

# ГЛАВА 2

# БОГИ ПОД МАСКАМИ

РАССКАЗ ПЕРЕВЕДЕН ЭМИЛИ ДЖИН

*И правда, и утро со временем проясняются.*

Африканская пословица

**Примечание Кай-Фу:** В этой истории рассказывается о молодом нигерийском видеопродюсере, которого наняли для создания серьезного дипфейка\*. Одним из основных направлений ИИ является так называемое компьютерное зрение — оно учит машины «видеть», и недавние прорывы в этой сфере позволяют получать невиданные доселе результаты. Этот рассказ — фантазия о мире будущего, характерной чертой которого стали беспрецедентные высокотехнологические игры в стиле «кошки-мышки» между мошенниками и их разоблачителями; между преступниками и потенциальными жертвами. Можем ли мы избежать такого мира, в котором все визуальные линии размыты и нечетки? Я исследую этот вопрос в своем комментарии: я опишу недавние и предстоящие открытия в области компьютерного зрения, биометрии и безопасности ИИ — трех технологических сферах, которые позволяют создавать подделки-дипфейки и другие подобные технологии.

---

\* Дипфейк (от deep learning [глубокое обучение] + фейк) — методика синтеза изображения, основанная на искусственном интеллекте. В видеороликах-дипфейках лицо одного человека подменяют лицом другого. *Прим. ред.*

**Поезд наземного метро** подтягивался к станции Яба, и Амака нажал кнопку открытия дверей. Они с визгом разъехались; парень, не дожидаясь полной остановки, спрыгнул на перрон. Он больше не мог вынести в этом черепашьем поезде, в его спертom воздухе, ни одного мгновения. Почти прижавшись к какому-то пожилому мужчине, Амака ловко проскользнул через турникет на выходе из вокзала — он опять проехался зайцем.

**Система распознавания лиц** автоматически списывала стоимость проезда со счета каждого проходящего мимо камер человека. Но Амака был в маске — камеры его не увидели, и он опять прокатился бесплатно.

Маски давно стали обычным аксессуаром для молодежи Лагоса\*. Для их родителей они были скорее ритуальными атрибутами, но для молодых людей, которых в последние десятилетия заметно прибавилось, маски стали модным трендом, а заодно и средством обмана камер наблюдения.

В Лагосе, крупнейшем городе Западной Африки, жило примерно от 27 до 33 миллионов человек — каждый метод подсчета давал свою цифру. Пять лет назад государство строго ограничило число приезжих, считая мигрантами даже тех, кто родился в других частях Нигерии. И странствующим мечтателям вроде Амаки пришлось постоянно находиться в состоянии поиска убежища — организовывать пристанище в нелегальных квартирах, в общежитиях, на рынках, на автобусных станциях или даже под эстакадами.

Амака встречал много бездомных, оказавшихся на улице по самым разным причинам. Дома одних снесли, чтобы расчистить место для новых торговых центров; другие недавно приехали в Нигерию из еще более бедных стран; третьи были просто нищими бродягами. Нигерия — «молодая» страна, средний возраст жителей всего 21 год: результат высокой рождаемости. Но в выигрыше от этого оказались далеко не все граждане третьего в мире по численности населения государства.

Некоторые части Лагоса задыхались и страдали от избытка молодежи, но район Яба процветал. Он получил название «Кремниевая долина Западной

---

\* Столица Нигерии до 1991 года. Крупнейший город страны и Африки. Население (в агломерации) — около 21 миллиона человек. *Прим. ред.*

Африки» и выгодно выделялся порядком, свежим воздухом и бытом, насыщенным высокими технологиями. Пешеход тут мог активировать мультяшного персонажа на рекламном щите и взаимодействовать с ним с помощью жестов. Роботы-уборщики неустанно шныряли по улицам, собирая и сортируя мусор — и район был невероятно чистым. Бытовые отходы отправляли в центры переработки и превращали в возобновляемые материалы и биотопливо. Модным трендом, а не только строительным материалом, вдруг стал экологически чистый бамбук — по крайней мере, для жителей Ябы.

Выйдя из вокзала, Амака поднял смартстрим на уровень глаз и наложил активную виртуальную карту на окружающий пейзаж. Следуя по проложенному маршруту, он добрался до тихого переулочка. Компания Ljele, которую он искал, находилась, судя по всему, на третьем этаже серого дома с номером 237.

Два дня назад он получил из Ljele загадочный анонимный мейл с предложением работы, которая, как сообщалось в послании, наверняка «придется ему по душе». Место уже практически его, было написано дальше, но надо лично явиться на собеседование — это пустая формальность, таков общий порядок.

Амака вошел в небольшую приемную на третьем этаже, и улыбчивая администраторша тут же указала на маску — ее надо было снять для проверки личности. Молодой человек на секунду замешкался, но команду выполнил. В объектив камеры попало юное лицо с гладкой кожей.

Амака носил самую простую маску, напечатанную на 3D-принтере. Изысканные дорогие модели ручной работы туристы с удовольствием раскупали как сувениры на рынке Лекки — они стоили абсурдно дорого. Но и дешевая маска Амака неплохо дурила большинство камер наблюдения в городе.

Благодаря грубой поделке с узором в форме бабочек Амака был в понимании искусственного интеллекта «человеком без лица», и это парня очень устраивало. Маска не только позволяла ему экономить, но и, что куда важнее, защищала от преследования властей — у него еще не было официального вида на жительство для мигрантов.

Администраторша провела Амаку в конференц-зал. Парень сел на стул и замер, обдумывая, что говорить о предыдущем опыте работы. *Придется врать*, решил он. *Тут, кажется, без вариантов.*

Прошло минут десять. Обещанный интервьюер все не появлялся. И вдруг прямо напротив Амаки засветилась проекционная стена — началось воспроизведение видеозаписи камеры наблюдения.

С первых кадров Амака узнал запись — он помнил ее мельчайшие подробности. Полночь. Тусклый желтый свет фонарей. Несколько бездомных лежат под эстакадой на самодельных матрасах, набитых мусором. Из темноты возникает щуплый силуэт. Мальчик подходит к спящим и всматривается в их лица.

Наезд камеры. Мальчику не больше пяти-шести лет, белый, одет в полосатую пижамку, личико бледное и невыразительное. Один из спящих, вздрогнув,

просыпается и встречается взглядом с мальчиком. Спрашивает, как его зовут и что он тут делает. Мальчик что-то бессвязно бормочет и дрожит всем телом.

Внезапно его лицо искажается жуткой гримасой, он разевает рот — обнажаются два ряда острых зубов. И вдруг ребенок вонзает клыки в шею бедного бездомного! Мужчина кричит от боли — просыпаются остальные. По губам и подбородку мальчика течет кровь, он отступает и быстро скрывается в темноте.

Видео было выложено в интернете и называлось «Белый вампир нападает на бездомных в Лагосе». В первые сутки ролик набрал миллионы просмотров на платформе обмена видео GаgгiV. Но через несколько дней платформа идентифицировала видео как фейк и удалила — как того и требовал закон. Аккаунт пользователя, загрузившего файл, Enitan0231, заблокировали, а все доходы от рекламы, связанные с роликом, заморозили.

Видео закончилось, Амака продолжал сидеть в одиночестве, но в комнате раздался громкий голос: «Отличная работа, Амака! Безупречное слияние мизансцены, актеров-любителей и съемки в реальном времени! Даже не верится, что это сделано в подпольной интернет-кафешке в Икедже», — невидимый мужчина говорил с сильным акцентом игбо\*.

От неожиданности Амака вскочил и с криком «Вы кто?!» инстинктивно начал озираться. Быстро обшарил взглядом пустую комнату и остановился на динамиках.

— Да расслабься, — хохотнул невидимый незнакомец, — можешь называть меня Чи. Так нужна тебе работа или нет?

Амака вздохнул, опустил на стул и сник. Чи был прав. Без вида на жительство ему никогда не найти в Лагосе нормальную работу. Таинственная компания под названием Ljele сегодня была его единственной надеждой.

— Почему вы выбрали меня? — спросил Амака.

— Мы видели твои работы. Ты талантливый и амбициозный, ты никогда бы не поехал в Лагос, если бы не жаждал прославиться. А нам нужен человек, которому можно доверять. *Один из наших.*

Амака сразу понял, что имел в виду таинственный Чи. В Нигерии больше 250 этнических групп, каждая со своими языками и обычаями, многие конфликтуют уже не первую сотню лет. Йоруба и игбо — вторая и третья по величине этнические группы в стране, именно их стычки в последние годы были особо жестокими, поскольку обе народности рвутся к власти.

В Лагосе преобладали йоруба, Амака был игбо с юго-востока, поэтому обычно не афишировал свою этническую принадлежность — кому нужны неприятности?

— И что за работа? Что надо делать? — Амака уже практически принял предложение, а вопросы — так, для проформы. Ответы были ему не очень важны.

— То, что у тебя лучше всего получается, — снимать фейковое видео.

— Это, надо полагать, незаконно?

