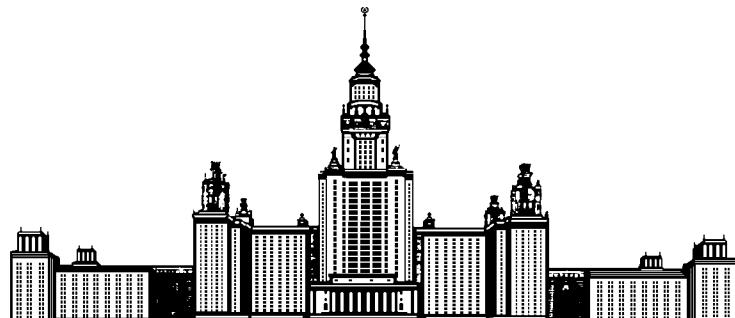


Московский государственный университет имени М. В. Ломоносова

Факультет Вычислительной Математики и Кибернетики

Кафедра Математических Методов Прогнозирования



студент группы 317

Машарипов Тимур Алишерович

Основные применения нейронных сетей

Эссе по курсу Deep Learning

Москва, 2021

Содержание

1 Аннотация	3
2 Введение	4
3 Основные применения нейронных сетей	5
3.1 Компьютерное зрение	5
3.1.1 Классификация изображений	5
3.1.2 Обнаружение объектов	6
3.1.3 Сегментация объектов	7
3.1.4 Предсказание геометрии пространства	9
3.1.5 Определение позы	10
3.1.6 Распознавание лиц	11
3.2 Обработка изображений	12
3.2.1 Генерация	12
3.2.2 Интерполяция	14
3.2.3 Трансляция	15
3.2.4 Повышение разрешения	16
3.2.5 Стилизация	16
3.2.6 Раскрашивание	17
3.2.7 Генерация анимаций	18
3.2.8 3D моделирование окружения	18
3.3 Понимание изображений	19
3.3.1 Аннотирование	19
3.3.2 Visual Reasoning	20
3.3.3 Visual Explanations	20
3.4 Text2Image	21
3.5 Обработка текста	22
3.5.1 Понимание смысла	22
3.5.2 Автоматический перевод	22
3.5.3 Диалоговые системы	22
3.5.4 Представления слов, предложений, текстов	23
3.5.5 Генерация текста	23
3.6 Обработка аудио	24
3.6.1 Распознавание речи	24
3.6.2 Генерация речи	24
3.7 Обработка видео	24
3.7.1 Чтение по губам	24
3.7.2 Deepfake	24
3.8 Автоматизация программирования	25
3.9 Искусственный интеллект в играх	26

3.10 Медицина	28
3.11 Продажи	30
3.11.1 Next-Item Prediction	30
3.11.2 Next-Basket Prediction	31
3.12 Неожиданные применения нейронных сетей	31
4 Заключение	33

1. Аннотация

В данной статье рассматриваются направления и задачи, в которых на данный момент развиваются технологии глубокого обучения и нейронных сетей в целом, сферы деятельности, в которых они уже активно применяются. Или не так активно, зато потенциал для развития весьма велик.

2. Введение

Несколько лет назад человечество и представить себе не могло, что применение глубокого обучения позволит создать беспилотные автомобили (как Tesla) и виртуальных помощников (как Алиса от Яндекса). Но сегодня эти творения являются частью нашей повседневной жизни. Глубокое обучение продолжает очаровывать нас своими бесконечными возможностями. Оно позволяет машинам решать относительно сложные задачи, несмотря на разнообразие и слабую структурированность данных.

Глубокое обучение - это форма машинного обучения, которая вдохновлена и смоделирована на основе того, как работает человеческий мозг.

Далее предлагается рассмотреть основные применения искусственных нейронных сетей в различных областях и задачах.

3. Основные применения нейронных сетей

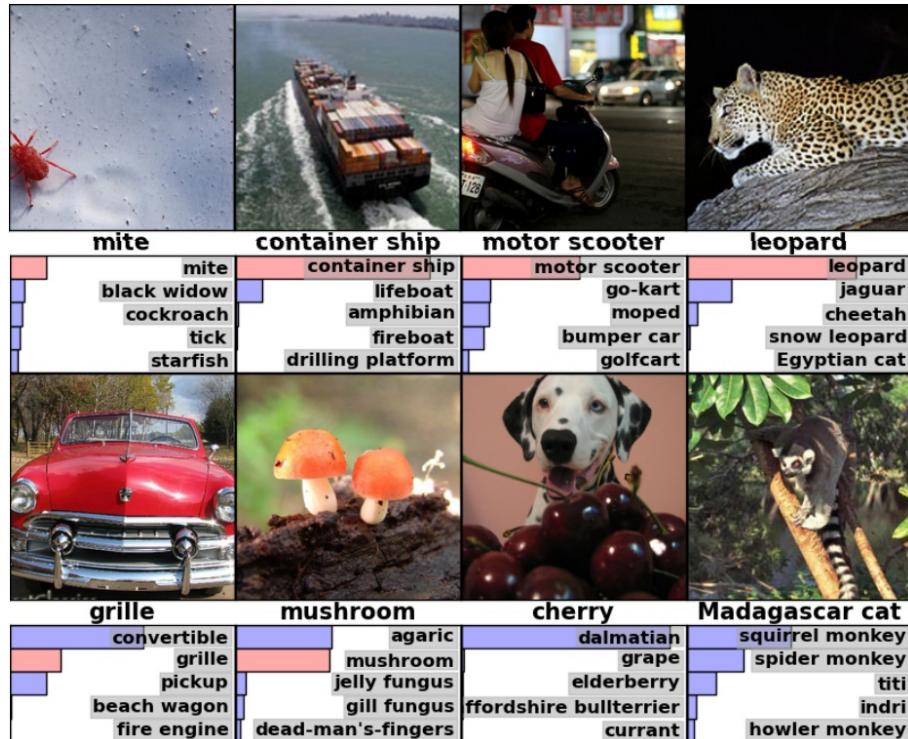
3.1. Компьютерное зрение

Стоит начать именно с компьютерного зрения, которое позволяет компьютерам и системам получать значимую информацию из цифровых изображений, ведь именно в этой области произошли одни из наиболее ярких прорывов в развитии технологии нейронных сетей.

3.1.1 Классификация изображений

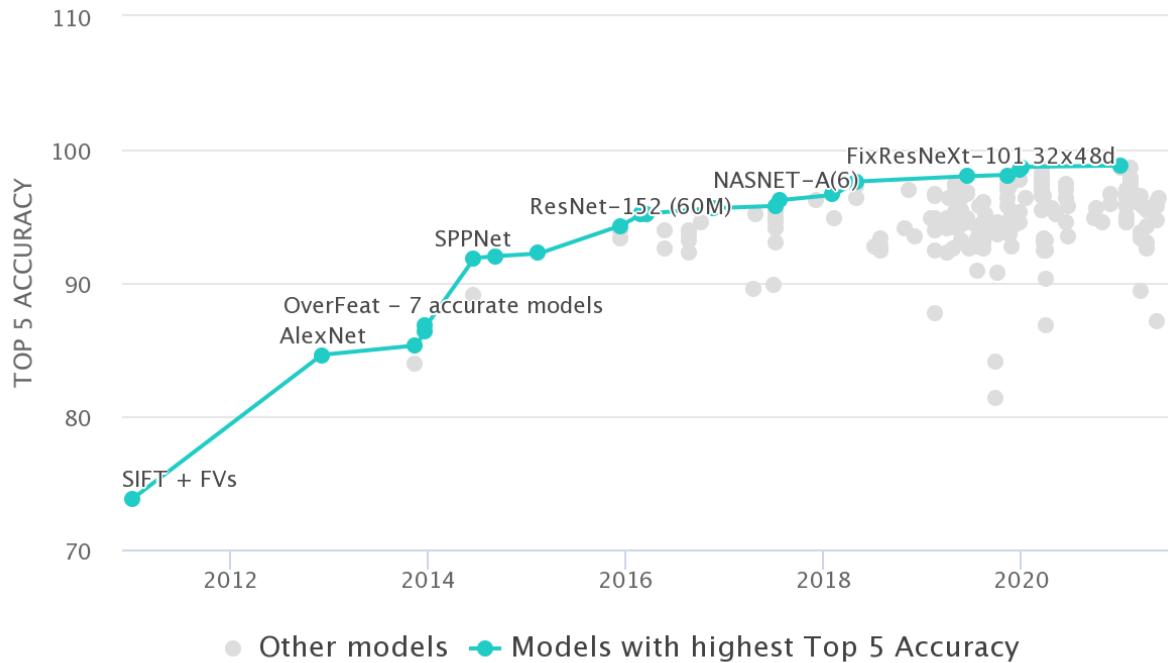
В 2012 году авторы статьи [1] представили миру революционную AlexNet - свёрточную нейронную сеть для решения задачи классификации изображений. Именно она победила в конкурсе по классификации изображений ImageNet.

