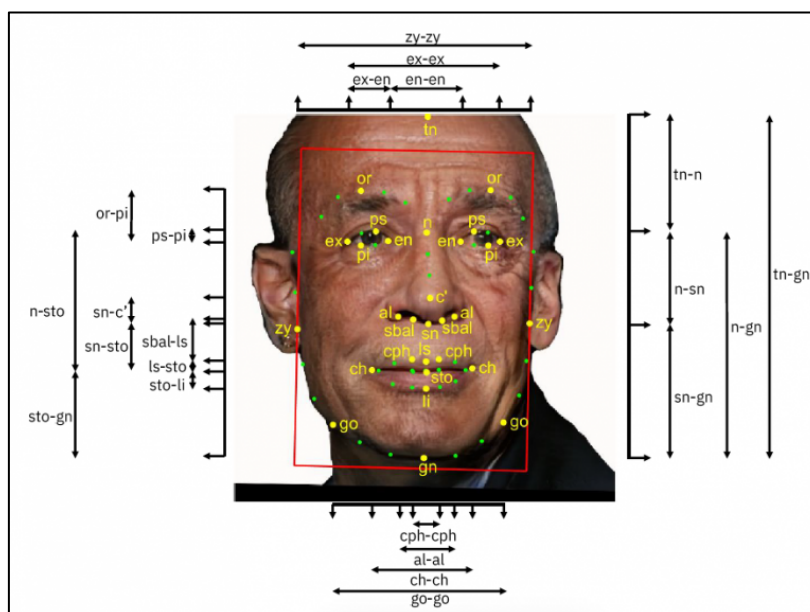


Рисунок 4 — Аннотированная фотография из набора данных Diversity in Faces от IBM

Поскольку алгоритмы распознавания лиц требуют огромного количества изображений, то методы ручного сбора фотографий никак не могут покрыть их потребности, так что исследователи начали массово собирать изображения с сайтов, таких как Flickr, Facebook, YouTube и других.

Статья [19] в подробностях раскрывает проблематику несанкционированного извлечения миллионов фотографий из Интернета для обучения соответствующих алгоритмов распознавания лиц. В январе 2019 г. IBM выпустила коллекцию почти миллиона фото с платформы Flickr, упомянув о намерении сократить объективные ошибки с помощью многообразного обучающего датасета, что в свою очередь вызвало серьезный резонанс в среде фотографов, чьи работы попали в эту коллекцию без уведомления как автора, так и модели. В частности были высказаны опасения, что собранные данные

могут быть использованы для ограничения основных прав и конфиденциальности, а также репрессивной и дискриминационной политики.

Такие случайные фото также становятся элементами информационного поля, связанного с теми или иными объектами, к примеру, с людьми, чьи лица можно определить на фотографии и далее установить их личность более детально.

Учесть такие данные в личном контроле невозможно, в отличие от той информации, которую человек самостоятельно размещает в открытом доступе. В частности люди, которые плотно работают с публикой удаленно с участием камер и микрофонов (даже как любители) фактически предоставляют в открытый доступ записи своего голоса, своей мимики, используемого лексикона и многое другое, что может быть использовано в различных целях. К примеру, телефонными мошенниками для вредоносных махинаций, которые способны «украсть голос» даже при помощи телефонного разговора.

Пользователями уже разрабатываются идеи защиты от подобных краж, к примеру, разработка голосовых аватаров для защиты абонента при звонках от неизвестных ему лиц [19].

Что касается крупных коммерческих нейросетевых систем по обработке изображений, которые сотрудничают с государственными и силовыми структурами всё тоже не так однозначно. В новаторском исследовании 2018 года, оказавшем значительное влияние на исследования ИИ, Джой Буоламвини и Тимнит Гебру [20] первыми выявили, что все популярные системы распознавания лиц с наибольшей точностью идентифицируют мужчин со светлой кожей (2,4% ошибок) и наиболее часто ошибаются при распознавании женщин с более темной кожей (61% ошибок). Вероятными причинами данного феномена являются следующие: меньшее количество фотографий женщин с темной кожей в базах данных, преимущественно белый мужской состав разработчиков таких систем, а также худшая работа датчиков камер при распознавании деталей в темных оттенках. Это усугубляется тем фактом, что некоторые коммерческие компании подходят к разработке нейросетевых алгоритмов с позиции «черного ящика», когда получают результаты, и сравнивают их с теми, которые хотели бы получить, не исследуя суть внутренних процессов.