\* Язык, распространенный на юго-востоке Нигерии. *Прим. пер.*

— Мы обеспечим тебя всем необходимым.

Зрачки Амаки сузились, ноздри раздулись.

— А если я откажусь? — он еще продолжал изображать сомнения. — Вы меня убьете?

— Убьем? — Чи опять издал какой-то странный смешок. — О, нет! Мы сделаем кое-что похуже.

И на проекционной стене опять началось видео.

Частный ночной клуб. Снимает камера в углу под потолком; она направлена на центр танцпола и берет крупный план. Несколько полубоуженных парней (на них только штаны — рубашек или футболок нет) томно извиваются под мигающими лазерными вспышками. План еще более укрупняется.

Один из парней — Амака. Он, вдруг извернувшись всем голым торсом, страстно целует соседа — другого парня, щеки которого светятся флуоресцентно-розовым светом. Следующий поцелуй Амаки не без труда (приходится изогнуться еще сильнее) достается темнокожему, танцующему позади. Стоп-кадр: три молодых лица, похожих на листья манго, наложены друг друга: тела переплетены — они буквально слились.

Амака смотрел на стену не мигая — лицо его ничего не выражало. Прошло несколько бесконечных секунд, и он усмехнулся. Безусловно, это дипфейк, только что сляпанный на скорую руку, — а «сырьё» взялось с камеры на стойке регистрации, ведь Амаке пришлось позволить просканировать свое лицо.

— Неплохо. Лицо, возможно, и мое, но вот с шеей вы лажанулись, — Амака стянул капюшон. Длинный розовый шрам шел от правого уха до левой ключицы. Память об одной уличной драке. — А потом, мы в Лагосе, здесь люди делают и вещи и побезумнее!

— Разумеется, бывает и похуже, — парировал Чи, — но это видео все равно может отправить тебя за решетку. И кстати, о семье подумай.

Амака подавленно замолчал. Любые проявления гомосексуальности караются в Нигерии серьезными тюремными сроками, с момента принятия в 2013 году закона о запрете однополых браков миновало несколько десятков лет, а нигерийское общество относится к сексуальным и гендерным меньшинствам враждебно и непримиримо — за десятилетия ничего не изменилось. Амаки отлично знал: если на него стукнут, то последует обвинение в уголовном преступлении; но даже если и нет, проблем с полицией точно не избежать. И под угрозой тюрьмы у него наверняка станут вымогать деньги.

Да, ведь еще родные... В последние несколько лет отношения Амаки с родственниками стали прохладными, но он не мог даже помыслить обрушить на их головы такую жесть. *И плевать, что видео фейковое.* Особенно жалко было отца — он возлагал на Амаку большие надежды.

Парень натянул капюшон и закусил нижнюю губу — создал иллюзию хоть какой-то защищенности.

— Мне нужен аванс. В криптовалюте. И подробности — что за объект? Не хочу тратить время на расследование.

— Как скажешь, дружище. Что же до цели... Ну, по этой точно не промажешь.

На проекционной стене возник нерезкий кадр. Когда проявились черты лица, Амака тихонечко присвистнул.

61

**Йоруба называют Лагос** «Еко», в переводе с их языка — «ферма». В Лагосе экваториальный муссонный климат, поэтому июнь — самый холодный месяц года с обильными ливнями. Амака лежал на узкой кровати в нелегально снятой комнате в общежитии и слушал звуки дождя. Новая работа была совершенно другого уровня, чем все предыдущие занятия, — детские шалости, похоже, закончились.

Амака надел XR-очки\* и принялся разбираться с новым гаджетом — темно-зеленым Illumiware Mark-V.

Новая работа представляла сложность совсем не из-за недостатка опыта в создании видеофейков. Все было как раз наоборот. В прошлом году Амака много вечеров выдавал себя на сайтах знакомств за девушек из пригорода. Чтобы не проколотся, следовало собрать как можно больше видеоданных с помощью робота-поисковика. Идеальными целями для Амаки были юные модницы йоруба в ярких кофтах-*бубах* с глубоким V-образным вырезом, в юбках-*иро* — экстремальном мини, по сути, в небольшом куске ткани, почти ничего не скрывавшем, и с волосами, убранными под платок-*геле*.

Желательно, чтобы видео снимали в комнатах с ярким и стабильным освещением — нужны четкие лица. Это позволяет искусственному интеллекту сделать достаточное количество цифровых стоп-кадров. Затем датасет объекта соединялся с набором данных по лицу самого Амаки — при другом освещении, под разными углами и с разнообразными выражениями, которые автоматически генерировал его смартстрим.

Далее Амака загружал оба датасета в облако и приступал к работе с гипергенеративно-состязательной нейросетью. Спустя несколько часов или дней он получал результат — модель DeepMask, сотканную из алгоритмов. Применив ее в видеоприложении, он мог «превратиться» в девушку, которую сам сложил из битов информации, — отличить этот дипфейк от настоящего видео невооруженным глазом было совершенно невозможно.

Если позволяла скорость интернета, Амака мог развлечения ради менять лица в реальном времени. Но чтобы было весело, приходилось попотеть. Обман в реальном времени удавался, если синхронно переводить с английского или игбо на йоруба, использовать прогу transVoice для имитации голоса девушки

\* От XR (eXtended Reality — «расширенная реальность») — технология, позволяющая погрузиться в виртуальный мир. *Прим. ред.*

йоруба и инструменты с открытым исходным кодом для синхронизации движения губ со звукорядом.

62 Однако если человек в чате раскошелится на высококачественный детектор дипфейков, то аномалии в видео автоматически выявлялись и помечались красными полупрозрачными квадратиками-предупреждениями.

На заре появления технологии дипфейков воспроизведение могло сбоить из-за недостаточной скорости интернета или преувеличенной мимики — изображение размывалось, губы двигались несинхронно. Сбой мог длиться какие-нибудь сотые доли секунды, но человеческий мозг, эволюционировавший миллионы лет, обычно улавливал какую-то неестественность. Однако к 2041 году софт DeepMask — преемник дипфейка — достиг такой степени правдоподобия и синхронизации образов, что научился обманывать и человеческое сознание.

Антифейковые детекторы стали неотъемлемой частью стандартной конфигурации кибербезопасности — в Европе, Америке и Азии даже приняли соответствующие законы, но в Нигерии верификацию требовали только основные контент-платформы и государственные веб-сайты. Объяснение простое: для таких детекторов необходимы огромные вычислительные мощности и профессиональные навыки, кроме того, они существенно тормозят скорость проигрывания видео. Если же пользователям приходится ждать, они, как известно, имеют обыкновение переключаться на что-нибудь другое.

Нигерийские соцсети и платформы обмена видео обновляли детекторы выборочно, в основном в моменты, совпадающие по времени с наиболее популярными алгоритмами создания дипфейков; чем больше пользователи шерили контент, тем пристальнее обращали на него внимание.

**После каждого такого «видеоосвидания»** Амаке обычно требовались тишина и темнота. Аскетичная обстановка съемной комнатки не давала забыть о суровой жизни. Однако вынырнув из виртуальной реальности, он позволял разуму еще чуть-чуть побыть с этими сладкими парнями, млея от их улыбок и нежных слов. *Они любят не меня*, напоминал он себе периодически, *они любят девушку йоруба с моим лицом*.

Когда Амака только появился на свет, местный прорицатель объявил отцу, что его новорожденный сын — реинкарнация женской души, заключенной в теле младенца мужского пола. Под клеймом «несоответствия души и тела» прошло все детство Амаки, и это стало позором для его семьи.

Амака подрастал и убеждался, что действительно не похож на других мальчиков. Во многом именно из-за этого он и уехал из деревни в продвинутый Лагос. Но и тут все оказалось непросто. Стоило увидеть привлекательного парня в метро или на улице, как Амака чувствовал: в теле и в душе что-то трепещет. Зажечь чувства порой мог даже простой зрительный контакт.

Амака знал: приватно встретиться в реале с парнями, с которыми кокетничал виртуально, он не осмелится никогда. И чем больше мощи он раскапывал в DeerMask — тем сильнее прикипал к маске. Она скрывала его настоящее лицо; благодаря ей он мог излить чувства и испытать облегчение, не подвергая себя опасности общественного порицания и не сгорая от стыда.

Амака изо всех сил старался сосредоточиться на заказе, на фейковом видео, но в дверь постучали. Пришла Озиома, хозяйка комнаты. В руках у нее была миска нарезанных семян ореха кола. Озиома тоже была игбо, но переехала в Лагос двадцать лет назад и совершенно ассимилировалась. Однако при первом знакомстве с Амакой она расслышала его тщательно скрываемый акцент игбо.

— Там, откуда я родом, раскалывать орехи кола разрешается только мужчинам, — заметил Амака, набивая рот угощением и смакуя знакомую горечь.

— Вот поэтому я оттуда и уехала! — хихикнула Озиома. — Йоруба называют это орех *оби*, а игбо — *одзи*. Но разве в названии дело? И то и другое, оказавшись у тебя во рту, решает все твои проблемы.