Рис. 1: Результаты работы AlexNet. 8 изображений и для каждого 5 меток, наиболее вероятных с точки зрения модели. Из статьи [1]



Успех модели повлёк за собой огромный рост числа исследований применения глубокого обучения в классификации изображений. Уже начиная с некоторого момента времени, подходы различных исследователей предлагали точность, превосходящую даже человеческую.

Рис. 2: Топ5 точность различных моделей в решении задачи классификации изображений из набора данных ImageNet [2]

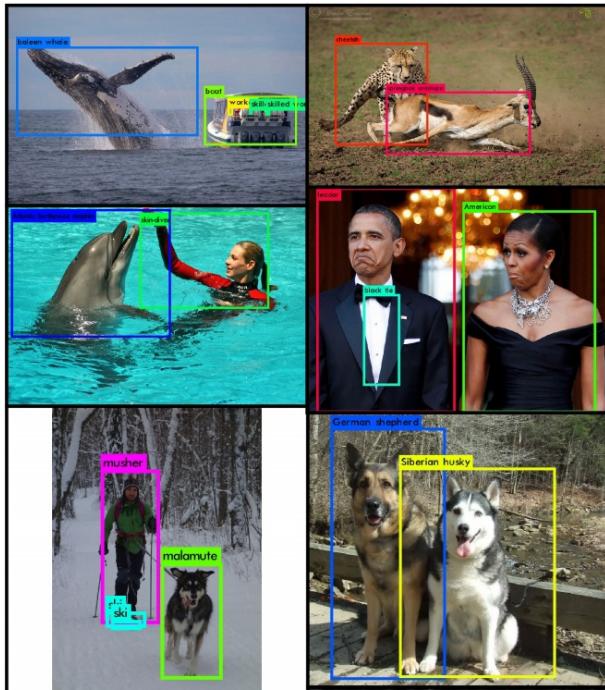


3.1.2 Обнаружение объектов

В отличие от классификации, задача обнаружения объектов подразумевает определение положения некоторых сущностей на изображении. То есть необходимо получить не только информацию о том, *кто* изображён на фото, но и *где*.

Всерьёз за решение этой задачи взялись авторы статьи [3]. Они не просто представили модель, которая способна обнаруживать объекты на изображении. Yolo9000 способна это делать в реальном времени. Вся суть намерений авторов отражена в первом же вступительном предложении из статьи: *"Обнаружение объектов должно быть быстрым, точным и способным распознавать широкий спектр объектов"* (в пер. на русский язык).

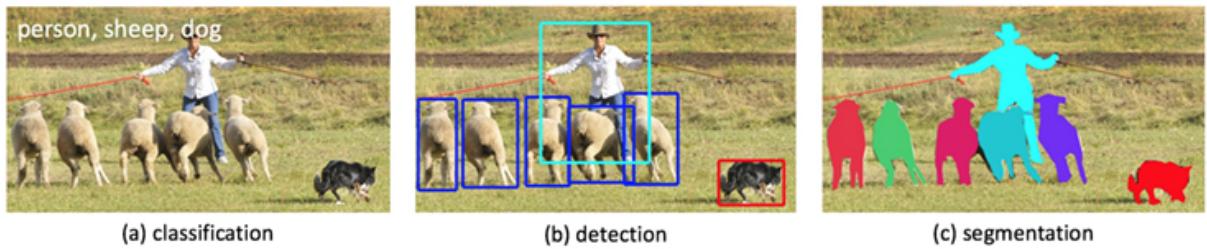
Рис. 3: Результаты работы Yolo9000. Обнаружение объектов на изображении. Из статьи [3]



3.1.3 Сегментация объектов

Любопытным обобщением обнаружения объектов является задача сегментации изображения. Она заключается в том, чтобы разделить цифровое изображение на несколько сегментов (множество пикселей, также называемых суперпикселями). То есть это процесс присвоения таких меток каждому пикселию изображения, что пиксели с одинаковыми метками имеют общие визуальные характеристики.

Рис. 4: Сравнение задач классификации, обнаружения и сегментации на примере одного изображения [4]



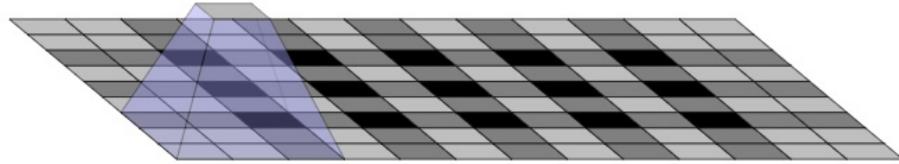
Глубокое обучение во многом упростило пайплайн выполнения сегментации объектов, одновременно показав впечатляющее качество.

Одна из самых простых и популярных архитектур, используемых для сегмен-

тации, это полносверточная сеть [5]. Несмотря на её эффективность, она обладает недостатками:

- наличие артефактов, расположенных в шахматном порядке, связанных с неравномерным перекрытием свёрток
- низкое разрешение по краям из-за потери информации в процессе кодирования

Рис. 5: Артефакт "шахматная доска". Из статьи [6]



Поэтому была разработана усовершенствованная версия простой FCN. Это сеть U-Net [7] использующая концепцию skip-connections. И наиболее естественным применением технологии сегментации является автоматическое удаление фона на изображении.

Рис. 6: Архитектура U-Net [7]