Несмотря на выявляемые проблемы, данные системы продолжают активно применяться в различных областях, включая правоохранительные органы России и Китая. Исследования подтверждают, что представители расовых меньшинств в этих государствах подвергаются большему риску быть ошибочно идентифицированными как преступники. Такая тенденция обусловлена тем, что алгоритмы системы чаще устанавливают совпадение с лицами, чьи черты схожи с чертами лица подозреваемого. Известный пример подобного ложного совпадения произошел в 2023 году, когда ученый-гидролог был арестован по делу об убийствах 20-летней давности на основе данных искусственного интеллекта (ИИ). По данным новостных статей [21], программа ИИ определила, что фото задержанного на 55% совпадало с изображением подозреваемого в убийствах, совершенных в 2003 году. Дело было прекращено лишь спустя год.

В связи с этим, становится очевидной необходимость регуляций в области разработки искусственного интеллекта (ИИ). Регуляции должны заставлять разработчиков ИИ следовать общим стандартам, чтобы те не экономили на безопасности. Хотя регуляции сами по себе не создают технических решений, они всё же могут дать мощный стимул к их разработке и внедрению. Компании будут более интенсивно вырабатывать меры безопасности, если без них нельзя будет продавать свои продукты, особенно если другие компании подчинены тем же стандартам. Какие-то компании, может, и регулировали бы себя сами, но государственная регуляция помогает предотвратить то, что менее аккуратные конкуренты на безопасности сэкономят. Регуляции должны быть проактивными, а не реактивными. Часто говорят, что в авиации регуляции «написаны кровью» — но тут их надо разработать до катастрофы, а не после. Они должны быть устроены так, чтобы давать конкурентное преимущество компаниям с лучшими стандартами безопасности, а не компаниям с большими ресурсами и лучшими адвокатами. Регуля-

торов надо набирать независимо, не из одного источника экспертов (например, больших компаний), чтобы они могли сосредоточиться на своей миссии для общего блага без внешнего влияния [22, 23].

Для повышения прозрачности и подотчетности ИИ-систем необходимо требовать от компаний документации данных, объясняющей, какие источники данных используются при обучении и развёртывании своих моделей. Принятые компаниями решения использовать датасеты, в которых есть персональные данные или агрессивный контент, повышают и без того бешеный темп разработки ИИ и мешают подотчётности. Документация должна описывать мотивацию выбора, устройство, процесс сбора, назначение и поддержку каждого датасета. Общественный контроль за ИИ-системами общего назначения также становится необходимым, учитывая риски, которые частные компании никогда в должной мере не учтут. Чтобы удостовериться, что они адекватно принимаются во внимание, может потребоваться прямой общественный контроль за такими системами.

Идеальный сценарий предполагал бы, что ИИ разрабатываются, тестируются, а потом развёртываются только тогда, когда все их катастрофические риски пренебрежимо малы и находятся под контролем. Прежде чем начать работу над новым поколением ИИ-систем, должны проходить годы тестирования, мониторинга и внедрения в общество предыдущего поколения.

5. Заключение

Быстрое развитие диффузионных моделей позволило достичь больших результатов в различных областях, от медиа до медицины и производства. Способность этих моделей к генерации реалистичных изображений и текстов открывает новые возможности, но также порождает немало проблем. Одним из основных рассматриваемых моментов является критическая роль данных и их качества в процессе обучения моделей, что напрямую влияет на их работоспособность и точность. Введено и обосновано понятие «информационное поле». Отмечена необходимость уделять внимание вопросам конфиденциальности и безопасности данных для обучения, так как с увеличением влияния нейросетей, возрастают и риски, требующие решения проблем, связанных с возможным коллапсом моделей и манипуляцией общественным мнением через цифровые копии и псевдоличности. Важно продолжать исследовать и развивать нейросетевые технологии, внедряя при этом необходимые меры для их безопасного использования.