— О-о-о, мудрость старейшин. Спасибо за угощение, — сказал Амака и собрался было выпроводить Озиому, но она схватила его за руку и кивнула на светившийся монитор. Выражение лица женщины вдруг стало озабоченным.

— У тебя ведь с ним... ничего? Я имею в виду, симпатичный парень, просто... Мне не нужны неприятности... Ну, ты понял, о чем я.

— Нет, конечно, я просто новости читал, — Амака натянул на лицо беззаботную улыбку. — Что я, дурак, что ли?! Мне ведь еще вид на жительство получать.

— Ну и молодец. Да благословит его бог — на чьей бы стороне он ни был.

Озиома ушла, прикрыв за собой дверь.

Амака облегченно вздохнул, опять запрыгнул на кровать и снова сосредоточился на мониторе.

Лицо мужчины излучало силу. Лоб и щеки были покрыты белой краской — это символ племенного духа. Глаза светились, будто два огня. Губы были слегка приоткрыты, уголки рта образовали полуулыбку — будто он собирался заговорить на божественном языке новой эпохи и взять штурмом весь мир.

Амака отлично знал этого нигерийца. Его звали Фела Кути, он был музыкантом, отцом жанра афробит, борцом за демократию и умер 45 лет назад.

**Амаке предстояло сделать фейковое еще более фейковым.**

Аватар — виртуальный образ с лицом Фелы Кути — возник в сети и начал постить видео на платформе GargiV. «Воскресший» Фела произвел в интернете фурор. Аватар называл себя АФАК — аббревиатура от «Аватар Фелы Аникулапо Кути». Видеоролики содержали едкие комментарии по поводу текущих социальных проблем, хотя точно определить его политическую принадлежность было трудно. Большинство людей восприняли происходящее как шутку — все знали, что настоящий Фела умер в 1997 году.

Технология переноса лиц в этих видеороликах была грубой до смехотворности. Платформы для обмена видео особо не заморачивались с запретом видео с АФАКом — просто помечали этот фейк как пародию.

64 Однако авторитет АФАКа рос, и вскоре его влияние стало совсем не забавным. Миллионы нигерийцев заходили в закрытые чат-группы и обсуждали его видеоролики — с лупой рассматривали каждый кадр, смаковали каждый слог. Текст даже стали переводить на разные диалекты — делали отличный дубляж и синхронизировали артикуляцию. Благодаря столь высокому качеству обработки послания АФАКа стали распространяться гораздо шире.

Официальный фонд Фелы Кути выступил с заявлением: организация озадачена происхождением аватара не меньше других. Однако не потребовал, чтобы загадочный персонаж, кто бы ни стоял за этим аккаунтом, перестал использовать образ музыканта.

Отследить человека, скрывающегося за ником, пока не удалось. Вся видеoinформация зашифровалась; учетные записи, с которых загружали видео, были одноразовыми и проходили через несколько прокси-серверов. Вполне логично возникла конспирологическая теория: а вдруг АФАК — дело рук антиправительственных активистов или какого-нибудь недружественного иностранного государства, посягающего на стабильность в Нигерии?

Компания Ljele, новый работодатель Амаки, оказалась ширмой, прикрытием подпольной группы «Слава игбо», а Чи — просто агентом, которого наняли завербовать Амаку и управлять его дальнейшими действиями. Группа проанализировала содержание видео с АФАКом и пришла к выводу: за аватаром стоят не иностранцы, а ультранационалисты йоруба.

Они хотели воспользоваться его популярностью и манипулировать сознанием людей, постепенно сделать видео с аватаром более «прйорубским» и перетянуть общественное мнение на свою сторону. Амаке было ясно: чем больше власти окажется в руках у и без того доминирующих йоруба, тем сильнее они станут притеснять другие этнические группы, и в первую очередь игбо.

Так, недавно АФАК призвал штаты, в которых преобладают игбо, отказаться от претензий на новое месторождение редкоземельных элементов и предложил сделать его «общей собственностью всех нигерийцев». Это была очередная и на текущий день последняя попытка отобрать у игбо ресурсы дарованной им богом земли. Народ игбо всегда чувствовал себя чем-то вроде хвоста ящерицы Нигерии: его отрезали, он отрастал, его рубили вновь — и так бесконечно. Болит при этом культя, кровоточит — никого не волновало.

Игбо вконец устали. Миссия Амаки была ключом к революции, к достижению цели, поставленной перед собой группой «Слава игбо». Чи поручил Амаке создать фейковые видеоролики, которые подорвали бы авторитет и влияние аватара — так предполагалось ослабить доверие к АФАКу и, соответственно, его контроль общественного мнения.

Технически все было не так уж сложно. С помощью H-GAN (гипергенеративно-состязательной нейросети) Амака легко скопировал машинную модель лица АФАКа. Модель Амаки оказалась пиксельной зеркальной копией аватара АФАКа — начиная с частоты моргания и движения губ и заканчивая довольно заметной несостыковкой контура рта с окружающей кожей. Амака знал, как установить параметры и точно соотнести каждое математическое значение дипфейка и оригинала. Он мог обмануть любой детектор фейков и уж тем более — любой человеческий глаз.

Труднее всего оказалось воспроизвести речь в стиле АФАКа. Темы видеороликов варьировались от социально-политических новостей до головолomных популистских заявок, которые так любят обыватели. В монологах аватар то цитировал известные слова настоящего Фелы Кути, то сыпал перлами народной мудрости. Амаке часто было трудно разобрать, что он там произносит, не говоря уже об имитации речи.

АФАК заявлял, что Нигерия остро нуждается в новом языке, выходящем за рамки этнических границ, — дабы «очистить наш разум и язык от колониального яда». Он сетовал, что нигерийские матери, которые «страдают больше всех и заслуживают искреннего и высочайшего поклонения», собственноручно «приветствуют падение и закапывают трупы» своих многочисленных детей.

Аватар провозглашал, что «музыка — это оружие будущего», что только когда образование и богатство будут «распределяться также равномерно, как барабанная дробь пронизывает воздух, биение человеческих сердец сможет слиться в единый устойчивый ритм».

Очень скоро слова АФАКа, словно ливень, обрушившийся на выжженную солнцем землю, начали утолять жажду и в сердце Амаки. Как бы парень ни сопротивлялся, он чувствовал, что идеи аватара воодушевляют его и вселяют в него надежду. Может, Чи ошибается в отношении АФАКа? Амака попытался отмахнуться от нахлынувших чувств. *Не нужно мне никакого вашего паршивого чувства принадлежности*, твердил он себе.

Нужно ему сейчас было только одно — идеальный фейк речи в стиле АФАКа, в который бы поверили люди.

**Карнавальное шествие** растеклось по улицам Центрального Лагоса. Из глубины своего балкона Амака наблюдал за группой молодых нигерийцев. Они раскачивались и кружились ритмично, грациозно и легко, как пылинки в солнечных лучах. Лица парней были расписаны белой краской в стиле Фелы Кути. Мускулы на их обнаженных спинах блестели под палящим солнцем. Следуя ритму, они в унисон воздевали руки и встряхивали кистями, будто бы читая закливания.

Музыка инструментов разных этнических групп образовывала гармоничную мелодию. Пронзительный грохот барабана бата и низкие стоны барабана дундун

народа йоруба; металлический звон колокольчика *огзи* и серебристая мелодия дудочки *они* народа игбо. Воздух звенел от музыки, словно тетива боевого лука, когда ее натягивает стрелок.

66 Танцоры, гибкие, как молодые побеги маниоки в сезон дождей, ритмично двигались под музыку. Амаке они казались не отдельными людьми, а единым существом — таким же мощным и неделимым, как повторяемая ими мантра: «Еди-ная Ниге-рия». Это был слоган видеокмпании с АФАКом.

Амака чувствовал, что сейчас лопнет от эмоций. С одной стороны, он завидовал танцорам и инстинктивно жаждал присоединиться к ним — но желание подавлял страх разоблачения и обвинения в предательстве. С другой стороны, действительно ли эти танцоры — последователи АФАКа и желают зла игбо, народу, от которого Амака отдалился, но который еще любил?

Но куда больше парня беспокоил стремительно приближающийся дедлайн, установленный Чи. С каждым днем Амака все больше убеждался, что поставленная перед ним задача абсолютно невыполнима.

Амаке стало казаться, что единственной идентичности АФАКа просто не существует. Аватар создавала команда, она опиралась на систему смарт-тегов платформы обмена видео и выдавала ролики, адаптированные к различным профилям пользователей. Она точно подгоняла основные темы, слоганы, тон и жесты под потребности каждой конкретной аудитории — как рекламное агентство, которое потворствует потребностям и желаниям определенной демографической группы.

Создание просто дипфейка — это одно, а дипфейка с множеством разных идентичностей — совсем другое; и это уже выходило за рамки умений Амаки. По какой-то непостижимой причине осознание этого принесло облегчение. Однако теперь парню предстояло столкнуться с последствиями провала миссии, порученной ему Чи.

