

# Компьютер и человек – кто кого?

Окончание. [Начало в № 7/2023](#)

Арсений Ворошилов

**И**так, в [первой статье](#) на тему искусственного интеллекта рассматривались больше экономические, правовые и эстетические вопросы. Пора хоть немного разобраться, каковы же возможности генеративного искусственного интеллекта (genAI) на сегодняшний день и чего ожидать в перспективе. А главное, почему не надо бояться победы искусственного разума над человеческим.

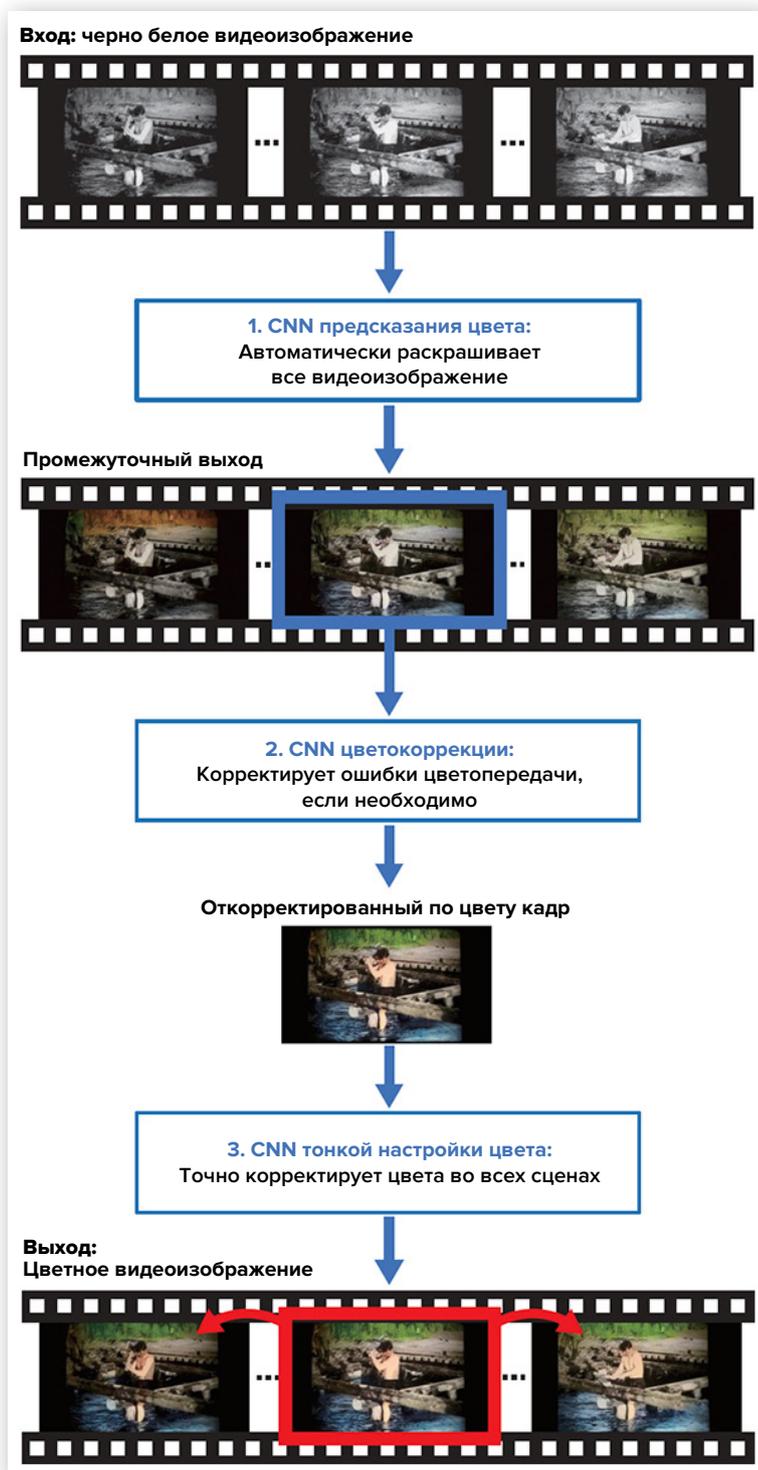
Тестирование возможностей искусственного интеллекта применительно к решению условно творческих задач предпринимаются уже не первый год. Почему «условно»? Потому что речь идет скорее о рутинных процедурах, все же предусматривающих определенное творчество. Например, преобразование старых черно-белых документальных кино материалов в цветную форму. Разве не творчество? Корпорация NHK уже демонстрировала результаты такого эксперимента и объясняла в общих чертах, на основе каких исходных данных он проводился. В частности, для этого применили сразу три сверточные нейросети CNN (Convolutional Neural Network), одна из которых служила для предсказания цвета, вторая – для цветокоррекции, а третья – для тонкой настройки цветапередачи на основе интерполяции.