6. Благодарности

Вычислительная работа проведена с помощью гибридного суперкомпьютера К-100, установленного в Центре коллективного пользования ИПМ им. М.В. Келдыша РАН.

Список источников

1. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative Adversarial Networks, 2014, <https://doi.org/10.48550/arXiv.1406.2661>
2. J. Ho, A. Jain, P. Abbeel, Denoising Diffusion Probabilistic Models, 2020, <https://doi.org/10.48550/arXiv.2006.11239> (дата обращения 29.03.2023)
3. C. Meng, Y. He, Y. Song, J. Song, J. Wu, J. Zhu, S. Ermon, SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations, 2022, <https://doi.org/10.48550/arXiv.2108.01073>
4. R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-Resolution Image Synthesis with Latent Diffusion Models, 2021, <https://doi.org/10.48550/arXiv.2112.10752>

5. Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A., 2017. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1125–1134.
6. Koh, J. Y., Baldridge, J., Lee, H., and Yang, Y., 2021. Text-to-image generation grounded by fine-grained user attention. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 237–246.
7. Ramesh A., Pavlov M., Goh G., Gray S., Voss C., Radford A., Chen M, Sutskever I., 2021. Zero-Shot Text-to-Image Generation, <https://doi.org/10.48550/arXiv.2102.12092>
8. Radford A., Jong W.K., Hallacy C., Ramesh A., Goh G., Agarwal S., Sastry G., Askell A., Mishkin P., Clark J., Krueger G., Sutskever I. 2021. Learning Transferable Visual Models From Natural Language Supervision. arXiv preprint arXiv:2103.00020 [cs.CV]. <https://doi.org/10.48550/arXiv.2103>.
9. Manda Bharadwaj, Dhayarkar Shubham, Mitheran Sai, V K, Viekash, Muthuganapathy, Ramanathan. 2021. 'CADSketchNet' - An Annotated Sketch dataset for 3D CAD Model Retrieval with Deep Neural Networks. Computers & Graphics. 99. 10.1016/j.cag.2021.07.001.
10. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. 2017. Attention Is All You Need, arXiv preprint arXiv:1706.03762 <https://doi.org/10.48550/arXiv.1706.03762>
11. Conceptual Captions Dataset. URL: <https://github.com/google-research-datasets/conceptual-captions> (дата обращения 27.08.2024)
12. YFCC100M, URL: <https://paperswithcode.com/dataset/yfcc100m> (дата обращения 27.08.2024)
13. Generate CAD from text prompts, URL: <https://zoo.dev/text-to-cad> (дата обращения 27.08.2024)
14. Бондарева Н.А.. Графические нейронные сети и проблемы верификации изображений // Труды 33-ой Международной конференции по компьютерной графике и машинному зрению ГрафиКон 2023, Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия, 19-21 сентября 2023 г., с.317-327, DOI: 10.20948/graphicon-2023-317-327 https://www.graphicon.ru/html/2023/papers/paper_031.pdf
15. Ломакина Я. «Первый в мире мертвый блогер» оказалась живой актрисой: что это вообще было, 2024, URL: <https://journal.tinkoff.ru/dead-blogger/> (дата обращения 27.08.2024)
16. Clone yourself, 2024, URL: <https://www.delphi.ai/> (дата обращения 27.08.2024)
17. Shumailov I., Shumaylov Z., Zhao Y., Gal Y, Papernot N., Anderson R., 2023. The Curse of Recursion: Training on Generated Data Makes Models Forget. arXiv preprint arXiv:2305.17493 <https://doi.org/10.48550/arXiv.2305.17493>
18. Рожков Р. Градация деградации: ожидает ли генеративный искусственный интеллект вырождение, 2023, URL: <https://www.forbes.ru/tekhnologii/491359-gradaciya-degradacii-ozidaet-li-generativnyj-iskusstvennyj-intellekt-vyrozhdenie> (дата обращения 27.08.2024)
19. Solon O. Facial recognition's 'dirty little secret': Millions of online photos scraped without consent, 2019, URL: <https://www.nbcnews.com/tech/internet/facial-recognition-s-dirty-little-secret-millions-online-photos-scraped-n981921> (дата обращения 27.08.2024)
20. Сообщение из телеграм канала, 2023, <https://t.me/sburyi/182> (дата обращения 26.08.2024)
21. Buolamwini J., Gebru T., Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification // Proceedings of the 1st Conference on Fairness, Accountability and Transparency, PMLR 81:77-91, 2018
22. ТАСС: Суд Москвы освободил ученого, арестованного с помощью ИИ по делу 20-летней давности, 2023, URL: <https://tass.ru/proisshestiya/19508893> (дата обращения 27.08.2024)