— А почему бы тебе к ним не присоединиться? — голос Озиомы раздался прямо над ухом. Хозяйка шагнула из комнаты на балкон, приветственно пихнув Амаку, закурила английскую сигарету, облокотилась на перила и посмотрела вниз.

— А ведь я когда-то была королевой танцев в нашей деревне, — взгляд Озиомы затуманился: она ностальгировала. — Правда-правда, я не хвастаюсь. Парни глаз не могли от меня оторвать. Отец ненавидел, когда я танцевала. Он каждый раз грозился отлупить меня, когда заставлял танцующей.

— И вы слушались?

Озиома от души расхохоталась.

— Да какой же ребенок откажется от того, что он любит, только из-за запрета родителей?! Я нашла способ, чтобы как минимум закончить танец.

— Какой?

— Прежде чем начать танцевать, я всегда надевала Агбого Ммуо.

— Чего-о-о?! — глаза Амаки вылезли из орбит. Агбого Ммуо — священная маска северных игбо, олицетворение девичьих духов, а также матери всего живого.

— Вот-вот, у отца, когда он видел меня в маске, становилось точно такое же лицо. Ему оставалось только кланяться и выражать почтение маске и богине, которую она олицетворяет. Конечно, потом, когда танец заканчивался и я снимала маску, мне доставалось на орехи, — Озиома сияла от гордости: воспоминания на миг вернули ее в юность, когда она была такой смелой и находчивой.

От истории Озиомы в голове Амаки шевельнулась какая-то идея. Она заворочалась, словно огромная рыбина на глубине, потом начала переваливаться с боку на бок и бить хвостом. Амака задумчиво наморщил нос и даже пробормотал вслух:

— Маска...

— Да, дитя мое. Маска — вот что давало мне силу, — эхом откликнулась Озиома. Она, оказывается, все еще была здесь.

— А если снять маску? *Содрать маску...* — бормотал, вдруг оглохнув, Амака.

Он вдруг бросился к Озиоме, подхватил на руки и звонко чмокнул в нос, прокричал: «Спасибо, спасибо, моя королева танцев!» — и ломанулся в комнату, оставив за спиной суматоху маскарадного шествия и совершенно сбитую с толку женщину.

— Если вложить в уста АФАКа явное вранье, это не заставит последователей отвернуться, он же кумир, — объяснял Амака в тот же день в видеочате своему куратору Чи. Парень был чрезвычайно взволнован неожиданным открытием. — Но если снять с него маску и разоблачить скрытую сущность кукловода, это наверняка поможет.

— Но кукловода никто не знает, — засомневался Чи.

— Вот именно! — Амака просиял. — Не понимаете? Раз его никто не знает, значит, кукловодом может быть *кто угодно!*

— То есть ты предлагаешь...

— ...сорвать с АФАКа маску и сделать его таким, каким вы захотите!

Чи довольно долго молчал, а потом наконец пробормотал: «Да ты просто гребаный гений».

— *Ндево*, — попрощался Амака на языке игбо, собираясь уже выйти из чата.

— Погоди, — окликнул его Чи, — но ведь тогда нужно создать лицо, которое существует в реальности.

— Да, именно так.

— И оно должно обмануть все антифейковые фильтры, — задумчиво добавил Чи. — Надо учесть искажение цвета, структуру помех, вариации степени компрессии, частоту моргания, биосигнал... Это вообще выполнимо?

— Мне нужно время, — кивнул Амака. — И неограниченные мощности облачного искусственного интеллекта.

— Свяжусь с тобой, — бросил Чи и отключился.

Экран погас, Амака не двинулся с места и разглядывал свое отражение. Адреналин, поначалу захлестнувший его, куда-то улетучился. Теперь на лице Амаки было не возбуждение, а усталость и тревога — он чувствовал себя так, будто только что предал духа-хранителя, наблюдающего за ним с небес.

**Теоретически создать фейковый образ** и вставить его в видео может кто угодно, да и обойти антифейковые фильтры — не особая проблема. Проблема в расходах — стоимость необходимых для этого вычислительных мощностей огромна.

Дипфейки и их детекторы пребывали в состоянии вечной борьбы, как Эрос и Танатос. Перед Амакой стояла невероятно сложная задача, но он был настроен достигнуть своей единственной цели: создать настоящее человеческое лицо.

Чи предложил новую схему: АФАКа следовало лишить цифровой маски Фелы Кути и показать всем известного Репо, политика йоруба, известного нескончаемыми нападками на другие этнические группы и главного врага движения «Единая Нигерия». Как только Чи и его команда покажут, что именно Репо дергает за ниточки, что это он стоит за харизматичной и вдохновляющей фигурой АФАКа, вера людей в аватара рассыплется в прах. Но первым делом сфабрикованное Амакой видео должно будет пройти проверку миллионами глаз — как людских, так и искусственного интеллекта, включая VIP-детектор.

Программу, прозванную VIP-детектором, разработали для защиты репутации общественных деятелей: политиков, госчиновников, знаменитостей, спортсменов, ученых. Эти выдающиеся личности оставляли слишком заметный цифровой след, и это делало их особо перспективными мишенями для дипфейков.

VIP-детектор должен защитить эти «суперточки» киберпространства от возможного мошенничества или других разрушительных для них действий — в противном случае социальному порядку может быть нанесен ущерб. Веб-сайты, публикующие изображения или видео известных людей, обязаны перед публикацией контента применять специальный алгоритм обнаружения дипфейков. В VIP-детекторе задействовано множество технологий, от распознавания лиц в сверхвысоком разрешении до датчиков распознавания жестов и геометрии ладони и пальцев, оценки речи и даже распознавания вен.

Эти данные вводились в систему глубокого обучения ИИ VIP-детектора. Если объект защиты был особо важен, VIP-детектор заносил в свою базу данных даже его историю болезни. Понятно, что Репо с его социальным статусом и неоднозначностью входил в эту категорию.

Но Амака был уверен, что и в VIP-детекторе есть изъяны. Сумей он раскодировать его принцип действия, он точно мог бы найти пробелы в пересекающихся цепочках входных данных и использовать их. И пусть эти лазейки совсем узкие — кто ищет, тот, как известно, всегда найдет. И протиснется бочком.

И вот Амака, взяв за основу реальное видео Репо, принялся, подобно какому-то создателю Франкенштейна XXI века, слой за слоем сшивать его лицо: губы, глаза, нос. Искусственный интеллект в помощь: каждое мельчайшее движение, мимолетный жест в фейковом видео исходил от самого Репо, и это сводило к минимуму вероятность того, что антифейковый детектор обнаружит подделку.

Амака начал с того, что с помощью технологии XR-зрения создал себе трехмерное рабочее пространство. Он махал руками, выбирал, перетаскивал объекты, увеличивал и уменьшал масштаб; иконки и фрагменты видеозаписи парили в воздухе и управлялись разными жестами. Амака предпочел бы видеть себя магом, творящим волшебство, но пока куда больше был похож на звезду кулинарного шоу за приготовлением блюда авторской кухни.

Для каждой части тела Репо Амака тщательно выбрал самую эффективную программу с открытым кодом — и это было похоже на то, как повар выбирает и откладывает в подходящую посуду сырые ингредиенты для будущего блюда. Затем, словно приправляя кушанье разными специями, он скорректировал параметры, модели и алгоритм обучения. И наконец, довел все это «до кипения» на облачной ИИ-платформе — на огне максимальной вычислительной мощности.

Каждый набор видеоресурсов, обработанных GAN, генерировал серию миниатюр, которые образовали в виртуальном рабочем пространстве бесконечную галерею — и это была длинная череда постеров с различными частями тела Репо.

А между тем за этой стеной из плакатов развернулась жесточайшая битва — в облаке, в полной тишине. Сражались позитивные и негативные полюса GAN: сеть-фальсификатор и сеть-расследователь. Первая стремилась переобучить и модернизировать себя для создания максимально реалистичных изображений, способных обмануть антифейковые детекторы. Эти изображения, базируясь на обратной связи от сети-расследователя и благодаря ей должны были минимизировать значение функции потерь сгенерированного образа.

Сеть-расследователь, напротив, стремилась увеличить значение функции потерь. Эта битва, ставки в которой повышаются с каждой миллисекундой, будет повторяться миллионы раз — пока стороны не достигнут определенного баланса.

Корректировка параметров, очередная итерация модели... С каждой последующей корректировкой Амака убеждался: видео становится все более реалистичным. Он уже почти ослеп от бесконечно мелькающих разноцветных пикселей, но продолжал всматриваться в кадры в поле XR-зрения, которые практически не отличались друг от друга. Пот струился по лицу парня, но ловким пальцам Амаки, порхавшим по клавиатуре, это не мешало.

Время от времени в ушах, отвлекая и раздражая, возникал противный шепот — словно голос злого духа *огбанже*, навечно застрявшего в подвешенном состоянии между жизнью и смертью. «Ты убиваешь бога собственными руками», — шептал этот голос.