Хотя сам эксперимент проводился лет пять тому назад, официальный отчет о нем был опубликован в журнале Broadcast Technology № 83, который вышел зимой 2021 года.

Как объясняли авторы отчета Реи Эндо, Йошихико Каваи и Такахиро Мокизуки, нейронная сеть, отвечавшая за предсказание цвета, выполняла полностью автоматизированное раскрашивание статичных изображений, опираясь на возможности, имеющиеся в широко распространенных системах преобразования черно-белого изображения в цветное. Далее сформированные на первой стадии метаданные передавались в CNN цветокоррекции. Эти метаданные содержали информацию не только о преобразовании черно-белого изображения в цветное, но и о границах каждого из цветов. Это позволяло точно раскрашивать изображения, не выходя за очерченные границы цветов. Причем вполне приемлемый результат достигался даже для изображений, снятых на монохромную пленку и содержащих большое количество белых и черных пятен, что обычно не позволяет CNN автоматически определять границы цветов.

«Компьютер делает не то, что вы хотите, а то, что вы ему приказываете»

Автор неизвестен



Тракт системы NHK для автоматического преобразования черно-белого изображения в цветное

А на этапе тонкой настройки цветопередачи третья нейросеть опиралась на концепцию, которую специалисты NHK разработали специально для данного проекта. В целом суть ее в следующем: при раскрашивании монохромного кадра CNN брала в качестве эталона цветной контент из соседних кадров, которые уже были раскрашены, а затем корректировала цвета в текущем кадре так, чтобы все кадры – текущий и соседние – имели одинаковую цветопередачу.

Таким образом, три сверточные нейросети сформировали систему преобразования черно-белого изображения в цветное, и эта система не только позволяет пользователям легко исправлять ошибки раскрашивания монохромных материалов, но также обеспечивает существенное уменьшение цветового дрейфа от кадра к кадру.

Все это, повторюсь, было представлено примерно лет пять тому назад, и с тех пор генAI претерпел удивительно стремительное развитие. Сегодня плоды «творчества» искусственного интеллекта уже повсюду. Но так ли совершенен генAI, чтобы действительно составить конкуренцию человеческому разуму? У ряда ученых в области AI есть ответ, но к нему мы подойдем чуть позже.

По мере совершенствования искусственного интеллекта представители многих творческих профессий вдруг обнаруживают, что оказались на совершенно незнакомой им территории. Максимально реалистичные изображения, на создание которых ранее уходило мно-

го труда и, если можно так выразиться, мук творчества, теперь появляются за считанные секунды, их уже миллионы и в таких количества они возникают чуть ли не ежедневно. Эти творения участвуют в конкурсах и даже побеждают в них. И при этом ни одно из них не создано человеком. Как тут не встревожиться? Ведь речь не только об изображениях. Подобным способом создаются музыка, новостные и даже художественные тексты, сценарии к сериалам и различному потоковому контенту. И это далеко не полный список.

И здесь хотелось бы снова сослаться на уже упоминавшегося в первой части цикла авторитетного ученого Ахмеда Эльгаммала – профессора компьютерных наук Ратгенского университета. Его соображения помогают лучше узнать о распространении AI в творческой среде и о том, что это несет человеческому творчеству в цифровую эру.

Прежде всего интересно было бы разобраться, как искусственный интеллект генерирует изображения. Как отметил Эльгаммал, пять лет тому назад произошел рывок в совершенствовании AI, состоявший в появлении генеративных состязательных сетей GAN (Generative Adversarial Network). Такая сеть способна получить на входе те или иные изображения и на их основе сгенерировать аналогичные. К примеру, если загрузить в сеть изображения собак, то она вернет полностью новые версии, на которых тоже будут собаки, но не те, что на



Созданная искусственным интеллектом картина *Théâtre D'opéra Spatial*, получившая первый приз в цифровой категории на конкурсе ярмарки штата Колорадо (США)

исходных изображениях. На тот момент сети GAN произвели настоящую революцию, и многие художники не преминули этим воспользоваться.

Но развитие AI, разумеется, на этом не остановилось. Если привести в пример человека, того же художника, то ему ведь не обязательно показать картинку, чтобы попросить его нарисовать нечто похожее – достаточно объяснить на словах. Например, одинокое дерево в поле пасмурным осенним днем. Нет сомнений, что художник вполне справится с задачей, причем вариантов ее решения будет множество, стремящееся к бесконечности.