23. Харитонова Ю.С., Савина В.С., Паньини Ф. Предвзятость алгоритмов искусственного интеллекта: вопросы этики и права // Вестник Пермского университета. Юридические науки. 2021. №53. URL: <https://cyberleninka.ru/article/n/predvzyatost-algoritmov-iskusstvennogo-intellekta-voprosy-etiki-i-prava> (дата обращения: 05.09.2024).
24. Hendrycks D., Mazeika M., Woodside T. An Overview of Catastrophic AI Risks arXiv preprint arXiv:2306.12001 <https://doi.org/10.48550/arXiv.2306.12001>

The Impact of Input Data Density on the Performance of Graphic Neural Networks

N.A. Bondareva^{1,A}

Keldysh Institute of Applied Mathematics RAS

¹ ORCID: 0000-0002-7586-903X, nicibond9991@gmail.com

Abstract

The paper provides a brief overview of generative neural networks and considers the role of information in training generative neural networks. In the digital environment, each object is surrounded by a vast information field, including unordered information and a set of references to it. The density of the object's information field determines the ability of technologies such as artificial intelligence to recreate its image based on the collected data. The more data is available, the more accurately and completely the digital image can be recreated. The paper considers a number of problems arising from the use of text-to-image networks and possible methods for solving them. The article considers various aspects of the role of personal data and possible ethical and social consequences in the era of generative technologies, as well as the prospects and risks of further development of generative neural networks in specialized areas such as medicine and manufacturing. The rapid development of neural network technologies can have a significant impact on education and social phenomena.

Keywords: Machine learning, computer vision and pattern recognition, neural network, computer graphics, information field density, Text-to-image.

References

1. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde -Farley, S. Ozair , A. Courville , Y. Bengio , Generative Adversarial Networks, 2014, <https://doi.org/10.48550/arXiv.1406.2661>
2. J. Ho, A. Jain, P. Abbeel, Denoising Diffusion Probabilistic Models, 2020, <https://doi.org/10.48550/arXiv.2006.11239> (date appeals 03/29/2023)
3. C. Meng, Y. He, Y. Song, J. Song, J. Wu, J. Zhu, S. Ermon, SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations, 2022, <https://doi.org/10.48550/arXiv.2108.01073>
4. R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-Resolution Image Synthesis with Latent Diffusion Models, 2021, <https://doi.org/10.48550/arXiv.2112.10752>
5. Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A.A., 2017. Image -toimage translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1125–1134.
6. Koh, J.Y., Baldridge, J., Lee, H., and Yang, Y., 2021. Text- image generation grounded by fine-grained user attention. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 237–246.
7. Ramesh A., Pavlov M., Goh G., Gray S., Voss C., Radford A., Chen M, Sutskever I., 2021. Zero-Shot Text-to-Image Generation, <https://doi.org/10.48550/arXiv.2102.12092>
8. Radford A., Jong WK, Hallacy C., Ramesh A., Goh G., Agarwal S., Sastry G., Askell A., Mishkin P., Clark J., Krueger G., Sutskever I. 2021. Learning Transferable Visual Models From Natural Language Supervision. arXiv preprint arXiv:2103.00020 [cs.CV]. <https://doi.org/10.48550/arXiv.2103>.

9. Manda Bharadwaj, Dhayarkar Shubham , Mitheran Sai, VK, Vieakash , Muthuganapathy , Ramanathan . 2021. ' CADSketchNet ' - An Annotated Sketch dataset for 3D CAD Model Retrieval with Deep Neural Networks. Computers & Graphics. 99. 10.1016/j.cag.2021.07.001.
10. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez AN, Kaiser L., Polosukhin I. 2017. Attention Is All You Need, arXiv preprint arXiv:1706.03762 <https://doi.org/10.48550/arXiv.1706.03762>
11. Conceptual Captions Dataset. URL: <https://github.com/google-research-datasets/conceptual-captions> (date appeals 08/27/2024)
12. YFCC 100 M, URL: <https://paperswithcode.com/dataset/yfcc100m> (accessed 27.08.2024)
13. Generate CAD from text prompts, URL: <https://zoo.dev/text-to-cad> (date appeals 08/27/2024)
14. Bondareva N.A. Graphic neural networks and image verification problems // Proceedings of the 33rd International Conference on Computer Graphics and Machine Vision GraphiCon 2023, V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia, September 19-21, 2023, pp. 317-327, DOI: 10.20948/graphicon-2023-317-327 https://www.graphicon.ru/html/2023/papers/paper_031.pdf
15. Lomakina Ya. "The World's First Dead Blogger " Turned Out to Be a Living Actress: What Was That Anyway, 2024, URL: <https://journal.tinkoff.ru/dead-blogger/> (date accessed 08/27/2024)
16. Clone yourself, 2024, URL: <https://www.delphi.ai/> (accessed 27.08.2024)
17. Shumailov I., Shumaylov Z., Zhao Y., Gal Y, Papernot N., Anderson R., 2023. The Curse of Recursion: Training on Generated Data Makes Models Forget. arXiv preprint arXiv:2305.17493 <https://doi.org/10.48550/arXiv.2305.17493>
18. Rozhkov R. Gradation of degradation: is generative artificial intelligence facing degeneration, 2023, URL: <https://www.forbes.ru/tekhnologii/491359-gradacia-degradacii-ozidaet-li-generativnyj-iskusstvennyj-intellekt-vyrozdenie> (date of access 27.08. .2024)
19. Solon O. Facial recognition's 'dirty little secret': Millions of online photos scraped without consent, 2019, URL: <https://www.nbcnews.com/tech/internet/facial-recognition-s-dirty-little-secret-millions-online-photos-scraped-n981921> (date appeals 08/27/2024)
20. Message from a telegram channel, 2023, <https://t.me/sburyi> / 182 (date of access 08/26/2024)
21. Buolamwini J., Gebru T., Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification // Proceedings of the 1st Conference on Fairness, Accountability and Transparency, PMLR 81:77-91, 2018
22. TASS: Moscow court releases scientist arrested with AI in 20-year-old case, 2023, URL: <https://tass.ru/proisshestviya/19508893> (date accessed 27.08.2024)
23. Kharitonova Yu.S., Savina V.S., Pagnini F. Bias of Artificial Intelligence Algorithms: Issues of Ethics and Law // Bulletin of Perm University. Legal Sciences. 2021. No. 53. URL : <https://cyberleninka.ru/article/n/predvzyatost-algoritmov-iskusstvennogo-intellekta-voprosy-etiki-i-prava> (date of access : 09/05/2024) .
24. Hendrycks D., Mazeika M., Woodside T. An Overview of Catastrophic AI Risks arXiv preprint arXiv:2306.12001 <https://doi.org/10.48550/arXiv.2306.12001>