«Он не мой бог. Он йоруба», — возражал Амака, заставляя себя не отвлекаться от дела.

Наконец фейковое видео блестяще обошло фильтр — сумело обмануть VIP-детектор. У Амаки хватило сил только шагнуть к кровати — он рухнул плашмя и тут же уснул.

**Незнакомый голос звал Амаку по имени.** В изножье кровати колыхалась темная тень. Амака в ужасе попытался нащупать выключатель — пальцы не слушались. Тень приближалась. И тут он разглядел лицо — это был АФАК.

70 — Чего тебе?! — Амака аж задохнулся.

— Не бойся, дитя мое. Я услышал твой зов и пришел посмотреть на тебя, — АФАК улыбался в темноте, причем, похоже, вполне искренне.

— Но я не... Я не хотел причинить тебе вред, — Амака шептал дрожащим голосом.

АФАК вдруг расхохотался — и это было похоже на рык африканского леопарда.

— Никто не способен причинить мне вред, дитя мое. Ни ты, ни они.

— Они?

— Люди, которые пытаются задушить будущее Нигерии и обманом заманить тебя в ночные джунгли.

— Мне очень жаль, АФАК, но у меня не было выбора.

— Нет, выбор есть, дитя мое. Отправляйся в Голливуд\*. Не ищи простого решения — расскажи миру настоящую нигерийскую историю.

Амака не мог вымолвить ни слова, не получалось и отвести взгляд от пиксельной фигуры, стоящей у кровати. *Я и вправду всегда хотел рассказать свою историю*, думал он, *историю игбо, который мечется между традиционной и изменяющейся реальностью.*

— Мой дух-хранитель покинул меня, потому что я оставил свою землю и уехал на землю йоруба, — запинаясь, с трудом выговорил Амака.

— Чушь какая! — перебил АФАК. И Амака вдруг понял, что знает этот голос. — Помнишь, как ты был маленьким?

— Я был маленьким?

— Я показывал тебе разных птиц и говорил, как они называются; я показал, из какого дерева получится самая хорошая рогатка; рассказал, как смастерить флейту из слоновой травы... Ты что, все забыл?

— Всему этому меня учил отец. — Амака изумленно замолк, глаза его расширились.

— Вот именно, дитя мое. Помнишь народную поговорку игбо? Когда человек говорит «да», его духу-хранителю ничего не остается, как тоже сказать «да». Это только люди покидают своего бога — бог их никогда не оставит.

— Но отец, я же не хочу тебя подвести, — Амака вспомнил об угрозе Чи опозорить всю его семью и почти шептал.

— Есть кое-что, о чем я никогда тебе не говорил.

— О чем ты?

\* Так, по аналогии с американским Голливудом и индийским Болливудом, называют всю нигерийскую киноиндустрию. *Прим. ред.*

— Мне все равно, что сказал при твоём рождении прорицатель. И мне неважно, чья душа живет в теле моего ребенка. Я хочу только, чтобы он был счастлив, чтобы он был добрым и чтит богов и духов.

— Отец... — Амака потянулся к силуэту АФАКа. Он хотел снять с него маску и увидеть родное обветренное лицо отца.

— Сходи в New Afrika Shrine\*. Я уверен, ты сумеешь сделать правильный выбор. А потом возвращайся ко мне.

Амака уже почти дотянулся — кончики пальцев вот-вот должны были коснуться мерцающего пиксельного лица. Но АФАК вдруг исчез. Амака проснулся. Горел ночник. С темно-зеленого экрана монитора Illumiware Mark-V улыбалось знакомое лицо.

**Расписанное граффити здание** клуба New Afrika Shrine в Икедже можно за просто принять за полуразрушенный гараж. Впрочем, недостаток архитектуры с лихвой компенсирует энергетика места. В здании вместимостью две тысячи человек еженедельно проходят концерты, в нем работают всевозможные киоски с едой и напитками — и это весьма процветающий бизнес.

Фела Кути открыл ночной клуб Afrika Shrine в отеле Empire, но в 1977 году клуб сожгла полиция. Здесь была уже реинкарнация клуба — его в 2000 году открыл в память об отце Феми, сын Фелы.

Амака много раз бывал в Afrika Shrine. Как и любой молодой человек в Лагосе, любящий хорошо провести время, он видел в нем не только подходящее для вечеринки место с едой и выпивкой, но и храм, место паломничества — тут можно соприкоснуться с мятежным и свободным духом живших полвека назад таких же молодых. В этом особенном месте люди волшебным образом забывали обо всех этнических и классовых конфликтах и радовались жизни, пусть бы и под парами алкоголя.

Но сегодня Амака пришел со всем этим проститься.

В Afrika Shrine — и в старом, который он не застал, и в новом — свято хранили память и превозносили черных богов и богинь: Кваме Нкруму, Мартина Лютера Кинга — младшего, Малкольма Икса, Тома Санкару, Нельсона Манделу, Эстер Ибанга, Чинуа Ачебе, Воле Шойинка, Флоренс Озор... — великие души, посвятившие свою жизнь свободе, демократии и равенству\*\*. Артисты в этом клубе часто прерывают выступления, чтобы отдать дань уважения своим культурным предкам.

Амака смотрел во все глаза — он молчал и старался получше запомнить все лица. Он молился, чтобы эти боги и духи оберегали его и впредь.

\* Развлекательный центр под открытым небом в Икедже, столице штата Лагос. *Прим. пер.*

\*\* Все имена принадлежат темнокожим — деятелям искусства, культуры, религии, борцам за права темнокожих в разных (в основном африканских) странах. *Прим. ред.*

Он уедет из Лагоса, вернется домой и все расскажет отцу. Амака еще не решил, что будет делать дальше. Возможно, его умение работать с GAN поможет ему найти хорошую работу, и ему больше не придется делать дипфейки, а он станет помогать людям. Или ему удастся устроиться в сфере здравоохранения, скажем, заниматься переносом лиц в медицинских датасетах для обучения ИИ?

А ведь он еще отлично умеет раскрашивать старые черно-белые фильмы и улучшать изображение! А вдруг он сможет еще сильнее расправить крылья и сделать то, о чем лишь изредка осмеливался мечтать: снять настоящий нолливудский фильм? У Амаки уже появилась отличная идея для сюжета.

В смартстриме зазвенели монеты — пришли деньги, обещанные Чи. Это означало, что его фейковое видео настолько реально, насколько это вообще возможно — оно разлетелось по интернету, и его влияние можно было сравнить по мощи с ядерным взрывом.

В последние годы видео, созданные с помощью ИИ, спровоцировали мятеж в Габоне и серьезные политические беспорядки в Малайзии. И Амаке была невыносима мысль, что его видео может сыграть такую же скверную роль в истории его родной Нигерии.

*Но выбор уже был сделан.*

Амака поднялся на сцену и встал в самом центре, прямо под висящим высоко над ней черно-белым портретом Фелы Кути, сначала воздел руки, а потом вытянул их вперед — как бы стараясь соединиться с силой богов и духов.

«Я стану хозяином своей судьбы и сам буду решать, когда смерть придет забрать меня», — серьезно прошептал парень, словно какое-то магическое заклинание. Это были слова самого Фелы Кути, объяснявшие его второе имя Аникулапо — на языке йоруба оно означает «тот, кто несет в суме смерть».

Амака ввел что-то в смартстрим и запустил его в ближайший мусорный бак. Вынул из кармана и надел свою верную грубую маску, напечатанную на 3D-принтере. Амака молился и мечтал убраться отсюда как можно дальше, прежде чем спохватится Чи. Он уедет из Лагоса, из огромного города, исписанного слоганами *Eko o ni baje* — «Лагос не испортит», и вернется домой, к запахам свежей земли.

Он уничтожит ложь, создав еще одну ложь.

Второе видео, снятое им с помощью софта DeepMask, он уже выложил в интернет — оно готово спровоцировать еще один взрыв. И в первом, и во втором ролике АФАК снимал цифровую маску, под которой оказывалось лицо Репо — идеальный дипфейк, легко преодолевший все антифейковые фильтры. Но второй ролик на этом не заканчивался — аватар продолжал снимать маски: за АФАКом следует маска Репо, затем — другая маска под этой маской, потом еще... Слой за слоем, до бесконечности.

Нигерийцы изумятся: за АФАКом — все боги и богини, почитаемые за все-гдамыми клуба New Afrika Shrine.

# АНАЛИЗ

## КОМПЬЮТЕРНОЕ ЗРЕНИЕ, СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ; ДИПФЕЙКИ; ГЕНЕРАТИВНО- СОСЯЗАТЕЛЬНЫЕ СЕТИ (GAN); БИОМЕТРИЯ; БЕЗОПАСНОСТЬ ИИ

«Боги под масками» — история о визуальном обмане, об обмане зрения. Научившись видеть, распознавать, понимать и синтезировать объекты, ИИ может также манипулировать ими, создавая изображения и видео, неотличимые от реальных. В рассказе описывается будущее, в котором люди больше не смогут полагаться на невооруженный глаз, если наверняка пожелают отличить подлинное видео от тонкого монтажа.