Те же возможности обрел и искусственный интеллект следующего поколения, получив способности генерировать изображения на основе описательного текста. К тому же это дало возможность человеку точнее управлять процессом генерации изображений за счет более точного описания желаемого результата, то есть исключения заведомо лишних вариантов. В основе процесса лежала тренировка модели на большом числе изображений и сопровождавших их текстовых описаний, чтобы научить нейросеть понимать, как слова соотносятся с изображениями. В процессе обучения AI получал, например, изображение с птицей на дереве и делал предположение, где на изображении дерево, а где птица, а нейросеть оценивала, правильный ли дан ответ. «Перелопатив» подобным образом миллиарды изображений, AI в итоге учится находить для слов соответствующие изображения.

Но просто выбрать правильное изображение по его словесному описанию – это еще далеко не все. Одно из слабых мест genAI – мелкие детали, что вполне объяснимо с технической точки зрения и описывается термином «класс точности». Дело в том, поясняет Эльгаммаль, что искусственному интеллекту сложно иметь дело с мелкими деталями. Нейросетям всегда трудно генерировать что-то очень малое, потому что они натренированы на оптимизацию через потерянную функцию, а это критерий, заставляющий AI оптимизировать изображение в целом, сделать так, чтобы оно в общем было правильным.

Поэтому AI пренебрегает мелкими деталями, которые обязательно замечают люди. К примеру, мимо человеческого внимания не пройдет четырехпалая рука или человек с тремя ногами. А для AI это останется незамеченным, как и многие другие мелкие детали на фоне.

Подобные ситуации были не раз описаны в произведениях тех же писателей-фантастов. Когда, к примеру, некий генератор вещей выдает по запросу человека что угодно, но не то, что тот просил. Очень и очень близкое, но не то. Потому что практически любая словесная формулировка оставляет слишком много пространства для интерпретации. Словесное описание – это ведь не чертеж, не принципиальная схема и не точная технологическая карта. Особо когда речь идет о творчестве.

К тому же стоит помнить, что нейросети состоят из высокопроизводительных компьютеров, которые довольно «прожорливы» – потребляют много электроэнергии и выделяют много тепла, в том числе в процессе тренировки тех или иных моделей AI. Чтобы добиться успешного машинного обучения, нужно «гонять» компьютеры с

мощными GPU неделями, пока они обрабатывают миллиарды изображений. А затем придется повторить все это много раз, чтобы оптимизировать процесс машинного творчества.

Но даже после такой интенсивной тренировки модели они должны постоянно работать на базе GPU с миллионами процессоров, а это очень энергоемкие устройства. Стало быть, если они будут работать довольно долго, то для этого потребуется много энергии, что, помимо чисто финансовых расходов, повлечет еще и вред окружающей среде.

Помимо пренебрежения мелкими деталями, у искусственного интеллекта есть и другие слабые места. Впрочем, как и у людей, что вполне закономерно, поскольку в основе искусственного интеллекта лежит модель человеческого разума. Правда, последняя тоже пока изучена далеко не полностью, и все неопределенности, само собой, перекочевали в модели AI.

Но вернемся ко второй слабости искусственного интеллекта. Она заключается в неспособности отличать правду от вымысла, в отсутствии защиты от предубежденности и дезинформации, которые эти модели волей-неволей усваивают, обучаясь на информации, получаемой из Интернета. Все настолько серьезно, что даже встала дилемма: нужно ли ограничить доступную для моделей AI информацию каким-то определенным перечнем либо дать им полный доступ ко всему.

В связи с этим Ахмед Эльгаммаль задается вопросом: «А как мы можем контролировать данные, передаваемые искусственному интеллекту?». На этот счет есть разные мнения, высказываемые в широком диапазоне от политики до религии и стиля жизни со всеми промежуточными вариантами. Цензура, в поддержку которой звучат определенные голоса, уж точно не выход. Равно как не применима она и к людям.

Хотя бы потому, что AI, и это естественно, должен отражать все мнения и точки зрения, имеющиеся в мире. А раз так, то этому всегда сопутствует большой объем дезинформации. Но ничего не поделаешь – таков мир, в котором мы живем. К тому же дезинформация далеко не всегда результат злонамеренности, очень часто – плод заблуждений. Но на заре цивилизации почти все было заблуждением, из которого потом выросла наука. А все попытки цензуры вели только к торможению прогресса и откату в прошлое. К тому же есть очень хорошее высказывание на этот счет: *«Дурак – не тот, кто заблуждается, а тот, кто упорствует в своем заблуждении».*

Возвращаясь к тому, что аналогично моделям AI, черпающим данные – как корректные, так и ложные – из Интернета, действуют и сами люди, получающие сведения из соцсетей. Далеко не всегда есть возможность точно отделить правду от лжи, зачастую приходится делать предположения и полагаться на интуицию, жизненный опыт, репутацию источника и т. д.