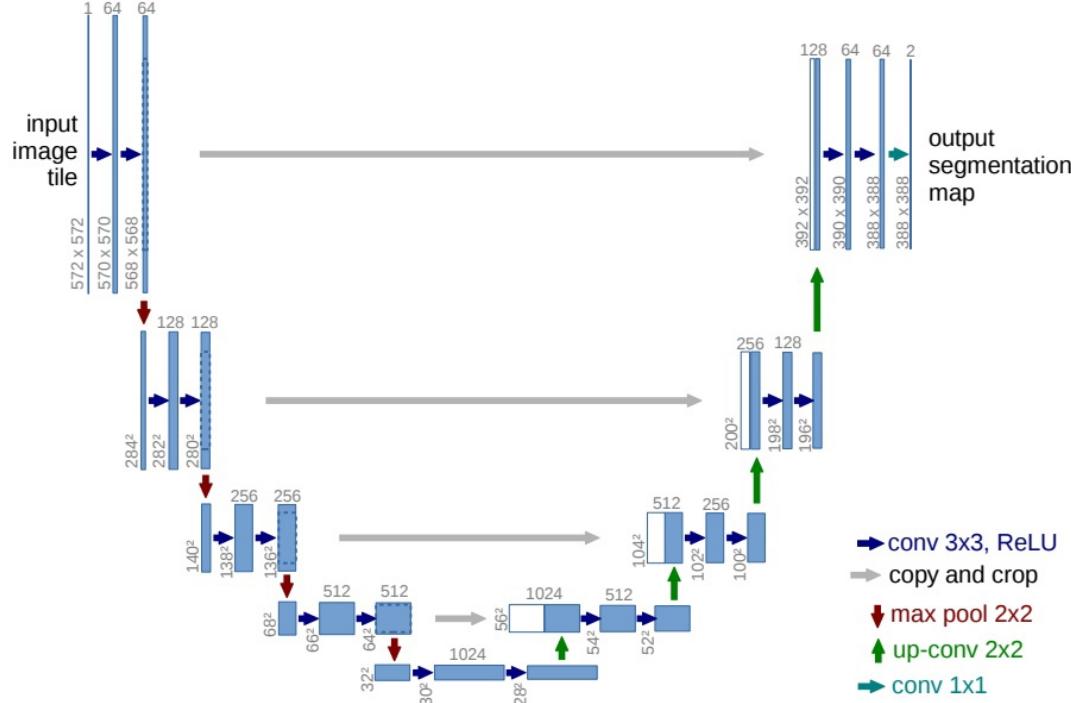


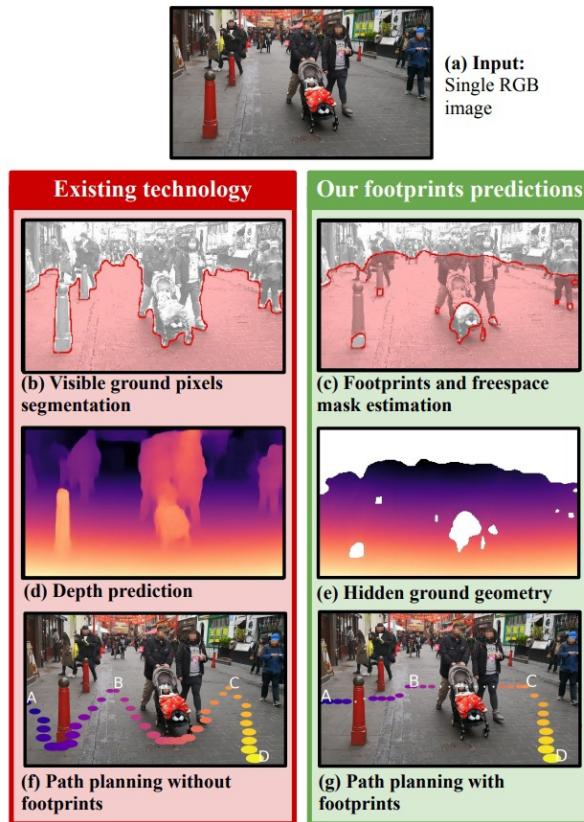
Рис. 7: Примеры масок для удаления фона [8]



3.1.4 Предсказание геометрии пространства

Не менее интересной задачей, в которой техника глубокого обучения также находит применение, является предсказание геометрии пространства, опираясь на изображение некоторой сцены. Авторы данной [9] работы предлагают модель, которая способна моделировать поверхности, которые даже скрыты от объектива камеры за другими объектами (экстраполяция). Исследователи провели сравнение между их подходом и традиционным, в котором модель сцены строилась только на основе видимой информации изображения и предсказания глубины (расстояния от точки на изображении до объектива камеры):

Рис. 8: Сравнение традиционного и предлагаемого [9] подходов, соответственно



3.1.5 Определение позы

Также одной из стандартных задач является задача определения позы человека. Например, для распознавания действий человека:

- лежит
- стоит
- бежит
- сидит

Или же для выявления нестандартного поведения человека в толпе.

Рис. 9: Результаты метода из статьи [10]



3.1.6 Распознавание лиц

Не менее успешных результатов глубокое обучение достигло и в сфере распознавания лиц

Рис. 10: Статистика использования технологии распознавания лиц в разных странах [11]



3.2. Обработка изображений

3.2.1 Генерация

Одной из важнейших задач машинного обучения является генерация некоторой синтетической выборки объектов, по свойствам и характеристика схожих с объектами генеральной совокупности. Ведь огромному классу алгоритмов необходимы данные для обучения, которые получить бывает очень трудно из-за различных ограничений:

- финансовых, если сбор данных требует денежных вложений
- временных, если сроки на сбор ограничены
- связанных с конфиденциальностью, например, паспортные данные, данные кредитных карт, транзакций и т.п.

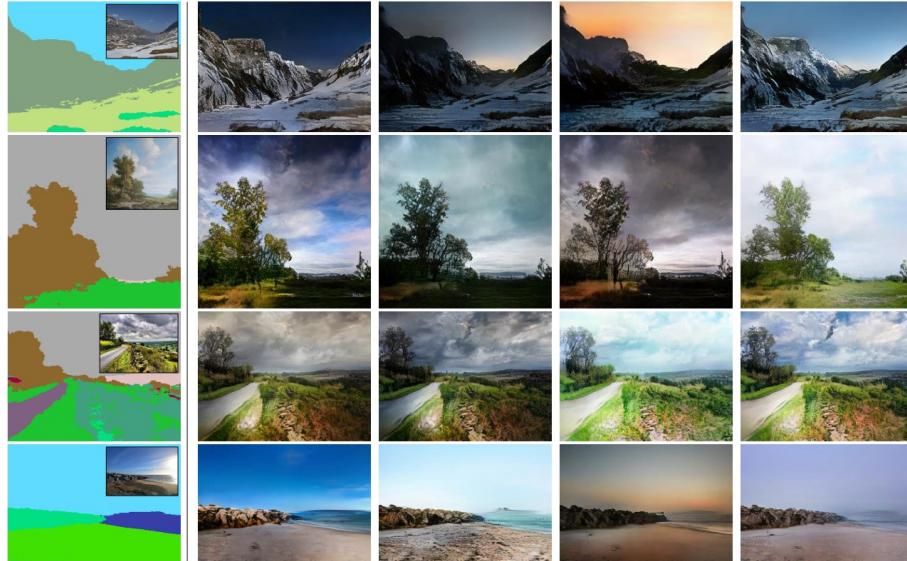
Так очень важной является задача генерации человеческих лиц, с которой на высшем уровне справилась компания Nvidia, описав свой подход в статье [12]. Они использовали генеративно-состязательную сеть (GAN), причём ключевая идея заключалась в постепенном росте как генератора, так и дискrimинатора: по мере прохождения обучения они добавляли всё новые и новые слои, что с каждым разом формировало всё более мелкие детали. Такое решение ускорило и стабилизировало обучение, по словам самих авторов.

Рис. 11: Примеры сгенерированных лиц алгоритмом [12]



Более того, разработчики из Nvidia предоставляют инструмент создания собственных реалистичных ландшафтов из рисунков, которые могут быть созданы в обычном Paint. Это [GauGAN](#).

Рис. 12: Ландшафты из примитивов. Из статьи [13]



Однако нейронные сети позволяют генерировать не только реальные изображения. Благодаря методам визуализации нейронных сетей, появились возможности генерировать абстрактные изображения [14], являющиеся комбинацией некоторых сущностей из реального мира.

Так раскрывается не только практическая, но и эстетическая сторона глубокого обучения.

Рис. 13: Примеры картин "написанных" нейронными сетями. Из статьи [14]



3.2.2 Интерполяция

Генерация изображений - это не просто игрушка, созданная трудами исследователей. Данное утверждение подтверждается работой [15]. Авторы статьи создали инструмент на основе свёрточных нейронных сетей, который моделирует интерполяцию в пространстве 3D моделей.