Сайтам и компьютерным приложениям придется (их обяжут законом) использовать антидипфейковый софт (как сегодня используются антивирусные программы), чтобы защитить пользователей от поддельных видео. В нашей истории перетягивание каната между создателями дипфейков и детекторов для их выявления уже переросло в гонку вооружений, и с переменным успехом побеждает тот, у кого больше вычислительных мощностей.

Действие рассказа происходит в 2041 году, но, вполне вероятно, подобная ситуация сложится и раньше — как только развитый мир сможет позволить себе огромные расходы на супердорогие компьютеры, программное обеспечение и ИИ-экспертов, необходимых для создания и обнаружения дипфейков и прочих манипуляций ИИ. И регулирующие эту сферу законы, скорее всего, будут приняты прежде всего в развитых странах. Наша история происходит в стране развивающейся — в ней негативные внешние эффекты дипфейков, по всей вероятности, проявятся несколько позже.

Итак, как же ИИ научается видеть — и через камеры, и через записанные видео? Как используются способности ИИ? Как работает создатель дипфейков, применяющий ИИ? Способны ли люди или ИИ обнаруживать дипфейки?

Наполнят ли наши соцсети фейковые видео? Как остановить поток дипфейков? Какими еще дырами в безопасности может быть чреват ИИ? И есть ли что-нибудь полезное в технологии, позволяющей создавать дипфейки?

74

## ЧТО ТАКОЕ КОМПЬЮТЕРНОЕ ЗРЕНИЕ?

В рассказе «Золотой слон» мы увидели потенциальные возможности глубокого обучения при работе с большими данными, в частности в интернете и финансах. Вряд ли вы удивляетесь тому, что ИИ опережает людей по эффективности обработки больших массивов данных. Но вот как насчет способностей, уникальных для человека или прочих живых существ, например восприятия?

Зрение — самый важный из органов чувств человека. Компьютерное (машинное) зрение — это отрасль ИИ, которая обучает компьютеры видеть. Это слово означает здесь не только оцифровку видео или изображения, но и осмысление того, что при этом «видит» компьютер. Алгоритмы компьютерного зрения позволяют реализовать следующие возможности (перечислены в порядке возрастания сложности):

- Получение изображений и их обработка — для реальных 3D-сцен в видео используются камеры и другие приборы-датчики. Каждое видео состоит из последовательности изображений, каждое изображение представляет собой двумерный массив чисел, представляющих конкретный цвет, где каждое число соответствует пикселю.
- Обнаружение объектов и сегментация изображения — изображение разбивается на области и определяется место расположения объектов.
- Распознавание объекта — распознается объект (например, собака) и, возможно, детальная информация (немецкая овчарка, темно-коричневый окрас, и так далее).
- Трекинг объекта — отслеживание движущихся объектов в последовательных изображениях или видео.
- Распознавание жестов и движений — движения идентифицируются, как, например, танцевальные па в игре для Xbox.
- Интерпретация сцены — обеспечивается понимание сцены в целом, включая неявные взаимоотношения и нюансы, например голодная собака, жадно глядящая на кость.

Инструменты для создания дипфейков, использованные Амакой в рассказе, содержат все эти компоненты. Чтобы отредактировать видео с АФАКом, Амаке сначала нужно было разбить его на фрагменты — 60 кадров в секунду, каждый из которых состоит из десятков миллионов пикселей. ИИ считывает эти пиксели и автоматически сегментирует тело АФАКа (или рисует вокруг него границу), которое далее сегментируется на лицо в маске, рот, руки и так далее. И так для каждого видеокadra.

Если видео длится 50 секунд, у нас будет три тысячи кадров (изображений). Кроме того, сопоставляется движение между кадрами, что позволяет выделить связи между объектами. Все это, заметьте, делается до того, как начнется редактирование видео.

Процесс представляется жутко трудоемким, но мы, люди, в своей жизни делаем это на раз-два. Нам достаточно лишь бросить взгляд, и все перечисленное уже проделано — тратится меньше секунды. Кроме того, люди обладают способностью к абстрактной и обобщающей интерпретации, причем даже если один и тот же объект выглядит по-разному под разными углами, при разном освещении, с разных расстояний или даже почти скрыт другими объектами.

Например, увидев человека, сидящего за столом в определенной позе, мы можем сделать вывод, что он держит ручку, нацелившись на лист бумаги, хотя на самом деле не видим ни того, ни другого.

Когда мы что-то «видим», мы, по сути, каждый раз задействуем все накопленные ранее знания о мире — все, что узнали за свою жизнь о перспективе, геометрии, здравом смысле и обо всем том, что видели до этого. У человека это происходит просто и естественно, но научить тому же самому компьютер крайне сложно.

Компьютерное зрение — область исследований, которая пытается преодолеть трудности и сделать так, чтобы компьютеры умели видеть и понимать то, что они видят.

## СПОСОБЫ ПРИМЕНЕНИЯ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Уже сегодня мы практически ежедневно соприкасаемся с технологиями компьютерного зрения.

Их можно использовать в режиме реального времени в разнообразных областях — от транспорта до сферы безопасности. Приведу примеры.

- В автомобилях устанавливаются системы помощи водителю «Анти-сон», не дающие уснуть за рулем.

- Магазины без касс и очередей — типа Amazon Go: установленные в них камеры сами распознают товар, который вы кладете в корзину или возвращаете на полку.
- Охрана аэропортов (подсчет количества людей, распознавание террористов).
- Распознавание жестов (баллы за движения в танцевальных играх для Xbox).
- Распознавание лиц (использование лица пользователя для разблокировки мобильного телефона).
- Смарт-камеры (портретный режим вашего iPhone распознает и выделяет людей на переднем плане, а затем красиво размывает фон для создания эффекта фото, сделанного зеркальным фотоаппаратом).
- Военное дело (определение, кто напротив — солдат противника или мирный житель).
- Автономные системы навигации дронов и автомобилей.

В самом начале рассказа «Боги под масками» мы видели, как системы распознавания лиц используются в реальном времени для автоматической оплаты проезда — путем идентификации пассажиров, проходящих через турникет вокзала. И узнаем, что пешеходы, используя жестикуляцию, могут взаимодействовать с мультяшными животными в уличной рекламе. А смартстрим Амаки использует компьютерное зрение для распознавания городских улиц и указания маршрута, помогает ему добраться до незнакомого пункта назначения.

Кроме того, компьютерное зрение можно применять к изображениям и видео — не такими прямолинейными способами, но не менее полезными и важными. Приведу примеры:

- Интеллектуальное редактирование фотографий и видео (такие инструменты, как Photoshop, широко используют компьютерное зрение для удаления эффекта красных глаз и улучшения качества селфи).
- Анализ медицинских снимков (для определения наличия злокачественной опухоли на КТ-снимке легких).
- Модерирование (цензурирование) контента (обнаружение порнографического и насильственного контента в социальных сетях).
- Подбор сопутствующей рекламы, основанный на содержании конкретного видео.
- Интеллектуальный поиск изображений (находит нужное по ключевым словам или другим изображениям).
- И, конечно же, создание дипфейков (подмена лиц в видео).

В «Богам под масками» рассказывается об инструменте для создания дипфейков, который, по сути, является ничем иным, как автоматическим инструментом для редактирования видео; он по частям подменяет одного человека другим, начиная с черт лица, пальцев, ладоней и голоса и заканчивая жестами, походкой и мимикой. Поговорим об этом подробнее.

## СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ (CNN) ДЛЯ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Заставить компьютерное зрение работать на основе стандартной нейронной сети оказалось очень непростой задачей — ведь любое изображение состоит из десятков миллионов пикселей, и научить систему глубокого обучения находить тончайшие подсказки и признаки в огромном количестве изображений — это, согласитесь, даже звучит устрашающе. В поисках вдохновения в деле усовершенствования этой технологии исследователи обратились к человеческому мозгу.

Зрительная зона коры головного мозга задействует нейроны, соответствующие множеству ограниченных областей (известных как рецептивные поля, или поля восприятия), в рамках которых наши глаза фиксируют изображение в любой конкретный момент времени. Рецептивные поля идентифицируют основные признаки видимых объектов: формы, линии, цвета или углы. Эти детекторы соединены с неокортексом, верхним слоем коры головного мозга (новой корой).

Неокортекс хранит информацию иерархически и обрабатывает выходы полей восприятия, преобразуя их в более сложную интерпретацию сцены.

Наблюдения за тем, как «видят» люди, и вдохновили разработчиков на изобретение так называемых сверточных нейронных сетей (CNN — convolutional neural networks). Самый низкий слой CNN состоит из большого числа фильтров, которые многократно применяются к изображению. Каждый из этих фильтров, как и рецептивные поля, может видеть только небольшие смежные участки изображения.