В настоящее время у AI нет шансов отделить факт от вымысла. Для него и то, и другое – просто слова, а как только речь заходит о фактах, начинаются проблемы. Поскольку и сами люди далеко не всегда способны определить, какие факты – действительно факты, и какие мнения верны.

Правда, здесь у AI есть своя сильная сторона. Пусть он не способен отличить правду от вымысла, но обладает возможностью быстрого поиска верифицирующих данных относительно того или иного факта, что очень полезно, например, при проверке информации для выпуска новостей.

Еще одна проблема, о которой вкратце шла речь в первой статье и от которой не удастся уйти здесь, это правовая основа применения искусственного интеллекта. Технология настолько нова и так быстро развивается, что законодательство за ней просто не поспевает, и это приводит к определенным коллизиям, особенно в сфере авторских прав.

Проблема авторских прав пришла вместе с инструментами создания изображений текущего поколения, которые в основном натренированы на миллиардах изображений. Однако еще пару лет назад, по мнению Эльгаммалю, этой проблемы не было, потому что в то время художники обычно применяли искусственный интеллект, обучая его модели с помощью собственных изображений, созданных этими художниками.

Проблема стала острой, когда для тренировки моделей AI стали использовать миллиарды изображений, взятых из Интернета, причем без ведома художников, их создавших. Но здесь надо понимать еще вот что: хотя использование чужих изображений неэтично, это далеко не всегда является нарушением законов об авторских правах. Ведь AI делает не точную копию чьего-то творения, а версии на его основе, поэтому закон об авторском праве не нарушается.

Тем не менее нужно постоянно напоминать всем, что действовать неэтично не следует. Искусственный интеллект, как практически все в мире, можно использовать и во благо, и во вред. Нормально, когда тот или иной пользователь применяет AI в сочетании со своим собственным контентом, не крадя плоды работы других людей.

И еще один вопрос, возможно, самый важный, который волнует очень и очень многих, заключается в следующем: «Способен ли искусственный интеллект научиться настоящему творчеству и обрести эмоции, столь необходимые для создания действительно великого произведения искусства?». Здесь Эльгаммалю вполне конкретен: «Возможности AI нынешнего поколения ограничены только копированием плодов человеческого творчества». Стало быть, чтобы этот искусственный интеллект создал что-то действительно полезное, требуется немалое участие людей в процессе создания этого «чего-то». Профессор Эльгаммалю уверен, что AI – это отличный инструмент, но не нечто, способное творить само по себе. В основе творческого соиздания лежит личное отношение человека к тому, что происходит в мире, а для этого человек должен быть осведомлен об этом и осознавать суть происходящего. У моделей AI таких возможностей просто нет.

Ахмед Эльгаммалю рассказал, что некоторое время назад он и его коллеги использовали AI в проекте по генерированию 10-й симфонии Бетховена. Модели AI прошли интенсивную тренировку на огромных объемах классической музыки, затем им показали наброски симфонии, которые оставил Бетховен, после чего искусственному интеллекту поставили задачу сгенерировать компиляцию



*Профессор Ратгерского университета Ахмед Эльгаммалю*

этих нот. И... ничего не получилось. Отличный пример того, что AI – это инструмент, но в его работе нет творчества. Такова ситуация с AI в мире в настоящее время. Творчество по-прежнему остается свойством, присущим человеку, а искусственный интеллект просто следует правилам, которые определили люди, чтобы создавать для них контент.

Что же касается мнения, что AI – это просто веяние моды, современное направление в искусстве по аналогии с импрессионизмом или модернизмом, то, как считает Эльгаммалю, в последние лет пять такое действительно имело место, но уже завершилось. У тех, кто начинал применять AI несколько лет назад, прослеживалась определенная эстетика в работах, что-то жуткое и нечеловеческое. Это был специфический стиль, теперь все пришло к фотореалистичности. С одной стороны, это хорошо, поскольку свидетельствует о повышении «мастерства» искусственного интеллекта, но с другой, исчезла неожиданность, сюрреализм. К тому же фотореалистичные изображения можно получить гораздо более простым и проверенным способом – с помощью съемки. Но и способности AI тоже будут полезны, а как – об этом можно прочитать в статье [«Нейронные поля яркости – сделать большое из малого»](#), опубликованной в [августовском \(06/136\) номере журнала](#).

Вот что еще сказал профессор Эльгаммалю: «Я думаю, что эра [AI как направления в искусстве] ушла. Теперь это больше доступный всем инструмент для генерирования фотореалистичного изображения, графических композиций и логотипов, а не уникальный предмет искусства».

Ну и в завершение – обещанный в начале этой статьи ответ на вопрос: «Способен ли AI составить конкуренцию человеческому разуму?». Ответ прозвучал из уст того же профессора Эльгаммалю: «Искусственному интеллекту всегда будет не хватать искры творчества». ■