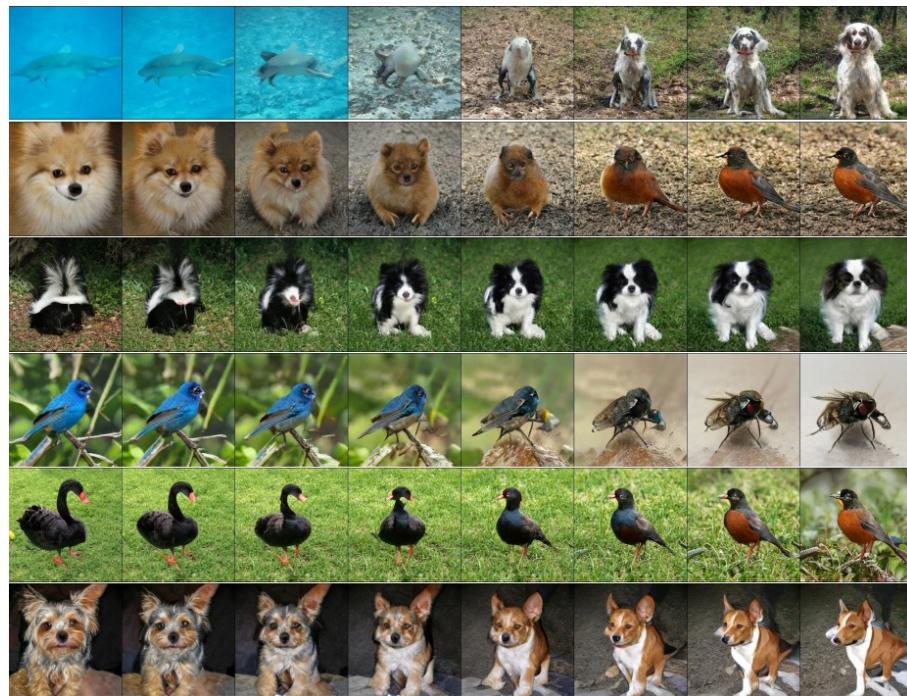
Это сильно облегчает работу дизайнераам, которые помимо заготовленных проектов могут демонстрировать промежуточные (интерполированные) проекты, являющиеся чем-то средним.

Рис. 14: Элементы декора и их различные интерполяции. Из статьи [15]



Аналогичным образом интерполяцию можно задать и в пространстве обычных изображений, что сделали авторы статьи [16]

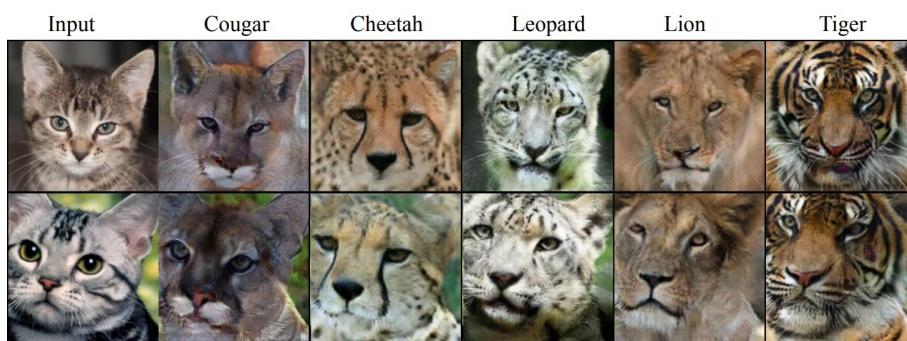
Рис. 15: Изображения животных и их различные интерполяции. Из статьи [16]



3.2.3 Трансляция

Трансляция или контролируемое изменение изображения позволяет перевести одно в "термины" другого. Так, например, авторы статьи [17] демонстрируют, как котёнок выглядит в телах различных кошачьих животных.

Рис. 16: Трансляция котят в образы различных кошачьих животных. Из статьи [17]



3.2.4 Повышение разрешения

Нейронные сети также нашли своё применение и в задачах повышения разрешения изображения (увеличение размера без потери качества). Так в статье [18] приводится сравнение подхода глубокого обучения с традиционными методами повышения разрешения:

Рис. 17: Сравнение результатов различных подходов повышения разрешения: бикубическая интерполяция, SRResNet, SRGAN. Справа оригинал



3.2.5 Стилизация

Задача стилизации изображения заключается в том, чтобы перенести стиль с одного на другое. То есть, имея пару: исходное изображение и изображение, выполненное в целевом стиле, - алгоритм должен сгенерировать стилизованное изображение.

Рис. 18: Демонстрация работы алгоритма [19]



3.2.6 Раскрашивание

Если нет необходимости менять стиль изображения, а достаточно просто перевести его из оттенков серого в реалистичные цвета, то нейронные сети справляются и с этим.

Рис. 19: Алгоритм [20] раскрашивает чёрно-белые изображения



3.2.7 Генерация анимаций

Удивительной является работа [21]. Авторы предлагают оживить статичное изображение.

Рис. 20: Алгоритм [21] анимирует персонажа статичного изображения



3.2.8 3D моделирование окружения

Глубокое обучение также позволяет создать 3d визуализацию помещения, опираясь на его изображение в 360°.

Рис. 21: 3D визуализация комнаты [22]



3.3. Понимание изображений

Нейронные сети исключительно хорошо справляются с задачей нахождения и генерирования паттернов. Поэтому у исследователей не было причин не воспользоваться этим преимуществом в задачах понимания изображений.

3.3.1 Аннотирование

Задача аннотирования схожа с задачей классификации, однако более сложна. Необходимо не просто назвать главный объект на изображении, нужно описать и его, и происходящие с ним в данном контексте.

Авторы статьи [23] демонстрируют ответы сети на некоторые входящие изображения. Как можно увидеть, в некоторых ситуациях модель довольно точна, однако также бывает, что ответ далёк от истины.

Рис. 22: Результаты аннотирования изображений [23]

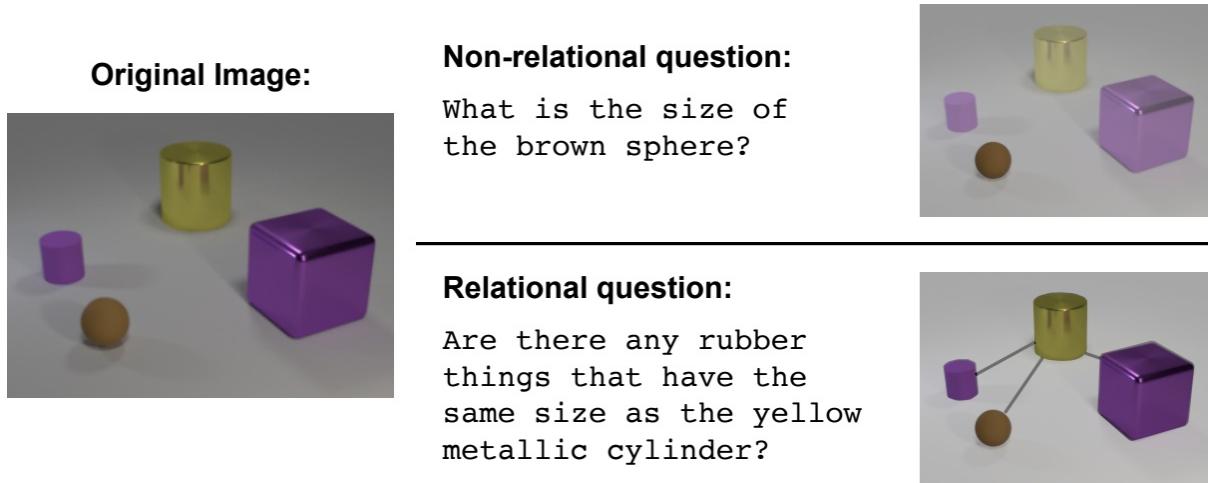


3.3.2 Visual Reasoning

Данная задача заключается в том, чтобы дать ответ на некоторый вопрос, связанный с контекстом исходного изображения.

Авторами статьи [24] была предложена сложная архитектура, использующая внутри себя как свёрточные, так и рекуррентные нейронные сети.

Рис. 23: Примеры вопросов из набора данных CLEVR [24]



3.3.3 Visual Explanations

Кроме того, в задачах? связанных с парой: изображение и описание, - интересно такое направление, как объяснение изображений (visual explanations). Данная задача является комбинацией классификации с некоторым пояснением.

Это может быть полезно, так как во многих алгоритмах классификации существует проблема интерпретации: не очевидна причина выбора метки для конкретного объекта. Однако, решив эту задачу, алгоритм сможет средствами естественного человеческого языка объяснить, почему он принял конкретное решение.

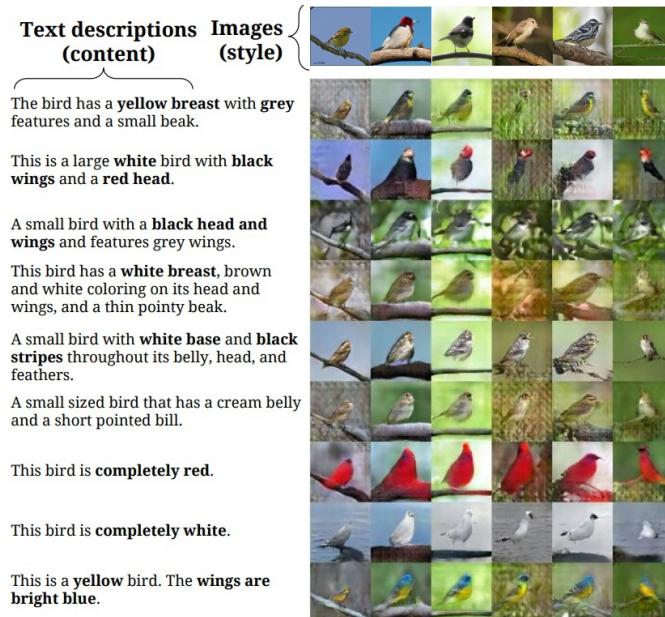
Рис. 24: Сгенерированные алгоритмом предложения для изображений [25]

<i>This is a Black-Capped Vireo because...</i>	<i>This is a White Pelican because...</i>
	
Description: this bird has a white belly and breast black and white wings with a white wingbar.	Description: this bird is white and black in color with a long curved beak and white eye rings.
Explanation-Dis: this is a bird with a white belly yellow wing and a black head .	Explanation: this is a large white bird with a long neck and a large orange beak .
<i>This is a Crested Auklet because...</i>	<i>This is a Geococcyx because...</i>
	
Description: this bird is black and white in color with a orange beak and black eye rings.	Description: this bird has a long black bill a white throat and a brown crown.
Explanation-Dis.: this is a black bird with a white eye and an orange beak.	Explanation-Dis.: this is a black and white spotted bird with a long tail feather and a pointed beak.
<i>This is a Green Jay because...</i>	<i>This is a Cape Glossy Starling because...</i>
	
Description: this bird has a bright blue crown and a bright yellow throat and breast.	Description: this bird is blue and black in color with a stubby beak and black eye rings.
Explanation-Dis.: this is a yellow bird with a blue head and a black throat .	Explanation-Dis.: this is a blue bird with a red eye and a blue crown.

3.4. Text2Image

Эта задача, в некотором смысле обратная задачам из нескольких предыдущих разделов. Имея некоторое описание контекста, алгоритм должен сформировать изображение.

Рис. 25: Сгенерированные алгоритмом [26] предложения для изображений



3.5. Обработка текста

Обработка текста (NLP) - область, железобетонно захваченная нейронными сетями.

3.5.1 Понимание смысла

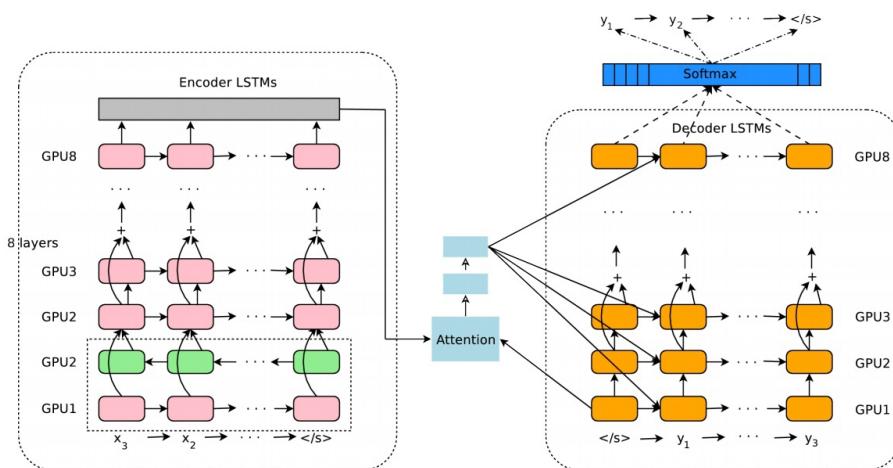
Эта задача аналогична задаче Visual Reasoning, только объектом выступает текст, а не изображение. То есть это задача о понимании того, какие объекты указаны в тексте, сколько их, чем различаются, как взаимодействуют и т.д.

3.5.2 Автоматический перевод

А данная задача заключается в том, чтобы перевести предложение с одного языка на другой. Учитывая все фразеологизмы, идиомы, образные выражения и т.п., эта задача становится очень сложной, ведь примитивный дословный перевод не просто неприятен для глаз человека, так ещё и может нести вредоносный характер: начиная от банального недопонимания, заканчивая негативно настроенной оскорбившейся стороной диалога из-за неправильно переведённого слова. Тем более часто перевод слова зависит от контекста предложения, в котором оно используется.

Поэтому для решения столь важной задачи взялась небезызвестная компания Google представившая миру нейронную систему машинного перевода [27].

Рис. 26: Архитектура модели от Google [27] для машинного перевода



3.5.3 Диалоговые системы

Также нейронные сети используются в разработке диалоговых систем. Основной проблемой диалога является то, что он должен семантически грамотно поддерживаться моделью. Она должна:

- понимать смысл предложения клиента

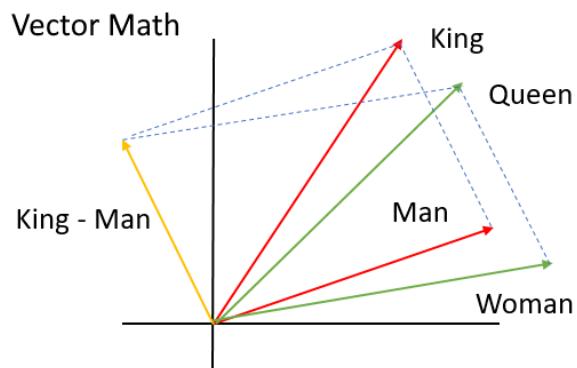
- не терять тему диалога, отвечая на очередной вопрос (существуют архитектуры обладающие памятью [28], [29], [30])
- в меру быстро отвечать

3.5.4 Представления слов, предложений, текстов ...

На самом деле, получение представлений - это отдельная тема (embeddings), однако наиболее исторически ярким событием в теории представлений являлось именно появление алгоритмов представлений слов (Word2Vec).

Как оказалось, все слова можно отобразить в некоторое многомерное пространство слов. Причём это отображение несёт за собой не только метрический смысл (возможность замерять расстояние между различными словами), но и алгебраический (арифметика над словами).

Рис. 27: Арифметика над векторами слов. Из статьи [31]



3.5.5 Генерация текста

В данной задаче модели машинного обучения должны продолжить текст, на основе уже имеющегося.

Довольно наглядным и интерактивным примером является модель Порфириевич [32].

Рис. 28: Пример продолженного предложения алгоритмом Порфириевич [32]

В данной задаче модели машинного обучения должны продолжить текст, на основе уже [созданных печатных пособий](#). Это позволит сократить круг слушателей лекций и сохранить при этом гибкость подготовки программного материала.

[▶ Дополнить \(Tab\)](#)

3.6. Обработка аудио

Подходы глубокого обучения не оставили без внимания и аудио-данные.

3.6.1 Распознавание речи

Ярчайшим представителем среди нейронных сетей, созданных, для распознавания человеческой речи, является модель распознавания внутри Yandex SpeechKit. Она использует нейронную сеть [33], обученную на массивах данных из различных сервисов Яндекса.

3.6.2 Генерация речи

Тот же Yandex SpeechKit предоставляет инструментарий для синтеза человеческой речи опять же на основе нейронных сетей [34].

Google не отстает и предлагает модель Tacotron 2 [35], на основе которого Nvidia, в свою очередь, разрабатывает Mellotron [36].

3.7. Обработка видео

На обработке статических изображений применение глубокого обучения не заканчивается.

3.7.1 Чтение по губам

Например, задача чтения по губам, которая подразумевает генерацию текста на основе движений губ и мышц лица, зафиксированных на видео, также имеет решение, описанное в статье [37]

3.7.2 Deepfake

Технология deepfake – это методика синтеза аудио или видео с использованием генеративно-состязательных нейронных сетей. Впервые такая нейросеть была создана в 2014 году студентом Стэнфордского университета Яном Гудфеллоу. Он автоматизировал процесс обучения двух нейронных сетей, когда одна из них генерирует лица или голоса, а другая анализирует полученные результаты и даёт заключение: похожи они на оригинал, или нет.

Рис. 29: Пример deepfake, в котором лицо актрисы заменяется лицом актёра [38]

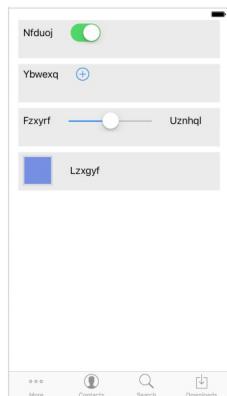


3.8. Автоматизация программирования

Исследователи со временем начинают внедрять нейронные сети в область программирования. А если более точно - написание кода. Об этом свидетельствует работа [39], в которой авторы взялись за задачу конвертирования между различными языками программирования.

Также существует подход [40], при котором генерируется код, который воссоздаёт графический интерфейс пользователя (GUI), изображённый на скриншоте.

Рис. 30: Сгенерированный код графического интерфейса с использованием технологии pix2code [40]



(a) iOS GUI screenshot

```
stack {
    row {
        label, switch
    }
    row {
        label, btn-add
    }
    row {
        label, slider, label
    }
    row {
        img, label
    }
}
footer {
    btn-more, btn-contact, btn-search, btn-download
}
```

(b) Code describing the GUI written in our DSL

3.9. Искусственный интеллект в играх

Искусственный интеллект основанный на глубоком обучении способен превзойти человека в такой развлекательной сфере, как игры. Это упорно доказывает Google Deepmind.

В 2013 она разрабатывает алгоритм, способный играть в классические игры 70-80ых годов для игровой консоли Atari 2600. Причём в 22е игры модель играла лучше человека, не обладая даже знаниями о правилах самих игр. Она использовала только изображение на входе и информацию об очках.

Следующее достижение - победа над сильнейшим игроком в мире в игре го Ли Седолем. Сам матч окончился со счётом 4 - 1 в пользу машины. Алгоритму дали название AlphaGo. Программа разрушила устоявшееся мнение о том, что ещё долгое время компьютеры не смогут играть в Го сильнее человека, так как дерево возможных партий слишком велико для метода перебора.

Затем линейка AlphaGo развивалась вплоть до современного MuZero, которому для обучения не нужны ни базы партий людей, ни эмпирические знания, программируемые людьми, ни даже правила самой игры.

Рис. 31: Эволюция AlphaGo [41]

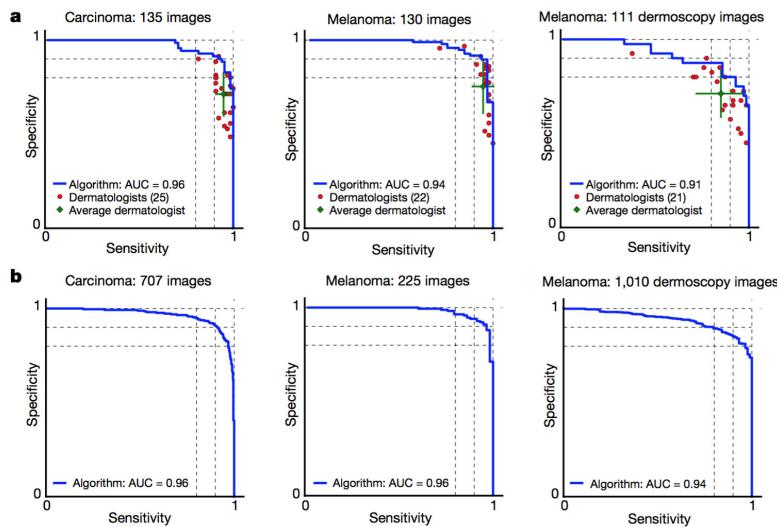


3.10. Медицина

Сложно представить сферу научных знаний более важную, чем медицина, ведь человеческое здоровье всегда на первом месте. От нейросетей ждут серьезных достижений в этой области — это и диагностика различных заболеваний, и разработка новых лекарств, и новые методы лечения.

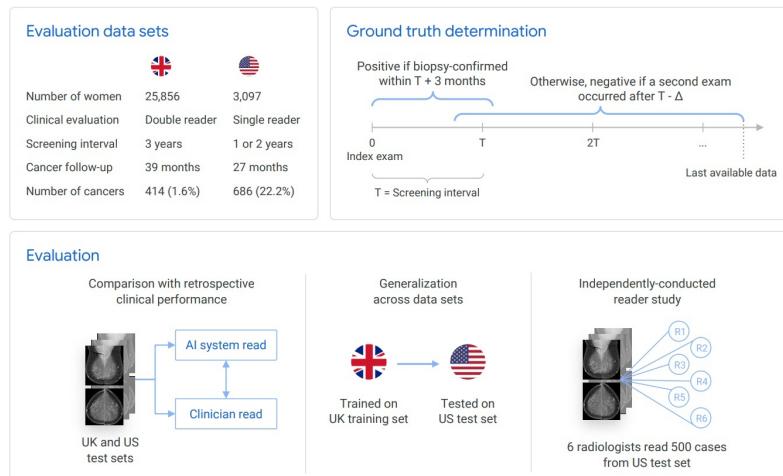
Авторы статьи [42] уже радуют своими достижениями в диагностировании рака кожи. Также работа [43] публикует результаты применения глубокого обучения в

Рис. 32: Классификация раковых опухолей. Из статьи [42]



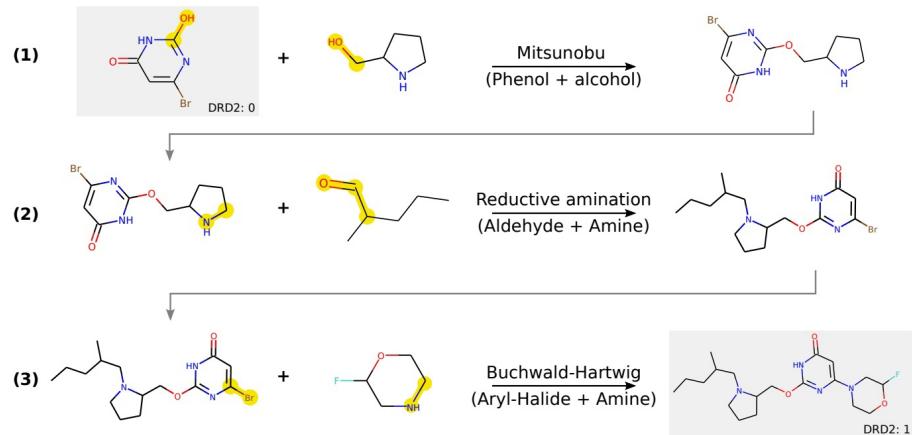
задаче скрининга мамограмм.

Рис. 33: Скриннинг мамограмм [43]



Интересно, что обучение с подкреплением, используемое в AlphaGo, полезно не только в играх. С помощью Reinforcement Learning авторы статьи [44] предсказывают свойства химических соединений, что положительно скажется на рынке создания лекарств.

Рис. 34: Предсказание свойств химических соединений [44]



3.11. Продажи

Бизнес - деятельность, направленная на систематическое получение прибыли. Высокая эффективность методов глубокого обучения не могла не привлечь людей, заинтересованных в открытии новых методов увеличения доходов. Как известно, спрос рождает предложение, что и побудило исследователей заняться вопросом применения нейронных сетей в продажах.

Поведение клиента описывает путь, который проходит потребитель, когда он исследует, выбирает и покупает продукт или услугу. Многие вещи влияют на поведение клиентов, к примеру:

1. Предыдущие покупки
2. Образ жизни
3. Культура
4. Образование
5. Занятость

Почему важно прогнозировать поведение потребителей? Когда компания предсказывает, как будут вести себя клиенты, перед ней открываются возможности:

- Сокращение оттока клиентов
- Выявление и таргетирование ценных клиентов
- Удовлетворение потребительского спроса
- Быстрый вывод продуктов на рынок
- Сокращение расходов на маркетинговую кампанию
- Персонализация клиентского опыта
- Улучшение качества обслуживания клиентов

3.11.1 Next-Item Prediction

Клиент - это всего лишь сущность, которая в некоторый момент времени выбирает товар и покупает его. Поэтому для предсказания поведения потребителя может быть достаточно предсказать, какой товар он купит, если уже известно какие товары он купил до этого. Эта задача о предсказании следующего товара (Next-Item Prediction).

Существует множество различных подходов [45], [46], [47].

3.11.2 Next-Basket Prediction

Однако клиент обычно не покупает один товар. Он набирает некоторую корзину товаров, затем оплачивает её содержимое целиком. Более того, корзина товаров содержит больше информации о вкусах и предпочтениях клиента. Поэтому имеет место задача о предсказании содержимого очередной корзины потребителя, на основе предыдущих корзин (Next-Basket Prediction).

Здесь подходы чаще всего предлагают использовать рекуррентные нейронные сети и сети с механизмом внимания [48], [49], [50].

3.12. Неожиданные применения нейронных сетей

Нейронные сети используют математику, как инструмент. Однако ведь можно использовать и нейронные сети, как инструмент для новых открытий в математике.

Так авторы статьи [51] создали алгоритм, который генерирует так называемые цепные дроби. У искушенные математикой сразу же возникли ассоциации с великим математиком Сринивасой Рамануджаном. Возникающие на выходе дроби весьма красивы, однако многие из них до сих пор не доказаны. В конце статьи указан перечень цепных дробей с пометкой, доказаны ли они или ещё нет. Более того, некоторые из этих дробей уже были известны науке.

Рис. 35: Некоторые равенства, которые сгенерировала сеть (все они верные и доказательства приведены в конце оригинальной статьи) [51]

$$\frac{4}{3\pi - 8} = 3 - \frac{1 \cdot 1}{6 - \frac{2 \cdot 3}{9 - \frac{3 \cdot 5}{12 - \frac{4 \cdot 7}{\dots}}}}$$

$$\frac{e}{e - 2} = 4 - \frac{1}{5 - \frac{2}{6 - \frac{3}{7 - \frac{4}{\dots}}}}$$

$$\frac{2}{\pi + 2} = 0 - \frac{1 \cdot (3 - 2 \cdot 1)}{3 - \frac{2 \cdot (3 - 2 \cdot 2)}{6 - \frac{3 \cdot (3 - 2 \cdot 3)}{9 - \frac{4 \cdot (3 - 2 \cdot 4)}{\dots}}}}$$

$$\frac{1}{e - 2} = 1 + \frac{1}{1 + \frac{-1}{1 + \frac{2}{1 + \frac{-1}{1 + \frac{3}{\dots}}}}}$$

Исследователи из Google Deepmind уже разработали программу, которая находит лучшие (возможно) ходы в шахматной позиции. А что если научить нейронную сеть предсказывать плохие ходы? Точнее - научить её играть, как несильные игроки. Это поможет любителям играть с машиной своего уровня силы. Именно это и сделали создатели модели Maia [52], внедрив свою технологию на всемирный шахматный онлайн-сервер Lichess. Здесь игроки могут сами выбрать уровень силы Maia (так называемый, рейтинг) и начать играть так, как будто с ними играет настоящий человек, а не робот.

Другое интересное применение для нейронных сетей нашла уже рассмотренная ранее компания Nvidia. Она предлагала превращать примитивные рисунки в фо-

тореалистичные. Теперь их новая модель GANcraft [53] делает то же самое, но с трёхмерными блочными мирами, подобными мирам известной игры Minecraft.

Рис. 36: Фотореалистичные миры от Nvidia на основе Minecraft [53]



4. Заключение

Подводя итоги, хочется сказать, что сегодня глубокое обучение, нейронные сети и всё, что с ними связано уже не являются уделом небольшой группы теоретиков. В нейросетевых технологиях заинтересованы инженеры и исследователи разных специальностей. Не может не радовать прогресс в построении удачных моделей глубокого обучения исследуемых явлений, полностью базирующихся на экспериментальных данных. Здесь наиболее полно проявляются замечательные свойства искусственных нейронных систем:

- массивная параллельность обработки информации
- ассоциативность памяти
- способность к обучению на опыте

Это открывает новые перспективы для систематизации многочисленной экспериментальной информации в таких областях знаний, где традиционно трудно приживаются математический формализм, например, в медицине, психологии и истории.

Список литературы

- [1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 25, 01 2012.
- [2] PapersWithCode. ImageNet Benchmark. [https://paperswithcode.com/sota/
image-classification-on-imagenet](https://paperswithcode.com/sota/image-classification-on-imagenet).
- [3] Joseph Redmon and Ali Farhadi. Yolo9000: Better, faster, stronger, 2016.
- [4] Piotr Dollar. Learning to Segment. <https://research.fb.com/learning-to-segment/>, 08 2016.
- [5] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation, 2016.
- [6] Augustus Odena, Vincent Dumoulin, and Chris Olah. Deconvolution and checkerboard artifacts. *Distill*, 2016.
- [7] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.
- [8] Yunan Wu. Go Selfies — How to do photo background removal using deep learning(segmentation + codes). [https://medium.com/@yunanwu2020/
go-selfies-how-to-do-photo-background-removal-using-deep-learning-segmentation-codes-6131d59a08](https://medium.com/@yunanwu2020/go-selfies-how-to-do-photo-background-removal-using-deep-learning-segmentation-codes-6131d59a08) 2019.
- [9] Jamie Watson, Michael Firman, Aron Monszpart, and Gabriel J. Brostow. Footprints and free space from a single color image, 2020.
- [10] Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh. Convolutional pose machines, 2016.
- [11] Visual Capitalist. Face Recognition World Map. [https://www.visualcapitalist.com/
wp-content/uploads/2020/05/Facial-Recognition-World-Map-Full-Size.html](https://www.visualcapitalist.com/wp-content/uploads/2020/05/Facial-Recognition-World-Map-Full-Size.html).
- [12] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation, 2018.
- [13] Taesung Park, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. Semantic image synthesis with spatially-adaptive normalization, 2019.
- [14] Alexander Mordvintsev, Christopher Olah, and Mike Tyka. Inceptionism: Going deeper into neural networks. [https://ai.googleblog.com/2015/06/
inceptionism-going-deeper-into-neural.html](https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html), 06 2015.

- [15] Alexey Dosovitskiy, Jost Tobias Springenberg, Maxim Tatarchenko, and Thomas Brox. Learning to generate chairs, tables and cars with convolutional networks, 2017.
- [16] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale gan training for high fidelity natural image synthesis, 2019.
- [17] Ming-Yu Liu, Thomas Breuel, and Jan Kautz. Unsupervised image-to-image translation networks, 2018.
- [18] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network, 2017.
- [19] Fujun Luan, Sylvain Paris, Eli Shechtman, and Kavita Bala. Deep photo style transfer, 2017.
- [20] Richard Zhang, Phillip Isola, and Alexei A. Efros. Colorful image colorization, 2016.
- [21] Chung-Yi Weng, Brian Curless, and Ira Kemelmacher-Shlizerman. Photo wake-up: 3d character animation from a single photo, 2018.
- [22] Chuhang Zou, Alex Colburn, Qi Shan, and Derek Hoiem. Layoutnet: Reconstructing the 3d room layout from a single rgb image, 2018.
- [23] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator, 2015.
- [24] Adam Santoro, David Raposo, David G. T. Barrett, Mateusz Malinowski, Razvan Pascanu, Peter Battaglia, and Timothy Lillicrap. A simple neural network module for relational reasoning, 2017.
- [25] Lisa Anne Hendricks, Zeynep Akata, Marcus Rohrbach, Jeff Donahue, Bernt Schiele, and Trevor Darrell. Generating visual explanations, 2016.
- [26] Scott Reed, Zeynep Akata, Xinchen Yan, Laajanugen Logeswaran, Bernt Schiele, and Honglak Lee. Generative adversarial text to image synthesis, 2016.
- [27] Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Lukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, and Jeffrey Dean. Google’s neural machine translation system: Bridging the gap between human and machine translation, 2016.
- [28] Alex Graves, Greg Wayne, and Ivo Danihelka. Neural turing machines, 2014.

- [29] Wojciech Zaremba and Ilya Sutskever. Reinforcement learning neural turing machines - revised, 2016.
- [30] Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, and Rob Fergus. End-to-end memory networks, 2015.
- [31] Edward Ma. From conte to entity type embeddings in natural language processing. <https://laptrinhx.com/from-conte-to-entity-type-embeddings-in-natural-language-processing-708004840/>, 04 2020.
- [32] Порфирьевич. <https://porfirevich.ru/>.
- [33] Yandex. Модели распознавания. <https://cloud.yandex.ru/docs/speechkit/stt/models>.
- [34] Yandex. Синтез речи. <https://cloud.yandex.ru/docs/speechkit/tts/>.
- [35] Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio, 2016.
- [36] Rafael Valle, Jason Li, Ryan Prenger, and Bryan Catanzaro. Mellotron: Multispeaker expressive voice synthesis by conditioning on rhythm, pitch and global style tokens, 2019.
- [37] Joon Son Chung, Andrew Senior, Oriol Vinyals, and Andrew Zisserman. Lip reading sentences in the wild, 2017.
- [38] Debankita Basu. Deepfake: Trivial or not? <https://medium.com/deep-learning-digest/deepfake-trivial-or-not-5b9e7c9ee080>, 01 2021.
- [39] Marie-Anne Lachaux, Baptiste Roziere, Lowik Chanussot, and Guillaume Lample. Unsupervised translation of programming languages, 2020.
- [40] Tony Beltramelli. pix2code: Generating code from a graphical user interface screenshot, 2017.
- [41] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, and David Silver. Muzero: Mastering go, chess, shogi and atari without rules. <https://medium.com/deep-learning-digest/deepfake-trivial-or-not-5b9e7c9ee080>, 12 20.
- [42] Andre Esteva, Brett Kuprel, Roberto Novoa, Justin Ko, Susan Swetter, and Helen Blau. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 01 2017.

- [43] Scott Mayer McKinney, M. Sieniek, Varun Godbole, Jonathan Godwin, N. Antropova, H. Ashrafi, T. Back, Mary Chesnut, Greg C. Corrado, A. Darzi, M. Etemadi, Florencia Garcia-Vicente, F. Gilbert, M. Halling-Brown, D. Hassabis, S. Jansen, A. Karthikesalingam, Christopher J. Kelly, Dominic King, J. Ledsam, David S. Melnick, H. Mostofi, Lily Peng, J. Reicher, B. Romera-Paredes, R. Sidebottom, Mustafa Suleyman, Daniel Tse, K. Young, J. Fauw, and S. Shetty. International evaluation of an ai system for breast cancer screening. *Nature*, 577:89–94, 2020.
- [44] Julien Horwood and Emmanuel Noutahi. Molecular design in synthetically accessible chemical space via deep reinforcement learning. *ACS Omega*, 5(51):32984–32994, Dec 2020.
- [45] Jianqing Zhang, Dongjing Wang, and Dongjin Yu. TLSAN: time-aware long-and short-term attention network for next-item recommendation. *Neurocomputing*, 441:179–191, 2021.
- [46] Bin Wu, Xiangnan He, Zhongchuan Sun, Liang Chen, and Yangdong Ye. ATM: an attentive translation model for next-item recommendation. *IEEE Trans. Ind. Informatics*, 16(3):1448–1459, 2020.
- [47] Chaoyue He, Yong Liu, Qingyu Guo, and Chunyan Miao. Multi-scale quasi-rnn for next item recommendation. *CoRR*, abs/1902.09849, 2019.
- [48] Tong Liu, Xianrui Yin, and Weijian Ni. Next basket recommendation model based on attribute-aware multi-level attention. *IEEE Access*, 8:153872–153880, 2020.
- [49] Binbin Che, Pengpeng Zhao, Junhua Fang, Lei Zhao, Victor S. Sheng, and Zhiming Cui. Inter-basket and intra-basket adaptive attention network for next basket recommendation. *IEEE Access*, 7:80644–80650, 2019.
- [50] Haoji Hu, Xiangnan He, Jinyang Gao, and Zhi-Li Zhang. Modeling personalized item frequency information for next-basket recommendation, 2020.
- [51] Gal Raayoni, Shahar Gottlieb, Yahel Manor, George Pisha, Yoav Harris, Uri Mendlovic, Doron Haviv, Yaron Hadad, and Ido Kaminer. Generating conjectures on fundamental constants with the ramanujan machine. *Nature*, 590(7844):67–73, Feb 2021.
- [52] Reid McIlroy-Young, Siddhartha Sen, Jon Kleinberg, and Ashton Anderson. Aligning superhuman ai with human behavior. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, Jul 2020.
- [53] Zekun Hao, Arun Mallya, Serge Belongie, and Ming-Yu Liu. Gancraft: Unsupervised 3d neural rendering of minecraft worlds, 2021.