Глубокое обучение за счет оптимизации параметров на множестве изображений решает, что «замечает» каждый фильтр. Каждый фильтр выдает на выходе уверенность, что видел конкретный признак, им представляемый (например, черную линию). Более высокие слои CNN организованы иерархически, как и неокортекс. Они принимают выход уверенности от более низких слоев и обнаруживают более сложные признаки.

Например, если в CNN загружается изображение зебры, фильтры нижнего слоя могут искать в каждой области изображения только черные и белые линии, а более высокие слои будут видеть полосы, уши и ноги в больших зонах.

Следующие слои могут увидеть множество полос, два уха и четыре ноги. На самом высоком слое части CNN могут целенаправленно пытаться отличить зебру от лошади или тигра.

78 Обратите внимание: все эти примеры иллюстрируют, что CNN *может* делать, но в реальной работе сеть сама решает, какие признаки (полосы, уши или вообще нечто лежащее за рамками человеческого понимания) будут использованы для максимизации целевой функции.

CNN — это специфическая и улучшенная архитектура глубокого обучения, разработанная для компьютерного зрения, с различными вариантами для изображений и видео. Идея таких сетей возникла в 1980-х годах, у ученых тогда не было ни достаточного объема данных, ни вычислительных мощностей, чтобы продемонстрировать, на что эти сети способны.

Только в 2012 году стало ясно, что данная технология превосходит все предыдущие подходы к компьютерному зрению. По счастливому совпадению, примерно в это же время пользователи сняли на сильно подешевевшие и из-за этого массово распространившиеся смартфоны огромное количество фото и видео и выложили их в социальные сети. Примерно тогда же стали более доступными быстродействующие компьютеры и хранилища данных большого объема. Все эти факторы в совокупности способствовали рывку в развитии и распространении этой замечательной технологии.

## ДИПФЕЙКИ

«Президент Трамп — полный и абсолютный ушлепок», — заявил однажды на камеру президент Обама... Ну или это был человек, который выглядел и говорил совсем как Барак Обама. В конце 2018 года видео стало вирусным, но это был дипфейк — поддельный ролик, смонтированный с помощью технологии глубокого обучения. Известны даже авторы — режиссер Джордан Пил и новостная интернет-медиакомпания BuzzFeed.

Пил записал фразу, ИИ преобразовал голос режиссера в голос Обамы. Затем ИИ взял реальное видео Обамы и отредактировал так, чтобы оно максимально соответствовало словам, включая артикуляцию, соответствующую мимику и выражение лица.

Джордан Пил сделал это с конкретной целью — так он хотел предупредить людей о скором наступлении дипфейков, что, собственно, и произошло. Уже в том же году в интернет оказались загружены несколько дипфейковых порнороликов со знаменитостями, это привело к скандалу и в конечном счете к новому закону, призванному решить эту проблему. Но появление дипфейков это уже не остановило.

В Китае в 2019 году было создано приложение, способное за считанные минуты превратить ваше селфи в главного героя какого-нибудь известного фильма. Оно сохраняет оригинальный саундтрек, что существенно снижает технологические требования к процессу.

Программа AvatariFu возглавила в 2021 году рейтинг Apple App Store. Она «оживляет» любую фотографию, заставляя человека на ней петь или смеяться.

Короче говоря, дипфейки вдруг стали мейнстримом, и уже любой юзер без особого труда может сам изготовить фейковое видео (пусть и дилетантское, и видимое для детекторов, а зачастую и для человеческого глаза).

Сказанное означает, что в будущем возможна подделка любого цифрового материала, будь то онлайн-видео, заранее записанная речь, кадры с камеры видеонаблюдения или видео с уликами из зала суда. В нашей истории Амака использует инструменты, гораздо более продвинутые, чем имелись в распоряжении Пила, создает сложное, высококачественное видео, подделку, которую не распознает ни человеческий глаз, ни обычное программное обеспечение для обнаружения дипфейков.

Сначала он использует инструмент для преобразования любого текста в звукоряд, который станет неотличим от голоса Репо. Затем он синхронизирует эту речь с лицом Репо с соответствующей мимикой и эмоциями. Далее это лицо накладывается на изображение АФАКа из ранее снятого видео с соответствующими руками, шеей, ступнями и даже его уникальными паттернами пульса и дыхания. А продвинутый ИИ образца 2041 года обеспечивает совершенно бесшовное соединение всех частей тела в нужных местах.

Помимо этого подхода к созданию «фейковых людей», основанного на видео, существует еще 3D-подход, включающий в себя построение трехмерной модели человека исключительно с помощью вычислений. Именно так, например, снимаются полнометражные анимационные картины вроде «Истории игрушек»\*.

Данный подход базируется на другой области информатики — на компьютерной графике. В ней все моделируется математически, и для волос, ветра, света, теней и проч. нужно разрабатывать реалистичные математические модели. Этот трехмерный подход обеспечивает «продюсера» гораздо большей свободой действий, позволяя ему манипулировать каждым персонажем, как только заблагорассудится, однако и сложность, и требования к вычислительным мощностям тут намного выше.

Компьютеры в 2021 году еще не умеют снимать полнометражные фильмы с использованием 3D-видео, способного обмануть человеческий глаз (вот почему люди в анимационных фильмах выглядят не вполне реалистично), не говоря уже о программах-детекторах. Но к 2041 году наверняка появятся

---

\* Первый полнометражный фильм, смоделированный на компьютере полностью трехмерным; американский анимационный фильм. Совместное производство киностудии Pixar и компании Уолта Диснея (режиссер Дж. Ласетер, 1995). *Прим. ред.*

фотореалистичные 3D-модели, о которых мы подробнее поговорим в других главах этой книги — «Воробьи-близнецы» и «Мой призрачный кумир».

80 Джордан Пил создал дипфейк ради развлечения и в назидание, а в нашем рассказе агент Чи вербует Амаку для изготовления дипфейка с конкретным злым умыслом. Помимо распространения сплетен дипфейки могут быть использованы для шантажа, преследования, клеветы, манипуляций на выборах.

Как делаются дипфейки? Как научить ИИ выявлять их? Кто победит — софт для создания дипфейков или программы для их обнаружения? Чтобы ответить на эти вопросы, необходимо разобраться в механизме, который генерирует дипфейки, — в генеративно-сопоставительных сетях (GAN — generative adversarial networks).

## ГЕНЕРАТИВНО-СОСТАВЛЯТЕЛЬНЫЕ СЕТИ (GAN)

Дипфейки базируются на технологии, называемой генеративными сопоставительными сетями. Как следует из названия, GAN — это пара «сопоставительных» нейронных сетей глубокого обучения. Первая сеть, генератор, пытается создать что-то, что выглядит реалистично (скажем, синтезированное изображение собаки), основываясь на миллионах изображений собак.

Вторая сеть, дискриминатор (сеть-детектив), сравнивает синтезированное изображение собаки из первой сети с подлинными изображениями собаки и определяет, является ли выход генератора подлинным или фальшивым.

Основываясь на обратной связи от дискриминатора, генератор переобучается так, чтобы в следующий раз обмануть дискриминатор. Он самокорректируется, минимизируя «функцию потерь», то есть различия между сгенерированным и реальным изображением. Затем переобучается и дискриминатор, и ему становится легче распознавать фейки — «функция потерь» максимизируется. Эти два процесса повторяются миллионы раз; обе сети совершенствуют свои навыки до тех пор, пока не возникает устойчивое равновесие.

Первая статья о GAN была опубликована в 2014 году. Ее автор Ян Гудфелло показал, как генератор сначала создал симпатичную, но явно фейковую «собаку-мячик», которую дискриминатор мгновенно обнаружил, и как первая сеть постепенно научилась создавать фейковые изображения собак, неотличимые от реальных. С тех пор эту технологию применили к видео, речи и многим другим типам контента, в том числе и для создания печально знаменитого видео Обамы, о котором я упомянул ранее.

Можно ли обнаружить дипфейки, созданные GAN? Из-за их относительно рудиментарной природы и ограничений современных вычислительных

мощностей большинство дипфейков сегодня обнаруживаются алгоритмами, а иногда и видны невооруженным человеческим глазом.

Facebook и Google уже вступили в состязание по разработке софта для обнаружения дипфейков. Эффективные детекторы дипфейков можно создать и сегодня, но это требует огромных вычислительных мощностей, что является весьма серьезной проблемой, если на ваш веб-сайт ежедневно загружаются миллионы видео и фото.

Однако в долгосрочной перспективе наибольшая сложность состоит в том, что GAN оснащена встроенным механизмом для апгрейда сети-генератора. Допустим, вы обучили генератор GAN, а кто-то другой придумал новый алгоритм для обнаружения вашего дипфейка. Вы можете просто переобучить генератор GAN, чтобы обмануть этот алгоритм-дискриминатор. Так начинается что-то вроде гонки вооружений — каждая из сторон старается переобучить лучшую модель на еще более мощном компьютере.

В нашем рассказе более ранний ролик (про мальчика-вампира) Амака изготовил с помощью относительно простых инструментов в интернет-кафе, на минимальных вычислительных мощностях. Видео получилось достаточно хорошим, чтобы обмануть людей; в 2041 году фейковые видео были уже вполне убедительными, чтобы люди не отличали их от реальных.

Однако видео не смогло обмануть GAN-детектор сайта, на котором его разместили (его обучали с использованием гораздо больших вычислительных мощностей), поэтому оно впоследствии было удалено с сайта и заблокировано.

А позже Чи предоставил Амаке мощный компьютер для обучения сложной GAN, умеющей генерировать не только лицо, но и кисти рук, пальцы, походку, жесты, голос и мимику. Кроме того, эта GAN обучалась на огромном количестве данных, доступных в сети по такой знаменитости, как Репе — мишень Амаки. В результате Амака сумел обмануть все обычные детекторы дипфейков.

Представьте себе ювелирный магазин с пуленепробиваемыми витринами, способными защитит от любых видов стрелкового оружия. Но если у преступника будет ручной противотанковый гранатомет (РПГ), никакое пуленепробиваемое стекло перед ним не устоит. Так что все дело в мощности компьютера.

К 2041 году антидипфейковые программы станут похожими на сегодняшний антивирусный софт. Правительственные, новостные и другие сайты, для которых достоверность информации критически важна, не потерпят фальшивого контента и потому будут устанавливать высококачественные детекторы дипфейков, предназначенные для выявления подделок с высоким разрешением, которые, в свою очередь, создадут крупные сети GAN, обученные на мощнейших компьютерах.

Сайты со слишком большими объемами видео- и фотоматериала (например, социальные сети и YouTube) столкнутся с проблемой высокой стоимости сканирования всего загружаемого контента с помощью высококачественных

детекторов дипфейков, поэтому они, скорее всего, начнут использовать для всего медиаконтента детекторы низкого качества, а более качественный софт применять, только когда популярность видео или изображения покажет экспоненциальный рост.

Поскольку фейковое видео Амаки должно было стать вирусным, его нужно было обучить на самом мощном компьютере с наибольшим объемом данных, чтобы его не могли обнаружить самые качественные антидипфейковые детекторы.

Следует ли считать цель стопроцентного обнаружения дипфейков безнадежной? В очень долгосрочной перспективе это может стать возможным благодаря применению совершенно другого подхода — аутентификации каждой фотографии и видео, когда-либо снятых каждой камерой или телефоном, с использованием в момент съемки технологии блокчейн (она гарантирует, что оригинал не изменялся).

Тогда любое фото, загруженное на сайт, должно будет сопровождаться блокчейн-авторизацией. Такой процесс действительно сделает дипфейки невозможными. Однако к 2041 году такой апгрейд еще будет немыслим — ведь для этого требуется, чтобы технология блокчейн использовалась абсолютно на всех устройствах (точно так же, как сегодня все AV-ресиверы используют Dolby Digital), и она должна стать достаточно быстрой для обработки данных в огромных масштабах.

Мы надеемся, что пока у нас нет такого долгосрочного решения, основанного на блокчейне или аналогичной технологии, инструменты для обнаружения дипфейков будут постоянно совершенствоваться. А поскольку они вряд ли будут идеальными, нам также потребуются законы, жестко наказывающие за создание злонамеренных дипфейков и тем самым сдерживающие пыл злоумышленников.

Так, в 2019 году в Калифорнии приняли закон, запрещающий использование дипфейков для порно и для манипуляций с видео с политическими кандидатами в преддверии выборов. Нам, по всей вероятности, потребуется научиться жить в новом мире (до тех пор пока не заработает блокчейн-решение), в котором любой онлайн-контент, независимо от его реалистичности, должен подвергаться сомнению.

Следует также отметить, что помимо создания дипфейков GAN можно использовать для вполне конструктивных задач — «состаривания» или «омоложения» фотографий, раскрашивания старых черно-белых фильмов и фото, создания анимированных живописных полотен (например, Моны Лизы). Можно увеличивать разрешение фотографий, обнаруживать глаукому, прогнозировать последствия изменения климата и даже изобретать новые лекарства.

Так что не стоит думать о GAN только в связи с дипфейками — как в случае с большинством новых прорывных новшеств, позитивные способы применения этой технологии со временем наверняка превзойдут по численности негативные.

## ВЕРИФИКАЦИЯ ЛЮДЕЙ С ПОМОЩЬЮ БИОМЕТРИИ

Биометрия — это область изучения физических характеристик человека с целью их дальнейшего использования для проверки и подтверждения личности. То, как сложная GAN использовалась в рассказе «Боги под масками», представляет собой одну из форм биометрической верификации. Там GAN объединяла важные признаки, включая распознавание лиц, походки, геометрии ладоней и пальцев, идентификацию говорящего, распознавание вен и жестов.

В реальности биометрия обычно применяется в режиме реального времени с помощью специальных датчиков, а не путем извлечения необходимых данных из видеозаписи, как это показано в нашем рассказе. Например, радужная оболочка человеческого глаза и отпечатки пальцев уникальны для каждого человека и идеально подходят для проверки личности.

Распознавание радужной оболочки глаза считается наиболее точным методом биометрической идентификации. Чтобы верифицировать личность таким способом, глаза объекта фотографируют в инфракрасном свете, а снимок сравнивают с радужной оболочкой человека, которого проверяют.

Распознавание отпечатков пальцев — также чрезвычайно точная методика. Однако и в том и в другом случае требуется согласие объекта и специальное оборудование с датчиками ближнего действия\*, поэтому в ситуации с видеозаписями из нашей истории эти методы использовать было бы невозможно.

Последние достижения в области глубокого обучения и GAN продвинули область биометрии далеко вперед. Получив любой биометрический параметр (например, голос или лицо), ИИ уже сегодня превосходит людей в деле верификации или распознавании личности. А в ситуациях, когда есть возможность собрать и объединить множество признаков, точность, по сути, будет абсолютной.

К 2041 году ИИ, без всяких сомнений, возьмет на себя рутинную задачу по распознаванию и верификации людей. Я также берусь утверждать, что в следующие двадцать лет использование интеллектуальной биометрии для уголовных расследований и криминалистики позволит нам раскрывать гораздо больше преступлений и существенно снизит уровень преступности.

---

\* Near-field sensor — это сенсор, который считывает показания, находясь рядом с объектом, а near field communication (NFC) — технология передачи данных на близком расстоянии. *Прим. науч. ред.*

## БЕЗОПАСНОСТЬ ИИ

84 Как известно, по ходу технического прогресса возникают новые области уязвимости и повышаются риски безопасности для всех вычислительных платформ: вирусы — для ПК, кража личных данных — для кредитных карт, спам — для электронной почты. А поскольку ИИ все больше становится мейнстримным явлением, ему тоже не избежать атак на особенно уязвимые места. И дипфейки — это лишь один из множества вариантов.

Уязвимым местом ИИ являются границы, разделяющие его решения, которые можно примерно оценить и использовать для маскировки входных данных, заставляя ИИ ошибаться. Например, один исследователь разработал солнцезащитные очки. Если их надеть, ИИ верифицирует тебя как Миллу Йовович. А другой наклеил на дорогу стикеры, и автопилот Tesla Model S сменил полосу движения и выехал на встречную полосу.

В самом начале нашего рассказа Амака использует маску, чтобы обмануть систему распознавания лиц на вокзале. Так вот, маскировка такого типа очень пригодилась бы военным — только представьте, что танк можно замаскировать под, скажем, машину скорой помощи!

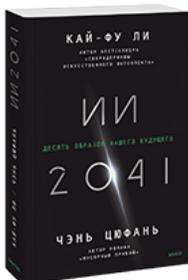
Еще один тип атаки называется отравлением (poisoning). Процесс обучения ИИ намеренно искажается за счет загрязнения обучающих данных, обученных моделей или самого процесса обучения. Это приводит к систематическому отказу ИИ в целом или к тому, что контроль над ИИ могут получить злоумышленники.

Только представьте себе боевые дроны, взломанные и перенаправленные террористами на их создателей! Такие атаки хуже обычного хакерства: ИИ-модели отладить очень нелегко, они представляют собой чрезвычайно сложные уравнения, имплементированные в тысячах слоев нейронных сетей, а не в детерминированном компьютерном коде.

Впрочем, несмотря на многие проблемы и трудности, мы уже сегодня можем предпринять некоторые меры: усилить безопасность среды обучения и исполнения; создать инструменты для автоматической проверки на наличие признаков отравления и разработки технологий, специально предназначенных для борьбы с подделкой данных.

Нам удалось с помощью технологических инноваций в основном преодолеть проблему спама и компьютерных вирусов, вот и безопасность ИИ обязательно будет усиливаться, хотя, конечно, время от времени атаки на него все же неизбежны (точно так же, как мы по-прежнему периодически подвергаемся атакам спама или вирусов).

Уязвимость, порождаемую техническим прогрессом, всегда устраняли или как минимум ослабляли благодаря новым технологическим решениям.



[Почитать описание, рецензии  
и купить на сайте](#)

Лучшие цитаты из книг, бесплатные главы и новинки:

